Mathematica Pannonica
9/2 (1998), 293–299

A GENERALIZATION OF THE CAUCHY–SCHWARTZ INEQUALITY

Pavel **Pech**

Umversity of South Bohemia, Pedagogical Faculty, Jeronýmova 10, 37115 České Budějovice, Czech Republic

Received: June 1998

MSC 1991: 52 A 40

Keywords: Cauchy-Schwarz inequality, angle of linear subspaces.

Abstract: A generalization of the Cauchy-Schwarz inequality for polyvectors according to Drozdov [1] is given. It leads to the definition of the angle of two vector subspaces with different dimensions.

Cauchy-Schwarz inequality, a well-known property of a scalar product in unitary vector spaces, is as follows:

If $V_1 = [\{\mathbf{u}\}]$ $V_1' = [\{\mathbf{v}\}]$ *are two one-dimensional subspaces of a real (or complex) vector space* V_n *with a scalar product, then*

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}|\ |\mathbf{v}|, \tag{1}$$

with equality if and only if $V_1 = V_1'$.

In this article we shall give a generalization of the above inequality (1). We extend this inequality on two subspaces V_r and V_s of a unitary vector space V_n, which are represented by their bases $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$, respectively.

First we will deal with the case when the subspaces V_r, V_s have the same dimension. This case was investigated by N. D. Drozdov [1] and we will give a brief overview of his results [1] (Th. 1). Thereafter we shall study the case when the subspaces V_r, V_s have different dimensions. These results are given in Th. 2. Since the Cauchy-Schwarz inequality holds for two vector subspaces of arbitrary dimensions, one can introduce the angle of these subspaces similarly as is done in the one-dimensional case. Th. 3 establishes that the angle of V_r, V_s is the same as the angle of their orthogonal complements. At the end of the article a comparison with the usual definition of the angle of two subspaces using eigenvalues of a characteristic matrix is given.

Preliminaries

Let the subspaces V_r and V_s of a real (or complex) unitary vector space V_n be represented by their bases $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$, $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s\}$, respectively. We shall call an ordered k-tuple of vectors k-vector or generally polyvector, an ordered k-tuple of linear independent vectors - complete k-vector or complete polyvector. Denote by $M = (m_{i,j})$, $i = 1, 2, \dots, s$, $j = 1, 2, \dots, r$ a real $r \times s$ matrix. Then the product $\mathbf{A}M$ of the r-vector $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$ and the matrix M is the s-vector

$$\mathbf{A}M = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\} \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1s} \\ m_{21} & m_{22} & \dots & m_{2s} \\ \dots & \dots & \dots & \dots \\ m_{r1} & m_{r2} & \dots & m_{rs} \end{pmatrix} =$$

$$= \Big\{ \sum_j m_{j1}\mathbf{a}_j, \sum_j m_{j2}\mathbf{a}_j, \dots, \sum_j m_{js}\mathbf{a}_j \Big\}.$$

The product of two polyvectors $\mathbf{A}=\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$, $\mathbf{B}=\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s\}$ is the $r \times s$ matrix $\mathbf{A} \cdot \mathbf{B}$ defined by

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1, & \mathbf{a}_1 \cdot \mathbf{b}_2, & \dots & \mathbf{a}_1 \cdot \mathbf{b}_s, \\ \mathbf{a}_2 \cdot \mathbf{b}_1, & \mathbf{a}_2 \cdot \mathbf{b}_2, & \dots & \mathbf{a}_2 \cdot \mathbf{b}_s, \\ \dots & \dots & \dots & \dots \\ \mathbf{a}_r \cdot \mathbf{b}_1, & \mathbf{a}_r \cdot \mathbf{b}_2, & \dots & \mathbf{a}_r \cdot \mathbf{b}_s, \end{pmatrix}$$

here $\mathbf{a}_i \cdot \mathbf{b}_j$ denotes the scalar product. It is easy to show that:

1) $\mathbf{A} \cdot \mathbf{B} = (\mathbf{B} \cdot \mathbf{A})^T$,

2) $\mathbf{A} \cdot (\mathbf{B}M) = (\mathbf{A} \cdot \mathbf{B})M$, $\mathbf{A}M \cdot \mathbf{B} = M^T(\mathbf{A} \cdot \mathbf{B})$.

Case r = s

A generalization of the Cauchy-Schwarz inequality for spaces with the same dimensions is as follows:

Theorem 1 (Drozdov [1]). *Let* $\mathbf{A}=\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$, $\mathbf{B}=\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_r\}$ *be complete r-vectors, which span subspaces V_r and V_r' of a unitary vector space V_n. Then*

$$\det{}^2(\mathbf{A}\cdot\mathbf{B}) \le \det(\mathbf{A}\cdot\mathbf{A})\det(\mathbf{B}\cdot\mathbf{B}), \tag{2}$$

with equality if and only if $V_r = V_r'$.

Remarks. 1) The inequality (2) enables us to define the angle of subspaces V_r and V_r' in the following way

$$\cos(V_r, V_r') = \frac{|\det(\mathbf{A}\cdot\mathbf{B})|}{\sqrt{\det(\mathbf{A}\cdot\mathbf{A})}\sqrt{\det(\mathbf{B}\cdot\mathbf{B})}}. \tag{3}$$

2) $V_r \perp V_r'$ if and only if $\det \mathbf{A}\cdot\mathbf{B} = 0$. See [1].

3) Denote by $\overline{V}_r$ and $\overline{V}_r'$ orthogonal complements of V_r and V_r' in V_n. Then [1]

$$\cos(\overline{V}_r, \overline{V}_r') = \cos(V_r, V_r').$$

Case r ≠ s

In this paragraph we shall assume that $r \le s$. The analogue of the Cauchy-Schwarz inequality to (1) and (2) for vector subspaces with different dimensions is described in the following:

Lemma 1. *Let the subspaces* V_r *and* V_s *are represented by complete polyvectors* $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$, $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s\}$. *Then*

$$V_r \perp V_s \iff \operatorname{rank}(\mathbf{A}\cdot\mathbf{B}) < r.$$

Proof. Let the subspaces V_r and V_s be orthogonal. Then there exist non zero vectors $\mathbf{a} \in V_r$ and $\mathbf{b} \in V_s$ such that $\mathbf{a} \perp V_s$, $\mathbf{b} \perp V_r$. Let $\mathbf{a} = \sum x_i \mathbf{a}_i$, $\mathbf{b} = \sum y_j \mathbf{b}_j$. Then

$$\begin{aligned} 0 &= \mathbf{a}\cdot\mathbf{b}_j = x_1\mathbf{a}_1\cdot\mathbf{b}_j + x_2\mathbf{a}_2\cdot\mathbf{b}_j + \dots + x_r\mathbf{a}_r\cdot\mathbf{b}_j \qquad j = 1, 2, \dots, s \\ 0 &= \mathbf{b}\cdot\mathbf{a}_i = y_1\mathbf{b}_1\cdot\mathbf{a}_i + y_2\mathbf{b}_2\cdot\mathbf{a}_i + \dots + y_s\mathbf{b}_s\cdot\mathbf{a}_i \qquad i = 1, 2, \dots, r \end{aligned}$$

holds. Each of the two systems of linear equation above has non zero solution, which gives the condition $\operatorname{rank}(\mathbf{A}\cdot\mathbf{B}) < r$. The converse implication is similar. ◇

Corollary. $V_r \perp V_s \iff \det(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{A}) = 0$.

Proof. Denote by $A\cdot B_{i_1,i_2,\dots,i_r}$ a matrix, which consists of the columns of the matrix $\mathbf{A}\cdot\mathbf{B}$ indexed by the numbers $i_1, i_2, \dots, i_r$. According to the Cauchy-Binet formula [2], [3] we get

$$\det(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{A}) = \sum_{1<i_1<i_2<\dots<i_r\le s} \det{}^2(A\cdot B_{i_1,i_2,\dots,i_r}).$$

Hence

$$\det(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{A}) = 0 \Longleftrightarrow \det(A\cdot B_{i_1,i_2,\dots,i_r}) = 0$$

for arbitrary $1 \leq i_1 < i_2 < \cdots < i_r \leq s$, which is equivalent to the fact that the rank of the matrix $\mathbf{A}\cdot\mathbf{B}$ is less than r. ◊

Theorem 2. *Let V_r and V_s be subspaces of V_n, which are represented by complete r-vector $\mathbf{A}$ and s-vector $\mathbf{B}$, respectively. Then*

$$0 \leq \det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})(\mathbf{A}\cdot\mathbf{A})^{-1}] \leq 1. \tag{4}$$

The equality on the left is attained if and only if the subspaces V_r and V_s are orthogonal, the equality on the right occurs if and only if $V_r \subseteq V_s$.

Proof. First we shall investigate the right inequality in (4). For this purpose denote $\mathbf{A}' = \mathbf{B}(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})$. Because $\mathbf{A}'$ is a complete r-vector then according to (2)

$$\det{}^2(\mathbf{A}\cdot\mathbf{A}') \leq \det(\mathbf{A}\cdot\mathbf{A})\det(\mathbf{A}'\cdot\mathbf{A}') \tag{5}$$

holds. This is equivalent to

$$\det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})(\mathbf{A}\cdot\mathbf{A})^{-1}] \leq 1. \tag{6}$$

The sign of equality in (6) occurs if and only if the subspaces $\mathbf{A}$ and $\mathbf{A}'$ are equal, i.e. iff there exist a regular $r\times r$ matrix C such that $\mathbf{A} = \mathbf{A}'C$. Denoting the $s\times r$ matrix $(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})C$ of rank r by M, we get $\mathbf{A} = \mathbf{B}M$, which means that $V_r \subseteq V_s$. As to the left inequality in (4), it suffices to prove

$$0 \leq \det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A}]). \tag{7}$$

We can write

$$(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A}) =$$

$$= [\mathbf{B}(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})]\cdot[\mathbf{B}(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})] = (\mathbf{B}M)^2,$$

where $M = (\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})$. From $\det(\mathbf{B}M)^2 \geq 0$ the inequality (7) follows. The sign of equality in (7) is attained iff the r-vector $\mathbf{B}M$ is not complete, which means that the rank of the matrix M is less than r. This occurs when the rank of the matrix $\mathbf{B}\cdot\mathbf{A}$ is less than r, which is equivalent — according to the previous Lemma — to the fact that the subspaces V_r, and V_s are orthogonal. ◊

Remarks. 1) As a special case of (4) for $r = s$ we get the inequality (2).

2) The inequality (4) enables us to define the angle of vector subspaces V_r and V_s with different dimensions in the following way:

$$\cos(V_r,\,V_s) = \sqrt{\det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})(\mathbf{A}\cdot\mathbf{A})^{-1}]}. \tag{8}$$

3) It is easy to prove that the angle of subspaces V_r and V_s does not depend on choosing the bases $\mathbf{A}$ and $\mathbf{B}$. Really, let $\mathbf{A}'$, $\mathbf{B}'$ be another bases such that $\mathbf{A}' = \mathbf{A}M$, $\mathbf{B}' = \mathbf{B}N$, where M, N are regular $r \times r$ and $s \times s$ matrices, respectively. Then

$$\begin{aligned}
&\det[(\mathbf{A}'\cdot\mathbf{B}')(\mathbf{B}'\cdot\mathbf{B}')^{-1}(\mathbf{B}'\cdot\mathbf{A}')(\mathbf{A}'\cdot\mathbf{A}')] = \\
&= \det[(\mathbf{A}M\cdot\mathbf{B}N)(\mathbf{B}N\cdot\mathbf{B}N)^{-1}(\mathbf{B}N\cdot\mathbf{A}M)(\mathbf{A}M\cdot\mathbf{A}M)^{-1}] = \\
&= \det[M^T(\mathbf{A}\cdot\mathbf{B})NN^{-1}(\mathbf{B}\cdot\mathbf{B})^{-1}(N^T)^{-1}N^T \\
&\qquad\qquad (\mathbf{B}\cdot\mathbf{A})MM^{-1}(\mathbf{A}\cdot\mathbf{A})^{-1}(M^T)^{-1}] = \\
&= \det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{B})^{-1}(\mathbf{B}\cdot\mathbf{A})(\mathbf{A}\cdot\mathbf{A})^{-1}].
\end{aligned}$$

4) If the polyvectors $\mathbf{A}$ and $\mathbf{B}$ form orthonormal bases in V_r and V_s respectively, then the inequality (4) has the form

$$0 \le \det[(\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{A})] \le 1.$$

Theorem 3. *The orthogonal complements $\overline{V}_{n-s}$ and $\overline{V}_{n-r}$ of V_r and V_s contain the same angle as the subspaces V_r and V_s, i.e.*

$$\cos(V_r, V_s) = \cos(\overline{V}_{n-s}\overline{V}_{n-r}).$$

Proof. Let V_r and V_s be represented by $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r\}$, $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_s\}$. Let $r < s$. Denote by $\overline{\mathbf{A}} = \{\mathbf{a}_{r+1}, \mathbf{a}_{r+2}, \ldots, \mathbf{a}_n\}$ and $\overline{\mathbf{B}} = \{\mathbf{b}_{s+1}, \mathbf{b}_{s+2}, \ldots, \mathbf{b}_n\}$ polyvectors representing the orthogonal complements $\overline{V}_{n-r}$ and $\overline{V}_{n-s}$. We may suppose that the bases $\mathbf{A}, \overline{\mathbf{A}}, \mathbf{B}, \overline{\mathbf{B}}$ are orthogonal. There exist matrices K, L, M, N, such that $\mathbf{B} = \mathbf{A}K + \overline{\mathbf{A}}L$ and $\overline{\mathbf{B}} = \mathbf{A}M + \overline{\mathbf{A}}N$, which can be written in the form

$$(\mathbf{B}, \overline{\mathbf{B}}) = (\mathbf{A}, \overline{\mathbf{A}}) \begin{pmatrix} K & M \\ L & N \end{pmatrix}.$$

Since the matrix $P = \left(\begin{smallmatrix} K & M \\ L & N \end{smallmatrix}\right)$ is orthogonal, then $P\cdot P^T = P^T\cdot P = I_n$, where I_n is the unit matrix. It implies

$$\begin{pmatrix} K & M \\ L & N \end{pmatrix}\begin{pmatrix} K^T & L^T \\ M^T & N^T \end{pmatrix} = \begin{pmatrix} KK^T + MM^T & KL^T + MN^T \\ LK^T + NM^T & LL^T + NN^T \end{pmatrix} = I_n$$

$$\begin{pmatrix} K^T & L^T \\ M^T & N^T \end{pmatrix}\begin{pmatrix} K & M \\ L & N \end{pmatrix} = \begin{pmatrix} K^TK + L^TL & K^TM + L^TN \\ M^TK + N^TL & M^TM + N^TN \end{pmatrix} = I_n$$

and we get the relations

$$KK^T + MM^T = I, \qquad KL^T + MN^T = 0,$$
$$LK^T + NM^T = 0, \qquad LL^T + NN^T = I_{n-r},$$
$$K^TK + L^TL = I_s, \qquad K^TM + L^TN = 0,$$
$$M^TK + N^TL = 0, \qquad M^TM + N^TN = I_{n-s}.$$

It suffices to prove that

$$\det[(\mathbf{A} \cdot \mathbf{B})(\mathbf{B} \cdot \mathbf{A})] = \det[(\overline{\mathbf{B}} \cdot \overline{\mathbf{A}})(\overline{\mathbf{A}} \cdot \overline{\mathbf{B}})],$$

which is equivalent to

$$\det(KK^T) = \det(N^TN)$$

and according to the above relations to

$$\det(I_r - MM^T) = \det(I_{n-s} - M^TM). \tag{9}$$

Let $r \leq n-s$ then the matrix M^TM has the same eigenvalues as the matrix MM^T and another $n-s-r$ eigenvalues, which equal zero [3]. Denote by $\lambda_1, \lambda_2, \dots, \lambda_r$ eigenvalues of the matrix MM^T. Then the eigenvalues of the $(n-s) \times (n-s)$ matrix M^TM are the numbers $\lambda_1, \lambda_2, \dots, \lambda_r, 0, \dots, 0$ with $n-s-r$ zeros and we have

$$\det(I_r - MM^T) = (1-\lambda_1)(1-\lambda_2)\dots(1-\lambda_r) = \det(I_{n-s} - M^TM). \qquad \Diamond$$

A comparison with the definition of the angle of subspaces using characteristic numbers

Assume that the polyvectors $\mathbf{A}$, $\mathbf{B}$ representing vector subspaces V_r and V_s form orthonormal bases. According to [4], the vector subspaces V_r, V_s, $r < s$ contain r angles $\varphi_1, \varphi_2, \dots, \varphi_r$ which satisfy

$$\cos\varphi_i = \sqrt{\lambda_i}, \qquad i = 1, 2, \dots, r, \tag{10}$$

where λ_i are characteristic numbers of the matrix $(\mathbf{A} \cdot \mathbf{B})(\mathbf{B} \cdot \mathbf{A})$. Hence λ_i are roots of the equation

$$\det[(\mathbf{A} \cdot \mathbf{B})(\mathbf{B} \cdot \mathbf{A}) - \lambda I] = 0. \tag{11}$$

From (11) we get

$$\lambda_1\lambda_2 \dots \lambda_r = \det[(\mathbf{A} \cdot \mathbf{B})(\mathbf{B} \cdot \mathbf{A})]$$

and in view of (8) and (10)

$$\cos\varphi = \cos\varphi_1 \cos\varphi_2 \dots \cos\varphi_r.$$

Finally, we get

$$\varphi \geq \operatorname{Max}\{\varphi_1, \varphi_2, \ldots, \varphi_r\}$$

Acknowledgement. I thank Professor Z. Nádenik for drawing my attention to the problem.

References

[1] DROZDOV, N. D.: Linejnaja algebra v teorii uravnivanija izmerenij, Nedra, Moskva, 1973.

[2] GANTMACHER, F. R.: Teorija matric, Nauka, Moskva 1988.

[3] HORN, R. A., JOHNSON, C. R.: Matrix analysis, Cambridge Univ. Press, Cambridge, 1986.

[4] VITNER, Č.: 0 úhlech lineárnich podprostorů v E^N, Čas. pro pěst. mat. **87** (1962), 415–422.