J. Math. Anal. Appl. 424 (2015) 685–695

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

On greedy algorithm approximating Kolmogorov widths in Banach spaces

P. Wojtaszczyk [1]

Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, 02-838 Warszawa, ul. Prosta 69, Poland

ARTICLE INFO

Article history:
Received 11 March 2014
Available online 22 November 2014
Submitted by D. Khavinson

Keywords:
Non-linear approximation
Greedy algorithm
Reduced basis method
Kolmogorov widths

ABSTRACT

The greedy algorithm to produce n-dimensional subspaces X_n to approximate a compact set $\mathcal{F}$ contained in a Hilbert space was introduced in the context of reduced basis method in [12,13]. The same algorithm works for a general Banach space and in this context was studied in [4]. In this paper we study the case $\mathcal{F} \subset L_p$. If Kolmogorov diameters $d_n(\mathcal{F})$ of $\mathcal{F}$ decay as $n^{-\alpha}$ we give an almost optimal estimate for the decay of $\sigma_n := \mathrm{dist}(\mathcal{F}, X_n)$. We also give some direct estimates of the form $\sigma_n \leq C_n d_n(\mathcal{F})$.

1. Introduction

Let X be a Banach space with norm $\|\cdot\| := \|\cdot\|_X$, and let $\mathcal{F}$ be one of its compact subsets. For notational convenience only, unless stated otherwise we shall assume that the elements f of $\mathcal{F}$ satisfy $\|f\|_X \leq 1$. We consider the following greedy algorithm for generating approximation spaces for $\mathcal{F}$. We first choose a function f_0 such that

$$\|f_0\| = \max_{f\in\mathcal{F}} \|f\|. \tag{1}$$

Assuming $\{f_0, \dots, f_{n-1}\}$ and $V_n := \mathrm{span}\{f_0, \dots, f_{n-1}\}$ have been selected, we then take $f_n \in \mathcal{F}$ such that

$$\mathrm{dist}(f_n, V_n)_X = \max_{f\in\mathcal{F}} \mathrm{dist}(f, V_n)_X, \tag{2}$$

and define

$$\sigma_n := \sigma_n(\mathcal{F}; X) := \mathrm{dist}(f_n, V_n)_X := \sup_{f\in\mathcal{F}} \inf_{g\in V_n} \|f - g\|. \tag{3}$$

E-mail address: wojtaszczyk@icm.edu.pl.

[1] The author was partially supported by Polish Narodowe Centrum Nauki grant DEC2011/03/B/ST1/04902.

http://dx.doi.org/10.1016/j.jmaa.2014.11.054
0022-247X/© 2014 Published by Elsevier Inc.

This greedy algorithm was introduced, for the case X is a Hilbert space, in the reduced basis method [12,13] for solving a family of PDEs. This algorithm and certain variants of it, known as *weak greedy algorithms*, are now numerically implemented with great success in the reduced basis method applied for various problems, see e.g. [3,6,5].

The study of this algorithm in the context of a general Banach space was carried out in [4]. In this paper we continue this line.

We feel that this greedy algorithm and our results can be of interest outside the scope of reduced basis methods. First there is a substantial interest in subspace approximation of large data sets in high dimensional spaces (see e.g. [8] and the references given there). In some cases this reduces to finding an almost optimal subspace for Kolmogorov width, which is the problem we deal with in the present paper. Second, in many cases an almost optimal subspace for Kolmogorov width is given by a random choice (see e.g. [11]). Having a constructive alternative may be of some theoretical interest.

We are interested in how well the space V_n approximates the elements of $\mathcal{F}$ and for this purpose we compare its performance with the best possible performance which is given by the Kolmogorov width $d_n(\mathcal{F}; X)$ of $\mathcal{F}$ defined for $n = 0, 1, 2, \ldots$ by

$$d_n := d_n(\mathcal{F}) := d_n(\mathcal{F}; X) := \inf_Y \sup_{f \in \mathcal{F}} \operatorname{dist}(f, Y)_X, \tag{4}$$

where the infimum is taken over all n dimensional subspaces Y of X. Note that if $\mathcal{F} \subset X \subset Y$ it may happen that $d_n(\mathcal{F}; Y) < d_n(\mathcal{F}; X)$. We refer the reader to [11] for a general discussion of Kolmogorov widths.

Of course, if $(\sigma_n)_{n \geq 0}$ decays at a rate comparable to $(d_n)_{n \geq 0}$, this would mean that the greedy selection provides essentially the best possible accuracy attainable by n-dimensional subspaces. Various comparisons have been given between σ_n and d_n. A first result in this direction, in the case that X is a Hilbert space $\mathcal{H}$, was given in [3] and improved in [2] where it was proved that

$$\sigma_n(\mathcal{F}; \mathcal{H}) \leq 2^{n+1} 3^{-1/2} d_n(\mathcal{F}; \mathcal{H}). \tag{5}$$

While this is an interesting comparison, it is only useful if $d_n(\mathcal{F})_{\mathcal{H}}$ decays to zero faster than 2^{-n}. Other estimates of this type were given in [2] in the Hilbert space setting and in [4] for general Banach spaces. It was shown in [2] that if $d_n(\mathcal{F}; \mathcal{H}) \leq C n^{-\alpha}$, $n = 1, 2, \ldots$, then

$$\sigma_n(\mathcal{F}; \mathcal{H}) \leq C'_\alpha n^{-\alpha}. \tag{6}$$

It was shown in [4] that if $d_n(\mathcal{F}; X) \leq C n^{-\alpha}$, $n = 1, 2, \ldots$, then for any $\epsilon > 0$

$$\sigma_n(\mathcal{F}; X) \leq C(\alpha, \epsilon) n^{-\alpha + \frac{1}{2} + \epsilon}. \tag{7}$$

A related results are known for sub-exponential decay, $d_n(\mathcal{F}) \leq C e^{-cn^\alpha}$.

The main aim of this paper is to explain the gap between (6) and (7) and provide an intermediate estimate for $X = L_p$. This is done in Theorem 2.3 and Corollary 2.5 below. We show

Corollary 2.5. *Let $\mathcal{F}$ be a compact subset of the unit ball of an L_p space $1 \leq p \leq \infty$. If $d_n(\mathcal{F}; L_p) \leq C_0 n^{-\alpha}$ for some $\alpha > \mu =: |\frac{1}{p} - \frac{1}{2}|$ then*

$$\sigma_n(\mathcal{F}; L_p) \leq C \left(\frac{\ln(n+2)}{n} \right)^\alpha n^\mu.$$

In this paper we will use a standard Banach space notation as explained for example in [17,1]. Let us make precise the notation dist already used above; for two subsets A, B of a Banach space X we denote

$$\operatorname{dist}(A, B) = \operatorname{dist}(A, B)_{\|\cdot\|} = \sup_{a \in A} \Big(\inf_{b \in B} \|a - b\|_X \Big). \tag{8}$$

2. Estimates for L_p-spaces

In this section we prove Theorem 2.3 which is our main technical result. The main tool, the Hadamard inequality, is the same as in [4] but the whole argument is quite involved.

General Banach spaces. Let us recall some well known notions from the Banach space theory, for details see e.g. [17,1]. For two Banach spaces X and Y we consider their Banach–Mazur distance:

$$d(X,Y) = \inf\{\|T\| \cdot \|T^{-1}\| : T : X \to Y \text{ is a 1–1 and onto linear map}\} \tag{9}$$

if the spaces are isomorphic and $d(X,Y) = \infty$ for non-isomorphic spaces. The name 'distance' is somewhat misleading (but very well established in tradition) since for Banach spaces X, Y, Z we have $d(X,Y) \leq d(X,Z)d(Z,Y)$ so actually $\log d(\cdot,\cdot)$ satisfies the triangle inequality.

For any pair of finite dimensional isomorphic (i.e. of the same dimension) Banach spaces, by compactness there exists $T : X \to Y$ such that $d(X,Y) = \|T\| \cdot \|T^{-1}\|$. We can additionally assume that $\|T^{-1}\| = 1$. Using such an operator we can define a new norm $\|\cdot\|_n$ on X by $\|x\|_n = \|Tx\|_Y$. Clearly it is a norm and T is an isometry between $(X, \|\cdot\|_n)$ and $(Y, \|\cdot\|_Y)$. Also

$$\|x\|_X \leq \|x\|_n \leq d(X,Y)\|x\|_X \tag{10}$$

because $\|x\|_X = \|T^{-1}Tx\|_X \leq \|Tx\|_Y = \|x\|_n \leq \|T\|\|x\|_X = d(X,Y)\|x\|_X$.

We will need also the notion of a quotient space. Let X be a Banach space and $F \subset X$ its closed subspace. For $x \in X$ its coset is

$$[x]_F =: \{x + F\} =: \{y \in X : y = x + f \text{ for some } f \in F\}.$$

Two cosets are either disjoint or equal. One can check that $[x]_F = [y]_F$ if and only if $x - y \in F$. The set of all cosets with natural operations, i.e. $\lambda[x]_F + \mu[y]_F = [\lambda x + \mu y]_F$, is a linear space denoted as X/F. The norm on X induces a natural norm on X/F given by

$$\big\|[x]_F\big\|_\sim =: \inf\{\|z\| : z \in [x]_F\} = \inf_{f\in F}\|x - f\|.$$

In our situation $(X/F, \|\cdot\|_\sim)$ is a Banach space which is called a quotient space. There is a natural map (quotient map) $q : X \to X/F$ defined as $q(x) = [x]_F$.

For a fixed (infinite dimensional) Banach space X we introduce a sequence of numbers

$$\gamma_n(X) = \sup\{d(V, \ell_2^n) : V \text{ is an } n\text{-dimensional subspace of } X\}.$$

The sequence $\gamma_n(X)$ is non-decreasing and $\gamma_1(X) = 1$. Also if $X_1 \subset X$ is a closed subspace, then $\gamma_n(X_1) \leq \gamma_n(X)$. It is known that for any n-dimensional space V we have $d(V, \ell_2^n) \leq \sqrt{n}$ (see e.g. [17, III.B.9]) so for any Banach space X, $\gamma_n(X) \leq \sqrt{n}$ for $n = 1, 2, \dots$ and it is also known that for any $L_p(\mu)$ space, $1 \leq p \leq \infty$, we have $\gamma_n \leq n^{|\frac{1}{2}-\frac{1}{p}|}$ (see e.g. [17, III.B.9]) and the order is correct.

Let us define a related concept

$$\tilde{\gamma}_n(X) = \sup\{\gamma_n(Z) : Z \text{ is a quotient space of } X\}. \tag{11}$$

Clearly $\gamma_n(X) \leq \tilde{\gamma}_n(X) \leq \sqrt{n}$. In the future we will need the following fact which is an easy (and probably known to specialists) consequence of classical but highly non-trivial results in Banach space theory.

Theorem 2.1. *There exists a constant C_p, $1 \leq p \leq \infty$, such that for every subspace of L_p we have $\tilde{\gamma}_n(X) \leq C_p n^{|\frac{1}{2}-\frac{1}{p}|}$.*

Proof. If $p = 1$ or $p = \infty$ the general estimate gives $\tilde{\gamma}_n(X) \leq \sqrt{n}$, i.e. we can take $C_1 = C_\infty = 1$. The case $1 < p < \infty$ is much more involved. We use the notion of type and cotype of Banach spaces. For definition and explanation we refer to [17] or [9]. It is known that any subspace of L_p has type $\min(2, p)$ and cotype $\max(2, p)$ (see [17, III.A.17]). Now let X be a subspace of L_p and let X/V be its quotient space. Then clearly $X/V \subset L_p/V$. By duality $(L_p/V)^*$ is a subspace of L_q with $\frac{1}{p} + \frac{1}{q} = 1$. This implies that $(L_p/V)^*$ has type $r = \min(2, q)$ and cotype $s = \max(2, q)$. Using duality between type and cotype (for explanation and references see [9, pp. 52–53]) we get that (L_p/V) and so X/V has cotype r^* with $\frac{1}{r} + \frac{1}{r^*} = 1$ and type s^* with $\frac{1}{s} + \frac{1}{s^*} = 1$. Now from Proposition 1.4 [16] we infer that $\tilde{\gamma}_n(X) \leq C_p n^{\frac{1}{s^*} - \frac{1}{r^*}}$. This gives the claim. $\square$

We know that $C_1, C_2, C_\infty \leq 1$; it seems to be an open question if $C_p \leq 1$ for other p's.

Basic calculations are summarised in the following

Proposition 2.2. *Let $\mathcal{K}$ be a subset of a Banach space $\mathcal{X}$ and let* $\sup\{\|f\| : f \in \mathcal{K}\} = B$. *Let $(\sigma_j)_{j=0}$ be given by the greedy algorithm applied to $\mathcal{K}$. Then for each $0 < m < N$ we have*

$$\left(\prod_{j=0}^{N} \sigma_j\right)^{\frac{1}{N+1}} \leq \sqrt{2}\gamma_{N+m+1}(\mathcal{X}) B^{\frac{m}{N}} d_m(\mathcal{K}; \mathcal{X})^{\frac{N-m}{N}}. \tag{12}$$

Proof. If we run the greedy algorithm for the set $\mathcal{K}$ we get sequence $f_0, f_1, \ldots, f_N$. Let $X = \operatorname{span}\{f_0, f_1, \ldots, f_N\}$. For $m < N + 1$ let us fix an arbitrary number $d > d_m(\mathcal{K}, \mathcal{X})$ and a subspace $T_m \subset \mathcal{X}$ which gives $\operatorname{dist}(\mathcal{K}, T_m) \leq d$. In particular $\|f_j - g_j\| \leq d$ for $j = 0, 1, \ldots, N$ and some $g_j \in T_m$.

We take $Y = \operatorname{span}(X, T_m)$, it has dimension $\leq N + 1 + m$ and X is a subspace of codimension $\leq m$ in Y. From (10) we infer that there exists a Euclidean norm $\|\cdot\|_e$ on Y such that $\|y\| \leq \|y\|_e \leq A\|y\|$ where $A \leq \gamma_{\dim Y}(\mathcal{X})$. Let Q be the orthogonal projection from Y onto X (in the norm $\|\cdot\|_e$). We get

$$\big\|f_j - Q(g_j)\big\|_e = \big\|Q(f_j - g_j)\big\|_e \leq \|f_j - g_j\|_e$$

We denote $\dim Q(T_m) = k$; obviously $k \leq m$.

First we fix a Gram–Schmidt orthogonalisation $(\phi_j)_{j=0}^N$ of $f_0, f_1, \ldots, f_N$ in the norm $\|\cdot\|_e$. Writing matrix $[\phi_j(f_k)]_{j,k=0}^N$ we get a triangular matrix and on the diagonal we have

$$\operatorname{dist}_{\|\cdot\|_e}\big(f_j, \operatorname{span}(f_0, \ldots, f_{j-1})\big) \geq \operatorname{dist}_{\|\cdot\|}\big(f_j, \operatorname{span}(f_0, \ldots, f_{j-1})\big) = \sigma_j$$

Now let us fix $(x_j)_{j=0}^N$, another orthonormal basis in X, such that $\operatorname{span}(x_j)_{j=0}^{k-1} = Q(T_m)$. If we look at the vector $[x_j(f_s)]_{j=0}^N$ we get

$$\sum_{j=0}^{k-1} \big|x_j(f_s)\big|^2 \leq \|f_s\|_e^2 \tag{13}$$

and

$$\sum_{j=k}^{N} \big|x_j(f_s)\big|^2 = \operatorname{dist}_{\|\cdot\|_e}\big(f_s; Q(T_m)\big). \tag{14}$$

Let k_j denote the j-th column of the matrix $[x_j(f_s)]_{j,s=0}^N$. Using first the Hadamard inequality, second the arithmetic geometric mean inequality and next (13) and (14) we get

$$
\begin{aligned}
\big(\det\big[x_j(f_s)\big]\big)^2 &\le \prod_{j=0}^{k-1}\|k_j\|_e^2\cdot\prod_{j=k}^{N}\|k_j\|_e^2\\
&\le \left(\frac{1}{k}\sum_{j=0}^{k-1}\|k_j\|_e^2\right)^k\left(\frac{1}{N+1-k}\sum_{j=k}^{N}\|k_j\|_e^2\right)^{N+1-k}\\
&= \left(\frac{1}{k}\sum_{j=0}^{k-1}\sum_{s=0}^{N}\big|x_j(f_s)\big|^2\right)^k\left(\frac{1}{N+1-k}\sum_{j=k}^{N}\sum_{s=0}^{N}\big|x_j(f_s)\big|^2\right)^{N+1-k}\\
&\le \left(\frac{1}{k}\sum_{s=0}^{N}\|f_s\|_e^2\right)^k\cdot\left(\frac{1}{N+1-k}\sum_{s=0}^{N}\operatorname{dist}_{\|\cdot\|_e}\big(f_s,Q(T_m)\big)^2\right)^{N+1-k}\\
&\le \left(\frac{N+1}{k}NB^2A^2\right)^k\left(\frac{N+1}{N+1-k}A^2d^2\right)^{N+1-k}\\
&= \left(\frac{N+1}{k}\right)^k\left(\frac{N+1}{N+1-k}\right)^{N+1-k}B^{2k}A^{2(N+1)}d^{2(N+1-k)}
\end{aligned}
$$

Note that $|\det[x_j(f_s)]| = |\det[\phi_j(f_s)]| = \prod_{j=0}^{N}|\phi_j(f_j)| \ge \prod_{j=0}^{N}\sigma_j$ so

$$
\left(\prod_{j=0}^{N}\sigma_j\right)^2 \le \left(\frac{N+1}{k}\right)^k\left(\frac{N+1}{N+1-k}\right)^{N+1-k}B^{2k}A^{2(N+1)}d^{2(N+1-k)}.
$$

Since $x^{-x}(1-x)^{x-1}\le 2$ for $x\in[0,1]$

$$
\left(\prod_{j=0}^{N}\sigma_j\right)^{\frac{1}{N+1}} \le A\sqrt{2}B^{k/(N+1)}d^{(N+1-k)/(N+1)} = A\sqrt{2}d\left(\frac{B}{d}\right)^{k/(N+1)}.
$$

Since d is arbitrary number $> d_m(\mathcal{K},\mathcal{X})$ and $k\le m$ we get (12). □

Passage to the quotient space. Suppose we have a set $\mathcal{F}\subset\mathcal{X}$ and we run a greedy algorithm to get $f_0, f_1, \ldots, f_N, f_{N+1}, \ldots$. We fix N and put $F = \operatorname{span}\{f_0, f_1, \ldots, f_{N-1}\}$ and consider $\tilde{\mathcal{X}} = \mathcal{X}/F$ with the quotient norm which we denote as $\|\cdot\|_\sim$. We put $\tilde{\mathcal{F}} = q(\mathcal{F})\subset\tilde{\mathcal{X}}$ where q is the quotient map. For $s\ge 0$ let us define $\tilde{f}_s = q(f_{N+s})$. Then for $s = 0,1,2,\ldots$ we have (we mean $\operatorname{span}\{\tilde{f}_j\}_{j<0} = \{0\}$)

$$
\begin{aligned}
\sigma_{N+s}(\mathcal{F};\mathcal{X}) &= \operatorname{dist}\big(f_{N+s},\operatorname{span}(f_j)_{j<N+s}\big)_{\mathcal{X}}\\
&= \inf\big\{\|f_{N+s}-g\| : g\in\operatorname{span}(f_j)_{j<N+s}\big\}\\
&= \inf_{g\in\operatorname{span}(f_j)_{j=N}^{N+s-1}}\ \inf_{f\in F}\|f_{N+s}-g-f\|\\
&= \inf_{g\in\operatorname{span}(f_j)_{j=N}^{N+s-1}}\big\|q(f_{N+s}-g)\big\|_\sim\\
&= \inf_{\tilde{g}\in\operatorname{span}(\tilde{f}_j)_{j=0}^{s=1}}\|\tilde{f}_s-\tilde{g}\|_\sim\\
&= \operatorname{dist}\big(\tilde{f}_s,\operatorname{span}(\tilde{f}_j)_{j=0}^{s-1}\big)_{\mathcal{X}}.
\end{aligned}
$$

Also $\sigma_{N+s}(\mathcal{F};\mathcal{X}) = \sup_{f\in\mathcal{F}}\operatorname{dist}(f,\operatorname{span}(f_j)_{j<N+s})_{\mathcal{X}}$ so using the above reasoning we get

$$\sigma_{N+s}(\mathcal{F};\mathcal{X}) = \sup_{f\in\mathcal{F}} \operatorname{dist}\big(q(f), \operatorname{span}(\tilde{f}_j)_{j=0}^{s-1}\big)_{\tilde{\mathcal{X}}}$$
$$= \operatorname{dist}\big(\tilde{\mathcal{F}}, \operatorname{span}(\tilde{f}_j)_{j=0}^{s-1}\big)_{\tilde{\mathcal{X}}}.$$

This means that there exists a realisation of a greedy algorithm for $\tilde{\mathcal{F}}$ in $\tilde{\mathcal{X}}$ which produces vectors $\tilde{f}_0, \tilde{f}_1, \ldots$ and numbers $\tilde{\sigma}_0, \tilde{\sigma}_1, \ldots$ such that $\tilde{\sigma}_s = \sigma_{N+s}$. Also for each m we have

$$d_{(N+m)}(\mathcal{F},\mathcal{X}) \le d_m(\tilde{\mathcal{F}},\tilde{\mathcal{X}}) \le d_m(\mathcal{F},\mathcal{X}). \tag{15}$$

Thus applying Proposition 2.2 we get for $m < K$

$$\left(\prod_{j=N}^{N+K} \sigma_j(\mathcal{F};\mathcal{X})\right)^{\frac{1}{K+1}} \le \sqrt{2}\tilde{\gamma}_{K+m+1}(\mathcal{X})\sigma_N(\mathcal{F};\mathcal{X})^{m/K} d_m(\mathcal{F};\mathcal{X})^{\frac{K-m}{K}} \tag{16}$$

which implies

$$\sigma_{N+K}(\mathcal{F};\mathcal{X}) \le \sqrt{2}\tilde{\gamma}_{K+m+1}(\mathcal{X})\sigma_N(\mathcal{F};\mathcal{X})^{\delta} d_m(\mathcal{F};\mathcal{X})^{1-\delta} \tag{17}$$

where $\delta = m/K$. In particular (take $K = N > m > 0$ so $\delta = m/N$ and use that $\tilde{\gamma}$ is an increasing sequence)

$$\sigma_{2N}(\mathcal{F};\mathcal{X}) \le \sqrt{2}\tilde{\gamma}_{2N}(\mathcal{X})\sigma_N(\mathcal{F};\mathcal{X})^{\delta} d_m(\mathcal{F};\mathcal{X})^{1-\delta} \tag{18}$$

Now let us prove our general estimate.

Theorem 2.3. *Suppose that we apply the greedy algorithm to the compact set $\mathcal{F}$ contained in the unit ball of a Banach space $\mathcal{X}$ such that $\tilde{\gamma}_n(\mathcal{X}) \le Cn^{\mu}$ for some $C > 0$ and $0 < \mu \le \frac{1}{2}$.*

1. *If for $\alpha > \mu$ we have $d_n(\mathcal{F};\mathcal{X}) \le C_0 n^{-\alpha}$ then*
$$\sigma_n(\mathcal{F};\mathcal{X}) \le C_1 \left(\frac{\ln(n+2)}{n}\right)^{\alpha} n^{\mu} \tag{19}$$
for some constant $C_1 = C_1(C, C_0, \mu, \alpha)$.
2. *If for some $\alpha > 0$ we have $d_n(\mathcal{F};\mathcal{X}) \le C_0 e^{-cn^{\alpha}}$ then*
$$\sigma_n \le \tilde{C}_0 n^{\mu} e^{-c_1 n^{\alpha}} \tag{20}$$
for some $\tilde{C}_0 \le 3^{\mu}\sqrt{2}C \max(1, C_0)$ and $c_1 = c \max_{0<\delta=\frac{m}{n}<1}(1-\delta)\delta^{\alpha}$.

Proof. *(1)* Let $s > 6$ be an integer such that the sequence $(n^{-\alpha+\mu}\ln^{\alpha}(n+2))_{n\ge s}$ is decreasing. Suppose that our theorem does not hold for $C_1 = D$ such that $D \ge \max_{1\le j\le s} j^{\alpha-\mu}\ln^{-\alpha}(j+2)$. Let M be the first integer such that

$$\sigma_M > DM^{-\alpha+\mu}\ln^{\alpha}(M+2). \tag{21}$$

Since $\sigma_n \le 1$ we get $M > s \ge 6$. Either $M = 2K$ or $M = 2K+1$ and $K \ge 3$. Using monotonicity of our sequences and (18) we get

$$DM^{-\alpha+\mu}\ln^{\alpha}(M+2) < \sigma_M \le \sigma_{2K} \le \sqrt{2}C(2K)^{\mu}\sigma_K^{\delta} d_m^{1-\delta} \tag{22}$$

for any $0 < m < K$ and $\delta = m/K$. Since $2K < M$ our assumptions give

$$DM^{-\alpha+\mu}\ln^{\alpha}(M+2) < \sqrt{2}CM^{\mu}\big[DK^{-\alpha+\mu}\ln^{\alpha}(K+2)\big]^{\delta}\big(C_0 m^{-\alpha}\big)^{1-\delta}$$

so using $K \leq M \leq 3K$ we get

$$C'D^{1-\delta} < K^{\delta\mu}\big(\ln(K+2)\big)^{-\alpha(1-\delta)}\delta^{-\alpha(1-\delta)} \tag{23}$$

where $C' > (CC_0^{1-\delta}3^{\alpha}\sqrt{2}\,)^{-1}$. Now we fix δ such that $\frac{1}{2\ln K} \leq \delta \leq \frac{1}{\ln K}$ (since $K \geq 3$ such a choice is possible) and we get

$$C'D^{1-\delta} \leq \big(K^{1/\ln K}\big)^{\mu}\left(\frac{2\ln K}{\ln(K+2)}\right)^{\alpha(1-\delta)} \tag{24}$$

$$\leq e^{\mu}2^{\alpha}. \tag{25}$$

This gives an upper estimate for D which proves *(1)*. The proof of *(2)* follows the arguments from [2,4] and is sketched here for completeness. We use (12) to get

$$\sigma_n \leq \sqrt{2}C3^{\mu}n^{\mu}\inf_{0<\delta=\frac{m}{n}<1} C_0^{1-\delta}e^{-c(1-\delta)\delta^{\alpha}n^{\alpha}} \leq \tilde{C}_0 n^{\mu}e^{-c_1 n^{\alpha}}. \qquad \square$$

Remark 2.4. Standard calculation gives that $\max_{0<t<1}(1-t)t^{\alpha} = \frac{t^t}{(1+t)^{1+t}}$ is attained for $t = \frac{\alpha}{1+\alpha}$ so for big n it suffices to take c_1 slightly bigger than $\frac{c_1\alpha^{\alpha}}{(1+\alpha)^{1+\alpha}}$.

Putting together Theorems 2.3 and 2.1 we get

Corollary 2.5. *Let $\mathcal{F}$ be a compact subset of the unit ball of an L_p space $1 \leq p \leq \infty$. If $d_n(\mathcal{F}; L_p) \leq C_0 n^{-\alpha}$ for some $\alpha > \mu =: |\frac{1}{p} - \frac{1}{2}|$ then*

$$\sigma_n(\mathcal{F}; \mathcal{X}) \leq C\left(\frac{\ln(n+2)}{n}\right)^{\alpha}n^{\mu}. \tag{26}$$

Remark 2.6. Let us recall that every separable Banach space $\mathcal{X}$ satisfies $\tilde{\gamma}_n(\mathcal{X}) \leq \sqrt{n}$ so Theorem 2.3 with $\mu = \frac{1}{2}$ improves Corollary 4.2(ii) from [4]. Our estimate (19) is somewhat better than in [4].

Remark 2.7. In the case when $\tilde{\gamma}_n$ is bounded the logarithm in (19) is not needed. We follow the above proof with (21) replaced by $\sigma_M > DM^{-\alpha}$ and (since $\mu = 0$) instead of (23) we get $C'D^{1-\delta} < \delta^{-\alpha(1-\delta)}$. Now the choice $\delta \sim \frac{1}{2}$ does the job. So for $p = 2$ there is no logarithm in (26).

Let us note that $\gamma_n(X)$ bounded implies that X is *isomorphic* to a Hilbert space (see [10]). This implies that the greedy algorithm in such a space can be interpreted as a realisation of a weak greedy algorithm (with appropriate weakness constant) in a Hilbert space. Thus this remark can also be seen as a corollary of [2, Theorem 3.1].

Next we present an example which shows that the estimate (26) is optimal up to a logarithmic factor. It is a natural modification of an example given in [4].

Let us take $\mathcal{F} =: \{n^{-\alpha}e_n\}_{n=1}^{\infty} \subset \ell_q$ with $2 < q < \infty$ and $\alpha > \frac{1}{q}$. Clearly $\sigma_n(\mathcal{F})_{\ell_q} = (n+1)^{-\alpha}$. In order to estimate the Kolmogorov widths of $\mathcal{F}$ we will use the following classical result (see (7.17) from [11, Chap. 14])

$$d_n\big(B_1^N; \ell_q^N\big) \leq C(q)N^{1/q}n^{-1/2} \tag{27}$$

where B_1^N is the unit ball from ℓ_1^N.

Let us fix an integer N and $\epsilon = 2^{-\alpha N}$. For $N > k > \alpha(\frac{1}{2} + \alpha - \frac{1}{q})^{-1} =: \mu N$ we consider vectors $\mathcal{F}_k =: \{n^{-\alpha} e_n\}_{n=2^k+1}^{2^{k+1}}$. From (27) we see that there exists a subspace $F_k \subset \operatorname{span}\{e_n\}_{n=2^k+1}^{2^{k+1}}$ of dimension n_k such that $\operatorname{dist}(\mathcal{F}_k, F_k) \le \epsilon$ and

$$n_k \le C(q)^2 2^{2N\alpha} 2^{2k(\frac{1}{q}-\alpha)} + 1. \tag{28}$$

We define the space

$$V = \operatorname{span}\biggl(\bigl\{e_n : n \le 2^{\mu N}\bigr\} \cup \bigcup_{N>k>\mu N} F_k\biggr).$$

We have $\operatorname{dist}(\mathcal{F}, V) \le \epsilon$ and $\dim V \le 2^{\mu N} + \sum_{N>k>\mu N} n_k$. Using (28) we obtain

$$\begin{aligned}\sum_{N>k>\mu N} n_k &\le N + C(q) 2^{2N\alpha} \sum_{N>k>\mu N} 2^{2k(\frac{1}{q}-\alpha)} \\ &\le C(q,\alpha) 2^{2N\alpha} 2^{2\mu N(\frac{1}{q}-\alpha)} = C(q,\alpha) 2^{\mu N}\end{aligned}$$

so $\dim V \le C'(q,\alpha) 2^{\mu N} =: m(N) =: m$. Thus $d_{m(N)}(\mathcal{F}; \ell_q) \le \epsilon$. On the other hand

$$\begin{aligned}\sigma_m(\mathcal{F}; \ell_q) &= (m+1)^{-\alpha} \ge d_m \epsilon^{-1} (m+1)^{-\alpha} \ge d_m C'' 2^{N\alpha} 2^{-\mu N\alpha} \\ &= C'' d_m 2^{(\frac{1}{2}-\frac{1}{q})\mu N} \ge C'' m^{\frac{1}{2}-\frac{1}{q}} d_m.\end{aligned}$$

3. Direct estimates

Let us note that the approximating subspace given by the greedy algorithm is always spanned by elements from $\mathcal{F}$ while there is no such requirement for subspaces used to calculate Kolmogorov widths. Thus it may seem more fair to compare σ_n with the following quantities:

$$\bar{d}_n(\mathcal{F}; X) = \inf_{f_1,\dots,f_n \in \mathcal{F}} \bigl\{\operatorname{dist}(\mathcal{F}, V) : V = \operatorname{span}\{f_1, \dots, f_n\}\bigr\}. \tag{29}$$

Examples that $\bar{d}_n > d_n$ are known, cf. [14, Chap. II.1] even for convex, centrally symmetric sets.

Theorem 3.1. *The following hold:*

(i) *For any compact set $\mathcal{F}$ in any Banach space X and any $n \ge 0$, we have $\bar{d}_n(\mathcal{F}) \le (n+1) d_n(\mathcal{F})$.*
(ii) *Given any $n > 0$ and $\epsilon > 0$, there is a set $\mathcal{F}$ such that $\bar{d}_n(\mathcal{F}) \ge (n - 1 - \epsilon) d_n(\mathcal{F})$.*

For X being the Hilbert space this theorem was proved in [2]. We find it somewhat surprising that exactly the same result holds for general Banach spaces.

Proof of (i). Assume first that X is finite dimensional and that $d_n(\mathcal{F}; X) < d_{n-1}(\mathcal{F}; X)$. Let $Y \subset X$ be an n-dimensional optimal Kolmogorov subspace, i.e. $\operatorname{dist}(\mathcal{F}, Y) = d_n(\mathcal{F})$, which exists because X is finite dimensional. If $d_n = 0$ then $\mathcal{F} \subset Y$ so $\bar{d}_n(\mathcal{F}) = 0$. Now assume that $d_n(\mathcal{F}) > 0$ and fix a basis $(\lambda_1, \dots, \lambda_n)$ in Y^*. For an element $f \in X$ let $P(f)$ denote an element from Y which realises $\inf_{y \in Y} \|f - y\|$. Such an element always exists, however it may be not unique. Moreover there may not exist a continuous selection of $P(f)$, see [15, C.6.3]. For arbitrary system $\{f_1, \dots, f_n\} \subset \mathcal{F}$ and arbitrary choice of $P(f_j)$'s we consider the determinant

$$D\big(P(f_1),\dots,P(f_n)\big)=\det\big[\lambda_i P(f_j)\big]_{i,j=1}^n$$

and put $\beta=\sup D(P(f_1),\dots,P(f_n))$ where the sup is taken over all choices of f_j's and P. Since

$$\big|\lambda_j\big(P(f_s)\big)\big|\le\|\lambda_j\|\cdot\|f_s\|\le\max_{1\le j\le n}\|\lambda_j\|\sup_{f\in\mathcal{F}}\|f\|$$

we infer that $\beta<\infty$.

Now we show that if $d_n<d_{n-1}$ then $\beta>0$. Note that $\beta=0$ means that for each system $\{f_1,\dots,f_n\}\subset\mathcal{F}$ and every choice of $P(f_j)$'s vectors $P(f_1),\dots,P(f_n)$ are linearly dependent. Let $k<n$ be the biggest dimension of the space they span. Let us fix vectors $f_1,\dots,f_k\in\mathcal{F}$ such that $P(f_1),\dots,P(f_k)$ span a space $V\subset Y$ of dimension k. Now let us take arbitrary $f\in\mathcal{F}$ and arbitrary $P(f)$. Vectors $P(f_1),\dots,P(f_k),P(f)$ span V so $P(f)\in V$. This implies that $d_k\le\operatorname{dist}(\mathcal{F};V)=d_n$.

We fix elements $f_1,\dots,f_n\in\mathcal{F}$ and their best approximations $P(f_1),\dots,P(f_n)$ such that $D(P(f_1),\dots,P(f_n))\ge\beta(1-\eta)$. Clearly for any $f\in\mathcal{F}$ and its best approximation $P(f)\in Y$ and any $i=1,2,\dots,n$ we have

$$\big|\beta^{-1}D\big(P(f_1),\dots,P(f_{i-1}),P(f),P(f_{i+1}),\dots,P(f_n)\big)\big|\le 1.$$

Let $f\in\mathcal{F}$ be an element where the distance from $\mathcal{F}$ to Y is achieved. Since $\beta>0$ elements $P(f_1),\dots,P(f_n)$ form a basis in Y. We write $Pf=\sum_{k=1}^n\alpha_k P(f_k)$. Note that for $i=1,\dots,n$

$$\begin{aligned}
&D\big(P(f_1),\dots,P(f_{i-1}),P(f),P(f_{i+1}),\dots,P(f_n)\big)\\
&=\sum_{k=1}^n(-1)^{k+i}D\big(P(f_1),\dots,P(f_{i-1}),P(f_k),P(f_{i+1}),\dots,P(f_n)\big)\\
&=(-1)^{2i}\alpha_i D\big(P(f_1),\dots,P(f_n)\big)
\end{aligned}$$

so $|\alpha_i|\le(1-\eta)^{-1}$. This gives

$$\begin{aligned}
\bar{d}_n(\mathcal{F})&\le\operatorname{dist}\big(\mathcal{F},\operatorname{span}(f_i)_{i=1}^n\big)\le\left\|f-\sum_{i=1}^n\alpha_i f_i\right\|\\
&\le\big\|f-P(f)\big\|+\left\|\sum_{i=1}^n\alpha_i\big[P(f_i)-f_i\big]\right\|\le(1-\eta)^{-1}(n+1)\operatorname{dist}(\mathcal{F},Y)\\
&\le\frac{n+1}{1-\eta}d_n(\mathcal{F}).
\end{aligned}$$

Since η is an arbitrary positive number we get (i) under our additional assumptions. Now if X is infinite dimensional and $\mathcal{F}\subset X$ let us fix $\epsilon>0$. Let us take an n-dimensional subspace $V_1\subset X$ such that $\operatorname{dist}(\mathcal{F};V_1)\le(1+\epsilon)d_n(\mathcal{F};X)$, a n-dimensional subspace V_2 spanned by elements of $\mathcal{F}$ such that $\operatorname{dist}(\mathcal{F};V_2)\le(1+\epsilon)\bar{d}_n(\mathcal{F};X)$ and a finite ϵ-net $\mathcal{F}_\epsilon\subset\mathcal{F}$ which contains a basis of V_2. Let $\hat{X}=\operatorname{span}\{\mathcal{F}_\epsilon\cup V_1\cup V_2\}$. Now let $k\le n$ be the biggest integer such that $d_k(\mathcal{F}_\epsilon;\hat{X})=d_n(\mathcal{F}_\epsilon;\hat{X})$. Now we have

$$\begin{aligned}
\bar{d}_k(\mathcal{F}_\epsilon;\hat{X})&\le(k+1)d_k(\mathcal{F}_\epsilon;\hat{X})\\
&=(k+1)\inf\big\{\operatorname{dist}(\mathcal{F}_\epsilon,V):V\subset\hat{X},\ \dim V=k\big\}\\
&\le(k+1)\big(\inf\big\{\operatorname{dist}(\mathcal{F},V):V\subset\hat{X},\ \dim V=k\big\}+\epsilon\big)\\
&\le(k+1)\big(\operatorname{dist}(\mathcal{F},V_1)+\epsilon\big)\\
&\le(n+1)\big((1+\epsilon)d_n(\mathcal{F};X)+\epsilon\big).
\end{aligned}$$

On the other hand

$$\begin{aligned}\bar{d}_k(\mathcal{F}_\epsilon;\hat{X}) &= \inf\{\operatorname{dist}(\mathcal{F}_\epsilon, V) : V \text{ spanned by } k \text{ elements of } \mathcal{F}_\epsilon\} \\ &\geq \inf\{\operatorname{dist}(\mathcal{F}_\epsilon, V) : V \text{ spanned by } k \text{ elements of } \mathcal{F}\} \\ &\geq \bar{d}_k(\mathcal{F};X) - \epsilon \geq \bar{d}_n(\mathcal{F};X) - \epsilon.\end{aligned}$$

Since ϵ is arbitrary the above estimates give us (i).

Examples proving (ii) were constructed, even for X being a Hilbert space, in [2, Theorem 4.1]. □

It is worth noting that for convex symmetric sets the situation is somewhat different. Namely we have the following

Proposition 3.2. *Let $\mathcal{F}$ be a convex, centrally symmetric subset of a Banach space $\mathcal{X}$. Then*

1. *if $\mathcal{X}$ is arbitrary Banach space then we have $\bar{d}_n(\mathcal{F};\mathcal{X}) \leq (\sqrt{n}+1)d_n(\mathcal{F};\mathcal{X})$,*
2. *if $\mathcal{X} = L_p$ for some $1 < p < \infty$ then we have $\bar{d}_n(\mathcal{F};\mathcal{X}) \leq (n^{|\frac{1}{p}-\frac{1}{2}|}+1)d_n(\mathcal{F};\mathcal{X})$,*
3. *if $\mathcal{X}$ is a Hilbert space then we have $\bar{d}_n(\mathcal{F};\mathcal{X}) = d_n(\mathcal{F};\mathcal{X})$.*

Proof. For a given n and $\epsilon > 0$ we fix a subspace $X \subset \mathcal{X}$ of dimension n such that $\operatorname{dist}(\mathcal{F}, X) \leq (1+\epsilon)d_n(\mathcal{F};\mathcal{X})$. Let F be a closed span of $\mathcal{F}$ and let F_1 be a closed linear span of $F \cup X$. Let us take a projection P from F_1 onto F and let $\tilde{X} = P(X)$. Since $\mathcal{F}$ is convex and symmetric set $\bigcup_{t>0} t\mathcal{F}$ is a linear subspace dense in F. This implies that there exists a subspace $\hat{X} \subset F$ such that $\dim \hat{X} = \dim \tilde{X}$ which is spanned by elements from $\mathcal{F}$ and $\operatorname{dist}(\hat{X}, \tilde{X}) \leq \epsilon$. Thus $\operatorname{dist}(\mathcal{F}, \hat{X}) \leq \epsilon + \operatorname{dist}(\mathcal{F}, \tilde{X}) \leq \epsilon + \|P\|(1+\epsilon)d_n(\mathcal{F},\mathcal{X})$. It is clear that for Hilbert space we can have $\|P\| = 1$ and the estimates for other cases are also known (see [17, III.B.11]). □

Now we will present a general direct comparison which is a generalisation of (5) to Banach spaces. Clearly it is only useful for d_n's decaying essentially faster than exponential, but then it may give better results than Proposition 2.2.

Theorem 3.3. *For a compact set $\mathcal{F} \subset L_p$ and $n = 0, 1, 2, \ldots$ we have*

$$\sigma_n(\mathcal{F};L_p) \leq Cn^{|\frac{1}{2}-\frac{1}{p}|}2^n d_n(\mathcal{F};L_p).$$

Proof. Applying the greedy algorithm to $\mathcal{F}$ we get vectors $f_0, f_1, \ldots, f_n$. Let $Y \subset L_p$ be a subspace of dimension n which almost attains $d_n(\mathcal{F})$, i.e. $\operatorname{dist}(\mathcal{F}, Y) \leq (1+\epsilon)d_n(\mathcal{F})$. There is a projection P from L_p onto $X = \operatorname{span}\{f_0, \ldots, f_n\}$ of norm $\leq n^{|\frac{1}{2}-\frac{1}{p}|}$, see e.g. [17, Theorem III.B.10]. Let $\tilde{Y} = P(Y) \subset X$. We will assume that $\dim \tilde{Y} = n$; we can enlarge $P(Y)$ if needed. We have

$$\begin{aligned}\max_j \operatorname{dist}(f_j, \tilde{Y}) &= \max_j \inf_{y\in Y} \|f_j - Py\| \leq \|P\| \max_j \inf_{y\in Y} \|f_j - y\| \\ &\leq (1+\epsilon)\|P\|d_n(\mathcal{F}).\end{aligned}$$

We fix $\lambda_0, \ldots, \lambda_n$ functionals on X such that $\|\lambda_j\| = 1$, $\lambda_j(f_s) = 0$ for $s < j$ and $\lambda_j(f_j) = \operatorname{dist}(f_j, V_j)$; such functionals exist by the Hahn–Banach Theorem (see e.g. [7, Chap. IV, Corollary 14.13]). Note that for $s > j$ we have $|\lambda_j(f_s)| \leq \operatorname{dist}(f_s, V_j) \leq \sigma_j$. Let $(e_j)_{j=0}^n$ be vectors in X biorthogonal to $(\lambda_j)_{j=0}^n$. Let ϕ be a functional on X, $\|\phi\| = 1$ and $\mathit{ker}\,\phi = \tilde{Y}$. We have

$$\sigma_n = \operatorname{dist}(f_n, V_j) = \lambda_n(f_n)$$

and for $j = 0, 1, \ldots, n$

$$\left|\sigma_n \phi(e_j)\right| \leq \sigma_j \left|\phi(e_j)\right| = \left|\phi(\sigma_j e_j)\right|.$$

We write $\sigma_j e_j = \sum_{s=0}^{n} \gamma_s^j f_s$. For each $k = 0, 1, \ldots, n$ we have

$$\sigma_j \delta_{k,j} = \sigma_j \lambda_k(e_j) = \sum_{s=0}^{n} \lambda_k(f_s).$$

Let us consider the following matrices; Σ which is diagonal with diagonal elements $\sigma_0, \sigma_1, \ldots, \sigma_n$, $\Gamma = [\gamma_s^j]$ and $\Lambda = [\lambda_k(f_s)]$. The above relations can be written as $\Sigma = \Gamma\Lambda$, so $\Gamma = \Sigma\Lambda^{-1}$. Since Λ is lower triangular with diagonal elements σ_j and elements in the j-th column at most σ_j in absolute value we infer that $\Lambda\Sigma^{-1}$ is a lower triangular matrix with diagonal elements 1 and elements in the j-th column at most 1 in absolute value. So Γ is lower triangular, i.e. $\sigma_j e_j = \sum_{s=j}^{n} \gamma_s^j f_s$, and calculating the inverse by back substitution we get $|\lambda_s^j| \leq 2^{s-j}$. This gives

$$\left|\phi(\sigma_j e_j)\right| \leq \sum_{s=j}^{n} \left|\gamma_s^j\right| \left|\phi(f_s)\right| \leq 2^{j+1} \operatorname{dist}(\mathcal{F}, \tilde{Y}) \leq 2^{j+1} \|P\|(1+\epsilon) d_n(\mathcal{F}).$$

Since $\phi = \sum_j \phi(e_j)\lambda_j$ we get $1 = \|\phi\| \leq \sum_j |\phi(e_j)|$ so

$$\begin{aligned} \sigma_n \leq \sigma_n \sum_{j=0}^{n} \left|\phi(e_j)\right| &\leq \sum_{j=0}^{n} 2^{j+1} \|P\|(1+\epsilon) d_n(\mathcal{F}) \\ &\leq 2^{n+2} n^{|\frac{1}{2} - \frac{1}{p}|}(1+\epsilon) d_n(\mathcal{F}). \end{aligned}$$

Since $\epsilon > 0$ is arbitrary the proof is completed. □

References

[1] F. Albiac, N.J. Kalton, Topics in Banach Space Theory, Grad. Texts in Math., vol. 233, Springer-Verlag, 2006.

[2] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis method, SIAM J. Math. Anal. 43 (3) (2011) 1457–1473.

[3] A. Buffa, Y. Maday, A.T. Patera, C. Prud'homme, G. Turinici, A priori convergence of the greedy algorithm for the parametrized reduced basis, Modél. Math. Anal. Numér. 46 (2012) 595–603.

[4] R. DeVore, G. Petrova, P. Wojtaszczyk, Greedy algorithms for reduced bases in Banach spaces, Constr. Approx. 37 (3) (2013) 455–466.

[5] B. Haasdonk, J. Salomon, B. Wohlmuth, A reduced basis method for parametrized variational inequalities, SIAM J. Numer. Anal. 50 (2) (2012) 2656–2676.

[6] H. Herrero, Y. Maday, F. Pla, RB (reduced basis) for RB (Rayleigh–Bénard), Comput. Methods Appl. Mech. Engrg. 261–262 (2013) 132–141.

[7] E. Hewitt, K. Stromberg, Real and Abstract Analysis, Springer, Berlin, 1969.

[8] M.A. Iwen, F. Krahmer, Fast subspace approximation via greedy least square, arXiv:1312.1413v1 [cs.CG].

[9] W.B. Johnson, J. Lindenstrauss, Basic concepts in the geometry of Banach spaces, in: W.B. Johnson, J. Lindenstrauss (Eds.), Handbook of the Geometry of Banach Spaces, vol. I, Elsevier, Amsterdam, 2001, pp. 1–84.

[10] J.T. Joichi, Normed linear spaces equivalent to inner product spaces, Proc. Amer. Math. Soc. 17 (2) (1966) 423–426.

[11] G.G. Lorentz, M. Golitschek, Y. Makovoz, Constructive Approximation, Advanced Problems, Grundlehren Math. Wiss., vol. 304, Springer-Verlag, Berlin, 1996.

[12] Y. Maday, A.T. Patera, G. Turinici, A priori convergence theory for reduced-basis approximations of single-parametric elliptic partial differential equations, J. Sci. Comput. 17 (2002) 437–446.

[13] Y. Maday, A.T. Patera, G. Turinici, Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations, C. R. Math. Acad. Sci. Paris 335 (2002) 2289–2294.

[14] A. Pincus, n-Widths in Approximation Theory, Springer-Verlag, 1985.

[15] D. Repovš, P.V. Semenov, Continuous Selections of Multivalued Mappings, Kluwer Academic Publishers, Dordrecht, 1998.

[16] N. Tomczak-Jaegermann, Banach–Mazur Distances and Finite-Dimensional Operator Ideals, Pitman Monogr. Surveys Pure Appl. Math., vol. 38, 1989.

[17] P. Wojtaszczyk, Banach Spaces for Analysts, Cambridge Stud. Adv. Math., vol. 25, Cambridge University Press, Cambridge, 1991.