Orthogonality of Matrices

Chi-Kwong Li[1]
Department of Mathematics, College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187-8795. (ckli@math.wm.edu)

Hans Schneider
Department of Mathematics, University of Wisconsin, Madison, WI 53706. (hans@math.wisc.edu)

Abstract

Let A and B be rectangular matrices. Then A is orthogonal to B if

$$\|A + \mu B\| \geq \|A\| \qquad \text{for every scalar } \mu.$$

Some approximation theory and convexity results on matrices are used to study orthogonality of matrices and answer an open problem of Bhatia and Šemrl.

---

1 Introduction

Let $(\mathbf{F}^{m\times n}, \|\cdot\|)$ be a normed matrix space over $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$. Suppose $A, B \in \mathbf{F}^{m\times n}$, we say that A is orthogonal to B (in the Birkhoff-James sense [4]) if

$$\|A + \mu B\| \geq \|A\| \qquad \text{for every } \mu \in \mathbf{F}.$$

The above condition can be interpreted in the context of approximation theory as follows. Suppose $A \in \mathbf{F}^{m\times n}$ is not in the linear subspace $\mathcal{W}$ spanned by the matrix $B \in \mathbf{F}^{m\times n}$. Then the zero matrix is the best approximation to A among all matrices in $\mathcal{W}$. In this note, we use some approximation theory and convexity results in matrix spaces to study orthogonality of matrices. Our results cover and extend those of other authors [1, 5].

We collect some preliminary results in Section 2, and use them to characterize matrix pairs which are orthogonal with respect to Schatten p-norms in Section 3. In the last section, we study orthogonal matrix pairs with respect to operator norms and give a counter-example to a conjecture of Bhatia and Šemrl [1].

We always assume that $\mathbf{F}^{m\times n}$ is equipped with the inner product $(A, B) = \mathrm{tr}\,(AB^*)$. This includes the special case when $\mathbf{F}^{n\times 1} = \mathbf{F}^n$ and $(x, y) = \mathrm{tr}\,(xy^*) = y^*x$. Denote by $\{e_1, \dots, e_n\}$ the standard basis for $\mathbf{F}^n$, and $\{E_{11}, E_{12}, \dots, E_{mn}\}$ the standard basis for $\mathbf{F}^{m\times n}$. Let $U_n(\mathbf{F})$ be the unitary or orthogonal group depending on $\mathbf{F} = \mathbf{C}$ or $\mathbf{R}$.

For notational convenience, we always consider $m\times n$ matrix with $m \leq n$ in our discussion; the case $m > n$ can be treated similarly. For $A \in \mathbf{F}^{m\times n}$, denote by $s_1(A) \geq \cdots \geq s_m(A)$ the singular values of A, which are the nonnegative square roots of the eigenvalues of the matrix AA^*. We always use the fact that every matrix $A \in \mathbf{F}^{m\times n}$ has a singular value decomposition, viz., $A = U^*(\sum_{j=1}^m s_j(A)E_{jj})V$ for some $U \in U_m(\mathbf{F})$ and $V \in U_n(\mathbf{F})$.

[1]Research partially supported by NSF.

2 Preliminary Results

Let $\|\cdot\|$ be a norm on $\mathbf{F}^{m\times n}$. The dual norm of $\|\cdot\|$ is defined by

$$\|X\|^D = \max\{|(X,Y)| : \|Y\| \le 1\}.$$

We have the following result, which is a special case of the general theorem of Singer in [9, p.170].

Proposition 2.1 *Let $\|\cdot\|$ be a norm on $\mathbf{F}^{m\times n}$. Suppose $A, B \in \mathbf{F}^{m\times n}$ are such that A is not a multiple of B. Then*

$$\|A + \mu B\| \ge \|A\| \qquad \textit{for all } \mu \in \mathbf{F}$$

if and only if there exist h extreme points $F_1, \ldots, F_h \in \mathbf{F}^{m\times n}$ of the unit ball $\{Y \in \mathbf{F}^{m\times n} : \|Y\|^D \le 1\}$ in the dual space $(\mathbf{F}^{m\times n}, \|\cdot\|^D)$ with $h \le 3$ in the complex case and $h \le 2$ in the real case, and positive numbers $t_1, \ldots, t_h$ with $t_1 + \cdots + t_h = 1$ such that

$$\sum_{j=1}^{h} t_j(F_j, B) = 0 \qquad \textit{and} \qquad (F_j, A) = \|A\|, \quad j = 1, \ldots, h. \tag{1}$$

The numerical range of a matrix $A \in \mathbf{F}^{n\times n}$ is defined by

$$W(A) = \{x^*Ax : x \in \mathbf{F}^n,\ x^*x = 1\},$$

which has been studied extensively, see [3, Chapter 1]. We have the following result.

Proposition 2.2 *Let $A \in \mathbf{F}^{n\times n}$. Then $W(A)$ is convex.*

Proof. For the complex case, see [3, Chapter 1]. For the real case, note that $W(A)$ can be viewed as the image of the unit sphere in $\mathbf{R}^n$ under the continuous map $x \mapsto x^*Ax$. Since the unit sphere in $\mathbf{R}^n$ is a compact connected set, the set $W(A)$ is a closed interval. □

3 The Schatten p-Norms

Suppose $1 \le p \le \infty$. The Schatten p-norm of $A \in \mathbf{F}^{m\times n}$ is defined by

$$S_p(A) = \begin{cases} \left\{\sum_{j=1}^n s_j(A)^p\right\}^{1/p} & \text{if } 1 \le p < \infty, \\ \max\{s_j(A) : 1 \le j \le m\} & \text{if } p = \infty\ . \end{cases}$$

We refer the readers to [7] for basic properties of the Schatten p-norms. Here we characterize $A \in \mathbf{F}^{m\times n}$ which are orthogonal to a given matrix $B \in \mathbf{F}^{m\times n}$ with respect to the Schatten p-norms. We shall use the basic fact that the dual space of $(\mathbf{F}^{m\times n}, S_p)$ is $(\mathbf{F}^{m\times n}, S_q)$, where

$1/p + 1/q = 1$. Moreover, $F \in \mathbf{F}^{m\times n}$ is an extreme point of the unit norm ball of $(\mathbf{F}^{m\times n}, S_p)$ if and only if

(i) $p = 1$ and $F = xy^*$ for some unit vectors $x \in \mathbf{F}^m$ and $y \in \mathbf{F}^n$;

(ii) $1 < p < \infty$ and $S_p(F) = 1$;

(iii) $p = \infty$ and $FF^* = I_m$.

In [1] (see also [5]) the authors obtained results for complex square matrices with $p > 1$ and partial results for $p = 1$ by different methods.

Theorem 3.1 *Let $A, B \in \mathbf{F}^{m\times n}$, where $m \le n$. The following conditions are equivalent.*

(a) $S_\infty(A + \mu B) \ge S_\infty(A)$ *for all $\mu \in \mathbf{F}$.*

(b) *There exist unit vectors $x \in \mathbf{F}^m$ and $y \in \mathbf{F}^n$ such that $S_\infty(A) = x^*Ay$ and $x^*By = 0$, equivalently, there is a unit vector $y \in \mathbf{F}^n$ such that $S_\infty(A) = l_2(Ay)$ and $(Ay, By) = 0$.*

(c) *For any $U \in \mathbf{F}^{m\times k}$ with orthonormal columns that form a basis for the eigenspace of AA^* corresponding to the largest eigenvalue, and $V = A^*U/S_\infty(A) \in \mathbf{F}^{n\times k}$, we have*

$$0 \in W(U^*BV) \qquad or \qquad 0 \in W(U^*BA^*U).$$

Proof. To prove the theorem, we use Proposition 2.1 with $\|\cdot\| = S_\infty$, and the fact that $(\mathbf{F}^{m\times n}, S_1)$ is the dual space of $(\mathbf{F}^{m\times n}, S_\infty)$.

Suppose (a) holds. By Proposition 2.1, there exist extreme points $F_j = x_j y_j^* \in \mathbf{F}^{m\times n}$ of the unit ball of $(\mathbf{F}^{m\times n}, S_1)$ with $1 \le j \le h$, and some positive constants $t_1, \dots, t_h$ with $t_1 + \cdots + t_h = 1$ so that $(F_j, A) = S_\infty(A)$ for $j = 1, \dots, h$, and $(\sum_{j=1}^h t_j F_j, B) = 0$. By our assumption on U, for each $j = 1, \dots, h$, there is a unit vector $v_j \in \mathbf{F}^k$ so that $x_j = Uv_j$, and $y_j = Vv_j$. Thus,

$$0 = (\sum_{j=1}^h t_j F_j, B) = \sum_{j=1}^h t_j v_j^*(U^*BV)v_j,$$

which is an element in the convex hull of $W(U^*BV)$, equivalently, $0 \in W(U^*BV)$ by Proposition 2.2. By the fact that $A^*U = S_\infty(A)V$, we see $0 \in W(U^*BV)$ if and only if $0 \in W(U^*BA^*U)$. Hence, condition (c) holds.

If (c) holds, and $v \in \mathbf{F}^k$ is a unit vector such that $0 = v^*U^*BA^*Uv$, then $x = Uv$ and $y = A^*x/S_\infty(A)$ are the unit vectors satisfying (b).

If (b) holds, then $F_1 = xy^* \in \mathbf{F}^{m\times n}$ is an extreme point of the unit ball of $(\mathbf{F}^{m\times n}, S_1)$; see (i). So condition (1) holds with $h = 1$. By Proposition 2.1, condition (a) holds. □

Note that condition (b) in the above theorem looks simpler than (c) as it does not depend on the construction of a basis for the eigenspace of AA^*. Nonetheless, in practice, it is easier to check condition (c) by studying $W(U^*BA^*U)$.

Theorem 3.2 *Let $1 < p < \infty$, $m \le n$. Suppose $A, B \in \mathbf{F}^{m\times n}$, where $A = HX$ for some positive semi-definite $H \in \mathbf{F}^{m\times m}$ and $X \in \mathbf{F}^{m\times n}$ with $XX^* = I_m$. Then*

$$S_p(A+\mu B) \ge S_p(A) \qquad \text{for all } \mu \in \mathbf{F}$$

if and only if for any $U \in U_m(\mathbf{F})$ and $V \in U_n(\mathbf{F})$ satisfying $UAV^* = \sum_{j=1}^m s_j(A)E_{jj}$ we have $\mathrm{tr}\,\left[U^*(\sum_{j=1}^m s_j(A)^{p-1}E_{jj})VB^*\right] = 0$, equivalently, $\mathrm{tr}\,(H^{p-1}XB^*) = 0$.

Proof. The theorem readily follows from Proposition 2.1 and the fact that if $U \in U_m(\mathbf{F})$ and $V \in V_n(\mathbf{F})$ are such that $A = U^*(\sum_{j=1}^m a_jE_{jj})V$ with $a_1 \ge \cdots \ge a_m \ge 0$, then $F = \gamma^{-1}U^*(\sum_{j=1}^m a_j^{p-1}E_{jj})V$ with

$$\gamma = \left\{\sum_{j=1}^m a_j^{(p-1)q}\right\}^{1/q} = \left\{\sum_{j=1}^m a_j^{p(1-1/p)q}\right\}^{1/q} = \left\{\sum_{j=1}^m a_j^p\right\}^{1/q}$$

is the unique extreme point of the dual norm ball in the dual space of $(\mathbf{F}^{m\times n}, S_p)$ satisfying $(F, A) = S_p(A)$. □

Theorem 3.3 *Let $A, B \in \mathbf{F}^{m\times n}$, where $m \le n$. The following conditions are equivalent.*

(a) $S_1(A+\mu B) \ge S_1(A)$ *for all $\mu \in \mathbf{F}$.*

(b) *There exists $F \in \mathbf{F}^{m\times n}$ such that $S_\infty(F) \le 1$, $\mathrm{tr}\,(AF^*) = S_1(A)$ and $\mathrm{tr}\,(BF^*) = 0$.*

(c) *For any $U \in U_m(\mathbf{F})$ and $V \in U_n(\mathbf{F})$ satisfying $UAV^* = \sum_{j=1}^m s_j(A)E_{jj}$, we have*

$$UBV^* = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22}\end{pmatrix} \qquad \text{and} \qquad |\mathrm{tr}\,(B_{11})| \le S_1(B_{22}),$$

where B_{11} is $k \times k$ with $k = \mathrm{rank}\,(A)$, and by convention $S_1(B_{22}) = 0$ if $m = k$.

Proof. Suppose (a) holds. By Proposition 2.1 and (iii), there exist extreme points $F_1, \ldots, F_h \in \mathbf{F}^{m\times n}$ satisfying $F_jF_j^* = I_m$, where $1 \le h \le 3$, and positive constants $t_1, \ldots, t_h$ with $t_1 + \cdots + t_h = 1$ such that $S_1(A) = (F_j, A)$ for $j = 1, \ldots, h$, and $(\sum_{j=1}^h t_jF_j, B) = 0$. Then condition (b) holds with $F = \sum_{j=1}^h t_jF_j$.

Suppose (b) holds. Let $U \in U_m(\mathbf{F})$ and $V \in U_n(\mathbf{F})$ satisfy $A = U^*(\sum_{j=1}^m s_j(A)E_{jj})V$. Furthermore, assume that $\mathrm{rank}\,(A) = k$. Since $\mathrm{tr}\,(AF^*) = S_1(A)$, we see that

$$UFV^* = \begin{pmatrix} I_k & 0 \\ 0 & G\end{pmatrix},$$

where $S_\infty(G) \le 1$. Thus,

$$0 = \mathrm{tr}\,(BF^*) = \mathrm{tr}\,\left[\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22}\end{pmatrix}\begin{pmatrix} I_k & 0 \\ 0 & G\end{pmatrix}^*\right]$$

implies that

$$|\mathrm{tr}\,(B_{11})| = |\mathrm{tr}\,(B_{22}G^*)| \le S_1(B_{22}),$$

i.e., condition (c) holds.

Finally, if (c) holds with $k = m$, then $F = U^*(\sum_{j=1}^m E_{jj})V$ is an extreme point of the unit ball of $(\mathbf{F}^{m\times n}, S_\infty)$ satisfying $(F, A) = S_1(A)$ and $(F_1, B) = 0$. By Proposition 2.1, condition (a) holds. If (c) holds with $k < m$, then by a result of Thompson [10] (see also [6]), there exists $G \in \mathbf{F}^{(m-k)\times(n-k)}$ such that $S_\infty(G) \le 1$ and $0 = \mathrm{tr}\,(B_{11}) + \mathrm{tr}\,(B_{22}G^*)$. Suppose $G = HX$ for some positive semi-definite H and $X \in \mathbf{F}^{(m-k)\times(n-k)}$ satisfying $XX^* = I_{m-k}$. Let $G_1 = (H + i\sqrt{I - H^2})X$, $G_2 = (H - i\sqrt{I - H^2})X$, and $F_1, F_2 \in \mathbf{F}^{m\times n}$ be such that

$$UF_jV^* = \begin{pmatrix} I_k & 0 \\ 0 & G_j \end{pmatrix} \qquad \text{for } j = 1, 2.$$

Then $F_jF_j^* = I_m$, $(F_j, A) = S_1(A)$ and $((F_1 + F_2)/2, B) = 0$. By Proposition 2.1, condition (a) holds. □

4 Operator norms, and a problem of Bhatia and Šemrl

Let ν be a norm on $\mathbf{F}^n$, and let $\|\cdot\|_\nu$ be the operator norm on $\mathbf{F}^{n\times n}$ induced by ν, i.e.,

$$\|A\|_\nu = \max\{\nu(Ax) : x \in \mathbf{F}^n,\ \nu(x) \le 1\}.$$

The dual norm of ν and $\|\cdot\|_\nu$ are defined as

$$\nu^D(x) = \max\{|(x, y)| : y \in \mathbf{F}^n,\ \nu(y) \le 1\},$$

and

$$\|A\|_\nu^D = \max\{|(A, B)| : B \in \mathbf{F}^{n\times n},\ \|B\|_\nu \le 1\},$$

respectively. We have the following result.

Proposition 4.1 *Let ν be a norm on $\mathbf{F}^n$. Denote by $\mathcal{E}$ and $\mathcal{E}^D$ the set of extreme points of the unit norm balls of ν and ν^D, respectively. Then A is an extreme point of the unit ball of $\|\cdot\|_\nu^D$ if and only if $A = xy^*$ such that $x \in \mathcal{E}^D$ and $y \in \mathcal{E}$.*

Proof. Let $\mathcal{E}_{\|\cdot\|_\nu^D}$ be the set of extreme points of the unit ball $\mathcal{B}_{\|\cdot\|_\nu^D}$ of $\|\cdot\|_\nu^D$. Since

$$\|A\|_\nu = \max\{|(A, X)| : X \in \mathbf{F}^{n\times n},\ \|X\|^D \le 1\},$$

and

$$\begin{aligned}\|A\|_\nu &= \max\{|x^*Ay| : \nu^D(x) = \nu(y) = 1\}\\ &= \max\{|\mathrm{tr}\,(Ayx^*)| : \nu^D(x) = \nu(y) = 1\}\\ &= \max\{|(A, xy^*)| : \nu^D(x) = \nu(y) = 1\},\end{aligned}$$

we see that

$$\mathcal{E}_{\|\cdot\|_\nu^D} \subseteq \{xy^* : \nu^D(x) = \nu(y) = 1\}.$$

If $y = (y_1 + y_2)/2$, then $xy^* = (xy_1^* + xy_2^*)/2$; if $x = (x_1 + x_2)/2$, then $xy^* = (x_1y^* + x_2y^*)/2$. Thus,

$$\mathcal{E}_{\|\cdot\|_\nu^D} \subseteq \{xy^* : x \in \mathcal{E}^D,\ y \in \mathcal{E}\}.$$

Suppose xy^* with $x \in \mathcal{E}^D$ and $y \in \mathcal{E}$. If xy^* is not an extreme point of $\mathcal{B}_{\|\cdot\|_\nu^D}$, then it is a convex combination of other matrices in $\{uv^* : u \in \mathcal{E}^D,\ v \in \mathcal{E}\}$, say,

$$xy^* = \sum_{j=1}^{m} t_j x_j y_j^*, \quad t_1, \ldots, t_m > 0,\ t_1 + \cdots + t_m = 1.$$

Let $u \in \mathcal{E}^D$ be such that $y^*u = 1$, and let $y_j^*u = \mu_j$ for $j = 1, \ldots, n$. Then $|\mu_j| \le 1$ and

$$x = xy^*u = \sum_{j=1}^{t} t_j x_j y_j^* u = \sum_{j=1}^{t} t_j \mu_j x_j.$$

Since $x \in \mathcal{E}^D$, it follows that $\mu_j x_j = x$ with $|\mu_j| = 1$ for all $j = 1, \ldots, m$. By a similar argument, we see that $\eta_j y_j = y$ for some η_j with $|\eta_j| = 1$ for all $j = 1, \ldots, m$. Hence

$$xy^* = xy^*(\sum_{j=1}^{m} t_j \mu_j \eta_j).$$

Thus, $\mu_j \eta_j = 1$ and $x_j y_j^* = xy^*$, which is a contradiction. So, $\mathcal{E}_{\|\cdot\|_\nu^D} = \{xy^* : x \in \mathcal{E}^D,\ y \in \mathcal{E}\}$ as asserted. □

Using the above result and Proposition 2.1, one readily deduces the following.

Proposition 4.2 *Suppose $\|\cdot\|_\nu$ is an operator norm on $\mathbf{F}^{n\times n}$ induced by the vector norm ν on $\mathbf{F}^n$. Given $A \in \mathbf{F}^{n\times n}$, let*

$$V(A) = \{xy^* : x \in \mathcal{E}^D,\ y \in \mathcal{E},\ (A, xy^*) = \|A\|_\nu\}.$$

Then $B \in \mathbf{F}^{n\times n}$ satisfies

$$\|A + \mu B\|_\nu \ge \|A\|_\nu \qquad \text{for all } \mu \in \mathbf{F}$$

if and only if there exist there exist h extreme points $x_1y_1^*, \ldots, x_hy_h^* \in V(A)$ with $h \le 3$ in the complex case and $h \le 2$ in the real case, and positive numbers $t_1, \ldots, t_h$ with $t_1 + \cdots + t_h = 1$ such that

$$\sum_{j=1}^{h} t_j (B, x_j y_j^*) = 0. \tag{2}$$

Suppose ν is a norm on $\mathbf{F}^n$ and $\|\cdot\|_\nu$ is the corresponding operator norm on $\mathbf{F}^{n\times n}$. Consider the following conditions for $A, B \in \mathbf{F}^{n\times n}$.

(I) $\|A+\mu B\|_\nu \geq \|A\|_\nu$ for all $\mu \in \mathbf{F}$.

(II) There exists a vector $y \in \mathbf{F}^n$ with $\nu(y)=1$ such that $\nu(Ay)=\|A\|_\nu$ and

$$\nu(Ay+\mu By) \geq \nu(Ay) \quad \text{for all } \mu \in \mathbf{F}.$$

In general, we have (II) implies (I). In [1], the authors conjectured that (I) also implies (II). We use Proposition 4.2 to show that this is not true in general.

Example 4.3 Let $\|\cdot\|$ be the operator norm on $\mathbf{F}^{n\times n}$ induced by the l_p norm with $p \neq 2$. Consider $A = A_1 \oplus 0_{n-2}$ and $B = I_2 \oplus 0_{n-2}$, where

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Define $V(A)$ as in the proof of Proposition 4.2. Clearly, if $Z \in V(A)$, then $Z = Z_1 \oplus 0_{n-2}$ for some $Z_1 = (a_1 a_2)^t(\bar{b}_1\bar{b}_2)$ such that

$$(A_1, Z_1) \geq |(A_1, uv^*)|$$

for any $u, v \in \mathbf{F}^2$ with $l_q(u) = l_p(v) = 1$, where $1/p+1/q=1$. It follows from [8, Proposition 2] that $Z_1 \in V(A_1)$ with

$$V(A_1) = \begin{cases} \{\gamma(1,1)^t(1,0), \gamma(1,-1)^t(0,1)\} & \text{if } p<2 \\ \{\gamma(1,0)^t(1,1), \gamma(0,1)^t(1,-1)\} & \text{if } p>2 \end{cases},$$

where $\gamma = 1/l_r((1,1)^t)$ with $r = \max\{p,q\}$. Let $V(A) = \{U_1 \oplus 0_{n-2}, U_2 \oplus 0_{n-2}\}$, where $V(A_1) = \{U_1, U_2\}$, and let $x_j, y_j \in \mathbf{F}^n$ satisfy $l_p(y_j) = l_q(x_j) = 1$ and $x_j y_j^* = U_j \oplus 0_{n-2}$ for $j = 1,2$. Then (2) holds with $t_1 = t_2 = 1/2$, and hence condition (I) follows.

Now, if $y \in \mathbf{F}^n$ satisfies $l_p(Ay) = \|A\|$, then $y = (b_1, b_2, 0, \ldots, 0)^t \in \mathbf{F}^n$ and

$$l_p(Ay) = l_p(A_1(b_1,b_2)^t).$$

By [8, Proposition 2], we see that

(i) $p<2$ and (b_1,b_2) is a multiple of $(1,0)$ or $(0,1)$, or
(ii) $p>2$ and (b_1,b_2) is a multiple of $(1,1)$ or $(1,-1)$.

In any case, it is impossible to have

$$l_p((A_1+\mu B_1)(b_1,b_2)^t) \geq l_p(A_1(b_1,b_2)^t) \qquad \text{for all } \mu \in \mathbf{F},$$

and thus (II) cannot hold.

References

[1] R. Bhatia and P. Šemrl, Orthogonality of matrices and some distance problems, Linear Algebra Appl. 287 (1999), 77-85.

[2] P.R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, Princeton, N.J., 1967.

[3] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.

[4] R.C. James, Orthogonality and linear functionals in normed linear spaces, Trans. Amer. Math. Soc. 61 (1947), 265-292.

[5] F. Kittaneh, On zero-trace matrices, Linear Algebra Appl. 151 (1991), 119-124.

[6] J. von Neumann, Some matrix inequalities and matrization of a matrix space, Tomsk Univ. Rev. 1 (1937), 286-300.

[7] R. Schatten, Norm ideals of completely continuous operators, Springer-Verlag, Berlin, 1960.

[8] H. Schneider and H.F. Weinberger, On matrices for which norm bounds are attained, Linear Algebra Appl. 275-276 (1998), 563-577.

[9] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Springer, New York, 1970.

[10] R.C. Thompson, The bilinear field of values, Monatsh. Math. 81 (1976), 153-167.