STATISTICAL PROBLEMS OF THE ELEMENTARY GAUSSIAN PROCESSES

Arató M. - Benczúr A. - Krámli A. - Pergel J.

Part II.

STATISTICS AND RELATED PROBLEMS.

Tanulmányok 41/1975

A kiadásért felel: Gertler János

ISBN 963 311 009 2

757731 MTA KESZ Sokszorosító F.v.: Szabó Gyula

Contents

Part II.

Chapter 1.

The elementary Gaussian processes

1.§ Definition and properties of elementary Gaussian processes.

A k dimensional stochastic process $\xi(t)$ is called elementary Gaussian process if it is Gaussian, stationary and Markov. In continuous time case we suppose that it is a diffusion process /see Part I. Ch.1/.

In the following we shall examine first the connection between the elementary Gaussian processes and the stochastic difference resp. stochastic differential equations. In the sequel we shall suppose that the process is not degenerated and it is linearly regular /see I. Ch. 2/. By the phrase $\xi(t)$ is not degenerated we mean that its components are pointwise linearly independent.

In the discrete time case the connection between the elementary Gaussian processes and the stochastic difference equations is characterized by the following two theorems. Let $\varepsilon(n)$ be normally and identically distributed independent k-dimensional random variables, with

$$E\,\varepsilon(n)=0, \quad E\left(\varepsilon(n),\, \varepsilon^*(n)\right)=B_\varepsilon,$$

where Rank $B_\varepsilon \geq 1$. Let Q denote a $k\times k$ matrix with eigenvalues λ_i, where $|\lambda_i| < 1$, $i=1,2,\ldots k$. Then the equation

$$(1.1) \quad B(0) = QB(0)Q^* + B_\varepsilon$$

has nonsingular positive definite solution B (0) /see Gantmaher [1] /. Let $\underline{\xi}(0)$ be normally distributed k-dimensional random vector variable with the parameters

$$E\,\underline{\xi}(0) = 0, \quad E(\underline{\xi}(0), \underline{\xi}^*(0)) = B(0)$$

and independent of $\underline{\varepsilon}(n)$.

Under these conditions there holds the following.

Theorem 1.1 Let $\underline{\xi}(n)$ be defined recursively by the equation

$$(1.2) \quad \underline{\xi}(n) = Q\,\underline{\xi}(n-1) + \underline{\varepsilon}(n), \qquad n = 1, 2, \ldots$$

where the eigenvalues of Q are in the inside the unit circle, $\underline{\varepsilon}(n)$ is independent of $\mathcal{F}_0^{n-1}(\xi) = \sigma(\xi(0), \ldots, \xi(n-1))$, $\xi(0)$ is normally distributed with $E\,\underline{\xi}(0) = 0$ and covariance B (0) satisfying (1.1). Then $\underline{\xi}(n)$ is an elementary Gaussian process with $E\,\underline{\xi}(n) = 0$ and covariance matrix

$$(1.3) \quad B(\ell) = E(\xi(n+\ell)\, \underline{\xi}^*(n)) = Q^\ell B(0).$$

Proof. The normality follows directly from the linearity of (1.2) by induction.

By repeated application of (1.2)

$$\underline{\xi}(n) = \underline{\varepsilon}(n) + Q\,\underline{\varepsilon}(n-1) + \dots + Q^{n-1}\underline{\varepsilon}(1) + Q^{n}\underline{\xi}(0)$$

from where (using the independence of the variables $\underline{\varepsilon}(n)$ and (1.1))

$$E(\underline{\xi}(n), \underline{\xi}^{*}(n-\ell)) = Q^{\ell}[B_{\varepsilon} + Q B_{\varepsilon} Q^{*} + \dots + Q^{n-\ell} B_{\varepsilon} Q^{*\,n-\ell}] =$$

$$= Q^{\ell}[B_{\varepsilon} + Q B_{\varepsilon} O^{*} + \dots + Q^{n-\ell-1}(B_{\varepsilon} + Q B(O) Q^{*})(Q^{*})^{n-\ell-1}]$$

and by induction (in ℓ)

$$E(\xi(n), \underline{\xi}^{*}(n-\ell)) = Q^{\ell} B(0)$$

which proves the stationarity. The Markov property will be proved after lemma 1.

<u>Lemma 1.</u>

If $\underline{\eta}(n)$ is a Gaussian process /n=0,1,...,/ with the properties

$E(\underline{\eta}(n) \mid \mathcal{F}_0^{n-1}(\eta)) = C(n)\,\underline{\eta}(n-1)$ **and**

$$E((\underline{\eta}(n) - E(\underline{\eta}(n) \mid \mathcal{F}_0^{n-1}(\underline{\eta}))).(\underline{\eta}^{*}(n) - E(\underline{\eta}(n) \mid \mathcal{F}_0^{n-1}(\underline{\eta}))^{*}) \mid \mathcal{F}_0^{n-1}(\underline{\eta})) = B(n),$$

where $C(n)$ and $B(n)$ are deterministic matrix functions, then $\underline{\eta}(n)$ is a Markov-process.

The proof is trivial, as a Gaussian distribution is determined by the first two moments.

<u>The proof of the Markov property</u>. As $\underline{\xi}(n) = A\,\underline{\xi}(n-1) + \underline{\varepsilon}(n)$, where $\underline{\varepsilon}(n)$ is independent of $\mathcal{F}_0^{n-1}(\xi)$

$$E(\underline{\xi}(n) | \mathcal{F}_0^{n-1}(\xi)) = A\underline{\xi}(n-1)$$

and

$$E(((\underline{\xi}(n) - E(\underline{\xi}(n) | \mathcal{F}_0^{n-1}(\underline{\xi})).(\underline{\xi}(n) - E(\xi(n) | \mathcal{F}_0^{n-1}(\xi)))^* | \mathcal{F}_0^{n-1}(\xi)) = B_\varepsilon$$

so the conditons of lemma 1. are satisfied.

In the precedings we examined the one side process /n > 0/ but it is interesting to consider a stationary process from $-\infty$.

Remark 1. Let $\underline{\varepsilon}(n)$ be a sequence of independent identically distributed /in the sequel i.i.d./ k-dimensional Gaussian random vectors with nondegenerated covariance matrix B_ε and Q a k x k matrix, such that its eigenvalues are all either inside the unit circle or outside the unit circle. Then the equation

$$(1.1) \quad \underline{\xi}(n) = Q\,\underline{\xi}(n-1) + \underline{\varepsilon}(n) \qquad (n = 0, \pm 1, \dots)$$

has a unique stationary solution with finite second moments. This solution is a regular Gaussian Markov process.

For the proof we need the following.

Lemma 2. The series

$$(1.4) \quad \sum_{n=0}^{\infty} Q^n \underline{\varepsilon}(n)$$

is convergent if and only if $|\lambda_i| < 1$, where λ_i are the eigenvalues of the matrix Q.

<u>Proof.</u> The convergence of (1.4) holds if and only if the series $\sum_{n=0}^{\infty} Q^n B_\varepsilon Q^{*n}$ of the variance matrices is convergent, /Kolmogorov's 3 series theorem is true also for vector variables/. The above series of variance matrices is convergent if and only if $|\lambda_i| < 1$ for every i.

<u>Proof of remark 1.</u> By lemma 2. the series $\underline{\xi}(n) = \sum_{i=0} Q^i \underline{\varepsilon}(n-i)$ / $\underline{\xi}'(n) = \sum_{i=0}^{\infty} Q^{-i} \underline{\varepsilon}(n+i)$ / is convergent if the eigenvalues of Q are inside /outside/ the unit circle. We can see directly that $\underline{\xi}(n)$ / $\underline{\xi}'(n)$ / is a regular stationary Gaussian Markov process and satisfies the equation (1.1') . Let $\underline{\eta}(n)$ / $\underline{\eta}'(n)$ / an another solution, then $\underline{\zeta}(n) = \underline{\xi}(n) - \underline{\eta}(n)$ / $\underline{\zeta}'(n) = \underline{\xi}'(n) - \underline{\eta}'(n)$ / satisfies the equation $\underline{\zeta}(n) = Q\underline{\zeta}(n-1)$ i.e. $\underline{\zeta}(n) = Q^n \underline{\zeta}(0)$ for every $n, -\infty < n < \infty$. Therefore there exists a real number $q > 1$ $(q' < 1)$ - the minimum (maximum) of the eigenvalues of the matrix Q (Q') such that

$$E(\underline{\zeta}(n), \underline{\zeta}^*(n)) \geq q^n E(\underline{\zeta}(0)\, \zeta^*(0)) \quad \text{for} \quad n > 0$$

$$(E(\zeta'(n), \zeta^*(n)) \geq q'^n E(\zeta'(0), \zeta^*(0)) \quad \text{for} \quad n < 0).$$

i.e. the uniqueness of the solution is proved.

<u>Remark 2.</u> From the proof we can see that - depending on the eigenvalues of Q - the best forward /backward/ extrapolation of $\underline{\xi}(n+1)$ is $Q\underline{\xi}(n)$ and the covariance matrix of the error is B_ε .

<u>Remark 3.</u> The random vector $\underline{\xi}(n)$ / $\underline{\xi}'(n)$ / is $\mathcal{F}_{-\infty}^{n}(\varepsilon)$ / $\mathcal{F}_{n}^{\infty}(\varepsilon)$ / measurable and therefore $\varepsilon(n+1)$ / $\varepsilon(n-1)$ / is

independent of $\mathcal{F}_{-\infty}^{n}(\xi)$ / $\mathcal{F}_{n}^{\infty}(\xi)$ /. The proof shows that $\mathcal{F}_{-\infty}^{n}(\xi) = \mathcal{F}_{-\infty}^{n}(\varepsilon)$. The $\underline{\varepsilon}(n)$ process is called innovation process.

<u>Theorem 1.2</u> Let $\underline{\xi}(n)$ / $n = 0, \pm 1, \dots$ / be a k-dimensional stationary Gaussian Markov process with 0 mean and covariance matrix function $B(\ell)$. Then there exists a k x k matrix Q with eigenvalues inside the unit circle and a sequence of i.i.d. Gaussian vectors $\underline{\varepsilon}(n)$ such that the equation (1.1') with Q and $\underline{\varepsilon}(n)$ holds for $\underline{\xi}(n)$.

<u>Proof.</u> As in the case of random variables with joint Gaussian distribution the regression is always linear, it follows from the Markovity that with some Q $E(\underline{\xi}(n) | \mathcal{F}_{-\infty}^{n-1}(\xi)) = Q\,\underline{\xi}(n-1)$. $\underline{\xi}(n)$ may be written in the form $Q\,\underline{\xi}(n-1) + \underline{\varepsilon}(n)$, where $\underline{\varepsilon}(n) = \underline{\xi}(n) - Q\,\underline{\xi}(n-1)$, $\underline{\varepsilon}(n)$ is independent of $\mathcal{F}_{-\infty}^{n-1}(\xi)$, therefore the random vectors $\underline{\varepsilon}(n)$ are independent. It follows from the stationarity that the matrix Q and the distribution of $\underline{\varepsilon}(n)$ do not depend on n.

<u>Remark 4.</u> This representation of elementary Gaussian processes shows that the process $\underline{\xi}'(n)$ in remark 1. satisfies another difference equation

$$\underline{\xi}'(n) = Q'\,\underline{\xi}'(n-1) + \underline{\delta}(n).$$

From the explicit form of the solution of equation (1.1') it is easy to see, that $Q' = Q^{-1}$ and

$$\delta(n) = Q^{-2}\,\underline{\varepsilon}(n-1) - \frac{Q^{2} - I}{Q^{3}} \sum_{i=0}^{\infty} Q^{-i}\,\underline{\varepsilon}(i+n).$$

It is well known that the reversed process $\underline{\xi}(-n)$ of a Markov process $\underline{\xi}(n)$ is also Markov. Therefore, on the basis of theorem 1.2 $\underline{\xi}(n)$ satisfies the equation

$$\underline{\xi}(n-1) = \tilde{Q}\,\underline{\xi}(n) + \underline{\tilde{\varepsilon}}(n)$$

where $\tilde{\varepsilon}(n)$ is a sequence of i.i.d. Gaussian random vectors with covariance matrix B_{ε} , and $\tilde{\varepsilon}(n)$ is independent of $\mathcal{F}_n^{\infty}(\xi)$. More about reversed processes see in Andel's paper [2] .

<u>Remark 5.</u> The parameter matrices $\tilde{Q}$ and $B_{\tilde{\varepsilon}}$ can be calculated from the matrices Q and B_{ε} of equation (1.1) by solving the system of equations

$$H = B_{\varepsilon}Q^{*} + QB_{\varepsilon}Q^{*2} + \dots + Q^{n}B_{\varepsilon}(Q^{*})^{n+1},$$
$$\tilde{Q} = (B_{\varepsilon} + QH) = H,$$
$$B_{\xi} = QB_{\xi}Q^{*} + B_{\varepsilon},$$
$$B_{\xi} = \tilde{Q}B_{\xi}\tilde{Q}^{*} + B_{\tilde{\varepsilon}}.$$

<u>Proof.</u> The proof is straightforward using the representation

$$\underline{\xi}(n) = \sum_{k=0}^{\infty} Q^{k}\,\underline{\varepsilon}(n-k).$$

Althought the observations of a real process give a discrete time process, it is useful to consider continuous time processes, because some phenomena can be described more adequately in that way, and also the results have simpler form. In theorem 1.5 - on the basis of Doob's results we shall formulate the exact correspondence between the two cases.

The analogon of the sequence of i.i.d. Gaussian vectors is the multidimensional white noise: the non-existing "derivative" of the multidimensional Wiener process. To stochastic difference equation corresponds the stochastic differential equation, which was introduced in Part I.

Let $(\underline{w}(t), \mathcal{F}_t)$ be a k-dimensional Brownian motion process, possibly degenerated, with the local parameters $E\,\underline{w}(t) = 0$, $E(d\underline{w}(t), d\underline{w}^*(t)) = B_w\,dt$, and let the k x k matrix A have eigenvalues only with negative real parts. Let us consider the stochastic differential equation

$$d\underline{\xi}(t) = A\underline{\xi}(t)\,dt + d\underline{w}(t) \tag{1.5}$$

and let $B(0)$ the unique solution of the matrix equation A matrix equation of the type $AX - XB = C$ is uniqually solvable if and only if A and B have no common eigenvalues (see Gantmakher [1]).

$$A\,B(0) + B(0)\,A^* = -B_w \tag{1.6}$$

<u>Theorem 1.3</u> The only stationary solution $\underline{\xi}(t)$ with continuous sample paths of (1.5) is an elementary Gaussian process. Its covariance matrix function has the form

$$B(\tau) = B(0)\, e^{A\tau}.$$

<u>Remark 5.</u> The solution of (1.5) has the integral representation (see example 3. too)

$$(1.7) \qquad \underline{\xi}(t) = \int_{-\infty}^{t} e^{+A(t-s)}\, d\underline{w}(s).$$

where the existence of integral (1.7) is equivalent to the finiteness of the integral $\int_{-\infty}^{0} e^{-As} B_W e^{-A^* s}\, ds$. The existence of the integral is equivalent to the condition $\lim\limits_{s\to\infty} e^{-sA} B_W = 0$. This representation is analogous to the sum in remark 1., and shows that $\underline{\xi}(t)$ is $\mathcal{F}_t$ measurable. $\underline{w}(t)$ is the innovation process of $\underline{\xi}(t)$. It is well known from the matrix theory that $\int_{-\infty}^{0} e^{-As} \bar{B}_W e^{-A^* s}\, ds$ exists, and gives the unique solution of equation (1.6) if and only if the real parts of the eigenvalues of A are all negative. The existence of integral (1.7) follows from the definition of the stochastic integral of a deterministic function on a finite interval.

<u>Proof of theorem 3. and remark 5.</u> First let us notice that $\underline{\xi}(t-\tau) = e^{A\tau}\, \underline{\xi}(t) + \int_{0}^{\tau} e^{A(\tau-s)}\, d\underline{w}(t+s)$ if $\underline{\xi}(t)$ is defined by (1.7) . As the second term on the right hand side

is independent of $\mathcal{F}_{\infty}^{t}(\xi)$ so $e^{A\tau}\xi(t)$ is the best extrapolation $E(\xi(t+\tau)/\mathcal{F}_{-\infty}^{t}(\xi))$ of the process $\xi(t)$.
From this representation and

$$E[(\xi(t+\tau)-E(\xi(t+\tau)|\mathcal{F}_{-\infty}^{t}(\xi)).(\xi(t+\tau)-E(\xi(t+\tau)|\mathcal{F}_{-\infty}^{t}(\xi)))^{*}|\mathcal{F}_{-\infty}^{t}(\xi)] =$$
$$= e^{At}\int_{0}^{\infty} e^{-Av} D_{w} e^{-A^{*}v} dv$$

the markovity of the process $\xi(t)$ and the formula for $B(\tau)$ is straightforward.

By a direct computation we may convince that the integral (1.7) satisfies the equation (1.5) :

$$A\int_{0}^{t}\xi(\tau)d\tau = A\int_{0}^{\tau}\int_{-\infty}^{\tau} e^{A(\tau-s)}dw(s)d\tau = A\int_{0}^{t}\int_{0}^{\tau} e^{A(t-s)}dw(s)d\tau +$$
$$+A\int_{0}^{t}\int_{-\infty}^{0} e^{A(\tau-s)}dw(s)d\tau = \int_{0}^{t}(e^{A(t-s)}-I)dw(s) + e^{At}\xi(0) - \xi(0) =$$
$$= \xi(t) - \xi(0) - w(t) + w(0),$$

where I denotes the unit matrix.

The stationarity and the continuity with probability 1 of $\xi(t)$ is obvious from the representation (1.7) . The uniqueness follows from general theorems for the stochastic differential equations /see Part I./, but it can be verified similarly to the discrete time case.

<u>Remark 6.</u> An analogous statement is true for matrices with eigenvalues having only positive real parts. Then $\xi(t) = = \int_{t}^{\infty} e^{A(t-s)} dw(s)$ will be the desired solution.*

---

*But $\xi(t)$ does not solve the equation (1.5) in strict sense, because $\xi(t)$ is not $\mathcal{F}_t$ measurable.

The converse of theorem 1.3 is also valid:

<u>Theorem 1.4</u> If the k-dimensional process $\underline{\xi}(t)$ with 0 mean and continuous sample paths is a stationary Gaussian Markov one, then there exists a matrix A with eigenvalues in the left halfplane and a Wiener process $\underline{w}(t)$, such that

$$d\underline{\xi}(t) = A\underline{\xi}(t) + d\underline{w}(t)$$

<u>Proof.</u> From Gauss-Markov property we get the existence of a matrix function $Q(t_2,t_1)$ two variables, satisfying the relation

$$(1.8)\quad E(\underline{\xi}(t_2) \mid \mathcal{F}_{-\infty}^{t_1}(\underline{\xi})) = Q(t_2,t_1)\,\underline{\xi}(t_1), \quad \text{for} \quad t_2 \geq t_1 .$$

Applying (1.8) succesively we can deduce the functional equation for $t_1 \leq t_2 \leq t_3$:

$$(1.9)\quad Q(t_3,t_1) = Q(t_3,t_2) \times Q(t_2,t_1) .$$

This relation is valid for non-stationary processes too. If moreover $\underline{\xi}(t)$ is stationary: $Q(t_2,t_1) = Q(t_2\ \ t_1)$.

As the process $\underline{\xi}(t)$ is continuous with probability 1, and therefore - being Gaussian - it is continuous in mean square too, the matrix function $Q(t_2-t_1)$ is continuous. The unique continuous solution of (1.8) under the initial condition $Q(0)=I$ is the matrix function $e^{A(t_2-t_1)}$ with some constant matrix A . For $\delta > 0$ let us take the sum

1.1o $$\sum_{i=1}^{[t/\delta]} [\underline{\xi}(i\delta) - e^{A\delta}\underline{\xi}((i-1)\delta)] = \underline{\xi}(\delta[\tfrac{t}{\delta}]) - \underline{\xi}(0) - \delta \sum_{i=1}^{[t/\delta]-1} \underline{\xi}(i\delta) + \mathcal{O}(\delta),$$

which almost surely tends to $\underline{\xi}(t) - \underline{\xi}(0) - A\int_0^t \underline{\xi}(s)ds$ if $\delta \to 0$. As, by (1.8) the terms on the left hand side are i.i.d. Gaussian random vectors, the limit process will be a multidimensional Gaussian process with independent, stationary increments, i.e. a multidimensional Wiener process. So we have proved that $\underline{\xi}(t)$ satisfies the equation (1.5) . Theorem 1.3 involves the condition on eigenvalues of matrix A .

Remark 7. It turns out from the above proof that a Gaussian process is elementary if and only if its covariance matrix function has the form $B(0)\,e^{A|t_1-t_2|}$, where eigenvalues of A are all in the left halfplane.

On the basis of 1.8 and 1.9 a stationary Gaussian process is Markov if and only if for $t_2 > t_1$

$$E(\underline{\xi}(t_2)) = e^{A(t_2-t_1)}\underline{\xi}(t_1).$$

For non-stationary Gaussian processes the necessary and sufficient condition of markovity is the relation

$$E(\underline{\xi}(t_1), \underline{\xi}^*(t_2)) = E(\underline{\xi}(t_1), \underline{\xi}^*(t_1)) \cdot Q(t_2, t_1)$$

where $Q(t_2, t_1)$ satisfies the equation (1.9) .

The following theorem explains the connection between the discrete and continuous time elementary Gaussian processes.

Theorem 1.5 /Doob's paper [1] / The continuous time process $\underline{\xi}(t)/-\infty < t < \infty/$ with continuous sample paths is an elementary Gaussian one if and only if for each $\delta > 0$ the discrete time process $\underline{\xi}(n\delta)$ is elementary Gaussian.

Proof. Necessity is trivial. For the proof of sufficiency let us first notice that the joint distribution of random vectors $\underline{\xi}(k_1\delta), \dots, \underline{\xi}(k_n\delta)$ for every $\delta > 0$ and every finite sequence $\{k_1, \dots, k_n\}$ of integers is Gaussian. Hence, by the continuity of sample paths the process $\underline{\xi}(t)$ is Gaussian. Stationarity is obvious. There remains to prove markovity. For this purpose - on the basis of remark 7. - it is sufficient to prove that for $t_2 > t_1$ $E(\underline{\xi}(t_2) \mid \underline{\xi}(t_1)) = e^{A(t_2-t_1)}\underline{\xi}(t_1)$. By Gaussity there exists a matrix function $Q(t_2-t_1)$ for which $E(\underline{\xi}(t_2) \mid \underline{\xi}(t_1) = Q(t_2-t_1)\,\underline{\xi}(t_1)$. As the process $\underline{\xi}(n\delta)$ is Markov for every $\delta > 0$. $Q(m\delta) \cdot Q(n\delta) = Q((m+n)\delta)$. Because of the continuity of sample paths $Q(t_2-t_1)$ is also continuous, and so - satisfying the equation (1.9) and the initial condition $Q(0) = I$ - has the desired form.

Theorem 3.2 of Part I. asserts that two k-dimensional Wiener process $\underline{W}^{(1)}(t)$ and $\underline{W}^{(2)}(t)$ can be distincted with probability 1 observing them on an arbitrary small interval $[0, T]$ because

$$\lim_{n\to\infty} \sum_{i=1}^{2^n} \left(w\left(\frac{Ti}{2^n}\right) - w\left(\frac{T(i-1)}{2^n}\right)\right)\left(w\left(\frac{Ti}{2^n}\right) - w\left(\frac{T(i-1)}{2^n}\right)\right)^* = B_w T.$$

We may ask now if distinction on this way is possible with probability 1 for any two elementary Gaussian processes. By theorem 1.6, see e.g. Baxter [1], answer is no; if their matrices of diffusion $B_W^{(1)}$ and $B_W^{(2)}$ are the same. Moreover, later we shall see that there is no possibility to distinct them almost surely on a finite interval.

<u>Theorem 1.6</u> Let $\underline{\xi}(t)$ be a k-dimensional elementary Gaussian process with parameters A and B_W. Then with probability 1 $\lim_{n\to\infty} \sum_{i=1}^{2^n} \left(\underline{\xi}(t_i) - \underline{\xi}(t_{i-1})\right)\left(\underline{\xi}(t_i) - \underline{\xi}(t_{i-1})\right)^* = B_W t,$
where $t_i = i\frac{T}{2^n}$.

<u>Proof.</u> On the basis of (1.10) we can write that

$$\sum_{i=1}^{2^n} \left(\underline{\xi}(t_i) - \underline{\xi}(t_{i-1})\right)\left(\underline{\xi}(t_i) - \underline{\xi}(t_{i-1})\right)^* = \sum_{i=1}^{2^n} \left(\underline{w}(t_i) - \underline{w}(t_{i-1})\right)\left(\underline{w}(t_i) - \underline{w}(t_{i-1})\right) +$$

$$+ \sum_{i=1}^{2^n} \int_{t_{i-1}}^{t_i} A\,\underline{\xi}(t)\,dt \left(\underline{w}(t_i) - \underline{w}(t_{i-1})\right)^* + \sum_{i=1}^{2^n} \left(\underline{w}(t_i) - \underline{w}(t_{i-1})\right) \int_{t_{i-1}}^{t_i} \underline{\xi}^*(t) A^*\,dt +$$

$$+ \sum_{i=1}^{2^n} \int_{t_{i-1}}^{t_i} A\,\underline{\xi}(t)\,dt \int_{t_{i-1}}^{t_i} \underline{\xi}^*(t) A\,dt$$

As for almost all sample path the vector functions $\int_0^t A\underline{\xi}(s)\,ds$ and $\int_0^t \underline{\xi}^*(s) A^*\,ds$ have bounded variation the last three terms tend to 0 with probability 1 if $n \to \infty$.

In the sequel we give some examples for non-stationary

Gaussian processes, defined by stochastic differential equations.

Example 1. If $A(t)$ and $D(t)$ are deterministic vector resp. matrix functions, then the process $\eta(t)$ with stochastic differential $d\underline{\eta}(t) = \underline{A}(t)\,dt + D(t)\,d\underline{w}(t)$

is a Gaussian one with independent increments, where

$$E(\underline{\eta}(t) - \underline{\eta}(0)) = \int_0^t A(s)\,ds, \quad D^2(\underline{\eta}(t) - \underline{\eta}(0)) = \int_0^t D(s)D^*(s)\,ds$$

Example 2. The solution of the homogeneous linear stochastic equation /if $\eta(0) \neq 0$/ $d\eta = B(t)\,\eta(t) + D(t)\,\eta(t)\,dw(t)$

has the following form

$$\eta(t) = \eta(0)\,\exp\left\{ \int_0^t [B(s) - \frac{1}{2}D^2(s)]\,ds + \int_0^t D(s)\,dw(s) \right\}.$$

The proof follows immediately from the Ito formula for the process $\xi(t) = \lg \eta(t)$, which states that

$$d\,\xi(t) = \frac{1}{\eta(t)}B(t)\,\eta(t)dt - \frac{1}{2}\frac{1}{\eta^2(t)}D^2(t)\,\eta^2(t)dt$$

$$+ \frac{1}{\eta(t)}D(t)\,\eta(t)\,dw(t)$$

and so

$$d\xi(t) = [B(t) - \frac{1}{2}D^2(t)]\,dt + D(t)\,dw(t),$$

$$\xi(t) = \xi(0) + \int_0^t [B(t) - \frac{1}{2}D^2(t)]\,dt + \int_0^t D(t)\,dw(t)$$

$$\eta(t) = \exp \xi(t)$$

The last formula is true until $\eta(t)$ does not become zero. But the right side does not do it in case $\eta(0) > 0$. And, so each solution may be written in this form. In case $\eta(0) < 0$ the situation is the same for $-\eta(t)$

Example 3. The solution of the inhomogenous linear stochastic equation

$$d\eta(t) = B(t)\ \eta(t)dt + F(t)dw(t)$$

may be written in the form

$$\eta(t) = \exp\left\{\int_0^t B(s)\,ds\right\}\left[\eta(0) + \int_0^t \exp\left\{-\int_0^s B(n)\,dn\right\} F(s)dw(s)\right]$$

To prove this let

$$\xi(t) = \eta_0(t)\,\eta(t),$$

where

$$\eta_0(t) = \exp\left\{-\int_0^t B(s)ds\right\}$$

It is easy to calculate that

$$d\eta_0(t) = -\eta(t)B(t)dt$$

and

$$d\xi(t) = \eta(t)\cdot d\eta_0(t) + \eta_0(t)d\eta(t) = \eta_0(t)F(t)dn(t)$$

From here we get

$$\xi(t) = \xi(0) + \int_0^t F(s)\eta_0(s)dw(s) = \xi(0) + \int_0^t F(s)\exp\left[-\int_0^t B(n)dn\right]dw(s),$$

and finally

$$\eta(t) = \frac{\xi(t)}{\eta_0(t)} = \exp\left\{\int_0^t B(s)ds\right\}\left\{\eta(0) + \int_0^t F(s)\exp\left[-\int_0^s B(u)du\right]dw(s)\right\}$$

Specially the one dimensional elementary Gaussian process (where $B(t) = -\lambda$ = const. and $F(t) = 1$ has the form (see remark 5.))

$$\eta(t) = e^{-\lambda t}\left\{\eta(0) + \int_0^t e^{\lambda s} dw(s)\right\}$$

Exercises

1. Compute the correlation function of a one dimensional Gaussian Markov process starting from the origin.
(Hint: use the representation $\xi(t) = \int_0^t e^{as} dw(s)$.)

2. Prove that if $\xi(t)$ is a one dimensional stationary Gaussian Markov process with parameters $a < 0$, and $b_w > 0$, then the process

$$w(t) = \sqrt{\frac{-2at}{b_w}}\ \xi\left(\frac{-1}{2a} \lg t\right)$$

is a standard Wiener process.
(Hint: compute the correlation function of $w(t)$.)

3. In the same way as in examples 2 - 3 prove that the stochastic differential equation

$$d\eta(t) = [A(t) + B(t)\eta(t)]dt + [F(t) + D(t)\eta(t)]dw(t)$$

has the solution

$$\eta(t) = \exp\left\{\int_0^t [B(s) - \tfrac{1}{2}D^2(s)]ds + \int_0^t D(s)dw(s)\right\}\Big[\eta(0) + \int_0^t \exp\Big\{-\int_0^s [B(u) - \tfrac{1}{2}D^2(u)]du - \int_0^s D(u)dw(u)\Big\}[A(s) - F(s)D(s)]ds + \int_0^t \exp\Big\{-\int_0^s [B(u) - \tfrac{1}{2}D^2(u)]du - \int_0^s D(u)dw(u)\Big\}F(s)dw(s)\Big].$$

4. For the multidimensional case prove that the process $\underline{\xi}(t)$ with differential

$$d\underline{\xi}(t) = B(t)\,\underline{\xi}(t)\,dt + d\underline{w}(t)$$

has the explicit form

$$\underline{\xi}(t) = \exp\left\{\int_0^t B(s)ds\right\}\left[\underline{\xi}(0) + \int_0^t \exp\left\{-\int_0^s B(u)du\right\} d\underline{w}(s)\right].$$

5. Let $\underline{\xi}(t)$ be - non necessarily stationary solution of (1.5) Prove that its mean value vector function $\underline{m}(t)$ satisfies the equation

$$\underline{m}'(t) = A\,\underline{m}(t)$$

and its variance-matrix function $R(t)$ satisfies the equation

$$R'(t) = A^* R(t) + R(t)A + B_W$$

6. Prove that the only homogeneous probability density function which describes a continuous Gaussian Markovian process has the form, for $t > s$,

$$p(y,t \mid x,s) = \frac{1}{\sqrt{2\pi(1-\varrho^2)\sigma^2}} \exp\left\{-\frac{(y-m-\varrho(x-m))^2}{2(1-\varrho^2)\sigma^2}\right\}, \qquad \varrho = e^{-A(t-s)}$$

where σ, m, A, are constants. This means that the process $\underline{\xi}(t)$ is the solution of the differential equation

$d\xi(t) = -A\xi(t) + A\,m\,dt + dw(t)$ (See Krámli [1])

7. Prove relation (1.6) assuming $\underline{\xi}(t)$ is a continuous solution of (1.5).

(Hint: Multiplying (1.5) by $\underline{\xi}^*(t)$ and taking the expectation we get

$$(*)\quad B(dt) = B(0) + A\,B(0)\,dt .$$

Using the fact that $E\,\underline{\xi}(t+dt)\,dw^*(t) = B_w\,dt$ multiplying the transposition of (1.5) by $\underline{\xi}(t+dt)$ and taking the expectation we get

$$(**)\quad B(0) - B(dt) = B(dt) + B_w\,dt ,$$

$(*)$ and $(**)$ prove the statement.)

8. Prove Theorem 1.3 by the differential equation (1.5) and the integral representation:

$$\underline{\xi}(t) = \underline{\xi}(\tau) + A\int_\tau^t \underline{\xi}(s)\,ds + \underline{w}(t) - w(\tau), \qquad 0 \le \tau \le t,$$

where $\underline{\xi}(\tau)$ is $\mathcal{F}_{-\infty}^{\tau}$ measurable. (Hint: Stationarity follows from

$$B(t,\tau) = E(\underline{\xi}(t), \xi^*(\tau)) = B(\tau,\tau) + A\int_\tau^t B(s,\tau)\,ds ,$$

with the only continuous solution

$$B(t,\tau) = e^{A(t-\tau)} B(\tau,\tau) .$$

Markovity is the consequence of

$$E(\underline{\xi}(t) \mid \mathcal{F}_{-\infty}^{\tau}) = \underline{\xi}(\tau) + \int_\tau^t E(\underline{\xi}(s) \mid \mathcal{F}_{-\infty}^{\tau})\,ds$$

with the solution

$$E(\underline{\xi}(t) \mid \mathcal{F}_\tau) = e^{A(t-\tau)}\,\underline{\xi}(\tau) .)$$

9. Prove theorem 1.4 using Levy's theorem (Part I. Ch.8.)

lo.Prove theorem 1.5 by the martingal convergence theorem (see Doob's paper [1]).

11.Prove directly that the solution of 1.5 is a diffusion type process (see exercise 3 in I Ch. 12.).

2.§ Radon-Nikodym derivatives with respect to Wiener measure.

In the statistics of elementary Gaussian processes - similarly to the statistics of independent observations - the maximum-likelihood principle has an important role. For this purpose it is desirable to determine the Radon-Nikodym derivative of the measure generated by this process with respect to some standard measure. Theorem 1.6 suggests that the elementary Gaussian processes with common matrix of diffusion generate equivalent measures, and these measures are equivalent to the Wiener-measure with the same local variance matrix. Theorem 2.1 expresses our heuristic argument in an exact form. Before this we introduce some notations.

Let $\mathcal{H}_k$ be the metric space of k-dimensional vector-valued continuous functions on the interval $[0,T]$ with the uniform metric$\| \|$. It will be convenient to assume $\mathcal{H}_k$ as a direct product of the space $C_k[0,T]$ of k-dimensional continuous functions $\underline{x}=\{\underline{x}(t), 0\le t\le T\}$ with the initial condition $\underline{x}(0)=\underline{0}$ and the k-dimensional Euclidean space R^k.

For the sake of generality we shall consider a Gaussian Markov process $\underline{\xi}(t)$ satisfying the stochastic differential equation (1.5) and having $f(x(0))$ as initial probability density function. Let μ be the probability measure on $\mathcal{H}_k$ generated by the above process $\underline{\xi}(t)$ and ν be the "conditional" product of the k-dimensional Lebesque-measure and the measure generated by the Wiener process on the right hand side of (1.5).

Theorem 2.1 The measures μ and ν are equivalent and their Radon-Nikodym derivative has the form

$$(2.1)\quad \frac{d\mu}{d\nu}(\underline{x}) = f(\underline{x}(0)) \exp\left\{\int_0^T (C\underline{x}(t), d\underline{x}(t)) - \frac{1}{2}\int_0^T (A\underline{x}(t), C\underline{x}(t))dt\right\}$$

where $C = B_W^{-1}A$.

The value of stochastic integral $\int_0^T (\underline{x}(t), d\underline{x}(t))$ can be determined for almost every Wiener sample path $\underline{w}(t)$. The symbol $\int_0^T (C\underline{x}(t), d\underline{x}(t))$ means this value; so formula (2.1) gives ν-almost everywhere the Radon-Nikodym derivative.

Proof. The proof is based on a variant of the invariance principle due to Prohorov [1], which will be cited in the course of the proof (see Arató [5] or Krámli-Pergel [2]).

First let us notice that the conditional measures $\bar{\mu}$ and $\bar{\nu}$ generated by processes $\underline{\xi}(t)$ and $\underline{w}(t)$ on the space $C_k^{(x)}(0,T)$ under the conditions $\underline{\xi}(0) = \underline{x}$, $w(0) = x$, may be treated in a simpler way. Our theorem will follow from the statement about these conditional measures:

The measures $\bar{\mu}$ and $\bar{\nu}$ are equivalent, and

$$(2.2)\quad \frac{d\bar{\mu}}{d\bar{\nu}}(\underline{x}(t)) = \exp\left\{\int_0^T (C\underline{x}(t), d\underline{x}(t)) - \frac{1}{2}\int_0^T (A\underline{x}(t), C\underline{x}(t))dt\right\}$$

Let $\{d^n\}$ be a sequence of divisions $0 = t_1^{(d_n)}, \dots, t_\ell^{(d_n)} = T$ of the interval $[0,T]$. Let us suppose that for $n > m$ d_n refines d_m. Introduce now the new stochastic process $\xi^{d_n}(t)$ recursively as follows:

$$\underline{\xi}^{dn}(0) = \underline{w}(0),$$

(2.3)
$$\underline{\xi}^{dn}(t) = \underline{\xi}^{dn}(t_{i-1}^{dn}) + A\,\underline{\xi}^{dn}(t_{i-1}^{dn})(t - t_{i-1}^{dn}) + \underline{w}(t) - \underline{w}(t_{i-1}^{dn}) \quad \text{if} \quad t_i \geq t > t_{i-1}$$

The process $\underline{\xi}^{dn}(t)$ is the so called Euler approximation of the process $\underline{\xi}(t)$. The sequence $\{\underline{\xi}^{dn}(t)\}$ $(n = 1, 2 \ldots)$ has two basic properties.

(i) Let $\Theta_1, \ldots, \Theta_\ell$ be a finite set of time points. The joint conditional distribution of random vectors $\underline{\xi}^{dn}(\Theta_1), \ldots, \underline{\xi}^{dn}(\Theta_\ell)$ under the condition $\underline{\xi}^{dn}(0) = \underline{x}$ tends to the corresponding distribution of $\underline{\xi}(\Theta_1), \ldots, \underline{\xi}(\Theta_\ell)$.

(ii) The formula (2.3) can be understood as a transformation Φ of the space $C_k^x[0, T]$ into itself: for every Wiener sample path corresponds a sample path of the process $\underline{\xi}(t)$. If K is a compactum of $C_k^x[0, T]$, then $\Phi(K)$ is a compactum too.

<u>Proof.</u> As the processes are Gaussian, for the proof of property (i) it is sufficient to show, that the conditional mean value vector and covariance matrix functions of the processes $\underline{\xi}^{dn}(t)$ under the conditions $\underline{\xi}^{dn}(0) = \underline{x}$ tend to the corresponding functions of $\underline{\xi}(t)$. This can be obtained by direct calculations.

For the proof of property (ii) we have to show - on the basis of the theorem of Arzela - the uniform boundedness and the equicontinuity of functions defined by (2.3) under the condition that the functions on the right hand side of (2.3) have these properties. The uniform boundedness follows from the inequality $\|\Phi(x(t))\| \leq \|x(t)\| + e^{\|A\|(T+\|x(t)\|)}$. Taking into account this the equicontinuity follows from the equicontinuity of functions $\underline{x}(t) \in K$.

From Levy's theorem on the modulus of continuity of the Wiener process (see exercise 6 in Part I Ch. 3) and from property (ii) we can derive the fundamental property.

(iii) For every $\varepsilon > 0$ there exists a compactum K_ε of the space $C_k^x[0,T]$ such that $\bar{\mu}_n(K_\varepsilon) > 1-\varepsilon$, where $\bar{\mu}_n$ is the conditional measure generated by the process $\underline{\xi}^{dn}(t)$ under the condition $\underline{\xi}^{dn}(0) = \underline{x}$.

<u>Proof.</u> From Levy's theorem follows the existence of such $K_\varepsilon^{?}$ for the Wiener process. Set $K_\varepsilon = \Phi(K_\varepsilon^{?})$.

The properties (i) and (iii) give by a variant of Prohorov's theorem (see Gikhman-Skorokhod [1] ch. IX. § 1.) a necessary and sufficient condition for the weak convergence of the measures $\bar{\mu}_n$ to $\bar{\mu}$. In other terms these properties provide that for every bounded continuous functional $f(\underline{x}(t))$ on $C_k^x[0,T]$ $\int_{C_k^x[0,T]} f(\underline{x}(t))\, d\bar{\mu}_n \longrightarrow \int_{C_k^x[0,T]} f(\underline{x}(t))\, d\mu$.

As we have collected all the necessary preliminaries to the proof of theorem 2.1, we can begin the proper calculation of the Radon-Nikodym derivative of measure $\bar{\mu}_n$ with respect to $\bar{\nu}$. The first step for this purpose is the following lemma.

<u>Lemma 2.1</u> The measure $\bar{\mu}_n$ is absolutely continuous with respect to $\bar{\nu}$ and

$$(2.4)\quad \frac{d\bar{\mu}_n}{d\bar{\nu}} = p_n(\underline{x}(t)) = \exp\left\{\sum_{j=1}^{m}(C\underline{x}_{j-1}, \Delta\underline{x}_j) - \frac{1}{2}\sum_{j=1}^{m}[A\underline{x}_{j-1}, C\underline{x}_{j-1}]\Delta t_j^{dn}\right\}$$

where $\Delta\underline{x}_j = \underline{x}(t_j^{dn}) - \underline{x}(t_{j-1}^{dn})$ **and** $\Delta t_j^{dn} = t_j^{dn} - t_{j-1}^{dn}$ **.**

<u>Proof.</u> Let $d^{n'}$ **be a refinement of** d^n **i.e.:**

$$0 = t_0^{dn} = t_0^{dn'} < t_1^{dn'} < \dots < t_{i_1}^{dn'} = t_1^{dn} < \dots < t_{i_m}^{dn} = t_m^{dn} = T.$$

By a direct calculation we get

$$p_{n,n'}(\underline{x}(t)) = \exp\frac{1}{2}\left\{-\sum_{j=1}^{m}\sum_{i>i_{j-1}}^{i_j}\frac{(B_w^{-1}\Delta\underline{x}_i^{dn'} - C\underline{x}_{j-1}^{dn}\Delta t_i^{dn'}, \Delta\underline{x}_i^{dn'} - A\underline{x}_{j-1}^{dn}\Delta t_i^{dn'}}{\Delta t_i^{dn}} + \right.$$

$$(2.5)\qquad \left. + \sum_{i=1}^{i_m}\frac{(B_w^{-1}\Delta\underline{x}_i^{dn'}, \Delta\underline{x}_i^{dn'}}{\Delta t_i^{dn}}\right\}, \Delta t_i^{dn'} = t_i^{dn'} - t_{i-1}^{dn'}, \Delta\underline{x}_i^{dn'} = \underline{x}(t_i^{dn}) - \underline{x}(t_{i-1}^{dn'}), \quad \underline{x}_j^{dn} = \underline{x}(t_j^{dn})$$

for the ratio of joint density functions of random vectors $\underline{\xi}(t_1^{dn'}), \dots, \underline{\xi}(t_{i_m}^{dn'})$ **and** $\underline{w}(t_1^{dn'}), \dots, \underline{w}(t_{i_m}^{dn'})$.
Letting n' tend to ∞ **(2.5) turns into (2.4) for almost every** $\underline{x}(t)$ **. As** $p_n(\underline{x}(t))$ **is the ratio of density functions of random vectors** $\underline{\xi}(t_1^{dn}), \dots, \underline{\xi}(t_m^{dn})$ **and** $\underline{w}(t_1^{dn}), \dots, w(t_m^{dn})$.
Therefore we can apply the martingale convergence theorem (theorem 7 of Part I.Ch lo.) to the sequence $p_{,n'}(\underline{x}(t))$, **which completes the proof of lemma 2.1 .**

As the terms in the exponent of formula (2.4) tend to the integrals $\int_0^T(C\underline{x}(t), d\underline{x}(t))$ **and** $-\frac{1}{2}\int_0^T(A\underline{x}(t), C\underline{x}(t)dt$ **in mean square norm, we can choose a subsequence** d^{n_ℓ} **of sequence** d^n **in such a way that the limit** $p(\underline{x}(t)) = \lim_{\ell=\infty} p_{n_\ell}(\underline{x}(t))$ **exists for a.e.** $\underline{x}(t)$.

Let us consider the compactum K_ε **such that**

(2.6) $$\bar{\mu}_n(K_\varepsilon) = \int_{K_\varepsilon} p_n(\underline{x}(t))\,d\bar{\nu} \geq 1-\varepsilon$$

for every n. As the elements of K_ε are uniformly bounded functions $\frac{1}{2}\left|\sum_{j=1}^{m}(A\underline{x}_{j-1}, C\underline{x}_{j-1})\Delta t_j^{dn}\right| \leq \frac{1}{2}N_\varepsilon^2\ \|A\|\|C\|$ (where N_ε is the common upper bound for the norms of $x(t)\in K_\varepsilon$). So we have $[p_n(\underline{x}(t))]^2 \leq e^{2N_\varepsilon^2\|A\|\|C\|}\,p_n^{(2)}(\underline{x}(t))$ for every $\underline{x}(t) \in K_\varepsilon$. ($p_n^{(2)}$ means the probability density function obtained in the same way for the process $\underline{\xi}^{(2)}(t)$ with parameters $2A$ and $B_{\underline{w}}$). From this inequality (and theorem 3 of Part I. Ch lo.) we can deduce the uniform integrability of the sequence $p_n(\underline{x}(t))$ on the compactum K_ε with respect to the measure $\bar{\nu}$. /We notice that the uniform integrability is valid on the whole space $C_k^x[0,T]$ but its verification is not so simple as on the compact subsets of $C_k^x[0,T]$; this is the advantage of the application of Prohorov's theorem.

From (2.6) on the basis of Fatou's theorem we get

$$\int_{K_\varepsilon} p(\underline{x}(t))\,d\bar{\nu} \geq 1-\varepsilon .$$

Let f^{K_ε} be a non-negative bounded continuous functional on $C_k^x[0,T]$. Also by Fatou's theorem we get

(2.7) $$\int_{C_k^x[0,T]} f(\underline{x}(t))p(\underline{x}(t))\,d\bar{\nu} \leq \lim_{\ell\to\infty}\int_{C_k^x[x,T]} f(\underline{x}(t))p_{n_\ell}(\underline{x}(t))\,dt .$$

Using (2.6) and (2.7) and the uniform integrability of the sequence $p_{n_\ell}(\underline{x}(t))$ on K_ε we obtain

$$\overline{\lim_{\ell\to\infty}} \int_{C_k^x[0,T]} f(\underline{x}(t))p_n(\underline{x}(t))\,d\bar{\nu} \le \overline{\lim_{\ell\to\infty}} \int_{K_\varepsilon} f(\underline{x}(t))p_{n_k}(\underline{x}(t))\,d\bar{\nu} + \varepsilon' =$$

(2.8)

$$= \int_{K_\varepsilon} f(\underline{x}(t))\, p((\underline{x}(t))\,d\bar{\nu} + \varepsilon' \le \int_{C_k^x[0,T]} f(\underline{x}(t))\,p(\underline{x}(t))\,d\bar{\nu} + \varepsilon'$$

where ε' is $\max\limits_{x(t)\in K_\varepsilon} |f(\underline{x}(t))|$.

Analogous considerations are valid for negative functionals too. So relations (2.7) and (2.8) involve

$$2.9 \quad \lim_{k\to\infty} \int_{C_k^x[0,T]} f(\underline{x}(t))\, p_{k_\ell}(\underline{x}(t))\,d\bar{\nu} = \int_{C_k^x[0,T]} f(\underline{x}(t))\,p(\underline{x}(t))\,d\bar{\nu}$$

for arbitrary continuous, bounded functional $f(\underline{x}(t))$ i.e. the measure $\mu(\cdot) = \int_{(\cdot)} p(\underline{x}(t))\,d\bar{\nu}$ is the weak limit of measures μ_n. On the other hand - as we have mentioned - from properties (i) and (iii) follows that the sequence μ_n has the weak limit $\bar{\mu}$. But a sequence of measures has no two different weak limits, so the measure $\tilde{\mu}$ generated by the density function $p(\underline{x}(t))$ coincides with $\bar{\mu}$. The equivalence of measures $\bar{\mu}$ and $\bar{\nu}$ follows from the fact that the stochastic integral $\int_0^T \underline{w}(t)\,d\underline{w}(t)$ is finite with probability 1.

Remark 1. In real applications we observe a trajectory of the process $\underline{\xi}(t)$. But, by the just proved equivalence of measures $\bar{\mu}$ and $\bar{\nu}$ the value of $\int C\underline{x}(t)\,d\underline{x}(t)$ for $\bar{\mu}$ ($\bar{\nu}$) almost every trajectory does not depend on the regarded measure on $\mathcal{H}_k$, as it is defined as an a.e. limit of a sequence of measurable functions on $\mathcal{H}_k$. In the literature there is often used the following formula for

$$\frac{d\bar{\nu}}{d\bar{\mu}} = \left(\frac{d\bar{\mu}}{d\bar{\nu}}\right)^{-1},$$

$$\frac{d\bar{\nu}}{d\bar{\mu}}(\underline{\xi}) = \exp\left\{-\int_0^T (C\underline{\xi}(t), d\underline{\xi}(t)) + \frac{1}{2}\int_0^T (A\underline{\xi}(t), C\underline{\xi}(t))dt\right.,$$

correctness of which is garanteed by the above remark.

Remark 2. The proof of theorem 2.1. may be carried out in the case when $\underline{\xi} = \{\underline{\xi}(t), \mathcal{F}_t\}$ is a diffusion type process, i.e.

$$\underline{\xi}(t) = \int_0^t \underline{\alpha}(s, \underline{\xi})ds + \underline{w}(t)$$

where $\underline{\alpha}(t, \underline{\xi})$ is a measurable vector functional not depending on the future and

$$P\left\{\int_0^T |\underline{\alpha}(t, \underline{\xi})|\, dt < \infty\right\} = 1$$

(see Lipcer - Shiryayev's book [1] or Benczúr - Szeidl [1]).

The concrete formulas - suitable for computational purposes are given in the following exercises.

Exercises.

1. Prove that in the one dimensional stationary case, when

$$d\underline{\xi}(t) = -\lambda\ \xi(t)dt + dw(t)$$

$$Ew(t) = 0,\ E(dw^2) = \sigma dt$$

$$E\underline{\xi}(0) = 0,\ E\underline{\xi}^2(0) = \frac{\sigma^2}{2\lambda}$$

$$\frac{d\mu}{d\nu}(x) = \sqrt{\frac{\lambda}{\pi}}\ \frac{1}{\sigma} \exp\left\{-\frac{\lambda^2}{2\sigma^2}\int_0^1 x^2(t)\,dt + \frac{\lambda T}{2} - \frac{\lambda}{2\sigma^2}[x^2(T)+x^2(0)]\right\}.$$

Obtain this formula from the ratio of probability density functions of Gaussian vectors $\left\{\xi\left(\frac{k}{n}\right)\right\}, \left\{w\left(\frac{k}{n}\right)\right\}$ $k=1,\dots,n$ letting n tend to infinity. This ratio can be calculated using the relation

$$\xi\left(\frac{k}{n}\right) = e^{-\frac{\lambda}{n}}\,\xi\left(\frac{k-1}{n}\right) + \frac{\varepsilon(k)}{\sqrt{n}} \qquad (D^2\varepsilon(k) = 1)$$

following from theorem 1.5. (See Arató [5] or Striebel [1])

2. Let $\xi(t)$ be as above, and

$\xi_m(t) = \xi(t) + m$. Let μ and μ_m the probability measures generated by processes $\xi(t)$ and $\xi_m(t)$ on $\mathcal{H}_1$ respectively. Prove that (see Grenander [1])

$$\frac{d\mu_m}{d\mu} = \exp\left\{-\frac{\lambda m}{\sigma^2}\left[x(0) + x(T) + \lambda\int_0^T x(t)\,dt + m\left(1+\frac{\lambda T}{2}\right)\right]\right\}$$

(Hint: Let ν_m be the direct product of the Lebesque measure and the measure generated by the Wiener-process $w(t)+m$. Notice that ν_m and ν coincide on $\mathcal{H}_1$ (i.e. $\frac{d\nu_m}{d\nu} = 1$) , and use the "chain-rule"

$$\frac{d\mu_m}{d\mu} = \frac{d\mu_m}{d\nu_m}\cdot\frac{d\nu_m}{d\nu}\cdot\frac{d\nu}{d\mu}.)$$

3. Prove that in the two-dimensional case, when

$$A = \begin{pmatrix} -\lambda & -\omega \\ \omega & -\lambda \end{pmatrix}, \qquad E(dw_i)^2 = \sigma^2 dt \qquad (i=1,2)$$

then

$$f_A(\underline{x}(0)) = \frac{\lambda}{\pi\sigma^2} \exp\left\{-\frac{\lambda}{\sigma^2} x_1^2(0) - \frac{\lambda}{\sigma^2} x_2^2(0)\right\}.$$

and

$$\frac{d\mu}{d\nu} = \frac{\lambda}{\pi\sigma^2} \exp\left\{-\frac{\lambda^2+\omega^2}{2\sigma^2}\int_0^T [x_1^2(t) + x_2^2(t)]dt - \frac{\lambda}{2\sigma^2}[x_1^2(T) + x_2^2(T) + \right.$$

$$\left. + x_1^2(0) + x_2^2(0)] + \lambda T + \int_0^T [x_1(t)dx_2(t) - x_2(t)dx_1(t)]\right\}.$$

4. In the previous example let we take the complex valued process $x(t) = |x(t)|e^{i\theta(t)}$ **where** $x(t) = x_1(t) + ix_2(t)$ $[\underline{x}(t)|^2 = x_1^2(t) + x_2^2(t)$. **Prove that**

$$\int_0^t [x_1(t)dx_2(t) - x_2(t)dx_1(t)] = \int_0^T |x(t)|^2 d\theta$$

and $$\frac{d\mu}{d\nu} = \frac{\lambda}{\pi\sigma^2}\exp\left\{-\frac{\lambda^2+\omega^2}{2\sigma^2}\int_0^T |x(t)|^2 dt + \frac{\omega}{\sigma^2}\int_0^T |x(t)|^2 d\theta + \lambda T - \right.$$

$$\left. - \frac{\lambda}{2\sigma^2}[|x(T)|^2 + |x(0)|^2]\right\}.$$

(**Hint. Use the relations**

$$\sum_j [x(t_j)\overline{x(t_{j-1})} - x(t_{j-1})\overline{x(t_j)}] =$$

$$= -2\sum_j [x_2(t_j)(x_1(t_j) - x_1(t_{j-1})) - x_1(t_j)(x_2(t_j) - x_2(t_{j-1}))],$$

and

$$\sum_j |x(t_j)||x(t_{j-1})|[e^{i(\theta(t_j)-\theta(t_{j-1}))} - e^{i(\theta(t_{j-1})-\theta(t_j))}] =$$

$$= \sum_j |x(t_j)||x(t_{j-1})| 2i\sin(\theta(t_j) - \theta(t_{j-1})) \sim 2i\sum_j |x(t_j)|^2(\theta(t_j) - \theta(t_{j-1})).$$

For further details see Arató [5] , Arató - Kolmogorov - Sinay [1] . The following exercises are concerned to the calculation and the asymptotic behaviour of the maximum-likelihood estimator for the matrix Q of the k-dimensional discrete time autoregressive process.

5. Assume that the components $\varepsilon_1(n), \varepsilon_2(n), \dots$ of the right hand side process of equation (1.1) are independent with dispersion (j=1, ... , k).

α. Prove that the joint conditional probability density function of random vectors $\underline{\xi}(1), \dots, \underline{\xi}(N)$ - under the condition $\underline{\xi}(0) = \underline{x}(0)$ has the following form

$$p(\underline{x}(1),\dots,\underline{x}(N) | \underline{\xi}(0) = \underline{x}(0)) = (2\pi)^{-\frac{Nk}{2}} \left(\prod_{i=0}^{k-1} \sigma_i\right)^{-N} \exp\left\{-\frac{1}{2} \cdot \sum_{i=0}^{k-1} \frac{1}{\sigma_i^2} \sum_{j=0}^{N-1} \left[x_i(j+1) - q_{i,0}\, x_0(j) + q_{i,1}\, x_1(j) + \dots + q_{i,k-1}\, x_{k-1}(j)\right]^2\right\}. \tag{2.10}$$

β. The conditional likelihood equations for parameters $q_{i,j}$ has the form:

$$\sum_{j=0}^{N-1} [\xi_i(j+1) - (q_{i,0}\, \xi_0(j) + \dots + q_{i,k-1}\, \xi_{k-1}(j))]\, \xi_0(j) = 0, \tag{2.11}$$

$$\vdots$$

$$\sum_{j=0}^{N-1} [\xi_i(j+1) - (q_{i,0}\, \xi_0(j) + \dots + q_{i,k-1}\, \xi_{k-1}(j))]\, \xi_{k-1}(j) = 0,$$

where $i = 0, \dots, k-1$.

6. Denote by $\hat{q}_{i,j}$ the solution of equation (2.11). Calculate the elements of $k^2 \times k^2$ covariance matrix of the random variables $\hat{q}_{i,j} - q_{i,j}$.

(Hint: Define the random variables

$\eta_{i,\ell}(N)$ as follows

$$\eta_{i,\ell}(N) = \frac{1}{\sqrt{N}}(\hat{q}_{i,0} - q_{i,0})\sum_{j=0}^{N-1} \xi_0(j)\xi_\ell(j) + \dots + \frac{1}{\sqrt{N}}(\hat{q}_{i,k-1} - q_{i,k-1})\sum_{j=0}^{N-1} \xi_{k-1}(j)\xi_\ell(j).$$

Express them by $\hat{q}_{i,j} - q_{i,j}$, and prove that their covariance matrix has the form

$$\begin{pmatrix} \sigma_1^2 B(0) & & 0 \\ & \ddots & \\ 0 & & \sigma_{k-1}^2 B(0) \end{pmatrix}, \text{ i.e. } \quad \begin{array}{l} E\,\eta_{i,\ell_1}(N)\eta_{i,\ell_2}(N) = \sigma_1^2 b_{\ell_1 \ell_2}, \\ E\,\eta_{i,\ell_1}(N),\eta_{j,\ell_2}(N) = 0 \text{ for } i \neq j \end{array}$$

where $B(0)$ satisfy the equation (1.2).

Remark.

From the ergodicity of the process $\underline{\xi}(n)$ follows the asymptotic efficiency of the conditional maximum-likelihood estimator. The strong mixing property, provide its asymptotic normality. This theorem was proved first by Mann-Wald [1].

7. On the basis (2.2) it is possible to get estimate for the unknown matrix $A(B_w$ is known). The method of least spuares minimizes the functional

$$\int_0^T (B_w^{-2} A\,\underline{\xi}(s), d\underline{\xi}(s)) - \frac{1}{2}\int_0^T (A\underline{\xi}(s), B_w^{-1} A\underline{\xi}(s))\,ds =$$

$$= \frac{1}{2}\int_0^T (B_w^{-1} A \underline{\xi}(s), A\underline{\xi}(s))\, ds + \int_0^T (B_w^{-1} A \underline{\xi}(s), d\underline{w}(s)).$$

Prove that the solution $(\hat{a}_{pq})$ of the linear system of equations

$$\frac{\partial}{\partial a_{pq}} \left\{ \int_0^T (B_w^{-1} A \underline{\xi}(s), A\underline{\xi}(s))\, ds + \int_0^T (A\underline{\xi}(s), B_w^{-1} d\underline{w}(s)) \right\} = 0$$

$$p, q = 0, 1, \dots, k-1, \quad (B_w^{-1} = (b_{pq}^{-1}))$$

gives the minimum, and

$$(*)\ \frac{1}{T}\int_0^T \xi_q(s) \sum_{j=0}^{k-1} b_{pj}^{-1} \sum_{i=0}^{k-1} \sqrt{T}(\hat{a}_{ji} - a_{ji})\, \xi_i(s)\, ds =$$

$$= \frac{1}{\sqrt{T}} \int_0^T \xi_q(s) \sum_j b_{pj}^{-1}\, dw_j(s) = \eta_{pq}(T), \quad (p, q = 0, \dots, k-1),$$

where

$$E\,\eta_{pq}(T) = 0,\ E\,\eta_{pq}(T)\eta_{rs}(T) = \frac{1}{T}\int_0^T b_{rp}^{-1} E\,\xi_q(t)\xi_s(t)\, dt.$$

As the elementary Gaussian process is ergodic

$$\frac{1}{T} b_{rp}^{-1} \int_0^T \xi_q(t)\xi_i(t)\, dt \longrightarrow b_{rp}^{-1} b_{qi}^{\xi} \qquad , \text{ if } T \longrightarrow \infty$$

and $\eta_{pq}(T)$ is asymptotically (if $T \longrightarrow \infty$) normally distributed (see J. Rozanov's book [1], Taraskin [1]). Prove that the random variables $\sqrt{T}(\hat{a}_{pq} - a_{pq})$ are asymptotically normally distributed with $\underline{0}$ mean and covariance matrix $B^{-1} = \{b_{rp}^{-1}\, b_{qi}^{\xi}\}^{-1}$ (see Arató [4], Pisarenko [1]).

3.§ Autoregressive-moving average processes.

Definition 1. We call a stationary Gaussian process $\xi(n)$ with discrete time an autoregressive moving average /ARMA/ process if it satisfies the equation

$$(3.1) \quad \xi(n) = \sum_{i=1}^{p} a_i \, \xi(n-i) + \sum_{i=1}^{q} b_i \, \varepsilon(n-i) + \varepsilon(n)$$

where $\{\varepsilon(n)\}$ is a sequence of i.i.d. Gaussian random variables, and $\varepsilon(n)$ is independent of $\mathcal{F}_{-\infty}^{n-1}(\xi)$.
In case $b_i = 0 \ (i \geq 1)$ the process is an autoregressive one and in case $a_i = 0 \ (i \geq 1)$ the process is a moving average one.

Theorem 3.1. Equation (1) has a unique stationary solution if and only if no one of the roots of the characteristic polynomial
$P_1(z) = z^p - \sum_{i=1}^{p} a_i z^{p-i}$ lies outside the unit circle $(|z| \geq 1)$.
In this case $\xi(n)$ is the first component of a k=max $\{p, q+1\}$ dimensional stationary Gaussian Markov process.

$$\xi(t) = \left\{ \xi^{(1)}(t), \dots, \xi^{(k)}(t) \right\}.$$

Proof. Let us assume that $\xi(n) = \xi^{(1)}(n)$ and consider the system of equations

$$\xi^{(i)}(n) = \xi^{(i+1)}(n-1) + C_{i-1} \, \varepsilon(n) \qquad \text{if} \qquad i \leq 1 < p,$$

$$\xi^{(p)}(n) = \sum_{i=1}^{p} a_{p+1-i} \, \xi^{(i)}(n-1) + \sum_{i=p+1}^{q+1} b_{i-1} \, \xi^{(i)}(n-1) + C_{p-1} \, \varepsilon(n),$$

$$\xi^{(p+1)}(n) = \varepsilon(n),$$

(3.2)

$$\xi^{(p+i)}(n) = \xi^{(p+i-1)}(n-1) \quad \text{if} \quad 1+p < i \leq q+1 .$$

/Naturally in the case of $q<p$ the suitable terms and equations are omitted./

If the constants c_j $(j=0,\dots,(p-1))$ satisfy the equations

$$c_0 = 1$$

(3.3) $$c_1 - a_1 \cdot c_0 = b_1$$

$$\vdots$$

$$c_{p-1} - a_1 c_{p-2} - \dots - a_{p-1} c_0 = b_{p-1},$$

then the system (2) is equivalent to the equation (1) . It is easy to see that the characteristic polynomial $P_2(z)$ of (2) is equal to $P_1(z)$ if $q<p$, and $z^{q+1}P_1(z)$ otherwise. So the system (2) of stochastic difference equations has a unique stationary solution, which is a k-dimensional Gaussian Markov process and its first component will be the unique stationary solution of the equation (1).

Q.E.D.

Remark 1. The solution of the equation (1) can be obtained in a constructive way similarly to the first order autoregressive process (see remark 1§ 1 and exercise 5. in this§).

(3.4) $$\xi(n) = \sum_{k=0}^{\infty} c_k \, \varepsilon(n-k).$$

Proof Indeed, if the coefficients c_k satisfy the infinite recursive system of equations

$$
(3.5)\quad
\begin{aligned}
& c_0 = 1 \\
& c_1 - a_1 \cdot c_0 = b_1 \\
& \vdots \\
& c_k - \sum_{i=1}^{p} a_i \cdot c_{k-i-1} = b_k \qquad \text{if} \qquad k \geq p \\
& \vdots
\end{aligned}
$$

/notice that the first p equations coincide with system (3) /, and $\sum_{k=1}^{\infty} |c_k|^2 < \infty$ then the process (4) is a correctly defined stationary Gaussian process satisfying (1).

As $b_k=0$ for $k > q$ and the roots of characteristic polynomial $P_1(z)$ are inside the unit circle system (5) has a unique solution of desired property.

Remember that a multidimensional k-dimensional Gaussian Markov process $\underline{\xi}(n)$ has the representation $\underline{\xi}(n) = \sum_{i=0}^{\infty} Q^i \underline{\varepsilon}(n-i)$. As the matrix Q satisfies its own characteristic equation:

$$Q^k - \sum_{i=1}^{k} a_i Q^{k-i} = 0$$

, all the elements of $\{Q^n\}$ satisfy a recursive system of equations similar to (5) therefore the components of $\underline{\xi}(n)$ are sums of ARMA processes. Notice that if $\xi(n) = \sum_{k=1}^{\ell} d_k \xi^{(k)}(n)$, where

$$\xi^{(k)}(n) = \sum_{i=1}^{p} a_i \xi(n-1) + \sum_{i=0}^{q} b^{(k)}(i)\, \varepsilon^{(k)}(n-i)$$

and $\{\varepsilon^{(k)}(n)\}$ is a sequence of i.i.d. Gaussian vectors, then $\xi(n)$ is ARMA process. So we get the converse of theorem 1:

<u>Theorem 3.2.</u> Any component of a multidimensional stationary Gaussian Markov process is ARMA process.

In the continuous time case the equation

$$(3.1')\ \xi^{(p)}(t) = \sum_{i=1}^{p} a_i\, \xi^{(p-i)}(t) + \sum_{i=1}^{q} b_i\, w^{(q+2+i)}t + w'(t)$$

would correspond to equation (1) . Before giving an exact meaning to (1') we try to solve it formally. For this purpose we need the following

<u>Lemma 1.</u> If function $f(t)$ is differentiable and

$$\int_0^{\infty} (|f(t)|^2 + |f'(t)|^2)\, dt < \infty, \qquad \text{then}$$

$$(3.6)\quad \int_{-\infty}^{t+h} f(t+h-s)\, dw(s) - \int_{-\infty}^{t} f(t-s)\, dw(s) = \int_t^{t+h} \int_{-\infty}^{\tau} f'(\tau - s)\, dw(s) + f(0)[w(t+h) - w(t)] .$$

The proof can be carried out by changing the order of integration. The relation (6) formally can be considered as a "rule of differentiation":

$$(3.7)\quad [\int_{-\infty}^{t} f(t-s)\, dw(s)] = \int_{-\infty}^{t} f'(t-s)\, dw(s) + f(0)\, w'(t) .$$

We are looking for a solution of (1') in the form $\xi(t) = \int_{-\infty}^{t} f(t-s)\, dw(s)$, suggested by the first order case. If $q < p$, then there exists a unique function $f(t)$ satisfying the homogeneous differential equation

$$(3.8) \quad f^{(p)}(t) - \sum_{i=1}^{p} a_i f^{(p-i)}(t) = 0,$$

and the initial conditions

$$(3.9) \quad \begin{aligned} & f(0) = 1 \\ & f'(0) - a_1 f(0) = b_1, \\ & \vdots \\ & f^{(p-1)}(0) - \sum_{i=1}^{p-1} a_i f^{(p-1-i)}(0) = b_{p-1}, \ (\text{if } i > q \quad b_i = 0). \end{aligned}$$

Using the formal differentiation rule (7) we may convinced that

$$(3.10) \quad \xi(t) = \int_{-\infty}^{t} f(t-s)\, dw(s)$$

is a formal solution of (1').

If the roots of the characteristic polynomial $P_1'(\lambda) = \lambda^p - \sum_{i=1}^{p} a_i \lambda^{p-i}$ has negative real parts, then $\int_0^\infty |f^{(i)}(t)|^2\, dt < \infty$ for every i=0,1,... . In this case the process $\xi(t)$ given by (10) is a correctly defined stationary Gaussian process. We may assume (10) as the definition of continuous time ARMA process. /We notice that for $q \geq p$ (1') has only generalized solution./ For continuous time ARMA processes theorems corresponding to theorems (1) and (2) are valid too:

<u>Theorem 3.3.</u> A continuous time process $\xi(t)$ is ARMA if and only if it is a component of a multidimensional stationary Gaussian process $\xi(t)$.

<u>Proof.</u> The first part of the proof is obvious.
The p-dimensional process $\left\{\xi^{(i)}(t)\right\} = \left\{\int_{-\infty}^{t} f^{(i)}(t-s)\,dw(s)\right\}$ $/i=0,\dots p-1/$ satisfies the system of equations

$$
(3.11)\qquad
\begin{aligned}
d\xi^{(i)} &= \xi^{(i+1)}(t)\,dt + c_i\,dw(t)\,, \qquad i=0,\dots,p-2\,,\\
d\xi^{(p-1)} &= \sum_{i=0}^{p-1} a_{p-i}\,\xi^{(i)}(t)\,dt + c_{p-1}\,dw(t)\,,
\end{aligned}
$$

where $c_i = f^{(i)}(0)$.

The converse assertion can be obtained similarly to the discrete time case, using the integral representation (1.5) of a multidimensional Gaussian Markov process, and the fact that the matrix function e^{At} satisfies the differential equation $\left(e^{At}\right)^{(p)} = \sum_{i=0}^{p-1} a_i \cdot e^{(At)^{(i)}}$ where the coefficients a_i coincide with the coefficients of the characteristic polynom of A.

<u>Remark 1.</u> If we suppose that $q \geq p$ we would have to add further equations to system (11) among them the equation $d\xi^{(p+1)}(t) = dw(t)$, which has no stationary solution. This is the reason of the additional condition $q < p$.

<u>Remark 2.</u> The system of equation (11) has the following visual meaning: an ARMA process $\xi(t)$ is not differentiable in general - but by the addition of a suitable Wiener process it becomes differentiable. This procedure can be continued up to the $(p-1)$-th derivative of $\xi(t)$.

<u>Remark 3.</u> Combining theorems 1., 2. and 3. with Doob's theorem (see Doob's paper [1]) we get that the discrete time

sample process $\xi(n\delta)$ of a continuous time ARMA process $\xi(t)$ is also ARMA. But, the sample process $\xi(n\delta)$ of a pure autoregressive process isn't generally a discrete time pure autoregressive process, because if a matrix A has the form

$$\begin{pmatrix} 0 & 1 & \dots & 0 \\ \cdot & 0 & 1 & \\ \cdot & 0 & 0 \dots & 1 \\ a_1 & \cdot & \cdot & a_p \end{pmatrix}$$

its exponent $e^{A\delta}$ has not the same one.

In this work we have avoided the spectral approach to stationary processes because of the necessity of deep analytic tools. But in some technical applications the spectral density function has a simple visual meaning and it can be easily measured. For this reason we breefly summarize without proofs the basic facts concerning to the ARMA processes. A regular discrete (continuous) time stationary Gaussian process has the representation (see Rozanov's book [1])

(3.12.) $\xi(n) = \int_0^{2\pi} e^{in\varphi} g(\varphi)\, dw(\varphi),$

(3.13.) $\xi(t) = \int_{-\infty}^{\infty} e^{it\lambda} h(\lambda)\, dw(\lambda)$

where $w(\varphi), w(\lambda)$ are standard Wiener processes "random measures" , and functions $g(\varphi)$ resp. $h(\lambda)$ can be analitically continued to the open unit circle resp. upper halfplane. The sequence of i.i.d. Gaussian random variables (resp. the white noise process) corresponds to the identically constant function on the interval $(0, 2\pi)$ (resp. $(-\infty, \infty)$) . Using this fact we can easily find the connection between the "moving-average" representations (4) and (10) and the spectral representations (12) and (13):

$$g(\varphi) = \sum_{n=0}^{\infty} c_n e^{in\varphi},$$
$$h(\lambda) = \int_{-\infty}^{0} f(-s) e^{i\lambda s} ds.$$

Using the formal correspondences

$$\xi(n) \sim g(\varphi) e^{in\varphi}, \qquad \xi(t) \sim h(\lambda) e^{i\lambda t}$$
$$\xi'(t) \sim h(\lambda) i\lambda e^{i\lambda t},$$

$w'(t) \sim e^{i\lambda t}$ we get for ARMA process the correspondencees

$$g(\varphi) = \frac{\sum_{n=0}^{q} b_n e^{-in\varphi}}{\sum_{n=0}^{p} a_n e^{-in\varphi}}, \quad h(\lambda) = \frac{\sum_{n=0}^{q} b_n (i\lambda)^n}{\sum_{n=0}^{p} a_n (i\lambda)^n}.$$

In continuous time case we can see from the form of $h(\lambda)$ that in the case $q \geq p$ the integral of the spectral density function $|h(\lambda)|^2$ would be infinite. By physical reasons such a system can't exist.

Exercises.

1. Prove Theorem 2 in another way:

   Writing the sequence of equations

$$\underline{\xi}(n) = Q\, \underline{\xi}(n-1) + \underline{\varepsilon}(n)$$

$$\underline{\xi}(n-1) = Q\, \underline{\xi}(n-2) + \underline{\varepsilon}(n-1)$$

$$\vdots$$

$$\underline{\xi}(n-k+1) = Q\, \underline{\xi}(n-k) + \underline{\varepsilon}(n-k+1)$$

for the k^2 unknowns $\xi^{(1)}(n), \xi^{(2)}(n), \xi^{(2)}(n-1), \dots, \xi^{(2)}(n-k)$, $\xi^{(3)}(n), \dots, \xi^{(k)}(n), \dots, \xi^{(k)}(n-k)$ we have k^2 equations. Show that they can be solved uniqually if det $(Q) \neq 0$. If Det $(Q)=0$, then the dimension of the elementary Gaussian process $\underline{\xi}(n)$ can be reduced.

2. We say that a k-dimensional process $\underline{\xi}(t)$ with discrete time-parameter is generalized autoregressive one if the equation

$$\underline{\xi}(n) = \sum_{j=1}^{p} A_j \underline{\xi}(n-j) + \underline{\varepsilon}(n) \qquad \text{holds}$$

with i.i.d sequence $\{\underline{\varepsilon}(n)\}$ of nondegenerated Gaussian vectors (see Andel [1]). Prove that equation (3.14.) has a unique solution $\underline{\xi}(n)$ which does not depend on $\underline{\varepsilon}(i)$'s for $i<n$ if and only if the zeros of the polynomial

$$\det \left(-I z^p + \sum_{j=1}^{p} A_j z^{p-j}\right) \qquad \text{are inside}$$

the unit circle.

(Hint: prove that the above zeros are the same as the characteristic roots of the $pk \times pk$ matrix

$$\begin{pmatrix} 0 & I & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & I \\ -I & -A_1 & \cdot & \cdot & \cdot & -A_p \end{pmatrix} .$$

3. Calculate the inverse of the covariance matrix

$$R = [E(\xi(n-i)\cdot\xi(n-j))] \quad \text{for a}$$

p -order discrete time parameter autoregressive process, if $i,j \le N$ and $p<N$. (Hint: consider $\xi(n)$ as a component of a p dimensional elementary Gaussian process and solve the equation for the left upper $p\times p$ minor $\tilde{R}^{-1}$ of R^{-1}

$$\tilde{R}^{-1} = Q^{*}\tilde{R}^{-1}Q - \tilde{R}^{-1}Q^{-1}B_{\varepsilon}\tilde{R}^{-1}Q \quad (\text{where} \quad Q = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \\ -1 & -q_1 & \cdots & -q_p \end{pmatrix}$$

and $B_2 = \begin{pmatrix} 0 & \cdot & 0 \\ 0 & \cdot & \sigma \end{pmatrix})$ derived from equation (1.1). Check the following result by direct calculation

$$R^{-1} = \left\{ r^{(-1)}_{i,j} \right\}, \qquad \text{where}$$

$$r^{(-1)}_{i,j} = \begin{cases} 0 & ,\text{if} \quad |j-i| > p+1 \\ \sum\limits_{\ell=0}^{\min(i,p-j+i)} q_\ell\, q_{\ell+j-i} & ,\text{if} \quad i<j \end{cases}$$

and $r^{(-1)}_{i,j} = r^{(-1)}$, $r^{(-1)}_{i,j} = r^{(-1)}_{N-i,N-j}$, $q_0 = 1$.

(See e.g. Siddiqui [1], Arató [1].)

4. Calculate $B^{-1}(0)$ for continuous time parameter autoregressive process too, where $B(0) = \{E(\xi^{(i)}(0), \xi^{(j)}(0))\}$, $i,j = 1,2,\dots,p$. (See e.g. Hajek [1], Arató [5].)

(Hint: check the following result: $B^{-1}(0) = b^{(-1)}_{i,j}$,

where

$$b_{i,j}^{(-1)} = \begin{cases} 0 , & \text{if} \quad i \equiv j+1 \pmod 2) \\ \frac{2}{\sigma} \sum (-1)^{\ell} a_{i-\ell}\, a_{j+1+\ell} & \text{if} \quad i \equiv j \pmod 2) , \end{cases}$$

$a_p = 1$ **and** $a_i = 0$ **for** $i > p$.)

5. Let $\xi(n)$ **be an autoregressive stationary process with discrete time**

$$\xi(n) = \sum_{i=1}^{p} a_i \xi(n-1) + \varepsilon(n)$$

where all the roots of

$$(*) \qquad P_1(z) = z^p - \sum_{i=1}^{p} a_i z^{p-i} \quad .$$

are inside of the unit circle. Prove the Wold's expansion

$$(**) \qquad \xi(n) = \sum_{\tau=0}^{\infty} c_\tau\, \varepsilon(n-\tau)$$

and calculate the c_τ **coefficients.** [**Hint: There are** p **roots, say** $x_1, \dots, x_p$ **of equation** $(*)$ **, with** $|x_i| < 1$ **, then** ($z_i = 1/x_i$ **and** $|z| < \min |z_i|$)

$$\frac{1}{1-\sum_{i=1}^{p} a_i z^i} = \frac{1}{\prod_{i=1}^{p} \left(1 - \frac{z}{z_i}\right)} = \prod_{i=1}^{p} \sum_{\nu=0}^{\infty} \left(\frac{z}{z_i}\right)^{\nu} = \sum_{\tau=0}^{\infty} c_\tau z^\tau$$

converges absolutely. From this for the coefficients we obtain the equations $(a_0 = 1)$

$$\begin{aligned} & 1 = a_0 c_0 = c_0 \\ (***) \qquad & 0 = a_0 c_1 - a_1 c_0 = c_1 - a_1 \\ & \vdots \\ & 0 = a_0 c_{p-1} - a_1 c_{p-2} - \dots a_{p-1} c_0 \\ (\text{IV}) \qquad & 0 = a_0 c_t - a_1 c_{t-1} - \dots a_p c_{t-p}, \qquad t = p, p+1, \dots \end{aligned}$$

If the roots (*) are different the general solution of the homogeneous difference equation (IV) is

$$c_r = \sum_{i=1}^{p} \tilde{c}_i \, x_i^r$$

The $\tilde{c}_i$ coefficients are determined from the boundary conditions (***) of c_r.]

6. Prove the Wold's expansion in the case of multiple roots in exercise 5.

7. In exercise 5. if $p=1$ we get $c_r = a_1^r$. If $p=2$ and x_1 and x_2 are different prove

$$c_1 = \frac{x_1}{x_1 - x_2} \, , \qquad \tilde{c}_2 = \frac{-x_2}{x_1 - x_2}$$

and

$$c_r = \frac{x_1^{r+1} - x_2^{r+1}}{x_1 - x_2} \, , \quad r = 0, 1, \ldots$$

In this case if the roots are complex, $x_1 = a e^{i\theta}$ and $x_2 = a e^{i\theta}$, then

$$\tilde{c}_1 = \frac{e^{i\theta}}{e^{i\theta} - e^{-i\theta}} \, , \qquad \tilde{c}_2 = - \frac{e^{-i\theta}}{e^{i\theta} - e^{-i\theta}}$$

and

$$c_r = \tilde{c}_1 \, x_1^r + \tilde{c}_2 \, x_2^r = a^r \, \frac{\sin (r+1)\theta}{\sin \theta}$$

8. Let $\xi(t)$ be a 2-order moving average process

$$\xi(t) = \varepsilon(t) + b_1 \, \varepsilon(t-1),$$

where $E\xi(t) = E\varepsilon(t) = 0$ and $E\varepsilon^2(t) = 1$, then

$$\sigma_\xi^2 = E\xi^2(t) = (1+b_1^2), \qquad \varrho = E\xi(t)\,\xi(t-1)\Big/\sigma_\xi^2 = \frac{b_1}{1+b_1^2}.$$

Prove that the density function of random variables $\xi(1),\dots,\xi(N)$ **has the form**

$$(*)\quad p(x_1,\dots,x_N) = \sigma_\xi^{-N}(2\pi)^{-N/2}|B_N|^{-1/2}\exp\left\{-\frac{1}{2\sigma_\xi^2}\sum_{i,j=1}^{N}\bar b_{ij}\,x_i x_j\right\},$$

where $|B_N| = $ **det** B_N **and** $B_N^{-1} = \{\bar b_{ij}\}_{i,j=\overline{1,N}}$ **. The elements**

$$\bar b_{ij} = (-1)^{j-i}\varrho^{j-i}|B_{i-1}||B_{N-j}|\frac{1}{|B_N|}, \quad \text{if} \quad i<j,$$

and

$$|B_i| = \frac{u_1^{i+1}-u_2^{i+1}}{u_1-u_2}$$

where

$$u^2-u+\varrho^2=0, \qquad \text{i.e. } u_1=\frac{1+\sqrt{1-4\varrho^2}}{2},\ u_2=\frac{1-\sqrt{1-4\varrho^2}}{2}.$$

The determinant $|B_N|$ **fulfils the difference equation**

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

From $(*)$ **we get that the sufficient statistics for parameters** (b_1,ϱ) **is the only** $(\xi(1),\dots,\xi(N))$ (**see Arató** [1] **, Shaman** [1] **,** [2]).

9. Let $\xi(t)$ be a second order autoregressive process satisfying the equation

$$d\xi'(t) = (-a_1\xi'(t) - a_0\xi(t))\,dt + dw(t).$$

It was proved (see theorem 3.3) that $\xi(t)$ is the first component of a two dimensional elementary Gaussian process

$\{\xi_1(t), \xi_2(t)\}$ with matrices

$$A = \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}, \quad B_w = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We suppose that the matrix A has complex eigenvalues $\lambda_1 = \lambda + i\omega$, $\lambda_2 = \lambda - i\omega$, where $\lambda = -\frac{a_1}{2}$, $\omega = \sqrt{-(\frac{a_1^2}{4} - a_0)}$. Then from equation $AB(0) + B(0)A^* = -B_W$ we get the following explicit form

$$B(0) = \begin{pmatrix} -\frac{1}{4\lambda(\lambda^2+\omega^2)} & 0 \\ 0 & -\frac{1}{4\lambda} \end{pmatrix}$$

In §1. we proved that $(\xi_1(n\delta), \xi_2(n\delta))$, the discrete time process, is also an elementary Gaussian one. Prove that the matrices Q and B_ε has the following form:

$$Q = e^{A\delta} = \frac{e^{\lambda\delta}}{\omega} \begin{pmatrix} \omega\cos\omega\delta - \lambda\sin\omega\delta & \sin\omega\delta \\ -(\lambda^2+\omega^2)\sin\omega\delta & \omega\cos\omega\delta + \lambda\sin\omega\delta \end{pmatrix},$$

$$B_\varepsilon = -e^{A\delta} B(0) e^{A^*\delta} + B(0).$$

Applications of this description the reader may find in papers Gy. Németh, T.[1] and Mehra, R. K.[2].

[1] Gy.Németh, T.: On estimates of parameters of the second order autoregressive process with continuous time, SzTAKI Közlemények, 10 (1973) 33-43 (in Hungarian).

[2] Mehra, R. K.: Optimal input signals for parameter estimation in dynamic systems - Survey and new results, IEEE Trans. Automatic Control, AC-19 (1974) No.6, 753-768.

4.§. Parametrization of the discrete time autoregressive process by partial correlations

The most natural parameters for autoregressive processes are the coefficients a_i figuring in equation (3.1) and the dispersion σ of the right hand side sequence of i.i.d. Gaussian variables. But the domain $\mathcal{U} \subset R^p$ of admissible parameters is very complicated, although in cases $p = 1, 2, 3$ it is possible to give rather simple criteria which are expressed directly in the coefficients a_i (see exercise 3.)

For this reason in the literature (see e.g. Ramsey's paper [1] or Box-Jenkins [1]) there is often assumed another natural parametrization of the autoregressive processes, namely the partial autocorrelation.

Let π_j denote the j-th partial autocorrelation $(j>0)$, i.e. the conditional correlation between $\xi(n)$ and $\xi(n-j)$ with respect to the σ-algebra $\mathcal{F}_{n-j+1}^{n-1}$ generated by $\xi(n-1)$, $\xi(n-2), \dots, \xi(n-j+1)$.

i.e.

$$\pi_j = \frac{E\left(\xi(n) - E(\xi(n)|\mathcal{F}_{n-j+1}^{n-1})\right)\left(\xi(n-j) - E(\xi(n-j)|\mathcal{F}_{n-j+1}^{n-1})\right)}{E|\xi(n) - E(\xi(n)|\mathcal{F}_{n-j+1}^{n-1})|^2}$$

It follows from the p-order Markov property, that for $j > p$ $\pi_j = 0$.

This parametrization has the advantage that the variation domain $\mathcal{P}$ of $\pi = \{\pi_1, \ldots, \pi_p\}$ is the simple product set $(-1,1) \underbrace{x \ldots x}_{\text{p-times}} (-1,1)$.

The mapping Φ, which transforms $a_1, \ldots, a_p$ to $\pi_1, \ldots, \pi_p$ can be easily given by a system of linear equations, analogous to the Yule-Walker equations.

Let us introduce the notations

$$E|\xi(n)|^2 = \gamma^2, \qquad \frac{E(\xi(n) . \overline{\xi(n-j)})}{\gamma^2} = \varrho_j ,$$

$(E\,\xi(n) = 0)$. By the normality of the process $\xi(n)$, we can write

$$(4.1) \quad E(\xi(n) | \mathcal{F}_{n-j}^{n-1}) = \sum_{i=1}^{j} \alpha_{j,i} \; \xi(n-i)$$

It is easy to calculate that $\pi_j = \frac{\alpha_{j,j}}{\gamma}$.
Multiplying (1) by $\overline{\xi(n-i)}$ for i=1,...,j and taking expectation we get the desired system:

$$(4.1') \quad \begin{aligned}
\alpha_{j,1} + \varrho_1 \alpha_{j,2} + \cdots + \varrho_{j-1} \alpha_{j,j} &= \varrho_1 , \\
\varrho_1 \alpha_{j,1} + \alpha_{j,2} + \cdots + \varrho_{j-2} \alpha_{j,j} &= \varrho_2 , \\
\vdots \qquad\qquad \vdots \qquad\qquad & \vdots \\
\varrho_{j-1} \alpha_{j1} + \varrho_{j-2} \alpha_{j2} + \cdots + \alpha_{j,j} &= \varrho_j ,
\end{aligned}$$

j : 1, ... , p

Especially - by the Yule-Walker equations $\alpha_{pi} = a_i$.
If the coefficients
$a_0, \dots, a_p$ are admissible
(the roots of polynomial $z^p - a_1 z^{p-1} - \dots - a_p$ are all inside the unit circle),
then there exists a unique stationary process satisfying (1). From this fact follows the existence and unicity of the solution of system (1) especially the existence of π_j's, moreover $|\pi_j| < 1$. (If for some j $\pi_j = 1$ the process $\xi(n)$ would be deterministic.)
We can prove the converse of this assertion.
Theorem 1. The mapping π, which transforms

$$\underline{a} = (a_1, \dots, a_p) \quad \text{to} \quad \underline{\pi} = (\pi_1, \dots, \pi_p)$$

one-to-one and onto to $\underbrace{(-1,1) \times \dots \times (-1,1)}_{p\text{-times}}$. Furthermore, both Π and its inverse Φ are continuously differentiable.

Before the proof we recite a criterion for the polynomial having zeros only inside the unit circle (see Duffin [1]).

<u>Criterion 1.</u>

Let $f(z)$ be the polynomial
$f(z) = a_0 + a_1 z + \dots a_n z^n$
where $a_0 \neq 0$, $a_n \neq 0$ and $n \neq 0$.
Let $\check{f}(z)$ the reduced polynomial

$$\check{f}(z) = (\bar{a}_n a_1 - a_0 \bar{a}_{n-1}) + (\bar{a}_n a_2 - a_0 \bar{a}_{n-2}) z + \ldots + (\bar{a}_n a_n - a_0 \bar{a}_0) z^{n-1}$$

of degree $n-1$ **. Then** $f(z)$ **has zeros only inside the unit circle if and only if** $|a_0| < |a_n|$ **and** $\check{f}(z)$ **do so.**

Proof. Consider the polynomial

$$\varphi(z) = 1 - a_1 z - \ldots - a_p z^p$$

It is easily seen that $\underline{a} \in \mathfrak{A}$ **if and only if** $\Psi(z) = z^p \varphi(\frac{1}{z})$ **has zeros only inside the unit circle and, according to Criterion 1,** this is equivalent to the oonditions

$$|a_p| < 1$$

$$\check{\Psi}(z) = (1 - a_p^2)(\beta_0 + \beta_1 z + \ldots + \beta_{p-2} z^{p-2} + z^{p-1})$$

has zeros only inside the unit circle, where

$$\beta_i = (a_{p-1-i} + a_{i+1} a_p)/(1 - a_p^2), \quad i = 0, 1, \ldots, p-2.$$

Next it will be proved that

$$(4.2) \quad \beta_i = \alpha_{p-1,\, p-1-i}, \quad i = 0, 1, \ldots, p-2$$

The equations

$$(4.3) \quad \alpha_{p,j} = \alpha_{p-1,j} - \alpha_{p,p}\, \alpha_{p-1,\, p-j}, \quad j = 1, \ldots, p-1,$$

determine, for $\alpha_{p,p}$ **fixed, a transformation of** $(\alpha_{p,1}, \ldots, \alpha_{p,p-1})$ **to** $(\alpha_{p-1,1}, \ldots, \alpha_{p-1,p-1})$.

with the Jacobian matrix

$$J_{p-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & -\pi_p \\ 0 & 1 & 0 & \cdots & -\pi_p & 0 \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ 0 & -\pi_p & 0 & & 1 & 0 \\ -\pi_p & 0 & 0 & & 0 & 1 \end{pmatrix}$$

(recall that $\pi_p = \alpha_{p,p}$), where the central element is $1-\pi_p$ if p is even. Since the value of $|J_{p-1}|$ is

$$(1-\pi_p)^{\left[\frac{p}{2}\right]}(1+\pi_p)^{[(p-1)/2]}$$

which is $\neq 0$ if and only if $\pi_p^2 \neq 1$, it suffices to prove that the β-s satisfy (3).

Thus (2) is established.

Observing that

$$\beta_0 = \alpha_{p-1,p-1} = \pi_{p-1}$$

a repetitive use of Criterion 1 immediately shows that $\Psi(z)$ has zeros only inside the unit circle if and only if

$$|\pi_1| < 1 \;\&\; |\pi_2| < 1 \;\&\ldots\&\; |\pi_p| < 1,,$$

and thus

$$\mathcal{P} = (-1,1) \underbrace{\times \ldots \times}_{p\text{-times}} (-1,1).$$

From the definition of partial correlations it follows that π_j's are uniqually determined by the coefficients a_j and dispersion σ.

On the other hand from recursive formulas due to Durbin (see Exercise 2) we get

$$\begin{aligned}\alpha_{p,j} &= \alpha_{p-1,j} - \alpha_{p,p}\,\alpha_{p-1,p-j} = \\ &= \alpha_{p-2,j} - \alpha_{p-1,p-1}\,\alpha_{p-2,p-1-j} - \alpha_{p,p}\,\alpha_{p-1,p-j} = \\ &= \cdots \\ &= \alpha_{j,j} - \alpha_{j+1,j+1}\,\alpha_{j,1} - \cdots - \alpha_{p,p}\,\alpha_{p-1,p-j}, \\ &\qquad j = 1,\ldots,p-1.\end{aligned} \tag{4.4}$$

The right hand side of (4) contains π_j's and $\alpha_{\ell,m}$'s with $m<l<p$ and we can continue using (4), ending up with a polynomial in the π_j's . So we get that the inverse ϕ of the mapping π can be correctly defined, and therefore π is one-to-one.

Since ϕ, defined by (4), is a polynomial in π_j's, continuosly differentiable, it is therefore, by the inverse function theorem, sufficient to show that

$$\left|\frac{\partial \phi}{\partial \pi^*}\right| \neq 0 \qquad \pi \in \mathcal{F} \tag{4.5}$$

(A^* denotes the transposed of A). As was pointed out by Daniels [1], the Jacobian (5) can be found by repetitive use of the transformation (3) , yielding

$$\left|\frac{\partial \phi}{\partial \pi^*}\right| = \prod_{j=1}^{p} (1-\pi_j)^{[j/2]}(1+\pi_j)^{[(j-1)/2]}.$$

Exercises

1. Prove that $\pi_j = \alpha_{j,j}/\gamma$

[Hint: use the relation

$$E(\xi(n)\,|\,\mathcal{F}_{n-j+1}^{n-1}) = \alpha_{j,j}\, E(\xi(n-2)\,|\,\mathcal{F}_{n-j+1}^{n-1})\,].$$

2. Prove the recursive formulas

$$\alpha_{j+1,i} = \alpha_{j,i} - \alpha_{j+1,j+1}\,\alpha_{j,j+1-i}$$

$$\alpha_{j+1,j+1} = \left(\varrho_{j+1} - \sum_{i=1}^{j} \alpha_{j,i}\,\varrho_{j+1-i}\right) \Big/ \left(1 - \sum_{i=1}^{j} \alpha_{j,i}\,\varrho_i\right)$$

which are due to Durbin [1].

[Hint: use the geometrical picture expressed by equation (1).]

3. Prove that in the 3-dimensional case the domain $\mathfrak{A}$ of admissible parameters is determined by the inequalities

$$\begin{aligned} a_1 + a_2 + a_3 &< 1 \\ -a_1 + a_2 - a_3 &< 1 \\ a_3(a_3 - a_1) - a_2 &< 1 \\ |a_3| &< 1 \end{aligned}$$

4. Prove the following statement: the wide sense stationary process $\xi(n)$ $(n = 0, \pm 1, \pm 2, \dots)$ with partial autocorrelation function π_j $(j > 0)$ is regular if and only if

(*) $\quad -1 < \pi_j < 1 \qquad$ for all $\quad j > 0$.

Furthermore the autocorrelation function $R(j)$ in regular case is strictly positive definite. So (*) is fulfilled if and only if $R(j)$ is strictly positive definite (see Ramsey [1]).

5. Prove that the stationary process is singular if and only if (Ramsey [1])

$$\sum_{j=1}^{\infty} \pi_j^2 = \infty$$

6. Prove that $\xi(n)$ is a p-order autoregressive process if and only if $\pi_j = 0$ when $j > p$.

7. Prove theorem 1 by the help of exercises 4. and 6. (Ramsey [1]).

8. Let $\xi(n)$ a stationary Gaussian process with

$$E\,\xi(n) = \mu, \quad E(\xi(n)-\mu)^2 = \gamma^2.$$

Prove that

$$E(\xi(k+1) \mid \mathcal{F}_1^k) = \mu + \sum_{j=1}^{k} \alpha_{k,j}(\xi(k+1-j)-\mu)$$

$$E([\xi(k+1) - E(\xi(k+1) \mid \mathcal{F}_1^k)]^2 \mid \mathcal{F}_1^k) = \gamma^2(1-\pi_1^2)\ldots(1-\pi_k^2),$$

where $\alpha_{k,j}$ are given in (4.1').

Bibliography

Books

Anderson, T. W. [1](1958) An Introduction to Multivariate Statistical Analysis, John Wiley, New York.

[2] (1970) Time Series Analysis, John Wiley, New York.

Arató M. [1](1962) Some statistical problems of the stationary Gauss-Markov processes, (in Russian), Dissertation, Moscow University.

[2](1972) Statistical problems of elementary Gaussian processes, Dissertation, Hungarian Academy of Sciences.

Arató, M. - Benczúr, A. - Krámli, A. - Pergel, J. (1974) Statistical problems of the elementary Gaussian processes. MTA SzTAKI Stochastic Processes, Part I.

Bartlett, M. S. (1966) An Introduction to Stochastic Processes with Special Reference to Methods and Applications, Second Edition, Cambridge Univ. Press, Cambridge.

Blackman, R. B. and Tukey, J. W. (1959) The Measurement of Power Spectra from the Point of View of Communications Engineering, Dover Publications , New York, (originally printed in 1958, Bell System Techn.J.37,185-282, 485-569).

Box, G. E. P. and Jenkins, G. M. (1970) Time Series Analysis Forecasting and Control, Holden Day, San Francisco.

Doob, J. L. (1953) Stochastic Processes, John Wiley, New York.

Feller, W. [1](1968) An Introduction to Probability Theory and its Applications, Vol.I, Third Edition, John Wiley, New York.

[2](1966) Introduction to Probability Theory and its Applications, Vol.II, John Wiley, New York.

Gantmaher, F. R. (1946) Theory of matrices, (in Russian), Nauka, Moskva.

Gihman, J. J. and Skorohod, A. V. (1965) Introduction to the Theory of Stochastic Processes, (in Russian), Nauka, Moskva.

Granger, C. W. J. and Hatanaka, M. (1964) Spectral Analysis of Economic Time Series, Princeton Univ. Press, Princeton, N.S.

Grenander, U. and Rosenblatt, M. (1957) Statistical Analysis of Stationary Time Series, John Wiley, New York.

Hannan, E. J. (1960) Time Series Analysis, Methuen and John Wiley, New York, London.

Kendall, M. G. and Stuart, A.[1](1961) The Advanced Theory of Statistics, Vol. II, Charles Griffin and Hafner, New York, London.

[2](1966) The Advanced Theory of Statistics, Vol.III, Charles Griffin and Hafner,

New York, London.

Kolmogorov, A. N. (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin (reprinted in 1946 by Chelsea, New York).

Lipcer, R. S. and Shiryayev, A. N. (1974) Statistics of Stochastic Processes. Moskva, Nauka (in Russian).

Rozanov, Yu. A.(1962) Stationary Random Processes (in Russian). English translation (1967) Holden Day, San Francisco.

Shiryayev, A. N. (1969) Statistical sequential analysis (in Russian),Nauka, Moskva.

Whittle, P. [1] Hypothesis Testing in Time Series Analysis, Thesis, Uppsala University, Almquist and Wiksell, Uppsala, Hafner, New York (1951).

[2] (1963) Prediction and Regulation, London.

Yaglom, A. M. (1962) An Introduction to the Theory of Stationary Random Functions (Revised English Edition, Richard A. Silverman), Prentice Hall.

References

Andel, J. [1](1971) On the multiple autoregressive series, Ann.Math.Stat., 42, No.2, 755-759.

[2](1972) Symmetric and reversed multiple stationary autoregressive series, Ann.Math.Stat.,43, No.4, 1197-1203.

Arató, M.[1](1961) On the sufficient statistics for stationary Gaussian random processes, Theory of Prob. and Its Appl., 6, 199-201.

[2](1962) Estimation of the parameters of a stationary Gauss-Markov process (in Russian), Dokl. Akad. Nauk, SSSR, 145, 13-16.

[3](1964) On the statistical investigation of Markov processes with continuous state space,I-IV, (in Hungarian), MTA III. Oszt. Közl.,14,13-14, 137-159, 317-33o., 15, 107-204, (1965).

[4](1970) On the parameter estimation of processes satisfying a linear stochastic differential equation (in Russian), Studia Sci. Math. Hung., 5, 11-16.

[5](1970) Exact formulas for density measure of elementary Gaussian processes (in Russian), Studia Sci. Math. Hung., 5, 17-27.

Arató, M. and Benczúr, A. (1970) Distribution function of the damping parameter of stationary Gaussian processes (in Russian), Studia Sci. Math. Hung., 5, 445-456.

Arató, M., Kolmogorov, A. N. and Sinai, Ya. G. (1962) On the parameter estimation for complex stationary Gauss-Markov processes (in Russian), Dokl. Akad. Nauk, SSSR, 146, 747-750.

Barndorff-Nielsen, O. and Schou, G. (1973) On the parametrization of autoregressive models by partial autocorrelations, J. Multivar. Anal.,3,

Baxter, G. (1956) A strong limit theorem for Gaussian processes, Proc. Amer. Math. Soc.,7, 522-525.

Benczúr, A. and Szeidl, L. (1974) On absolute continuity of measures defined by multidimensional diffusion processes with respect to the Wiener measure, MTA SzTAKI Közlemények, 13, 5-10.

Daniels, H. E. (1956) The approximate distribution of serial correlation coefficients, Biometrika, 43, 169-185.

Duffin, E. J. (1969) Algorithms for classical stability problems, SIAM Review, 11, 196-213.

Durbin, J. (1960) The fitting of time series models, Rev. Int. Stat. Inst., 28, 233-244.

Grenander, U. (1950) Stochastic processes and statistical inference, Årkiv för Matematik, 1,195-277.

Hajek, J. (1962) On linear statistical problems in stochastic processes, Czechoslovak Math.Journ.,12,404-444.

Krámli, A. (1971) On homogeneous Gaussian Markovian processes (in Russian), Studia Sci. Math. Hung.,6, No.1-2, 167-168.

Krámli, A. and Pergel, J.[1](1974) The connection between Gaussian Markov Processes and autoregressive--moving average processes, MTA SzTAKI Közlemények, 13, 53-58.

[2] (1974) On the Random-Nikodym derivative of elementary Gaussian processes.(in Hungarian) Alk. Mat. Lapok, 1, No.1 (in print).

Lipcer, R. S. and Sirjaev, A. N. (1968) Non-linear filtration for diffusional Markov process (in Russian), Trudy Mat. Inst., AN SSSR, 104, 135-150.

Mann, H. B. and Wald, A. (1943) On the stochastical treatment stochastic difference equations, Econometrica, 11, 173-220.

Pisarenko, V. F. (1963) On the estimation of parameters of a stationary Gauss process with spectral density (in Russian), Litovsk. Mat. Sbornik, 2, 159-167.

Prohorov, J. V. (1956) Convergence of stochastic processes and limit theorems of probability theory (in Russian), Theory of Prob. and Its Appl., 1, 177-238.

Ramsey, Fr. (1974) Characterization of the partial autocorrelation function, Ann. Stat., 2, No.6, 1296-1301.

Shaman, P. [1](1969) On the inverse of the covariance matrix of a first.order moving average process, Biometrika,56, 595-600.

[2](1973) On the inverse of the covariance matrix for an autoregressive-moving average process, Biometrika, 60, No.1, 193-196.

Siddiqui, M. M. (1957) On the inversion of the sample covariance matrix in a stationary autoregressive process, Ann. Math. Stat., 29, 585-588.

Striebel, Ch. (1959) Densities for stochastic processes, Ann. Math. Stat., 30, 559-567.

Taraskin, A. F. (1973) Some limit theorems for stochastic integrals (in Russian), Theory of Stoch.Proc., 119-134.

Jelen dolgozat az 5.11.1. és az 5.11.3.sz. kutatási téma keretében készült.