ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
21,2 (1991), pp. 225–233

W. WYSOCKI (Warszawa)

MAXIMAL CORRELATION IN PATH ANALYSIS

Abstract. The notions of multiple correlation coefficient and correlation ratio are generalized in this paper to the case of a system of random vectors. By geometrical means, using both linear regression and regression, a certain formula used in path analysis is also generalized to the vector case. The formulae obtained allow one to compute the generalized partial correlation coefficient for a pair of random vectors (Y_1, Y_2) after the linear impact or impact of random vectors X_i, $i = 1, \ldots, n$, has been eliminated.

1. Introduction. In the sequel we use the index of stochastic dependence between a pair of random vectors (Z_1, Z_2) introduced in [6].

The index is given by the following formula:

$$\varrho_\Lambda(Z_1, Z_2) = \frac{\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_1, Z_2))}{(\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_1, Z_1))^{1/2}(\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_2, Z_2))^{1/2}},$$

where Λ is a given, symmetric and positive definite $k \times k$ matrix.

In Section 2 we maximize ϱ_Λ on two sets of random vectors $L(Y) \times \mathcal{H}_{X_1,\ldots,X_n}$ and $L(Y) \times L^2_{X_1,\ldots,X_n}$, where Y and X_i, $i = 1, \ldots, n$, are k-dimensional, centered random vectors with all coordinates being random variables with finite second moments. $L(Y)$ denotes the space spanned by Y. The space $\mathcal{H}_{X_1,\ldots,X_n}$ consists of all random vectors X of the form $X = \sum_{i=1}^n A_i X_i$, where A_i, $i = 1, \ldots, n$, are $k \times k$ matrices. The space $L^2_{X_1,\ldots,X_n}$ contains all random vectors having the form $f(X_1, \ldots, X_n)$, where f is a Borel vector-valued function whose second order moments of coordinates are finite.

Maximization of ϱ_Λ on the sets described above leads us to new indices of stochastic dependence between systems of random vectors

---

1991 *Mathematics Subject Classification*: Primary 62H20. Secondary 62J05.

Key words and phrases: correlation ratio, partial correlation, multiple correlation, linear regression, cosine formula, measure of dependence.

$(Y, (X_1, \dots, X_n))$, namely:

$$\begin{aligned}\varrho_\Lambda(Y; X_1, \dots, X_n) &= \sup\{\varrho_\Lambda(bY, X) : bY \in L(Y), X \in \mathcal{H}_{X_1,\dots,X_n}\} \\ &= \varrho_\Lambda(Y, X_0), \\ \widetilde{\varrho}_\Lambda(Y; X_1, \dots, X_n) &= \sup\{\varrho_\Lambda(bY, Z) : bY \in L(Y), Z \in L^2_{X_1,\dots,X_n}\} \\ &= \varrho_\Lambda(Y, X_0').\end{aligned}$$

The random vectors X_0 and X_0' are eigenvectors of certain operators.

The indices $\varrho_\Lambda(Y; X_1, \dots, X_n)$ and $\widetilde{\varrho}_\Lambda(Y; X_1, \dots, X_n)$ are natural generalizations of the multiple correlation coefficient and the correlation ratio, respectively, to the case of a system of random vectors $(Y, (X_1, \dots, X_n))$.

In Section 3 a certain formula known in path analysis is generalized to the vector case. The formula was originally given in [6]. A principle of path analysis [4] was used to prove it. This principle claims that the correlation coefficient between two random variables is the sum of all paths connecting them on a suitable diagram. The formula unifies ordinary correlation, multiple correlation, partial correlation and path coefficients.

Using some geometric interpretation (given in [2]) we generalize the formula to the vector case. Two cases are considered, the first one uses the regression of Y_i on $\mathcal{H}_{X_1,\dots,X_n}$ (linear), the other one that on $L^2_{X_1,\dots,X_n}$ for $i = 1, 2$. The following formulae can be obtained:

$$\begin{aligned}(1.1)\qquad \varrho_\Lambda(Y_1, Y_2) &= \varrho_\Lambda(Y_1; X_1, \dots, X_n)\varrho_\Lambda(Y_1', Y_2')\varrho_\Lambda(Y_2; X_1, \dots, X_n) \\ &\quad + u_1 \varrho_\Lambda(U_1, U_2) u_2,\end{aligned}$$

where $u_i = (1 - \varrho_\Lambda^2(Y_i; X_1, \dots, X_n))^{1/2}$, and $U_i \in \mathcal{H}^\perp_{X_1,\dots,X_n}$ for $i = 1, 2$, with Y_i' such that $Y_i = Y_i' + U_i$. It should be noted here that the random vectors U_i are the remainders of Y_i, $i = 1, 2$, after subtracting the linear impacts X_j, $j = 1, \dots, n$. Next,

$$\begin{aligned}(1.2)\qquad \varrho_\Lambda(Y_1, Y_2) &= \widetilde{\varrho}_\Lambda(Y_1; X_1, \dots, X_n)\varrho_\Lambda(Y_1', Y_2')\widetilde{\varrho}_\Lambda(Y_2; X_1, \dots, X_n) \\ &\quad + u_1 \varrho_\Lambda(U_1, U_2) u_2,\end{aligned}$$

where $u_i = (1 - \widetilde{\varrho}_\Lambda^2(Y_i; X_1, \dots, X_n))^{1/2}$ and $U_i \in (L^2_{X_1,\dots,X_n})^\perp$, with Y_i' satisfying $Y_i = Y_i' + U_i$. The random vectors U_i are the remainders of Y_i after eliminating the impacts X_j, for $i = 1, 2$ and $j = 1, \dots, n$.

The numbers u_1, u_2 appearing in (1.1) and (1.2) are path coefficients. In the case described above they are also correlations in the sense of the index ϱ_Λ due to the fact that $\varrho_\Lambda(Y_i, U_i) = 0$ for $i = 1, 2$.

The indices $\varrho_\Lambda(U_1, U_2)$ appearing in (1.1) and (1.2) are generalizations of a partial correlation coefficient with linear impact, the impact of the random vectors $X_1, \dots, X_n$ removed. The formulae can be employed in order to calculate the partial correlation $\varrho_\Lambda(U_1, U_2)$.

2. Maximal correlation for a system of random vectors. In this and the next sections we consider, with certain restrictions, k-dimensional random vectors. They are defined as follows.

Let $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$ denote the space of all k-dimensional centered random vectors defined on a probability space $(\Omega, \mathcal{A}, P)$ such that all coordinates have their second moments finite. For a given symmetric and positive definite $k \times k$ matrix Λ we introduce in $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$ the scalar product $\langle Z_1, Z_2 \rangle_\Lambda = E(Z_1^T \Lambda^{-1} Z_2) = \operatorname{tr}(\Lambda^{-1} \operatorname{cov}(Z_1, Z_2))$ for $Z_1, Z_2 \in L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$.

The norm generated by this product is denoted by $\| \ \|_\Lambda$. It is complete. For purposes of this paper some subspaces of $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$ are defined. Let X_i, $i = 1, \ldots, n$, and Y be in $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$. Then

(2.1) $L(Y)$ is the subspace spanned by Y,

(2.2) $$L^2_{X_1,\ldots,X_n} = \{f(X_1, \ldots, X_n) \in L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k) : f : \mathbf{R}^{nk} \to \mathbf{R}^k \text{ is any Borel function whose all coordinates have the second moments finite}\},$$

(2.3) $$\mathcal{H}_{X_1,\ldots,X_n} = \left\{X \in L^2_{X_1,\ldots,X_n} : X = \sum_{i=1}^{n} A_i X_i, \text{ where } A_i \text{ are } k \times k \text{ matrices}\right\}.$$

Of course, (2.3) is a subspace of (2.2) and its dimension is at most nk^2. The spaces (2.1), (2.2), (2.3) are closed in $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$. From the Schwarz inequality one can see that the number

(2.4) $$\varrho_\Lambda(Z_1, Z_2) = \frac{\langle Z_1, Z_2 \rangle_\Lambda}{\|Z_1\|_\Lambda \|Z_2\|_\Lambda} = \frac{\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_1, Z_2))}{\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_1, Z_2))^{1/2}(\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Z_2, Z_2))^{1/2}}$$

can be considered as the cosine of the angle between the vectors Z_1, Z_2 in the space $(L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k), \langle \, , \, \rangle_\Lambda)$.

This suggests that the scalar product $\langle Z_1, Z_2 \rangle_\Lambda$ can be viewed as the covariance of the random vectors Z_1, Z_2, and the numbers $\|Z_1\|_\Lambda$ and $\|Z_2\|_\Lambda$ as dispersions of these random vectors. The formula (2.4) was introduced by Sampson [6] as an index of stochastic dependence between random vectors. This index will be maximized on the following spaces of random vectors:

(2.5) $$L(Y) \times \mathcal{H}_{X_1,\ldots,X_n},$$

(2.6) $$L(Y) \times L^2_{X_1,\ldots,X_n}.$$

Let $X_i = (X_{i1}, \ldots, X_{ik})^T$, $i = 1, \ldots, n$, and $Y = (Y_1, \ldots, Y_n)^T$ be random vectors in $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$. Denote by X the column vector consisting of X_i's, $i = 1, \ldots, n$:

$$X = (X_1^T, \ldots, X_n^T)^T . \tag{2.7}$$

We also introduce the following covariance matrices:

$$\Sigma_0 = \mathrm{cov}(Y, Y) , \tag{2.8}$$

$$\Sigma_i = \mathrm{cov}(X_i, Y) , \qquad i = 1, \ldots, n , \tag{2.9}$$

$$\Sigma_{ij} = \mathrm{cov}(X_i, Y_j) , \qquad i, j = 1, \ldots, n , \tag{2.10}$$

$$\Sigma = \mathrm{cov}(X, X) = [\Sigma_{ij}] , \qquad i, j = 1, \ldots, n , \tag{2.11}$$

$$\widetilde{\Sigma} = [\Sigma_i] , \qquad i = 1, \ldots, n . \tag{2.12}$$

Let B denote the block matrix

$$B = [B_i^T] , \quad i = 1, \ldots, n . \tag{2.13}$$

The facts necessary to prove that the maximization of ϱ_Λ on the sets (2.5) and (2.6) is correct are gathered in the following lemma:

LEMMA 1.a. *If the matrix Σ* (cf. (2.11)) *is nonsingular then the orthogonal projector $\mathcal{P}_2 : L(Y) \to \mathcal{H}_{X_1,\ldots,X_n}$ has the form*

$$\mathcal{P}_2(bY) = b\widetilde{\Sigma}^T \Sigma^{-1} X . \tag{2.14}$$

LEMMA 1.b. *The conditional expectation, given the random vector X, is the orthogonal projection $L(Y) \to L^2_{X_1,\ldots,X_n}$ given by*

$$\mathcal{P}_2(bY) = bE(Y \mid X)Y . \tag{2.15}$$

LEMMA 1.c. *The orthogonal projection $\mathcal{P}_1 : L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k) \to L(Y)$ has the form*

$$\mathcal{P}_1(Z) = \frac{\mathrm{tr}(\Lambda^{-1}\mathrm{cov}(Z, Y))}{\mathrm{tr}(\Lambda^{-1}\Sigma_0)} Y . \tag{2.16}$$

Proof of Lemma 1.a. Suppose that the random vector

$$\sum_{i=1}^n B_i X_i = B^T X , \tag{2.17}$$

where B_i, $i = 1, \ldots, n$, are some $k \times k$ matrices, is a projection of Y on $\mathcal{H}_{X_1,\ldots,X_n}$. We have used (2.7) and (2.13) here.

For every system $A_1, \ldots, A_n$ of $k \times k$ matrices the following holds:

$$\Big\langle Y - \sum_{i=1}^n B_i X_i , \sum_{j=1}^n A_j X_j \Big\rangle_\Lambda = \mathrm{tr}\Big(\Lambda^{-1}\mathrm{cov}\Big(Y - \sum_{i=1}^n B_i X_i , \sum_{j=1}^n A_j X_j\Big)\Big) = 0 . \tag{2.18}$$

We shall now find the last covariance matrix:

$$\text{(2.19)}\quad \operatorname{cov}\Big(Y-\sum_{i=1}^{n}B_iX_i\,,\ \sum_{j=1}^{n}A_jX_j\Big)$$
$$=\sum_{i=1}^{n}\Sigma_i^TA_i^T-\sum_{i=1}^{n}B_i\Big(\sum_{j=1}^{n}\Sigma_{ij}A_j^T\Big)=\sum_{i=1}^{n}\Big(\Sigma_i^T-\sum_{j=1}^{n}B_j\Sigma_{ji}\Big)A_i^T\,.$$

It is clear that a necessary condition for equality (2.14) to hold is that the covariance matrix (2.19) is the null matrix,

$$\sum_{i=1}^{n}\Big(\Sigma_i^T-\sum_{j=1}^{n}B_j\Sigma_{ji}\Big)A_i^T=0\,.$$

The last equality is equivalent to the system of n equalities:

$$\text{(2.20)}\qquad \sum_{j=1}^{n}B_j\Sigma_{ji}=\Sigma_i^T\,,\qquad i=1,\ldots,n\,,\quad\text{or}$$
$$\sum_{j=1}^{n}\Sigma_{ij}B_j^T=\Sigma_i\,,\qquad i=1,\ldots,n\,.$$

Using (2.7) and (2.9)–(2.13) we can write the system (2.20) as $\Sigma B=\widetilde{\Sigma}$. Since Σ is nonsingular, we have $B=\Sigma^{-1}\widetilde{\Sigma}$. This, together with (2.17), implies the formula (2.14) for the orthogonal projector.

P r o o f o f L e m m a 1.b. Set $Z_0=E(Y\mid X)$. First we show that

$$\text{(2.21)}\qquad \|Y-Z\|_\Lambda^2=\|Y-Z_0\|_\Lambda^2+\|Z_0-Z\|_\Lambda^2\quad\text{for } Z\in L^2_{X_1,\ldots,X_n}\,.$$

Note that the random vectors $Y-Z_0$ and Z_0-Z are orthogonal in $(L^2(\Omega,\mathcal{A},P,\mathbf{R}^k),\langle\,,\rangle_\Lambda)$:

$$\begin{aligned}\langle Y-Z_0,Z_0-Z\rangle_\Lambda&=\operatorname{tr}(\Lambda^{-1}\operatorname{cov}(Y-Z_0,Z_0-Z))\\&=\operatorname{tr}(\Lambda^{-1}E(Y-Z_0)(Z_0-Z)^T)\\&=\operatorname{tr}(\Lambda^{-1}E(E((Y-Z_0)(Z_0-Z)^T\mid X)))\\&=\operatorname{tr}(\Lambda^{-1}E(E((Y-Z_0)\mid X)(Z_0-Z)^T))=0\,.\end{aligned}$$

Squaring both sides of the equality $Y-Z=(Y-Z_0)+(Z_0-Z)$ we obtain (2.21).

The equality $\inf\{\|Y-Z\|_\Lambda : Z\in L^2_{X_1,\ldots,X_n}\}=\|Y-E(Y|X)\|_\Lambda$ is a consequence of (2.21). Hence the orthogonal projection theorem and the last equality yield the assertion of Lemma 1.b.

P r o o f o f L e m m a 1.c. The proof is obvious.

In the case where the matrix Σ is singular, we replace the inverse Σ^{-1} in the formula (2.14) by the Moore–Penrose inverse [3].

Maximization of ϱ_Λ on (2.5) and (2.6) will be performed via the following theorem (see [7]):

THEOREM. *Let $\mathcal{X}$ and $\mathcal{Y}$ be (closed) subspaces of a real Hilbert space* (1) $(H, \langle , \rangle)$, *and let* $\mathcal{P}_1 : X \to Y$, $\mathcal{P}_2 : \mathcal{Y} \to \mathcal{X}$ *be the orthogonal projections. If* $\mathcal{P}_1 \circ \mathcal{P}_2$ *and* $\mathcal{P}_2 \circ \mathcal{P}_1$ *are compact then there exist vectors* x_0 *and* y_0 *in* $\mathcal{X}$ *and* $\mathcal{Y}$, *respectively, such that*

$$\frac{\langle x_0, y_0 \rangle}{\|x_0\| \|y_0\|} = \sup \left\{ \frac{\langle x, y \rangle}{\|x\| \|y\|} : x \in \mathcal{X}, y \in \mathcal{Y} \right\} = a ,$$

where a is the square root of the maximal eigenvalue of these operators. Furthermore, x_0 *and* y_0 *are the respective eigenvectors.*

Maximization of ϱ_Λ on $L(Y) \times \mathcal{H}_{X_1,\ldots,X_n}$. From (2.14) and (2.16) we can obtain the following form of the operator $\mathcal{P}_1 \circ \mathcal{P}_2$ acting in $L(Y)$:

$$(\mathcal{P}_1 \circ \mathcal{P}_2)(bY) = \frac{\operatorname{tr}(\Lambda^{-1} \widetilde{\Sigma}^T \Sigma^{-1} \widetilde{\Sigma})}{\operatorname{tr}(\Lambda^{-1} \Sigma_0)} Y .$$

This operator has a unique eigenvalue a^2, equal to the maximal eigenvalue of $\mathcal{P}_2 \circ \mathcal{P}_1$.

Let Y and X_0 denote respective eigenvectors of the two operators. We obtain

$$\begin{aligned} (2.22) \qquad \varrho_\Lambda(Y; X_1, \ldots, X_n) &= \left(\frac{\operatorname{tr}(\Lambda^{-1} \widetilde{\Sigma}^T \Sigma^{-1} \widetilde{\Sigma})}{\operatorname{tr}(\Lambda^{-1} \Sigma_0)} \right)^{1/2} \\ &= \sup\{\varrho_\Lambda(bY, X)\} = \varrho_\Lambda(Y, X_0) , \end{aligned}$$

where the supremum is taken over $bY \in L(Y)$ and $X \in \mathcal{H}_{X_1,\ldots,X_n}$.

Maximization of ϱ_Λ on $L(Y) \times L^2_{X_1,\ldots,X_n}$. From (2.15) and (2.16) we obtain

$$\mathcal{P}_1 \circ \mathcal{P}_2(bY) = \frac{\operatorname{tr}(\Lambda^{-1} \operatorname{cov}(Y, E(Y \mid X)))}{\operatorname{tr}(\Lambda^{-1} \Sigma_0)} Y$$

and by an analogous argument combined with the easily verified fact

$$\operatorname{cov}(Y, E(Y \mid X)) = \operatorname{cov}(E(Y \mid X), E(Y \mid X)) ,$$

we have

$$\begin{aligned} (2.23) \quad \widetilde{\varrho}_\Lambda(Y; X_1, \ldots, X_n) &= \left(\frac{\operatorname{tr}(\Lambda^{-1} \operatorname{cov}(E(Y \mid X), E(Y \mid X)))}{\operatorname{tr}(\Lambda^{-1} \Sigma_0)} \right)^{1/2} \\ &= \sup\{\varrho_\Lambda(bY, Z) : bY \in L(Y) ,\ Z \in L^2_{X_1,\ldots,X_n}\} \\ &= \varrho_\Lambda(Y, X_0') . \end{aligned}$$

(1) The original proof of the theorem was given under the separability assumption imposed on the Hilbert space. Actually, it remains valid without it.

Note that for $k = 1$, or if X_i, $i = 1,\dots,n$, and Y are simply random variables, (2.22) reduces to the multiple correlation coefficient and (2.23) to the correlation ratio provided $n = 1$ [6]. Thus we clearly see that the indices $\varrho_\Lambda(Y; X_1,\dots,X_n)$ and $\widetilde{\varrho}_\Lambda(Y; X_1,\dots,X_n)$ are generalizations of the multiple correlation coefficient and the correlation ratio, respectively, to the system of random vectors $(Y,(X_1,\dots,X_n))$. The index $\varrho_\Lambda(Y; X_1,\dots,X_n)$ is the Sampson correlation ratio for $n = 1$.

3. Maximal correlation in path analysis. In this section we attempt to generalize a certain formula that is used in path analysis and was originally obtained algebraically in [4, 5]. There one of the principles of path analysis was used that claims the equality between the correlation coefficient and the sum of all paths between corresponding random variables on a suitable diagram. The formula obtained unifies ordinary correlation, multiple correlation, partial correlation and path coefficients.

Let $(H,\langle\ ,\ \rangle)$ be a real Hilbert space. In the sequel it is treated as an affine space (with adjoint vector space H). We recall a lemma of [2], in a slightly more general formulation. The proof is still valid without change.

LEMMA 2. *Let q_1 and q_2 denote the orthogonal projections of points p_1 and p_2, respectively, in the affine space H onto its closed subspace H_0. Then the measures of angles*

$$\alpha_i = m\angle(0p_i, 0q_i)\,, \quad i = 1,2\,, \tag{3.1}$$

$$\beta = m\angle(0q_1, 0q_2)\,, \tag{3.2}$$

$$\gamma = m\angle(0p_1, 0p_2)\,, \tag{3.3}$$

$$\varphi = m\angle(p_1q_1, p_2q_2)\,, \tag{3.4}$$

satisfy

$$\cos\gamma = \cos\alpha_1 \cos\beta_2 \cos\alpha_2 + \sin\alpha_1 \cos\varphi \sin\alpha_2\,. \tag{3.5}$$

Note that if $\beta = \varphi = 0$ the formula (3.5) is simply

$$\cos(\alpha_1 - \alpha_2) = \cos\alpha_1 \cos\alpha_2 + \sin\alpha_1 \sin\alpha_2\,.$$

We will apply Lemma 2 to the subspaces $H_0 = \mathcal{H}_{X_1,\dots,X_n}$ and $H_0 = L^2_{X_1,\dots,X_n}$, introduced in Section 2 by (2.2) and (2.3), respectively. Let Y_1, Y_2, X_i, $i = 1,\dots,n$, be in $L^2(\Omega, \mathcal{A}, P, \mathbf{R}^k)$.

Case $H_0 = \mathcal{H}_{X_1,\dots,X_n}$. Denote by $\widetilde{\Sigma}_i$, $i = 1,2$, the matrix defined by (2.12) for $Y = Y_i$, $i = 1,2$, and suppose that the matrix Σ (see (2.11)) is nonsingular. Then from (2.14) we see that the orthogonal projections of Y_i on H_0 have the form $Y_i' = \widetilde{\Sigma}_i^T \Sigma' X$ for $i = 1,2$. Hence $Y_i = Y_i' + U_i$, $U_i \in \mathcal{H}^{\perp}_{X_1,\dots,X_n}$, $\varrho_\Lambda(Y_i', U_i) = 0$ for $i = 1,2$. We call U_1, U_2 the remainders of Y_1, Y_2 after eliminating the linear impact of $X_1,\dots,X_n$. Use now Lemma

2 with $p_i = Y_i$, $i = 1,2$. We have $q_i = Y_i'$ and $p_i q_i = U_i$ for $i = 1,2$. The cosines of the angles (3.1)–(3.4) are

$$\cos\alpha_i = \varrho_\Lambda(Y_i, Y_i') = \varrho_\Lambda(Y_i; X_1, \ldots, X_n), \quad i = 1,2,$$
$$\cos\beta = \varrho_\Lambda(Y_1', Y_2'), \quad \cos\gamma = \varrho_\Lambda(Y_1, Y_2), \quad \cos\varphi = \varrho_\Lambda(U_1, U_2),$$

In this case (3.5) takes the form

$$\text{(3.6)} \quad \varrho_\Lambda(Y_1, Y_2) = \varrho_\Lambda(Y_1; X_1, \ldots, X_n)\varrho_\Lambda(Y_1', Y_2')\varrho_\Lambda(Y_2; X_1, \ldots, X_n) + u_1\varrho_\Lambda(U_1, U_2)u_2,$$

where $u_i = (1 - \varrho_\Lambda^2(Y_i; X_1, \ldots, X_n))^{1/2}$ for $i = 1,2$. The numbers u_i, $i = 1,2$, correspond to the path coefficients, and in this case they are the correlations $u_i = \varrho_\Lambda(Y_i, U_i)$, $i = 1,2$. Note that $\varrho_\Lambda(Y_i; X_1, \ldots, X_n)$, $i = 1,2$, in (3.6) is the multiple correlation coefficient for the systems of random vectors $(Y_i, (X_1, \ldots, X_n))$, $i = 1,2$ (see (2.22)). $\varrho_\Lambda(U_1, U_2)$ is a generalization of the partial correlation coefficient. This is a correlation in the sense of the index (2.4) for the pair (Y_1, Y_2) with the linear impact of $X_1, \ldots, X_n$ being eliminated.

Case $H_0 = L^2_{X_1,\ldots,X_n}$. Due to (2.15) the form of the orthogonal projection of Y_i on $L^2_{X_1,\ldots,X_n}$ is the following: $Y_i' = E(Y_i|X)$ for $i = 1,2$. Hence $Y_i = Y_i' + U_i$, $U_i \in (L^2_{X_1,\ldots,X_n})^\perp$ and $\varrho_\Lambda(Y_i', U_i) = 0$ for $i = 1,2$.

By an argument analogous to that used in the previous case, U_i, $i = 1,2$, are called the remainders of Y_i after eliminating the impact of X_j, $j = 1, \ldots, n$. Use now Lemma 2 with $p_i = Y_i$, then $q_i = Y_i'$ and $p_i q_i = U_i$ for $i = 1,2$. The cosines of the angles (3.1)–(3.4) are

$$\cos\alpha_i = \varrho_\Lambda(Y_i, Y_i') = \varrho_\Lambda(Y_i; X_1, \ldots, X_n), \quad i = 1,2,$$
$$\cos\beta = \varrho_\Lambda(Y_1', Y_2'), \quad \cos\gamma = \varrho_\Lambda(Y_1, Y_2), \quad \cos\varphi = \varrho_\Lambda(U_1, U_2).$$

In this case (3.5) takes the form

$$\text{(3.7)} \quad \varrho_\Lambda(Y_1, Y_2) = \widetilde{\varrho}_\Lambda(Y_1; X_1, \ldots, X_n)\varrho_\Lambda(Y_1', Y_2')\widetilde{\varrho}_\Lambda(Y_2; X_1, \ldots, X_n) + u_i\varrho_\Lambda(U_1, U_2)u_2,$$

where $u_i = (1 - \widetilde{\varrho}_\Lambda^2(Y_i; X_1, \ldots, X_n))^{1/2}$ for $i = 1,2$.

The numbers u_i, $i = 1,2$, correspond to the path coefficients, and in this case they are the correlations in the sense of ϱ_Λ. The numbers $\widetilde{\varrho}_\Lambda(Y_i; X_1, \ldots, X_n)$ appearing in (3.7) are the correlation ratios for the systems $(Y_i; X_1, \ldots, X_n)$ of random vectors, $i = 1,2$ (cf. (2.23)). As in the previous case the index $\varrho_\Lambda(U_1, U_2)$ is a generalization of the partial correlation coefficient for the pair (Y_1, Y_2) with the impact of $X_1, \ldots, X_n$ being eliminated. For $k = 1$ or for the system of random variables $(Y_1, Y_2, X_1, \ldots, X_n)$ formula (3.6) reduces to that obtained in [1]. (3.6) and (3.7) can be of some help when calculating the partial correlation coefficient for a system of random vectors $(Y_1, Y_2, (X_1, \ldots, X_n))$. The first of them uses linear regression while the second uses regression.

In [7] the notion of angle between linear subspaces and its measure were introduced. The values $\varrho_\Lambda(Y_i; X_1, \ldots, X_n)$ and $\widetilde{\varrho}_\Lambda(Y_i; X_1, \ldots, X_n)$ for $i = 1, 2$ are the cosines of the angles between $L(Y_i)$ and $\mathcal{H}_{X_1,\ldots,X_n}$ in the first case and between $L(Y_i)$ and $L^2_{X_1,\ldots,X_n}$ in the second case. This yields a geometric interpretation of (3.6) and (3.7).

References

[1] C. C. Li, S. Mazumdar and B. R. Rao, *Partial correlation in terms of path coefficients*, Amer. Statist. 29 (1975), 89–90.

[2] B. R. Rao and C. C. Li, *The geometry of path coefficients and correlations*, Biometrical J. 21 (1982), 673–678.

[3] C. R. Rao, *Linear Statistical Inference and its Applications*, Wiley, New York 1973.

[4] S. Wright, *The theory of path coefficients—a reply to Niks' criticism*, Genetics 8 (1923), 239–255.

[5] —, *The method of path coefficients*, Ann. Math. Statist. 5 (1934), 161–215.

[6] A. R. Sampson, *A multivariate correlation ratio*, Statist. Probab. Lett. 2 (1984), 77–81.

[7] W. Wysocki, *Geometrical aspects of measures of dependence for random vectors*, this issue, 211–224.

WLODZIMIERZ WYSOCKI
INSTITUTE OF COMPUTER SCIENCE
POLISH ACADEMY OF SCIENCES
P.O. BOX 22
00-901 WARSZAWA, POLAND

Received on 15.9.1989