Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Approximation Theory 140 (2006) 178 – 190
Journal of Approximation Theory
www.elsevier.com/locate/jat

Approximation characteristics for diagonal operators in different computational settings ☆

Fang Gensun*, Qian Lixin

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

Received 20 May 2005; accepted 9 December 2005

Communicated by Allan Pinkus
Available online 3 February 2006

Abstract

First we study several extremal problems on minimax, and prove that they are equivalent. Then we connect this result with the exact values of some approximation characteristics of diagonal operators in different settings, such as the best n-term approximation, the linear average and stochastic n-widths, and the Kolmogorov and linear n-widths. Most of these exact values were known before, but in terms of equivalence of these extremal problems, we present a unified approach to give them a direct proof.

MSC: 41A46; 42A61; 47A58

Keywords: Extremal problem; Diagonal operator; n-term approximation; n-width; Average case setting; Randomized setting

1. Introduction

Let $\ell_p,\ 0 < p \leqslant \infty$, be the classical real sequence space with the usual norm (or quasi-norm) and $e_i = (0, 0, \ldots, 1, 0, \ldots)$ where the ith coordinate is one and the others are zeros. The Fourier coefficients of $f \in \ell_p$ are denoted by $f_i := \langle f, e_i \rangle$, where $\langle , \rangle$ means the usual inner product in ℓ_2. Consider an operator $T : \ell_p \mapsto \ell_q$ $(0 < p \leqslant q < \infty)$, which is defined by

$$T\{e_i\}_{i=1}^{\infty} = \{\lambda_i e_i\}_{i=1}^{\infty}, \quad i = 1, 2, \ldots, \tag{1}$$

☆ Project supported by the Natural Science Foundation of China (Grant No. 10371009) and Research Fund for the Doctoral Program Higher Education (Grant No. 20050027007).

* Corresponding author.
E-mail address: fanggs@bnu.edu.cn (F. Gensun).

0021-9045/$ - see front matter
doi:10.1016/j.jat.2005.12.005

where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant \cdots \geqslant 0$ and $\lim_{n\mapsto\infty} \lambda_n = 0$. Then T is a diagonal operator from ℓ_p to ℓ_q. Clearly, in a certain sense, the diagonal operators will share some properties of eigenvalues for compact operators on the Hilbert space.

The purpose of this paper is to study several extremal problems which are in close relation with various approximation methods in the worst, average case setting, and the randomized setting. Using these results, we give a new and simple proof to the exact value of the best n-term approximation of the diagonal operators on the space ℓ_p, $0 < p < \infty$, and we continue the research of Mathé [8,9], and Novak [11] to determine the exact constants of the linear average n-widths of diagonal operators on the space ℓ_p $(1 < p < 2)$. Finally, it is also the most important contribution of this paper that we discuss the relations between the various n-widths and the best n-term approximation in ℓ_2 space. These approximation quantities are equal, and they were obtained before by different authors [4,9,10,11,15,16,17,19] in different methods, but we present a unified approach to give them a direct proof in terms of the equivalence of these extremal problems without using the exact value of the extremal problems.

The organization of the paper is as follows. In Section 2, we study several extremal problems on minimax. Section 3 consists of four subsections. Each of the first three subsections is connected with one or two extremal problems, and in Section 3.4, we review some related results of Gel'fand and information n-widths. In the last section, we give a summary.

2. Exact solutions of some extremal problems

In this section, we study several extremal problems which are connected with various n-widths and the best n-term approximation.

Let $n \in \mathbb{N}$ be given, $B = \left\{b = \{b_i\}_{i=1}^{\infty} : b_i \geqslant 0,\ \sum_{i=1}^{\infty} b_i \leqslant 1\right\}$, and

$$\mathscr{X}_n = \left\{\xi = \{\xi_i\}_{i=1}^{\infty} :\ 0 \leqslant \xi_i \leqslant 1,\ \sum_{i=1}^{\infty} \xi_i \leqslant n\right\}.$$

It is clear that B and $\mathscr{X}_n$ are closed convex sets. Let us denote by Ext(B) (Ext($\mathscr{X}_n$)) the set of all extreme points of B ($\mathscr{X}_n$). Clearly, $\mathrm{Ext}(B) = \{e_i,\ i = 1, 2, \ldots\} \bigcup \{0\}$, and $\mathrm{Ext}(\mathscr{X}_n) = \bigcup_{k=0}^{n} \{\xi = \{\xi_i\}_{i=1}^{\infty},\ \xi_i \in \{0, 1\},\ \sum_{i=1}^{\infty} \xi_i = k\}$.

We start with a proposition, which plays a key role in connecting to two different extremal problems, and can be obtained from much more result about properties of a continuous, convex and real function defined on a compact and convex set. Here, we prefer to give a elementary proof.

Proposition 1. *Let $c_i \geqslant 0,\ i = 1, 2, \ldots,$ and $\sum_{i=1}^{\infty} c_i < \infty$. Then*

$$\sup_{\xi\in\mathscr{X}_n} \sum_{i=1}^{\infty} c_i \xi_i = \sup_{\xi\in\mathrm{Ext}(\mathscr{X}_n)} \sum_{i=1}^{\infty} c_i \xi_i. \tag{2}$$

Proof. Let $c_{j_1}, c_{j_2}, \ldots, c_{j_n}$ be the first n numbers which are not equal to zero and let $M := \min\{c_{j_1}, c_{j_2}, \ldots, c_{j_n}\} > 0$. Since $\sum_{i=1}^{\infty} c_i < \infty$, there exits a natural number $N > n$ such that $\sum_{k=N+1}^{\infty} c_k < M$. So $c_k < M$ for all $k \geqslant N + 1$. We rearrange the numbers $c_1, c_2, \ldots, c_N$ and obtain $c_{i_1} \geqslant c_{i_2} \geqslant \cdots \geqslant c_{i_N}$. Let $\Gamma_n^* = \{i_1, \ldots, i_n\}$, $\xi^* = \{\xi_i^*\}_{i=1}^{\infty}$ where $\xi_i^* = 1$ if $i \in \Gamma_n^*$ and

$\xi_i^* = 0$ if $i \notin \Gamma_n^*$, then

$$\sup_{\xi \in \mathrm{Ext}(\mathscr{X}_n)} \sum_{i=1}^{\infty} c_i \xi_i = \sum_{i=1}^{\infty} c_i \xi_i^*.$$

It is clear that

$$\sup_{\xi \in \mathscr{X}_n} \sum_{i=1}^{\infty} c_i \xi_i \geqslant \sup_{\xi \in \mathrm{Ext}(\mathscr{X}_n)} \sum_{i=1}^{\infty} c_i \xi_i = \sum_{i=1}^{\infty} c_i \xi_i^*. \tag{3}$$

On the other hand, if $i \notin \Gamma_n^*$, then $c_i \leqslant c_{i_n}$, thus, for any $\xi \in \mathscr{X}_n$, we have

$$\begin{aligned} \sum_{i=1}^{\infty} c_i \xi_i &= \sum_{i \in \Gamma_n^*} c_i \xi_i + \sum_{i \notin \Gamma_n^*} c_i \xi_i \leqslant \sum_{i \in \Gamma_n^*} c_i \xi_i + c_{i_n} \left(n - \sum_{i \in \Gamma_n^*} \xi_i \right) \\ &= \sum_{i \in \Gamma_n^*} c_i \xi_i + c_{i_n} \sum_{i \in \Gamma_n^*} (1 - \xi_i) \leqslant \sum_{i \in \Gamma_n^*} c_i \xi_i + \sum_{i \in \Gamma_n^*} c_i (1 - \xi_i) \\ &= \sum_{i \in \Gamma_n^*} c_i = \sum_{i=1}^{\infty} c_i \xi_i^*, \end{aligned} \tag{4}$$

which together with (3) completes the proof of Proposition 1. □

A family $\mathscr{F} = \{\Phi_\beta : \beta \in B\}$ of real functions defined on some set K is said to be concave, if given any collection $\Phi_1, \Phi_2, \ldots, \Phi_n$ in $\mathscr{F}$ and $c_1, c_2, \ldots, c_n \geqslant 0$ such that $\sum_{i=1}^{n} c_i = 1$, there is a $\Phi \in \mathscr{F}$ satisfying

$$\Phi(x) \geqslant \sum_{i=1}^{n} c_i \Phi_i(x), \quad \forall x \in K.$$

Lemma 1 (*Pietsch [13, p. 40, Ky Fan Lemma]*). *Let K be a compact convex subset of a linear topological Hausdorff space and let $\mathscr{F}$ be a concave collection of lower semi-continuous convex real functions Φ on K. Suppose that for every $\Phi \in \mathscr{F}$ there exists an $x \in K$ with $\Phi(x) \leqslant \rho$. Then we can find an $x_0 \in K$ such that $\Phi(x_0) \leqslant \rho$ for all $\Phi \in \mathscr{F}$, simultaneously.*

Using Ky Fan's lemma (or minimax theorem) above, we get

Proposition 2. *Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_i \geqslant \cdots \geqslant 0$ and $\lim_{i \to \infty} \lambda_i = 0$. Then*

$$\sup_{b \in B}, \inf_{\xi \in \mathscr{X}_n} \sum_{i=1}^{\infty} \lambda_i b_i (1 - \xi_i) = \inf_{\xi \in \mathscr{X}_n}, \sup_{b \in B} \sum_{i=1}^{\infty} \lambda_i b_i (1 - \xi_i). \tag{5}$$

Proof. Let $\mathscr{X}_n$ be considered as a subset of ℓ_2. Then it is closed and convex. Moreover, it is compact with respect to the weak topology. For every $b \in B$, we denote

$$\Phi_b(\xi) := \sum_{i=1}^{\infty} \lambda_i b_i (1 - \xi_i), \quad \xi \in \mathscr{X}_n.$$

So the function $\Phi_b(\xi)$ is continuous on ξ. Let $\mathscr{F} = \{\Phi_b,\ b \in B\}$, $m \in \mathbb{N}$ and $c_1, c_2, \ldots, c_m \geqslant 0$ such that $\sum_{i=1}^m c_i = 1$. Let $b^* = \sum_{i=1}^m c_i e_i$. Then $b^* \in B$ and

$$\Phi_{b^*}(\xi) = \sum_{i=1}^{m} c_i \Phi_{e_i}(\xi), \qquad \xi \in \mathscr{X}_n.$$

So $\mathscr{F}$ is concave. Let $\varepsilon > 0$ and put $\rho := \sup_{b \in B} \inf_{\xi \in \mathscr{X}_n} \Phi_b(\xi)$. For all $b \in B$, we could find a $\xi \in \mathscr{X}_n$ with $\Phi_b(\xi) \leqslant \rho + \varepsilon$. Ky Fan's lemma implies the existence of a $\xi_0 \in \mathscr{X}_n$, such that for all $b \in B$ we have $\Phi_b(\xi_0) \leqslant \rho + \varepsilon$. This means

$$\inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \Phi_b(\xi) \leqslant \rho + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we obtain

$$\inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \Phi_b(\xi) \leqslant \rho.$$

Hence

$$\inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \Phi_b(\xi) \leqslant \sup_{b \in B} \inf_{\xi \in \mathscr{X}_n} \Phi_b(\xi).$$

On the other hand, it is clear

$$\sup_{b \in B} \inf_{\xi \in \mathscr{X}_n} \Phi_b(\xi) \leqslant \inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \Phi_b(\xi).$$

So the proof is completed. □

Propositions 1 and 2 give the main result of this section, which states that the values of following four extremal problems on minimax are equal.

Theorem 1. *Let* $0 < p < \infty$, $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_i \geqslant \cdots \geqslant 0$ *and* $\lim_{i \to \infty} \lambda_i = 0$. *Then*

$$\sup_{b \in B} \inf_{\xi \in \mathrm{Ext}(\mathscr{X}_n)} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i) \tag{6a}$$

$$= \sup_{b \in B} \inf_{\xi \in \mathscr{X}_n} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i) \tag{6b}$$

$$= \inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i) \tag{6c}$$

$$= \inf_{\xi \in \mathscr{X}_n} \sup_{b \in \mathrm{Ext}(B)} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i). \tag{6d}$$

Proof. Proposition 1 implies that (6a)=(6b). From Proposition 2, we have (6b)=(6c). Obviously, (6d)$\leqslant$(6c) and

$$\inf_{\xi \in \mathscr{X}_n} \sup_{b \in \mathrm{Ext}(B)} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i) = \inf_{\xi \in \mathscr{X}_n} \sup_{i \in \mathbb{N}} \lambda_i^p (1 - \xi_i). \tag{7}$$

So (6c)$\leqslant$(6d), and the proof of the theorem is complete. □

Following some ideas of [4], [14, pp. 207–210], we give the solution of the extremal problem (6d).

Proposition 3. *Let* $0 < p < \infty$, $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_i \geqslant \cdots \geqslant 0$ *and* $\lim_{i\to\infty} \lambda_i = 0$. *Then*

$$\inf_{\xi \in \mathscr{X}_n} \sup_{b \in \mathrm{Ext}(B)} \sum_{i=1}^{\infty} \lambda_i^p b_i (1-\xi_i) = \max_{m>n} \frac{m-n}{\sum_{i=1}^m \lambda_i^{-p}}. \tag{8}$$

Proof. From the equality (7), we only need to prove that

$$\inf_{\xi \in \mathscr{X}_n} \sup_{i \in \mathbb{N}} \lambda_i^p (1-\xi_i) = \max_{m>n} \frac{m-n}{\sum_{i=1}^m \lambda_i^{-p}}.$$

If $\lambda_m > 0$, we put

$$\lambda_n(m) = \frac{m-n}{\sum_{i=1}^m \lambda_i^{-p}} \quad \text{for } m = n+1, \ldots .$$

Otherwise let $\lambda_n(m) = 0$. It follows from

$$\lambda_k^{-p} \sup_{i\in\mathbb{N}} \left\{(1-\xi_i)\lambda_i^p\right\} \geqslant (1-\xi_k) \quad \text{for all } k \in \mathbb{N}$$

that

$$\sum_{k=1}^m \lambda_k^{-p} \sup_{i\in\mathbb{N}} \left\{(1-\xi_i)\lambda_i^p\right\} \geqslant \sum_{k=1}^m (1-\xi_k) \geqslant m-n.$$

Consequently,

$$\sup_{i\in\mathbb{N}} \left\{(1-\xi_i)\lambda_i^p\right\} \geqslant \lambda_n(m) \quad \text{for } m = n+1, \ldots,$$

which implies that

$$\inf_{\xi \in \mathscr{X}_n} \sup_{i \in \mathbb{N}} \lambda_i^p (1-\xi_i) \geqslant \sup_{m>n} \lambda_n(m). \tag{9}$$

Now we prove that the sup in (9) can be attained. Put $s = [m/2]$. For $m > n$, we have

$$\frac{m-n}{\sum_{i=1}^m \lambda_i^{-p}} \leqslant \frac{m-n}{s\lambda_1^{-p} + s\lambda_s^{-p}} \to 0 \quad \text{as } m \to \infty,$$

hence, there exists an $m_0 \geqslant n+1$ such that

$$\lambda_n(m_0) = \max_{m>n} \lambda_n(m) = \sup_{m>n} \lambda_n(m).$$

By an easy computation, it follows from

$$\frac{m_0-1-n}{\sum_{i=1}^{m_0-1} \lambda_i^{-p}} \leqslant \frac{m_0-n}{\sum_{i=1}^{m_0} \lambda_i^{-p}} \quad \text{and} \quad \frac{m_0-n}{\sum_{i=1}^{m_0} \lambda_i^{-p}} \geqslant \frac{m_0+1-n}{\sum_{i=1}^{m_0+1} \lambda_i^{-p}},$$

that

$$\lambda_{m_0}^p \geqslant \lambda_n(m_0) \geqslant \lambda_{m_0+1}^p. \tag{10}$$

Put

$$\xi_i^* = \begin{cases} 1 - \lambda_n(m_0)\lambda_i^{-p} & \text{for } i = 1, \ldots, m_0, \\ 0 & \text{for } i = m_0 + 1, \ldots . \end{cases}$$

Then $0 \leqslant \xi_i^* \leqslant 1$, and $\sum_{i=1}^{\infty} \xi_i^* = \sum_{i=1}^{m_0} (1 - \lambda_n(m_0)\lambda_i^{-p}) = m_0 - (m_0 - n) = n$. By virtue of (10) and

$$(1 - \xi_i^*)\lambda_i^p = \begin{cases} \lambda_n(m_0) & \text{for } i = 1, \ldots, m_0, \\ \lambda_i^p & \text{for } i = m_0 + 1, \ldots, \end{cases}$$

we have

$$\sup_{i \in \mathbb{N}} \left\{(1 - \xi_i^*)\lambda_i^p\right\} = \lambda_n(m_0),$$

consequently,

$$\inf_{\xi \in \mathscr{X}_n} \sup_{i \in \mathbb{N}} \lambda_i^p (1 - \xi_i) \leqslant \lambda_n(m_0), \tag{11}$$

which combining with (9) completes the proof of Proposition 3. □

From Proposition 3, we get

Corollary 1. *The four extremal problems in Theorem* 1 *have the same value, i.e.,*

$$(6a) = (6b) = (6c) = (6d) = \max_{m > n} \frac{m - n}{\sum_{i=1}^{m} \lambda_i^{-p}}.$$

3. Connections to various approximation quantities

This section includes four subsections. First, in Section 3.1, we connect the best n-term approximation of the diagonal operator on the space ℓ_p, $0 < p < \infty$, with the extremal problem (6a). Second, in Section 3.2, we study the linear average and stochastic n-widths, which are related to the extremal problems (6b) and (6c). Then in Section 3.3, we investigate the relations of the Kolmogorov and linear n-widths with the extremal problem (6d). Finally, in Section 3.4, we review some related results of Gel'fand and information n-widths. Most results of this section were known before, except the exact value of the linear average n-width of the diagonal operator from ℓ_p to ℓ_2, $1 < p < 2$ (see Theorem 2 and Remark 2), our main contribution of this section is to study the best n-term approximation, and various n-widths of the diagonal operator in terms of the equivalence of the extremal problems (6a)–(6d).

3.1. Best n-term approximation

This subsection is devoted to the best n-term approximation due to Stechkin [18] in the studying absolute convergence of the Fourier sequence [17]. The problem of n-term approximation is closely related to some other nonlinear approximation problems and also to the applications in the image and signal processing (for example, see [1] and the survey [2]).

Now we recall the definition of the best n-term approximation. Assume that Γ_n is an arbitrary collection of n natural numbers, $a_i,\ i \in \Gamma_n$, are real numbers, and $P_{\Gamma_n} = \sum_{i\in\Gamma_n} a_i e_i$. Then the quantity

$$\sigma_n(f)_p = \inf_{a_i,\Gamma_n} \|f - P_{\Gamma_n}\|_p = \left(\|f\|_p^p - \sup_{\Gamma_n} \sum_{i\in\Gamma_n} |f_i|^p \right)^{1/p} \tag{12}$$

is called the best n-term approximation of $f \in \ell_p$ in the space ℓ_p. For the operator T denoted by (1), the numbers $\sigma_n(T)$ can be presented as

$$\begin{aligned}\sigma_n(T) &:= \sigma_n(T : \ell_p \mapsto \ell_q)\\ &= \sup_{f\in B\ell_p} \left(\|T(f)\|_q^q - \sup_{\Gamma_n} \sum_{i\in\Gamma_n} \lambda_i^q |f_i|^q \right)^{1/q}\\ &= \sup_{f\in B\ell_p} \left(\sum_{i=1}^{\infty} \lambda_i^q |f_i|^q - \sup_{\Gamma_n} \sum_{i\in\Gamma_n} \lambda_i^q |f_i|^q \right)^{1/q}\\ &= \sup_{f\in B\ell_p} \inf_{\Gamma_n} \left(\sum_{i\notin\Gamma_n} \lambda_i^q |f_i|^q \right)^{1/q}, \quad 0 < p \leqslant q < \infty. \end{aligned} \tag{13}$$

Because Γ_n can be identified with $\mathrm{Ext}(\mathscr{X}_n)$, we have

Lemma 2. *Let* $0 < p < \infty$. *Then*

$$\sigma_n^p(T : \ell_p \mapsto \ell_p) = \sup_{b\in B} \inf_{\xi\in \mathrm{Ext}(\mathscr{X}_n)} \sum_{i=1}^{\infty} \lambda_i^p b_i (1 - \xi_i).$$

Lemma 2 and Corollary 1 imply

Corollary 2 (*Stepanets [19]*). *Let* $0 < p < \infty$. *Then*

$$\sigma_n(T : \ell_p \mapsto \ell_p) = \max_{m>n} \left(\frac{m-n}{\sum_{i=1}^m \lambda_i^{-p}} \right)^{1/p}.$$

Remark 1. The above result was obtained by Stepanets in [19]. However, our proof is different.

3.2. Linear average and stochastic n-widths

In this subsection, we continue the research of Novak [11] and Mathé [9] to determine the exact value of the linear average, and stochastic n-widths for the diagonal operator T from ℓ_p, $1 \leqslant p \leqslant 2$, to ℓ_2.

Following Mathé [8,9] and Novak [11], we begin with the definitions of the linear average, and stochastic n-widths. Let X, Y be normed spaces and T be a linear and continuous operator from X to Y. We approximate T by

$$T_n = \phi \circ N \quad \text{where } N : X \mapsto \mathbb{R}^n \text{ and } \phi : \mathbb{R}^n \mapsto Y.$$

One can study different classes of admissible information N and algorithms φ. However, in this paper, we always assume the information N is linear and continuous. The approximation of T by T_n can be measured in the different settings, such as the worst case setting, the average case setting and the randomized setting (see the monograph Traub et al. [21] for the history of this problem and further information).

Denote by BX be the closed unit ball of the space X. Let $\mathscr{B}(BX)$ be the smallest σ-algebras on BX containing all open sets in BX, and μ be a probability measure on $[BX, \mathscr{B}(BX)]$. The average error of T by T_n with respect to the measure μ is defined by

$$\triangle_\mu(T_n) = \left(\int_{BX} \|T(f) - T_n(f)\|_Y^2 \, d\mu(f) \right)^{1/2}. \tag{14}$$

The number

$$\begin{aligned} a_n^{\mu-\mathrm{avg}}(T) &:= a_n^{\mu-\mathrm{avg}}(T : X \mapsto Y) \\ &= \inf \left\{ \triangle_\mu(T_n) | N \text{ linear and continuous}, \ \phi \text{ linear} \right\} \\ &= \inf \left\{ \triangle_\mu(T_n) | T_n \text{ linear and continuous with } \dim(T_n(X)) \leqslant n \right\} \end{aligned} \tag{15}$$

is called μ linear average n-width [8,21].

Let $\mathcal{P}(BX)$ be the set of all probability measures on $[BX, \mathscr{B}(BX)]$. Following Mathé [8], we define the linear average n-width by

$$a_n^{\mathrm{avg}}(T) := a_n^{\mathrm{avg}}(T : X \mapsto Y) = \sup_{\mu \in \mathcal{P}(BX)} a_n^{\mu-\mathrm{avg}}(T). \tag{16}$$

Now we consider randomized (or Monte Carlo) methods of the form $T_n^\omega = \phi^\omega \circ N^\omega$, i.e., the mapping $N : X \mapsto \mathbb{R}^n$ and $\phi : \mathbb{R}^n \mapsto Y$ are randomly chosen. The error of T_n^ω is defined by

$$\triangle_{\max}(T_n^\omega) = \sup_{f \in BX} \left(E(\|T(f) - T_n^\omega(f)\|_Y^2) \right)^{1/2}, \tag{17}$$

where E means the expectation of a random variable. We always assume that $(f, \omega) \mapsto \|T(f) - T_n(f)\|_Y$ is measurable, i.e., $\triangle_{\max}(T_n^\omega)$ is well defined. Similarly, we define the linear stochastic n-width $a_n^{\mathrm{ran}}(T)$ (see [3,8,11]) by

$$\begin{aligned} a_n^{\mathrm{ran}}(T) &:= a_n^{\mathrm{ran}}(T : X \mapsto Y) \\ &= \inf\{\triangle_{\max}(T_n^\omega) | \ N^\omega \text{ linear and continuous}, \ \phi^\omega \text{ linear}\} \\ &= \inf\{\triangle_{\max}(T_n^\omega) | \ T_n^\omega \text{ linear and continuous with } \dim(T_n^\omega(X) \leqslant n)\}. \end{aligned} \tag{18}$$

The following lemma gives a relation between the linear average and stochastic n-widths and whose proof can be immediately obtained by Fubini's theorem.

Lemma 3.

$$a_n^{\mathrm{avg}}(T) \leqslant a_n^{\mathrm{ran}}(T). \tag{19}$$

Consider the diagonal operator T defined by Eq. (1). The two equalities in the following lemma can be found in [9,11] and relate the n-widths to the extremal problems (6b) and (6c) from Theorem 1.

Lemma 4.

$$(a_n^{\mathrm{avg}}(T : \ell_1 \mapsto \ell_2))^2 = \sup_{b \in B} \inf_{\xi \in \mathscr{X}_n} \sum_{i=1}^{\infty} \lambda_i^2 b_i (1 - \xi_i), \tag{20}$$

$$(a_n^{\mathrm{ran}}(T : \ell_2 \mapsto \ell_2))^2 = \inf_{\xi \in \mathscr{X}_n} \sup_{b \in B} \sum_{i=1}^{\infty} \lambda_i^2 b_i (1 - \xi_i). \tag{21}$$

Note that the embedding $\ell_p \hookrightarrow \ell_q$ $(1 \leqslant p < q \leqslant \infty)$ is measurable (see [22, p. 15]), using Lemmas 3 and 4, we get the main result of this subsection as follows:

Theorem 2. *Let* $1 \leqslant p \leqslant 2$. *Then*

$$a_n^{\mathrm{avg}}(T : \ell_p \mapsto \ell_2) = a_n^{\mathrm{ran}}(T : \ell_p \mapsto \ell_2) = \max_{m > n} \left(\frac{m - n}{\sum_{i=1}^{m} \lambda_i^{-2}} \right)^{1/2}. \tag{22}$$

Proof. Since $\ell_1 \subset \ell_p \subset \ell_2$ $(1 < p < 2)$, by Lemma 3, we have

$$\begin{aligned} a_n^{\mathrm{avg}}(T : \ell_1 \mapsto \ell_2) &\leqslant a_n^{\mathrm{avg}}(T : \ell_p \mapsto \ell_2) \\ &\leqslant a_n^{\mathrm{ran}}(T : \ell_p \mapsto \ell_2) \leqslant a_n^{\mathrm{ran}}(T : \ell_2 \mapsto \ell_2), \quad 1 < p < 2. \end{aligned} \tag{23}$$

Using the results of Theorem 1, Lemma 4 and Corollary 1, we obtain

$$a_n^{\mathrm{avg}}(T : \ell_1 \mapsto \ell_2) = a_n^{\mathrm{ran}}(T : \ell_2 \mapsto \ell_2) = \max_{m > n} \left(\frac{m - n}{\sum_{i=1}^{m} \lambda_i^{-2}} \right)^{1/2},$$

which together with (23) completes the proof. □

Remark 2. The exact value of $a_n^{\mathrm{ran}}(T : \ell_2 \mapsto \ell_2)$ is due to Novak [11], $a_n^{\mathrm{avg}}(T : \ell_1 \mapsto \ell_2)$ and $a_n^{\mathrm{ran}}(T : \ell_p \mapsto \ell_2)$, $1 \leqslant p < 2$, were calculated by Mathé in [9] and [10], the exact value of $a_n^{\mathrm{avg}}\ (T : \ell_p \mapsto \ell_2)$, $1 < p < 2$, is new.

Remark 3. In [20], Stepanets proved

$$\sigma_n(T : \ell_p \mapsto \ell_2) = \max_{m > n} \frac{(m - n)^{1/2}}{\left(\sum_{i=1}^{m} \lambda_i^{-p}\right)^{1/p}}, \quad 1 \leqslant p \leqslant 2.$$

So, it is interesting to note that for any $p \in [1, 2]$,

$$\sigma_n(T : \ell_2 \mapsto \ell_2) = a_n^{\text{avg}}(T : \ell_p \mapsto \ell_2) = a_n^{\text{ran}}(T : \ell_p \mapsto \ell_2).$$

However, if $1 \leqslant p < 2$, then

$$\sigma_n(T : \ell_p \mapsto \ell_2) < a_n^{\text{avg}}(T : \ell_p \mapsto \ell_2) = a_n^{\text{ran}}(T : \ell_p \mapsto \ell_2).$$

3.3. *Kolmogorov and linear n-widths*

First we recall the definitions of the Kolmogorov and linear n-widths (see [14]). Let X and Y be normed linear spaces, BX be the closed unit ball of X, and $T : X \mapsto Y$ be a bounded linear operator. The Kolmogorov n-width of $T(BX)$ in Y is defined by

$$d_n(T : X \mapsto Y) := d_n(T(BX), Y) = \inf_{X_n} \sup_{x \in BX} \inf_{y \in X_n} \|T(x) - y\|_Y,$$

where X_n runs over all subspaces of Y of dimension n or less.

The linear n-width of $T(BX)$ in Y is given by

$$a_n(T : X \mapsto Y) := a_n(T(BX), Y) = \inf_{P_n} \sup_{x \in BX} \|T(x) - P_n(x)\|_Y,$$

where the infimum is taken over all continuous linear operators P_n of X into Y of rank at most n.

As in [13, pp. 161–162], for the diagonal operator T defined by (1), we have

Lemma 5.

$$\begin{aligned}(d_n(T : \ell_1 \mapsto \ell_2))^2 &= (a_n(T : \ell_1 \mapsto \ell_2))^2 \\ &= \inf_{\xi \in \mathscr{X}_n} \sup_{b \in \operatorname{Ext}(B)} \sum_{i=1}^{\infty} \lambda_i^2 b_i (1 - \xi_i). \end{aligned} \tag{24}$$

This lemma relates to the quantity (6d) of Theorem 1. So from Lemmas 2, 4 and 5, applying Theorem 1, we obtain

Theorem 3.

$$\begin{aligned}\sigma_n(T : \ell_2 \mapsto \ell_2) &= a_n^{\text{avg}}(T : \ell_1 \mapsto \ell_2) = a_n^{\text{ran}}(T : \ell_2 \mapsto \ell_2) \\ &= d_n(T : \ell_1 \mapsto \ell_2) = a_n(T : \ell_1 \mapsto \ell_2).\end{aligned}$$

Remark 4. (a) Let I be the identity operator, i.e., let $\lambda_1 = \cdots \lambda_M = 1$ and $\lambda_j = 0$ for all $j \geqslant M+1$ in Theorem 3 (finite-dimensional case). We have

$$d_n(I : \ell_1^M \mapsto \ell_2^M) = (1 - n/M)^{1/2}, \quad 1 \leqslant n \leqslant M. \tag{25}$$

The upper estimate in (25) was proved in the paper of Kolmogorov, Petrov, Smirnov in 1947 [5], and the low bound was obtained by Maltsev [6]. But the authors of these papers did not actually state that they had calculated n-width $d_n(I : \ell_1^M \mapsto \ell_2^M)$. This was noted by Stechkin in [17].

The exact value of the Kolmogorov n-width $d_n(T : \ell_1^M \mapsto \ell_2^M)$ of the diagonal operator T was due to Sofman in 1969 (see [15]). Then in 1973 he extended this result to the infinite-dimensional case in [16] (see also [4]).

(b) Sun first noted in [23] that from

$$d_n(T:\ell_1 \mapsto \ell_2) = \max_{m>n} \left(\frac{m-n}{\sum_{k=1}^m \lambda_k^{-2}} \right)^{1/2}, \tag{26}$$

and the result of Stepanets [19]

$$\sigma_n(T:\ell_2 \mapsto \ell_2) = \max_{m>n} \left(\frac{m-n}{\sum_{k=1}^m \lambda_k^{-2}} \right)^{1/2}, \tag{27}$$

one has

$$d_n(I:\ell_1 \mapsto \ell_2) = \sigma_n(I:\ell_2 \mapsto \ell_2). \tag{28}$$

He pointed out that the equality (28) is interesting in its own and conjectured that there should be a direct proof without using the exact values. Our Theorem 3 confirms this claim.

(c) The exact values of the best n-term approximation, and other n-widths in Theorem 3 were also known before [19,11,9,14]. However, our proof is new and does not depend on the knowledge of the values of these quantities.

3.4. Gel'fand and information n-widths

In this subsection, we review some related results of Gel'fand and information n-widths (see [7,12]). Let X and Y be normed linear spaces, BX be the closed unit ball of X, and $T: X \mapsto Y$ be a bounded linear operator.

The Gel'fand n-widths of $T(BX)$ in Y is given by

$$d^n(T:X \mapsto Y) := d^n(T(BX), Y) = \inf_{L^n} \sup_{x \in BL^n} \|T(x)\|_Y,$$

where BL^n is the closed unit ball of L^n and the infimum is taken over all subspaces L^n of X of codimension n.

The information n-widths of $T(BX)$ in Y is defined by

$$i_n(T:X \mapsto Y) := i_n(T(BX), Y) = \inf_{\substack{Z \supset BX \\ l_1,\ldots,l_n \in Z^*}} \inf_{S:M^n \mapsto Y} \sup_{f \in BX} \|T(f) - S(l_1 f, \ldots, l_n f)\|_Y,$$

where the infimum runs over all normed linear spaces Z containing BX, here $M = \mathbb{R}$ or $\mathbb{C}$ and Z^* is the dual space of Z.

Lemma 6 (*Mathé [7, Theorem 5]; Pietsch [13, p. 161]; Osipenko [12, pp. 155–159]*). *Let T be the diagonal operator defined by* (1), $n \in \mathbb{N}$. *Then*

$$\begin{aligned} a_n(T:\ell_2 \mapsto \ell_\infty) &= d^n(T:\ell_2 \mapsto \ell_\infty) \\ &= i_n(T:\ell_2 \mapsto \ell_\infty) = \max_{m>n} \left(\frac{m-n}{\sum_{k=1}^m \lambda_k^{-2}} \right)^{1/2}. \end{aligned}$$

4. Summary

In Section 3, we connect four extremal problems on minimax with some approximation quantities which have the same value. Now we summarize all considerations as in the following corollary.

Corollary 3. *Let* $1 \leqslant p \leqslant 2$. *Then*

$$\begin{aligned} & a_n^{\mathrm{avg}}(T : \ell_p \mapsto \ell_2) = a_n^{\mathrm{ran}}(T : \ell_p \mapsto \ell_2) = \sigma_n(T : \ell_2 \mapsto \ell_2) \\ & \quad = d_n(T : \ell_1 \mapsto \ell_2) = a_n(T : \ell_1 \mapsto \ell_2) = a_n(T : \ell_2 \mapsto \ell_\infty) \\ & \quad = d^n(T : \ell_2 \mapsto \ell_\infty) = i_n(T : \ell_2 \mapsto \ell_\infty) \\ & \quad = \max_{m>n} \left(\frac{m-n}{\sum_{k=1}^m \lambda_k^{-2}} \right)^{1/2}. \end{aligned}$$

Acknowledgments

The authors would like to thank the anonymous referees and Professors P. Mathé, E. Novak, A. Pinkus for their valuable comments, remarks and suggestions which greatly helped us to improve the presentation of our paper and make it more readable.

References

[1] S. Dahkle, E. Novak, W. Sickel, Optimal approximation of elliptic problems by linear and nonlinear mappings I, J. Complexity 22 (2006) 29–49.
[2] R.D. DeVore, Nonlinear approximation, Acta Numerica 7 (1998) 51–150.
[3] S. Heinrich, Lower bounds for the complexity of Monte Carlo function approximation, J. Complexity 8 (1992) 277–300.
[4] C.V. Hutton, J.S. Morrell, J.R. Retherford, Diagonal operators, approximation numbers, and Kolmogorov diameters, J. Approx. Theory 16 (1976) 48–80.
[5] A.N. Kolmogorov, A.A. Petrov, Yu.M. Smirnov, A formula of Gauss in the theory of the method of least squares, Izv. Acad. Nauk SSSR Ser. Mat. 11 (1947) 561–566.
[6] A.I. Maltsev, A remark on the paper of A.N. Kolmogorov, A.A. Petrov, and Yu.M. Smirnov, A formula of Gauss in the theory of the method of least squares, Izv. Acad. Nauk SSSR Ser. Mat. 11 (1947) 567–568.
[7] P. Mathé, S-numbers in information-based complexity, J. Complexity 6 (1990) 41–66.
[8] P. Mathé, Random approximation of Sobolev embeddings, J. Complexity 7 (1991) 261–281.
[9] P. Mathé, A minimax principle for the optimal error of Monte Carlo methods, Constr. Approx. 9 (1993) 23–29.
[10] P. Mathé, Approximation Theory of Stochastic Numerical Methods, Habilitationsschrift, Fachbereich Mathematik, Freie Universität Berlin, 1994.
[11] E. Novak, Optimal linear randomized methods of linear operators in Hilbert spaces, J. Complexity 8 (1992) 22–36.
[12] K.Y. Osipenko, Optimal Recovery of Analytic Functions, Nova Science Publishers, New York, 2000.
[13] A. Pietsch, Operator Ideals, North-Holland, Amsterdam, 1980.
[14] A. Pinkus, n-Widths in Approximation Theory, Springer, New York, 1985.
[15] L.B. Sofman, Diameters of octahedra, Mat. Zametki 5 (1969) 429–436; English translation Math. Notes 5, 258–262.
[16] L.B. Sofman, Diameters of an infinite-dimensional octahedra, Vestn. Mowscow Univ. Mat 28 (1973) 54–56; English translation Moscow Univ. Math Bull. 28, 45–47.
[17] S.B. Stechkin, On the best approximation of given classes by any polynomials, Uspekhi Mat. Nauka 9 (1954) 133–134.
[18] S.B. Stechkin, On the absolute convergence of orthogonal series, Dokl. Akad. Nauk SSSUR 102 (1955) 37–40 (in Russian).

[19] A.I. Stepanets, Approximation characteristics of spaces S^p_φ, Ukr. Mat. Zh. 53 (2001) 392–416; English translation Ukr. Math. J. 53 (2001) 446–475.
[20] A.I. Stepanets, Approximation characteristics of the space S^p_φ in different metrics, Ukr. Mat. Zh. 8 (2001) 1121–1146; English translation Ukr. Math. J. 53 (2001) 1340–1374.
[21] J.F. Traub, G.W. Wasilkowski, H. Woźniakowski, Information-Based Complexity, Academic Press, New York, 1988.
[22] N.N. Vakhania, Probability Distributions on Linear Spaces, Elsevier, North-Holland, New York, 1981.
[23] Yongsheng Sun, Some remarks on the paper approximative characteristics of the space S^p_φ of Stepanets, 2000, preprint.