Industrial Mathematics Institute

2000:13

On convergence of weak greedy algorithms

E.D. Livshitz and V.N. Temlyakov

IMI

Preprint Series

Department of Mathematics
University of South Carolina

On convergence of Weak Greedy Algorithms[1]

E.D. Livshitz and V.N.Temlyakov

Moscow State University, Moscow, Russsia//
University of South Carolina, Columbia, SC, USA

1. Introduction

This paper is devoted to investigation of Weak Greedy Algorithms (WGA) introduced in [T]. We remind some notations and definitions from the theory of greedy algorithms. Let H be a real Hilbert space with an inner product $\langle\cdot,\cdot\rangle$ and the norm $\|x\| := \langle x,x\rangle^{1/2}$. We say a set $\mathcal{D}$ of functions (elements) from H is a dictionary if each $g\in\mathcal{D}$ has norm one ($\|g\|=1$) and $\overline{\text{span}}\mathcal{D}=H$. We give now the definition of WGA (see [T]). Let a sequence $\tau=\{t_k\}_{k=1}^\infty$, $0\le t_k\le 1$, be given.

Weak Greedy Algorithm. *We define $f_0^\tau := f$. Then for each $m\ge 1$, we inductively define:*

1). $\varphi_m^\tau\in\mathcal{D}$ is any satisfying

$$|\langle f_{m-1}^\tau,\varphi_m^\tau\rangle| \ge t_m \sup_{g\in\mathcal{D}} |\langle f_{m-1}^\tau, g\rangle|;$$

2).

$$f_m^\tau := f_{m-1}^\tau - \langle f_{m-1}^\tau,\varphi_m^\tau\rangle\varphi_m^\tau;$$

3).

$$G_m^\tau(f,\mathcal{D}) := \sum_{j=1}^m \langle f_{j-1}^\tau,\varphi_j^\tau\rangle\varphi_j^\tau.$$

The following theorem has been proven in [T].

Theorem A. *Assume*

$$\sum_{k=1}^\infty \frac{t_k}{k} = \infty. \tag{1.1}$$

Then for any dictionary $\mathcal{D}$ and any $f\in H$ we have

$$\lim_{m\to\infty} \|f - G_m^\tau(f,\mathcal{D})\| = 0.$$

In Section 2 of this paper we prove the following theorem.

---

[1]This research was supported by the National Science Foundation Grant DMS 9970326 and by ONR Grant N00014-91-J1343

Theorem 1. *In the class of monotone sequences $\tau = \{t_k\}_{k=1}^\infty$, $1 \ge t_1 \ge t_2 \ge \dots \ge 0$, the condition (1.1) is necessary and sufficient for convergence of Weak Greedy Algorithm for each f and all Hilbert spaces H and dictionaries $\mathcal{D}$.*

In Section 3 we consider another particular case of sequences τ. Let $\mathcal{N} := \{n_k\}_{k=1}^\infty$, $n_1 < n_2 < \dots$, be a given subsequence of natural numbers and $0 < t \le 1$. We define

$$\tau(\mathcal{N}, t) := \{t_n \quad : \quad t_{n_k} = t \quad \text{and} \quad t_n = 0, \quad n_k < n < n_{k+1}, \quad k = 1, 2 \dots\}.$$

It is convenient for us to impose some regularity restrictions on $\mathcal{N}$. We consider the class $\mathcal{M}$ of sequences

$$\mathcal{M} := \{\{n_k\}_{k=1}^\infty \quad : \quad n_{k+1} - n_k \ge n_k - n_{k-1}; \quad n_{k+1}n_{k-1} \le n_k^2, \quad k = 2, \dots\}.$$

We prove in Section 3 the following theorem.

Theorem 2. *In the class of sequences $\tau(\mathcal{N}, t)$, $\mathcal{N} \in \mathcal{M}$, the condition*

$$\text{(I)} \qquad \sum_{k=1}^\infty \frac{(n_{k+1} - n_k)^{1/2}}{n_k} = \infty$$

is necessary and sufficient for convergence of Weak Greedy Algorithm for each f and all Hilbert spaces H and dictionaries $\mathcal{D}$.

2. Proof of Theorem 1

The sufficiency of condition (1.1) for convergence follows from Theorem A. We prove here the necessity part in Theorem 1. Let H be a Hilbert space with an orthonormal basis $\{e_j\}_{j=1}^\infty$. For two elements e_i, e_j, $i \ne j$, and for a positive number $t \le 1/3$ we define the procedure which we call "equalizer" and denote $E(e_i, e_j, t)$.

Equalizer $E(e_i, e_j, t)$**.** Denote $f_0 := e_i$ and $g_1 := \alpha_1 e_i - (1-\alpha_1^2)^{1/2} e_j$ with $\alpha_1 := t$. Then $\|g_1\| = 1$ and $\langle f_0, g_1\rangle = t$. We define the sequences $f_1, \dots, f_N$; $g_2, \dots, g_N$; $\alpha_2, \dots, \alpha_N$ inductively:

$$f_n := f_{n-1} - \langle f_{n-1}, g_n\rangle g_n; \quad g_{n+1} := \alpha_{n+1} e_i - (1 - \alpha_{n+1}^2)^{1/2} e_j$$

with α_{n+1} satisfying

$$\langle f_n, g_{n+1}\rangle = t, \quad n = 1, 2, \dots.$$

Let $f_n = a_n e_i + b_n e_j$ and $N := N_t$ be the number such that

$$a_{N-1} - b_{N-1} \ge \sqrt{2}t, \quad a_N - b_N < \sqrt{2}t.$$

Then we modify the N-th step as follows. We take $g_N := 2^{-1/2}(e_i - e_j)$ and

$$f_N = f_{N-1} - \langle f_{N-1}, g_N\rangle g_N.$$

It is clear that than $a_N = b_N$ and

$$t \le \langle f_{N-1}, g_N\rangle \le 2t.$$

We list here the following simple relations

$$a_{n+1}=a_n-t\alpha_{n+1};\quad b_{n+1}=b_n+t(1-\alpha_{n+1}^2)^{1/2},\quad n<N-1;$$

(2.1) $$a_{n+1}-b_{n+1}=a_n-b_n-t(\alpha_{n+1}+(1-\alpha_{n+1}^2)^{1/2}),\quad n<N-1;$$

$$\|f_{n+1}\|^2=\|f_n\|^2-t^2,\quad n<N-1.$$

Relation (2.1) and the inequality $1\le x+(1-x^2)^{1/2}\le 2^{1/2}$, $0\le x\le 1$, imply that

(2.2) $$N\le 1/t$$

and

$$\|f_N\|^2\ge\|f_{N-1}\|^2-4t^2\ge\|f\|^2-t-3t^2.$$

This gives for $t\le 1/3$ that

$$\|f_N\|^2\ge\|f\|^2-2t.$$

It is clear that $E(e_i,e_j,t)$ is a WGA with regard to the dictionary $e_i,g_1,g_2,\dots,g_N$ with the "weakness" parameter t.

Let $1/3\ge t_1\ge t_2\ge\dots\ge 0$ be such that

(2.3) $$\sum_{k=1}^{\infty}\frac{t_k}{k}<\epsilon$$

with $\epsilon>0$ to be chosen later. Then

$$\sum_{s=0}^{\infty}t_{2^s}<2\epsilon.$$

We define WGA and a dictionary $\mathcal{D}$ as follows. We begin with $f:=e_1$ and apply $E(e_1,e_2,t)$. After $N_{t_1}\ge 1$ steps we get $g_1^0,\dots,g_{N_{t_1}}^0$ and

$$f^1=c_1(e_1+e_2)$$

with the properties

$$\|f^1\|^2\ge\|f\|^2-2t_1;\quad (c_1)^2\le 1/2.$$

We use now $E(e_1,e_3,t_2)$ and $E(e_2,e_4,t_2)$. After $2N_{t_2}\ge 2$ steps we obtain $g_1^1,\dots,g_{2N_{t_2}}^1$ and

$$f^2=c_2(e_1+\dots+e_4)$$

with the properties

$$\|f^2\|^2\ge\|f^1\|^2-2t_2;\quad (c_2)^2\le 2^{-2}.$$

After s iterations we get

$$f^s=c_s(e_1+\dots+e_{2^s})$$

and apply $E(e_i, e_{i+2^s}, t_{2^s})$, $i = 1, 2, \dots, 2^s$. We make $2^s N_{t_{2^s}} \geq 2^s$ steps and get $g_1^s, \cdots, g^s_{2^s N_{t_{2^s}}}$ and

$$f^{s+1} = c_{s+1}(e_1 + \cdots + e_{2^{s+1}})$$

with the properties

$$\|f^{s+1}\|^2 \geq \|f\|^2 - 2t_1 - 2t_2 - \cdots - 2t_{2^s} \geq 1 - 2\sum_{s=1}^{\infty} t_{2^s} \geq 1 - 4\epsilon.$$

$$(c_{s+1})^2 \leq 2^{-s-1}.$$

Choosing $\epsilon = \frac{3}{16}$ we see that $\|f^s\| \geq 1/2$ for all s.

Thus we get that the WGA with τ satisfying (2.3) does not converge for $f = e_1$ with regard to the dictionary

$$\mathcal{D} = \bigcup_{k\in\mathbb{N}} e_k \cup \bigcup_{s\geq 0; 1\leq l\leq 2^s N_{t_{2^s}}} g_l^s.$$

We will show now how the general case

$$\sum_{k=1}^{\infty} \frac{t_k}{k} < \infty$$

can be reduced to the case (2.3). We find n such that

$$\sum_{s=n}^{\infty} t_{2^s} < \epsilon,$$

take $f = e_1 + \cdots + e_{2^n}$ and pick at the first $2^n - 1$ steps $e_1, \dots, e_{2^n-1}$ as approximating elements from the dictionary. Then we use the described above procedure with $f = e_{2^n}$ instead of e_1 with the natural change in indices.

3. Proof of Theorem 2

We consider here the case of $\tau = \{t_n\}_{n=1}^{\infty}$ of the form

$$t_{n_k} = t \quad \text{and} \quad t_n = 0, \quad n_k < n < n_{k+1}, \quad k = 1, 2 \dots, \tag{3.1}$$

for a given subsequence $n_1 < n_2 < \dots$. Theorem A implies that the Weak Greedy Algorithm with the above τ converges if

$$\sum_{k=1}^{\infty} 1/n_k = \infty.$$

Theorem 2 shows that the above condition can be replaced by a weaker one. We begin with the proof of the sufficiency part of Theorem 2.

Lemma 3.1. *In the class $\mathcal{M}$ of sequences*

$$\mathcal{M} := \{\{n_k\}_{k=1}^{\infty} \quad : \quad n_{k+1} - n_k \geq n_k - n_{k-1}; \quad n_{k+1}n_{k-1} \leq n_k^2, \quad k = 2, \dots\}$$

the following two conditions are equivalent

$$\text{(I)} \qquad \sum_{k=1}^{\infty} \frac{(n_{k+1} - n_k)^{1/2}}{n_k} = \infty;$$

$$\text{(II)} \qquad \forall \{a_j\} \in l_2 \quad \liminf_{k\to\infty} a_{n_k} \sum_{j=1}^{n_k} a_j = 0.$$

Remark 3.1. We point out here that in the proof of $(I) \Rightarrow (II)$ in Lemma 3.1 we use only the property of boundedness of n_{k+1}/n_k:

$$\text{(B)} \qquad \exists C \quad : \quad \forall k \in \mathbb{N}, \quad \frac{n_{k+1}}{n_k} \leq C.$$

Thus in the sufficiency part of Theorem 2 the assumption $\{n_k\}_{k=1}^{\infty} \in \mathcal{M}$ can be replaced by a weaker assumption (B). We note also that in the proof of $(II) \Rightarrow (I)$ in Lemma 3.1 we use only the property of convexity:

$$\text{(C)} \qquad n_{k+1} - n_k \geq n_k - n_{k-1}.$$

Proof of Lemma 3.1. Let us prove first that (I) implies (II). We will prove the following a little stronger statement than (II)

$$\text{(3.2)} \qquad \sum_{k=1}^{\infty} (n_{k+1} - n_k)^{1/2} n_k^{-1} |a_{n_k}| \sum_{j=1}^{n_k} |a_j| < \infty.$$

It is known (see [Z],Ch.1,S.9) that $\{a_j\}_{j=1}^{\infty} \in l_2$ implies that

$$\text{(3.3)} \qquad \{b_n\}_{n=1}^{\infty} \in l_2 \quad \text{with} \quad b_n := \frac{1}{n} \sum_{j=1}^{n} |a_j|.$$

We observe first that

$$\text{(3.4)} \qquad \sum_{k=1}^{\infty} (n_{k+1} - n_k) b_{n_k}^2 < \infty.$$

Indeed, for any $m > n$ we have

$$m b_m \geq n b_n$$

and for $n_k < m < n_{k+1}$ we have for $\{n_k\}_{k=1}^{\infty} \in \mathcal{M}$

$$b_{n_k} \leq \frac{n_{k+1}}{n_k} b_m \leq C b_m$$

with a constant C independent of k and m. Therefore

$$(n_{k+1}-n_k)b_{n_k}^2 \le C^2 \sum_{m=n_k}^{n_{k+1}-1} b_m^2. \tag{3.5}$$

Combining (3.3) and (3.5) we get (3.4).

We return to (3.2)

$$\sum_{k=1}^{\infty}(n_{k+1}-n_k)^{1/2} n_k^{-1}|a_{n_k}| \sum_{j=1}^{n_k}|a_j| = \sum_{k=1}^{\infty}(n_{k+1}-n_k)^{1/2}|a_{n_k}|b_{n_k} \le$$

$$(\sum_{k=1}^{\infty} a_{n_k}^2)^{1/2} (\sum_{k=1}^{\infty}(n_{k+1}-n_k)b_{n_k}^2)^{1/2} < \infty.$$

Let us prove now that (II) implies (I). Assume the contrary that for $\{n_k\}_{k=1}^{\infty} \in \mathcal{M}$ we have

$$\sum_{k=1}^{\infty}(n_{k+1}-n_k)^{1/2} n_k^{-1} < \infty. \tag{3.6}$$

We will construct a sequence $\{a_j\}_{j=1}^{\infty} \in l_2$ such that

$$\liminf_{k\to\infty} a_{n_k} \sum_{j=1}^{n_k} a_j > 0. \tag{3.7}$$

Define

$$a_{n_k} := (n_{k+1}-n_k)^{1/4} n_k^{-1/2}; \quad a_j := (n_{k+1}-n_k)^{-1/4} n_k^{-1/2},$$

for

$$j \in (n_k, n_{k+1}), \quad k = 1, 2, \dots.$$

Then (3.6) implies that $\{a_j\}_{j=1}^{\infty} \in l_2$. We have from the definition of $\{a_j\}_{j=1}^{\infty}$ that

$$\sum_{j\in[n_k,n_{k+1})} a_j \ge (n_{k+1}-n_k)^{3/4} n_k^{-1/2}.$$

Next, we obtain from here

$$\sum_{j=1}^{n_{k+1}} a_j \ge \sum_{l=1}^{k} \sum_{j\in[n_l,n_{l+1})} a_j \ge \sum_{l=1}^{k}(n_{l+1}-n_l)^{3/4} n_l^{-1/2} \ge$$

$$\sum_{l=1}^{k}(n_{k+1}-n_k)^{-1/4}(n_{l+1}-n_l) n_l^{-1/2} \ge (n_{k+1}-n_k)^{-1/4} \sum_{l=1}^{k}(n_{l+1}^{1/2} - n_l^{1/2}) =$$

$$(n_{k+1}-n_k)^{-1/4}(n_{k+1}^{1/2} - n_1^{1/2}).$$

This estimate and the definition of a_{n_k} implies (3.7). Lemma 3.1 is proved now.

Lemma 3.2. *Assume that a sequence* $\{n_k\}_{k=1}^\infty \in \mathcal{M}$ *satisfies (I). Let* $0 < t \le 1$ *and* $\tau = \{t_n\}_{n=1}^\infty$ *satisfies (3.1). Then* $\{f^\tau_{n_k-1}\}_{k=1}^\infty$ *converges.*

This lemma combined with the following simple modification of Lemma 2.1 from [T] give the sufficient part of the conclusion of Theorem 2.

Lemma 3.3. *Assume that for some* $\{n_k\}_{k=1}^\infty$

$$\sum_{k=1}^\infty t^2_{n_k} = \infty.$$

Then if $\{f^\tau_{n_k-1}\}_{k=1}^\infty$ *conveges it converges to zero.*

Proof of Lemma 3.2. This proof is similar to the corresponding arguments from [T]. We present it here for selfcompleteness of this paper. It is easy to derive from the definition of WGA the following two relations

$$f^\tau_m = f - \sum_{j=1}^m \langle f^\tau_{j-1}, \varphi^\tau_j\rangle \varphi^\tau_j, \tag{3.8}$$

$$\|f^\tau_m\|^2 = \|f\|^2 - \sum_{j=1}^m |\langle f^\tau_{j-1}, \varphi^\tau_j\rangle|^2. \tag{3.9}$$

Denote $a_j := |\langle f^\tau_{j-1}, \varphi^\tau_j\rangle|$. We get from (3.9) that

$$\sum_{j=1}^\infty a_j^2 \le \|f\|^2.$$

We take any two indecies $n < m$ and consider

$$\|f^\tau_n - f^\tau_m\|^2 = \|f^\tau_n\|^2 - \|f^\tau_m\|^2 - 2\langle f^\tau_n - f^\tau_m, f^\tau_m\rangle.$$

Denote

$$\theta^\tau_{n,m} := |\langle f^\tau_n - f^\tau_m, f^\tau_m\rangle|.$$

Using (3.8) and the definition of the WGA we get for all $n < m$ that

$$\theta^\tau_{n,m} \le \sum_{j=n+1}^m |\langle f^\tau_{j-1}, \varphi^\tau_j\rangle||\langle f^\tau_m, \varphi^\tau_j\rangle| \le \frac{a_{m+1}}{t_{m+1}} \sum_{j=1}^{m+1} a_j. \tag{3.10}$$

Specifying $n = n_l - 1$ and $m = n_k - 1$ we get from (3.10) that for any $l < k$

$$\theta^\tau_{n_l-1,n_k-1} \le t^{-1} a_{n_k} \sum_{j=1}^{n_k} a_j. \tag{3.11}$$

The relation (3.11) and Lemma 3.1 imply that

$$\varliminf_{k\to\infty} \max_{l<k} \theta^\tau_{n_l-1,n_k-1} = 0.$$

It remains to use the following simple lemma (see [T]).

Lemma 3.4. *Let in a Banach space X a sequence $\{x_n\}_{n=1}^\infty$ be given. Assume that for any k,l we have*

$$\|x_k - x_l\|^2 = y_k - y_l + \theta_{k,l},$$

with $\{y_n\}_{n=1}^\infty$ is a convergent sequence of real numbers and $\theta_{k,l}$ satisfying the property

$$\varlimsup_{l\to\infty} \max_{k<l} \theta_{k,l} = 0.$$

Then $\{x_n\}_{n=1}^\infty$ converges.

We proceed now to the necessity part of Theorem 2. We will need the following simple properties of sequences from $\mathcal{M}$. Denote

$$\Delta n_k := n_{k+1} - n_k.$$

Then by monotonicity of $\{\Delta n_k\}$ we have

$$n_{2^s} - n_{2^{s-1}} = \sum_{k=2^{s-1}}^{2^s-1} \Delta n_k \le 2^{s-1}\Delta n_{2^s} \tag{3.12}$$

and

$$n_{2^s} - n_{2^{s-1}} \ge \sum_{k=1}^{2^{s-1}-1} \Delta n_k = n_{2^{s-1}} - n_1,$$

$$n_{2^s} - n_{2^{s-1}} \ge (n_{2^s} - n_1)/2. \tag{3.13}$$

Combining (3.12) and (3.13) we get

$$\Delta n_{2^s} \ge 2^{-s}(n_{2^s} - n_1)$$

and

$$(\Delta n_{2^s})^{-1/2} \le 2^s (\Delta n_{2^s})^{1/2}(n_{2^s} - n_1)^{-1}. \tag{3.14}$$

Next, $\{(\Delta n_k)^{1/2}/n_k\}$ is a monotone sequence:

$$(\Delta n_k)^{1/2}/n_k = n_k^{-1/2}(\Delta n_k/n_k)^{1/2}, \quad n_k \uparrow, \quad \Delta n_k/n_k \downarrow .$$

Thus the following two conditions are equivalent

$$\sum_{k=1}^\infty \frac{(\Delta n_k)^{1/2}}{n_k} < \infty,$$

$$\sum_{s=0}^\infty 2^s \frac{(\Delta n_{2^s})^{1/2}}{n_{2^s}} < \infty. \tag{3.15}$$

It is clear that (3.14) and (3.15) imply

$$\sum_{s=0}^{\infty}(\Delta n_{2^s})^{-1/2} < \infty.$$

The construction of the corresponding counterexample is similar to that from Section 2. We assume that

$$\sum_{s=0}^{\infty}(\Delta n_{2^s})^{-1/2} < \epsilon$$

with small enough ϵ, say, $\epsilon = 3/16$. Define a new sequence $\tau' := \{t'_n\}_{n=1}^{\infty}$ with

$$t'_n = (\Delta n_{2^s})^{-1/2}, \quad n \in [2^s, 2^{s+1}), \quad s = 0, 1, \dots .$$

Then $t'_n \downarrow 0$ and the WGA from Section 2 with τ' and $f := e_1$ reduces the square of the norm of f by at most 2ϵ. We modify now the above WGA. At each step n_k we replace WGA by Pure Greedy Algorithm (PGA) what means that at each step n_k we throw away a term $c_l e_j$ with, say, the biggest j. Let us estimate how much do we reduce the square of the norm of f in this way. After s iterations we have

$$f^s = c_s(e_1 + \dots + e_{N_s}), \quad N_s \le 2^s, \quad (c_s)^2 \le 2^{-s}.$$

Relation (2.2) implies, that working on $(s+1)$-st iteration we will make at most N_s/t'_{2^s} steps of the algorithm. It is easy to see that after s iterations we have made at least 2^s steps. Therefore we will use PGA at most

$$\frac{N_s}{t'_{2^s}\Delta n_{2^s}} + 1 \le \frac{2^s}{(\Delta n_{2^s})^{1/2}} + 1$$

times during $(s+1)$-st iteration. Thus the norm $\|\cdot\|^2$ will be reduced by at most

$$\frac{1}{(\Delta n_{2^s})^{1/2}}(1 + 2^{-s})$$

what gives the total reduction due to PGA steps at most 2ϵ. This reduction combined with the reduction of WGA at other steps sums up to at most 4ϵ. Choosing ϵ small enough (say $\epsilon = 3/16$) we get divergence of the defined above WGA.

Remark 3.2. In the proof of the necessity part of Theorem 2 we have used the convexity property (C) and also the monotonicity of $\{(\Delta n_k)^{1/2}/n_k\}$ what is weaker than the assumption $\{n_k\}_{k=1}^{\infty} \in \mathcal{M}$.

References

[T] V.N. Temlyakov, *Weak Greedy Algorithms*, Advances in computational Mathematics **12** (2000), 213–227.

[Z] A. Zygmund, *Trigonometric series*, University Press, Cambridge, 1959.