Beiträge zur Algebra und Geometrie
Contributions to Algebra and Geometry
Volume 42 (2001), No. 1, 289-300.

Principal Values and Principal Subspaces of Two Subspaces of Vector Spaces with Inner Product

Ice B. Risteski **Kostadin G. Trenčevski**

Institute of Mathematics, St. Cyril and Methodius University
P.O.Box 162, 91000 Skopje, Macedonia
e-mail: kostatre@iunona.pmf.ukim.edu.mk

Abstract. In this paper is studied the problem concerning the angle between two subspaces of arbitrary dimensions in Euclidean space E_n. It is proven that the angle between two subspaces is equal to the angle between their orthogonal subspaces. Using the eigenvalues and eigenvectors of corresponding matrix representations, there are introduced principal values and principal subspaces. Their geometrical interpretation is also given together with the canonical representation of the two subspaces. The canonical matrix for the two subspaces is introduced and its properties of duality are obtained. Here obtained results expand the classic results given in [1,2].

MSC 2000: 15A03 (primary), 51N20 (secondary)
Keywords: angles between subspaces, principal values, principal subspaces, principal directions

1. Angle between two subspaces in E_n

We prove the following theorem which will enable us to define the angle between two subspaces of arbitrary dimensions of the Euclidean space E_n.

Theorem 1.1. *Let* $\mathbf{a}_1, \dots, \mathbf{a}_p$ *and* $\mathbf{b}_1, \dots, \mathbf{b}_q$ *are bases of two subspaces* Σ_1 *and* Σ_2 *of Euclidean space* E_n *with inner product* $(,)$ *respectively and suppose that* $p \leq q \leq n$. *Then the* *following inequality holds*

$$(1.1)\qquad det(MM^T) \le \begin{vmatrix} (\mathbf{a}_1,\mathbf{a}_1) & (\mathbf{a}_1,\mathbf{a}_2) & \cdots & (\mathbf{a}_1,\mathbf{a}_p) \\ (\mathbf{a}_2,\mathbf{a}_1) & (\mathbf{a}_2,\mathbf{a}_2) & \cdots & (\mathbf{a}_2,\mathbf{a}_p) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ (\mathbf{a}_p,\mathbf{a}_1) & (\mathbf{a}_p,\mathbf{a}_2) & \cdots & (\mathbf{a}_p,\mathbf{a}_p) \end{vmatrix} \times$$

$$\times \begin{vmatrix} (\mathbf{b}_1,\mathbf{b}_1) & (\mathbf{b}_1,\mathbf{b}_2) & \cdots & (\mathbf{b}_1,\mathbf{b}_q) \\ (\mathbf{b}_2,\mathbf{b}_1) & (\mathbf{b}_2,\mathbf{b}_2) & \cdots & (\mathbf{b}_2,\mathbf{b}_q) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ (\mathbf{b}_q,\mathbf{b}_1) & (\mathbf{b}_q,\mathbf{b}_2) & \cdots & (\mathbf{b}_q,\mathbf{b}_q) \end{vmatrix},$$

where

$$M = \begin{bmatrix} (\mathbf{a}_1,\mathbf{b}_1) & (\mathbf{a}_1,\mathbf{b}_2) & \cdots & (\mathbf{a}_1,\mathbf{b}_q) \\ (\mathbf{a}_2,\mathbf{b}_1) & (\mathbf{a}_2,\mathbf{b}_2) & \cdots & (\mathbf{a}_2,\mathbf{b}_q) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ (\mathbf{a}_p,\mathbf{b}_1) & (\mathbf{a}_p,\mathbf{b}_2) & \cdots & (\mathbf{a}_p,\mathbf{b}_q) \end{bmatrix}$$

and moreover equality holds if and only if Σ_1 *is subspace of* Σ_2.

Proof. The inequality (1.1) is invariant under any elementary row operation. Without loss of generality we can assume that $\{\mathbf{a}_1,\dots,\mathbf{a}_p\}$ is an orthonormal system and also $\{\mathbf{b}_1,\dots,\mathbf{b}_q\}$ is an orthonormal system. Then we should prove that

$$det(MM^T) \le 1.$$

Let denote

$$\mathbf{c}_i = ((\mathbf{a}_i,\mathbf{b}_1),(\mathbf{a}_i,\mathbf{b}_2),\dots,(\mathbf{a}_i,\mathbf{b}_q)) \in \mathbf{R}^q \quad (1 \le i \le p).$$

Since $\{\mathbf{b}_i\}$ and $\{\mathbf{a}_i\}$ are orthonormal systems we get that $\|\mathbf{c}_i\| \le 1$ with respect to the Euclidean metric in $\mathbf{R}^q$.

Let $\mathbf{c}_{p+1},\dots,\mathbf{c}_q$ be an orthonormal system of vectors such that each of them is orthogonal to $\mathbf{c}_1,\dots,\mathbf{c}_p$. Then

$$det(MM^T) = \begin{vmatrix} (\mathbf{c}_1\cdot\mathbf{c}_1) & (\mathbf{c}_1\cdot\mathbf{c}_2) & \cdots & (\mathbf{c}_1\cdot\mathbf{c}_p) \\ (\mathbf{c}_2\cdot\mathbf{c}_1) & (\mathbf{c}_2\cdot\mathbf{c}_2) & \cdots & (\mathbf{c}_2\cdot\mathbf{c}_p) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ (\mathbf{c}_p\cdot\mathbf{c}_1) & (\mathbf{c}_p\cdot\mathbf{c}_2) & \cdots & (\mathbf{c}_p\cdot\mathbf{c}_p) \end{vmatrix} =$$

$$= \begin{vmatrix} (\mathbf{c}_1 \cdot \mathbf{c}_1) & (\mathbf{c}_1 \cdot \mathbf{c}_2) & \cdots & (\mathbf{c}_1 \cdot \mathbf{c}_q) \\ (\mathbf{c}_2 \cdot \mathbf{c}_1) & (\mathbf{c}_2 \cdot \mathbf{c}_2) & \cdots & (\mathbf{c}_2 \cdot \mathbf{c}_q) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ (\mathbf{c}_q \cdot \mathbf{c}_1) & (\mathbf{c}_q \cdot \mathbf{c}_2) & \cdots & (\mathbf{c}_q \cdot \mathbf{c}_q) \end{vmatrix}$$

which is the square of the volume of the parallelotop in $\mathbf{R}^q$ generated by the vectors $\mathbf{c}_1, \dots, \mathbf{c}_q$. Since $\|\mathbf{c}_i\| \leq 1$, $(1 \leq i \leq q)$ we obtain $det(MM^T) \leq 1$.

Moreover, equality holds if and only if $\mathbf{c}_1, \dots, \mathbf{c}_q$ is an orthonormal system. But $\|\mathbf{c}_i\| = 1$ implies that $\mathbf{a}_i$ belongs to the subspace Σ_2. Thus $\Sigma_1 \subseteq \Sigma_2$. Conversely, if $\Sigma_1 \subseteq \Sigma_2$ then it is trivial that equality holds in (1.1). □

Under the assumptions of Theorem 1.1 we define the angle φ between Σ_1 and Σ_2 by

$$\cos\varphi = \frac{\sqrt{det(MM^T)}}{\sqrt{\Gamma_1} \cdot \sqrt{\Gamma_2}} \tag{1.2}$$

where the matrix M was defined in Theorem 1.1 and Γ_1 and Γ_2 are the Gram's determinants obtained by the vectors $\mathbf{a}_1, \dots, \mathbf{a}_p$ and $\mathbf{b}_1, \dots, \mathbf{b}_q$ respectively.

Note that $det(MM^T) \geq 0$; considering both values of $\sqrt{det(MM^T)}$, we obtain two angles φ and $\pi - \varphi$. Note that $det(MM^T) = 0$ if $q < p$.

In this paper we give some deeper results concerning the Theorem 1.1. Indeed, some theorems which yield to principal directions on both subspaces Σ_1 and Σ_2 and common principal values are proven.

In the next research will be used the following result.

Theorem 1.2. *Let U be any $p \times q$ matrix. Any nonzero scalar λ is an eigenvalue of the square matrix UU^T if and only if it is eigenvalue of the square matrix U^TU and moreover the multiplicities of λ for both matrices UU^T and U^TU are equal.*

Proof. Assume that $\lambda \neq 0$ is an eigenvalue of UU^T with geometrical multiplicity r and assume that $\mathbf{x}_1, \dots, \mathbf{x}_r$ are linearly independent eigenvectors corresponding to λ. Then we will prove that the vectors

$$\mathbf{y}_i = U^T\mathbf{x}_i, \quad (1 \leq i \leq r)$$

are linearly independent eigenvectors for the matrix U^TU. Indeed,

$$U^TU\mathbf{y}_i = (U^TU)U^T\mathbf{x}_i = U^T(UU^T\mathbf{x}_i) = \lambda U^T\mathbf{x}_i = \lambda\mathbf{y}_i$$

and thus $\mathbf{y}_i$ are eigenvectors of U^TU corresponding to the eigenvalue λ.

Now let us assume that $\alpha_1\mathbf{y}_1 + \cdots + \alpha_r\mathbf{y}_r = 0$, then multiplying this equality by U from left we obtain

$$\lambda\alpha_1\mathbf{x}_1 + \cdots + \lambda\alpha_r\mathbf{x}_r = 0.$$

Since $\lambda \neq 0$ we obtain

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_r\mathbf{x}_r = 0$$

and hence $\alpha_1 = \cdots = \alpha_r = 0$ because $\mathbf{x}_1, \ldots, \mathbf{x}_r$ are linearly independent vectors.

Hence the geometric multiplicity of λ for the matrix UU^T is smaller or equal to the geometric multiplicity of λ for the matrix U^TU. Analogously, the geometric multiplicity of λ for the matrix U^TU is smaller or equal to the geometric multiplicity of λ for the matrix UU^T. Thus these two geometrical multiplicities are equal. Since UU^T and U^TU are symmetric non-negative definite matrices, we obtain that their geometrical multiplicities are equal to the algebraic multiplicities. □

Now we are enabled to prove the following theorem.

Theorem 1.3. *If Σ_1 and Σ_2 are any subspaces of the Euclidean vector space E_n and Σ_1^* and Σ_2^* are their orthogonal complements, then*

$$\varphi(\Sigma_1, \Sigma_2) = \varphi(\Sigma_1^*, \Sigma_2^*).$$

Proof. Assume that $dim\Sigma_1 = p$ and $dim\Sigma_2 = q$. Without loss of generality we assume that $p \leq q$ and assume that Σ_1 is generated by $\mathbf{e}_i$, $(1 \leq i \leq p)$ and Σ_1^* is generated by $\mathbf{e}_j$, $(p+1 \leq j \leq n)$ where $\mathbf{e}_i$, $(1 \leq i \leq n)$ is the standard basis of E_n. Further without loss of generality we can assume that Σ_2 is generated by $\mathbf{a}_i$, $(1 \leq i \leq q)$ and Σ_2^* is generated by $\mathbf{a}_j$, $(q+1 \leq j \leq n)$, where $\mathbf{a}_i$, $(1 \leq i \leq n)$ is an orthonormal system of vectors. Let $\mathbf{a}_i$ have coordinates $(a_{i1}, a_{i2}, \ldots, a_{in})$, $(1 \leq i \leq n)$ and the matrix with row vectors $\mathbf{a}_1, \cdots, \mathbf{a}_n$ will be denoted by A. We denote by X, Y and Z the following submatrices of A: X is the submatrix of A with elements a_{ij}, $(1 \leq i \leq p;\ 1 \leq j \leq q)$; Y is the submatrix of A with elements a_{ij}, $(1 \leq i \leq p;\ q+1 \leq j \leq n)$; Z is the submatrix of A with elements a_{ij}, $(p+1 \leq i \leq n;\ q+1 \leq j \leq n)$. According to these assumptions

$$\cos^2 \varphi(\Sigma_1, \Sigma_2) = det(XX^T)$$

and

$$\cos^2 \varphi(\Sigma_1^*, \Sigma_2^*) = det(Z^TZ)$$

and we should prove that

$$det(XX^T) = det(Z^TZ).$$

Since A is an orthogonal matrix, it holds

$$XX^T = I_{p\times p} - YY^T \quad \text{and} \quad Z^TZ = I_{(n-q)\times(n-q)} - Y^TY$$

and we should prove that

$$det(I_{p\times p} - YY^T) = det(I_{(n-q)\times(n-q)} - Y^TY).$$

Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of YY^T and $\mu_1, \ldots, \mu_{n-q}$ be the eigenvalues of Y^TY. According to Theorem 1.2, the matrices YY^T and Y^TY have the same non-zero eigenvalues with the same multiplicities and hence

$$det(I_{p\times p} - YY^T) = (1-\lambda_1)\cdots(1-\lambda_p) =$$

$$= (1-\mu_1)\cdots(1-\mu_q) = det(I_{(n-q)\times(n-q)} - Y^TY).$$ □

2. Principal values and principal subspaces

First we prove the following statement.

Theorem 2.1. *Let Σ_1 and Σ_2 be two vector subspaces of the Euclidean space E_n of dimensions p and q, $(p \leq q)$ and let A_1 and A_2 be $n \times p$ and $n \times q$ matrices whose vector rows generate the subspace Σ_1 and Σ_2 respectively. Then the eigenvalues of the matrix*

$$f(A_1, A_2) = A_1 A_2^T (A_2 A_2^T)^{-1} A_2 A_1^T (A_1 A_1^T)^{-1}$$

are p canonical squares $\cos^2 \varphi_i$, $(1 \leq i \leq p)$ and moreover

$$\cos^2 \varphi = \prod_{i=1}^{p} \cos^2 \varphi_i,$$

where φ is the angle between the subspaces Σ_1 and Σ_2.

Proof. The transition of the base of Σ_j to another base corresponds to multiplication of A_j by nonsingular matrix P_j, i.e. $A_j \to P_j A_j$, where P_1 is $p \times p$ matrix and P_2 is $q \times q$ matrix. By direct calculation one verifies that

$$f(P_1 A_1, P_2 A_2) = P_1 f(A_1, A_2) P_1^{-1}$$

and thus the eigenvalues are unchanged. Moreover, $f(A_1, A_2)$ is unchanged under the transformation of form $A_j \to A_j R$ where R is any orthogonal matrix of n-th order, which means that $f(A_1, A_2)$ is invariant under the change of the rectangular Cartesian coordinates in the Euclidean space E_n.

Since $A_1 A_1^T$ and $A_2 A_2^T$ are positive definite matrices, there exist symmetric positive definite matrices P_1 and P_2 of orders p and q respectively such that

$$P_1 A_1 A_1^T P_1^T = B_1 B_1^T = I_{p \times p} \quad \text{and} \quad P_2 A_2 A_2^T P_2^T = B_2 B_2^T = I_{q \times q},$$

where B_1 and B_2 correspond to another bases of Σ_1 and Σ_2. Since $S = (B_1 B_2^T)(B_1 B_2^T)^T$ is non-negative definite matrix, there exists a symmetric non-negative definite orthogonal matrix Q_1 of order p such that $Q_1 S Q_1^{-1}$ is diagonalized, i.e.

$$Q_1 S Q_1^{-1} = (C_1 B_2^T)(C_1 B_2^T)^T = diag(c_1^2, c_2^2, \ldots, c_p^2), \quad (c_1 \geq c_2 \geq \cdots \geq c_p \geq 0)$$

where $C_1 = Q_1 B_1$ corresponds to another basis of Σ_1. Having in mind that each c_i is an inner product of two unimodular vectors, we get $c_i = \cos \varphi_i$, $0 \leq \varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_p \leq \pi/2$. The vector rows of $C_1 B_2^T$ are mutually orthogonal, which means that there exists an orthogonal matrix Q_2 of order q, such that

$$C_1 B_2^T Q_2^T = C_1 C_2^T = \cos \varphi_i \delta_{ik},$$

where $C_2 = Q_2 B_2$ corresponds to another orthonormal base of Σ_2. This shows that the ordered set of angles $\varphi_1, \varphi_2, \ldots, \varphi_p$ is canonical and its invariance follows from the decomposition

$$det[\lambda I_{p \times p} - f(C_1, C_2)] = \prod_{i=1}^{p} (\lambda - \cos^2 \varphi_i) = det[\lambda I_{p \times p} - f(A_1, A_2)].$$

Finally note that according to the chosen bases of Σ_1 and Σ_2, we obtain

$$\cos^2\varphi = det(f(C_1, C_2)) = det(f(A_1, A_2)) = \prod_{i=1}^{p} \cos^2\varphi_i$$

where φ is the angle between the subspaces Σ_1 and Σ_2. □

Note that if the bases of Σ_1 and Σ_2 are orthonormal then $A_1A_1^T = A_2A_2^T = I$ and $f(A_1, A_2) = A_1A_2^T(A_1A_2^T)^T$.

Now let us consider the case $p \geq q$. Instead of the matrix $f(A_1, A_2)$ we should consider the matrix $f(A_2, A_1)$ which is of type $q \times q$. Analogously to Theorem 2.1 the eigenvalues of $f(A_2, A_1)$ are q canonical squares of cosine functions but the product of them is equal to zero if $p > q$. Now we prove the following theorem considering the mutually eigenvalues of $f(A_1, A_2)$ and $f(A_2, A_1)$.

Theorem 2.2. *Any nonzero scalar λ is an eigenvalue of $f(A_1, A_2)$ if and only if it is eigenvalue of $f(A_2, A_1)$ and moreover the multiplicities of λ for both matrices $f(A_1, A_2)$ and $f(A_2, A_1)$ are equal.*

Proof. Let C_1 and C_2 have the same meanings like in the Theorem 2.1. According to Theorem 1.2 we obtain that any nonzero scalar λ is an eigenvalue of $f(C_1, C_2)$ if and only if it is eigenvalue of $f(C_2, C_1)$ and moreover the multiplicities of λ for both matrices $f(C_1, C_2)$ and $f(C_2, C_1)$ are equal, because $f(C_1, C_2) = (C_1C_2^T)(C_1C_2^T)^T$. On the other hand, $f(A_1, A_2)$ is the same eigenvalues as $f(C_1, C_2)$ with the same multiplicity and $f(A_2, A_1)$ is the same eigenvalues as $f(C_2, C_1)$ with the same multiplicity. □

Note that $\lambda = 0$ is eigenvalue for the matrix $f(A_2, A_1)$ if $q > p$, but $\lambda = 0$ may not be eigenvalue for the matrix $f(A_1, A_2)$.

The common eigenvalues will be called *principal values.* According to the Theorems 2.1 and 2.2 there are unique decompositions of the subspaces Σ_1 and Σ_2 into the orthogonal eigenspaces for the common non-negative eigenvalues and for the zero eigenvalue if such exists. These eigenspaces are called *principal subspaces* or *principal directions* for the eigenvalues with multiplicity 1. The geometrical interpretation of the principal values and principal subspaces will be given after the proof of the Theorem 2.3.

Theorem 2.3. *The function $\cos^2\varphi$, where φ is the angle between any vector $\mathbf{x} \in \Sigma_1$ and the subspace Σ_2, has maximum if and only if the vector $\mathbf{x}$ belongs to a principal subspace of Σ_1 which corresponds to the maximal principal value. The maximal value of $\cos^2\varphi$ is the maximal principal value.*

Proof. According to the proof of Theorem 2.1, without loss of generality we can suppose that Σ_1 is generated by the orthonormal vectors $\mathbf{a}_i$, $(1 \leq i \leq p)$ and Σ_2 is generated by the orthonormal vectors $\mathbf{b}_j$, $(1 \leq j \leq q)$ such that $(\mathbf{a}_i, \mathbf{b}_j) = 0$, $(i \neq j;\ 1 \leq i \leq p,\ 1 \leq j \leq q)$. Let $\mathbf{x} = \alpha_1\mathbf{a}_1 + \cdots + \alpha_p\mathbf{a}_p$, let $\lambda_1^2 = \mathbf{a}_1\mathbf{b}_1$ be the maximal principal value and the corresponding subspace of Σ_1 be generated by $\mathbf{a}_1, \dots, \mathbf{a}_r$. Then for the angle φ between $\mathbf{x}$ and Σ_2 it holds

$$\cos^2\varphi = \frac{(\alpha_1\lambda_1)^2 + \cdots + (\alpha_p\lambda_s)^2}{\alpha_1^2 + \cdots + \alpha_p^2} =$$

$$= \frac{\lambda_1^2(\alpha_1^2 + \cdots + \alpha_r^2) + \lambda_{r+1}^2(\cdots) + \cdots}{\alpha_1^2 + \cdots + \alpha_p^2} \le \lambda_1^2$$

and equality holds if and only if $\alpha_{r+1} = \cdots = \alpha_p = 0$, i.e. if and only if $\mathbf{x}$ belongs to the eigenspace corresponding to λ_1. □

Note that an analogous statement like Theorem 2.3 holds also if we consider $\mathbf{x}$ as vector of Σ_2 and φ is the angle between $\mathbf{x}$ and Σ_1. Thus we obtain the following geometrical interpretation:

Among all values $\cos^2\varphi$ where φ is angle between any vector $\mathbf{x} \in \Sigma_1$ and any vector $\mathbf{y} \in \Sigma_2$, the maximal value λ_1^2 is the first (maximal) principal value. Then

$$\Sigma_{11} = \{\mathbf{x} \in \Sigma_1 | \cos^2(\mathbf{x}, \Sigma_2) = \lambda_1^2\}$$

is the the principal subspace of Σ_1. Analogously

$$\Sigma_{21} = \{\mathbf{y} \in \Sigma_2 | \cos^2(\mathbf{y}, \Sigma_1) = \lambda_1^2\}$$

is the principal subspace of Σ_2 and moreover $dim\Sigma_{11} = dim\Sigma_{21}$. Now let us consider the subspaces Σ_1' and Σ_2' where Σ_1' is orthogonal complement of Σ_{11} in Σ_1 and Σ_2' is orthogonal complement of Σ_{21} in Σ_2. Among all values $\cos^2\varphi$ where φ is angle between any vector $\mathbf{x} \in \Sigma_1'$ and any vector $\mathbf{y} \in \Sigma_2'$, the maximal value λ_2^2 is the second principal value. Then

$$\Sigma_{12} = \{\mathbf{x} \in \Sigma_1' | \cos^2(\mathbf{x}, \Sigma_2') = \lambda_2^2\}$$

is the principal subspace of Σ_1'. Analogously

$$\Sigma_{22} = \{\mathbf{y} \in \Sigma_2' | \cos^2(\mathbf{y}, \Sigma_1') = \lambda_2^2\}$$

is the principal subspace of Σ_2' and moreover $dim\Sigma_{12} = dim\Sigma_{22}$. Continuing this procedure we obtain the decompositions of orthogonal principal subspaces

$$\Sigma_1 = \Sigma_{11} + \Sigma_{12} + \cdots + \Sigma_{1,s+1}$$

$$\Sigma_2 = \Sigma_{21} + \Sigma_{22} + \cdots + \Sigma_{2,s+1}$$

where $dim\Sigma_{1i} = dim\Sigma_{2i}$, $(1 \le i \le s)$. The subspaces $\Sigma_{1,s+1}$ and $\Sigma_{2,s+1}$ correspond for the possible value 0 as a principal value.

Example. Let Σ_1 be generated by the vectors $(1,0,0,0)$ and $(0,1,0,0)$ and Σ_2 be generated by $(\cos\varphi, 0, \sin\varphi, 0)$ and $(0, \cos\varphi, 0, \sin\varphi)$. Then $\cos^2\varphi$ is unique principal value, its multiplicity is 2 and Σ_1 and Σ_2 are principal subspaces themselves.

At the end we prove a theorem which determines the orthogonal projection of any vector $\mathbf{x}$ on any subspace of E_n.

Theorem 2.4. *In the n-dimensional Euclidean space E_n let be given a subspace Σ generated by k linearly independent vectors $\mathbf{a}_i$, $(1 \le i \le k;\ k \le n-1)$. The orthogonal projection $\mathbf{x}'$ of* *an arbitrary vector* $\mathbf{x}$ *of* E_n *is given by*

$$\mathbf{x}' = -\frac{1}{\Gamma}\begin{vmatrix} \mathbf{0} & (\mathbf{x},\mathbf{a}_1) & (\mathbf{x},\mathbf{a}_2) & \cdots & (\mathbf{x},\mathbf{a}_k) \\ \mathbf{a}_1 & (\mathbf{a}_1,\mathbf{a}_1) & (\mathbf{a}_1,\mathbf{a}_2) & \cdots & (\mathbf{a}_1,\mathbf{a}_k) \\ \mathbf{a}_2 & (\mathbf{a}_2,\mathbf{a}_1) & (\mathbf{a}_2,\mathbf{a}_2) & \cdots & (\mathbf{a}_2,\mathbf{a}_k) \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \mathbf{a}_k & (\mathbf{a}_k,\mathbf{a}_1) & (\mathbf{a}_k,\mathbf{a}_2) & \cdots & (\mathbf{a}_k,\mathbf{a}_k) \end{vmatrix}, \tag{2.1}$$

where Γ *is the Gram's determinant of the vectors* $\mathbf{a}_i$, $(1 \le i \le k)$.

Proof. According to (2.1) it is obvious that

$$\mathbf{x} - \mathbf{x}' = \frac{1}{\Gamma}\begin{vmatrix} \mathbf{x} & (\mathbf{x},\mathbf{a}_1) & (\mathbf{x},\mathbf{a}_2) & \cdots & (\mathbf{x},\mathbf{a}_k) \\ \mathbf{a}_1 & (\mathbf{a}_1,\mathbf{a}_1) & (\mathbf{a}_1,\mathbf{a}_2) & \cdots & (\mathbf{a}_1,\mathbf{a}_k) \\ \mathbf{a}_2 & (\mathbf{a}_2,\mathbf{a}_1) & (\mathbf{a}_2,\mathbf{a}_2) & \cdots & (\mathbf{a}_2,\mathbf{a}_k) \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \mathbf{a}_k & (\mathbf{a}_k,\mathbf{a}_1) & (\mathbf{a}_k,\mathbf{a}_2) & \cdots & (\mathbf{a}_k,\mathbf{a}_k) \end{vmatrix}.$$

By scalar multiplication of this equality by $\mathbf{a}_i$, $(1 \le i \le k)$ the first column is equal to the $(i+1)$-st column and thus

$$(\mathbf{x} - \mathbf{x}', \mathbf{a}_i) = 0, \quad (1 \le i \le k).$$

Since $\mathbf{x}'$ is a linear combination of the vectors $\mathbf{a}_i$, $(1 \le i \le k)$ then the vector $\mathbf{x}'$ lies in Σ. Moreover, $\mathbf{x} - \mathbf{x}'$ is orthogonal to the base vectors of Σ, we obtain that $\mathbf{x}'$ is the required orthogonal projection of $\mathbf{x}$ on the subspace Σ. □

3. Principle of duality and canonical form

In this section we will consider the duality principle like in the Theorem 1.3 and as a crown of all previous research will be given the canonical form of two subspaces Σ_1 and Σ_2. Now let Σ_i^* denote the orthogonal subspace of Σ_i, $(i = 1, 2)$ in the Euclidean space E_n. We saw that $\varphi(\Sigma_1, \Sigma_2) = \varphi(\Sigma_1^*, \Sigma_2^*)$ and now the same conclusions for the eigenvalues and principal subspaces (principal directions) also hold for the subspaces Σ_1^* and Σ_2^*.

Theorem 3.1. *If* Σ_1 *and* Σ_2 *are any subspaces of the Euclidean vector space* E_n *and* Σ_1^* *and* Σ_2^* *are their orthogonal complements, then the nonzero and different from* 1 *principal values for the pair* (Σ_1, Σ_2) *are the same for the pair* (Σ_1^*, Σ_2^*) *with the same multiplicities and conversely.*

If $p + q \le n$*, then the multiplicity of* 1 *for the pair* (Σ_1^*, Σ_2^*) *is bigger for* $n - p - q$ *than the multiplicity of* 1 *for the pair* (Σ_1, Σ_2).

If $p+q \geq n$, then the multiplicity of 1 *for the pair* (Σ_1, Σ_2) *is bigger for* $p+q-n$ *than the multiplicity of* 1 *for the pair* (Σ_1^*, Σ_2^*).

Proof. We use the same notations and assumptions as in the proof of the Theorem 1.3. Specially, the matrices X, Y and Z are the same. Assume that $p+q \leq n$. The case $n > p+q$ can be discussed analogously.

We will prove the following identity

$$det(\lambda I_{p\times p} - XX^T) \cdot (\lambda - 1)^{n-q-p} = det(\lambda I_{(n-q)\times(n-q)} - Z^T Z)$$

and hence the proof will be finished.

Since A is an orthogonal matrix, it holds

$$XX^T = I_{p\times p} - YY^T \quad \text{and} \quad Z^T Z = I_{(n-q)\times(n-q)} - Y^T Y$$

and we should prove that

$$det((\lambda-1)I_{p\times p} + YY^T) \cdot (\lambda-1)^{n-q-p} = det((\lambda-1)I_{(n-q)\times(n-q)} + Y^T Y).$$

Multiplying this equality by $(-1)^{n-q}$ and putting $1-\lambda = \mu$, we should prove that

$$det(\mu I_{p\times p} - YY^T) \cdot \mu^{n-q-p} = det(\mu I_{(n-q)\times(n-q)} - Y^T Y).$$

Let $\mu_1, \ldots, \mu_p$ be the eigenvalues of YY^T. According to Theorem 1.2, both sides of the last equality are equal to

$$(\mu-\mu_1)(\mu-\mu_2)\cdots(\mu-\mu_p)\mu^{n-q-p}. \qquad \square$$

According to Theorem 3.1 we obtain the following consequence.

Corollary 3.2. *According to the notations of Theorem 3.1,*

i) *the number of nonzero and nonunit principal values (each value counts as many times as its multiplicity) of the pair* (Σ_1, Σ_2) *is less or equal to* $n/2$;

ii) *if* n *is an odd number and* $p = q$, *then at least one of the pairs* (Σ_1, Σ_2) *and* (Σ_1^*, Σ_2^*) *has a principal value* 1, *i.e. they have a common subspace of dimension* ≥ 1.

Now we are able to give the canonical form of two subspaces. In order to avoid many indices we assume that the considered subspaces of E_n are Σ and Π with dimensions p and q respectively. We denote by Σ^* and Π^* the orthogonal subspaces of E_n. Without loss of generality we assume that $p \leq q$. Since the canonical form is according to these four subspaces, we can also assume that $p+q \leq n$. Indeed, if $p+q > n$ then $(n-p)+(n-q) < n$ and we can consider the subspaces Σ^* and Π^*.

Assume that $1 = c_0 > c_1 > c_2 > \cdots > c_s > c_{s+1} = 0$ be the principal values for the pair (Σ, Π) with multiplicities $r_0, r_1, \ldots, r_{s+1}$ respectively, such that $p = r_0 + r_1 + \cdots + r_{s+1}$. Let Σ be generated by the following orthonormal vectors

$$\mathbf{a}_{01}, \ldots, \mathbf{a}_{0r_0}, \mathbf{a}_{11}, \ldots, \mathbf{a}_{1r_1}, \ldots, \mathbf{a}_{s1}, \ldots, \mathbf{a}_{sr_s}, \mathbf{a}_{s+1,1}, \ldots, \mathbf{a}_{s+1,r_{s+1}},$$

such that the vectors $\mathbf{a}_{i1}, \dots, \mathbf{a}_{ir_i}$ generate the principal subspace for the principal value c_i, $(0 \le i \le s+1)$. The pair of subspaces (Σ^*, Π^*) have the same principal values $1 = c_0 > c_1 > c_2 > \cdots > c_s > c_{s+1} = 0$ with multiplicities $r_0' = r_0 + n - p - q, r_1, \dots, r_{s+1}$. Assume that Σ^* is generated by the following orthonormal vectors

$$\mathbf{a}_{01}^*, \dots, \mathbf{a}_{0r_0'}^*, \mathbf{a}_{11}^*, \dots, \mathbf{a}_{1r_1}^*, \dots, \mathbf{a}_{s1}^*, \dots, \mathbf{a}_{sr_s}^*, \mathbf{a}_{s+1,1}^*, \dots, \mathbf{a}_{s+1,r_{s+1}}^*, \mathbf{a}_1^*, \dots, \mathbf{a}_{q-p}^*$$

where the vectors $\mathbf{a}_{i1}, \dots, \mathbf{a}_{ir_i}$ generate the principal subspace for the principal value c_i, $(1 \le i \le s+1)$, $\mathbf{a}_{01}, \dots, \mathbf{a}_{0r_0'}$ generate the principal subspace for the principal value 1 and $\mathbf{a}_1^*, \dots, \mathbf{a}_{q-p}^*$ be the remaining $q-p$ orthonormal vectors.

Now we chose the orthonormal vectors of Π as follows. We chose

$$\mathbf{b}_{01}, \dots, \mathbf{b}_{0r_0}, \mathbf{b}_{11}, \dots, \mathbf{b}_{1r_1}, \dots, \mathbf{b}_{s1}, \dots, \mathbf{b}_{sr_s}, \mathbf{b}_{s+1,1}, \dots, \mathbf{b}_{s+1,r_{s+1}}, \mathbf{b}_1, \dots, \mathbf{b}_{q-p}$$

such that $\mathbf{b}_{0i}$ coincides with $\mathbf{a}_{0i}$, $(1 \le i \le r_0)$, $\mathbf{b}_{i1}, \dots, \mathbf{b}_{ir_i}$ generate the principal subspace for the principal value c_i, $(1 \le i \le s)$ and such that $(\mathbf{a}_{iu}, \mathbf{b}_{iv}) = \delta_{uv} c_i$. The vectors $\mathbf{b}_{s+1,1}, \dots, \mathbf{b}_{s+1,r_{s+1}}$ generate the same subspace as the vectors $\mathbf{a}_{s+1,1}^*, \dots, \mathbf{a}_{s+1,r_{s+1}}^*$ and we can choose $\mathbf{b}_{s+1,i} = \mathbf{a}_{s+1,i}^*$, $(1 \le i \le r_{s+1})$. The vectors $\mathbf{b}_1, \dots, \mathbf{b}_{q-p}$ generate the same space as the vectors $\mathbf{a}_1^*, \dots, \mathbf{a}_{q-p}^*$ and we can choose $\mathbf{b}_i = \mathbf{a}_{q-p+1-i}^*$, $(1 \le i \le q-p)$.

Finally we determine the orthonormal vectors of Π^*

$$\mathbf{b}_{01}^*, \dots, \mathbf{b}_{0r_0'}^*, \mathbf{b}_{11}^*, \dots, \mathbf{b}_{1r_1}^*, \dots, \mathbf{b}_{s1}^*, \dots, \mathbf{b}_{sr_s}^*, \mathbf{b}_{s+1,1}^*, \dots, \mathbf{b}_{s+1,r_{s+1}}^*$$

as follows. The vectors $\mathbf{b}_{01}^*, \dots, \mathbf{b}_{0r_0'}^*$ can be chosen such that $\mathbf{b}_{0i}^* = \mathbf{a}_{0i}^*$, $(1 \le i \le r_0')$. The vectors $\mathbf{b}_{i1}^*, \dots, \mathbf{b}_{ir_i}^*$ generate the principal subspace for the principal value c_i, $(1 \le i \le s)$, and the vectors $\mathbf{b}_{i1}^*, \dots, \mathbf{b}_{ir_i}^*$ can uniquely be chosen such that $(\mathbf{a}_{iu}^*, \mathbf{b}_{iv}^*) = \delta_{uv} c_i$. The vectors $\mathbf{b}_{s+1,1}^*, \dots, \mathbf{b}_{s+1,r_{s+1}}^*$ generate the same subspace as the vectors $\mathbf{a}_{s+1,1}^*, \dots, \mathbf{a}_{s+1,r_{s+1}}^*$ and thus we can choose $\mathbf{b}_{s+1,i}^* = \mathbf{a}_{s+1,i}^*$, $(1 \le i \le r_{s+1})$.

Moreover, the vectors $\mathbf{a}_{11}^*, \dots, \mathbf{a}_{1r_1}^*, \dots, \mathbf{a}_{s1}^*, \dots, \mathbf{a}_{sr_s}^*$ can be chosen such that

$$(\mathbf{a}_{iu}^*, \mathbf{b}_{iv}) = -\delta_{uv}\sqrt{1 - c_i^2}, \quad (1 \le i \le s).$$

Now we know some of the inner products between the base vectors of Σ and Σ^* and the base vectors of Π and Π^*. The matrix P of all such $n \times n$ inner products must be orthogonal and can uniquely be obtained from the above inner products. Considering the base vectors of Σ in the mentioned order together with the base vectors of Σ^* in the opposite order and on the other side the base vectors of Π in the mentioned order together with the base vectors of Π^* in the opposite order we obtain the following

$$(r_0 + r_1 + r_2 + \cdots + r_s + r_{s+1} + (q-p) + r_{s+1} + r_s + \cdots + r_2 + r_1 + r_0')\times$$

$$\times(r_0 + r_1 + r_2 + \cdots + r_s + r_{s+1} + (q-p) + r_{s+1} + r_s + \cdots + r_2 + r_1 + r_0')$$

matrix as *canonical matrix* for the subspaces Σ and Π:

$$P = \begin{bmatrix} I & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & c_1 I & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & d_1 I' & 0 \\ 0 & 0 & c_2 I & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & d_2 I' & 0 & 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 & 0 & 0 & \cdots & c_s I & 0 & 0 & 0 & d_s I' & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & I' & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & I & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & I' & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & -d_s I' & 0 & 0 & 0 & c_s I & \cdots & 0 & 0 & 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 & 0 & -d_2 I' & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & c_2 I & 0 & 0 \\ 0 & -d_1 I' & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & c_1 I & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & I \end{bmatrix},$$

where $d_i = \sqrt{1 - c_i^2}$, $(1 \le i \le s)$ and I' denotes the matrix with 1 on the opposite diagonal of the main diagonal and the other elements are zero.

Note that the principal values for the pair (Σ, Π^*) (also (Σ^*, Π)) are the numbers $d_i^2 = 1 - c_i^2 = \sin^2 \varphi_i$ with the same multiplicities as c_i^2. Moreover the previous canonical matrix P is also canonical matrix for the pair (Σ, Π^*) (also (Σ^*, Π)) if we permute its rows and columns. Then the order $q - p$ converts into $n - p - q$ and vice versa.

The previous consideration yields to the following statement.

Theorem 3.3. *Let n, p, q be positive integers such that $n \le p + q$ and $p \le q$. Then for any p values $c_1^2, \dots, c_p^2$, $(0 \le c_i \le 1)$ there exist two subspaces Σ_1 and Σ_2 of E_n with dimensions p and q such that $c_1^2, \dots, c_p^2$ are principal values for the pair (Σ_1, Σ_2). The existence of the subspaces Σ_1 and Σ_2 is uniquely up to orthogonal motion in E_n.*

Proof. Let n, p, q be positive integers such that $n \le p + q$ and $p \le q$ and let be given p values c_i^2, $(0 \le c_i \le 1)$. We choose arbitrary orthonormal base $\mathbf{a}_1, \dots, \mathbf{a}_p, \mathbf{a}_{n-p}^*, \dots, \mathbf{a}_1^*$ of E_n. Then we introduce q vectors $\mathbf{b}_1, \dots, \mathbf{b}_q$ whose coordinates with respect to $\mathbf{a}_1, \dots, \mathbf{a}_p, \mathbf{a}_{n-p}^*, \dots, \mathbf{a}_1^*$ are given by the first q columns of the matrix P. Then it is obvious that the principal values for the pair (Σ_1, Σ_2) where Σ_1 is generated by $\mathbf{a}_1, \dots, \mathbf{a}_p$ and Σ_2 is generated by the vectors $\mathbf{b}_1, \dots, \mathbf{b}_q$ are just the given numbers $c_1^2, \dots, c_p^2$.

Let (Σ_1, Σ_2) and (Σ_1', Σ_2') be two pairs of subspaces with the same principal values. Without loss of generality we assume that both of them are given in canonical form given by the same canonical matrix P. Let

$$\{\mathbf{a}_1, \dots, \mathbf{a}_p, \mathbf{a}_1^*, \dots, \mathbf{a}_{n-p}^*\} \quad \text{and} \quad \{\mathbf{a}_1', \dots, \mathbf{a}_p', \mathbf{a}_1'^*, \dots, \mathbf{a}_{n-p}'^*\}$$

be the base vectors of $\Sigma_1 + \Sigma_1^*$ and $\Sigma_1' + \Sigma_1'^*$ corresponding to their canonical forms. Since the base vectors of $\Sigma_2 + \Sigma_2^*$ and $\Sigma_2' + \Sigma_2'^*$ are determined uniquely, it is sufficient to choose the orthogonal transformation φ which maps the mentioned base of $\Sigma_1 + \Sigma_1^*$ into the mentioned base of $\Sigma_1' + \Sigma_1'^*$ and then $\varphi(\Sigma_1) = \Sigma_1'$ and $\varphi(\Sigma_2) = \Sigma_2'$. $\square$

Theorem 3.4. *Let A be a symmetric matrix of n-th order. Assume that the linear subspace L of E_n such that A is positive definite matrix in L and A^{-1} is positive definite matrix in the orthogonal complement L^*, then A is positive definite matrix.*

Proof. If $A|L$ denotes the restriction of A to L, and by $ind(A|L)$ is denoted the number of negative eigenvalues of V^TAV, where V is the matrix of the base of L, then the following lemma holds.

Lemma 3.5. *Let A be a symmetric nonsingular matrix of n-th order, and let L and L^* be the same notations as in Theorem 3.4. If $A^{-1}|L^*$ is a nonsingular restriction, then also the restriction $A|L$ is nonsingular and moreover*

$$ind(A|E_n) = ind(A|L) + ind(A^{-1}|L^*).$$

The Theorem 3.4 obtains for the special case

$$ind(A|L) = ind(A^{-1}|L^*) = 0.$$

Proof of Lemma 3.5. Let V and W denote the matrices from the bases of L and L^* respectively. Then $B = AVW$ is nonsingular matrix. Indeed, it is supposed that $AV\mathbf{x} = W\mathbf{y}$ for the vectors $\mathbf{x}$ and $\mathbf{y}$. Multiplying this equality by W^*A^{-1} from left, we obtain $W^*A^{-1}W\mathbf{y} = 0$, because $V^*W = 0$. This implies $\mathbf{y} = 0$ which means that $W^*A^{-1}W$ is nonsingular matrix. Consequently, $W\mathbf{x} = A^{-1}W\mathbf{y} = 0$ implies $\mathbf{x} = 0$. It implies that

$$ind(A|E_n) = ind(A^{-1}|E_n) = ind(B^TA^{-1}B|E_n) =$$

$$= ind(A|L) + ind(A^{-1}|L^*).$$ $\square$

References

[1] Halmos, P. R.: *Finite Dimensional Vector Spaces.* 2nd ed. Van Nostrand Reinhold, New York 1958.

[2] Kurepa, S.: *Finite Dimensional Vector Spaces and Applications.* Sveučilišna Naklada Liber, Zagreb 1979 (in Croatian).

Received January 31, 2000