Matrices and n-Widths

Allan Pinkus
Department of Mathematics
Technion, I.I.T.
Haifa, Israel

Submitted by Hans Schneider

ABSTRACT

This paper is concerned with a collection of ideas and problems in approximation theory which lead to some solved and unsolved problems in matrix theory. As an example, consider the problem of approximating the identity matrix by matrices of fixed rank where the norm is taken to be the maximum of the absolute value of the elements of the matrix. This problem is unsolved.

1. INTRODUCTION

Some of the easy-to-state and yet difficult problems of approximation theory are concerned with matrices. This paper is intended as a review of a collection of ideas and problems which arise in approximation theory and which lead to some interesting solved and unsolved problems in matrix theory.

We start with the statement of the problems from an approximation theorist's viewpoint. The linear algebraist is encouraged to bear with us, as we shall shortly specialize the problems concerned.

One of the central problems of approximation theory and the one around which this paper turns is the question of the extent to which a given class of functions may be approximated by n-dimensional subspaces. This idea was introduced and formalized by Kolmogorov [11] in 1936, when he defined the concept of n-widths. Specifically, let X be a normed linear space and $\mathfrak{A}$ a subset of X. The n-width of $\mathfrak{A}$ with respect to X, in the sense of Kolmogorov, is defined as

$$d_n(\mathfrak{A};X) = \inf_{X_n} \sup_{x\in\mathfrak{A}} \inf_{y\in X_n} \|x-y\|_X.$$

 0024-3795/79/050245+34$01.75

In other words, we first measure the distance of X_n to a given $x \in \mathfrak{A}$, then determine the distance of X_n to $\mathfrak{A}$ in terms of the supremum of the distance of X_n to each $x \in \mathfrak{A}$, and finally attempt to minimize the above quantity by taking the infimum over all n-dimensional subspaces X_n of X. If there exists an n-dimensional subspace X_n^* for which $d_n(\mathfrak{A}, X)$ is attained, i.e.,

$$d_n(\mathfrak{A}; X) = \sup_{x \in \mathfrak{A}} \inf_{y \in X_n^*} \|x - y\|_X,$$

then X_n^* is said to be optimal.

We wish of course to determine the value d_n and find optimal subspaces. This problem is in general intractable. We must specialize to some extent if we expect to obtain any concrete results.

A customary choice of the subset $\mathfrak{A}$ of X is

$$\mathfrak{A} = \{Ax : \|x\|_Y \leqslant 1\},$$

where A is some compact linear map of Y into X, and Y is also a normed linear space. While this problem is solvable for certain choices of X, Y, and A, much yet remains to be done; see e.g. Tichomirov [28] and Micchelli and Pinkus [16].

Let us further specialize by assuming that A is a real $m \times k$ matrix and X and Y are R^m and R^k, respectively, equipped with one of the l_p norms, $1 \leqslant p \leqslant \infty$, but not necessarily the same l_p norm for both X and Y. Thus we are interested in

$$d_n(A; l_p^k; l_q^m) = \min_{X_n} \max_{\|x\|_p \leqslant 1} \min_{y \in X_n} \|Ax - y\|_q.$$

(In this case the sup and inf are attained and are therefore replaced by max and min.) Again this is too difficult a general problem. The only case in which at present a solution is always obtainable is when $p = q = 2$. In this case $d_n(A; l_2^k; l_2^m)$ is the $n+1$st singular value of A.

THEOREM 1.1. *Let A be a real $m \times k$ matrix. Let $\lambda_1 \geqslant \cdots \geqslant \lambda_m \geqslant 0$ denote the m eigenvalues of the positive semidefinite matrix AA^T, and let $x^1, \dots, x^m$ denote a corresponding set of orthonormal eigenvectors. Then*

$$d_n(A; l_2^k; l_2^m) = \begin{cases} \lambda_{n+1}^{1/2}, & n < m, \\ 0, & n \geqslant m. \end{cases}$$

Furthermore if $n < m$, then an optimal subspace for d_n is $\operatorname{span}\{x^1,\dots,x^n\}$.

Proof. It is a standard duality result (essentially Hölder's inequality and equality therein) that

$$\min_{y\in X_n}\|Ax-y\|_q=\max_{\substack{z\perp X_n\\ \|z\|_{q'}\le 1}}(z,Ax),$$

where $1/q+1/q'=1$ and $z\perp X_n$ means that $(z,y)=0$ for all $y\in X_n$. Thus

$$\begin{aligned}
d_n(A;l_2^k;l_2^m)&=\min_{X_n}\max_{\|x\|_2\le 1}\min_{y\in X_n}\|Ax-y\|_2\\
&=\min_{X_n}\max_{\|x\|_2\le 1}\max_{\substack{z\perp X_n\\ \|z\|_2\le 1}}(z,Ax)\\
&=\min_{X_n}\max_{\substack{z\perp X_n\\ \|z\|_2\le 1}}\max_{\|x\|_2\le 1}(A^Tz,x)\\
&=\min_{X_n}\max_{\substack{z\perp X_n\\ \|z\|_2\le 1}}\|A^Tz\|_2\\
&=\min_{X_n}\max_{z\perp X_n}\left[\frac{(A^Tz,A^Tz)}{(z,z)}\right]^{1/2}\\
&=\left[\min_{y^1,\dots,y^n}\max_{\substack{(z,y^i)=0\\ i=1,\dots,n}}\frac{(AA^Tz,z)}{(z,z)}\right]^{1/2}.
\end{aligned}$$

That the last expression is the $n+1$st eigenvalue of AA^T is the content of the Courant-Fischer min-max theorem. The theorem follows. ■

The space l_2^m is of course special. It is a Hilbert space, and among other properties of a Hilbert space, one of the characterizing properties is that the best approximation operator is a linear (projection) operator. On this basis, it is easily seen that

$$d_n(A;l_2^k;l_2^m)=\min_{\operatorname{rank}P\le n}\max_{\|x\|_2\le 1}\|Ax-Px\|_2,$$

where the minimum is taken over all $m\times k$ real matrices P whose rank is at most n. In this form the results of the above Theorem 1.1 are known, since the theorem then simply says that the singular value decomposition provides a best rank n approximation to a given matrix in the induced (l_2^k, l_2^m) norm; see e.g. Gohberg and Krein [4, p. 28]. This same rank n matrix is optimal if we use the matrix (Frobenius) norm $\|A\|_2 = [\sum_{i=1}^m \sum_{j=1}^k |a_{ij}|^2]^{1/2}$ in place of the above induced operator norm. This result, which seems to be due to Eckart and Young [3], is just a matrix version of the corresponding result proven for integral operators by E. Schmidt [19].

The n-width, in the sense of Kolmogorov, is only the first in a series of n-widths which have been introduced over the past forty years. In the course of this article only four n-widths will be discussed. The reader who wishes to pursue these matters is referred to Tichomirov [28], Pietsch [18], and the references therein. One of the more interesting and important of these n-widths is what we shall here call the linear n-width. (It is sometimes referred to, especially by those in Banach space theory, as the approximation number.) We define it by

$$\delta_n(A; l_p^k; l_q^m) = \min_{\operatorname{rank} P \leqslant n} \max_{\|x\|_p \leqslant 1} \|Ax - Px\|_q,$$

where, as above, the minimum is taken over all $m \times k$ matrices of rank at most n. Since δ_n is a measure of the best linear approximation from an n-dimensional subspace (the range of P), it is obvious that $d_n \leqslant \delta_n$. (A nonobvious result is that $\delta_n \leqslant C\sqrt{n}\, d_n$ for some constant C, independent of n; see e.g. Ha [5], Hutton, Morrell, and Retherford [6], or Pietsch [18].) One is also interested in knowing if $d_n = \delta_n$, i.e., if linear methods suffice.

As was previously mentioned, only for the case $p = q = 2$ is there a complete solution. One other known case, which will not be dealt with in this paper (its proof may be found in Micchelli and Pinkus [14]), is where $p = q = \infty$ and A is a totally positive (T.P.) matrix. A matrix is T.P. if all its minors are nonnegative. In this case the n-widths d_n and δ_n are equal, and furthermore, there exist n column vectors of the matrix A whose span is an optimal subspace for d_n. This result is very much connected with the T.P. property of the matrix. If A is T.P. and $p = q = 1$, then d_n and δ_n are also equal and known. The statement of this result is slightly more complicated; see Micchelli and Pinkus [15].

In general, there is very little else which may be said concerning n-widths of arbitrary matrices. Of course, the 0-width is easily obtained as

$$\delta_0(A; l_p^k; l_q^m) = d_0(A; l_p^k; l_q^m) = \max_{x \neq 0} \frac{\|Ax\|_q}{\|x\|_p}.$$

While the above result is trivial, there is a correspondingly simple statement for the $(k-1)$-width whose proof is not trivial, namely

$$\delta_{k-1}(A;l_p^k;l_q^m) = d_{k-1}(A;l_p^k;l_q^m) = \min_{x\neq 0} \frac{\|Ax\|_q}{\|x\|_p}.$$

We shall not prove this result here. Its proof may be found in Brown [2] and Micchelli and Pinkus [17]. For the particular cases dealt with here, the result will be reproven.

In this paper we discuss the situation wherein A is a diagonal matrix. Since the matrix A is henceforth taken to be diagonal, we refer to A as D and without loss of generality, assume that D is an $m\times m$ diagonal matrix with diagonal entries $(D_1,\ldots,D_m)$, and $D_1 \geqslant D_2 \geqslant \cdots \geqslant D_m > 0$. It is sufficient, by the symmetry of the l_p norm, to consider only diagonal matrices of the above form.

What is initially surprising is that the n-widths $d_m(D;l_p^m;l_q^m)$ and $\delta_n(D;l_p^m;l_q^m)$ are unknown for many $p,q\in[1,\infty]$. As an example of a seemingly simple problem which is unsolved, let $D=I$, $p=1$, and $q=\infty$. It is an easy matter to check that

$$d_n(I;l_1^m;l_\infty^m) = \delta_n(I;l_1^m;l_\infty^m) = \min_{\operatorname{rank} P < n}\ \max_{i,j=1,\ldots,m} |\delta_{ij} - p_{ij}|,$$

where δ_{ij} are the elements of I and $P=(p_{ij})_{i,j=1}^m$. Except for the cases $n=1$ and $n=m-1$, the above quantities are not known. What is also interesting is that this problem has a different solution if $p_{ij}\in R$, $i,j=1,\ldots,m$ or if the p_{ij}'s are permitted to be in C. This latter result is rather surprising, since in the cases where d_n and δ_n are known, there is no difference in the result if we consider C^m rather than R^m. This example is discussed in some detail in Sec. 5.

Before discussing the organization of the paper, let us note various topics which have not been included in this work. Firstly, we have not considered this subject from the point of view of Banach space theory, nor have we discussed the case where D is an infinite diagonal matrix. Results of these types may be found in Ha [5], Hutton, Morrell, and Retherford [6], Pietsch [18], and Sofman [21]. Secondly, we have limited ourselves to a consideration of a very few of the n-width concepts, and in particular, to the more analytic ones. Results have been obtained concerning the Aleksandrov and Urysohn n-widths of finite diagonal matrices. These n-widths are more geometric and topological in nature, and it was felt that a discussion of these quantities was best deferred. The interested reader is here referred to Stesin [24, 25] and Tichomirov [28].

The organization of this paper runs as follows. Section 2 is concerned with the introduction of the two remaining n-widths, namely the n-widths in the sense of Gel'fand (g_n) and of Bernstein (b_n), and the relations among the four n-widths. In Sec. 3, we consider the n-widths d_n, δ_n, and g_n for $p \geqslant q$, show that they are all equal, obtain their common value, and identify an optimal subspace which here is simply the span of the first n columns of D. In Sec. 4 we discuss the case $p<q$. Only for $p=1$, $q=2$ is d_n known explicitly for all n. Section 5 is concerned with a further consideration of the n-widths when $D=I$, the identity matrix. In particular, we return to the case $p=1$, $q=\infty$ and obtain some lower and upper bounds on $d_n(I;l_1^m;l_\infty^m)$.

2. PRELIMINARIES

Let $D=\text{diag}\,(D_1,\ldots,D_m)$ denote the real $m\times m$ diagonal matrix with diagonal entries $D_1, D_2,\ldots,D_m$ where $D_1 \geqslant D_2 \geqslant \cdots \geqslant D_m > 0$. As usual, for $x\in R^m$ we set

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^{m} |x_i|^p\right)^{1/p}, & 1\leqslant p<\infty \\ \max_{i=1,\ldots,m} |x_i|, & p=\infty \end{cases}.$$

This section deals with the relationships between the various n-widths. An attempt has been made to properly reference, where possible, the results of this section. Originally these results were of a general nature. They have here been specialized to our particular situation.

Before introducing the additional n-widths, let us note that δ_n possesses the following simple duality property.

LEMMA 2.1 (Ismagilov [8]; Hutton, Morrell, and Retherford [6]). *For* $p,q\in[1,\infty]$,

$$\delta_n(D;l_p^m;l_q^m) = \delta_n(D;l_{q'}^m;l_{p'}^m),$$

where

$$\frac{1}{p}+\frac{1}{p'}=\frac{1}{q}+\frac{1}{q'}=1.$$

Proof. Since $D = D^T$,

$$\begin{aligned}
\delta_n(D; l_p^m; l_q^m) &= \min_{\operatorname{rank} P < n} \max_{\|x\|_p \le 1} \|(D-P)x\|_q \\
&= \min_{\operatorname{rank} P < n} \max_{\|x\|_p \le 1} \max_{\|y\|_{q'} \le 1} ((D-P)x, y) \\
&= \min_{\operatorname{rank} P < n} \max_{\|y\|_{q'} \le 1} \max_{\|x\|_p \le 1} (x, (D-P^T)y) \\
&= \min_{\operatorname{rank} P < n} \max_{\|y\|_{q'} \le 1} \|(D-P)y\|_{p'} \\
&= \delta_n(D; l_{q'}^m; l_{p'}^m). \qquad \blacksquare
\end{aligned}$$

In Sec. 5, we shall give an example which shows that d_n does not exhibit the above duality property, i.e., there exist $p, q \in [1, \infty]$ and diagonal D for which $d_n(D; l_p^m; l_q^m) \neq d_n(D; l_{q'}^m; l_{p'}^m)$. The duality relationship which does however exist is the motivation for this next definition.

In the literature of n-widths, one additional n-width concept is invariably mentioned, and this is the n-width in the sense of Gel'fand. It is defined as follows.

Let X be a normed linear space and $\mathfrak{A}$ a subset of X. The n-width in the sense of Gel'fand is given by

$$g_n(\mathfrak{A}; X) = \inf_{L^n} \sup_{x \in \mathfrak{A} \cap L^n} \|x\|_X,$$

where the infimum is taken over all subspaces L^n of X of codimension at most n. In the particular case studied in this paper it may be shown that

$$g_n(D; l_p^m; l_q^m) = \min_{X_n} \max_{\substack{x \perp X_n \\ \|x\|_p \le 1}} \|Dx\|_q,$$

where X_n is any n-dimensional subspace of R^m, and $x \perp X_n$ means that $(x, y) = 0$ for all $y \in X_n$. As indicated above, the following result is valid.

LEMMA 2.2 (Ioffe and Tichomirov [7]). *For* $p, q \in [1, \infty]$

$$d_n(D; l_p^m; l_q^m) = g_n(D; l_{q'}^m; l_{p'}^m)$$

where

$$\frac{1}{p}+\frac{1}{p'}=\frac{1}{q}+\frac{1}{q'}=1.$$

Proof. The proof is based on the duality statement

$$\min_{y\in X_n}\|x-y\|_q=\max_{\substack{z\perp X_n\\ \|z\|_{q'}\leqslant 1}}(z,x)$$

as well as on the fact that D is symmetric:

$$\begin{aligned}
d_n(D;l_p^m;l_q^m) &= \min_{X_n}\max_{\|x\|_p\leqslant 1}\min_{y\in X_n}\|Dx-y\|_q\\
&= \min_{X_n}\max_{\|x\|_p\leqslant 1}\max_{\substack{z\perp X_n\\ \|z\|_{q'}\leqslant 1}}(z,Dx)\\
&= \min_{X_n}\max_{\substack{z\perp X_n\\ \|z\|_{q'}\leqslant 1}}\max_{\|x\|_p\leqslant 1}(Dz,x)\\
&= \min_{X_n}\max_{\substack{z\perp X_n\\ \|z\|_{q'}\leqslant 1}}\|Dz\|_{p'}\\
&= g_n(D;l_{q'}^m;l_{p'}^m).
\end{aligned}$$

■

It is often the case that n-widths are evaluated by obtaining upper and lower bounds which are the same. Upper bounds are obtained by determining the distance of the class from a judiciously chosen subspace. Lower bounds are not so easy to come by. Various techniques have been developed to deal with this problem. One such method is the evaluation of a lower bound known as the n-width in the sense of Bernstein, denoted by b_n, which is defined for diagonal matrices by

$$b_n(D;l_p^m;l_q^m)=\max_{X_{n+1}}\min_{\substack{\|x\|_p=1\\ x\in X_{n+1}}}\|Dx\|_q,$$

where X_{n+1} is any $(n+1)$-dimensional subspace of R^m. Note that we maximize over such subspaces. The Bernstein n-width has the obvious property that $b_n(I;l_p^m;l_p^m)=1$ for any $p\in[1,\infty]$ and any $n=0,1,\ldots,m-1$.

A natural ordering among the various n-widths is contained in the following proposition.

PROPOSITION 2.1. *For* $p,q\in[1,\infty]$

$$\delta_n(D;l_p^m;l_q^m) \geqslant d_n(D;l_p^m;l_q^m), g_n(D;l_p^m;l_q^m) \geqslant b_n(D;l_p^m;l_q^m).$$

Proof. The proof of the fact that $\delta_n(D;l_p^m;l_q^m)\geqslant d_n(D;l_p^m;l_q^m)$ is a direct result of the definitions of δ_n and d_n. The proof of $\delta_n(D;l_p^m;l_q^m)\geqslant g_n(D;l_p^m;l_q^m)$ follows from the above result and Lemmas 2.1 and 2.2. It may also be proven directly.

Let X_{n+1} be any $(n+1)$-dimensional (not n-dimensional) subspace of R^m. Then

$$\begin{aligned} g_n(D;l_p^m;l_q^m) &= \min_{X_n} \max_{\substack{x\perp X_n\\ \|x\|_p\leqslant 1}} \|Dx\|_q \\ &\geqslant \min_{X_n} \max_{\substack{x\perp X_n,\, x\in X_{n+1}\\ \|x\|_p\leqslant 1}} \|Dx\|_q. \end{aligned}$$

Since X_n and X_{n+1} are of dimension n and $n+1$, respectively, there exists an $x\in X_{n+1}$, $x\neq 0$, for which $x\perp X_n$. Thus,

$$\begin{aligned} g_n(D;l_p^m;l_q^m) &\geqslant \min_{X_n} \min_{\substack{\|x\|_p=1\\ x\in X_{n+1}}} \|Dx\|_q \\ &= \min_{\substack{\|x\|_p=1\\ x\in X_{n+1}}} \|Dx\|_q. \end{aligned}$$

Since this is true for any $(n+1)$-dimensional subspace, we have $g_n(D;l_p^m;l_q^m)$ $\geqslant b_n(D;l_p^m;l_q^m)$.

Now,

$$\begin{aligned} d_n(D;l_p^m;l_q^m) &= \min_{X_n} \max_{\|x\|_p\leqslant 1} \min_{y\in X_n} \|Dx-y\|_q \\ &\geqslant \min_{X_n} \max_{\substack{\|x\|_p\leqslant 1\\ x\in X_{n+1}}} \min_{y\in X_n} \|Dx-y\|_q \end{aligned}$$

for any given $(n+1)$-dimensional subspace X_{n+1}. We shall use the fact, which will be proven below, that there exists an $x \in X_{n+1}$, $x \neq 0$, for which the best approximation to Dx from X_n in l_q^m is the zero vector, i.e.,

$$\min_{y \in X_n} \|Dx - y\|_q = \|Dx\|_q.$$

On the assumption that this result is correct,

$$d_n(D; l_p^m; l_q^m) \geqslant \min_{X_n} \min_{\substack{\|x\|_p = 1 \\ x \in X_{n+1}}} \|Dx\|_q$$

$$= \min_{\substack{\|x\|_p = 1 \\ x \in X_{n+1}}} \|Dx\|_q.$$

Since this is valid for any X_{n+1}, we have $d_n(D; l_p^m; l_q^m) \geqslant b_n(D; l_p^m; l_q^m)$.

The crux of the proof of the lemma is contained in the fact which we shall now prove and which is due to Krein, Krasnosel'ski, and Milman [12]. Let us assume for the moment that $q \in (1, \infty)$. Then the best approximation to each Dx from X_n in l_q^m is unique. Let $\Gamma x \in X_n$ denote the element of best approximation to Dx in l_q^m, i.e.,

$$\min_{y \in X_n} \|Dx - y\|_q = \|Dx - \Gamma x\|_q.$$

It follows from the uniqueness that Γx is a continuous function of x. Furthermore, Γ is odd, i.e., $\Gamma(-x) = -\Gamma x$. Restricting the map Γ to $X_{n+1} \cap \{x : \|x\|_p = 1\}$, we have a continuous odd map from the surface of an $(n+1)$-sphere into an n-dimensional space. Thus by the Borsuk Antipodensatz [1], there exists an $x \in X_{n+1} \cap \{x : \|x\|_p = 1\}$ for which $\Gamma x = 0$. To obtain this same result for $q = 1$ and $q = \infty$, we slightly perturb the respective norms to obtain norms which are strictly convex. On these new norms we have the continuity of Γ, so that the result then holds for these perturbed norms. We now limit back. ■

Aside from the above inequalities, there are certain situations in which equalities naturally occur. We formulate these equalities only between δ_n and d_n, although we could also use Lemmas 2.1 and 2.2 to formulate the corresponding equalities for δ_n and g_n.

LEMMA 2.3. *For* $1 \leqslant p \leqslant \infty$

$$\delta_n(D; l_p^m; l_2^m) = d_n(D; l_p^m; l_2^m).$$

Proof. As noted in the introduction, this is the statement that the best approximation operator in a Hilbert space is a linear (projection) operator. ■

LEMMA 2.4 (Hutton, Morrell, and Retherford [6]). *For* $1 \leqslant q \leqslant \infty$

$$\delta_n(D;l_1^m;l_q^m) = d_n(D;l_1^m;l_q^m).$$

Proof. Let e^j denote the jth unit vector in R^m. Since $\{\pm e^j\}_{j=1}^m$ are the extreme points of the unit ball in l_1^m, we have

$$d_n(D;l_1^m;l_q^m) = \min_{X_n} \max_{\|x\|_1 \leqslant 1} \min_{y \in X_n} \|Dx - y\|_q$$

$$= \min_{X_n} \max_{j=1,\ldots,m} \min_{y \in X_n} \|D_j e^j - y\|_q.$$

Let $y^j \in X_n$ denote a best approximation to $D_j e^j$ from X_n in l_q^m, i.e.,

$$\min_{y \in X_n} \|D_j e^j - y\|_q = \|D_j e^j - y^j\|_q.$$

Let P_n be the $m \times m$ matrix whose jth column vector is y^j, $j=1,\ldots,m$. Thus $\operatorname{rank} P_n \leqslant \dim X_n \leqslant n$. Furthermore,

$$\max_{\|x\|_1 \leqslant 1} \|Dx - P_n x\|_q = \max_{j=1,\ldots,m} \|(D-P_n)e^j\|_q$$

$$= \max_{j=1,\ldots,m} \|D_j e^j - y^j\|_q.$$

It now easily follows that $d_n(D;l_1^m;l_q^m) \geqslant \delta_n(D;l_1^m;l_q^m)$, which, with Proposition 2.1, proves the lemma. ■

The following interesting result can only hold in a matrix setting.

LEMMA 2.5. *For* $p,q \in [1,\infty]$,

$$b_n(D;l_p^m;l_q^m) \cdot d_{m-n-1}(D^{-1};l_{p'}^m;l_{q'}^m) = 1,$$

$n=0,1,\ldots,m-1$, *where* $1/p + 1/p' = 1/q + 1/q' = 1$.

Proof. From Lemma 2.2,

$$\begin{aligned} d_{m-n-1}(D^{-1}; l_{p'}^m; l_{q'}^m) &= g_{m-n-1}(D^{-1}; l_q^m; l_p^m) \\ &= \min_{X_{m-n-1}} \max_{\substack{x \perp X_{m-n-1} \\ \|x\|_q \leq 1}} \|D^{-1}x\|_p \\ &= \min_{X_{m-n-1}} \max_{\substack{x \perp X_{m-n-1} \\ x \neq 0}} \left[\frac{\|x\|_q}{\|D^{-1}x\|_p} \right]^{-1}. \end{aligned}$$

Setting $y = D^{-1}x$, and since D is invertible, this is

$$\begin{aligned} &= \min_{X_{m-n-1}} \max_{\substack{y \perp X_{m-n-1} \\ y \neq 0}} \left[\frac{\|Dy\|_q}{\|y\|_p} \right]^{-1} \\ &= \left[\max_{X_{m-n-1}} \min_{\substack{y \perp X_{m-n-1} \\ y \neq 0}} \frac{\|Dy\|_q}{\|y\|_p} \right]^{-1} \\ &= \left[\max_{X_{n+1}} \min_{\substack{y \in X_{n+1} \\ \|y\|_p = 1}} \|Dy\|_q \right]^{-1} \\ &= \left[b_n(D; l_p^m; l_q^m) \right]^{-1}. \end{aligned}$$

■

Corollary 2.1. *For $p, q \in [1, \infty]$,*

$$d_n(D; l_p^m; l_q^m) d_{m-n-1}(D^{-1}; l_{p'}^m; l_{q'}^m) \geq 1,$$

and

$$d_n(D; l_p^m; l_q^m) d_{m-n-1}(D^{-1}; l_q^m; l_p^m) \geq 1.$$

Proof. Proposition 2.1 and Lemmas 2.2 and 2.5. ■

3. THE CASE $p \geq q$

Undoubtedly the easiest determination of the n-widths introduced above is in the case $p = q$. Because the proof is so simple, we include it below as a separate result although we shall also obtain it as a particular case of the more general result for $p \geq q$.

THEOREM 3.1. *For* $1 \leqslant p \leqslant \infty$,

$$\delta_n(D; l_p^m; l_p^m) = d_n(D; l_p^m; l_p^m) = g_n(D; l_p^m; l_p^m)$$
$$= b_n(D; l_p^m; l_p^m) = D_{n+1}$$

for $n = 0, 1, \dots, m-1$. *Furthermore, an optimal subspace for* d_n *is the span of the first* n *unit vectors, and* $P_n = \text{diag}(D_1, \dots, D_n, 0, \dots, 0)$ *is an optimal best rank* n *matrix.*

Proof. Since we have shown that $\delta_n \geqslant d_n$, $g_n \geqslant b_n$, it suffices to prove that $\delta_n(D; l_p^m; l_p^m) \leqslant D_{n+1}$ and $b_n(D; l_p^m; l_p^m) \geqslant D_{n+1}$. The proof of these facts is simple. Let

$$P_n = \begin{bmatrix} D_1 & & & & & \\ & \ddots & & 0 & & \\ & & D_n & & & \\ & 0 & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix},$$

i.e., P_n is the $m \times m$ diagonal matrix with diagonal entries $(D_1, \dots, D_n, 0, \dots, 0)$. Thus,

$$\delta_n(D; l_p^m; l_p^m) \leqslant \max_{\|x\|_p \leqslant 1} \|(D - P_n)x\|_p$$

$$= \max_{x \neq 0} \frac{\left(\sum_{k=n+1}^{m} |D_n x_n|^p \right)^{1/p}}{\left(\sum_{k=1}^{m} |x_n|^p \right)^{1/p}}$$

$$\leqslant D_{n+1} \max_{x \neq 0} \frac{\left(\sum_{k=n+1}^{m} |x_n|^p \right)^{1/p}}{\left(\sum_{k=1}^{m} |x_n|^p \right)^{1/p}}$$

$$= D_{n+1},$$

since $D_{n+1} \geqslant \cdots \geqslant D_m > 0$. To prove that $b_n(D; l_p^m; l_p^m) \geqslant D_{n+1}$, it suffices to recall that $b_n(D; l_p^m; l_p^m) = [d_{m-n-1}(D^{-1}; l_{p'}^m; l_{p'}^m)]^{-1}$ and to note that the above argument has also given us an upper bound on $d_{m-n-1}(D^{-1}; l_{p'}^m; l_{p'}^m)$. The inequality $b_n(D; l_p^m; l_p^m) \geqslant D_{n+1}$ may also be proven directly by noting that

$$b_n(D; l_p^m; l_p^m) \geqslant b_n(D'; l_p^m; l_p^m) = b_n(D''; l_p^{n+1}; l_p^{n+1}) = D_{n+1},$$

where $D' = \mathrm{diag}(D_{n+1}, \ldots, D_{n+1}, 0, \ldots, 0)$ and D'' is the $(n+1)\times(n+1)$ matrix of the form $D'' = \mathrm{diag}(D_{n+1}, \ldots, D_{n+1})$. ∎

The fact that $P_n = \mathrm{diag}(D_1, \ldots, D_n, 0, \ldots, 0)$ is the optimal rank n approximation to D is hardly surprising. After all, what else could be? In this next theorem we prove that if $p > q$, then P_n remains optimal for δ_n, d_n, and g_n. Unfortunately this result is no longer valid if $p < q$.

THEOREM 3.2. *Given* $1 \leqslant q \leqslant p \leqslant \infty$. *Let* $1/r = 1/q - 1/p$. *Then*

$$\delta_n(D; l_p^m; l_q^m) = d_n(D; l_p^m; l_q^m) = g_n(D; l_p^m; l_q^m) = \left(\sum_{k=n+1}^{m} D_k^r \right)^{1/r}.$$

An optimal subspace for d_n *is* $\mathrm{span}\{e^i\}_{i=1}^n$, *and* $P_n = \mathrm{diag}(D_1, \ldots, D_n, 0, \ldots, 0)$ *is an optimal rank* n *matrix.*

This is a fairly recent theorem proven independently by Stesin [25] and Pietsch [18]. The proof given herein is that of Pietsch. Stesin's proof uses the Aleksandrov n-widths and is rather laborious.

In the proof of the theorem we use the following two lemmas.

LEMMA 3.1. *Let* X_n *be any* n*-dimensional subspace in* R^m, $n < m$. *There exists an* $x \in R^m$ *for which* $\|x\|_\infty = 1$, $x \perp X_n$, *and at most* n *components of* x *are not equal to* 1 *in absolute value.*

Proof. Let $E = \{x : \|x\|_\infty \leqslant 1,\ x \perp X_n\}$. E is a closed, convex, nonempty set in R^m and hence has extreme points. Let x^* be an extreme point of E. Assume that x^* does not satisfy the hypothesis of the lemma. Thus there exist $n+1$ distinct integers $\{i_k\}_{k=1}^{n+1}$ in $\{1, \ldots, m\}$ for which $|(x^*)_{i_k}| < 1$, $k = 1, \ldots, n+1$. Let e^i denote the ith unit vector in R^m, and let $G = \mathrm{span}\{e^{i_1}, \ldots, e^{i_{n+1}}\}$. There exists a $g \in G$ for which $g \perp X_n$, $g \neq 0$. Thus for ε sufficiently small, $x^* \pm \varepsilon g \in E$, contradicting the extreme point property of x^*. This proves the lemma. ∎

LEMMA 3.2. *Let* $1 \leqslant s < r \leqslant \infty$, $a_1, a_2, \dots, a_{k+1} > 0$, *and* $b_j \geqslant b_{k+1} > 0$, $j = 1, \dots, k$. *Then*

$$\frac{\left(\sum_{j=1}^{k+1} b_j^s a_j\right)^{1/s}}{\left(\sum_{j=1}^{k+1} b_j^r a_j\right)^{1/r}} \geqslant \frac{\left(\sum_{j=1}^{k} b_j^s a_j\right)^{1/s}}{\left(\sum_{j=1}^{k} b_j^r a_j\right)^{1/r}}.$$

Proof. Let $\alpha^r = \sum_{j=1}^{k} b_j^r a_j$ and $\beta^s = \sum_{j=1}^{k} b_j^s a_j$. Since $b_j \geqslant b_{k+1} > 0$, $j = 1, \dots, k$, and $s < r$,

$$\left(\frac{b_j}{b_{k+1}}\right)^s \leqslant \left(\frac{b_j}{b_{k+1}}\right)^r, \qquad j = 1, \dots, k,$$

and thus

$$\frac{\sum_{j=1}^{k} b_j^s a_j}{b_{k+1}^s} \leqslant \frac{\sum_{j=1}^{k} b_j^r a_j}{b_{k+1}^r}.$$

Hence $b_{k+1}^s / \beta^s \geqslant b_{k+1}^r / \alpha^r$, and

$$\left[1 + \frac{b_{k+1}^s}{\beta^s} a_{k+1}\right]^{r/s} \geqslant \left[1 + \frac{b_{k+1}^s}{\beta^s} a_{k+1}\right] \geqslant \left[1 + \frac{b_{k+1}^r}{\alpha^r} a_{k+1}\right].$$

Therefore,

$$\frac{\left(\sum_{j=1}^{k+1} b_j^s a_j\right)^{1/s}}{\left(\sum_{j=1}^{k=1} b_j^r a_j\right)^{1/r}} = \frac{(\beta^s + b_{k+1}^s a_{k+1})^{1/s}}{(\alpha^r + b_{k+1}^r a_{k+1})^{1/r}}$$

$$= \frac{\beta}{\alpha} \frac{(1 + b_{k+1}^s a_{k+1}/\beta^s)^{1/s}}{(1 + b_{k+1}^r a_{k+1}/\alpha^r)^{1/r}} \geqslant \frac{\beta}{\alpha}$$

$$= \frac{\left(\sum_{j=1}^{k} b_j^s a_j\right)^{1/s}}{\left(\sum_{j=1}^{k} b_j^r a_j\right)^{1/r}}. \qquad \blacksquare$$

Proof of Theorem 3.2. To prove Theorem 3.2 we will first show that $\delta_n(D;l_p^m;l_q^m) \leq (\sum_{k=n+1}^m D_k^r)^{1/r}$, and then since $1/r = 1/q - 1/p = 1/p' - 1/q'$ and by Lemma 2.2, it will suffice to prove that $g_n(D;l_p^m;l_q^m) \geq (\sum_{k=n+1}^m D_k^r)^{1/r}$.

As in Theorem 3.1, let $P_n = \operatorname{diag}(D_1,\ldots,D_n,0,\ldots,0)$. Then

$$\delta_n(D;l_p^m;l_q^m) \leq \max_{x\neq 0} \frac{\|(D-P)x\|_q}{\|x\|_p}$$

$$= \max_{x\neq 0} \frac{\left(\sum_{k=n+1}^m |D_k x_k|^q\right)^{1/q}}{\left(\sum_{k=1}^m |x_k|^p\right)^{1/p}}.$$

Since $q<p$, we obtain from Hölder's inequality

$$\left(\sum_{k=n+1}^m |D_k x_k|^q\right)^{1/q} \leq \left(\sum_{k=n+1}^m D_k^r\right)^{1/r}\left(\sum_{k=n+1}^m |x_k|^p\right)^{1/p}.$$

Thus $\delta_n(D;l_p^m;l_q^m) \leq (\sum_{k=n+1}^m D_k^r)^{1/r}$.

Now

$$g_n(D;l_p^m;l_q^m) = \min_{X_n} \max_{\substack{x\perp X_n\\ x\neq 0}} \frac{\|Dx\|_q}{\|x\|_p}$$

$$= \min_{X_n} \max_{\substack{x\perp X_n\\ x\neq 0}} \frac{\left(\sum_{k=1}^m |D_k x_k|^q\right)^{1/q}}{\left(\sum_{k=1}^m |x_k|^p\right)^{1/p}}.$$

Put $z_k = x_k D_k^{-q/(p-q)}$ and note that $pq/(p-q) = r$. Since $D_k \neq 0$, $k=1,\ldots,m$,

$$g_n(D;l_p^m;l_q^m) = \min_{X_n} \max_{\substack{z\perp X_n\\ z\neq 0}} \frac{\left(\sum_{k=1}^m |z_k|^q D_k^r\right)^{1/q}}{\left(\sum_{k=1}^m |z_k|^p D_k^r\right)^{1/p}}.$$

Applying Lemmas 3.1 and 3.2, we obtain

$$g_n(D;l_p^m;l_q^m) \geq \frac{\left(\sum_{k=1}^{m-n} D_{i_k}^r\right)^{1/q}}{\left(\sum_{k=1}^{m-n} D_{i_k}^r\right)^{1/p}} = \left(\sum_{k=1}^{m-n} D_{i_k}^r\right)^{1/r}$$

for some $1 \leq i_1 < \cdots < i_{m-n} \leq m$. Since $D_1 \geq \cdots \geq D_m > 0$, $g_n(D;l_p^m;l_q^m) \geq (\sum_{k=n+1}^{m} D_k^r)^{1/r}$. The theorem is proven. ■

4. THE CASE $p<q$

The case $p<q$ is not well understood. A lower bound on the n-width which may be obtained as a consequence of Proposition 2.1, Theorem 3.2, and Lemma 2.5 is the following.

PROPOSITION 4.1. *For* $1 \leq p < q \leq \infty$,

$$\delta_n(D;l_p^m;l_q^m) \geq d_n(D;l_p^m;l_q^m), g_n(D;l_p^m;l_q^m)$$

$$\geq b_n(D;l_p^m;l_q^m) = \left(\sum_{k=1}^{n+1} D_k^r\right)^{1/r},$$

where $1/r = 1/q - 1/p$. *(Note that* $r<0$.)

This lower bound is rarely attained by d_n or g_n. The one general case where it is attained is if $n = m-1$.

PROPOSITION 4.2. *For* $1 \leq p < q \leq \infty$,

$$\delta_{m-1}(D;l_p^m;l_q^m) = d_{m-1}(D;l_p^m;l_q^m) = g_{m-1}(D;l_p^m;l_q^m)$$

$$= b_{m-1}(D;l_p^m;l_q^m) = \left(\sum_{k=1}^{m} D_k^r\right)^{1/r},$$

where $1/r = 1/q - 1/p$.

Proof. As a consequence of Proposition 4.1, it remains to prove that $\delta_{m-1}(D;l_p^m;l_q^m) \leqslant (\sum_{k=1}^m D_k^r)^{1/r}$. It is in fact possible, using Lemma 2.5, to prove the equality for d_{m-1} and g_{m-1}. However, we must also prove this fact in any case for δ_{m-1}. It is also possible to prove the proposition by drawing upon a general result of Brown [2], which in our case states that

$$\delta_{m-1}(D;l_p^m;l_q^m) = \min_{\|x\|_p=1} \|Dx\|_q,$$

which in turn equals the Bernstein $(m-1)$-width and is $(\sum_{k=1}^m D_k^r)^{1/r}$. However, we have no wish to reproduce Brown's general result here and will instead explicitly construct the requisite rank $m-1$ matrix P.

Let $x_k^* = D_k^{r/p}$, $k=1,\dots,m$. Note that $x^* = (x_1^*,\dots,x_m^*)$ is a strictly positive vector and is such that equality is attained in the application of Hölder's inequality given by

$$\left(\sum_{k=1}^m (x_k^*)^p\right)^{1/p} = \left(\sum_{k=1}^m D_k^r\right)^{-1/r}\left(\sum_{k=1}^m (D_k x_k^*)^q\right)^{1/q}$$

(recall that $r<0$).

For $1 \leqslant p < \infty$, choose any $m-1$ linearly independent vectors $\{x^i\}_{i=1}^{m-1}$ which satisfy

$$\sum_{k=1}^m (x_k^*)^{p-1}(x_k^i) = 0, \qquad i = 1,\dots,m-1.$$

If $p=\infty$, then since $\|x^*\|_\infty = x_m^*$, we choose $x^i = e^i$, the ith unit vector, $i=1,\dots,m-1$. In both cases the choice of the $\{x^i\}_{i=1}^{m-1}$ is such that

$$\|x^*\|_p = \min_{a_1,\dots,a_{m-1}} \left\| x^* - \sum_{i=1}^{m-1} a_i x^i \right\|_p.$$

Note that $\{x^1,\dots,x^{m-1},x^*\}$ also forms a basis for R^m. Let P^* be the $m\times m$ matrix defined by

$$P^*x^i = Dx^i, \quad i = 1,\dots,m-1, \quad \text{and} \quad P^*x^* = 0.$$

Since $x^* \neq 0$, P^* is singular and thus $\operatorname{rank} P^* \leqslant m-1$.

Each $x \in R^m$ may be written in the form $x = \sum_{i=1}^{m-1} b_i x^i + b^* x^*$. Since $(D - P^*)x = b^* D x^*$, we may assume in what follows that $b^* \neq 0$. Since

$$|b^*|\,\|x^*\|_p \leqslant \left\| b^* x^* + \sum_{i=1}^{m-1} b_i x^i \right\|_p = \|x\|_p,$$

it follows that

$$\begin{aligned} \delta_{m-1}(D; l_p^m; l_q^m) &\leqslant \max_{x \neq 0} \frac{\|(D - P^*)x\|_q}{\|x\|_p} \\ &\leqslant \frac{|b^*|\,\|Dx^*\|_q}{|b^*|\,\|x^*\|_p} \\ &= \left(\sum_{k=1}^{m} D_k^r \right)^{1/r}. \end{aligned}$$

The proposition is proven. ■

If D is the identity matrix, then for all $1 \leqslant p < q \leqslant \infty$, P^* is simply the matrix whose diagonal entries are all $(m-1)/m$ and whose off diagonal entries are all $-1/m$.

COROLLARY 4.1. *If* $1 \leqslant p < q \leqslant \infty$, $1/r = 1/q - 1/p$, *and* $D_{n+2} \leqslant (\sum_{k=1}^{n+1} D_k^r)^{1/r}$, *then*

$$\begin{aligned} \delta_n(D; l_p^m; l_q^m) &= d_n(D; l_p^m; l_q^m) = g_n(D; l_p^m; l_q^m) \\ &= b_n(D; l_p^m; l_q^m) = \left(\sum_{k=1}^{n+1} D_k^r \right)^{1/r}. \end{aligned}$$

Proof. Again, on the basis of Proposition 4.1, it is only necessary that we prove the upper bound $\delta_n(D; l_p^m; l_q^m) \leqslant (\sum_{k=1}^{n+1} D_n^r)^{1/r}$. The idea of the proof is to use the P^* of the previous proposition.

Let P^* be the $(n+1) \times (n+1)$ matrix constructed in Proposition 4.2, and let P_n denote the $m \times m$ matrix whose first $n+1$ rows and columns agree with P^* and which is zero elsewhere. Thus $\operatorname{rank} P_n \leqslant \operatorname{rank} P^* \leqslant n$. Similarly, let x^* and $\{x^i\}_{i=1}^n$ be the x^* and $\{x^i\}_{i=1}^n$ of Proposition 4.2 with added zero terms. Any $x \in R^m$ may be written in the form

$$x = b^* x^* + \sum_{i=1}^{n} b_i x^i + \sum_{k=n+2}^{m} x_k e^k.$$

If $b^* = 0$, then

$$\frac{\|(D-P_n)x\|_q}{\|x\|_p} = \frac{\left(\sum_{k=n+2}^{m} |D_k x_k|^q \right)^{1/q}}{\|x\|_p} \leqslant D_{n+2} \leqslant \left(\sum_{k=1}^{n+1} D_k^r \right)^{1/r}.$$

If $b^* \neq 0$, then we may assume $b^* = 1$, and as previously,

$$\frac{\|(D-P_n)x\|_q}{\|x\|_p} \leqslant \frac{\left[\|Dx^*\|_q^q + \sum_{k=n+2}^{m} |D_k x_k|^q \right]^{1/q}}{\left[\|x^*\|_p^p + \sum_{k=n+2}^{m} |x_k|^p \right]^{1/p}}$$

$$\leqslant \frac{\left[\|x^*\|_p^q \left(\sum_{k=1}^{n+1} D_k^r \right)^{q/r} + \sum_{k=n+2}^{m} |D_k x_k|^q \right]^{1/q}}{\left[\|x^*\|_p^p + \sum_{k=n+2}^{m} |x_k|^p \right]^{1/p}}.$$

Since $1 \leqslant p < q \leqslant \infty$, the above quantity is less than or equal to

$$\max\left\{ \left(\sum_{k=1}^{n+1} D_k^r \right)^{1/r}, D_{n+2}, \ldots, D_m \right\} = \left(\sum_{k=1}^{n+1} D_k^r \right)^{1/r}.$$

The corollary is proven. ■

The corollary is exact in that there exist $p < q$ for which $d_n(D; l_p^m; l_q^m) = (\sum_{k=1}^{n+1} D_k^r)^{1/r}$ iff $(\sum_{k=1}^{n+1} D_k^r)^{1/r} \geqslant D_{n+2}$. For an example of this, see the next theorem.

The only case in which all the n-widths $d_n(D; l_p^m; l_q^m)$ are known for $p < q$ is when $p = 1$ and $q = 2$. The result (Theorem 4.1) was originally announced but never published by S. A. Smolyak. The first published proof is due to

Sofman [20]. Hutton, Morrell, and Retherford [6] re-proved the result, unaware of the work of Sofman. The proof presented here is an amalgamation of these two proofs. When $D=I$, then the theorem may also be found in Stechkin [23] and Solomjak and Tichomirov [22].

THEOREM 4.1.

$$d_n(D;l_1^m;l_2^m) = \delta_n(D;l_1^m;l_2^m) = \max_{n<k<m} \sqrt{\frac{k-n}{\sum_{i=1}^{k} D_i^{-2}}} .$$

Note that no claim is made concerning the n-widths in the sense of Gel'fand. This is for the excellent reason that equality does not hold. We shall use the following lemmas in the proof of the theorem.

LEMMA 4.1. *For given A_j, $j=1,\ldots,m$, satisfying $0\leqslant A_j\leqslant 1$, $j=1,\ldots,m$, and $\sum_{j=1}^{m}A_j=n$, there exist n orthonormal vectors $x^i\in R^m$, $i=1,\ldots,n$, i.e., $(x^i,x^j)=\delta_{ij}$, for which $A_j=\sum_{i=1}^{n}|(x^i)_j|^2$, $j=1,\ldots,m$.*

Proof. The proof is by induction on m. If $m=n$, it suffices to take any set of orthonormal vectors in R^m, since then $A_j=\sum_{i=1}^{m}|(x^i)_j|^2=1$, $j=1,\ldots,m$.

Assume that $m>n$ and that for any given $\{A_j\}_{j=1}^{m-1}$ satisfying $0\leqslant A_j\leqslant 1$, $j=1,\ldots,m-1$, $\sum_{j=1}^{m-1}A_j=n$, there exist $x^i\in R^{m-1}$, $i=1,\ldots,n$, orthonormal and such that $A_j=\sum_{i=1}^{n}|(x^i)_j|^2$, $j=1,\ldots,m-1$. To advance the induction, let $\{B_j\}_{j=1}^{m}$ be given satisfying the conditions of the lemma, and assume, without loss of generality, that $B_m=\min_{j=1,\ldots,m}B_j$. Let $\{A_j\}_{j=1}^{m-1}$ be any sequence for which $0\leqslant B_j\leqslant A_j\leqslant 1$, $j=1,\ldots,m-1$, and $\sum_{j=1}^{m-1}A_j=n$. Set $A_m=0$. By the induction hypothesis there exists a set of n orthonormal vectors $\{y^i\}_{i=1}^{n}$ in R^m for which $(y^i)_m=0$, $i=1,\ldots,n$, and such that $A_j=\sum_{i=1}^{n}|(y^i)_j|^2$, $j=1,\ldots,m$. The idea of the proof is to move from the sequence $(A_1,\ldots,A_{m-1},A_m)$ to the sequence $(B_1,\ldots,B_m)$, where at the lth step we decrease A_l to B_l while increasing the previous value of A_m, i.e. $\sum_{i=1}^{l-1}(A_i-B_i)$, by the amount A_l-B_l, and where at each step we maintain an orthonormal set of n vectors which satisfy the lemma.

The required transformation is the following: Given any vector $x\in R^m$, let $x[t;l]\in R^m$, where

$$(x[t;l])_i = \begin{cases} (x)_i, & i\neq l,m \\ (x)_l\cos t+(x)_m\sin t, & i=l \\ -(x)_l\sin t+(x)_m\cos t, & i=m. \end{cases}$$

Note that this transformation applied to any set $x^1,\ldots,x^n$ of orthonormal vectors in R^m preserves their orthonormality. Furthermore, if we set

$$A_j(t) = \sum_{i=1}^{n} |(y^i[t;l])_j|^2, \qquad j = 1,\ldots,m,$$

then for all t, $A_j(t)=A_j(0)$, $j\neq l,m$, $\sum_{j=1}^m A_j(t)=n$, and $0\leqslant A_j(t)\leqslant 1$, $j=1,\ldots,m$. Thus $A_l(t)+A_m(t)=A_l(0)+A_m(0)$ for any t and $A_l(\pi/2)=A_m(0)$, $A_m(\pi/2)=A_l(0)$.

The result now follows by the intermediate value theorem, where we make use of the fact that $B_m=\min_{j=1,\ldots,m}B_j$, which implies that as we go from $\{A_j\}_{j=1}^m$ to $\{B_j\}_{j=1}^m$, the minimality property of the last term of the sequence is always preserved. The lemma is proven. ■

LEMMA 4.2. *If $(k-n)/\sum_{i=1}^k D_i^{-2}\geqslant D_{k+1}^2$ for some k, $k>n$, then for all $l>k$,*

$$\frac{k-n}{\sum_{i=1}^{k} D_i^{-2}} \geqslant \frac{l-n}{\sum_{i=1}^{l} D_i^{-2}}.$$

Proof. $(k-n)/\sum_{i=1}^k D_i^{-2}\geqslant D_{k+1}^2$ is equivalent to the fact that $(k-n)/D_{k+1}^2\geqslant \sum_{i=1}^k D_i^{-2}$. Multiplying each side by $l-k$ and using the fact that $D_{k+1}\geqslant\cdots\geqslant D_m>0$, we obtain

$$(k-n)\sum_{i=k+1}^{l} D_i^{-2} \geqslant (l-k)\sum_{i=1}^{k} D_i^{-2},$$

which is equivalent to

$$(k-n)\sum_{i=1}^{l} D_i^{-2} \geqslant (l-n)\sum_{i=1}^{k} D_i^{-2}.$$

The lemma follows. ■

LEMMA 4.3. *For $n<k<m$,*

$$\frac{k-n}{\sum_{i=1}^{k} D_i^{-2}} \leqslant D_{k+1}^2 \quad \textit{iff} \quad \frac{k-n}{\sum_{i=1}^{k} D_i^{-2}} \leqslant \frac{k+1-n}{\sum_{i=1}^{k+1} D_i^{-2}}.$$

Proof. The proof of this lemma follows very much the same reasoning as the proof of Lemma 4.2. Since we take $l=k+1$ in Lemma 4.2, all statements in the proof are "if and only if" statements. ∎

Proof of Theorem 4.1. As a result of either Lemma 2.3 or 2.4, $\delta_n(D;l_1^m;l_2^m)=d_n(D;l_1^m;l_2^m)$. Let X_n be any n-dimensional subspace of R^m, and let $x^1,\ldots,x^n$ be an orthonormal basis for X_n. Set

$$\rho_j(X_n) = \min_{y\in X_n} \|D_je^j - y\|_2, \qquad j = 1,\ldots,m$$

where e^j is the jth unit vector. Since $\{\pm e^j\}_{j=1}^m$ are the extreme points of the unit ball in l_1^m,

$$d_n(D;l_1^m;l_2^m) = \min_{X_n} \max_{j=1,\ldots,m} \rho_j(X_n).$$

The orthonormality of the x^i implies that

$$\rho_j(X_n) = \min_{\alpha_1,\ldots,\alpha_n} \left\| D_je^j - \sum_{i=1}^n \alpha_i x^i \right\|_2$$

$$= \left[|(D_je^j,D_je^j)| - \sum_{i=1}^n |(D_je^j,x^i)|^2 \right]^{1/2}$$

$$= D_j\left[1 - \sum_{i=1}^n |(x^i)_j|^2 \right]^{1/2}.$$

Set $A_j=\sum_{i=1}^n|(x^i)_j|^2$. Lemma 4.1 implies that varying over all n-dimensional subspaces is exactly the same as varying over all $\{A_j\}_{j=1}^m$ satisfying the hypotheses of Lemma 4.1. Thus,

$$d_n(D;l_1^m;l_2^m) = \min_{\substack{0\leq A_j\leq 1\\ \sum_{j=1}^m A_j=n}} \max_{j=1,\ldots,m} D_j\left[1-A_j\right]^{1/2}$$

$$= \min_{\substack{0\leq C_j\leq 1\\ \sum_{j=1}^m C_j=m-n}} \max_{j=1,\ldots,m} D_jC_j^{1/2}.$$

Let k be the smallest integer, $n+1 \leqslant k \leqslant m$, such that $(k-n)/\sum_{i=1}^{k} D_i^{-2} \geqslant D_{k+1}^2$. Let $D_{m+1}=0$, so that such a k necessarily exists. Put

$$C_j^* = \frac{\dfrac{k-n}{D_j^2}}{\displaystyle\sum_{i=1}^{k} D_i^{-2}}, \qquad j=1,\ldots,k,$$

and $C_j^*=1$, $j=k+1,\ldots,m$.

Obviously, $C_j^* \geqslant 0$ and $\sum_{j=1}^{m} C_j^* = m-n$. To show that $\{C_j^*\}_{j=1}^{m}$ is admissible, we must prove that $C_j^* \leqslant 1$. Since $D_1 \geqslant \cdots \geqslant D_k$, it suffices to prove that $C_k^* \leqslant 1$. Now, $C_k^* \leqslant 1$ if and only if $(k-n)D_k^{-2} \leqslant \sum_{i=1}^{k} D_i^{-2}$, which, in turn, is seen to be equivalent to

$$\frac{k-1-n}{\displaystyle\sum_{i=1}^{k-1} D_i^{-2}} \leqslant D_k^2.$$

This latter inequality is valid by the choice of k (and trivially so if $k=n+1$), so that $\{C_j^*\}_{j=1}^{m}$ is admissible.

Since $D_j C_j^{*1/2} = [(k-n)/\sum_{i=1}^{k} D_i^{-2}]^{1/2}$ for $j=1,\ldots,k$, and $D_k \leqslant [(k-n)/\sum_{i=1}^{k} D_i^{-2}]^{1/2}$, it follows that $d_n(D;l_1^m;l_2^m) \leqslant [(k-n)/\sum_{i=1}^{k} D_i^{-2}]^{1/2}$.

To prove the lower bound, assume the existence of an n-dimensional subspace X_n for which $\rho_j(X_n) < [(k-n)/\sum_{i=1}^{k} D_i^{-2}]^{1/2}$ for all $j=1,\ldots,m$. Since $\sum_{j=1}^{m} \rho_j^2(X_n)/D_j^2 = m-n$ and $0 \leqslant \rho_j^2(X_n)/D_j^2 \leqslant 1$, we have

$$k-n \leqslant \sum_{j=1}^{k} \frac{\rho_j^2(X_n)}{D_j^2} < \sum_{j=1}^{k} \frac{(k-n)/\sum_{i=1}^{k} D_i^{-2}}{D_j^2} = k-n.$$

From this contradiction, it follows that $d_n(D;l_1^m;l_2^m) = [(k-n)/\sum_{i=1}^{k} D_i^{-2}]^{1/2}$.

It remains to prove that

$$\sqrt{\frac{k-n}{\displaystyle\sum_{i=1}^{k} D_i^{-2}}} = \max_{n<l\leqslant m} \sqrt{\frac{l-n}{\displaystyle\sum_{i=1}^{l} D_i^{-2}}}\,.$$

This fact is an immediate application of Lemmas 4.2 and 4.3. The theorem is proven. ■

5. $D=I$

In the case where we choose D to be the identity matrix I, things are somewhat simpler. We restate the previous results in this setting and then deal with certain extensions.

THEOREM 5.1. *Assume* $1 \leqslant p,q \leqslant \infty$. *Then*

(i) *for* $p \geqslant q$

$$\delta_n(I;l_p^m;l_q^m) = d_n(I;\ l_p^m;l_q^m) = g_n(I;l_p^m;l_q^m) = (m-n)^{1/q-1/p},$$

(ii) $\delta_n(I;l_1^m;l_2^m) = d_n(I;l_1^m;l_2^m) = ((m-n)/m)^{1/2}$.

(iii) *for* $p \leqslant q$

$$\delta_{m-1}(I;l_p^m;l_q^m) = d_{m-1}(I;l_p^m;l_q^m) = g_{m-1}(I;l_p^m;l_q^m) = m^{-(1/p-1/q)}$$

(iv) *for* $p<q$

$$\delta_n(I;l_p^m;l_q^m) \geqslant d_n(I;l_p^m;l_q^m), g_n(I;l_p^m;l_q^m) \geqslant (n+1)^{-(1/p-1/q)}.$$

We also have

PROPOSITION 5.1. *For* $1<q \leqslant \infty$

$$\delta_1(I;l_1^m;l_q^m) = d_1(I;l_1^m;l_q^m) = \frac{(m-1)^{1/q}}{\left[1+(m-1)^{1/(q-1)}\right]^{1-1/q}}.$$

Proof. The equality between δ_1 and d_1 is a result of Lemma 2.4. Let $e=(1,1,\ldots,1)$. Then

$$d_1(I;l_1^m;l_q^m) \leqslant \max_{j=1,\ldots,m} \min_{\alpha} \|e^j - \alpha e\|_q.$$

Since $\|e^j-\alpha e\|_q = (|1-\alpha|^q + (m-1)|\alpha|^q)^{1/q}$, it follows that for $q>1$, the optimal α is $\alpha = [1+(m-1)^{1/(q-1)}]^{-1}$, from which we obtain

$$d_1(I;l_1^m;l_q^m) \leqslant \frac{(m-1)^{1/q}}{\left[1+(m-1)^{1/(q-1)}\right]^{1-1/q}}.$$

Let $y=(y_1,\dots,y_m)$ be any vector in R^m, and let us assume, without loss of generality, that $|y_1|=\min_{j=1,\dots,m}|y_j|$. If $y_1=0$, then

$$\min_\alpha \|e^1-\alpha y\|_q = 1 > \frac{(m-1)^{1/q}}{\left[1+(m-1)^{1/(q-1)}\right]^{1-1/q}}.$$

Thus we may assume that $y_1\neq 0$. Now,

$$\begin{aligned}\min_\alpha \|e^1-\alpha y\|_q &= \min_\alpha \left[|1-\alpha y_1|^q + |\alpha|^q \sum_{j=2}^m |y_j|^q\right]^{1/q}\\ &\geqslant \min_\alpha \left[|1-\alpha y_1|^q + |\alpha|^q(m-1)|y_1|^q\right]^{1/q}\\ &= \min_\beta \left[|1-\beta|^q + (m-1)|\beta|^q\right]^{1/q}\\ &= \frac{(m-1)^{1/q}}{\left[1+(m-1)^{1/(q-1)}\right]^{1-1/q}}.\end{aligned}$$

Thus $d_1(I;l_1^m;l_q^m)\geqslant (m-1)^{1/q}/[1+(m-1)^{1/(q-1)}]^{1-1/q}$. ∎

Note that the above result is very much simpler in the cases $q=2$ and $q=\infty$. For $q=\infty$, we obtain $d_1(I;l_1^m;l_\infty^m)=\frac{1}{2}$. It may in fact be shown that

$$d_1(D;l_1^m;l_\infty^m) = \frac{D_1D_2}{D_1+D_2}.$$

In this same vein, we have

PROPOSITION 5.2 (Hutton, Morrell, and Retherford [6]). *For* $1\leqslant p\leqslant\infty$,

$$d_1(I;l_p^m;l_\infty^m) = 2^{-1/p}.$$

Proof. From Theorem 5.1(iv), we have $d_1(I;l_p^m;l_\infty^m)\geqslant 2^{-1/p}$. Let $e=(1,1,\dots,1)$. Now,

$$\min_\alpha \|x-\alpha e\|_\infty = \frac{\max x_i - \min x_i}{2}.$$

Thus

$$\begin{aligned} d_1(I;l_p^m;l_\infty^m) &\leqslant \max_{x\neq 0}\min_{\alpha}\frac{\|x-\alpha e\|_\infty}{\|x\|_p} \\ &= \max_{x\neq 0}\frac{1}{2}\frac{\left[\max x_i-\min x_i\right]}{\left[\sum_{i=1}^{m}|x_i|^p\right]^{1/p}} \\ &= \max_{x\neq 0}\frac{1}{2}\frac{\max x_i-\min x_i}{\left[|\max x_i|^p+|\min x_i|^p\right]^{1/p}} \\ &= 2^{-1/p}. \end{aligned}$$

■

Note that no mention is made of the linear n-width $\delta_1(I;l_p^m;l_\infty^m)$. The value $2^{-1/p}$ is a lower bound for $\delta_1(I;l_p^m;l_\infty^m)$ and not a very good lower bound, since by Lemma 2.1, $\delta_n(I;l_p^m;l_\infty^m)=\delta_n(I;l_1^m;l_{p'}^m)$ and from Proposition 5.1,

$$\delta_1(I;l_p^m;l_\infty^m)=\frac{(m-1)^{1-1/p}}{\left[1+(m-1)^{p-1}\right]^{1/p}}.$$

As was noted in the introduction, we have considered, until now, the l_p^m and l_q^m norms on R^m. If we consider these norms on C^m, then the results so far obtained remain unchanged. This fortunate situation is no longer valid if $p=1$, $q=\infty$, as will be seen at the end of this section, and it is therefore necessary that we differentiate between these two choices. Since we shall only be dealing with d_n and δ_n, we shall write δ_n^R, d_n^R or δ_n^C, d_n^C to denote that the underlying space is R^m or C^m, respectively.

A considerable effort has been devoted to the determination of upper bounds for $d_n^R(I;l_1^m;l_\infty^m)$ and $d_n^R(I;l_2^m;l_\infty^m)$ as both n and m increase. These upper bounds are used to obtain exact asymptotic estimates of the n-width of the Sobolev spaces $W_p^r[0,1]$ in $L^q[0,1]$. The main work in this area was done by Kashin [10] (see also Maiorov [13]), who proved

THEOREM 5.2 (Kashin [10]). *There exist constants C_1 and C_2, independent of n and m, such that for all $1\leqslant n<m<\infty$,*

$$d_n^R(I;l_1^m;l_\infty^m)\leqslant\frac{C_1}{\sqrt{n}}\left(1+\ln\frac{m}{n}\right)^{1/2},$$

$$d_n^R(I;l_2^m;l_\infty^m)\leqslant\frac{C_2}{\sqrt{n}}\left(1+\ln\frac{m}{n}\right)^{3/2}.$$

We cannot, in the limited space available, consider reproducing the proof of the above results. The idea of the proof of (say) the second inequality is to construct an $m\times n$ matrix A such that if a^i denotes the ith row of A, $i=1,\ldots,m$, then

(1) every n rows of A are linearly independent;
(2) for every $n+1$ rows $a^{i_1},\ldots,a^{i_{n+1}}$ for which

$$a^{i_{n+1}} = \sum_{k=1}^{n} \alpha_k a^{i_k},$$

$\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)$ satisfies

$$\frac{\|\boldsymbol{\alpha}\|_2+1}{\|\boldsymbol{\alpha}\|_1} \leqslant \frac{C}{\sqrt{n}}\left(1+\ln\frac{m}{n}\right)^{3/2}.$$

If one can construct such a matrix, one then takes as X_n the span of the column vectors of A. It is then not difficult to obtain the desired result. The problem thus reduces to constructing a matrix A with the above properties. The proof of the existence of such an A is probabilistic rather than constructive in character and is rather complicated.

The case $m=2n$ has held particular attraction to many. Szarek [27] for example re-proved, by different means, Kashin's result where $m=2n$, thereby obtaining $d_n^R(I;l_2^{2n};l_\infty^{2n})\leqslant C/\sqrt{n}$. Szarek's interest in this result comes from Banach space theory. From the duality of Lemma 2.2, the above result implies the existence of an n-dimensional subspace X_n of R^{2n} such that for any $x\in X_n$,

$$C\sqrt{n}\ \|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{2n}\ \|x\|_2.$$

(The right hand inequality is simple. It is the left hand inequality which is of interest.) In fact Szarek shows that the orthogonal complement of X_n also has this same property.

As was previously mentioned, the n-width $d_n(I;l_1^m;l_\infty^m)$ is of interest in and of itself, since it is easily shown that

$$d_n(I;l_1^m;l_\infty^m) = \delta_n(I;l_1^m;l_\infty^m) = \min_{\operatorname{rank} P\leqslant n}\ \max_{i,j=1,\ldots,m} |\delta_{ij}-p_{ij}|,$$

where $I=(\delta_{ij})_{i,j=1}^m$ and $P=(p_{ij})_{i,j=1}^m$, P a real or complex $m\times m$ rank n matrix, depending on the space. The first to consider this case was Ismagilov [8], who showed by explicit construction that $d_n^C(I;l_1^{2n};l_\infty^{2n})\leqslant C/\sqrt{n}$ for some constant C, independent of n. Kashin [9] then showed that if $n<m<n^\lambda$ for some constant λ, then there exists a constant C_λ for which $d_n^R(I;l_1^m;l_\infty^m)\leqslant C_\lambda/\sqrt{n}$. It should be noted that it is easy to prove that $d_n^C(I;l_1^m;l_\infty^m)\leqslant d_n^R(I;l_1^m;l_\infty^m)$.

One of the simpler upper bounds is the following. This result was proven with the assistance of R. Loewy.

PROPOSITION 5.3. *Let n be such that there exists a $k\times k$ Hadamard matrix (i.e., a $k\times k$ matrix H all of whose entries are ± 1, and such that $HH^T=kI$) where $n\leqslant k\leqslant n+\sqrt{n}$. Then*

$$d_n^R\left(I;l_1^{2n};l_\infty^{2n}\right)=\delta_n^R\left(I;l_1^{2n};l_\infty^{2n}\right)\leqslant\frac{1}{1+\sqrt{n}}.$$

Proof. Let H denote the $k\times k$ Hadamard matrix. Let A denote an $n\times n$ matrix obtained from H by deleting any $k-n$ rows and columns. Let

$$B=\begin{pmatrix} n^{-1/2}A\\ I\end{pmatrix},$$

where I is the $n\times n$ identity matrix. Thus B is a $2n\times n$ matrix. Consider

$$P=\frac{\sqrt{n}}{1+\sqrt{n}}BB^T.$$

Obviously $\operatorname{rank}P\leqslant n$. Furthermore, since

$$BB^T=\begin{pmatrix} n^{-1}AA^T & n^{-1/2}A\\ n^{-1/2}A^T & I\end{pmatrix},$$

it follows that $p_{ii}=\sqrt{n}/(1+\sqrt{n})$ for all $i=1,\dots,2n$, and $|p_{ij}|\leqslant 1/(1+\sqrt{n})$ for all $i\neq j$ if at least one of the i,j is greater than n. We claim that $|p_{ij}|\leqslant 1/(1+\sqrt{n})$ for all $i\neq j$.

Since A is obtained from H by deleting $k-n$ rows and columns, and since $HH^T=kI$,

$$|(AA^T)_{ij}|\leqslant k-n\qquad\text{for any}\quad i\neq j.$$

Thus for $i \neq j$, $i,j=1,\ldots,n$,

$$|p_{ij}| \leqslant \frac{\sqrt{n}}{1+\sqrt{n}}\,\frac{k-n}{n} = \frac{k-n}{(1+\sqrt{n}\,)\sqrt{n}} \leqslant \frac{1}{1+\sqrt{n}},$$

where we have used the fact that $k \leqslant n+\sqrt{n}$.

Hence

$$\max_{i,j=1,\ldots,2n} |\delta_{ij}-p_{ij}| = \frac{1}{1+\sqrt{n}},$$

which proves the proposition. ■

Thus we are reduced to considering the existence of k for which $k\times k$ Hadamard matrices may be constructed. It is simple to prove that if k is the order of a Hadamard matrix, then $k=1$, 2, or a multiple of 4 (see Street and Wallis [26]). It is a well-known conjecture that the above condition is sufficient for the existence of Hadamard matrices. The conjecture is known to be valid to $k=264$. Hadamard matrices are obtainable if $k=2^m$ for some m, or $k=4m=p+1$, where p is a prime. If the above conjecture is valid, then Proposition 5.3 would hold for all n, except $n=5$, by the above proof. To prove the case $n=5$, let A be the 5×5 matrix whose diagonal entries are all 1 and whose off diagonal entries are all -1.

The result $\delta_n^R(I;l_1^{2n};l_\infty^{2n}) \leqslant 1/(1+\sqrt{n}\,)$ for $n=2^l$ was proved by König in his *Habilitationsschrift*.

Almost no results had been obtained concerning lower bounds for $d_n(I;l_1^m;l_\infty^m)$. From the fact that $d_n(I;l_1^m;l_2^m)=\sqrt{(m-n)/m}$, it is easy to show that $d_n^C(I;l_1^m;\ l_\infty^m) \geqslant \sqrt{m-n}\,/m$. We have as well, from Proposition 4.1, the lower bound $d_n^R(I;l_1^m;l_\infty^m) \geqslant 1/(n+1)$. A sharper lower bound is the following.

PROPOSITION 5.4. *For all* $1 \leqslant n < m < \infty$,

$$d_n^C(I;l_1^m;l_\infty^m) = \delta_n^C(I;l_1^m;l_\infty^m) \geqslant \frac{1}{1+\sqrt{(m-1)n/(m-n)}}$$

Proof. Let P be any matrix of rank at most n, and let $Q=I-P$. Since P has the eigenvalue zero with geometric multiplicity at least $m-n$, the matrix Q possesses the eigenvalue 1 with geometric multiplicity at least $m-n$. Let

$u^1,\dots,u^{m-n}$ be linearly independent eigenfunctions of the eigenvalue 1 of Q. Thus for any $w\in\text{span}\{u^1,\dots,u^{m-n}\}=[u^1,\dots,u^{m-n}]$, we have

$$\sum_{j=1}^{m} q_{ij}w_j = w_i,$$

where $Q=(q_{ij})_{i,j=1}^m$. Let $\max_{i,j}|q_{ij}|=d$. Now,

$$(1-q_{ii})w_i = \sum_{\substack{j=1\\ j\neq i}}^{m} q_{ij}w_j,$$

and since $|1-q_{ii}|\geqslant 1-d$, $|q_{ij}|\leqslant d$, we have

$$|1-q_{ii}||w_i| \leqslant \left\{\sum_{\substack{j=1\\ j\neq i}}^{m} |q_{ij}|^2\right\}^{1/2}\left\{\sum_{\substack{j=1\\ j\neq i}}^{m} |w_j|^2\right\}^{1/2},$$

which implies

$$(1-d)|w_i| \leqslant (m-1)^{1/2}d\left\{\sum_{\substack{j=1\\ j\neq i}}^{m} |w_j|^2\right\}^{1/2}.$$

Thus

$$\frac{\left[(1-d)^2+(m-1)d^2\right]^{1/2}|w_i|}{(m-1)^{1/2}d} \leqslant \left(\sum_{j=1}^{m}|w_j|^2\right)^{1/2}$$

for all $i=1,\dots,m$, which in turn implies that

$$\frac{\left[(1-d)^2+(m-1)d^2\right]^{1/2}}{(m-1)^{1/2}d} \leqslant \frac{\|w\|_2}{\|w\|_\infty}.$$

Therefore,

$$\max_{w\in[u^1,\dots,u^{m-n}]}\frac{\|w\|_\infty}{\|w\|_2} \leqslant \frac{(m-1)^{1/2}d}{\left[(1-d)^2+(m-1)d^2\right]^{1/2}}.$$

We now minimize both sides of the inequality over all possible $m-n$ dimensional subspaces of C^m. Since the right hand side of the inequality is monotone in d, we obtain

$$g_n(I;l_2^m;l_\infty^m) \leqslant \frac{(m-1)^{1/2}d^*}{\left[(1-d^*)^2+(m-1)d^{*2}\right]^{1/2}}$$

where $d^* = d_n^C(I;l_1^m;l_\infty^m)$. Since $g_n(I;l_2^m;l_\infty^m) = d_n(I;l_1^m;l_2^m) = \sqrt{(m-n)/m}$, we have

$$\frac{(m-n)^{1/2}}{m^{1/2}} \leqslant \frac{1}{\left[(1-d^*)^2/(m-1)d^{*2}+1\right]^{1/2}}.$$

This quickly reduces to

$$\left[d^*\right]^2 m(n-1) + 2d^*(m-n) - (m-n) \geqslant 0,$$

which implies that

$$d_n^C(I;l_1^m;l_\infty^m) = d^* \geqslant \frac{1}{1+\sqrt{\dfrac{(m-1)n}{m-n}}}. \qquad \blacksquare$$

Note that this lower bound is exact for $n=1$ and $n=m-1$. The first nontrivial (i.e., $1<n<m-1$) case occurs when $m=4$ and $n=2$. By the above proposition

$$d_2^C(I;l_1^4;l_\infty^4) \geqslant \frac{1}{1+\sqrt{3}}.$$

This lower bound is attained by the rank 2 matrix

$$P = \frac{1}{1+\sqrt{3}}\begin{pmatrix} \sqrt{3} & i & i & i \\ -i & \sqrt{3} & i & -i \\ -i & -i & \sqrt{3} & i \\ -i & i & -i & \sqrt{3} \end{pmatrix}.$$

However, if we consider $d_2^R(I; l_1^4; l_\infty^4)$, it is possible to prove that the upper bound given in Proposition 5.3, i.e. $1/(1+2)$, is sharp.

The author is indebted to Messrs. C. de Boor, R. Loewy, and A. Melkman for their many valuable suggestions.

REFERENCES

1 K. Borsuk, Drei Sätze über die n-dimensionale euklidische Sphäre, *Fund. Mat.* 20:177–191 (1933).
2 A. L. Brown, Best n-dimensional approximation to sets of functions, *Proc. London Math. Soc.* 14:577–594 (1964).
3 C. Eckart and G. Young, The approximation of one matrix by another of lower rank, *Psychometrika* 1:211–218 (1936).
4 I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. of Math. Monographs, Vol. 18, Amer. Math. Soc., Providence, R.I., 1969.
5 C. W. Ha, Approximation numbers of linear operators and nuclear spaces, *J. Math. Anal. Appl.* 46:292–311 (1974).
6 C. V. Hutton, J. S. Morrell, and J. R. Retherford, Diagonal operators, approximation numbers and Kolmogorov diameters, *J. Approximation Theory* 16:48–80 (1976).
7 A. D. Ioffe and V. M. Tichomirov, Duality of convex functions and extremal problems, *Russian Math. Surveys* 23:53–124 (1968).
8 R. S. Ismagilov, Diameter of sets in normed linear spaces and the approximation of functions by trigonometric polynomials, *Russian Math. Surveys* 29:169–186 (1974).
9 B. S. Kashin, On diameters of octahedra, *Uspehi Mat. Nauk* 30:251–252 (1975).
10 B. S. Kashin, The widths of certain finite dimensional sets and classes of smooth functions, *Izv. Akad. Nauk SSSR* 41:334–351 (1977).
11 A. Kolmogorov, Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse, *Annals of Math.* 37:107–110 (1936).
12 M. G. Krein, M. A. Krasnoselski, and D. P. Milman, On deficiency numbers of linear operators in Banach spaces and on some geometric problems, *Sb. Trudov Inst. Mat. Akad. Nauk SSSR* 11:97–112 (1948).
13 V. E. Maiorov, Discretization of the problem of diameters, *Uspehi Mat. Nauk* 30:179–180 (1975).
14 C. A. Micchelli and A. Pinkus, On n-widths in L^∞, *Trans. Amer. Math. Soc.* 234:139–174 (1977).
15 C. A. Micchelli and A. Pinkus, Total positivity and the exact n-width of certain sets in L^1, *Pacific J. Math.* 71:499–515 (1977).
16 C. A. Micchelli and A. Pinkus, Some problems in the approximation of functions of two variables and n-widths of integral operators, *J. Approximation Theory*, 24:51–77 (1978).

17 C. A. Micchelli and A. Pinkus, The n-widths of rank $n+1$ kernels, *J. Integral Equations*, to appear.
18 A. Pietsch, s-numbers of operators in Banach spaces, *Studia Math.*, 51:201–223 (1974).
19 E. Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen I., *Math. Ann.* 63:433–476 (1907).
20 L. B. Sofman, Diameters of octahedra, *Math. Notes* 5:258–262 (1969).
21 L. B. Sofman, Diameters of an infinite-dimensional octahedron, *Moscow Univ. Math. Bull.* 28:45–47 (1973).
22 M. E. Solomjak and V. M. Tichomirov, Some geometric characteristics of the embedding map from W_p^α into C, *Izv. Vysš. Učebn. Zaved. Matematika* 10:76–82 (1967).
23 S. R. Stechkin, The best approximation of given classes of functions, *Uspehi Math. Nauk* 9:133–134 (1954).
24 M. I. Stesin, On Aleksandrov diameters of balls, *Soviet Math. Dokl.* 15:1011–1014 (1974).
25 M. I. Stesin, Aleksandrov diameters of finite-dimensional sets and classes of smooth functions, *Soviet Math. Dokl.* 16:252–256 (1975).
26 A. P. Street and W. D. Wallis, *Combinatorial Theory: An Introduction*, CBRC, Winnipeg, 1977.
27 S. J. Szarek, A short proof and strengthening of a result of Kashin, preprint.
28 V. M. Tichomirov, *Some Problems in the Theory of Approximation* (Russian), Nauka Moscow, 1976.

Received 30 December 1978