APPLICATIONES MATHEMATICAE
24,3 (1997), pp. 267–280

K. ZIĘTAK (Wrocław)

STRICT SPECTRAL APPROXIMATION OF A MATRIX AND SOME RELATED PROBLEMS

Abstract. We show how the strict spectral approximation can be used to obtain characterizations and properties of solutions of some problems in the linear space of matrices. Namely, we deal with

(i) approximation problems with singular values preserving functions,
(ii) the Moore–Penrose generalized inverse.

Some properties of approximation by positive semi-definite matrices are commented.

1. Introduction. Let $\mathcal{M}$ be a nonempty closed convex subset of the normed linear space $\mathbb{C}^{m\times n}$ of $m \times n$ complex matrices. We consider the following problem:

$$\min_{X\in\mathcal{M}} \|A - X\|_\infty, \tag{1}$$

where $\|\cdot\|_\infty$ is the spectral norm and $A \in \mathbb{C}^{m\times n}$ is given. Let $\sigma_j(X)$ denote the jth singular value of X. The singular values $\sigma_j(X)$ are defined by saying that the eigenvalues of the Hermitian positive semi-definite matrix X^HX are $\sigma_j^2(X)$ [see for example Golub and Van Loan (1989)]. We assume that the singular values are ordered decreasingly:

$$\sigma_1(X) \ge \ldots \ge \sigma_t(X) \ge 0 \quad (t = \min\{m, n\}).$$

The vector of ordered singular values of X is denoted by $\sigma(X)$. The spectral norm $\|X\|_\infty = \sigma_1(X)$ is a particular case of the c_p-norm which is the l_p-norm of the vector of singular values. The c_p-norms are unitarily invariant, i.e. $\|UAV\|_p = \|A\|_p$ for all unitary matrices U and V. The properties

1991 *Mathematics Subject Classification*: 15A60, 15A09, 15A48, 15A57.
Key words and phrases: strict spectral approximation of a matrix, c_p-minimal approximation, singular values preserving functions, positive semi-definite matrix, Moore–Penrose generalized inverse.

of unitarily invariant norms are presented for example in Horn and Johnson (1986), and Stewart and Sun (1990).

A matrix $\widehat{X} \in \mathcal{M}$ for which the minimum (1) is reached is called a *spectral approximation* to A by elements from $\mathcal{M}$. It may not be unique in the general case because the spectral norm is not strictly convex. Among all spectral approximations we select the best one defined as follows [see Ziętak (1995)]. A matrix $A^{(\mathrm{st})} \in \mathcal{M}$ is a *strict spectral approximation* of A by elements from $\mathcal{M}$ if the vector $\sigma(A - A^{(\mathrm{st})})$ is minimal with respect to the lexicographic ordering $\leq_l$ in the set $\{\sigma : \sigma = \sigma(A - X), X \in \mathcal{M}\}$. For example, if $u = [3, 3, 2, 0]$ and $v = [3, 2, 2, 2]$ then $v \leq_l u$. The strict spectral approximation always exists and it is unique [see Ziętak (1995)]. If $\mathcal{M}$ is the set of all Hermitian positive semi-definite matrices, which is known to be convex, then $A^{(\mathrm{st})}$ is called the *strict spectral positive approximant* to A.

The strict spectral approximation of a matrix is a generalization of the strict approximation of a vector, introduced by Rice (1962) [see also Huotari and Li (1994)]. The definition of $A^{(\mathrm{st})}$ is connected with an order in the space of matrices. We say $X \leq_{\mathrm{st}} Y$ if the vector $\sigma(X)$ is equal to or smaller than $\sigma(Y)$ in the lexicographic order on $\mathbb{R}^t$, $\sigma(X) \leq_l \sigma(Y)$. The ordering $\leq_{\mathrm{st}}$ is called the *strict spectral ordering* and it was used in Ziętak (1995). The same ordering was applied by Young (1986) to distinguish some solution of the Nevanlinna–Pick problem for matrix-valued functions, called *superoptimal approximation* [compare Davis (1976), Woerdeman (1994)].

On the space of matrices one considers also other orders. The most popular is the *Loewner order* $\preceq$ on Hermitian matrices: $X \preceq Y$ if and only if $Y - X$ is positive semi-definite. The c_p-*minimality*, introduced by Rogers and Ward (1981), can also be considered as an ordering. In this paper we prove that the c_p-minimal ordering coincides with the strict spectral ordering in some sense. This implies that if a matrix $\widehat{X}$ is a common best approximation to A with respect to the c_p-norm for every p then $\widehat{X}$ is the strict spectral approximation. Using this we show that some specific matrices are strict spectral approximants.

We also investigate the properties of approximations, with respect to the spectral norm, of a matrix by matrices from a linear subspace $\mathcal{M}$ described by linear, singular values preserving functions. The Hermitian matrices are an example of such a real linear subspace. We describe all approximants of a matrix and we give a necessary and sufficient condition for uniqueness. We deal with that in Section 3.

In the last section some characterizations of the Moore–Penrose generalized inverse are presented.

The problem (1) is a particular case of matrix nearness problems. A survey and applications of nearness problems, in areas including control theory, numerical analysis, statistics and optimization, are given in Higham (1989). For example, approximation by symmetric matrices occurs in optimization when approximating a Hessian matrix by finite differences of a gradient vector [see Gill, Murray and Wright (1981), p. 116]. The most well-known application of approximation by positive semi-definite matrices is in detecting and modifying an indefinite Hessian matrix in Newton methods for optimization [see Gill, Murray and Wright (1981), Sec. 4.4.2].

2. Strict approximation of matrices. The strict spectral approximation $A^{(\mathrm{st})}$ is the unique matrix for which we have

$$A - A^{(\mathrm{st})} \le_{\mathrm{st}} A - X \quad \text{for all } X \in \mathcal{M}.$$

From this we conclude that if the singular values of $A - A^{(\mathrm{st})}$ are all equal then the problem (1) has a unique solution. This conclusion helps us to obtain a necessary and sufficient condition under which the problem (1) has a unique solution for some special cases when we know $A^{(\mathrm{st})}$ explicitly. We will present such examples.

We now prove another characterization of the strict spectral approximant $A^{(\mathrm{st})}$.

THEOREM 1. *A matrix $\widehat{X} \in \mathcal{M}$ is the strict spectral approximant to A if and only if for every $X \in \mathcal{M}$, $X \ne \widehat{X}$, we have*

$$\|A - X\|_p > \|A - \widehat{X}\|_p \quad \textit{for all } p \textit{ sufficiently large.} \tag{2}$$

Proof. Let $\widehat{X}$ be the strict approximant to A, $\widehat{X} = A^{(\mathrm{st})}$, and let $X \in \mathcal{M}, X \ne \widehat{X}$. Then the vector $\sigma(A - X)$ is not minimal on $\{\sigma : \sigma = \sigma(A - Z),\ Z \in \mathcal{M}\}$ in the ordinary lexicographic ordering on $\mathbb{R}^t$. Therefore there exists an index $j < t$ such that

$$\sigma_k(A - X) = \sigma_k(A - \widehat{X}), \quad k = 1, \ldots, j-1, \tag{3}$$

$$\sigma_j(A - X) > \sigma_j(A - \widehat{X}). \tag{4}$$

Thus we have

$$\begin{aligned}\|A - X\|_p^p - \|A - \widehat{X}\|_p^p &= \sum_{k=j}^{t} [\sigma_k^p(A - X) - \sigma_k^p(A - \widehat{X})] \\ &\ge \sum_{k=j}^{t} \sigma_k^p(A - X) - (t + 1 - j)\sigma_j^p(A - \widehat{X}).\end{aligned}$$

The inequality (4) implies that for all p large enough we have

$$\sigma_j^p(A - X) > (t + 1 - j)\sigma_j^p(A - \widehat{X}) \tag{5}$$

and consequently $\|A - X\|_p^p - \|A - \widehat{X}\|_p^p > 0$. Therefore the condition (2) is satisfied.

Assume to the contrary that $\widehat{X}$ satisfying (2) is not the strict approximant $A^{(\mathrm{st})}$. Then there exists j such that for $\widehat{X}$ and $A^{(\mathrm{st})}$ conditions analogous to (3) and (4) are satisfied:

$$\sigma_k(A-\widehat{X}) = \sigma_k(A-A^{(\mathrm{st})}), \qquad k=1,\ldots,j-1,$$
$$\sigma_j(A-\widehat{X}) > \sigma_j(A-A^{(\mathrm{st})}).$$

Thus for all p large enough we have [compare (5)]

$$\|A-A^{(\mathrm{st})}\|_p^p - \|A-\widehat{X}\|_p^p \le (t+1-j)\sigma_j^p(A-A^{(\mathrm{st})}) - \sigma_j^p(A-\widehat{X}) < 0,$$

which contradicts (2) for $X = A^{(\mathrm{st})}$. ∎

The property of strict spectral approximation described in Theorem 1 is similar to that of the strict approximation of a vector [see Lemma 2.1 in Houtari and Li (1994)].

We say that a matrix $\widehat{X}\in\mathcal{M}$ is *c_p-minimal* if (2) holds. This notion was introduced in Rogers and Ward (1981) to construct a c_p-minimal positive approximant P_m of an operator in a finite-dimensional complex Hilbert space. They show that each operator A has a c_p-minimal positive approximant and they state that P_m seems to be the operator analogue of the strict approximant of a vector. Moreover, they show that $A-P_m$ is normal. Theorem 1 implies that the c_p-minimal positive approximant of a matrix is exactly the strict spectral positive approximant. We stress that the ordering which leads to the definition of $A^{(\mathrm{st})}$, i.e. the strict spectral ordering, is more natural than the one defined in the theorem, i.e. the c_p-minimality. However, in some special cases the property (2) is helpful when verifying if a matrix is a strict spectral approximant. Theorem 1 extends the result of Rogers and Ward to the spectral approximation of a matrix by elements from an arbitrary convex subset $\mathcal{M}$. Namely, Theorem 1 implies that the c_p-minimal approximation (i.e. the strict spectral approximation) always exists for every convex subset $\mathcal{M}$ and that it is unique.

Let $\mathcal{M}$ and A be such that the same matrix $\widehat{X}$ is the approximation to A by matrices from $\mathcal{M}$ with respect to the c_p-norm for every p. Then $\widehat{X}$ is necessarily the strict spectral approximation because (2) holds. Therefore we can formulate the following corollary.

Corollary 2. *If for* $\widehat{X}\in\mathcal{M}$,

$$\|A-\widehat{X}\|_p = \min_{X\in\mathcal{M}}\|A-X\|_p \quad \textit{for every } p, \tag{6}$$

then $\widehat{X}$ is the strict spectral approximation to A by elements from $\mathcal{M}$. If additionally the singular values of $A-\widehat{X}$ are all equal then the spectral approximation of A is unique.

Corollary 2 can lead to new characterizations of some specific matrices. In the next sections we illustrate this on some special cases of approximation of matrices. Now we show how Corollary 2 can be used to explain some known properties of positive semi-definite approximants of a matrix.

Let A be a complex matrix of order n and let $\mathcal{M}$ be the set of all $n \times n$ Hermitian positive semi-definite matrices. Let $B + iC$ be the *Cartesian representation* of A, i.e. $B = B^H$ and $C = C^H$. Then [see Halmos (1972)]

$$\delta_A \equiv \min_{P \in \mathcal{M}} \|A - P\|_\infty = \min\{r \geq 0 : r^2 I - C^2 \succeq 0 \text{ and } B + (r^2 I - C^2)^{1/2} \succeq 0\}$$

where $X^{1/2}$ denotes the positive square root of the matrix X and r is a real number. The matrix $P^{(\mathrm{hl})}$, called the *Halmos approximant*, defined by

$$P^{(\mathrm{hl})} = B + (\delta_A^2 I - C^2)^{1/2}$$

is a positive approximant to A with respect to the spectral norm. This approximant was shown by Bouldin (1973) to be maximal in the sense of the Loewner ordering, among all positive approximants P to A, i.e. $P \preceq P^{(\mathrm{hl})}$ for every positive approximant P. Unfortunately, there need not be a positive approximant minimal in the Loewner ordering [see Rogers and Ward (1981)]. We now show that it is better to compare the differences $A - P$ with respect to the strict spectral ordering.

Let $P^{(\mathrm{st})}$ denote the strict spectral positive approximant to A. Then $A - P^{(\mathrm{st})}$ is minimal in the sense of the strict spectral ordering among $A - P$ for every positive approximant P of A. It is easily seen that $A - P^{(\mathrm{hl})}$ is maximal in the strict spectral ordering because every singular value of $A - P^{(\mathrm{hl})}$ is equal to δ_A, i.e. $A - P^{(\mathrm{hl})}$ is a multiple of a unitary matrix. Hence we have

$$A - P^{(\mathrm{st})} \leq_{\mathrm{st}} A - P \leq_{\mathrm{st}} A - P^{(\mathrm{hl})} \quad \text{for every positive approximant } P \text{ of } A.$$

If every singular value of $A - P^{(\mathrm{st})}$ is equal to δ_A then by the uniqueness of the strict spectral approximation we obtain $P^{(\mathrm{st})} = P^{(\mathrm{hl})}$ and in this case the positive approximation is unique. On the other hand, if the positive approximant is unique then it has to be equal to $P^{(\mathrm{hl})}$, so $P^{(\mathrm{st})} = P^{(\mathrm{hl})}$. Therefore the condition that $A - P^{(\mathrm{st})}$ is a multiple of a unitary matrix is necessary and sufficient for uniqueness. From this we can obtain new proofs of some known results. We explain this in the case of normal matrices.

If A is normal then, in each unitarily invariant norm, the *positive part* $B^{(+)}$ of B,

$$B^{(+)} = \tfrac{1}{2}[(B^H B)^{1/2} + B], \tag{7}$$

is a positive approximant to A [see Bhatia and Kittaneh (1992)]. Therefore we have (6) for $\widehat{X} = B^{(+)}$. Hence Corollary 2 implies that the positive part $B^{(+)}$ of B is the strict spectral positive approximant. Since in this case the strict spectral positive approximant is given explicitly, we can easily obtain the necessary and sufficient conditions characterizing normal matrices A which have a unique positive spectral approximant. Namely, a normal matrix A has a unique positive approximation if and only if $B^{(+)} = P^{(\mathrm{hl})}$. This leads to a new proof of the result of Ando *et al.* (1973) [compare Bouldin (1973)] that $B^{(+)}$ is the unique spectral positive approximation to A if and only if $A - B^{(+)}$ is a multiple of a unitary matrix, i.e. the singular values of $A - B^{(+)}$ are all equal. Moreover, the zero matrix is the unique spectral positive approximation of A if and only if $B^{(+)} = 0$ and A is a multiple of a unitary matrix [see Ando *et al.* (1973), compare Bhatia and Kittaneh (1992)].

3. Approximation problems with singular values preserving functions. We say that a function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ is *singular values preserving* if X and $f(X)$ have the same singular values. We now recall very interesting results of Li and Tsing (1987).

THEOREM 3 (Li and Tsing). *Let a function* $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ *satisfy*

- f *is singular values preserving*,
- f *is real linear*, *i.e.* $f(\lambda X + \mu Y) = \lambda f(X) + \mu f(Y)$ *for any* $X, Y \in \mathbb{C}^{n\times n}$ *and* $\lambda, \mu \in \mathbb{R}$,
- $f = f^{-1}$, *i.e.* $f^2(X) = X$ *for any* $X \in \mathbb{C}^{n\times n}$.

If $A \in \mathbb{C}^{n\times n}$ *and*

$$\mathcal{S}_f = \{X \in \mathbb{C}^{n\times n} : f(X) = X\}$$

then

$$A^{(f)} = \tfrac{1}{2}[A + f(A)] \in \mathcal{S}_f, \tag{8}$$

and for any unitarily invariant norm,

$$\|A - A^{(f)}\| = \min_{X\in\mathcal{S}_f} \|A - X\|. \tag{9}$$

It is known that f is real linear and singular values preserving if and only if there exist unitary matrices P and Q such that

$$f(X) = PX^{\square}Q \quad \text{for all } X \in \mathbb{C}^{n\times n}, \tag{10}$$

where $X^{\square}$ stands for $X, X^T, \overline{X}$ or X^H [see Li and Tsing (1987)].

The set $\mathcal{S}_f$ is a real linear space. The matrix $A^{(f)}$ is the best approximation to A over $\mathcal{S}_f$ for every unitarily invariant norm. Therefore it is the strict spectral approximation.

In the general case the solution of (9) is not unique for the spectral norm because this norm is not strictly convex. Higham (1989) mentions that the uniqueness of the approximation by Hermitian matrices is an open question for the spectral norm. We will answer this question. For this purpose we will characterize all solutions of (9) for the following functions f (see (10)):

$$f(X) = X^H, \tag{11}$$

$$f(X) = X^T, \tag{12}$$

$$f(X) = \overline{X}. \tag{13}$$

For $f(X) = X^H$ the set $\mathcal{S}_f$ is the set of all Hermitian matrices.

Let f satisfy the assumptions of Theorem 3 and let $g = -f$. Then $\mathcal{S}_f \cap \mathcal{S}_g = \{0\}$ and $\mathbb{C}^{n\times n} = \mathcal{S}_f \oplus \mathcal{S}_g$. Therefore each $A \in \mathbb{C}^{n\times n}$ can be uniquely expressed in the form $A = A_1 + A_2$, where $A_1 \in \mathcal{S}_f$ and $A_2 \in \mathcal{S}_g$. It is easy to verify that

$$A_1 = (A + f(A))/2, \qquad A_2 = (A - f(A))/2. \tag{14}$$

Moreover, we have

$$A_2^{(f)} = (A_2 + f(A_2))/2 = 0.$$

Hence the zero matrix is the best approximation to A_2 over $\mathcal{S}_f$. Let X be the best approximation to A over $\mathcal{S}_f$ and let $B = A + C$, $C \in \mathcal{S}_f$. Then $X + C$ is the best approximation to B over $\mathcal{S}_f$ because we approximate by elements from a linear space. Therefore it is sufficient to consider only the case $A \in \mathcal{S}_g$.

We now describe all best approximations to $A \in \mathcal{S}_g$ over $\mathcal{S}_f$, with respect to the spectral norm.

THEOREM 4. *Let f satisfy the assumptions of Theorem* 3, $g = -f$ *and let $A \in \mathcal{S}_g$ have the following singular value decomposition* (*SVD*):

$$A = U\Sigma V^H = [U_1, U_2]\,\mathrm{diag}(\sigma_1 I_s, \Sigma_2)\begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix}, \tag{15}$$

where the singular values σ_j of A are ordered and

$$\sigma_1 = \sigma_2 = \ldots = \sigma_s > \sigma_{s+1} \ge \sigma_{s+2} \ge \ldots \ge \sigma_n \ge 0,$$
$$\Sigma_2 = \mathrm{diag}(\sigma_{s+1}, \ldots, \sigma_n),$$

U_1, U_2, V_1 *and* V_2 *are blocks of U and V, respectively*, $U_1, V_1 \in \mathbb{C}^{n\times s}$. *Then every solution of* (9) *for the spectral norm has the form*

$$\widehat{X} = A - \sigma_1 U_1 V_1^H - U_2 Z V_2^H = U \begin{bmatrix} 0 & 0 \\ 0 & G \end{bmatrix} V^H, \tag{16}$$

where $G = \Sigma_2 - Z$, and Z is such that $\widehat{X} \in \mathcal{S}_f$ and $\|Z\|_\infty \le \sigma_1$. Moreover, the matrix $U_1V_1^H$ is uniquely determined.

P r o o f. Since the zero matrix is the strict spectral approximation to A, the residue matrix $\widehat{R} = A - \widehat{X}$ has at least s singular values equal to σ_1 for every solution $\widehat{X}$ of (9). By similar arguments to those used in the proof of Theorem 4.3 in Ziętak (1993), it is easy to show that every $\widehat{X}$ has the form (16). We omit the details. ■

Let $A \in \mathcal{S}_g$. We now specialize conditions under which the problem (9) has a unique solution. If $s = n$ in Theorem 4 then the problem (9) has a unique solution for the spectral norm as an obvious consequence of the properties of strict spectral approximation. We now verify that this condition is also necessary. Namely, we show that if $s < n$ then there exists a nonzero matrix G such that $\widehat{X}$ given by (16) is a solution of (9). We show this for functions (11)–(13).

Case 1. Let f be as in (11), $g = -f$, and let $A \in \mathcal{S}_g$. Then $\mathcal{S}_g$ is the set of all skew-Hermitian matrices. Therefore A has the form

$$A = iQ\,\mathrm{diag}(\lambda_j)Q^H,$$

where Q is unitary, and λ_j are ordered real numbers

$$|\lambda_1| = |\lambda_2| = \ldots = |\lambda_s| > |\lambda_{s+1}| \ge \ldots \ge |\lambda_n| \ge 0.$$

Let $D = \mathrm{diag}(d_j)$ with $d_j = \mathrm{sgn}(\lambda_j)$ for $\lambda_j \neq 0$, and $d_j = 1$ for $\lambda_j = 0$. Then A has the SVD (15) with $U = iQD$, $V = Q$ and $\Sigma = \mathrm{diag}(|\lambda_j|)$. Thus every solution $\widehat{X}$ of (9) has the form (see (16))

$$\widehat{X} = iQD\begin{bmatrix} 0 & 0 \\ 0 & G \end{bmatrix}Q^H$$

for appropriate $G \in \mathbb{C}^{(n-s)\times(n-s)}$ such that $\|\widehat{X} - A\|_\infty = \sigma_1$. Since $\widehat{X}$ has to be Hermitian, we have

$$\widehat{X} = \widehat{X}^H = -iQ\begin{bmatrix} 0 & 0 \\ 0 & G^H \end{bmatrix}DQ^H.$$

This implies

$$D\begin{bmatrix} 0 & 0 \\ 0 & G \end{bmatrix} = -\begin{bmatrix} 0 & 0 \\ 0 & G^H \end{bmatrix}D. \tag{17}$$

We choose $G = \alpha i I_{n-s}$, $\alpha \in \mathbb{R}, \alpha \neq 0$. Thus $G \in \mathcal{S}_g$ and (17) holds. The parameter α has to be chosen such that

$$\|\Sigma_2 - \alpha i D_2\|_\infty = \sigma_{s+1} + |\alpha| \le \sigma_1.$$

Such a nonzero α exists because we have assumed $\sigma_s > \sigma_{s+1}$. In this way we have proven that if $s < n$ then there exists a nonzero approximation $\widehat{X}$ of A. Therefore $A \in \mathcal{S}_g$ has a unique spectral approximation by Hermitian matrices if and only if the singular values of A are all equal, i.e., A is a multiple of a matrix H, $A = \sigma H$, where σ is a positive number and H is a unitary and skew-Hermitian matrix.

Case 2. Let f be as in (12). Then $\mathcal{S}_g$ is the set of skew-symmetric complex matrices. Let $A \in \mathcal{S}_g$ have the SVD (15). Then

$$(U\Sigma V^H)^T = \overline{V}\Sigma U^T = -U\Sigma V^H.$$

By the properties of the singular decomposition [see for example de Sá (1994)] there exist unitary matrices $D_{11} \in \mathbb{C}^{s\times s}$ and $D_{12}, D_{22} \in \mathbb{C}^{(n-s)\times(n-s)}$ such that

$$\overline{V} = -UD_1, \quad \overline{U} = VD_2,$$

where $D_1 = \operatorname{diag}(D_{11}, D_{12})$ and $D_2 = \operatorname{diag}(D_{11}, D_{22})$. We obtain $\overline{V} = -\overline{V}\overline{D}_2 D_1$. Therefore $\overline{D}_2 D_1 = -I_n$ and $\overline{D}_{22} D_{12} = -I_{n-s}$. Thus we must have (see (16))

$$\widehat{X} = \widehat{X}^T = \overline{V}\begin{bmatrix} 0 & 0 \\ 0 & G^T \end{bmatrix} U^T = -UD_1 \begin{bmatrix} 0 & 0 \\ 0 & G^T \end{bmatrix} D_2^H V^H.$$

Hence

$$\begin{bmatrix} 0 & 0 \\ 0 & G \end{bmatrix} = -D_1 \begin{bmatrix} 0 & 0 \\ 0 & G^T \end{bmatrix} D_2^H. \tag{18}$$

Therefore G has to fulfil $G = -D_{12} G^T D_{22}^H$. We take $G = \alpha D_{12}$, $\alpha \in \mathbb{R}$, $\alpha \neq 0$. Then G satisfies (18) because $-\alpha D_{12} D_{12}^T D_{22}^H = \alpha D_{12}$. We now verify that there exists $\alpha \neq 0$ such that $\|\Sigma_2 - G\|_\infty \le \sigma_1$. Namely, we can select α satisfying

$$\|\Sigma_2 - \alpha D_{12}\|_\infty \le \|\Sigma_2\|_\infty + |\alpha| \cdot \|D_{12}\|_\infty = \sigma_{s+1} + |\alpha| \le \sigma_1$$

since $\sigma_{s+1} < \sigma_1$. Thus we have proven that $A \in \mathcal{S}_g$ has a unique spectral approximation by symmetric complex matrices if and only if the singular values of A are all equal, i.e., A is a multiple of a matrix H, $A = \sigma H$, where σ is a positive number and H is unitary and skew-symmetric.

Case 3. Let f be as in (13) and let $A \in \mathcal{S}_g$. Now $\mathcal{S}_f$ is the set of all real matrices and $\mathcal{S}_g$ is the set of all purely imaginary matrices. Therefore the SVD of A has the form ($U = iQ$)

$$A = iQ\Sigma V^T,$$

where Q and V are real orthogonal. Thus every approximation $\widehat{X}$ of A by elements from $\mathcal{S}_f$ satisfies (see (16))

$$\widehat{X} = iQ \begin{bmatrix} 0 & 0 \\ 0 & G \end{bmatrix} V^T = -iQ \begin{bmatrix} 0 & 0 \\ 0 & \overline{G} \end{bmatrix} V^T. \tag{19}$$

We choose $G = i\operatorname{diag}(\mu_j)$, where $\mu_j \in \mathbb{R}$, $\mu_1 \ge \mu_2 \ge \ldots > 0$ and μ_1 satisfies

$$\|\Sigma_2 - G\|_\infty = \sigma_{s+1} + \mu_1 \le \sigma_1.$$

Then G satisfies (19). Thus we have shown that $A \in \mathcal{S}_g$ has a unique spectral approximation by real matrices if and only if A is a multiple of a matrix H, $A = \sigma H$, where σ is a positive number and H is a unitary and purely imaginary matrix.

The above conditions for uniqueness have a common form. Namely, we have the following corollary.

Corollary 5. *Let f be one of the functions* (11)–(13). *Then $A \in \mathcal{S}_g$ has a unique approximation, with respect to the spectral norm, by elements from $\mathcal{S}_f$ if and only if $A = \sigma H$, where σ is a positive number and $H \in \mathcal{S}_g$ is unitary.*

We recall that if $A \in \mathcal{S}_g$ then $A^{(f)} = 0$ is the approximation of A. Therefore Corollary 5 means in fact that the zero matrix is the unique approximation of $A \in \mathcal{S}_g$ if and only if the singular values of A are all equal.

Let $A \in \mathbb{C}^{n \times n}$ be arbitrary and let A_1, A_2 be determined as in (14). Corollary 5 implies that A has a unique spectral approximation by elements from $\mathcal{S}_f$ if and only if A_2 has all singular values equal. Of course, this unique best approximation is equal to A_1. Therefore we have the following corollary.

Corollary 6. *Let f be one of the functions* (11)–(13). *The zero matrix is the unique approximation to $A = A_1 + A_2$, $A_1 \in \mathcal{S}_f$, $A_2 \in \mathcal{S}_g$, by elements from $\mathcal{S}_f$, with respect to the spectral norm, if and only if $A_1 = 0$ and A_2 is a multiple of a unitary matrix.*

The above conditions for uniqueness are analogous to those in the last part of the previous section, where we consider approximation by positive semi-definite matrices. This is not surprising because the function

$$h(X) = (X^H X)^{1/2},$$

where we take the positive square root, is also singular values preserving, but not linear [see Li and Tsing (1987)]. If A is normal then the positive approximant $B^{(+)}$ of A has the form $B^{(+)} = \frac{1}{2}(A + h(A))$ (see (7), compare with $A^{(f)}$ given in (8)).

4. Moore–Penrose generalized inverse. Let $\mathcal{S}$ denote the set of all g-inverses of a complex $m \times n$ matrix A,

$$\mathcal{S} = \{X : AXA = A\}.$$

Let A have the following SVD:

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^H,$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $r = \operatorname{rank}(A)$ and σ_j are the nonzero singular values of A and U and V are unitary matrices. Then $X \in \mathcal{S}$ if and only if X has the form [see Rao (1973)]

$$X = V \begin{bmatrix} \Sigma^{-1} & L \\ M & N \end{bmatrix} U^H, \tag{20}$$

where L, M and N are arbitrary. If we take $L = 0, M = 0$ and $N = 0$ then

X is equal to the Moore–Penrose generalized inverse $A^\dagger$ [see Penrose (1955)]

$$A^\dagger = V \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^H.$$

The Moore–Penrose generalized inverse $A^\dagger$ is uniquely determined by the well known conditions [for another characterization see for example Fiedler and Markham (1993)]

$$AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (AA^\dagger)^H = AA^\dagger, \quad (A^\dagger A)^H = A^\dagger A.$$

The set $\mathcal{S}$ of all g-inverses of A is convex. Consider the following problem for an arbitrary unitarily invariant norm $\|\cdot\|$:

(21) $$\mu(A) = \min_{X\in\mathcal{S}} \|X\|.$$

By the pinching property of unitarily invariant norms we have for all $X \in \mathcal{S}$ as in (20) [see for example Stewart and Sun (1990), pp. 86–88]

(22) $$\|X\| \ge \|\mathrm{diag}(\Sigma^{-1}, N)\| \ge \|\mathrm{diag}(\Sigma^{-1}, 0)\| = \|A^\dagger\|.$$

Therefore

(23) $$\mu(A) = \|A^\dagger\|.$$

It is a well known classical result of von Neumann that every unitarily invariant norm corresponds to an apropppriate symmetric gauge function ϕ. A unitarily invariant norm associated with a symmetric gauge function ϕ is denoted by $\|\cdot\|_\phi$. Then

$$\|X\|_\phi = \phi(\sigma(X)),$$

where $\sigma(X)$ denotes the vector of ordered singular values $\sigma_j(X)$ of X. The unitarily invariant norm is *strictly convex* if and only if the symmetric gauge function ϕ is strictly convex [see for example Ziętak (1988)]. We recall that the c_p-norm is strictly convex for $1 < p < \infty$.

If the norm $\|\cdot\|$ is strictly convex then for $X \in \mathcal{S}$ we have

(24) $$\|X\| < \|A^-\| \quad \text{for } A^- \ne X,\ A^- \in \mathcal{S},$$

if and only if $X = A^\dagger$. The matrix $A^\dagger$ is distinguished by the condition (24) which holds for every strictly convex unitarily invariant norm.

The property (24) was proven by Maher (1990) for the c_p-norm with $1 < p < \infty$ [for the case $p = 2$ see Kalman (1976), Penrose (1956)]. Unfortunately, (24) is not true for the spectral norm. Let A^- be as in (20) with $L = 0, M = 0$ and N such that $\|N\|_\infty \le \|\Sigma^{-1}\|_\infty$. Then $\|A^-\|_\infty = \|A^\dagger\|_\infty$.

A symmetric gauge function ϕ is an *absolute norm* in $\mathbb{R}^q, q = \min\{m, n\}$, so it is *monotonic* [for the properties of absolute norms see Bauer *et al.* (1961)]. Therefore for $x, y \in \mathbb{R}^q$ we have

$$|x| \le |y| \quad \text{implies} \quad \phi(x) \le \phi(y).$$

We say that a symmetric gauge function ϕ is *strictly monotonic* if

$$0 \le x \le y,\ \phi(x) = \phi(y) \quad \text{implies} \quad x = y.$$

The l_p-norm, $1 < p < \infty$, is strictly monotonic. The trace norm $\|\cdot\|_1$ is not strictly convex, but the l_1-norm is strictly monotonic. That is the reason for which the condition (24) characterizes the Moore–Penrose generalized inverse also for the trace norm. This follows immediately from the following lemma [compare So (1990), Thompson (1972)].

LEMMA 7. *Let matrices $X, Y \in \mathbb{C}^{m\times n}$ have the block forms*

$$X = \begin{bmatrix} K_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} K_1 & K_2 \\ K_3 & K_4 \end{bmatrix},$$

where the block K_1 has order r. Then

$$\sigma_j(X) \le \sigma_j(Y) \quad \text{for } j = 1, \ldots, r, \tag{25}$$

and for every unitarily invariant norm $\|\cdot\|_\phi$ we have $\|X\|_\phi \le \|Y\|_\phi$. Moreover, if the symmetric gauge function ϕ is strictly monotonic then

$$\|X\|_\phi = \|Y\|_\phi \Leftrightarrow K_2, K_3, K_4 \text{ are zero blocks.} \ \blacksquare \tag{26}$$

Proof. The proof of (25) is given in Thompson (1972). Identity (26) is proven by So (1990) for the trace norm. We now verify that (26) holds for unitarily invariant norms associated with strictly monotonic symmetric gauge functions ϕ. Let $\|X\|_\phi = \|Y\|_\phi$. Then

$$\phi(\sigma(X)) = \phi(\sigma(Y)), \quad 0 \le \sigma(X) \le \sigma(Y),$$

which implies that $\sigma(X) = \sigma(Y)$ because ϕ is strictly monotonic. Therefore $\|X\|_1 = \|Y\|_1$. Thus $Y = X$ by the result of So for the trace norm. $\blacksquare$

If we apply the above considerations to the matrix (20) then we immediately deduce that $A^\dagger$ satisfies (24) for the unitarily invariant norms $\|\cdot\|_\phi$ with ϕ strictly monotonic, in particular for the trace norm. The inequalities (25) applied to $K_1 = \Sigma^{-1}$ and $Y = A^-$ imply that $A^\dagger$ is minimal with respect to an order which is stronger than the strict spectral ordering in $\mathcal{S}$ (see also Corollary 2 and (23)). Thus we have proven the following theorem characterizing $A^\dagger$.

THEOREM 8. *Let $X \in \mathcal{S}$. Then the following statements are equivalent*:

(a) $X = A^\dagger$,

(b) *X satisfies* (24) *for every strictly convex unitarily invariant norm*,

(c) *X satisfies* (24) *for every unitarily invariant norm $\|\cdot\|_\phi$ with ϕ strictly monotonic*, *in particular for the trace norm*,

(d) *the singular values of X satisfy $\sigma_j(X) \le \sigma_j(A^-)$ for all j and every $A^- \in \mathcal{S}$.*

The property (d) means that $A^\dagger$ is the strict spectral approximation to the zero matrix by matrices from $\mathcal{S}$. Therefore Theorem 8 shows that in some special cases a strict spectral approximation can satisfy a stronger condition than is assumed in its definition.

Remark. After this work was written, the author has learned that the inequality

$$\sigma_j(A^\dagger) \le \sigma_j(A^-) \quad \text{for all } j$$

was proven by Bapat and Ben-Israel in *Singular values and maximum rank minors of generalized inverses*, Linear and Multilinear Algebra 40 (1995), 153–161.

Acknowledgements. The author wishes to thank the referee for his remarks which helped to improve the presentation. She also thanks Professor J. Zemánek (Warsaw) for drawing her attention to the problem (21).

References

T. Ando, T. Sekiguchi and T. Suzuki (1973), *Approximation by positive operators*, Math. Z. 131, 273–282.

F. L. Bauer, J. Stoer and C. Witzgall (1961), *Absolute and monotonic norms*, Numer. Math. 3, 257–264.

R. Bhatia and F. Kittaneh (1992), *Approximation by positive operators*, Linear Algebra Appl. 161, 1–9.

R. Bouldin (1973), *Positive approximants*, Trans. Amer. Math. Soc. 177, 391–403.

C. Davis (1976), *An extremal problem for extensions of a sesquilinear form*, Linear Algebra Appl. 13, 91–102.

M. Fiedler and T. L. Markham (1993), *A characterization of the Moore–Penrose inverse*, Linear Algebra Appl. 179, 129–133.

P. E. Gill, W. Murray and M. H. Wright (1981), *Practical Optimization*, Academic Press, London.

G. H. Golub and C. Van Loan (1989), *Matrix Computations*, J. Hopkins Univ. Press, Baltimore.

P. R. Halmos (1972), *Positive approximants of operators*, Indiana Univ. Math. J. 21, 951–960.

N. J. Higham (1989), *Matrix nearness problems and applications*, in: *Application of Matrix Theory*, M. J. C. Gover and S. Barnett (eds.), Oxford Univ. Pres, New York, 1–27.

R. A. Horn and Ch. R. Johnson (1986), *Matrix Analysis*, Cambridge Univ. Press, Cambridge.

R. Huotari and W. Li (1994), *Continuity of metric projection*, *Pólya algorithm*, *strict best approximation*, *and tubularity of convex sets*, J. Math. Anal. Appl. 182, 836–856.

R. E. Kalman (1976), *Algebraic aspects of the generalized inverse of a rectangular matrix*, in: *Generalized Inverses and Applications*, M. Z. Nashed (ed.), Academic Press, New York, 111–124.

C.-K. Li and N.-K. Tsing (1987), *On the unitarily invariant norms and some related results*, Linear and Multilinear Algebra 20, 107–119.

P. J. Maher (1990), *Some operator inequalities concerning generalized inverses*, Illinois J. Math. 34, 503–514.
R. Penrose (1955), *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc. 51, 406–413.
R. Penrose (1956), *On best approximate solutions of linear matrix equations*, ibid. 52, 17–19.
C. R. Rao (1973), *Linear Statistical Inference and Its Applications*, Wiley, New York.
J. R. Rice (1962), *Tchebycheff approximation in a compact metric space*, Bull. Amer. Math. Soc. 68, 405–410.
D. D. Rogers and J. D. Ward (1981), *C_p-minimal positive approximants*, Acta Sci. Math. (Szeged) 43, 109–115.
E. M. de Sá (1994), *Faces of the unit ball of a unitarily invariant norm*, Linear Algebra Appl. 197/198, 451–493.
W. So (1990), *Facial structures of Schatten p-norms*, Linear and Multilinear Algebra 27, 207–212.
G. W. Stewart and J.-G. Sun (1990), *Matrix Perturbation Theory*, Academic Press, Boston.
R. C. Thompson (1972), *Principal submatrices IX: Interlacing inequalities for singular values of submatrices*, Linear Algebra Appl. 5, 1–12.
H. J. Woerdeman (1994), *Superoptimal completions of triangular matrices*, Integral Equations Operator Theory 20, 492–501.
N. J. Young (1986), *The Nevanlinna–Pick problem for matrix-valued functions*, J. Operator Theory 15, 239–269.
K. Ziętak (1988), *On characterization of the extremal points of the unit sphere of matrices*, Linear Algebra Appl. 106, 57–75.
K. Ziętak (1993), *Properties of linear approximations of matrices in the spectral norm*, ibid. 183, 41–60.
K. Ziętak (1995), *Strict approximation of matrices*, SIAM J. Matrix Anal. Appl. 16, 232–234.

Krystyna Ziętak
Institute of Computer Science
University of Wrocław
ul. Przesmyckiego 20
51-151 Wrocław, Poland
E-mail: zietak@ii.uni.wroc.pl

Received on 29.11.1995;
revised version on 11.9.1996