CORE

Metadata, citation and similar papers at core.ac.uk

Provided by Elsevier - Publisher Connector

NORTH-HOLLAND

On Approximation Problems With Zero-Trace Matrices

Krystyna Ziętak*

Institute of Computer Science
University of Wrocław
ul. Przesmyckiego 20
51151 Wrocław, Poland

ABSTRACT

We consider some approximation problems in the linear space of complex matrices with respect to unitarily invariant norms. We deal with special cases of approximation of a matrix by zero-trace matrices. Moreover, some characterizations of zero-trace matrices are given by means of matrix approximation problems.

1. INTRODUCTION

Let $A = [a_{ij}] \in \mathscr{C}^{n\times n}$ be a complex matrix. The trace of A is equal to

$$\operatorname{tr}(A) = \sum_j a_{jj}.$$

It is well known that $\operatorname{tr}(A) = 0$ if and only if A is a commutator, that is, $A = XY - YX$ for some matrices X and Y. In this paper we consider some approximation problems, involving zero-trace matrices, with respect to an arbitrary unitarily invariant norm $\|\cdot\|$. A norm $\|\cdot\|$ is unitarily invariant if $\|UA\| = \|AV\| = \|A\|$ for all unitary matrices U and V. The most popular unitarily invariant norms are the c_p-norms $\|\cdot\|_p$. Let $\sigma_j(A)$ denote the jth singular value of A. We assume that singular values are ordered decreasingly:

$$\sigma_1(A) \geqslant \cdots \geqslant \sigma_n(A) \geqslant 0.$$

* E-mail: zietak@ii.uni.wroc.pl.

LINEAR ALGEBRA AND ITS APPLICATIONS 247:169–183 (1996)

© Elsevier Science Inc., 1996
655 Avenue of the Americas, New York, NY 10010

0024-3795/96/$15.00
SSDI 0024-3795(95)00098-C

The c_p-norm $\|A\|_p$ is equal to the l_p-norm of the vector $\sigma(A) = [\sigma_1(A), \dots, \sigma_n(A)]^T$. For $p = \infty$ we have the spectral norm, for $p = 1$ the trace norm.

We now formulate two approximation problems involving zero-trace matrices.

PROBLEM I. Let $\mathscr{Z}$ denote the set of all zero-trace matrices of order n. Show that for every unitarily invariant norm

$$\mu(A) \equiv \min_{X \in \mathscr{Z}} \|A - X\| = \frac{|\mathrm{tr}(A)|}{\|I\|^*}, \tag{1}$$

where $\|\cdot\|^*$ denotes the dual norm to the norm $\|\cdot\|$, and describe all $X \in \mathscr{Z}$ satisfying (1).

PROBLEM II. Prove that if $\mathrm{tr}(B) = 0$, then for every unitarily invariant norm

$$\gamma(B) \equiv \min_{z \in \mathscr{C}} \|I + zB\| = \|I\|, \tag{2}$$

and investigate when (2) implies $\mathrm{tr}(B) = 0$.

In the problem (2) we approximate I by elements from the linear subspace span$\{B\}$. For this particular case the approximation error is equal to the norm of the matrix which is approximated, i.e., the zero matrix is an approximation. We will show that if $A = I$ then the zero matrix is also a solution of the problem (1).

The problems (1) and (2) were considered by Kittaneh [4, 5] for the c_p-norms. Namely, he has proven that for the c_p-norm, $\mu(A) = |\mathrm{tr}(A)|/n^{1/q}$ whenever $1/p + 1/q = 1$. The result of Kittaneh for the problem (2) can be stated as follows. Let B be a complex matrix of order n and let $1 \leqslant p < \infty$. Then B has zero trace if and only if $\|I + zB\|_p \geqslant n^{1/p}$ for all complex numbers z. Kittaneh has given an example that this is not true for the spectral norm $\|\cdot\|_\infty$. We explain this failure below.

The purpose of this paper is a generalization of the results known for Problems I and II to the case of arbitrary unitarily invariant norms. Moreover, we describe all solutions of the problem (1). Some conditions concerning uniqueness of the solutions are also given. As a consequence of the results, presented in the paper, a new characterization of $\mathscr{Z}$ is given by means of a matrix approximation problem.

2. CHARACTERIZATION OF APPROXIMATIONS OF MATRICES

Let us consider $\mathscr{C}^{n\times n}$ as a normed linear space over the complex field $\mathscr{C}$ endowed with an arbitrary unitarily invariant norm $\|\cdot\|$. This linear space has dimension n^2. However, we will use some standard results from convex analysis of real linear spaces to investigate the properties of the solutions of Problems I and II. In this situation it is better to interpret $\mathscr{C}^{n\times n}$ as a real linear normed space in the following way. We write $A \in \mathscr{C}^{n\times n}$ in the form $A = A_1 + iA_2$ where A_1 and A_2 are real, $i = \sqrt{-1}$. Then $\mathscr{C}^{n\times n}$ can be identified with a real linear space of real matrices with the block form $[A_1, A_2]$. Hence the dimension of $\mathscr{C}^{n\times n}$ over the real field $\mathscr{R}$ is equal to $2n^2$. In the linear space $\mathscr{C}^{n\times n}$ over $\mathscr{C}$, the usual inner product is defined by

$$\operatorname{tr}(B^H A);$$

but in $\mathscr{C}^{n\times n}$ over $\mathscr{R}$ we take the following inner product:

$$\operatorname{Re}\operatorname{tr}(B^H A) = \operatorname{tr}\big([B_1, B_2]^T[A_1, A_2]\big) = \operatorname{tr}\big(B_1^T A_1 + B_2^T A_2\big).$$

Let $\mathscr{M}$ be a linear subspace of complex matrices over $\mathscr{C}$. We approximate a complex matrix $A = A_1 + iA_2$ by matrices from $\mathscr{M}$ with respect to an arbitrary unitarily invariant norm $\|\cdot\|$:

$$\delta(A) = \min_{X\in\mathscr{M}} \|A - X\|. \tag{3}$$

Let $\mathscr{M}$ over $\mathscr{C}$ have the dimension m, and let $\mathscr{M}$ be spanned by matrices $A_j = A_{1j} + iA_{2j}$, where A_{1j} and A_{2j} are real, $j = 1,\dots,m$. Then the problem (3) can be written as

$$\begin{aligned}\delta(A) &= \min_{z_j\in\mathscr{C}} \Big\|A - \sum_j z_j A_j\Big\| \\ &= \min_{x_j,\, y_j\in\mathscr{R}} \Big\|A_1 + iA_2 - \sum_j x_j(A_{1j} + iA_{2j}) - \sum_j y_j(-A_{2j} + iA_{1j})\Big\|.\end{aligned}$$

This can be interpreted as an approximation of the real matrix $[A_1, A_2]$ by matrices from the real linear subspace

$$\mathscr{M}_{\mathscr{R}} = \operatorname{span}\{[A_{11}, A_{21}], [-A_{21}, A_{11}], \dots, [A_{1m}, A_{2m}], [-A_{2m}, A_{1m}]\}$$

with respect to the norm which for $[A_1, A_2]$ is equal to the unitary invariant norm of $A = A_1 + iA_2$. Therefore, in order to obtain a characterization of solutions of (3), we apply the standard results from convex analysis to $[A_1, A_2]$ and the subspace $\mathcal{M}_{\mathcal{R}}$. For this purpose we need the orthogonal complement $\mathcal{M}_{\mathcal{R}}^{\perp}$ and the subdifferential $\partial\|X\|$. The definition of the subdifferential is given in Rockafellar [8]. Some properties of the subdifferentials of unitarily invariant norms will be recalled later.

The orthogonal complement of $\mathcal{M}_{\mathcal{R}}$ is the set of all real matrices $[F_1, F_2]$ such that for every j we have

$$[F_1, F_2] \perp [A_{1j}, A_{2j}], \qquad [F_1, F_2] \perp [-A_{2j}, A_{1j}].$$

These conditions are equivalent to

$$\operatorname{tr}\left((F_1 + iF_2)^H (A_{1j} + iA_{2j})\right) = 0 \qquad (j = 1, \dots, m).$$

Let

$$\mathcal{M}^{\perp} = \left\{F \in \mathcal{C}^{n\times n} : \operatorname{tr}(F^H X) = 0 \text{ for all } X \in \mathcal{M}\right\}.$$

Then $F = F_1 + iF_2 \in \mathcal{M}^{\perp}$ if and only if $[F_1, F_2] \in \mathcal{M}_{\mathcal{R}}^{\perp}$. Similar considerations imply that the subdifferential of the norm of $[A_1, A_2]$ is equivalent to the subdifferential of $\|A_1 + iA_2\|$.

Using standard results in convex analysis of real linear spaces to the subspace $\mathcal{M}_{\mathcal{R}}$ and $[A_1, A_2]$, we obtain the following characterization of solutions of the problem (3) (compare Watson [12], Ziętak [15] for a case of approximation in $\mathcal{R}^{m\times n}$).

COROLLARY 1. *A complex matrix $\tilde{X} \in \mathcal{M}$ solves* (3) *if and only if there exists a matrix F such that*

$$F \in \partial\|A - \tilde{X}\|, \qquad F \in \mathcal{M}^{\perp}. \tag{4}$$

Let $\|\cdot\|$ be an arbitrary unitarily invariant norm. The subdifferential of $\|X\|$ is equal to

$$\partial\|X\| = \left\{Y : \operatorname{Re}\operatorname{tr}(Y^H X) = \|X\|, \|Y\|^* \leqslant 1\right\}. \tag{5}$$

A matrix $Y \in \partial\|X\|$ is called the subgradient of $\|X\|$. If $X \neq 0$ then $\|Y\|^* = 1$ in (5).

We now can specialize the conditions (4). Namely, $\tilde{X}$ is a solution of the problem (3) if and only if there exists a matrix F such that

$$\operatorname{Re}\operatorname{tr}\big(F^H(A-\tilde{X})\big) = \|A-\tilde{X}\|, \qquad \|F\|^* = 1, \quad F \in \mathcal{M}^{\perp}. \tag{6}$$

Of course, $\operatorname{Re}\operatorname{tr}(F^H(A-\tilde{X})) = \operatorname{Re}\operatorname{tr}(F^H A)$ and $\operatorname{tr}(F^H X) = 0$ for all $X \in \mathcal{M}$. Therefore in Corollary 1 we can choose a common matrix F for all solutions of (3). Moreover, we have

$$\frac{1}{\|A-\tilde{X}\|}(A-\tilde{X}) \in \partial\|F\|^*,$$

by properties of the subdifferential (see below). This implies the following corollary.

COROLLARY 2. *There exists* $F \in \mathcal{M}^{\perp}$, $\|F\|^* = 1$, *such that every solution* $\tilde{X}$ *of* (3) *has the form*

$$\tilde{X} = A - \delta(A)G \tag{7}$$

for appropriate $G \in \partial\|F\|^*$. *If the subgradient of* $\|F\|^*$ *is unique, then the problem* (3) *has a unique solution.*

REMARKS. In the general case not every $G \in \partial\|F\|^*$ in (7) gives a solution of (3), because $\tilde{X}$ has to belong to $\mathcal{M}$. However, in the next section we show that for Problem I every G determines the solution.

It is a well-known result of von Neumann [6] that there is a correspondence between unitarily invariant norms and the symmetric gauge functions g (see for example Horn and Johnson [3], Stewart and Sun [10]). Let $\|\cdot\|_g$ be the unitarily invariant norm associated with the symmetric gauge function g. Then $\|X\|_g = g(\sigma(X))$. Let X have the following singular-value decomposition (SVD):

$$X = U\Sigma V^H, \tag{8}$$

where U and V are unitary matrices and $\Sigma = \operatorname{diag}(\sigma_j(X))$ is diagonal. As before, we assume the singular values are ordered decreasingly. The dual norm $\|\cdot\|_g^*$ to $\|\cdot\|_g$ is given by (see Stewart and Sun [10, pp. 57, 78])

$$\|X\|_g^* = \max_{\|Y\|_g \leqslant 1} |\operatorname{tr}(Y^H X)| = \max_{\|Y\|_g \leqslant 1} \operatorname{Re}\operatorname{tr}(Y^H X) = \|X\|_{g^*} = g^*(\sigma(X)), \tag{9}$$

where g^* denotes the polar to g,

$$g^*(x) = \max_{y \in \mathscr{R}^n,\ g(y) \leqslant 1} y^T x.$$

We recall that the c_q-norm is dual to the c_p-norm with p and q satisfying $1/p + 1/q = 1$.

Let $\|\cdot\|$ be an arbitrary unitarily invariant norm, and let $\mathscr{Y}(X; \|\cdot\|)$ denote the set of all matrices Y for which we have [see (9)]

$$\|X\|^* = \operatorname{Re}\operatorname{tr}(Y^H X), \qquad \|Y\| = 1. \tag{10}$$

Then

$$\partial\|X\| = \mathscr{Y}(X; \|\cdot\|^*).$$

The matrix $Y \in \mathscr{Y}(X; \|\cdot\|)$ is called the $\|\cdot\|$-dual matrix to X (see de Sá [9], Ziętak [14–16]). We recall that $(\|\cdot\|^*)^* = \|\cdot\|$.

The relations (9) imply that for matrices Y satisfying (10) we have

$$\operatorname{Im}\operatorname{tr}(Y^H X) = 0. \tag{11}$$

Therefore $(X \neq 0)$

$$\text{if} \quad Y \in \partial\|X\|^* \quad \text{then} \quad \operatorname{tr}(Y^H X) = \|X\|^*, \quad \|Y\| = 1. \tag{12}$$

It is known that (see de Sá [9], Ziétak [16]; for the case of real matrices see Watson [11, 12], Ziętak [14])

$$\partial\|X\|_g = \{UDV^H: X = U\Sigma V^H \text{ is any SVD of } X\},$$

where $D = \operatorname{diag}(d_j)$ with diagonal elements $d_1, \dots, d_n$ formed by the components of the vector d being a subgradient of $g(\sigma(X))$, $d \in \partial(g(\sigma(X)))$. If the norm $\|\cdot\|$ is strictly convex, then the subdifferential contains exactly one matrix. The c_p-norm for $1 < p < \infty$ is strictly convex. However, the spectral and trace norms are not strictly convex (see for example Ziętak [14]).

We now recall the forms of matrices from $\partial\|X\|_p$ for $p = 1$ and $p = \infty$ (see de Sá [9], Watson [11], Ziętak [14, 16]; for the case $p = \infty$ compare Berens and Finzel [1]). Let X have the SVD (8), and let s be the number of singular values of X equal to $\sigma_1(X)$. A matrix $Y \in \partial\|X\|_\infty$ if and only if Y has the form

$$Y = U\begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}V^H, \tag{13}$$

where S is a Hermitian positive semidefinite matrix of order s, $\|S\|_1 = 1$. The subgradient of $\|X\|_\infty$ is unique if and only if $\sigma_1(X) > \sigma_2(X)$ (see Ziętak [15]).

Every matrix Y in $\partial\|X\|_1$ has the form

$$Y = U\begin{bmatrix} I_r & 0 \\ 0 & Z \end{bmatrix}V^H, \tag{14}$$

where Z is arbitrary matrix, $\|Z\|_\infty \leqslant 1$, and I_r is the identity matrix of order $r = \operatorname{rank} X$. The subgradient of $\|X\|_1$ is unique if and only if X is of full rank (see Ziętak [15]).

Let $|||\cdot|||_k$ denote the Ky Fan k-norm

$$|||X|||_k = \sum_{j=1}^{k} \sigma_j(X) \qquad (1 \leqslant k \leqslant n).$$

It is a well-known result of Ky Fan that the following statements are equivalent (see for example Stewart and Sun [10, p. 86])

(i) $\|X\| \leqslant \|Y\|$ for every unitarily invariant norm $\|\cdot\|$,
(ii) $|||X|||_k \leqslant |||Y|||_k$ for $k = 1, \dots, n$.

This property is very useful for example when we prove that some matrix is an approximation to a given matrix with respect to every unitarily invariant norm. We will have such situations in the next sections.

Let g_k be the symmetric gauge function corresponding to $|||\cdot|||_k$ ($x = [x_1, \dots, x_n]^T$)

$$g_k(x) = \max_{1 \leqslant j_1 \leqslant \cdots \leqslant j_k \leqslant n} \left\{|x_{j_1}| + \cdots + |x_{j_k}|\right\}.$$

Then (see Horn and Johnson [3, p. 214])

$$g_k^*(x) = \max\left\{\frac{1}{k}\|x\|_1, \|x\|_\infty\right\},$$

where $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote the usual l_1- and l_∞-norms of a vector. Hence

$$|||X|||_k^* = \max\left\{\frac{1}{k}\|X\|_1, \|X\|_\infty\right\}. \tag{15}$$

3. PROBLEM I

We now prove a characterization of all solutions of (3) for $\mathscr{M}$ an arbitrary linear subspace of $\mathscr{C}^{n\times n}$ of dimension $n^2 - 1$ over $\mathscr{C}$ (compare Ziętak [13]; for the spectral norm see also the example in Ziętak [15]).

THEOREM 1. *Let $\mathscr{M}$ be a linear subspace of $\mathscr{C}^{n\times n}$ over $\mathscr{C}$ endowed with an arbitrary unitarily invariant norm $\|\cdot\|$. Let $\mathscr{M}$ have dimension $n^2 - 1$, and let $\mathscr{M}^{\perp} =$ span$\{E\}$. Then $\tilde{X}$ is a solution of (3) if and only if $\tilde{X}$ has the form*

$$\tilde{X} = A - \nu G, \tag{16}$$

where

$$\nu = \operatorname{tr}(E^H A)/\|E\|^*$$

and $G \in \partial\|E\|^*$. Moreover, we have

$$\delta(A) = |\nu|.$$

Proof. Let $\tilde{X}$ have the form (16). Then

$$\operatorname{tr}(E^H\tilde{X}) = \operatorname{tr}(E^H A) - \frac{\operatorname{tr}(E^H A)\operatorname{tr}(E^H G)}{\|E\|^*}.$$

Since $G \in \partial\|E\|^*$, we have [see (11) and (12)]

$$\operatorname{tr}(E^H G) = \|E\|^*, \qquad \|G\| = 1.$$

Thus $\tilde{X} \in \operatorname{span}\{E\}^{\perp} = \mathscr{M}$ and $\|A - \tilde{X}\| = |\nu|$. Therefore $\tilde{X} \in \mathscr{M}$. We now verify that $|\nu| = \delta(A)$. By the properties of the dual norm $\|\cdot\|^*$ we have for every $X \in \mathscr{M}$ [see (9)]

$$\|A - X\| \geqslant \frac{|\operatorname{tr}(E^H A)|}{\|E\|^*} = |\nu| = \|A - \tilde{X}\|.$$

Therefore $\tilde{X}$ determined as in (16) is a solution of (3), and $\delta(A) = |\nu|$; this completes the first part of the proof.

Let now $\tilde{X}$ be a solution of (3). Then there exists a matrix F such that $\mathscr{M}^{\perp} = \operatorname{span}\{F\}$ and $\tilde{X}$ has the form (7) for some $G \in \partial\|F\|^*$. Therefore

$$0 = \operatorname{tr}(F^H \tilde{X}) = \operatorname{tr}(F^H A) - \delta(A)\operatorname{tr}(F^H G).$$

Thus $\delta(A) = \operatorname{tr}(F^H A)/\|F\|^*$ because $\operatorname{tr}(F^H G) = \|F\|^* = 1$ [see (10)–(12)] and consequently $\tilde{X}$ has the form (16). This completes the proof. ■

Theorem 1 means that every $G \in \partial\|E\|^*$ determines a solution of (3). This is a consequence of the assumption $\dim \mathscr{M} = n^2 - 1$. A matrix A has a unique approximation by elements from $\mathscr{M}$ if and only if the subgradient of $\|E\|^*$ is unique.

We now consider Problem I. The set $\mathscr{Z}$ of all zero-trace matrices is a linear subspace over $\mathscr{C}$ of dimension $n^2 - 1$. Therefore Problem I is a particular case of the problem considered in Theorem 1. We now prove that (1) holds for every unitarily invariant norm. For the c_p-norm it was done by Kittaneh [4].

THEOREM 2. *Let $\mathscr{Z}$ be the set of all zero-trace matrices of order n, and let $\|\cdot\|$ be an arbitrary unitarily invariant norm. Then* (1) *holds, and*

$$\tilde{X} = A - \frac{\operatorname{tr}(A)}{n} I \tag{17}$$

is the approximation to A by zero-trace matrices.

Proof. The first part of the theorem is an easy consequence of Theorem 1. Namely, $\mathscr{Z}^{\perp} = \operatorname{span}\{I\}$. Therefore $\mu(A) = |\operatorname{tr}(A)|/\|I\|^*$. Thus we have proven (1).

We now show that the matrix (17) is a solution of Problem I for every unitarily invariant norm. For this purpose we apply the Ky Fan k-norms $|||\cdot|||_k$, $1 \leqslant k \leqslant n$. Let $\tilde{X}$ be determined as in (17). Then $\tilde{X} \in \mathscr{Z}$ and

$$|||A - \tilde{X}|||_k = \frac{|\operatorname{tr}(A)|}{n} |||I|||_k = \frac{|\operatorname{tr}(A)|}{|||I|||_k^*} = \mu(A),$$

because $\|I\|_k^* = n/k$ [see (15)]. Thus $\tilde{X}$ is a solution of Problem I for the Ky Fan k-norms. Therefore for every $X \in \mathscr{Z}$ we have

$$|||A - X|||_k \geqslant |||A - \tilde{X}|||_k \qquad (1 \leqslant k \leqslant n).$$

By the result of Ky Fan [see (i) and (ii) in Section 2] $\tilde{X}$ is a solution of (1) for every unitarily invariant norm. This completes the proof. ∎

The formula (17) was given by Kittaneh [4] for the c_p-norms. Immediately from (17) we have $\mu(A) = \|A - \tilde{X}\| = |\mathrm{tr}(A)|\,\|I\|/n$ for every unitarily invariant norm. However, we have (1). This implies the following corollary.

COROLLARY 3. *For every unitarily invariant norm we have*

$$\|I\|\,\|I\|^* = n$$

and consequently

$$\frac{1}{\|I\|} I \in \partial\|I\|^*. \tag{18}$$

The orthogonal complement of $\mathscr{Z}$ is spanned by I. Therefore from (13), (14), and (16) we obtain the following corollary for Problem I.

COROLLARY 4. *A matrix $\tilde{X}$ is a solution of Problem I if and only if $\tilde{X}$ has the form*

$$\tilde{X} = A - \frac{\mathrm{tr}(A)}{\|I\|^*} G, \qquad G \in \partial\|I\|^*.$$

In particular, if $\|\cdot\|$ is the trace norm, then

$$\tilde{X} = A - \mathrm{tr}(A)\,G, \qquad G \text{ Hermitian, positive semidefinite}, \quad \|G\|_1 = 1.$$

For the spectral norm the matrix (17) *is the unique solution of Problem I.*

If $A = \alpha I$, $\alpha \in \mathscr{C}$, then $\tilde{X} = 0$ is the solution of Problem I for each unitarily invariant norm. Thus we have for every unitarily invariant norm (compare Problem II)

$$\|\alpha I - X\| \geqslant \|\alpha I\|, \qquad X \in \mathscr{Z}. \tag{19}$$

For the spectral norm the equality holds in (19) if and only if $X = 0$.

From Theorem 2 we obtain immediately the following corollary, which generalizes Corollary 1 in Kittaneh [4].

COROLLARY 5. *Let A and B be arbitrary. Then for every unitarily invariant norm we have*

$$\|A\| \geqslant \min_{X} \|A - (BX - XB)\| \geqslant \min_{X \in \mathscr{Z}} \|A - X\| = \frac{|\operatorname{tr}(A)|}{\|I\|^*}.$$

In particular, for $A = \alpha I$ we obtain (compare (19))

$$\|\alpha I - (BX - XB)\| \geqslant \|\alpha I\|. \tag{20}$$

The inequality (20) is a generalization of the result of Halmos [2], given by him for operators. The result of Halmos can be formulated for matrices in the following way. Let $\|\cdot\|$ be the spectral norm. Then

$$\|I - (BX - XB)\| \geqslant \|I\|.$$

We stress that the inequality (20) holds for every unitarily invariant norm and says that the zero matrix is an approximation to αI by matrices of the form $BX - XB$.

Let A and the norm $\|\cdot\|$ be such that $\|A\| \, \|I\|^* = |\operatorname{tr}(A)|$. Then the zero matrix is an approximation to A by matrices $BX - XB$. If A is a Hermitian positive semidefinite matrix, then $\|A\|_1 = \operatorname{tr}(A)$. Thus the following corollary is obtained immediately from Corollary 5.

COROLLARY 6. *Let A be Hermitian positive semidefinite. Then for every B and for the trace norm we have*

$$\min_{X} \|A - (BX - XB)\|_1 = \min_{X \in \mathscr{Z}} \|A - X\|_1 = \|A\|_1.$$

The linear subspace $\mathscr{Z}$ is distinguished among all subspaces $\mathscr{M}$ of dimension $n^2 - 1$. It is easy to verify that a linear subspace $\mathscr{M}$ of dimension $n^2 - 1$ over $\mathscr{C}$ is equal to $\mathscr{Z}$ if and only if $\mathscr{M}^{\perp} = \operatorname{span}\{I\}$. We now give a new geometrical characterization of $\mathscr{Z}$, completing the characterizations presented in Petz and Zemánek [7]. For this purpose we consider the problem (3) for $A = I$.

THEOREM 3. *Let $\mathscr{M}$ be a linear subspace of $\mathscr{C}^{n\times n}$ of dimension n^2-1 over $\mathscr{C}$, and let an arbitrary unitarily invariant norm $\|\cdot\|$ be such that the subgradient of $\|I\|$ is unique. Then $\mathscr{M}=\mathscr{Z}$ if and only if*

$$\min_{X\in\mathscr{M}} \|I-X\| = \|I\|. \tag{21}$$

Proof. If $\mathscr{M}=\mathscr{Z}$, then (21) holds for every unitarily invariant norm because we have (19).

Let now the relation (21) hold, and let $\mathscr{M}^{\perp}=\operatorname{span}\{E\}$. Then there exists F such that (see Corollary 1)

$$F\in\partial\|I\|,\qquad F\in\mathscr{M}^{\perp}.$$

Thus $\operatorname{span}\{E\}=\operatorname{span}\{F\}$. Moveover, the subgradient F of $\|I\|$ is unique. Therefore F has to be equal to $(1/\|I\|^*)I$ [see (18)]. This completes the proof. ∎

In Theorem 3 the uniqueness of the subgradient of $\|I\|$ plays a crucial role. We recall that the subgradient of $\|I\|_p$ is unique for $1\leqslant p<\infty$. Unfortunately, this is not true for the spectral norm. Therefore, for the spectral norm F can be different from $(1/\|I\|_1)I$. Thus it is impossible to characterize $\mathscr{Z}$ by means of the problem (21). The following example shows that (21) can hold for $\mathscr{M}$ different from $\mathscr{Z}$ and the spectral norm. Let $\mathscr{M}^{\perp}=\operatorname{span}\{E\}$ where E is nonsingular Hermitian positive definite, for example, $E=\operatorname{diag}(1,2,\dots,n)$. Then we have $\nu=1$, and $\tilde{X}=0$ is the solution of (21) [see (16)]. Therefore (21) holds although $\mathscr{M}$ is different from $\mathscr{Z}$. In the next section we give a characterization of a zero-trace matrix by means of another approximation problem with the same unitarily invariant norms as in Theorem 3.

4. PROBLEM II

Let B be a complex matrix of order n. We have mentioned in the Introduction that Kittaneh [5] has proven that (2) holds for the c_p-norms, $1\leqslant p<\infty$, if and only if $\operatorname{tr}(B)=0$. To explain better why this does not hold for the spectral norm, we now give a new proof of the result of Kittaneh. For this purpose we use a standard technique from convex analysis (see Section 2).

The problem (2) with arbitrary complex matrix B is a particular case of the problem (3) with $\mathscr{M} = \text{span}\{B\}$ and $A = I$. Let us consider (2) for the c_p-norm with $1 \leqslant p \leqslant \infty$. The condition (2), i.e., $\gamma(B) = \|I\|_p$, is satisfied if and only if there exists a matrix F such that [see (6)]

$$\|I\|_p = \text{Re}\,\text{tr}(F^H I), \qquad \|F\|_q = 1, \tag{22}$$

$$\text{tr}(F^H B) = 0, \tag{23}$$

where

$$1/p + 1/q = 1.$$

The matrix F satisfying (22) is the subgradient of $\|I\|_p$. Let $1 < p < \infty$. The strict convexity of $\|\cdot\|_p$ implies that the subgradient of $\|\cdot\|_p$ is unique. It is easy to verify that in this case $F = n^{(-1+1/p)}I$. Thus the condition (23) is equivalent to $\text{tr}(B) = 0$. Let $p = 1$. Then $F = I$ is the unique matrix satisfying (22), because I is of full rank [see (14)]. Hence the condition (23) is fulfilled if and only if $\text{tr}(B) = 0$. In this way we have proven the result of Kittaneh. His result holds because the subgradient of $\|I\|_p$, $1 \leqslant p < \infty$, is unique. For $p = \infty$ the situation is completely different because F satisfying (22) is not unique. Namely, the conditions (22) are fulfilled for every Hermitian positive semidefinite matrix F, $\|F\|_1 = 1$ [see (13)]. Therefore the condition (23) can be satisfied not only by zero-trace matrices B.

We now deal with Problem II for arbitrary unitarily invariant norms.

THEOREM 4. *Let* $\text{tr}(B) = 0$. *Then for every complex number* z *and arbitrary unitarily invariant norm we have*

$$\|I + zB\| \geqslant \|I\|, \tag{24}$$

i.e., (2) *holds.*

Proof. As in the proof of Theorem 2, it is sufficient to prove (24) for the Ky Fan k-norms.

Let $F = (1/|||\,I\,|||_k^*)I$. Then $F \in \partial\,|||\,I\,|||_k$ and $F \in \text{span}\{B\}^\perp$, because $\text{tr}(B) = 0$. This means that $z = 0$ is a solution of (2) for the Ky Fan k-norms $(1 \leqslant k \leqslant n)$, because the conditions formulated in Corollary 1 are satisfied for the problem (2). Therefore we have for every $z \in \mathscr{C}$

$$|||\,I + zB\,|||_k \geqslant |||\,I\,|||_k.$$

Hence for every unitarily invariant norm we have by the properties of the unitarily invariant norms

$$\|I + zB\| \geq \|I\|.$$

This completes the proof. ■

The above considerations imply that the characterization of a zero-trace matrix by means of the problem (2) for the norm $\|\cdot\|$ is possible if the subgradient of $\|I\|$ is unique, because then F has to be equal to $(1/\|I\|^*)I$ [see (18)]. The condition $\operatorname{tr}(B) = 0$ is sufficient to have $\gamma(B) = \|I\|$ for every unitarily invariant norm [see (2)]. We now prove that it is also necessary if $\|\cdot\|$ satisfies the assumptions of Theorem 3.

THEOREM 5. *Let a norm* $\|\cdot\|$ *satisfy the assumptions of Theorem* 3. *Then* $\operatorname{tr}(B) = 0$ *if and only if*

$$\min_{z \in \mathscr{C}} \|I + zB\| = \|I\|. \tag{25}$$

Proof. The necessity follows from Theorem 4. We now prove the sufficiency. The condition (25) is satisfied if and only if there exists F such that [see (4)]

$$F \in \partial\|I\|, \qquad F \in \operatorname{span}\{B\}^{\perp}.$$

Since the subgradient of $\|I\|$ is unique, the matrix F has to be equal to $(1/\|I\|^*)I$ because we have (18). Therefore B has zero trace because

$$\operatorname{tr}(B) = \|I\|^* \operatorname{tr}(F^H B) = 0.$$

This completes the proof. ■

The characterization, given in Theorem 5, of a zero-trace matrix is a generalization of the above-mentioned result of Kittaneh [5].

The author would like to thank Professor J. Zemánek from the Polish Academy of Sciences (Warsaw) for References [4] and [5] and for helpful discussions which inspired her to write this paper. She would like to thank Professor Chandler Davis for interesting conversation. She is also grateful to an anonymous referee for the reference to R. Bhatia, Some inequalities for

norm ideals II, J. Indian Math. Soc. 50:131–138 (1986), where Theorem 2 and the first part of Corollary 3 were proved in a different manner.

REFERENCES

1 H. Berens and M. Finzel, A continuous selection of the metric projection in matrix spaces, in Numerical Methods of Approximation Theory (G. Meinardus and G. Nürnberger, Eds.), ISBN 81, Birkhäuser, Basel, 1987.
2 P. R. Halmos, Commutators of operators, II, *Amer. J. Math.* 76:191–198 (1954).
3 R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge U.P., Cambridge, 1991.
4 F. Kittaneh, Inequalities for the Schatten p-norm II, *Glasgow Math. J.* 29:99–104 (1987).
5 F. Kittaneh, On zero-trace matrices, *Linear Algebra Appl.* 151:119–124 (1991).
6 J. von Neumann, Some matrix inequalities and metrization of metric spaces, *Tomsk Univ. Rev.* 1:286–300 (1937); in *Collected Works*, Pergamon, Oxford, 1962, pp. 205–219.
7 D. Petz and J. Zemánek, Characterizations of the trace, *Linear Algebra Appl.* 111:43–52 (1988).
8 R. T. Rockafellar, *Convex Analysis*, Princeton U.P., Princeton, N.J., 1970.
9 E. M. de Sá, Exposed faces and duality for symmetric and unitarily invariant norms, *Linear Algebra Appl.* 197/198:429–450 (1994).
10 G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic, Boston, 1990.
11 G. A. Watson, Characterization of the subdifferentials of some matrix norms, *Linear Algebra Appl.* 170:33–45 (1992).
12 G. A. Watson, On matrix approximation problems with Ky Fan k norms, *Numer. Algorithms* 5:263–272 (1993).
13 K. Ziętak, On the linear approximation of matrices, Report N-189, Inst. of Computer Science, Univ. of Wrocław, Wrocław, 1988.
14 K. Ziętak, On the characterization of the extremal points of the unit sphere of matrices, *Linear Algebra Appl.* 106:57–75 (1988).
15 K. Ziętak, Properties of linear approximations of matrices in the spectral norm, *Linear Algebra Appl.* 183:41–60 (1993).
16 K. Ziętak, Subdifferentials, faces and dual matrices, *Linear Algebra Appl.* 185:125–141 (1993).

Received 6 September 1994; final manuscript accepted 20 January 1995