arXiv:2304.06423v1 [math.NA] 13 Apr 2023

贪心算法的收敛速度

V.N. Temlyakov*

摘要

我们证明了关于贪心算法收敛速度的一些结果，这些算法能够给出展开式。我们同时考虑希尔伯特空间的情形以及更一般的巴拿赫空间的情形。本文的新要点在于，我们通过两个范数的乘积来界定逼近误差——即 f 的范数与 A_1-范数的乘积，其中f。通常只使用 A_1-范数 f。特别地，我们证明，在这种新的收敛速度意义下，一些贪婪算法（纯贪婪算法（PGA）及其推广）与正交贪婪算法（OGA）同样有效；而众所周知，在标准意义下，PGA 的性能远逊于 OGA。

1 引言

我们先对这个问题作一般性描述。设 X 是一个范数为 $\|\cdot\|_X$ 的巴拿赫空间， $Y \subset X$ 是 X 的一个子空间，其范数 $\|f\|_Y \geq \|f\|_X$更强， $f \in Y$。考虑一个齐次逼近算子（线性或非线性） $G\colon Y \to X$, $G(af) = aG(f)$, $f \in Y$, $a \in \mathbb{R}$，以及逼近误差

$$e(B_Y, G)_X := \sup_{f \in B_Y} \|f - G(f)\|_X, \quad B_Y := \{f : \|f\|_Y \leq 1\}.$$

*俄罗斯科学院斯捷克洛夫数学研究所，俄罗斯莫斯科；罗蒙诺索夫莫斯科国立大学；莫斯科基础与应用数学中心；南卡罗来纳大学。

那么，对于任意 $f \in Y$，我们有

$$\|f - G(f)\|_X \le e(B_Y, G)_X \|f\|_Y. \tag{1.1}$$

特征 $e(B_Y, G)_X$ 在逼近理论中发挥着重要作用，许多经典空间 X 和 Y 都涉及这一特征；例如，$X = L_p$ 和 Y 是类似索оболев、尼古拉斯基或 Besov 空间的光滑性空间。

本文关注以下版本的不等式 (1.1)：求最佳的 $\gamma(\alpha, G, X, Y)$，使得该不等式

$$\|f - G(f)\|_X \le \gamma(\alpha, G, X, Y) \|f\|_X^{1-\alpha} \|f\|_Y^{\alpha}, \quad \alpha \in [0, 1], \tag{1.2}$$

对所有 $f \in Y$ 都成立。显然，$\gamma(1, G, X, Y) = e(B_Y, G)_X$。此外，显然，在假设 $\|f - G(f)\|_X \le \|f\|_X$, $f \in Y$ 下，我们得到平凡的界

$$\gamma(\alpha, G, X, Y) \le e(B_Y, G)_X^{\alpha}.$$

本文讨论相对于给定字典的贪婪逼近，并证明了关于 $\gamma(\alpha, G, X, Y)$ 的一些非平凡不等式，既考虑 X 为希尔伯特空间的情形，也考虑 X 为巴拿赫空间的情形。特别地，我们证明，某些贪心算法（纯贪婪算法（PGA）及其推广）在不等式 (1.2) 的意义下与正交贪婪算法（OGA）同样有效；然而，众所周知，在不等式 (1.1) 的意义下，PGA 的效果远不如 OGA（定义和精确表述见下文）。

设 H 是一个实希尔伯特空间，其内积为 $\langle \cdot, \cdot \rangle$，范数为 $\|\cdot\|$。若元素（函数）集合 $\mathcal{D}$ 来自 H，且其中每个 $g \in \mathcal{D}$ 的范数均为 1（$\|g\| = 1$），并且张成 $\mathcal{D} = H$，则称其为字典（对称字典）。此外，为方便起见，我们假设其具有对称性：

$$g \in \mathcal{D} \quad \text{implies} \quad -g \in \mathcal{D}.$$

我们定义纯贪婪算法（PGA）。我们针对一般字典 $\mathcal{D}$ 描述此算法。如果 $f \in H$，我们令 $g(f) \in \mathcal{D}$ 为 $\mathcal{D}$ 中使 $\langle f, g \rangle$ 最大化的元素。为简便起见，我们假设这样的最大化元素存在；若不存在，则需要对下文的算法进行适当修改（见下文的弱贪婪算法）。我们定义

$$G(f, \mathcal{D}) := \langle f, g(f) \rangle g(f) \quad \text{and} \quad R(f, \mathcal{D}) := f - G(f, \mathcal{D}).$$

纯贪婪算法（PGA）。我们定义 $f_0 := f$ 和 $G_0(f,\mathcal{D}) := 0$。然后，对于每个 $m \geq 1$，我们归纳地定义

$$G_m(f,\mathcal{D}) := G_{m-1}(f,\mathcal{D}) + G(f_{m-1},\mathcal{D})$$

$$f_m := f - G_m(f,\mathcal{D}) = R(f_{m-1},\mathcal{D}).$$

N请注意，对于给定的元素 f，序列 $\{G_m(f,\mathcal{D})\}$ 可能并不唯一。

从收敛性和收敛速度的角度来看，这一算法已经得到了充分研究。读者可以在 [6], 第 2 章中找到相应的结果及历史评论。在本文中，我们关注收敛速度。通常，在逼近理论中，我们针对特定的类定义收敛速度。在经典逼近理论中，这些类是光滑性类。显然，在具有任意 H 和 $\mathcal{D}$ 的一般设置下，我们没有类似于经典函数光滑性的光滑性概念。事实证明，由几何方式定义的类，即 $\mathcal{D}$的凸包的闭包（记作 $A_1(\mathcal{D})$），是一个非常自然的类。对于每个 $f \in H$，我们定义如下范数

$$\|f\|_{A_1(\mathcal{D})} := \inf\{M : f/M \in A_1(\mathcal{D})\}.$$

显然，$\|f\| \leq \|f\|_{A_1(\mathcal{D})}$。于是，PGA 的收敛速度问题可以表述如下（见 [6], 第 95 页）。求序列的衰减阶

$$\gamma_m(H) := \sup_{\{G_m(f,\mathcal{D})\},f,\mathcal{D}} \frac{\|f - G_m(f,\mathcal{D})\|}{\|f\|_{A_1(\mathcal{D})}},$$

其中，上确界取遍所有可能的 $\{G_m(f,\mathcal{D})\}$ 选择、所有元素 $f \in H$, $f \neq 0$, $\|f\|_{A_1(\mathcal{D})} < \infty$，以及所有字典 $\mathcal{D}$。这是希尔伯特空间中贪婪逼近的核心理论问题，目前仍未解决。我们在此提及一些已知结果；关于该问题的详细历史，请读者参见 [6], Ch.2。显然，对于任意 $f \in H$，若满足 $\|f\|_{A_1(\mathcal{D})} < \infty$，则有

$$\|f - G_m(f,\mathcal{D})\| \leq \gamma_m(H)\|f\|_{A_1(\mathcal{D})}.$$

本文讨论渐近特征 γ_m(H 的如下扩展：对于 $\alpha \in (0,1]$ 定义

$$\gamma_m(\alpha,H) := \sup_{\{G_m(f,\mathcal{D})\},f,\mathcal{D}} \frac{\|f - G_m(f,\mathcal{D})\|}{\|f\|^{1-\alpha}\|f\|^{\alpha}_{A_1(\mathcal{D})}}.$$

显然,

$$\gamma_m(1,H)=\gamma_m(H),\quad \gamma_m(\alpha,H)\geq\gamma_m(\beta,H)\quad\text{if}\quad\alpha\leq\beta. \tag{1.3}$$

$\gamma_m(H)$ 的第一个上界是在 [2] 中得到的

$$\gamma_m(H)\leq m^{-1/6}.$$

实际上， [2] 中的证明（另见 [6], 第92–93页）给出了

$$\gamma_m(1/3,H)\leq m^{-1/6}.$$

我们在此建立以下界

$$\frac{1}{2}m^{-\alpha/2}\leq\gamma_m(\alpha,H)\leq m^{-\alpha/2},\quad\alpha\leq 1/3. \tag{1.4}$$

此外，我们还将在第2节中找到渐近特征的正确行为，它类似于 $\gamma_m(\alpha,H)$; 研究对象是比 PGA 更一般的算法，即参数为 b 的弱贪婪算法。

比较 PGA 与正交贪婪算法（OGA）的收敛速率很有意思。现在我们简要给出 OGA 的定义。我们定义 $f_0^o:=f$, $G_0^o(f,\mathcal{D})=0$，并对 $m\geq 1$ 归纳地定义 $G_m^o(f,\mathcal{D})$：它是 f 在由 $g(f_0^o)$, ...,$g(f_{m-1}^o)$ 张成的空间上的正交投影，并令 $f_m^o:=f-G_m^o(f,\mathcal{D})$。OGA 中与特征 $\gamma_m(H)$ 和 $\gamma_m(\alpha,H)$ 对应的量分别记为 $\gamma_m^o(H)$ 和 $\gamma_m^o(\alpha,H)$。以下界在 [2] 中得到证明（另见 [6], 第93页）

$$\gamma_m^o(H)\leq m^{-1/2}. \tag{1.5}$$

已知（见 [4]）， $\gamma_m(H)$ 比 $m^{-0.1898}$ 衰减得更慢。因此，从特征 $\gamma_m(H)$ 和 $\gamma_m^o(H)$ 的角度来看，OGA 远优于 PGA。我们在此建立以下界

$$\frac{1}{2}m^{-\alpha/2}\leq\gamma_m^o(\alpha,H)\leq m^{-\alpha/2},\quad\alpha\leq 1. \tag{1.6}$$

这意味着，从特征 $\gamma_m(\alpha,H)$ 和$\gamma_m^o(\alpha,H)$ 的角度来看，OGA 对于 $\alpha\leq 1/3$ 与 PGA 相同。这是一个非常令人惊讶的事实。

我们不知道 (1.4) 中的上界是否对 $\alpha > 1/3$ 成立。不过，(1.3) 中的不等式以及 $\gamma_m(H)$ 的下界表明

$$\gamma_m(\alpha, H) \geq \gamma_m(H) \geq cm^{-0.1898}.$$

T因此，式 (1.4) 中的上界无法扩展到 α_0 之外：$= 0.3796$。

第 3 节讨论巴拿赫空间 X 的情形。巴拿赫空间情形的结果与希尔伯特空间的结果相似，但不如后者的对应结果精确。

2 希尔伯特空间。带参数 b 的弱贪婪算法

给定一个序列 $\tau = \{t_k\}_{k=1}^{\infty}$、$0 \leq t_k \leq 1$ 和一个参数 $b \in (0, 1]$。我们定义带参数 b 的弱贪婪算法。

带参数 b 的弱贪婪算法（WGA(τ, b)）我们定义 $f_0 := f^{\tau,\ 0}{}^{b} := f$。于是，对每个 $m \geq 1$，我们归纳地定义：

1) $\varphi_m := \varphi^{\tau,\ m}{}^{b} \in \mathcal{D}$ 是满足以下条件的任意对象

$$\langle f_{m-1}, \varphi_m \rangle \geq t_m \sup_{g \in \mathcal{D}} \langle f_{m-1}, g \rangle;$$

2)

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

3)

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

在 $t_k = t$、$k = 1$、2、..... 的情况下，我们使用 t 的记号，而不是 τ。

下面讨论收敛速度。以下定理 2.1 已在 [5] 中得到证明。

定理 2.1. 设 $\mathcal{D}$ 是 H 中的任意字典。假设 $\tau := \{t_k\}_{k=1}^{\infty}$ 是一个非递增序列，并且 $b \in (0, 1]$。那么对于 $f \in A_1(\mathcal{D})$，我们有

$$\|f - G_m^{\tau,b}(f, \mathcal{D})\| \leq e_m(\tau, b) \tag{2.1}$$

其中

$$e_m(\tau, b) := \left(1 + b(2-b) \sum_{k=1}^{m} t_k^2\right)^{-\frac{(2-b)t_m}{2(2+(2-b)t_m)}}. \tag{2.2}$$

定理 2.1 对任意 f 和任意 $\mathcal{D}$ 蕴含以下不等式

$$\frac{\|f - G_m^{\tau,b}(f,\mathcal{D})\|}{\|f\|_{A_1(\mathcal{D})}} \le e_m(\tau,b). \tag{2.3}$$

现在我们推广定理 2.1，以给出一个界

$$\gamma_m^{t,b}(\alpha,H) := \sup_{\mathcal{D}} \sup_{f} \sup_{G_m^{t,b}(f,\mathcal{D})} \frac{\|f - G_m^{t,b}(f,\mathcal{D})\|}{\|f\|^{1-\alpha}\|f\|_{A_1(\mathcal{D})}^{\alpha}}. \tag{2.4}$$

我们在 $t_k = t$、 $k = 1$、2、.....的情形下证明以下定理 2.2。

定理 2.2。对于任意希尔伯特空间 H ，我们有

$$\gamma_m^{t,b}(\alpha,H) \le (1 + mb(2-b)t)^{-\alpha/2}, \tag{2.5}$$

只要 $\alpha \le \frac{(2-b)\,t}{(2-b)\,t+2}$。

证明。该定理的证明思路与 [5] 中定理 2.1 的证明一致。设 $\mathcal{D}$ 是 H中的一个字典。我们引入一些记号:

$$f_k := f_k^{t,b}, \quad \varphi_k := \varphi_k^{t,b}, \quad k = 0,1,\dots,$$

$$a_m := \|f_m\|^2, \quad y_m := \langle f_{m-1}, \varphi_m\rangle, \quad m = 1,2,\dots,$$

并考虑如下定义的序列 $\{B_n\}$

$$B_0 := \|f\|_{A_1(\mathcal{D})}, \quad B_m := B_{m-1} + by_m, \quad m = 1,2,\dots.$$

显然, $\|f_n\|_{A_1(\mathcal{D})} \le B_n$, $n = 0$, 1, ... 根据 [DT] 中的引理 3.5（另见 [6], 第 91 页，引理 2.17），我们得到

$$\sup_{g\in\mathcal{D}} \langle f_{m-1}, g\rangle \ge \|f_{m-1}\|^2 / B_{m-1}. \tag{2.6}$$

由此以及等式

$$\|f_m\|^2 = \|f_{m-1}\|^2 - b(2-b)\langle f_{m-1}, \varphi_m\rangle^2$$

我们得到以下关系

$$a_m = a_{m-1} - b(2-b)y_m^2, \tag{2.7}$$

$$B_m = B_{m-1} + by_m, \tag{2.8}$$

$$y_m \geq ta_{m-1}/B_{m-1}. \tag{2.9}$$

由 (2.7) 和 (2.9) 得到

$$a_m \leq a_{m-1}(1 - b(2-b)t^2 a_{m-1} B_{m-1}^{-2}).$$

利用 $B_{m-1} \leq B_m$，我们由此推得

$$a_m B_m^{-2} \leq a_{m-1} B_{m-1}^{-2}(1 - b(2-b)t^2 a_{m-1} B_{m-1}^{-2}). \tag{2.10}$$

我们需要下面这个简单的已知引理（例如，见 [6], 第 91 页，适用于 $C_1 = C_2$ 的情形）。

引理 2.1. 设 $\{x_m\}_{m=0}^{\infty}$ 是一个满足下列不等式的非负数序列

$$x_0 \leq C_1, \quad x_{m+1} \leq x_m(1 - x_m C_2), \quad m = 0, 1, 2, \ldots, \quad C_1, C_2 > 0.$$

那么，对每个 m，我们都有

$$x_m \leq (C_1^{-1} + C_2 m)^{-1}.$$

证明. 证明采用关于 m 的归纳法。根据假设，$m = 0$ 时命题成立。假设 $x_m \leq (C_1^{-1} + C_2 m)^{-1}$，并证明 $x_{m+1} \leq (C_1^{-1} + C_2(m+1))^{-1}$。若 $x_{m+1} = 0$，则该命题显然成立。因此假设 $x_{m+1} > 0$。于是我们有

$$x_{m+1}^{-1} \geq x_m^{-1}(1 - x_m C_2)^{-1} \geq x_m^{-1}(1 + x_m C_2) = x_m^{-1} + C_2 \geq C_1^{-1} + (m+1)C_2,$$

这意味着 $x_{m+1} \leq (C_1^{-1} + C_2(m+1))^{-1}$。 □

我们将引理 2.1 应用于 $x_m := a_m B_m^{-2}$。那么，不等式 $\|f\| \leq \|f\|_{A_1(\mathcal{D})}$ 表明我们可以取 $C_1 = 1$。令 $C_2 = b(2-b)t^2$，由 (2.10) 和引理 2.1 得到

$$a_m B_m^{-2} \leq (1 + mb(2-b)t^2)^{-1}. \tag{2.11}$$

关系式 (2.7) 和 (2.9) 意味着

$$a_m \leq a_{m-1} - b(2-b)y_m t a_{m-1}/B_{m-1} = a_{m-1}(1 - b(2-b)t y_m/B_{m-1}). \tag{2.12}$$

现在我们需要以下简单的不等式：对于任意 $x<1$ 和任意 $a>0$，有

$$(1-x)(1+x/a)^a \le 1. \tag{2.13}$$

将 (2.8) 改写为

$$B_m = B_{m-1}(1+by_m/B_{m-1}) \tag{2.14}$$

并对 $x=b(2-b)ty_m/B_{m-1}$ 和 $a=(2-b)t$使用不等式 (2.13)，由 (2.12) 和 (2.14) 可得

$$a_m B_m^{(2-b)t} \le a_{m-1}B_{m-1}^{(2-b)t} \le \cdots \le \|f\|^2\|f\|_{A_1(\mathcal{D})}^{(2-b)t}. \tag{2.15}$$

结合 (2.11) 和 (2.15)，我们得到

$$a_m^{(2-b)t+2} \le \|f\|^4\|f\|_{A_1(\mathcal{D})}^{2(2-b)t}(1+mb(2-b)t^2)^{-(2-b)t},$$

这就完成了定理 2.2 在 $\alpha_0 := {}_{(2-b)t}\ {}_{(2-b)t+2}$ 情形下的证明。~~情形~~ $\alpha \le \alpha_0$ 由下面的引理 2.2 得出。 □

下界。设 H 是一个无限维希尔伯特空间，$\{e_k\}_{k=1}^{\infty}$ 是 H中的一个正交归一系统。假设我们的对称字典 $\mathcal{D}$ 由 $\pm e_k$、$k=1$、2、.....组成，且其他元素 $g\in\mathcal{D}$ 具有性质 $\langle g$、$e_k\rangle=0$、$k=1$、2、.....。我们在 $t=1$ 的情形下给出一个例子。设 $b\in(0,1]$ 和 m 给定。我们考虑两种情形：(I) $b\le 1/4$ 和 (II)$b\in(1/4,1]$。

(I). 设 $m' := [2bm]+1$ 和

$$f=\sum_{k=1}^{m'} e_k.$$

那么在每次迭代中，WGA(1, b) 都会选取系数最大的 e_k、$k\in[1,m']$之一。经过 m 次迭代后，我们将得到

$$f_m=\sum_{k=1}^{m'} c_k e_k, \qquad c_k \ge 1-(m/m'+1)b \ge 1/4.$$

因此，我们得到

$$\|f_m\| \ge (m')^{1/2}/4, \quad \|f\|=(m')^{1/2}, \quad \|f\|_{A_1(\mathcal{D})} \le m'.$$

因此，对于任意 $\alpha \in [0,1]$，我们有

$$\frac{\|f_m\|}{\|f\|^{1-\alpha}\|f\|^{\alpha}_{A_1(\mathcal{D})}} \geq (m')^{-\alpha/2}/4. \tag{2.16}$$

（II）集合

$$f = \sum_{k=1}^{2m} e_k.$$

Then

$$f_m = \sum_{k=1}^{2m} c_k e_k, \quad c_k = 1,\, k \in G, \quad c_k = 1-b,\, k \notin G, \quad |G| = m.$$

因此，我们得到

$$\|f_m\| \geq (m)^{1/2}, \quad \|f\| = (2m)^{1/2}, \quad \|f\|_{A_1(\mathcal{D})} \leq 2m.$$

因此，对于任意 $\alpha \in [0,1]$，我们都有

$$\frac{\|f_m\|}{\|f\|^{1-\alpha}\|f\|^{\alpha}_{A_1(\mathcal{D})}} \geq 2^{-1/2}(2m)^{-\alpha/2}. \tag{2.17}$$

界(2.16)和 (2.17)表明，在 $t = 1$ 不等式(2.5) 的情形下，就对 m 和 b 的依赖性而言是紧的，当 m 趋于 ∞ 且 b 趋于 0 时。

(1.6)的证明。(1.6) 中的下界由 (2.17) 得出，因为在正交归一系统的情况下，算法 PGA 和 OGA 是一致的。现在我们证明上界。

引理 2.2. 令 $g_m(\alpha, H)$ 表示 $\gamma_m(\alpha, H)$ 或 $\gamma^o_m(\alpha, H)$。假设对于某个 $\beta \in (0,1]$，我们有

$$g_m(\beta, H) \leq C\varphi(m)^{-\beta/2}.$$

那么，对于任意 $\alpha \in (0, \beta)$，我们有

$$g_m(\alpha, H) \leq C^{\alpha/\beta}\varphi(m)^{-\alpha/2}. \tag{2.18}$$

证明。根据算法 PGA 和 OGA 的定义，我们有 $\|f_m\| \leq \|f\|$ and $\|f_m^o\| \leq \|f\|$。对于算法 PGA 和 OGA，证明完全相同。下面以 PGA 为例进行证明。根据我们的假设，对于任意 f、任意字典 $\mathcal{D}$ 以及 $G_m(f,\mathcal{D})$ 的任意一种实现，

$$\|f_m\| = \|f - G_m(f,\mathcal{D})\| \leq \|f\|^{1-\beta}\|f\|^{\beta}_{A_1(\mathcal{D})} C\varphi(m)^{-\beta/2}.$$

因此，对于任意 $a \in [0,1]$，我们有

$$\|f_m\| \leq \|f\|^{1-a}(\|f\|^{1-\beta}\|f\|^{\beta}_{A_1(\mathcal{D})} C\varphi(m)^{-\beta/2})^a. \tag{2.19}$$

取 $a = \alpha/\beta$，由 (2.19) 可得 (2.18)。 □

(1.6) 中的上界由 (1.5) 和引理 2.2 得出。

3 巴拿赫空间中的贪心展开

本节将第 2 节中的结果从希尔伯特空间推广到巴拿赫空间。我们先给出一些定义。设 X 是一个范数为 $\|\cdot\|$ 的实巴拿赫空间。与上文一样，如果来自 X 的元素（函数）集合 $\mathcal{D}$ 满足每个 $g \in \mathcal{D}$ 的范数为 1（$\|g\| = 1$），且 $\operatorname{span}\mathcal{D} = X$，我们称其为字典（~~对称~~字典）。此外，为方便起见，我们假设该字典是对称的

$$g \in \mathcal{D} \quad \text{implies} \quad -g \in \mathcal{D}.$$

本文研究关于 $\mathcal{D}$ 的、能够提供贪心展开的贪心算法。对于非零元素 $f \in X$，我们将 F_f 记为 f 的赋范（峰值）泛函：

$$\|F_f\| = 1, \qquad F_f(f) = \|f\|.$$

这种泛函的存在性由哈恩-巴拿赫定理保证。记

$$r_{\mathcal{D}}(f) := \sup_{F_f} \sup_{g\in\mathcal{D}} F_f(g).$$

我们注意到，一般而言，范数化泛函 F_f 并不唯一。这就是为什么我们在 $r_{\mathcal{D}}(f)$ 的定义中对 f 的所有范数化泛函取上确界 $\sup_{F_f}$。众所周知，在一致光滑巴拿赫空间（这是我们主要

研究的对象）中，范数化泛函 F_f 是唯一的。在这种情况下，我们不需要在 $r_{\mathcal{D}}(f)$ 的定义中取上确界 $\sup^{F_f}$。

我们在此考虑一致光滑巴拿赫空间中的逼近。对于巴拿赫空间 X，我们定义光滑模。

$$\rho(u) := \rho(u, X) := \sup_{\|x\|=\|y\|=1} (\frac{1}{2}(\|x + uy\| + \|x - uy\|) - 1).$$

一致光滑巴拿赫空间是指具有以下性质的空间

$$\lim_{u\to 0} \rho(u)/u = 0.$$

众所周知（例如参见 [1], 引理 B.1），在 $X = L_p, 1 \le p < \infty$ 的情况下，我们有

$$\rho(u, L_p) \le \begin{cases} u^p/p & \text{if} \quad 1 \le p \le 2, \\ (p-1)u^2/2 & \text{if} \quad 2 \le p < \infty. \end{cases} \tag{3.1}$$

现在我们给出 DGA(τ, b, μ)、$\tau = \{t_k\}_{k=1}^{\infty}$、$t_k \in (0, 1]$ 的定义，该定义引自 [5]（另见 [6], 第 6 章）。

带参数（τ，b，μ）的对偶贪心算法（DGA（τ，b，μ））。设 X 是一致光滑巴拿赫空间，其光滑模为ρ（u），且 μ（u）是 ρ（u）的一个主控函数：ρ（u）$\le$ μ（u），$u \in [0, \infty)$。对于序列 $\tau = \{t_k\}_{k=1}^{\infty}$，$t_k \in (0, 1]$）和参数 $b \in (0, 1]$），我们归纳地定义序列 $\{f_m\}_{m=0}^{\infty}$、$\{\varphi_m\}_{m=1}^{\infty}$、$\{c_m\}_{m=1}^{\infty}$ 和 $\{G_m\}_{m=0}^{\infty}$。设 $f_0 := f$ 且 $G_0 := 0$。若对 $m \ge 1$ $f_{m-1} = 0$ 成立，则令 $f_j = 0$（$j \ge m$）并停止。若 $f_{m-1} \ne 0$，则进行以下三个步骤：

1) 任取满足以下条件的 $\varphi_m \in \mathcal{D}$

$$F_{f_{m-1}}(\varphi_m) \ge t_m r_{\mathcal{D}}(f_{m-1}); \tag{3.2}$$

2) 从方程中选择 $c_m > 0$

$$\|f_{m-1}\| \mu(c_m/\|f_{m-1}\|) = \frac{t_m b}{2} c_m r_{\mathcal{D}}(f_{m-1}); \tag{3.3}$$

3) 定义

$$f_m := f_{m-1} - c_m \varphi_m, \qquad G_m := G_m^{\tau,b,\mu} := G_{m-1} + c_m \varphi_m. \tag{3.4}$$

除了算法 DGA(τ, b, μ) 外，我们还考虑它的一个小幅改动：在步骤 2) 中，根据方程求出 c_m （见 [5], 注 3.1）

$$\|f_{m-1}\|\mu(c_m/\|f_{m-1}\|) = \frac{b}{2}c_m F_{f_{m-1}}(\varphi_m). \tag{3.5}$$

我们将这种修改记为 DGA(τ, b, μ)*。

接下来，我们研究在光滑模具有幂型上界的一致光滑巴拿赫空间中，DGA(τ, b, μ) 的收敛速度： $\rho(u) \le \mu(u) = \gamma u^q$, $1 < q \le 2$。下面的定理 3.1 引自 [5] （另见 [6], 第372页）。

定理 3.1. 设 $\tau := \{t_k\}_{k=1}^{\infty}$ 是一个非增序列 $1 \ge t_1 \ge t_2 \cdots > 0$，且 $b \in (0,1)$。假设 X 的光滑模为$\rho(u) \le \gamma u^q$, $q \in (1, 2]$。记 $\mu(u) = \gamma u^q$。那么，对于任意字典 $\mathcal{D}$ 和任意 $f \in A_1(\mathcal{D})$，DGA(τ, b, μ) 的收敛速度由下式给出

$$\|f_m\| \le C(b, \gamma, q)\left(1 + \sum_{k=1}^{m} t_k^p\right)^{-\frac{t_m(1-b)}{p(1+t_m(1-b))}}, \quad p := \frac{q}{q-1}.$$

注 3.1. *[5],* 注 3.2 中指出，定理 3.1 对算法 DGA(τ, b, μ)* 同样成立。

定理 3.1 是定理 2.1 的一个类似结果。现在我们证明定理 2.2 的一个类似结果。我们将定理 3.1 推广，以给出一个界

$$\gamma_m^{t,b,\mu}(\alpha, X) := \sup_{\mathcal{D}} \sup_{f} \sup_{G_m^{t,b,\mu}(f,\mathcal{D})} \frac{\|f - G_m^{t,b,\mu}(f, \mathcal{D})\|}{\|f\|^{1-\alpha}\|f\|_{A_1(\mathcal{D})}^{\alpha}}. \tag{3.6}$$

算法 DGA(τ, b, μ)* 的相应特征记为 $\gamma^{t,\ b,\ m}\ ^{\mu}(\alpha, X)$*。我们在 $t_k = t$,$k = 1$, 2, . . . 的情形下证明以下定理 3.2。

定理 3.2. 对于任意具有光滑模 $\rho(u, X) \le \gamma u^q$, $1 < q \le 2$, $p := \frac{q}{q-1}$ 的巴拿赫空间 X，有

$$\gamma_m^{t,b,\mu}(\alpha, X) \le (1 + mct^p)^{-\alpha/p}, \quad c := (1-b)\left(\frac{b}{2\gamma}\right)^{\frac{1}{q-1}}, \tag{3.7}$$

给定 $\alpha \le {}^{t}\!/_{(1-b)1+t(1-b)}$。相同的不等式也适用于 $\gamma^{t,\ b,\ m}\ ^{\mu}(\alpha, X)$*。

证明。对于两种特征 $\gamma^{t,\ b,\ m}\mu(\alpha, X)$ 和 $\gamma^{t,\ b,\ m}\mu(\alpha, X)^*$，证明过程完全相同。我们以 $\gamma^{t,\ b,\ m}\mu(\alpha, X)$ 为例进行证明。根据光滑模的定义，我们有

$$\|f_{n-1} - c_n\varphi_n\| + \|f_{n-1} + c_n\varphi_n\| \leq 2\|f_{n-1}\|(1 + \rho(c_n/\|f_{n-1}\|)). \tag{3.8}$$

根据 φ_n的定义:

$$F_{f_{n-1}}(\varphi_n) \geq tr_{\mathcal{D}}(f_{n-1}) \tag{3.9}$$

我们得到

$$\begin{aligned} \|f_{n-1} + c_n\varphi_n\| &\geq F_{f_{n-1}}(f_{n-1} + c_n\varphi_n) \\ &= \|f_{n-1}\| + c_n F_{f_{n-1}}(\varphi_n) \geq \|f_{n-1}\| + c_n tr_{\mathcal{D}}(f_{n-1}). \end{aligned} \tag{3.10}$$

结合 (3.8) 和 (3.10)，我们得到

$$\|f_n\| = \|f_{n-1} - c_n\varphi_n\| \leq \|f_{n-1}\|(1 + 2\rho(c_n/\|f_{n-1}\|)) - c_n tr_{\mathcal{D}}(f_{n-1}). \tag{3.11}$$

根据 c_m 的选取，我们由此得到

$$\|f_m\| \leq \|f_{m-1}\| - t(1-b)c_m r_{\mathcal{D}}(f_{m-1}). \tag{3.12}$$

因此，我们需要从下方估计 $c_m r_{\mathcal{D}}(f_{m-1})$。显然,

$$\|f_{m-1}\|_{A_1(\mathcal{D})} = \|f - \sum_{j=1}^{m-1} c_j\varphi_j\|_{A_1(\mathcal{D})} \leq \|f\|_{A_1(\mathcal{D})} + \sum_{j=1}^{m-1} c_j. \tag{3.13}$$

记 $B_n := \|f\|_{A_1(\mathcal{D})} + \sum_{j=1}^{n} c_j$。由此，根据 (3.13) 可得

$$\|f_{m-1}\|_{A_1(\mathcal{D})} \leq B_{m-1}.$$

接下来，由 [T3] 中的引理 2.2（另见 [6], 第 343 页，引理 6.10）可得

$$\begin{aligned} r_{\mathcal{D}}(f_{m-1}) &= \sup_{g\in\mathcal{D}} F_{f_{m-1}}(g) = \sup_{\varphi\in A_1(\mathcal{D})} F_{f_{m-1}}(\varphi) \\ &\geq \|f_{m-1}\|^{-1}_{A_1(\mathcal{D})} F_{f_{m-1}}(f_{m-1}) \geq \|f_{m-1}\|/B_{m-1}. \end{aligned} \tag{3.14}$$

将 (3.14) 代入 (3.12)，得到

$$\|f_m\| \leq \|f_{m-1}\|(1 - t(1-b)c_m/B_{m-1}). \tag{3.15}$$

根据 B_m 的定义，我们发现

$$B_m = B_{m-1} + c_m = B_{m-1}(1 + c_m/B_{m-1}).$$

使用不等式

$$(1+x)^\alpha \le 1 + \alpha x, \quad 0 \le \alpha \le 1, \quad x \ge 0,$$

我们得到

$$B_m^{t(1-b)} \le B_{m-1}^{t(1-b)}(1 + t(1-b)c_m/B_{m-1}). \tag{3.16}$$

将 (3.15) 与 (3.16) 相乘可得

$$\|f_m\| B_m^{t(1-b)} \le \|f_{m-1}\| B_{m-1}^{t(1-b)} \le \cdots \le \|f\| \|f\|_{A_1(\mathcal{D})}^{t(1-b)}. \tag{3.17}$$

函数 $\mu(u)/u = \gamma u^{q-1}$ 在 $[0,\infty)$ 上递增。因此，(3.3) 中的 c_m 大于或等于（见 (3.14)）中的 c'_m。

$$\gamma \|f_{m-1}\| (c'_m/\|f_{m-1}\|)^q = \frac{tb}{2} c'_m \|f_{m-1}\|/B_{m-1}, \tag{3.18}$$

$$c'_m = \left(\frac{tb}{2\gamma}\right)^{\frac{1}{q-1}} \frac{\|f_{m-1}\|^{\frac{q}{q-1}}}{B_{m-1}^{\frac{1}{q-1}}}. \tag{3.19}$$

使用记号

$$p := \frac{q}{q-1}, \qquad c := (1-b)\left(\frac{b}{2\gamma}\right)^{\frac{1}{q-1}},$$

由 (3.12)、(3.14) 和 (3.19)，我们得到

$$\|f_m\| \le \|f_{m-1}\| \left(1 - ct^p \frac{\|f_{m-1}\|^p}{B_{m-1}^p}\right). \tag{3.20}$$

注意到 $B_m \ge B_{m-1}$，由 (3.20) 可得

$$(\|f_m\|/B_m)^p \le (\|f_{m-1}\|/B_{m-1})^p (1 - ct^p (\|f_{m-1}\|/B_{m-1})^p). \tag{3.21}$$

考虑到 $\|f\| \le \|f\|_{A_1(\mathcal{D})}$，由引理 2.1 和 $C_1 = 1$、$C_2 = ct^p$，我们从 (3.21) 得到

$$(\|f_m\|/B_m)^p \le (1 + mct^p)^{-1}. \tag{3.22}$$

结合 (3.17) 和 (3.22)，我们得到

$$\|f_m\| \le \|f\|^{1-\alpha_0}\|f\|_{A_1(\mathcal{D})}^{\alpha_0}(1+mct^p)^{-\alpha_0/p}, \quad p := \frac{q}{q-1}, \quad \alpha_0 := \frac{t(1-b)}{1+t(1-b)}.$$

这就完成了定理 3.2 在 $\alpha = \alpha_0$ 情况下的证明。$\alpha < \alpha_0$ 的情形可由 $\alpha = \alpha_0$ 的情形及引理 2.2 的相应类似结论推出。 □

我们来讨论定理 3.2 在希尔伯特空间情形下的一个应用。众所周知且易于验证，对于希尔伯特空间 H，有

$$\rho(u) \le (1+u^2)^{1/2} - 1 \le u^2/2.$$

我们来看看 DGA($t, b, u^2/2$) 在希尔伯特空间中的工作方式。考虑它的第 m 步。令 $\varphi_m \in \mathcal{D}$ 取自 (3.2)，其中 $t_m = t$ （当 $t = 1$ 时，我们假设其存在）。那么显然，对于 φ_m，我们有

$$\langle f_{m-1}, \varphi_m\rangle \ge t\|f_{m-1}\| r_{\mathcal{D}}(f_{m-1}) = t \sup_{g\in\mathcal{D}} \langle f_{m-1}, g\rangle.$$

在这一步，WGA(t, 1) 会使用系数为 $\langle f_{m-1}$ 的 φ_m，$\varphi_m\rangle$。DGA($t, b, u^2/2$)* 与 WGA(t, b) 类似，使用相同的 φ_m，但只取 $\langle f_{m-1}$ 的一部分，$\varphi_m\rangle$:

$$c_m = b\|f_{m-1}\| F_{f_{m-1}}(\varphi_m) = b\langle f_{m-1}, \varphi_m\rangle. \tag{3.23}$$

因此，(3.23) 中的选择 $b = 1$ 对应于 WGA。然而，从上述讨论可以清楚地看出，我们为一般巴拿赫空间设计的技术在 $b = 1$ 情形下并不适用。由定理 3.2，令$\mu(u) = u^2/2$，DGA(t, b, μ) 和 DGA(t, b, μ)* 可得到以下误差估计

$$\|f_m\| \le \|f\|^{1-\alpha}\|f\|_{A_1(\mathcal{D})}^{\alpha}(1+mct^2)^{-\alpha/2}, \quad \alpha \le \alpha_0 := \frac{t(1-b)}{1+t(1-b}. \tag{3.24}$$

注意，不等式 (3.24) 与由定理 2.2 推出的相应不等式类似，适用于 $\alpha \le \alpha_1$: $= {}^{t}{}_{(2-b)\ 2+t\ (2-b)}$。容易验证 $\alpha_0 < \alpha_1$，这意味着定理 2.2 给出的结果强于定理 3.2 的相应推论。

关于下界的说明。在第 2 节中，我们得到了下界，该下界对于参数 m 和 b 都是紧的。显然，最

重要的参数是 m。在这里，我们给出了关于 m 的下界，这些下界适用于任何在经过 m 次迭代后提供 m 项近似的算法。回顾一下，相对于给定字典 $\mathcal{D}$， m 项近似这一概念的定义。给定一个整数 $m \in \mathbb{N}$，我们用 $\Sigma_m(\mathcal{D})$ 表示相对于 $\mathcal{D}$ 的所有 m 项近似的集合：

$$\Sigma_m(\mathcal{D}) := \left\{ h \in X \,:\, h = \sum_{i=1}^{m} c_i g_i, \quad g_i \in \mathcal{D},\, c_i \in \mathbb{R},\, i = 1, \dots, m \right\}.$$

定义巴拿赫空间 X

$$\sigma_m(f, \mathcal{D})_X := \inf_{h \in \Sigma_m(\mathcal{D})} \|f - h\|_X$$

作为关于 $\mathcal{D}$、在 X 范数下对 $f \in X$ 的最佳 m 项逼近。

令 $1 < q \le 2$。考虑 $X = \ell_q$。已知（[3], 第 67 页）， ℓ_q,$1 < q \le 2$ 是一个一致光滑巴拿赫空间，其光滑模$\rho(u)$ 为幂型 q： $\rho(u) \le \gamma u^q$。选取 $\mathcal{D} := \mathcal{E}$ 作为对称化标准基 $\{\pm e_j\}_{j=1}^{\infty}$, $e_j := (0, \dots, 0, 1, 0, \dots)$，其中 ℓ_q。对于给定的 $m \in \mathbb{N}$，设

$$f := \sum_{i=1}^{2m} e_i.$$

那么，以下关系显而易见。

$$\|f\|_{\ell_q} = (2m)^{1/q}, \quad \|f\|_{A_1(\mathcal{E})} = 2m, \quad \sigma_m(f, \mathcal{E})_{\ell_q} = m^{1/q}.$$

因此，对于任意 $\alpha \in [0, 1]$

$$\frac{\sigma_m(f, \mathcal{E})_{\ell_q}}{\|f\|_{\ell_q}^{1-\alpha} \|f\|_{A_1(\mathcal{E})}^{\alpha}} \ge \frac{1}{2} m^{-\alpha/p}, \quad p := \frac{q}{q-1}. \tag{3.25}$$

这意味着，对于任意固定的参数 t 和 b，定理 3.2 给出的上界在 m中是尖锐的。更准确地说，对于每个 $q \in (1, 2]$，都存在一个巴拿赫空间 X，满足 $\rho(u, X) \le \gamma u^q$，以及一个字典 $\mathcal{D} \subset X$，使得下列不等式成立

$$\frac{\sigma_m(f, \mathcal{D})_X}{\|f\|_X^{1-\alpha} \|f\|_{A_1(\mathcal{D})}^{\alpha}} \ge \frac{1}{2} m^{-\alpha/p}, \quad p := \frac{q}{q-1}. \tag{3.26}$$

不等式（3.26）直接由（3.25）结合 $X = \ell_q$ 和 $\mathcal{D} = \mathcal{E}$ 得出。

备注 3.2。不等式（3.26）给出了最佳 m 项逼近的下界。已知（参见 *[6],* 第 6 章），存在贪婪型算法，例如弱切比雪夫贪婪算法和带自由松弛的弱贪婪算法，其弱参数为 $t \in (0,1]$, ，它们为 $f \in X$ （其中 ρ（u、X）$\leq \gamma u^q$、$1 < q \leq 2$ 提供了如下的收敛速度，

$$\|f_m\|_X \leq C(q,\gamma)(1+mt^p)^{-1/p}\|f\|_{A_1(\mathcal{D})}, \quad p := \frac{q}{q-1}.$$

这意味着，（3.26）中的上界（按阶的意义）可以通过一种贪心型算法实现。

注意，对于特定的 X 和 $\mathcal{D}$，不等式（3.26）可能得到改进。我们以 $X = \ell_q$的例子加以说明，其中有 $q \in (2,\infty)$ 和 $\mathcal{D} = \mathcal{E}$。不失一般性，我们可以假设

$$f = \sum_{i=1}^{\infty} c_i e_i, \quad c_1 \geq c_2 \geq \cdots \geq 0.$$

Then

$$\|f\|_{\ell_q} = \left(\sum_{i=1}^{\infty} c_i^q\right)^{1/q}, \quad \|f\|_{A_1(\mathcal{E})} = \|f\|_{\ell_1} = \sum_{i=1}^{\infty} c_i. \tag{3.27}$$

现在我们从上方进行估计

$$\sigma_m(f,\mathcal{E})_{\ell_q} = \left(\sum_{i=m+1}^{\infty} c_i^q\right)^{1/q}.$$

我们关于 $\{c_i\}$ 的单调性假设意味着

$$c_m \leq m^{-1/q}\|f\|_{\ell_q}, \quad c_m \leq m^{-1}\|f\|_{\ell_1}$$

因此，对于任意 $\beta \in [0,1]$

$$c_m \leq m^{-(1/q)(1-\beta)-\beta}\|f\|_{\ell_q}^{1-\beta}\|f\|_{\ell_1}^{\beta}. \tag{3.28}$$

令 $\alpha := \beta(1-1/q)+1/q$，由此取 $p := \frac{q}{q-1}$

$$\sigma_m(f,\mathcal{E})_{\ell_q} \leq c_m^{1-1/q}\|f\|_{\ell_1}^{1/q} \leq m^{-\alpha/p}\|f\|_{\ell_q}^{1-\alpha}\|f\|_{\ell_1}^{\alpha}. \tag{3.29}$$

因此，对于任意 $\alpha \in [0,1]$，对所有 $f \in \ell_q$ 均有

$$\frac{\sigma_m(f,\mathcal{E})_{\ell_q}}{\|f\|_{\ell_q}^{1-\alpha}\|f\|_{A_1(\mathcal{E})}^{\alpha}} \le m^{-\alpha/p}, \quad p := \frac{q}{q-1}. \tag{3.30}$$

注意到，已知（见上文 (3.1)），空间 ℓ_q with $q \in [2,\infty)$具有幂型 2 的光滑模。因此，定理3.2 给出了 (3.30) 的一个类似结果，其衰减速率 $m^{-\alpha/2}$ 比 (3.30) 中的 $m^{-\alpha/p}$ 更弱。

我们简要讨论另一个具有 $q \in (2,\infty)$ 的 L_q 型空间示例。考虑实数 2π—周期函数空间 $L_q(0,2\pi)$，并取 $\mathcal{D} = \mathcal{T}$为实三角系统。对于给定的 $m \in \mathbb{N}$ 集

$$f := \sum_{k=1}^{2m} \cos(2^k x).$$

那么，众所周知，

$$\|f\|_{L_q} \le C(q)(2m)^{1/2}, \quad \|f\|_{A_1(\mathcal{T})} = 2m.$$

此外，

$$\sigma_m(f,\mathcal{T})_{L_q} \ge \sigma_m(f,\mathcal{T})_{L_2} \ge Cm^{1/2}.$$

因此，对于任意 $\alpha \in [0,1]$，我们得到

$$\frac{\sigma_m(f,\mathcal{T})_{L_q}}{\|f\|_{L_q}^{1-\alpha}\|f\|_{A_1(\mathcal{T})}^{\alpha}} \ge C'(q)m^{-\alpha/2}. \tag{3.31}$$

致谢。本研究得到俄罗斯科学基金会（项目编号：23-71-30001）在莫斯科国立罗蒙诺索夫大学的资助。

参考文献

[1] M. Donahue, L. Gurvits,C.Darken, E. Sontag, 非希尔伯特空间中凸逼近的收敛速率, Constr. Approx., 13(1997),187– 220.

[2] R.A. DeVore 和 V.N. Temlyakov, 《关于贪心算法的一些评述》, Advances in Comp. Math., 5 (1996), 173–187。

[3] J. Lindenstrauss 和 L. Tzafriri, 《经典巴拿赫空间 I》, Springer-Verlag, 柏林, 1977。[4] E.D. Livshitz, 《关于贪心算法收敛速度下界的估计》, Izv. RAN, Ser. Matem., 73 (2009), 125–144。[5] V.N. Temlyakov, 《巴拿赫空间中的贪心展开》, Advances in Comput. Math. 26 (2007), 431–449。[6] V.N. Temlyakov, 《贪婪逼近》, 剑桥大学出版社, 2011。