ISSN 0001-4346, Mathematical Notes, 2022, Vol. 112, No. 2, pp. 171–176. © Pleiades Publishing, Ltd., 2022.
Russian Text © The Author(s), 2022, published in Matematicheskie Zametki, 2022, Vol. 112, No. 2, pp. 163–169.

Conical Greedy Algorithm

M. A. Valov[1,2*]

[1] Lomonosov Moscow State University, Moscow, 119991 Russia

[2] Moscow Center for Fundamental and Applied Mathematics, Moscow, 119991 Russia

Received January 19, 2022; in final form, March 26, 2022; accepted April 20, 2022

Abstract—A weak conical greedy algorithm is introduced with respect to an arbitrary positive complete dictionary in a Hilbert space; this algorithm gives an approximation of an arbitrary space element by a combination of dictionary elements with nonnegative coefficients. The convergence of this algorithm is proved and an estimate of the convergence rate for the elements of the convex hull of the dictionary is given.

DOI: 10.1134/S0001434622070203

Keywords: *greedy algorithm, cone, convergence, dictionary, approximation.*

Let H be a real Hilbert space with norm $\|\cdot\|$ and inner product $\langle\cdot\,,\cdot\rangle$. The set C is called a *cone* in H if, for any $v, w \in C$ and $\lambda_{1,2} \geq 0$, the inclusion $\lambda_1 v + \lambda_2 w \in C$ is valid.

For any set $C \subset H$, its *polar cone* is defined as follows:

$$C^0 = \{w \in H : \langle v, w\rangle \leq 0 \ \forall\, v \in C\}.$$

Let $\rho(x, A) = \inf\{\|x - v\| : v \in A\}$ be the distance from the point $x \in H$ to the set $A \subset H$, and let $P_A x = \{y \in A : \|x - y\| = \rho(x, A)\}$ be the metric projection of a point $x \in H$ to a set $A \subset H$.

It is well known that, in a Hilbert space, the metric projection on a closed convex set is a singleton. Metric projections on mutually polar cones are related by the following statement.

Theorem A (Moro [1]). *Let C be a closed convex cone in H, and let C^0 be its polar cone, $x, y, z \in H$. Then the following conditions are equivalent*:

1) $z = x + y$, $x \in C$, $y \in C^0$ *and* $\langle x, y\rangle = 0$;

2) $x = P_C z$ *and* $y = P_{C^0} z$.

A subset D of the unit sphere $S(H)$ is called a *dictionary* if linear combinations of D elements are dense in H. If linear combinations of D elements with nonnegative coefficients are dense in H, then we will call D a *positive complete dictionary*.

For each $x = x_0 \in H$ and dictionary $D \subset S(H)$, the *pure greedy algorithm* inductively defines the subsequence

$$x_{n+1} = x_n - \langle x_n, g_n\rangle g_n, \qquad n = 0, 1, \dots,$$

where the element $g_n \in D$ is chosen so that

$$\langle x_n, g_n\rangle = \max\{|\langle x_n, g\rangle| : g \in D\}. \qquad (*)$$

For each $x = x_0 \in H$, the *orthogonal greedy algorithm* with respect to the dictionary D defines the subsequence

$$x_{n+1} = x - P_{\operatorname{span}\{g_0,\dots,g_n\}} x, \qquad n = 0, 1, \dots,$$

where the element $g_n \in D$ is also selected from condition $(*)$.

*E-mail: mv01t@yandex.ru

In both algorithms, the maximum attainability condition $\max\{|\langle x, g\rangle| : g \in D\}$ for each $x \in H$ is an additional condition imposed on the dictionary.

A greedy algorithm is said to *converge* if the residuals x_n converge to 0 in norm as $n \to \infty$. It is known [2, Chap. 2] that pure greedy and orthogonal greedy algorithms converge for any dictionary that satisfies the existence condition for a maximum and, for any initial element $x \in H$, estimates of the convergence rate are also known.

The problem of approximating an element of a Hilbert space by linear combinations of elements of a dictionary with nonnegative coefficients was considered by Livshits in [3], [4]. He proposed a special recursive greedy algorithm providing such approximations. A natural generalization of the pure greedy algorithm is a positive greedy algorithm in which the next element g_n is chosen from the maximization condition for the inner product $\langle x_n, g\rangle$, not its module, but this algorithm may diverge for a positive complete dictionary [5].

In this paper, we propose a natural modification of the orthogonal greedy algorithm, namely, the conical greedy algorithm approximating an arbitrary element of the space by a linear combination of dictionary elements with nonnegative coefficients. Its convergence was proved for each positive complete dictionary D and any initial element; also the convergence rate for elements of special form was estimated.

Let $\operatorname{cone}\{g_0, g_1, \dots, g_n\}$ denote the minimal (with respect to inclusion) cone containing elements $g_0, g_1, \dots, g_n$, i.e., the following set

$$\left\{x \in H : x = \sum_{k=0}^{n} \lambda_k g_k, \lambda_k \ge 0\right\}.$$

For each $x = x_0 \in H$ and any positive complete dictionary $D \subset S(H)$, the *conical greedy algorithm* inductively defines the sequence

$$x_{n+1} = x - P_{\operatorname{cone}\{g_0,\dots,g_n\}} x, \qquad n = 0, 1, \dots,$$

where the element $g_n \in D$ is chosen so that

$$\langle x_n, g_n\rangle = \max\{\langle x_n, g\rangle : g \in D\}.$$

At the same time, the requirement is also imposed on the dictionary that $\max\{\langle x, g\rangle : g \in D\}$ exist for all $x \in H$.

We also define the more general *weak conical greedy algorithm*. Let us fix the sequence $\{t_n\}_{n=0}^{\infty}$, $0 < t_n \le 1$. For each $x = x_0 \in H$, the *weak conical greedy algorithm with weakness parameters* $\{t_n\}_{n=0}^{\infty}$ inductively defines the sequence

$$x_{n+1} = x - P_{\operatorname{cone}\{g_0,\dots,g_n\}} x, \qquad n = 0, 1, \dots,$$

where the element $g_n \in D$ is chosen to satisfy the condition

$$\langle x_n, g_n\rangle \ge t_n \sup\{\langle x_n, g\rangle : g \in D\}.$$

For $t_n < 1$, the weak conical greedy algorithm works for any positive complete dictionary.

Theorem 1. *Let $D \subset S(H)$ be a positive complete dictionary in H. Then the weak conical greedy algorithm with weakness parameters $\{t_n\}_{n=0}^{\infty}$ converges for any initial element $x \in H$ if*

$$\sum_{k=0}^{\infty} t_k^2 = \infty. \tag{1}$$

This theorem is analogous to Theorem 2.1 from [2, Chap. 2] on the convergence of a weak orthogonal greedy algorithm.

Proof. Let $C_n = \operatorname{cone}\{g_0, \dots, g_n\}$. By Theorem A, we have

$$x_n = x - P_{C_{n-1}} x = P_{C_{n-1}^0} x,$$

and $C_0^0 \supset C_1^0 \supset C_2^0 \supset \cdots$. We will need the following result.

Lemma 1. *Let $K_1 \supset K_2 \supset K_3 \supset \cdots$ be decreasing (with respect to inclusion) closed cones in H, and let $x \in H$. Then the sequence $\{x_n = P_{K_n}x\}$ is fundamental.*

Proof. For an arbitrary $y \in K_n$, by Theorem A, we have $\langle x - x_n, y\rangle \le 0$ and $\langle x - x_n, x_n\rangle = 0$. Therefore,

$$\begin{aligned}\|x-y\|^2 &= \|x-x_n\|^2 + \|x_n - y\|^2 + 2\langle x - x_n, x_n - y\rangle \\ &= \|x-x_n\|^2 + \|x_n - y\|^2 - 2\langle x - x_n, y\rangle \ge \|x-x_n\|^2 + \|x_n-y\|^2.\end{aligned}$$

Substituting $y = x_m$, $m > n$, we obtain

$$\|x_n - x_m\|^2 \le \rho(x, K_m)^2 - \rho(x, K_n)^2;$$

since the sequence $\rho(x, K_n)$ is nondecreasing and bounded ($\rho(x, K_n) \le \|x\|$), it follows that x_n is fundamental. The lemma is proved. □

By Lemma 1, we have $x_n \to z$ for some $z \in H$.

If $z = 0$, then the theorem is proved.

If $z \neq 0$, then the positive completeness condition for the dictionary implies that there exists an element $g \in D$, such that $\langle z, g\rangle > \delta > 0$. Therefore, there exists an m such that, for each $n \ge m$, the inequality $\langle x_n, g\rangle > \delta$ holds. The further proof requires the following technical lemmas.

Lemma 2. *Let $C \subset H$ be a closed convex cone, and let $x, y \in H$. Then*

$$\|P_C(x+y)\| \le \|P_C x\| + \|P_C y\|.$$

Proof. By Theorem A, for any $z \in H$, we have

$$\|P_C z\| = \|z - P_{C^0} z\| = \rho(z, C^0).$$

Therefore,

$$\begin{aligned}\|P_C(x+y)\| &= \rho(x+y, C^0) = \inf_{w\in C^0} \|x + y - w\| = \inf_{v,u\in C^0} \|x+y-(v+u)\| \\ &= \inf_{v,u\in C^0} \|(x-v)+(y-u)\| \le \inf_{v,u\in C^0} (\|x-v\| + \|y-u\|) \\ &= \inf_{v\in C^0} \|x-v\| + \inf_{u\in C^0} \|y-u\| = \rho(x, C^0) + \rho(y, C^0) \\ &= \|P_C(x)\| + \|P_C(y)\|.\end{aligned}$$

The lemma is proved. □

Lemma 3. *Let $C_1, C_2 \subset H$ be closed convex cones, and let $C_1 \subset C_2$. Then*

$$\|P_{C_1}(z)\| \le \|P_{C_1} P_{C_2}(z)\|$$

for each $z \in H$.

Proof. By Theorem A, we have $z = P_{C_2} z + P_{C_2^0} z$. Also note that if $C_1 \subset C_2$, then $C_1^0 \supset C_2^0$, which means $P_{C_1} v = 0$ for any v from C_2^0.

Applying Lemma 2, we obtain

$$\|P_{C_1} z\| = \|P_{C_1}(P_{C_2} z + P_{C_2^0} z)\| \le \|P_{C_1} P_{C_2} z\| + \|P_{C_1} P_{C_2^0} z\| = \|P_{C_1} P_{C_2} z\|.$$

The lemma is proved. □

Let us return to the proof of the theorem.

Applying Lemma 3 and taking n large enough, we obtain a contradiction, namely,

$$\begin{aligned}0 \le \|x_{n+1}\|^2 &= \|P_{C_n^0} x\|^2 \le \|P_{C_n^0} P_{C_{n-1}^0} x\|^2 \\ &= \|P_{C_n^0} x_n\|^2 = \|x_n\|^2 - \|P_{C_n} x_n\|^2 \le \|x_n\|^2 - \langle g_n, x_n\rangle^2 \\ &\le \|x_n\|^2 - t_n^2\delta^2 \le \cdots \le \|x_m\|^2 - \delta^2 \sum_{k=m}^{n} t_k^2 < 0,\end{aligned}$$

because the series $\sum_{k=0}^{\infty} t_k^2$ diverges. Theorem 1 is proved. □

Let us show that equality (1) is also necessary for the convergence of the weak conical greedy algorithm for any positive complete dictionary D.

Consider the space ℓ_2 with orthonormal basis $\{e, e_0, e_1, \dots\}$. Let

$$\sum_{k=0}^{\infty} t_k^2 < \infty.$$

Consider the weak conical greedy algorithm for the element

$$x_0 = e + \sum_{k=0}^{\infty} t_k e_k$$

with respect to the symmetric positive complete dictionary $D = \{\pm e, \pm e_0, \pm e_1, \dots\}$.

It is easy to prove by induction that, for the next g_n in the weak conical greedy algorithm, we can take the element $g_n = e_n$. In this case, the current residual can be expressed as $x_n = e + \sum_{k=n}^{\infty} t_k e_k$, and the algorithm does not converge.

Theorem 2. *Let*

$$x_0 \in A_1^+(M, D) = \overline{\left\{ \sum_{k=0}^{N} \lambda_k g_k : g_k \in D,\ N \in \mathbb{N},\ \lambda_k \ge 0,\ \sum_{k=0}^{N} \lambda_k \le M \right\}}.$$

Then, for the sequence $\{x_n\}$ of residuals of the weak conical greedy algorithm with weakness parameters $\{t_n\}_{n=0}^{\infty}$, the following inequalities hold:

$$\|x_n\| \le \frac{M}{\sqrt{1 + \sum_{k=0}^{n-1} t_k^2}}, \qquad n = 1, 2, \dots . \tag{2}$$

This assertion is an analogue of Theorem 2.20 from [2, Chap. 2] on the convergence rate of a weak orthogonal greedy algorithm for elements of the convex hull of a symmetric dictionary.

Proof. Let $C_n = \operatorname{cone}\{g_0, \dots, g_n\}$. As already noted, we have $C_{n-1}^0 \supset C_n^0$ for each $n \ge 1$. Using Lemma 3, we obtain

$$\begin{aligned}\|x_{n+1}\|^2 &= \|P_{C_n^0} x\|^2 \le \|P_{C_n^0} P_{C_{n-1}^0} x\|^2 = \|P_{C_n^0} x_n\|^2 = \|x_n\|^2 - \|P_{C_n} x_n\|^2 \\ &\le \|x_n\|^2 - \|\langle g_n, x_n\rangle g_n\|^2 = \|x_n\|^2 - \langle g_n, x_n\rangle^2 \le \|x_n\|^2 - t_n^2 \Big(\sup_{g\in D} \langle g, x_n\rangle\Big)^2.\end{aligned}$$

Lemma 4. *Let D be a positive complete dictionary, and let $x \in A_1^+(D, M)$. Then, for each $z \in H$ such that $\langle x - z, z\rangle = 0$, the following inequality holds*:

$$\sup_{g\in D} \langle z, g\rangle \ge \frac{\|z\|^2}{M}.$$

Proof. It suffices to prove the lemma for

$$x = \sum_{k=0}^{N} \lambda_k g_k, \qquad \text{where} \quad g_0, \dots, g_N \in D, \quad \lambda_k \geq 0, \quad \sum_{k=0}^{N} \lambda_k \leq M.$$

We have

$$\begin{aligned} \|z\|^2 = \langle z, z \rangle = \langle x, z \rangle - \langle x - z, z \rangle = \langle x, z \rangle = \sum_{k=0}^{N} \lambda_k \langle g_k, z \rangle \\ \leq \sup_{g \in D} \langle g, z \rangle \sum_{k=0}^{N} \lambda_k \leq M \sup_{g \in D} \langle g, z \rangle. \end{aligned}$$

The lemma is proved. □

Applying Lemma 4 to $z = x_n = P_{C^0_{n-1}} x_0$, we obtain

$$\begin{aligned} \|x_{n+1}\|^2 \leq \|x_n\|^2 - t_n^2 \Big(\sup_{g \in D} \langle g, x_n \rangle \Big)^2 \\ \leq \|x_n\|^2 - t_n^2 \frac{\|x_n\|^4}{M^2} = \|x_n\|^2 \bigg(1 - \frac{t_n^2 \|x_n\|^2}{M^2} \bigg). \end{aligned}$$

Now we need the following numerical lemma.

Lemma A [6]. *Let $\{c_n\}_{n=0}^{\infty}$ be a sequence such that*

$$c_0 \leq A, \qquad c_n \geq 0, \qquad c_{n+1} \leq c_n \bigg(1 - \frac{\alpha_n c_n}{A} \bigg), \quad n = 0, 1, 2, \dots,$$

for some sequence $\{\alpha_n\}_{n=0}^{\infty}$ of positive numbers and some number $A > 0$. Then

$$c_n \leq \frac{A}{1 + \sum_{k=0}^{n-1} \alpha_k}, \qquad n = 1, 2, \dots .$$

Obviously, $0 \leq \|x_n\| \leq \|x_0\| \leq M$ for each natural n. This allows us to apply Lemma A to the sequences

$$c_n = \|x_n\|^2, \qquad \alpha_n = t_n^2, \qquad A = M^2,$$

so that we obtain

$$\|x_n\|^2 \leq \frac{M^2}{1 + \sum_{k=0}^{n-1} t_k^2}, \qquad n = 1, 2, \dots .$$

Theorem 2 is proved. □

For the conical greedy algorithm in Theorem 2, we obtain an estimate for the norm of the residuals $\|x_n\| \leq M(n+1)^{-1/2}$ and the exponent $-1/2$ in this estimate is sharp.

Indeed, taking $H = \ell_2$, $D = \{\pm e_0, \pm e_1, \pm e_2, \dots\}$ and

$$x_0 = \sum_{k=0}^{\infty} \frac{1}{(k+1)^{(1+\varepsilon)}} e_k \in A_1(D, M),$$

where

$$M = \sum_{k=0}^{\infty} \frac{1}{(k+1)^{(1+\varepsilon)}}$$

for an arbitrary $\varepsilon > 0$, we obtain

$$\|x_n\| = \sqrt{\sum_{k=n}^{\infty} \frac{1}{(k+1)^{(2+2\varepsilon)}}} \geq \frac{1}{\sqrt{1+2\varepsilon}\,(n+1)^{1/2+\varepsilon}}\,.$$

Note that, for initial elements from $A_1^+(1, D)$, there exists a so-called incremental algorithm having convergence rate of the same order [2, Chap. 6, Sec. 6].

ACKNOWLEDGMENTS

The author wishes to extend gratitude to P. A. Borodin for posing the problem and valuable comments.

REFERENCES

1. J. J. Moreau, "Décomposition orthogonale d'un espace hilbertien selon deux cónes mutuellement polaires," C. R. Acad. Sci. Paris **255**, 238–240 (1962).
2. V. N. Temlyakov, *Greedy Approximation* (Cambridge Univ. Press, Cambridge, 2011).
3. E. D. Livshits, "On the recursive greedy algorithm," Izv. Math. **70** (1), 87–108 (2006).
4. E. D. Livshits, "On n-term approximation with positive coefficients," Math. Notes **82** (3), 332–340 (2007).
5. P. A. Borodin, "Example of divergence of a greedy algorithm with respect to an asymmetric dictionary," Math. Notes **109** (3), 379–385 (2021).
6. V. N. Temlyakov, "Weak greedy algorithms," Adv. Comput. Math. **12**, 213–227 (2000).