



Soviet-era science, translated into English

Reports of the Academy of Sciences of the USSR

O. V. SARMANOV

1958

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-195801.91874>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

Reports of the Academy of Sciences of the USSR
1958. Volume 121, No. 1

MATHEMATICS

O. V. SARMANOV

THE MAXIMAL CORRELATION COEFFICIENT

(THE NONSYMMETRIC CASE)

(Presented by Academician S. N. Bernstein on 21 II 1958)

1. Let $F(x, y)$ be the density of a distribution defining the correlation between the random variables x and y in the rectangular domain

$$\Omega = [a \leq x \leq b; a_1 \leq y \leq b_1],$$

which may also be infinite. By

$$p(x) = \int_{a_1}^{b_1} F(x, y) dy, \quad P(y) = \int_a^b F(x, y) dx$$

we denote the a priori densities, respectively, of x and y , and suppose that the square of the kernel

$$K(x, y) = \frac{F(x, y)}{\sqrt{p(x)P(y)}}$$

is integrable in both variables.

The nonsymmetric density $F(x, y)$ defines two symmetric densities

$$F_1(x, y) = \int_{a_1}^{b_1} \frac{F(x, t)F(y, t)}{P(t)} dt; \quad F_2(x, y) = \int_a^b \frac{F(t, x)F(t, y)}{p(t)} dt \quad (1)$$

and two symmetric kernels

$$K_1(x, y) = \frac{F_1(x, y)}{\sqrt{p(x)p(y)}}; \quad K_2(x, y) = \frac{F_2(x, y)}{\sqrt{P(x)P(y)}}. \quad (2)$$

The kernels (2) are positive and have identical spectra of eigenvalues

$$1 < \lambda_1^2 \leq \lambda_2^2 \leq \dots \leq \lambda_k^2 \leq \dots \quad (3)$$

and, generally speaking, different spectra of eigenfunctions

$$\{\varphi_i(x)\}, \quad \{\psi_i(x)\}, \quad i = 1, 2, \dots \quad (4)$$

According to the Hilbert-Schmidt theory of eigenfunctions, the spectrum of the kernel $K(x, y)$ has the form

$$\sqrt{p(x)}, \quad \sqrt{P(y)}, \quad \sqrt{p(x)}\varphi_i(x), \quad \sqrt{P(y)}\psi_i(y), \quad i = 1, 2, \dots, \quad (5)$$

where the bilinear expansion

$$K(x, y) \sim \sqrt{p(x)}\sqrt{P(y)} + \sum_{i=1}^{\infty} \frac{\varphi_i(x)\sqrt{p(x)}\psi_i(y)\sqrt{P(y)}}{\lambda_i} \quad (6)$$

converges in the mean to $K(x, y)$ in the domain Ω .

2. Definition. We shall call $R^* = \frac{1}{\lambda_1}$ the **maximal** (in absolute value) **correlation coefficient corresponding to the density** $F(x, y)$. This definition is analogous to the concept, introduced in the work ⁽¹⁾, of the maximal coefficient for a symmetric density—

In particular, $R^* = \frac{1}{\lambda_1^2}$ is the maximal correlation coefficient for symmetric densities ⁽¹⁾.

3. To compute the maximal correlation coefficient, one may recommend the following process of successive approximations, whose convergence is proved, for example, in ⁽²⁾.

As the “zero approximation” $r_0(y)$ one may take any function having variance (without loss of generality, we shall assume that the mean value of $r_0(y)$ is equal to zero). Put

$$r_{2k+1}(x) = \int_{a_1}^{b_1} r_{2k}(y) \frac{F(x, y)}{p(x)} dy, \quad r_{2k+2}(y) = \int_a^b r_{2k+1}(x) \frac{F(x, y)}{P(y)} dx, \quad (7)$$

$$k = 0, 1, 2, \dots$$

Then, as follows from ⁽²⁾, the first pair of eigenfunctions, to within the normalizing factors e_1 and g_1 , is determined by the equalities

$$e_1\psi_1(y) = \lim_{k \rightarrow \infty} r_{2k}(y)\lambda_1^{2k}, \quad g_1\varphi_1(x) = \lim_{k \rightarrow \infty} r_{2k+1}(x)\lambda_1^{2k}. \quad (8)$$

If k is sufficiently large, then

$$R^* = \frac{1}{\lambda_1^2} \simeq \frac{r_{2k}(y)}{r_{2k-2}(y)} \simeq \frac{r_{2k+1}(x)}{r_{2k-1}(x)}. \quad (9)$$

The sign of the correlation coefficient R^* between the functions found, $\varphi_1(x)$ and $\psi_1(y)$, is easily determined by direct computation.

4. The advisability of introducing the concept of the maximal correlation coefficient is justified by the following theorems.

Theorem 1. *For the independence of random variables x, y , it is necessary and sufficient that the maximal correlation coefficient vanish.*

Proof. Necessity is obvious, since if x and y are independent, $\varphi_1(x)$ and $\psi_1(y)$ are also independent, and the correlation coefficient between them is equal to zero, i.e. $R^* = 0$.

Now let $R^* = \frac{1}{\lambda_1} = 0$; then, according to (5), the kernel $K(x, y)$ has no eigenfunctions except $1 \cdot \sqrt{p(x)}$ and $1 \cdot \sqrt{P(y)}$, and from the mean-square convergence of the bilinear expansion (6) it follows that

$$\int_a^b \int_{a_1}^{b_1} \left[\frac{F(x, y)}{\sqrt{p(x)P(y)}} - \sqrt{p(x)P(y)} \right]^2 dx dy = 0,$$

i.e.

$$F(x, y) = p(x)P(y) \quad (10)$$

for almost all x and y in the domain Ω , as was required to prove.

Theorem 2. *If the correlation is rectilinear, then the ordinary correlation coefficient R between x and y coincides with the maximal correlation coefficient R^* .*

Proof. In this case, in the spectrum (5) there is a pair of linear and, consequently, monotone eigenfunctions,

$$\varphi_1(x) = \frac{x - c}{\sigma}, \quad \psi_1(y) = \frac{y - c_1}{\sigma_1}, \quad (11)$$

where c and c_1 are the means, and σ^2 and σ_1^2 are the variances of x and y , respectively.

It is known (see, for example, (3)) that if in the spectrum of an asymmetric stochastic kernel there is a pair of monotone functions, then they always belong to the first eigenvalue; therefore the correlation coefficient

between x and y , equal to the correlation coefficient between their linear functions $\frac{x-c}{\sigma}$ and $\frac{y-c_1}{\sigma_1}$, coincides with the maximal correlation coefficient, as was required to prove.

5. For discrete random variables the maximal correlation coefficient is defined analogously.

Let the correlation dependence between discrete random variables be defined by the rectangular matrix

$$\{p_{ij}\}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad (12)$$

where

$$0 \leq p_{ij} = P\{x = x_i; y = y_j\}, \quad \sum_{ij} p_{ij} = 1, \\ p_i = \sum_{j=1}^m p_{ij} = P\{x = x_i\}, \quad P_j = \sum_{i=1}^n p_{ij} = P\{y = y_j\}. \quad (13)$$

With the aid of (12) form two square symmetric matrices

$$\left\{ \frac{p_{ij}^{(1)}}{\sqrt{p_i p_j}} \right\}, \quad i, j = 1, 2, \dots, n; \quad \left\{ \frac{p_{ij}^{(2)}}{\sqrt{P_i P_j}} \right\}, \quad i, j = 1, 2, \dots, m. \quad (14)$$

where

$$p_{ij}^{(1)} = \sum_{k=1}^m \frac{p_{ik} p_{jk}}{P_k}, \quad p_{ij}^{(2)} = \sum_{k=1}^n \frac{p_{ki} p_{kj}}{p_k}, \quad (15)$$

Then the correlation coefficient between the first eigenvectors of the matrices (14) is called the **maximal correlation coefficient R^* between random variables with correlation table** (12).

Let us note that with a matrix (12) having n rows and m columns, one can associate an infinite set of pairs of vectors $X\{x_1, x_2, \dots, x_n\}$ and $Y\{y_1, y_2, \dots, y_m\}$, and for each pair one can compute the correlation coefficient. If, however, one takes the pair of first eigenvectors of the matrices (14), then the correlation coefficient between them has the maximal (in absolute value) value. The square of the maximal correlation coefficient is the first eigenvalue of both matrices (14).

The process of successive approximations for finding R^* is analogous to that described in Sec. 3.

Let $r_0(y)$ be an arbitrary vector with coordinates $\{y_1^{(0)}, y_2^{(0)}, \dots, y_m^{(0)}\}$ such that

$$\sum_{j=1}^m y_j^{(0)} P_j = 0.$$

Put

$$x_i^{(2k+1)} = \sum_{j=1}^m \frac{P_{ij}}{P_i} y_j^{(2k)}, \quad i = 1, 2, \dots, n, \quad k = 0, 1, 2, \dots,$$

$$y_j^{(2k+2)} = \sum_{i=1}^n \frac{P_{ij}}{P_j} x_i^{(2k+1)}, \quad j = 1, 2, \dots, m, \quad k = 0, 1, 2, \dots$$

If k is sufficiently large, then

$$R^{*2} = \frac{1}{\lambda_1^2} \approx \frac{y_j^{(2k)}}{y_j^{(2k-2)}} \approx \frac{x_i^{(2k+1)}}{x_i^{(2k-1)}}$$

for all i and j .

The coordinates of the first eigenvectors of the matrices (14) are determined from the conditions

$$\begin{aligned} \xi_i &= \lim_{k \rightarrow \infty} x_i^{(2k+1)} \lambda_1^{2k}, & i = 1, 2, \dots, n; \\ \eta_j &= \lim_{k \rightarrow \infty} y_j^{(2k)} \lambda_1^{2k}, & j = 1, 2, \dots, m. \end{aligned} \tag{16}$$

V. A. Steklov Mathematical Institute
Academy of Sciences of the USSR

Received
20 II 1958

References

¹ O. V. Sarmanov, DAN, **120**, No. 4 (1958). ² O. V. Sarmanov, DAN, **53**, No. 9 (1946). ³ M. K. Nomokonov, DAN, **72**, No. 6 (1950).

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

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.