arXiv:2602.01421v1 [math.FA] 1 Feb 2026

希尔伯特空间中带自适应松弛的贪心算法收敛性分析

帕布罗·M·伯恩´a* 安德烈娅·加西亚´ıa†

摘要

幂松弛贪心算法（PRGA）是对 DeVore 和 Temlyakov 提出的所谓松弛贪婪算法的一种推广，其通过将松弛参数 $1/m$ 替换为 $1/m^{\alpha}$，旨在提高收敛速率。尽管 $\alpha \leq 1$ 这一情形已得到充分理解，但该算法在 $\alpha > 1$ 时的表现仍是一个未解决的问题。在本文中，我们回答了这一问题，并进一步提出一种松弛贪婪算法，在每次迭代时通过精确线搜索选择最优步长。

1 引言

贪心算法在非线性逼近理论中得到了广泛研究，尤其是在涉及相对于冗余系统进行近似的问题中；这类系统通常称为字典（例如，参见 [1, 3]）。与标准正交基不同，字典不要求线性无关，因此在表示希尔伯特空间中的元素时具有更大的灵活性。然而，这种额外的灵活性使高效近似方法的设计更具挑战性，而贪心算法在此情境下提供了一种自然且直观的方法。

最早出现且研究最广泛的程序之一是纯贪婪算法。该算法通过迭代构造近似：在每一步中，选择与当前残差相关性最大的字典元素。随后，以直接的方式将选中的原子加入近似中。由于其简单易行，纯贪婪算法易于实现且计算效率高。然而，现在人们已经充分认识到，对于一般字典而言，这种方法可能导致收敛速率缓慢或非最优。

为克服这些不足，DeVore 和 Temlyakov 提出了松弛贪婪算法。该方法的主要改进是引入一个松弛步骤，在此步骤中，新的近似由先前的近似与新选出的字典元素的凸组合构成。这种松弛降低了更新的激进程度，从而使逼近过程更加稳定。该算法的一项重要结果是，对于一类重要的函数，即具有有界原子范数的函数，它能够达到最优收敛速率。

松弛贪婪算法的提出，是贪婪近似方法发展历程中的重要一步。它表明，即使面对高度冗余的字典，只要在贪心选择与松弛之间取得适当平衡，仍能恢复最优性能。此后，这一思想在多种广义贪婪算法的发展中发挥了作用，也凸显了它与 Frank–Wolfe 算法等优化技术之间的联系。

松弛贪婪算法的有效性自然而然地引出了更多问题，例如如何选择松弛参数，以及是否存在能够在特定情形下提升性能的替代策略。这些考量促成了对原始算法的若干扩展与变体，也构成了本研究的动机。

本文结构如下。在第 2 节中，我们给出本文所用不同算法和字典的主要定义。在第 3 节中，我们证明定理2.4，该定理针对 $\alpha > 1$ 的取值，对 PRGA 的收敛性给出了否定答案。在

*数学系，CUNEF大学，28040 马德里（西班牙），pablo.berna@cunef.edu

†圣巴勃罗CEU大学，CEU大学联盟，28003 马德里（西班牙）；以及数学系，CUNEF大学，28040 马德里（西班牙），andrea.garciapons@usp.ceu.es

第 4 节中，我们展示了使用 Python 实现的 PRGA 在 $\alpha > 1$ 的不同取值下的数值模拟，表明在这些取值下，该算法确实不一定收敛。最后，在第 5 节中，我们介绍 CRGA 并建立一个收敛结果。

2 符号与主要结果

本节介绍本文研究的主要背景：希尔伯特空间中的字典。此外，我们还给出纯贪婪算法和松弛贪婪算法的精确定义，以引出本文研究的主要算法。

本文考虑一个希尔伯特空间：设 $\mathbb{H}$是一个实或复向量空间，并赋有一个内积

$$\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{R} \text{ or } \mathbb{C},$$

该内积对所有 $x, y, z \in \mathbb{H}$ 以及所有标量 $\alpha \in \mathbb{R}$ 或 $\mathbb{C}$均满足以下性质 ：

1. 正定性： $\langle x, x \rangle \geq 0$，且当且仅当 $x = 0$时等号成立。
2. 共轭对称性： $\langle x, y \rangle = \overline{\langle y, x \rangle}$。
3. 第一变量齐次性： $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$。
4. 第一变量线性： $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$。

内积在 $\mathbb{H}$ 上诱导出一个范数，其定义为

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

如果空间 $\mathbb{H}$ 关于该范数所诱导的度量是完备的，则称其为希尔伯特空间。

一个字典 $\mathcal{D} \subset \mathbb{H}$ 是满足以下条件的元素集合：

1. 每个元素 $g \in \mathcal{D}$都具有单位范数，称为一个原子，即 $\|g\| = 1$；且当 $g \in \mathcal{D}$时，也有 $-g \in \mathcal{D}$。

2. $\mathcal{D}$ 的闭线性包在 $\mathbb{H}$中稠密。等价地，对于任意 $x \in \mathbb{H}$ 和任意 $\varepsilon > 0$，都存在元素 g_1、…、 $g_n \in \mathcal{D}$ 以及标量 a_1、…、 a_n，使得

$$\left\| x - \sum_{i=1}^{n} a_i g_i \right\| < \varepsilon.$$

作为一个部分例子，若 $\{h_k\}_{k=1}^{\infty}$ 是 $\mathbb{H}$的一组正交归一基，则集合 $\mathcal{D} = \{\pm h_k\}_{k=}$ [1] 构成一个字典，其含义如上所述。

2.1 PGA 和 RGA

设 $\mathbb{H}$ 为实希尔伯特空间， $\mathcal{D} \subset \mathbb{H}$ 为字典。对于 $\mathcal{D}$，并且给定 $\tau > 0$，我们将逼近类 $A_\tau(\mathcal{D})$ 定义为集合在 $\mathbb{H}$ 中的闭包：

$$\left\{ f \in \mathbb{H} : \ f = \sum_{k \in \Lambda} a_k g_k, \quad g_k \in \mathcal{D}, \ |\Lambda| < \infty, \text{ and } \sum_{k \in \Lambda} |a_k|^\tau \leq 1 \right\}.$$

为了介绍 [3], 中研究的松弛贪婪算法新版本，我们先简要回顾纯贪婪算法。

给定 $f \in \mathbb{H}$，令 $g = g(f) \in \mathcal{D}$为字典中使 $\langle f, g \rangle$达到最大值的元素（为简单起见，我们假设这样的最大值可以达到）。我们定义

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

纯贪婪算法（PGA）。设 $R_0(f) = R_0(f, \mathcal{D}) := f$和 $G_0(f) = G_0(f, \mathcal{D}) := 0$。对于每个$m \geq 1$，递归定义

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

$$R_m(f) = R_m(f,\mathcal{D}) := f - G_m(f) = R\big(R_{m-1}(f)\big).$$

松弛贪心算法（RGA）。设 $R_0^r(f) = R_0^r(f,\ \mathcal{D}) := f$和 $G_0^r(f) = G_0^r(f,\ \mathcal{D}) := 0$。对于$m = 1$，我们取

$$G_1^r(f) := G_1(f), \qquad R_1^r(f) := R_1(f).$$

对于 $m \geq 2$，我们定义

$$G_m^r(f) = G_m^r(f,\mathcal{D}) := \left(1 - \frac{1}{m}\right) G_{m-1}^r(f) + \frac{1}{m}\, g\big(R_{m-1}^r(f)\big),$$

$$R_m^r(f) = R_m^r(f,\mathcal{D}) := f - G_m^r(f).$$

利用该算法，[3]证明了如下估计。

定理 2.1. 设 $\mathcal{D}$ 为希尔伯特空间中的一个字典。则对每一个 $f \in A_1(\mathcal{D})$ 以及每一个$m = 1,\ 2,\ \ldots\ldots$，

$$\|f - G_m^r(f)\| \leq \frac{2}{\sqrt{m}}.$$

2.2 幂松弛贪婪算法（PRGA）

由 RGA 的定义自然产生如下问题：为何在凸组合中选用 $1/m$，而不是带有任意正实数 α 的 $1/m^\alpha$？这样的幂次能否改进 DeVore 与 Temlyakov 所证明的收敛结果?

这一自然问题促使 A. Garc´ıa 在 [4] 中引入了 RGA 的如下推广，称为幂松弛贪婪算法。

设 $\mathcal{R}_0^r(f) = \mathcal{R}_0^r(f,\mathcal{D}) := f$且 $\mathcal{T}_0^r(f) = \mathcal{T}_0^r(f,\mathcal{D}) := 0$。对于 $m = 1$，我们定义

$$\mathcal{T}_1^r(f) := G_1(f), \qquad \mathcal{R}_1^r(f) := R_1(f).$$

对于 $m \geq 2$，我们令

$$\mathcal{T}_m^r(f) = \mathcal{T}_m^r(f,\mathcal{D}) := \left(1 - \frac{1}{m^\alpha}\right) \mathcal{T}_{m-1}^r(f) + \frac{1}{m^\alpha}\, g\big(\mathcal{R}_{m-1}^r(f)\big),$$

$$\mathcal{R}_m^r(f) = \mathcal{R}_m^r(f,\mathcal{D}) := f - \mathcal{T}_m^r(f).$$

该算法的收敛性由下述结果描述，其证明见 [4]。

定理 2.2. 设 $\mathcal{D}$ 为希尔伯特空间中的一个字典。若 $\alpha \leq 1$，则对每一个 $f \in A_1(\mathcal{D})$ 以及每一个$m = 1,\ 2,$ $\ldots\ldots$，

$$\|f - \mathcal{T}_m^r(f)\|^2 \leq \frac{4}{m^\alpha}.$$

当然，正如该结果所示，对于 $\alpha \leq 1$，最优情形是 $\alpha = 1$，从而恢复了 RGA 以及 DeVore 与 Temlyakov 所证明的结果。于是问题出现了：对于 $\alpha > 1$ 会发生什么？这正是 [4]中提出的问题。

问题 2.3 ([4, 问题 1]**).** 是否可以找到 $\alpha_0 > 1$ 和 $c > 0$，使得

$$\|f - \mathcal{T}_m^r(f)\| \leq \frac{c}{m^{\alpha_0/2}}, \qquad m = 1, 2, \ldots,$$

对每个 $f \in A_1(\mathcal{D})$ 都成立吗?

在这里，我们给出了否定的答案。

定理 2.4。取 $\alpha > 1$ 作为（*PRGA*），则存在一个元素 $f \in A_1(\mathcal{D})$，其满足 $\mathcal{D}$ 某个特定字典的 $\mathbb{R}^2$，使得

$$\inf_m \|f - \mathcal{T}_m^r(f)\| > 0.$$

2.3 凸松弛贪心算法

如前所述，在 RGA 的定义中，我们在 $G^r_{m-1}(f)$ 和 $g(R^r_{m-1}(f))$ 之间使用因子为$1/m$ 的凸组合。这里我们要探讨的问题是：如果改变因子 $1/m$，并尝试研究哪种凸组合最优，会怎么样？为了解答这一问题，我们引入以下算法，称为**凸松弛贪心算法**（CRGA）：设 $\mathcal{D}$ 为实希尔伯特空间 $\mathbb{H}$中的一个对称字典，且 $f \in \mathbb{H}$。我们定义

$$f_0 := 0, \qquad r_0 := f.$$

对于每个 $m \geq 1$，假设 f_{m-1} 和 r_{m-1} 已定义，我们按如下步骤进行：

1. **贪心选择**。 我们选择 $g_m \in \mathcal{D}$ 使得

$$\langle r_{m-1}, g_m \rangle = \sup_{g \in \mathcal{D}} \langle r_{m-1}, g \rangle.$$

2. **最优松弛步骤**。 我们定义

$$f_m = (1 - \gamma_m) f_{m-1} + \gamma_m g_m,$$

其中，参数 $\gamma_m \in [0, 1]$ 选为以下问题的一个解

$$\gamma_m \in \arg\min_{\gamma \in [0,1]} \left\| f - \big((1-\gamma) f_{m-1} + \gamma g_m\big) \right\|^2.$$

3. **残差**。 我们定义

$$r_m := f - f_m.$$

注意到， 由构造可知，对所有 m，均有 $f_m \in$conv($\mathcal{D}$)。对于 CRGA，我们有如下结果。

定理 2.5. 设 $\mathcal{D}$ 为实希尔伯特空间中的一个字典。若 $f \in A_1(\mathcal{D})$，则由前述算法生成的序列 $(f_m)_{m \geq 0}$满足

$$\|f - f_m\| \leq \frac{2}{\sqrt{m+4}}, \qquad m = 0, 1, 2, \ldots$$

3 定理2.4的证明

为证明定理 2.4，我们需要证明如下技术性引理。

引理 3.1. 对于 $\alpha > 1$ ， 我们有

$$P_\alpha := \prod_{k=2}^{\infty} \Big(1 - \frac{1}{k^\alpha}\Big) \in (0, 1).$$

证明。 我们将乘积改写为对数的形式：

$$\ln P_\alpha = \sum_{k=2}^{\infty} \ln\Big(1 - \frac{1}{k^\alpha}\Big).$$

乘积中的每个因子都满足 $0 < 1 - 1/k^\alpha < 1$，因此

$$\ln\Big(1 - \frac{1}{k^\alpha}\Big) < 0 \qquad (k \geq 2),$$

从而部分和

$$S_N := \sum_{k=2}^{N} \ln\Big(1 - \frac{1}{k^\alpha}\Big)$$

构成一个单调递减序列。为了得出 $\ln P_\alpha$ 是一个实数，只需证明 (S_N) 有下界。

为此，我们使用对数的一个标准界。对于 $x \in (0,1)$，有

$$\ln(1-x) \geq -\frac{x}{1-x}.$$

特别地，如果 $0 < x \leq \frac{1}{2}$，那么 $\frac{1}{1-x} \leq 2$，因此

$$\ln(1-x) \geq -2x.$$

取 $x = 1/k^\alpha$，对所有 $k \geq 2$ 都有 $1/k^\alpha \leq \frac{1}{2}$，因此

$$\ln\Big(1 - \frac{1}{k^\alpha}\Big) \;\geq\; -\frac{2}{k^\alpha}, \qquad k \geq 2.$$

因此，对于每个 $N \geq 2$,

$$S_N = \sum_{k=2}^{N} \ln\Big(1 - \frac{1}{k^\alpha}\Big) \;\geq\; -2\sum_{k=2}^{N} \frac{1}{k^\alpha}.$$

由于 $\alpha > 1$，级数 $\sum_{k=2}^{\infty} k^{-\alpha}$ 收敛，因此部分和 $\sum_{k=2}^{N} k^{-\alpha}$ 在 N 中一致有界。所以，存在一个常数 $C \in \mathbb{R}$，使得

$$S_N \geq C \qquad \text{for all } N \geq 2.$$

这表明序列 (S_N) 单调递减且有下界，因此收敛到某个极限 $S \in \mathbb{R}$:

$$\sum_{k=2}^{\infty} \ln\Big(1 - \frac{1}{k^\alpha}\Big) = S.$$

最后，

$$P_\alpha = \prod_{k=2}^{\infty}\Big(1 - \frac{1}{k^\alpha}\Big) = \exp\Big(\sum_{k=2}^{\infty} \ln\Big(1 - \frac{1}{k^\alpha}\Big)\Big) = e^S,$$

由于 S 是一个实数，我们有 $e^S > 0$。此外，由于所有因子 $1 - 1/k^\alpha$ 都严格小于 1，因此可得 $P_\alpha < 1$。综上，$P_\alpha \in (0,\ 1)$。 □

定理 2.4 的证明。固定 $\alpha > 1$。令 $\mathbb{H} = \mathbb{R}^2$，其内积为欧几里得内积 nd

$$e_1 = (1,0), \qquad e_2 = (0,1), \qquad \mathcal{D} = \{\pm e_1, \pm e_2\}.$$

取 $b \in (0,\ 1/2)$ 并定义

$$f := (1-b)e_1 + be_2.$$

则 $f \in A_1(\mathcal{D})$，并且对于步长为 $\lambda_m := m^{-\alpha}$ 的幂松弛贪心算法，我们将证明

$$\|f - \mathcal{T}_m^r(f)\|_2 \;\geq\; \frac{b\,P_\alpha}{\sqrt{2}} \;> 0, \qquad m \geq 1,$$

其中

$$P_\alpha := \prod_{k=2}^{\infty}\Big(1 - \frac{1}{k^\alpha}\Big) \in (0,1)$$

利用引理 3.1。特别地，$\|f - \mathcal{T}_m^r(f)\|_2 > 0$ 对每个 m，并且定理 2.2 的估计对 $\alpha > 1$ 不成立。

由于 $f = (1-b)e_1 + be_2$ 满足 $(1-b) + b = 1$，我们有 $f \in A_1(\mathcal{D})$。此外，由于 $b < 1/2$，我们有 $\langle f,\ e_1\rangle = 1 - b > b = \langle f,\ e_2\rangle$，因此第一个贪心原子唯一地为 $g(f) = e_1$，从而

$$\mathcal{T}_1^r(f) = G_1(f) = \langle f, e_1\rangle e_1 = (1-b)e_1.$$

对于 $m \geq 2$，算法的形式为

$$\mathcal{T}_m^r(f) = (1-\lambda_m)\mathcal{T}_{m-1}^r(f) + \lambda_m\, g(\mathcal{R}_{m-1}^r(f)), \qquad \lambda_m = \frac{1}{m^\alpha},$$

且 $g(\mathcal{R}^r_{m-1}(f)) \in \mathcal{D}$。注意， $\mathcal{D}$ 中的每个原子的 $\|\cdot\|_1$-范数都等于 1，即 $\|d\|_1 = 1$ 对所有$d \in \mathcal{D}$。

设 $s_m := \|\mathcal{T}^r_m(f)\|_1$。利用 $\|\cdot\|_1$ 的三角不等式，

$$s_m = \|(1-\lambda_m)\mathcal{T}^r_{m-1}(f)+\lambda_m g(\mathcal{R}^r_{m-1}(f))\|_1 \le (1-\lambda_m)s_{m-1}+\lambda_m\|g(\mathcal{R}^r_{m-1}(f))\|_1 = (1-\lambda_m)s_{m-1}+\lambda_m.$$

定义 $u_m := 1 - s_m$。于是

$$u_m = 1 - s_m \;\ge\; 1 - \big[(1-\lambda_m)s_{m-1} + \lambda_m\big] = (1-\lambda_m)(1-s_{m-1}) = (1-\lambda_m)u_{m-1}.$$

从 $m = 2$ 迭代可得

$$u_m \ge u_1 \prod_{k=2}^{m}(1-\lambda_k).$$

由于 $u_1 = 1 - \|\mathcal{T}^r_1(f)\|_1 = 1 - (1-b) = b$，我们得到

$$1 - \|\mathcal{T}^r_m(f)\|_1 = u_m \;\ge\; b\prod_{k=2}^{m}\Big(1-\frac{1}{k^\alpha}\Big) \;\ge\; b\prod_{k=2}^{\infty}\Big(1-\frac{1}{k^\alpha}\Big) = bP_\alpha.$$

现在考虑单位向量 $v = (1,1)/\sqrt{2}$。对于任意 $x \in \mathbb{R}^2$，

$$\langle x, v\rangle = \frac{x_1+x_2}{\sqrt{2}} \le \frac{|x_1|+|x_2|}{\sqrt{2}} = \frac{\|x\|_1}{\sqrt{2}}.$$

因此，

$$\langle \mathcal{T}^r_m(f), v\rangle \le \frac{\|\mathcal{T}^r_m(f)\|_1}{\sqrt{2}} \le \frac{1-bP_\alpha}{\sqrt{2}}.$$

但 $\langle f,\ v\rangle = \frac{(1-b)+b}{\sqrt{2}} = \frac{1}{\sqrt{2}}$，所以

$$\langle f - \mathcal{T}^r_m(f), v\rangle = \langle f, v\rangle - \langle \mathcal{T}^r_m(f), v\rangle \ge \frac{bP_\alpha}{\sqrt{2}}.$$

由于 $\|v\|_2 = 1$，根据柯西–施瓦茨不等式

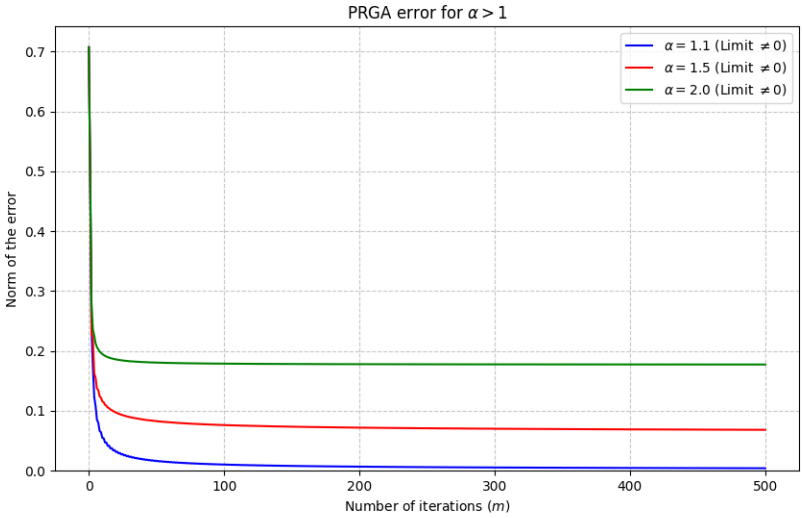
$$\|f - \mathcal{T}^r_m(f)\|_2 \;\ge\; \langle f - \mathcal{T}^r_m(f), v\rangle \;\ge\; \frac{bP_\alpha}{\sqrt{2}} \;> 0,$$

对每个 $m \ge 1$ 都成立，正如所述。 □

4 数值模拟

为了通过数值说明算法不收敛的现象，基于上一节的内容，我们考虑 $f = (1/2,\ 1/2)$，并针对不同的 α取值对算法进行模拟。下面给出的结果见 [2]。

--- 模拟结果 --- Alpha=1.1：最终误差 ||f - T_500|| = 0.003805 Alpha=1.5：最终误差 ||f - T_500|| = 0.068021 Alpha=2.0：最终误差 ||f - T_500|| = 0.177130



5 凸松弛贪心算法

在研究 CRGA 的收敛结果之前，我们将给出定义该算法的参数γ 的一个明确值。为此，我们需要以下引理。

引理 5.1（松弛参数的最优选择）。 设 $r \in \mathbb{H}$ 和 $d \in \mathbb{H}$ 满足 $d \neq 0$。考虑函数

$$\Phi(\gamma) := \|r - \gamma d\|^2, \qquad \gamma \in \mathbb{R}.$$

则：

1. Φ 是严格凸的，且其在 $\mathbb{R}$ 上的唯一全局最小化元为

$$\gamma^* = \frac{\langle r, d\rangle}{\|d\|^2}.$$

2. 约束问题

$$\min_{\gamma \in [0,1]} \Phi(\gamma)$$

有唯一解 $\gamma_m \in [0,1]$，其表达式为

$$\gamma_m = \Pi_{[0,1]}(\gamma^*) = \min\{1, \max\{0, \gamma^*\}\}.$$

3. 对于每个 $\gamma \in [0,1]$，以下不等式成立：

$$\Phi(\gamma_m) \leq \Phi(\gamma).$$

证明。 </style}函数 Φ 是关于 γ 的二次多项式，其首项系数为 $\|d\|^2 > 0$，因此它是严格凸的。求导可得

$$\Phi'(\gamma) = -2\langle r, d\rangle + 2\gamma\|d\|^2,$$

令导数等于零，可得唯一的临界点

$$\gamma^* = \frac{\langle r, d\rangle}{\|d\|^2},$$

它是 $\mathbb{R}$ 上 Φ 的全局最小化元。

由于 Φ 是凸函数，其在闭凸区间 $[0,1]$ 上的最小化可通过将全局最小化元投影到该区间上得到。由此可得表达式

$$\gamma_m = \Pi_{[0,1]}(\gamma^*).$$

最后，由于 γ_m 是 Φ 在 $[0,1]$, 上的极小值点，我们有

$$\Phi(\gamma_m) \le \Phi(\gamma), \qquad \forall\, \gamma \in [0,1],$$

至此，证明完成。 □

接下来，对于 CRGA，我们定义

$$d_m := g_m - f_{m-1}.$$

对于 $\gamma \in [0,1]$, 我们有

$$f - \big((1-\gamma)f_{m-1} + \gamma g_m\big) = r_{m-1} - \gamma d_m.$$

因此，我们要最小化的泛函为

$$\Phi(\gamma) := \|r_{m-1} - \gamma d_m\|^2 = \|r_{m-1}\|^2 - 2\gamma\langle r_{m-1}, d_m\rangle + \gamma^2\|d_m\|^2.$$

这是关于 γ 的严格凸二次函数。其导数为

$$\Phi'(\gamma) = -2\langle r_{m-1}, d_m\rangle + 2\gamma\|d_m\|^2.$$

令导数等于零，可得无约束最小化元

$$\gamma^* = \frac{\langle r_{m-1}, d_m\rangle}{\|d_m\|^2},$$

前提是 $d_m \neq 0$。由于问题的约束区间为 $[0,1]$, ，所需的最小化元就是将 γ^* 投影到该区间上所得的结果:

$$\gamma_m = \min\left\{1, \max\left\{0, \frac{\langle r_{m-1}, d_m\rangle}{\|d_m\|^2}\right\}\right\}.$$

若 $d_m = 0$, 则 $g_m = f_{m-1}$, 任意 $\gamma_m \in [0,1]$ 都是最优的；此时取 $\gamma_m = 0$.

此外，为了证明 CRGA 的收敛性，我们还需要用到以下性质。

引理 5.2。 设 $\mathcal{D}$ 是希尔伯特空间 $\mathbb{H}$中的一个对称字典，令 $f \in A_1(\mathcal{D})$，并设 $(f_m)_{m\ge 0}$是由采用最优步长的松弛贪婪算法生成的序列，其中 $f_0 = 0$。定义

$$a_m := \|f - f_m\|^2, \qquad m \ge 0.$$

则序列 $(a_m)_{m\ge 0}$ 满足以下性质:

1. (a_m) 单调不增，即

$$a_m \le a_{m-1} \qquad \textit{for all } m \ge 1.$$

2. (a_m) 一致有界，事实上

$$0 \le a_m \le a_0 = \|f\|^2 \le 1 \qquad \textit{for all } m \ge 0.$$

证明。根据算法的定义，参数 $\gamma_m \in [0,1]$ 的选取旨在使下式最小

$$\|f - \big((1-\gamma)f_{m-1} + \gamma g_m\big)\|^2 \quad \text{over } \gamma \in [0,1].$$

将 $\gamma = 0$ 代入上式，得到

$$a_m = \min_{\gamma\in[0,1]} \|f - \big((1-\gamma)f_{m-1} + \gamma g_m\big)\|^2 \le \|f - f_{m-1}\|^2 = a_{m-1},$$

这证明了单调性。

由于对所有 $g \in \mathcal{D}$ 均有 $f \in A_1(\mathcal{D})$ 与 $\|g\| = 1$, 故得 $\|f\| \le 1$。因此

$$a_m \le a_0 = \|f\|^2 \le 1 \quad \text{for all } m \ge 0,$$

至此证明完成。 □

定理的证明 2.5。 我们定义

$$a_m := \|r_m\|^2 = \|f - f_m\|^2.$$

步骤 *1.* 由于 $f \in A_1(\mathcal{D})$ 且泛函 $g \mapsto \langle r_{m-1}, g\rangle$ 是线性的，我们有

$$\langle r_{m-1}, g_m\rangle = \sup_{g\in\mathcal{D}} \langle r_{m-1}, g\rangle \geq \langle r_{m-1}, f\rangle.$$

将两边同时减去 $\langle r_{m-1}, f_{m-1}\rangle$ ，得到

$$\langle r_{m-1}, g_m - f_{m-1}\rangle \geq \langle r_{m-1}, f - f_{m-1}\rangle = \|r_{m-1}\|^2 = a_{m-1}.$$

令 $d_m := g_m - f_{m-1}$，则有

$$\langle r_{m-1}, d_m\rangle \geq a_{m-1}. \tag{1}$$

步骤 *2.* 由于 $f_{m-1} \in \text{conv}(\mathcal{D})$ 且对每个 $g \in \mathcal{D}$ 都有 $\|g\| = 1$ ，因此 $\|f_{m-1}\| \leq 1$。于是，

$$\|d_m\| = \|g_m - f_{m-1}\| \leq \|g_m\| + \|f_{m-1}\| \leq 2. \tag{2}$$

步骤 *3.* 对于 $\gamma \in [0, 1]$，我们定义

$$\Phi(\gamma) := \|r_{m-1} - \gamma d_m\|^2.$$

根据引理 5.1，由算法定义的参数 γ_m 满足

$$a_m = \Phi(\gamma_m) \leq \Phi(\gamma), \qquad \forall\, \gamma \in [0, 1].$$

利用 (1) 和 (2)，对于所有 $\gamma \in [0, 1]$ ，我们得到

$$\Phi(\gamma) = a_{m-1} - 2\gamma\langle r_{m-1}, d_m\rangle + \gamma^2\|d_m\|^2 \leq a_{m-1} - 2\gamma a_{m-1} + 4\gamma^2.$$

利用引理 5.2，我们知道

$$\gamma = \frac{a_{m-1}}{4} \in [0, 1].$$

然后，取此值，我们可得

$$a_m \leq a_{m-1} - \frac{a_{m-1}^2}{4}.$$

步骤 *4.* 由前面的不等式，并利用 $\dfrac{1}{1-t} \geq 1 + t$ 对任意 $t \in [0, 1)$，可得

$$\frac{1}{a_m} \geq \frac{1}{a_{m-1}} + \frac{1}{4}.$$

迭代并利用 $a_0 = \|f\|^2 \leq 1$这一事实，我们得到

$$\frac{1}{a_m} \geq \frac{m+4}{4},$$

这意味着

$$a_m \leq \frac{4}{m+4}.$$

□

这一结果之所以最优，是因为已知界限 $m^{-1/2}$ 无法改进。事实上，已知对于任意贪心型算法，都有以下下界；如果读者对此并不熟悉，我们将在此给出证明。

命题 5.3. 对于每个满足 $m \geq 2$ 的 $m \in \mathbb{N}$，都存在一个希尔伯特空间 $\mathbb{H}$、一个字典 $\mathcal{D} \subset \mathbb{H}$，以及一个元素 $f \in A_1(\mathcal{D})$，使得任何至多使用 $\mathcal{D}$ 中 m 个元素的近似 s 都满足

$$\|f - s\| \;\geq\; \frac{m^{-1/2}}{\sqrt{2}}.$$

证明。固定 $m \in \mathbb{N}, m \geq 2$，并令 $n := 2m$。设 $\mathbb{H} = \mathbb{R}^n$ 配备欧几里得范数，并令

$$\mathcal{D} := \{\pm e_1, \pm e_2, \ldots, \pm e_n\},$$

其中 $(e_i)_{i=1}^n$ 表示 $\mathbb{R}^n$ 的标准基。

定义

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

那么 f 可以表示为

$$f = \sum_{i=1}^{2m} a_i g_i, \qquad a_i = \frac{1}{2m}, \quad g_i = e_i \in \mathcal{D},$$

因此

$$\sum_{i=1}^{2m} |a_i| = 2m \cdot \frac{1}{2m} = 1.$$

根据原子类 $A_1(\mathcal{D})$ 的定义，这表明 $f \in A_1(\mathcal{D})$。此外，

$$\|f\|_2^2 = \sum_{i=1}^{2m} \left(\frac{1}{2m}\right)^2 = 2m \cdot \frac{1}{4m^2} = \frac{1}{2m} \leq 1,$$

因此，f 属于 $\mathbb{H}$的单位球。

设 $s \in \mathbb{R}^n$ 为任意一个可表示为$\mathcal{D}$中至多 m 个元素的线性组合的向量。等价地，s 至多有 m 个非零坐标。记 $\mathrm{supp}(s)$ 为 s的支撑，其中$|\mathrm{supp}(s)| \leq m$。

对于 $i \notin \mathrm{supp}(s)$，有 $s_i = 0$ 且 $f_i = 1/(2m)$，因此

$$\|f - s\|_2^2 = \sum_{i \notin \mathrm{supp}(s)} |f_i - s_i|^2 + \sum_{i \in \mathrm{supp}(s)} |f_i - s_i|^2 \;\geq\; \sum_{i \notin \mathrm{supp}(s)} |f_i|^2.$$

由于 $\mathrm{supp}(s)$ 中至多有 m个指标，至少有 m 个指标不属于其中，因此

$$\|f - s\|_2^2 \;\geq\; m \left(\frac{1}{2m}\right)^2 = \frac{1}{4m}.$$

因此

$$\|f - s\|_2 \geq \frac{1}{2\sqrt{m}},$$

对于每个可表示为 $\mathcal{D}$中至多 m 个元素的线性组合的 s，均有

由于任何贪心型算法产生的第 m 次迭代至多使用来自 $\mathcal{D}$ 的 m 个元素，因此相应的近似也适用同一下界。

□

参考文献

[1] J. D. Blanchar, J. Tanner，压缩感知中贪心算法的性能比较，《数值线性代数及其应用》**22** (2015)，254–282。[2] P. M. Berná, A. García，*PRGA* 的数值模拟，Zenodo（2026）https://doi.org/10.5281/zenodo.18443949[3] R. A. DeVore, V. N. Temlyakov，关于贪心算法的一些评述，《计算数学进展》**5** (1996)，第 2-3 期，173–187。[4] A. García，贪心算法：综述与开放问题，《不等式与应用杂志》**2025**，11（2025）。

资助

两位作者均获得 PID2022-142202NB-I00资助项目（国家研究署，西班牙）的支撑。