EXTRAPOLATION OF RANDOM VECTOR PROCESSES AND THE AMOUNT OF INFORMATION CONTAINED IN ONE VECTOR STATIONARY RANDOM PROCESS RELATIVE TO ANOTHER STATIONARILY CONNECTED WITH IT

1958

SovietRxiv

Abstract

Full Text

MATHEMATICS

M. S. PINSKER

EXTRAPOLATION OF RANDOM VECTOR PROCESSES AND THE AMOUNT OF INFORMATION CONTAINED IN ONE VECTOR STATIONARY RANDOM PROCESS RELATIVE TO ANOTHER STATIONARILY CONNECTED WITH IT

(Presented by Academician A. N. Kolmogorov on 1 III 1958)

1°. $\xi = (\{\xi(t)\}) = (\{\xi_1(t)\}, \{\xi_2(t)\}, \dots, \{\xi_n(t)\})$ is a vector stationary random process if $\xi_1 = (\{\xi_1(t)\}), \xi_2 = (\{\xi_2(t)\}), \dots, \xi_n = (\{\xi_n(t)\})$ are stationary and stationarily connected random processes ([7]). It is assumed that $M\xi_i(t) = 0,\ D\xi_i(t) < \infty,\ i = 1, \dots, n$.

In the cases when t runs through all integer values or all real values, we shall call ξ a random vector process, respectively, of discrete or continuous argument.

Let H_t be the closed linear span of the random variables $\xi_i(\tau)^*$, $\tau \le t$; $i = 1, \dots, n$, and let $\hat{\xi}_i(s)$, $s > 0$, be the perpendiculars dropped from $\xi_i(t+s)$ onto H_t. The process ξ has rank $k \le n$ if among $\hat{\xi}_i(s)$, $i = 1, \dots, n$, there are exactly k linearly independent random variables; it is easy to see that k does not depend on s or t.

Theorem 1. *If $\xi = (\{\xi(t)\}) = (\{\xi_1(t)\}, \dots, \{\xi_n(t)\})$ is a stationary vector random process of discrete argument and*

$$\hat{r}_{ij} = M\hat{\xi}_i(1)\hat{\xi}_j(1),$$

then

$$\left(\det \|\hat{r}_{ij}\|_{i,j=1,\dots,n}\right)^{1/2} = (2\pi)^{n/2} \exp\left\{\frac{1}{4\pi}\int_{-\pi}^{\pi} \log\det \|f_{\xi_i\xi_j}(\lambda)\|_{i,j=1,\dots,n}\, d\lambda\right\}; \quad (1)$$

$f_{\xi_i\xi_j}(\lambda) = F'_{\xi_i\xi_j}(\lambda)$; $F_{\xi_i\xi_j}(\lambda)$ are the spectral functions and cross-spectral functions of the processes ξ_i and ξ_j

It follows from the theorem that the vector process ξ has rank n if and only if the integral

$$\int_{-\pi}^{\pi} \log \det \|f_{\xi_i \xi_j}(\lambda)\|_{i,j=1,\dots,n} \, d\lambda$$

converges.

From this the theorem of Zasuhin ([5],[6]) easily follows:

In order that the vector process ξ be regular of rank n, it is necessary and sufficient that the spectral and cross-spectral functions of the processes ξ_i, $i = 1, \dots, n$, be absolutely continuous and that the integral

$$\int_{-\pi}^{\pi} \log \det \|f_{ij}(\lambda)\|_{i,j=1,\dots,n} \, d\lambda \tag{2}$$

converge.

---

* Convergence is understood in the mean-square sense, i.e. $\zeta_i \to \zeta$ if $M|\zeta_i - \zeta|^2 \to 0$.

in the case when ξ is a process of a discrete argument t, and of the integral

$$\int_{-\infty}^{\infty} \frac{\log \det \|f_{ij}(\lambda)\|_{i,j=1,\dots,n}}{1+\lambda^2} \, d\lambda \tag{3}$$

in the case when ξ is a process of a continuous argument t.

2°. Stationary vector random processes

$$\xi = ((\xi(t))) = (\{\xi_1(t)\}, \dots, \{\xi_n(t)\})$$

and

$$\eta = ((\eta(t))) = (\{\eta_1(t)\}, \dots, \{\eta_m(t)\}) = (\{\xi_{n+1}(t)\}, \dots, \{\xi_{n+m}(t)\})$$

are stationary-related if the random processes $\xi_1, \dots, \xi_{n+m}$ are stationary-related.

In accordance with (1-4), we introduce two variants of the definition of the rate of creation of information about the process η when observing the process ξ:

$$\vec{I}(\xi, \eta) = \frac{1}{T} MI(\xi_T, \eta \mid \xi_-); \tag{4}$$

$$\bar{I}(\xi, \eta) = \lim_{T\to\infty} \frac{1}{T} I(\xi_T, \eta_T), \tag{5}$$

where ξ_T and ξ_- are the parts of the vector process $\xi = ((\xi(t)))$, respectively, for $0 \le \tau \le T$ and $-\infty < \tau \le 0$.

In the case of a continuous argument t, it is useful to introduce into consideration processes of discrete argument

$$\xi^{(h)} = ((\xi^{(h)}(l))) = (\{\xi_1^{(h)}(l)\}, \dots, \{\xi_n^{(h)}(l)\}),$$

where $\xi_i^{(h)}(l) = \xi_i(hl)$, $i = 1, \dots, n$, and for processes of continuous argument one may propose two more variants of the definition of the rate of creation of information:

$$\vec{I}^{(\mathrm{d})}(\xi,\eta) = \lim_{h\to+0} \frac{1}{h}\, \vec{I}(\xi^{(h)}, \eta^{(h)}); \tag{6}$$

$$\bar{I}^{(\mathrm{d})}(\xi,\eta) = \lim_{h\to+0} \bar{I}(\xi^{(h)}, \eta^{(h)}). \tag{7}$$

Theorem 2.

$$\bar{I}(\xi,\eta) \ge \bar{I}^{(\mathrm{d})}(\xi,\eta) \ge \vec{I}^{(\mathrm{d})}(\xi,\eta) \ge \vec{I}(\xi,\eta).$$

3°. Put

$$M = \int \log \frac{\widetilde{A}_\xi(\lambda)\widetilde{A}_\eta(\lambda)}{\widetilde{A}_{\xi\eta}(\lambda)}\, d\lambda, \tag{8}$$

where $\widetilde{A}_\xi(\lambda)$ and $\widetilde{A}_\eta(\lambda)$ are nonzero minors of the highest k-th and r-th orders of the matrices

$$\|f_{\xi_i\xi_j}(\lambda)\|_{i,j=1,\dots,n} \quad \text{and} \quad \|f_{\eta_i\eta_j}(\lambda)\|_{i,j=1,\dots,m} = \|f_{\xi_i\xi_j}(\lambda)\|_{i,j=n+1,\dots,n+m}.$$

$\widetilde{A}_{\xi\eta}(\lambda)$ is a minor of order $(k+r)$ of the matrix

$$\|f_{\xi_i\xi_j}(\lambda)\|_{i,j=1,\dots,n+m},$$

containing the minors $\widetilde{A}_\xi(\lambda)$ and $\widetilde{A}_\eta(\lambda)$. If $k = 0$ or $r = 0$, then we also set the integrand equal to zero.

For vector processes of a discrete argument, the integral is taken over the limits from $-\pi$ to π, and for vector processes of a continuous argument, over the limits from $-\infty$ to ∞.

Remark. Although the choice of $\widetilde{A}_\xi(\lambda)$ and $\widetilde{A}_\eta(\lambda)$ may be made nonuniquely, i.e., for a given λ there may exist several nonzero minors $\widetilde{A}_\xi(\lambda)$ and $\widetilde{A}_\eta(\lambda)$, respectively, of orders k and r, nevertheless the value of the integrand does not depend on $\widetilde{A}_\xi(\lambda)$ and $\widetilde{A}_\eta(\lambda)$.

Now, using Theorem 1, it is easy to prove the following theorems, which are generalizations of the theorems proved in (8).

Theorem 3. If ξ, η are stationary and stationarily connected vector random processes, the pair (ξ, η) has a normal distribution*, and ξ is a regular process, then

$$\vec{I}(\xi, \eta) = M. \tag{9}$$

If, however, ξ is a singular vector process, then

$$\vec{I}(\xi, \eta) = 0.$$

When the rank of the vector process ξ is equal to n, Theorem 3 can be strengthened, namely:

Theorem 4. If ξ, η are stationary and stationarily connected vector processes, the pair (ξ, η) has a normal distribution and the rank of the process $\xi = (\{\xi(t)\}) = (\{\xi_1(t)\}, \dots, \{\xi_n(t)\})$ is equal to n, then

$$\bar{I}^{(n)}(\xi, \eta) = \vec{I}^{(n)}(\xi, \eta) = \vec{I}(\xi, \eta) = M. \tag{10}$$

If all the functions $f_{\xi_i \xi_j}(\lambda)$ are rational, then

$$\bar{I}(\xi, \eta) = M. \tag{11}$$

We note that equality (11) can also be established in a number of other cases; apparently, it will always be true when one of the vector processes ξ or η is regular.

Further, if the vector process—the noise $\zeta(t) = \xi(t) - \eta(t)$—is independent of $\eta(t)$, then from (8) it follows that

$$M = \int \log \frac{\widetilde{\Lambda}_\xi(\lambda)}{\widetilde{\Lambda}_\zeta(\lambda)} \, d\lambda; \tag{12}$$

in particular, for $n = m = 1$ we obtain the well-known [4,8] expression

$$M = \int \log \left(1 + \frac{f_{\xi\xi}(\lambda)}{f_{\zeta\zeta}(\lambda)} \right) d\lambda. \tag{13}$$

Received
26 II 1958

CITED LITERATURE

1. I. M. Gel' fand, A. N. Kolmogorov, A. M. Yaglom, DAN, **111**, 745 (1956).

2. A. N. Kolmogorov, *Proceedings of the Session on Automatic Problems of Production*, Publishing House of the Academy of Sciences of the USSR, 1956.

3. A. N. Kolmogorov, A. M. Yaglom, I. M. Gel' fand, Report at the 3rd Mathematical Congress, **3**, 1956.

4. C. E. Shannon, W. Weaver, *The Mathematical Theory of Communications*, 1949, p. 3.

5. A. N. Kolmogorov, Bull. Moscow State Univ., **2**, 6 (1941).

6. V. N. Zasukhin, DAN, **33**, 435 (1941).

7. A. M. Yaglom, Uspekhi Mat. Nauk, **7**, No. 5, 3 (1952).

8. M. S. Pinsker, DAN, **99**, 213 (1954).

* That is, the joint distribution of the random variables $\xi_1(t_1), \dots, \xi_{n+m}(t_s)$ is normal.

Note: Figure translations are in progress. See original paper for figures.