AN X-GREEDY ALGORITHM WITH WEAKNESS PARAMETERS

RYAN CAUSEY, S. J. DILWORTH, S. GOGYAN, AND DENKA KUTZAROVA

Abstract. We prove that in the space \( \left( \sum_{n=1}^{\infty} F_n \right)_{\ell^p} \), \( 1 < p < \infty \), where \( F_n \) are finite-dimensional smooth spaces, the X-Greedy Algorithm with weakness parameters converges for some special dictionaries.

1. Introduction

For background information on greedy approximation and the history of the subject we refer the reader to the recent monograph [T3]. Greedy algorithms in Hilbert space are known to have good convergence properties. The first general result in this direction was obtained by Huber [H], who proved convergence of the Pure Greedy Algorithm (PGA) in the weak topology. Jones [J] proved later that the PGA converges strongly.

There are several different ways to define greedy algorithms in a general Banach space \( X \) (see [T1, T2]). It is somewhat surprising that so few results have been obtained for the most natural generalization of the PGA to Banach spaces, namely the X-Greedy Algorithm (XGA). We give the precise definition of this algorithm in the next section.

In [DKSTW] it was shown that the XGA converges in the weak topology in uniformly convex Banach spaces with some additional properties. Livshits [L1] proved strong convergence in \( L^p(0,1) \), \( 1 < p < \infty \), for the dictionary consisting of the indicator functions of the dyadic intervals. On the other hand, strong convergence in \( L^p(0,1) \) for the

---

This material is based upon work of the second author supported by the National Science Foundation under Grant Number DMS-1361461. The third author was supported by State Committee Science MES RA, in frame of the research project N SCS13-1A313. The third author was partially supported by the Bulgarian National Science Fund under Grant DFNI/I02/10. The second and fourth authors were supported by the Workshop in Analysis and Probability at Texas A&M University in 2012 and 2013.
Haar basis dictionary is an open question for arbitrary target functions: Livshits [L2] proved strong convergence for functions belonging to certain approximation classes, and exact recovery results were obtained in [DOSZ] for functions with finite support with respect to the Haar basis. Currently there are no known strong convergence results for the XGA for arbitrary dictionaries in general Banach spaces.

In the present paper we show strong convergence of the XGA for a natural class of Banach spaces, namely $\ell_p$-sums of smooth finite-dimensional spaces, for dictionaries with a structure reflecting the finite-dimensional decomposition of the space.

2. Definitions and main result

Let $D \subset X$ be a fundamental set in $X$ (i.e., $D$ has dense linear span in $X$). We shall call $D$ a dictionary for $X$. Let us recall the definition of the XGA. For $x \in X$, let

$$\sigma_1(x) := \inf\{\|x - \lambda d\|: d \in D, \lambda \in \mathbb{R}\}$$

be the error in the best one-term approximation to $x$ from $D$. If we impose the requirement that, for each $x \in X$, the infimum in the definition of $\sigma_1(x)$ is attained by some $d \in D$ and $\lambda \in \mathbb{R}$, then we can define a sequence $(x_n)_{n=0}^{\infty}$ of residuals as follows: $x_0 = x$, and, for $n \geq 0$,

$$x_{n+1} = x_n - \lambda_{n+1}d_{n+1}$$

and

$$\|x_{n+1}\| = \sigma_1(x_n).$$

We say that the algorithm converges (strongly) if $\|x_n\| \to 0$. In this case we obtain the series representation $x = \sum_{n=1}^{\infty} \lambda_n d_n$.

To be well-defined the XGA requires the existence of a best approximation to each $x \in X$ from the set $\{\lambda d: d \in D, \lambda \in \mathbb{R}\}$. Since this is not always possible we define a more flexible algorithm using weakness parameters for which we obtain convergence results.

Fix a sequence $(t_n)_{n=1}^{\infty}$ with $t_n > 1$ for all $n$. We now define a sequence $(x_n)_{n=0}^{\infty}$ of residuals, with $x_0 = x$, and, for $n \geq 0$, $x_{n+1} = x_n - \lambda_{n+1}d_{n+1}$, for any choice of $d_{n+1} \in D$ and $\lambda_{n+1} \in \mathbb{R}$ which satisfy the following:

(i) If $\sigma_1(x_n) = 0$ then $\|x_{n+1}\| \leq \frac{1}{2}\|x_n\|;

(ii) If $\sigma_1(x_n) \neq 0$, then $\|x_{n+1}\| \leq \min\{\|x_n\|, t_{n+1}\sigma_1(x_n)\}$.

We say that the algorithm converges if $\|x_n\| \to 0$, and, just as for the XGA, we obtain a series representation $x = \sum_{n=1}^{\infty} \lambda_n d_n$ in this case. Note that if best approximations exist then we can take $\|x_{n+1}\| = \sigma_1(x_n)$.
and recover the XGA. So this algorithm is indeed a generalization of the XGA.

Now suppose that $1 < p < \infty$ and that $X = \left( \sum_{n=1}^{\infty} F_n \right)_{\ell_p}$ is an $\ell_p$-sum of smooth finite-dimensional spaces, $(F_n, \| \cdot \|)$, equipped with the norm

$$\| (z_n)_{n=1}^{\infty} \| = \left( \sum_{n=1}^{\infty} \| z_n \|^p \right)^{1/p} \quad (z_n \in F_n).$$

Recall that a Banach space $Y$ is called smooth if for all $x, y \in Y$, $\|x\| = \|y\| = 1$, we have

$$\lim_{t \to 0} \left( \frac{\|x + ty\| + \|x - ty\| - 2}{t} \right) = 0.$$

Let $D = \bigcup_{n=1}^{\infty} D_n$ be a structured dictionary for $X$ satisfying the following:

(a) $D_n \subseteq \left( \sum_{j=1}^{n} F_j \right)_{\ell_p}$ and $D_n$ is a dictionary for $\left( \sum_{j=1}^{n} F_j \right)_{\ell_p}$;

(b) $D_n \subseteq D_{n+1}$ for all $n$;

(c) $P_n(D_{n+1}) \subseteq \bigcup_{\lambda \in \mathbb{R}} \lambda \cdot D_n$ for all $n$, where $P_n((z_j)_{j=1}^{\infty}) = (z_1, z_1, \ldots, z_n, 0, 0, \ldots)$ is the natural projection of $X$ onto its subspace $\sum_{j=1}^{n} F_j$.

**Theorem 1.** The algorithm converges for every $x \in X$ provided $K_p := \sum_{m=1}^{\infty} (t_m^p - 1) < 1$.

**Proof.** We may assume that $\|x_0\| = 1$. We shall prove that there exists $N_1 \geq 1$ such that

$$\|x_{N_1}\|^p \leq \frac{2 + K_p}{3} < 1.$$

By repeated application of this fact there exist $N_1 < N_2 < \ldots$ such that

$$\|x_{N_j}\|^p \leq \left( \frac{2 + K_p}{3} \right)^j,$$
and hence $\|x_n\| \to 0$ as required. First we choose $k \geq 1$ such that
\[
\|(I - P_k)x_0\|^p \leq \frac{1}{3}(1 - K_p).
\]
Since $\left( \sum_{j=1}^{k} F_j \right)_{\ell_p}$ is a smooth finite-dimensional space, a compactness argument (see [DOSZ, Proposition 3.1]) yields $\delta > 0$ such that if $y \in \left( \sum_{j=1}^{k} F_j \right)_{\ell_p}$, then
\[
\inf \{ \|y - \lambda d\| : d \in \mathcal{D}_k, \lambda \in \mathbb{R} \} \leq (1 - \delta)\|y\|.
\]
Let $m \geq 0$. If $\sigma_1(x_m) = 0$ then
\[
\|x_{m+1}\| \leq \frac{1}{2}\|x_m\| \leq \frac{1}{2} (\text{by (i)}) and we can take $N_1 = m + 1$.
\]
So suppose $\sigma_1(x_m) > 0$ for all $m$. Then, using (ii) for the first inequality and the fact that $\sigma_1(x_m) \leq 1$ for the second, we have
\[
\|x_{m+1}\|^p \leq \inf \{ \|x_m - \lambda d\|^p : d \in \mathcal{D}_k, \lambda \in \mathbb{R} \}.
\]
Now for $d \in \mathcal{D}_k$ and $\lambda \in \mathbb{R}$, we have
\[
\|x_m - \lambda d\|^p = \|P_k(x_m - \lambda d)\|^p + \|(I - P_k)(x_m)\|^p
\]
since $\mathcal{D}_k \subseteq \sum_{j=1}^{k} F_j$.

On the other hand, since $P_k(d_{m+1}) = \lambda d$ for some $d \in \mathcal{D}_k$ and $\lambda \in \mathbb{R}$ (by (b) and (c)), we have
\[
\|P_k(x_{m+1})\| \geq \inf \{ \|P_k(x_m - \lambda d)\| : d \in \mathcal{D}_k, \lambda \in \mathbb{R} \}.
\]
Hence combining (2), (3) and (4), we get
\[
\|(I - P_k)(x_{m+1})\|^p = \|x_{m+1}\|^p - \|P_k(x_{m+1})\|^p \leq (t_{m+1}^p - 1) + \|(I - P_k)(x_m)\|^p,
\]
and hence, by iteration, for all $m \geq 1$, we have
\[
\|(I - P_k)(x_m)\|^p \leq \sum_{m=1}^{\infty} (t_{m+1}^p - 1) + \|(I - P_k)(x_0)\|^p \leq K_p + \frac{1}{3}(1 - K_p).
\]
It suffices to prove that for some $N_1 \geq 1$, $\|P_k(x_{N_1})\|^p < \frac{1}{3}(1 - K_p)$. For then
\[
\|x_{N_1}\|^p = \|P_k(x_{N_1})\|^p + \|(I - P_k)(x_{N_1})\|^p \leq \frac{1}{3}(1 - K_p) + K_p + \frac{1}{3}.
AN $X$-GREEDY ALGORITHM WITH WEAKNESS PARAMETERS

So suppose that $\|P_k(x_m)\|^p \geq \frac{1}{3}(1 - K_p)$. Then, using (1),

$$\|x_{m+1}\|^p \leq t_{m+1}^p \inf\{\|x_m - \lambda d\|^p : d \in D_k, \lambda \in \mathbb{R}\}$$

$$= t_{m+1}^p \inf\{\|P_k(x_m) - \lambda d\|^p : d \in D_k, \lambda \in \mathbb{R}\}$$

$$+ t_{m+1}^p \|(I - P_k)(x_m)\|^p$$

$$\leq t_{m+1}^p ((1 - \delta)^p \|P_k(x_m)\|^p + \|(I - P_k)(x_m)\|^p)$$

$$\leq t_{m+1}^p (\|x_m\|^p - (1 - (1 - \delta)^p)\|P_k(x_m)\|^p)$$

$$\leq t_{m+1}^p \left(1 - (1 - (1 - \delta)^p) \frac{1 - K_p}{3}\right) \|x_m\|^p$$

since $\|P_k(x_m)\|^p \geq \frac{1 - K_p}{3}$ and $\|x_m\| \leq 1$. Since $t_m \to 1$ as $m \to \infty$ there exists $N \geq 1$ such that

$$\left(\prod_{m=1}^{N} t_m^p\right) \left(1 - (1 - (1 - \delta)^p) \frac{1 - K_p}{3}\right)^N < \frac{1}{3}(1 - K_p).$$

Hence, for some $1 \leq N_1 \leq N$, we have $\|P_k(x_{N_1})\|^p < \frac{1}{3}(1 - K_p)$ as desired. \hfill \Box

**Remark 1.** The proof uses the $p$-additivity of the norm for disjointly supported vectors in $\ell_p$ in an essential way. We do not know if the result holds for more general norms, e.g. Orlicz norms.

**Corollary 1.** Let $D$ be any dictionary for $X = \left(\sum_{n=1}^{\infty} F_n\right)_{\ell_p}$. Then $D^* = \bigcup_{n=1}^{\infty} P_n(D)$ satisfies (a) – (c) (with $D_n := \bigcup_{k=1}^{n} P_k(D)$) and hence the XGA with weakness parameters converges provided $K_p < 1$.

As an application we obtain a discrete version of the main result of [L1].

**Corollary 2.** The XGA converges in $\ell_p$ $(1 < p < \infty)$ for the dyadic dictionary $H = \bigcup_{n=0}^{\infty} H_n$, where $H_n = \left\{\sum_{i=m \cdot 2^n + 1}^{(m+1)2^n} e_i : m = 0, 1, \ldots\right\}$. 

Remark 2. Note that best approximations exist for the dyadic dictionary, so the XGA is well-defined.

Proof. Let $D_n = \{ h \in H : \text{supp}(h) \subseteq [1, 2^n] \}$. Then $H = \bigcup_{n=0}^{\infty} D_n$, $D_n \subseteq D_{n+1}$, $D_n$ is a dictionary for

$$P^{2^n}(\ell_p) = \{ x \in \ell_p : \text{supp}(x) \subseteq [1, 2^n] \},$$

and $P_{2^n}(D_{n+1}) = D_n$ (here $(P_k)^{\infty}_{k=1}$ is the sequence of basis projections for the standard basis $(e_n)^{\infty}_{n=1}$ of $\ell_p$). So $H$ satisfies the assumptions of the theorem for the blocking $\sum_{n=0}^{\infty} F_n$ of $\ell_p$ corresponding to the natural projections $(P_{2^n})^{\infty}_{n=0}$, i.e. $\sum_{n=0}^{k} F_n = P_{2^k}(\ell_p)$. □

References


Department of Mathematics, University of South Carolina, Columbia, SC 29208, U.S.A.
E-mail address: causey@math.sc.edu

Department of Mathematics, University of South Carolina, Columbia, SC 29208, U.S.A.
E-mail address: dilworth@math.sc.edu

Institute of Mathematics, Armenian Academy of Sciences, 24b Marshal Baghramian Ave., Yerevan, 0019, Armenia
E-mail address: gogyan@instmath.sci.am

Institute of Mathematics, Bulgarian Academy of Sciences, Sofia, Bulgaria.
Current address: Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.
E-mail address: denka@math.uiuc.edu