

# $A^\dagger$ 分级硬阈值追踪

施章磊 李维国

(中国石油大学(华东)理学院, 山东 青岛 266580)

摘要: 受到硬阈值追踪算法(HTP)的启发,提出了用于求解压缩感知问题的 $A^\dagger$ 硬阈值追踪算法( $A^\dagger$  graded hard thresholding pursuit algorithm, APGHTP),并在约束等距条件下给出了该算法的理论保证。在数值实验中,不论测量值是否包含误差,APGHTP都表现较好,证明了该算法的稀疏恢复能力。在恢复稀疏向量时,APGHTP所需的迭代数与稀疏向量的稀疏度相同。

关键词: 压缩感知; 稀疏解; 硬阈值追踪; Moore-Penrose 逆

中图分类号: O242 文献标志码: A

引用格式: 施章磊, 李维国.  $A^\dagger$  分级硬阈值追踪[J]. 山东大学学报(理学版) 2017, 52(8): 58-64, 69.

## $A^\dagger$ graded hard thresholding pursuit algorithm

SHI Zhang-lei, LI Wei-guo

(College of Science, China University of Petroleum, Qingdao 266580, Shandong, China)

**Abstract:** Inspired by hard thresholding pursuit algorithm (HTP),  $A^\dagger$  graded hard thresholding pursuit algorithm (APGHTP) was proposed for solving compressive sensing problems. The theoretical guarantees of the new algorithm were given under restricted isometry property (RIP) condition. In the numerical experiment, regardless of whether the measured value contains error, APGHTP performance is better, which proves the sparse recovery ability of the algorithm. When recovering sparse vectors, the number of iterations required for APGHTP is the same as that of sparse vectors.

**Key words:** compressive sensing; sparse solution; hard thresholding pursuit; Moore-Penrose inverse

## 0 前言

令  $x \in \mathbf{R}^N$  是具有  $k$  个非零元素的实向量 ( $x$  称为  $k$  稀疏信号或向量),  $f \in \mathbf{R}^m$  是满足  $Ax = f$  的测量值, 其中  $m \ll N$ ,  $A \in \mathbf{R}^{m \times N}$  压缩感知问题 (compressive sensing) 可表示如下:

$$\begin{aligned} \min_{x \in \mathbf{R}^N} \|x\|_0 \\ \text{s. t. } f = Ax + e, \end{aligned} \quad (1)$$

其中  $e \in \mathbf{R}^m$  是测量误差。

求解压缩感知也即寻找满足欠定线性方程组  $Ax = f$ ,  $m \ll N$ ,  $A \in \mathbf{R}^{m \times N}$  的最稀疏的解  $x$ , 当  $m$  正比于  $k \log\left(\frac{N}{k}\right)$  时, 即可求得  $x$ 。然而模型 (1) 是 NP-Hard 问题<sup>[1]</sup>, 参考文献 [2-5] 表明  $l_0$  范数应该被松弛为凸范数, 一个非常流行的方式是将  $l_0$  范数松弛为  $l_1$  范数:

收稿日期: 2016-10-13; 网络出版时间: 2017-04-26 16:01

网络出版地址: <http://kns.cnki.net/kcms/detail/37.1389.N.20170426.1601.010.html>

基金项目: 中央高校基本科研基金资助项目(15CX08011A); 国家自然科学基金资助项目(60971132)

作者简介: 施章磊(1991—)男, 硕士研究生, 研究方向为信号与图形处理. E-mail: shizhanglei2010@163.com

$$\begin{aligned} & \min_{x \in \mathbf{R}^N} \|x\|_1 \\ \text{s. t. } & f = Ax + e, \end{aligned} \quad (2)$$

其中  $e \in \mathbf{R}^m$  是测量误差。

许多基于不同策略的算法可以用来进行稀疏恢复,其中直接求解模型(1)的贪婪迭代类算法包括正交匹配追踪(orthogonal matching pursuit, OMP)<sup>[6]</sup>、压缩采样匹配追踪(compressive sampling matching pursuit, CoSaMP)<sup>[7]</sup>等。同时,迭代硬阈值算法(iterative hard thresholding algorithm, IHT)<sup>[8]</sup>也可以直接求解模型(1)。迭代软阈值算法(iterative soft thresholding algorithm, IST)<sup>[9]</sup>、Bregman 迭代算法(Bregman iteration algorithm, BIA)及其变体线性化 Bregman 算法(linearized Bregman algorithm, IBA)<sup>[10-11]</sup>等可以有效地求解  $l_1$ -极小化问题(模型(2))。硬阈值追踪算法(hard thresholding pursuit, HTP)<sup>[12]</sup>及分级硬阈值追踪算法(graded hard thresholding pursuit, GHTP)<sup>[13]</sup>结合了贪婪算法以及硬阈值算法思想。受上述算法的启发我们引入 Moore-Penrose 逆<sup>[14]</sup>提出了  $A^\dagger$  硬阈值追踪算法(APHTP)。

文中所用记号为:  $A^\dagger$  是矩阵  $A$  的 Moore-Penrose 逆,  $A'$  是  $A$  的转置,  $I$  为单位矩阵;  $\text{supp}(x)$  是  $x$  的支撑集  $S$  是  $x$  的  $k$  个模量最大的指标集;  $x_T$  是  $x$  对应于指标集  $T$  的子向量, 同样  $A_T$  是  $A$  的子矩阵;  $\bar{T}$  是集合  $T$  在  $\{1, 2, \dots, N\}$  的余集,  $\bar{T} = \{1, 2, \dots, N\} \setminus T$ ;  $T \Delta T'$  是集合  $T, T'$  的对称差集,  $T \Delta T' = (T \setminus T') \cup (T' \setminus T)$ 。

## 1 算法框架

### 1.1 硬阈值追踪算法框架

贪婪迭代算法的关键是从  $A'(f - Ax^n)$  中选择  $x$  的支撑集, 然后在所选择的支撑集下寻找最适合测量值的稀疏解。然而在某些条件下从  $x^n + A'(f - Ax^n) \approx x$  中选择  $x$  的支撑集似乎更自然, 这即是硬阈值追踪算法(HTP), 如算法 1 所示。

算法 1 HTP

参数:  $A, f, k, \text{threshold} > 0$ 。

初始化:  $x^0 = 0, r_0 = f, S^0 = \emptyset$

While  $\|r_n\|_2 > \text{threshold}$  do

(HTP<sub>1</sub>) 计算  $S^{n+1} = \{ |x^n + A'(f - Ax^n)| \text{ 中 } k \text{ 个模量最大值对应的指标集} \}$

(HTP<sub>2</sub>) 计算  $x^{n+1} = \arg \min \|f - Az\|_2, \text{supp}(z) \subseteq S^{n+1}$

(HTP<sub>3</sub>)  $r_{n+1} = f - Ax^{n+1}$

end while

return  $x^{n+1}$

当矩阵的  $3k$  阶约束等距条件满足  $\delta_{3k} < \frac{1}{\sqrt{3}}$  时, 硬阈值追踪算法在有限次迭代之后可以找到该系统的所有  $k$  稀疏解<sup>[12]</sup>。矩阵  $A \in \mathbf{R}^{m \times N}$  的  $k$  阶约束等距常数是满足式(3)的最小  $\delta_k > 0$ :

$$(1 - \delta_k) \|z\|_2^2 \leq \|Az\|_2^2 \leq (1 + \delta_k) \|z\|_2^2. \quad (3)$$

### 1.2 新算法框架

本小节包含了建立新算法的关键性结论。

引理 1.1<sup>[14]</sup> 向量  $u$  是  $Ax = f$  的最小二乘解当且仅当  $u$  是  $A^\dagger Ax = A^\dagger f$  的解。并且对于任意  $z \in \mathbf{R}^N$ , 最小二乘解为

$$u = A^\dagger f + (I - A^\dagger A)z. \quad (4)$$

若  $u = \arg \min_{x \in \mathbf{R}^N} \|Ax - f\|_2$ , 则由引理 1.1 可知对于所有  $z \in \mathbf{R}^N$  有下式成立:

$$u = A^\dagger f + (I - A^\dagger A)z. \quad (5)$$

引理 1.2 向量  $u$  是  $Ax = f$  的最小二乘解, 即满足  $u = \arg \min_{x \in \mathbf{R}^N} \|Ax - f\|_2$ , 当且仅当  $u$  满足式(6):

$$A^\dagger Au = A^\dagger f. \quad (6)$$

证明 由引理 1.1 可知若  $u$  是  $Ax=f$  的最小二乘解, 则  $u=A^\dagger f+(I-A^\dagger A)z$ , 因此有

$$A^\dagger Au=A^\dagger A(A^\dagger f+(I-A^\dagger A)z)=A^\dagger f.$$

反之, 若  $A^\dagger Au=A^\dagger f$ , 则

$$u=u-A^\dagger Au+A^\dagger f=A^\dagger f+(I-A^\dagger A)u.$$

上述说明等式(5)成立。

基于引理 1.2, 方程(1)等价于以下极小化问题:

$$\begin{aligned} \min_{x \in \mathbb{R}^N} \|x\|_0 \\ \text{s. t. } A^\dagger f=A^\dagger Ax. \end{aligned} \quad (7)$$

证明了式(1)与式(7)的等价性之后, 基于 HTP 算法的思想建立  $A^\dagger$  硬阈值追踪( $A^\dagger$  hard thresholding pursuit, APHTP)。与 HTP 相比新算法是将 HTP 中的  $A$  替换为 Moore-Penrose 逆  $A^\dagger$ 。需要指出的是  $x^n$  是稀疏的,  $x^n-A^\dagger Ax^n$  比  $x^n-Ax^n$  更逼近 0, 且该格式降低了对矩阵  $A$  的要求(矩阵  $A$  满足极强的 RIP 条件才能使  $A^\dagger A$  十分接近  $I$ )。此时从  $x^n+A^\dagger(f-Ax^n) \approx A^\dagger f = \arg \min \|f-Ax\|_2$ , 也即从近似的最小二乘解中选择  $x$  的支撑集。如算法 2 所示。

### 算法 2 APHTP

参数:  $A, f, k, \text{threshold} > 0$

初始化:  $x^0=0, r_0=f, S^0=\emptyset$

While  $\|r_n\|_2 > \text{threshold}$  do

(APHTP<sub>1</sub>) 计算  $S^{n+1} = \{ |x^n + A^\dagger(f - Ax^n)| \text{ 中 } k \text{ 个模量最大值对应的指标集} \}$

(APHTP<sub>2</sub>) 计算  $x^{n+1} = \arg \min \|f - Az\|_2, \text{supp}(z) \subseteq S^{n+1}$

(APHTP<sub>3</sub>)  $r_{n+1} = f - Ax^{n+1}$

end while

return  $x^{n+1}$

与 HTP 一样, APHTP 算法需要稀疏度  $k$  的先验估计, 然而在大多数实际问题中稀疏度  $k$  是未知的。因此, 本文进一步采用 OMP 算法的分级策略提出了  $A^\dagger$  分级硬阈值追踪算法( $A^\dagger$  Graded Hard Thresholding Pursuit, APGHTP), 如算法 3 所示。

### 算法 3 APGHTP

参数:  $A, f, k, \text{threshold} > 0$

初始化:  $x^0=0, r_0=f, S^0=\emptyset$

While  $\|r_n\|_2 > \text{threshold}$  do

(APGHTP<sub>1</sub>) 计算  $S^{n+1} = \{ |x^n + A^\dagger(f - Ax^n)| \text{ 中 } n+1 \text{ 个模量最大值对应的指标集} \}$

(APGHTP<sub>2</sub>) 计算  $x^{n+1} = \arg \min \|f - Az\|_2, \text{supp}(z) \subseteq S^{n+1}$

(APGHTP<sub>3</sub>)  $r_{n+1} = f - Ax^{n+1}$

end while

return  $x^{n+1}$

## 2 理论保证

在给出这节的主要结论之前, 先给出定理 2.1, 该定理是证明算法收敛的关键性结论。

定理 2.1 假设  $A \in \mathbb{R}^{m \times N}$  满足  $k$  阶 RIP 条件, 则

$$\|I - A^\dagger A\|_2 \leq \frac{2\delta_k}{1 + \delta_k}. \quad (8)$$

证明 矩阵  $A$  满足  $k$  阶 RIP 条件, 即任意  $k$  稀疏向量  $x \in \mathbb{R}^N, (1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k) \|x\|_2^2$ , 因此得

$$\begin{aligned} (1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 = \|AA^\dagger Ax\|_2^2 = \|A(A^\dagger Ax)\|_2^2 \leq (1 + \delta_k) \|A^\dagger Ax\|_2^2, \\ \frac{1 - \delta_k}{1 + \delta_k} \|x\|_2^2 \leq \|A^\dagger Ax\|_2^2. \end{aligned} \quad (9)$$

不等式 (9) 说明

$$\frac{1 - \delta_k}{1 + \delta_k} \mathbf{x}' \mathbf{x} \leq (\mathbf{A}^\dagger \mathbf{A} \mathbf{x})' (\mathbf{A}^\dagger \mathbf{A} \mathbf{x}) \leq \mathbf{x}' \mathbf{x},$$

则可得

$$0 \leq \mathbf{x}' \mathbf{x} - \mathbf{x}' (\mathbf{A}^\dagger \mathbf{A}) \mathbf{x} \leq \mathbf{x}' \mathbf{x} - \frac{1 - \delta_k}{1 + \delta_k} \mathbf{x}' \mathbf{x} = \frac{2\delta_k}{1 + \delta_k} \mathbf{x}' \mathbf{x},$$

于是有式 (8) 成立。

APHTP 和 APGHTP 的第一步和第二步具有相似的结果 因此我们给出以下一般结果。

第一步结果

若  $\mathbf{x} \in \mathbf{R}^N$  是  $k$  稀疏的,  $\mathbf{f} = \mathbf{A}\mathbf{x} + \mathbf{e}$  且  $\mathbf{e} \in \mathbf{R}^m$ ,  $\mathbf{x}'$  是  $k$  稀疏的,  $T$  是  $|\mathbf{x}' + \mathbf{A}^\dagger(\mathbf{f} - \mathbf{A}\mathbf{x}')|$  的  $t \geq k$  个模量最大值对应的指标集 则

$$\|\mathbf{x}_{\bar{T}}\|_2 \leq 2\sqrt{2} \frac{\delta_{k+t+k'}}{1 + \delta_{k+t+k'}} \|\mathbf{x} - \mathbf{x}'\|_2 + \sqrt{2} \|(\mathbf{A}^\dagger \mathbf{e})_{T\Delta S}\|_2, \quad (10)$$

其中  $S = \text{supp}(\mathbf{x})$ 。

由  $S = \text{supp}(\mathbf{x})$  与  $T$  是  $|\mathbf{x}' + \mathbf{A}^\dagger(\mathbf{f} - \mathbf{A}\mathbf{x}')|$  的  $t \geq k$  个模量最大值对应的指标集 可得

$$\|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_T\|_2^2 \geq \|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_S\|_2^2, \quad (11)$$

消除  $S \cap T$  元素的影响

$$\|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{T \setminus S}\|_2 \geq \|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{S \setminus T}\|_2 \quad (12)$$

式 (12) 左端满足

$$\|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{T \setminus S}\|_2 = \|(\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e}\|_{T \setminus S},$$

同时式 (12) 右端满足

$$\|(\mathbf{x}' + \mathbf{A}^\dagger \mathbf{A}(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{S \setminus T}\|_2 \geq \|\mathbf{x}_{S \setminus T}\|_2 - \|(\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e}\|_{S \setminus T},$$

则

$$\begin{aligned} \|\mathbf{x}_{S \setminus T}\|_2 &\leq \|((\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{S \setminus T}\|_2 + \|((\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{S \setminus T}\|_2 \leq \\ &\sqrt{2} \|((\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}') + \mathbf{A}^\dagger \mathbf{e})_{T\Delta S}\|_2 \leq \\ &\sqrt{2} (\|(\mathbf{A}^\dagger \mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x}')_{T\Delta S}\|_2 + \|(\mathbf{A}^\dagger \mathbf{e})_{T\Delta S}\|_2) \leq \\ &2\sqrt{2} \frac{\delta_{k+t+k'}}{1 + \delta_{k+t+k'}} \|\mathbf{x} - \mathbf{x}'\|_2 + \sqrt{2} \|(\mathbf{A}^\dagger \mathbf{e})_{T\Delta S}\|_2. \end{aligned}$$

第二步结果

若  $\mathbf{x} \in \mathbf{R}^N$  是  $k$  稀疏的,  $\mathbf{f} = \mathbf{A}\mathbf{x} + \mathbf{e}$  且  $\mathbf{e} \in \mathbf{R}^m$ ,  $\mathbf{x}'$  是  $k$  稀疏的,  $T$  是大小为  $t$  的支撑集,  $\mathbf{x}'$  是  $\|\mathbf{f} - \mathbf{A}\mathbf{z}\|$  对应于  $\text{supp}(\mathbf{z}) \subseteq T$  的极小值解 则

$$\|\mathbf{x} - \mathbf{x}'\|_2 \leq \sqrt{\frac{1}{1 - 2\delta_{k+t}^2}} \|\mathbf{x}_T\|_2 + \sqrt{\frac{2}{1 - 2\delta_{k+t}^2}} \|(\mathbf{A}'\mathbf{e})_T\|_2. \quad (13)$$

由  $\mathbf{x}'$  的定义  $\mathbf{x}'$  满足  $(\mathbf{A}'(\mathbf{f} - \mathbf{A}\mathbf{x}'))_T = 0$  则

$$\begin{aligned} \|(\mathbf{x} - \mathbf{x}')_T\|_2^2 &= \langle (\mathbf{x} - \mathbf{x}')_T, (\mathbf{x} - \mathbf{x}')_T \rangle = \langle (\mathbf{x} - \mathbf{x}' - \mathbf{A}'(\mathbf{f} - \mathbf{A}\mathbf{x}'))_T, (\mathbf{x} - \mathbf{x}')_T \rangle \leq \\ &|\langle (\mathbf{I} - \mathbf{A}'\mathbf{A})(\mathbf{x} - \mathbf{x}')_T, (\mathbf{x} - \mathbf{x}')_T \rangle| + |\langle \mathbf{A}\mathbf{e}_T, (\mathbf{x} - \mathbf{x}')_T \rangle| \leq \\ &\|(\mathbf{I} - \mathbf{A}'\mathbf{A})(\mathbf{x} - \mathbf{x}')_T\|_2 \|\mathbf{x} - \mathbf{x}'\|_2 + \|(\mathbf{A}\mathbf{e})_T\|_2 \|\mathbf{x} - \mathbf{x}'\|_2 \leq \\ &\delta_{k+t} \|(\mathbf{x} - \mathbf{x}')_T\|_2 \|\mathbf{x} - \mathbf{x}'\|_2 + \|(\mathbf{A}\mathbf{e})_T\|_2 \|\mathbf{x} - \mathbf{x}'\|_2. \end{aligned} \quad (14)$$

式 (14) 两边同时除以  $\|(\mathbf{x} - \mathbf{x}')_T\|_2$ , 可得

$$\|(\mathbf{x} - \mathbf{x}')_T\|_2 \leq \delta_{k+t} \|\mathbf{x} - \mathbf{x}'\|_2 + \|(\mathbf{A}\mathbf{e})_T\|_2.$$

这说明

$$\|\mathbf{x} - \mathbf{x}'\|_2^2 = \|(\mathbf{x} - \mathbf{x}')_T\|_2^2 + \|(\mathbf{x} - \mathbf{x}')_{\bar{T}}\|_2^2 \leq (\delta_{k+t} \|\mathbf{x} - \mathbf{x}'\|_2 + \|(\mathbf{A}\mathbf{e})_T\|_2)^2 + \|\mathbf{x}_{\bar{T}}\|_2^2. \quad (15)$$

式 (15) 表明

$$\|\mathbf{x} - \mathbf{x}'\|_2^2 \leq 2(\delta_{k+t}^2 \|\mathbf{x} - \mathbf{x}'\|_2^2 + \|(\mathbf{A}\mathbf{e})_T\|_2^2) + \|\mathbf{x}_{\bar{T}}\|_2^2,$$

则有式(13)成立。

有了上述结果,接下来证明 APHTP 和 APGHTP 的收敛性。

**定理 2.2** 对于满足  $f = Ax + e$  的任意  $k$  稀疏向量  $x \in \mathbf{R}^N$ ,若矩阵  $A$  的  $3k$  阶约束等距常数满足:

$$\delta_{3k} < 0.4068,$$

则由 APHTP 生成的序列  $\{x^n\}$  趋于  $x$  且

$$\|x - x^n\|_2 \leq \rho_{3k} \|x - x^{n-1}\|_2 + \tau_{2k} \|e\|_2, \rho_{3k} < 1, \forall n \geq 1, \quad (16)$$

$$\text{其中 } \rho_{3k} = \frac{2\delta_{3k}}{1 + \delta_{3k}} \sqrt{\frac{2}{1 - 2\delta_{3k}^2}}, \tau_{2k} = \sqrt{\frac{2}{1 - 2\delta_{2k}^2}} \left( \sqrt{1 + \delta_{2k}} + \frac{1}{\sqrt{1 - \delta_{2k}}} \right).$$

证明 在式(10)中令  $x' = x^{n-1}$ ,  $T = S^n$ ,可得

$$\|x_{S^n}\|_2 \leq 2\sqrt{2} \frac{\delta_{3k}}{1 + \delta_{3k}} \|x - x^{n-1}\|_2 + \sqrt{2} \|(A^\dagger e)_{S^n \Delta S}\|_2. \quad (17)$$

相似地,将  $x' = x^n$ ,  $T = S^n$  代入式(13)则

$$\|x - x^n\|_2 \leq \sqrt{\frac{1}{1 - 2\delta_{2k}^2}} \|x_{S^n}\|_2 + \sqrt{\frac{2}{1 - 2\delta_{2k}^2}} \|(A^\dagger e)_{S^n}\|_2. \quad (18)$$

将式(17)和式(18)合并整理可得

$$\|x - x^n\|_2 \leq \frac{2\sqrt{2}\delta_{3k}}{1 + \delta_{3k}} \sqrt{\frac{1}{1 - 2\delta_{2k}^2}} \|x - x^{n-1}\|_2 + \sqrt{\frac{2}{1 - 2\delta_{2k}^2}} \|(A^\dagger e)_{S^n \cup S}\|_2 + \sqrt{\frac{2}{1 - 2\delta_{2k}^2}} \|(A^\dagger e)_{S^n}\|_2.$$

最后由  $\|(A^\dagger e)_{S^n}\|_2 \leq \sqrt{1 + \delta_{2k}} \|e\|_2$  及  $\|(A^\dagger e)_{S^n \cup S}\|_2 \leq \frac{1}{\sqrt{1 - \delta_{2k}}} \|e\|_2$  可得式(16),且令  $\rho_{3k} < 1$  知

$$\delta_{3k} < 0.4068.$$

接下来给出 APGHTP 的收敛定理。

**定理 2.3** 对于满足  $f = Ax + e$  的任意  $k$  稀疏向量  $x \in \mathbf{R}^N$ ,如果  $\{x^n\}$  是 APGHTP 生成的序列,则

$$\|x - x^n\|_2 \leq \rho_{k+2n-1} \|x - x^{n-1}\|_2 + \tau_{k+n} \|e\|_2, \forall n \geq k, \quad (19)$$

其中:

$$\rho_{k+2n-1} = \frac{2\delta_{k+2n-1}}{1 + \delta_{k+2n-1}} \sqrt{\frac{2}{1 - 2\delta_{k+2n-1}^2}}, \tau_{k+n} = \sqrt{\frac{2}{1 - 2\delta_{k+n}^2}} \left( \sqrt{1 + \delta_{k+n}} + \frac{1}{\sqrt{1 - \delta_{k+n}}} \right).$$

证明 式(10)中令  $x' = x^{n-1}$ ,  $T = S^n$ . 稀疏向量  $x'$  包含  $n-1$  个非零元素且支撑集  $T$  包含  $n$  个指标. 式(13)中令  $x' = x^n$ ,  $T = S^n$ ,则  $x'$  的稀疏度和  $T$  的基数均为  $n$ . 剩余证明类似于定理 2.2.

### 3 数值实验

这部分将给出 APHTP、APGHTP、HTP、GHTP、CoSaMP、IHT 及线性化 Bregman 算法在加噪以及无噪的压缩感知问题下的表现. IHT 是用于求解压缩感知问题的 IHT. 在所有测试中,矩阵  $A \in \mathbf{R}^{m \times N}$  为  $m = 200$ ,  $N = 1000$  的高斯矩阵且所有向量为高斯向量. 在实验中记录各算法成功重建的总次数,该次数与总实验次数的比值即为各算法的恢复成功率.

#### 3.1 无噪情况下的表现

本小节给出各算法在无噪情况下的表现,所有结果均为 500 次实验后的平均结果. 各算法的停机准则为  $\|f - Ax^n\|_2 \leq 10^{-8}$  或达到最大迭代步数 500,且一次重建的相对误差  $\|x - x^n\|_2 \leq 10^{-4} \|x\|_2$  则为成功.

图(1)给出了恢复成功率. 显然 APGHTP 的表现最好且略优于 GHTP, CoSaMP 与 HTP 表现相近. 同时, APGHTP、GHTP 在区间  $[0, 60]$  上恢复成功率近乎为 100%, 而 IHT、HTP、CoSaMP 恢复成功率为 100% 的区间大致为  $[0, 45]$ . APHTP 与线性化 Bregman 算法的区间相近为  $[0, 20]$ .

接下来给出恢复稀疏向量时各算法所需迭代数. 图(2)说明在 APGHTP、IHT、CoSaMP 三种算法恢复成功率均为 100% 在区间  $[0, 45]$  上, APGHTP、GHTP 恢复  $k$  稀疏向量时所用迭代次数为  $k$  且与 IHT 所需迭代数相差无几,同时 CoSaMP 与 HTP 在该区间上所需迭代数相当. 在整个区间上, HTP 所需的迭代数最少, APGHTP 所需迭代数不超过 200, 而线性化 Bregman 算法所需的迭代数是最多的.

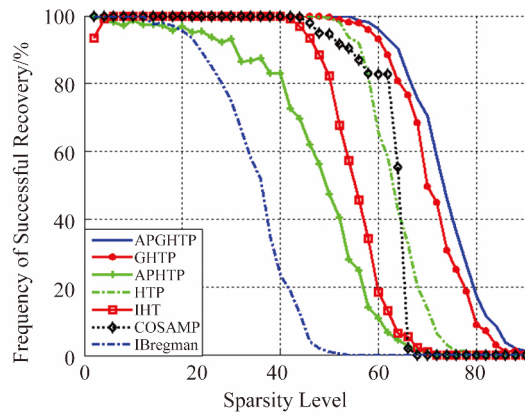
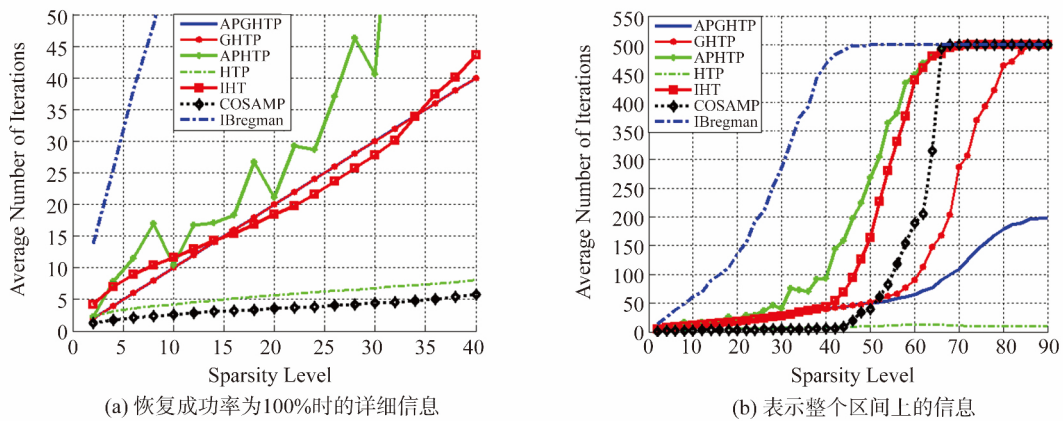


图 1 无噪情况下各种算法恢复成功率

Fig. 1 Numerical test of different algorithms on frequency of success in ideal situation



(a) 恢复成功率为100%时的详细信息

(b) 表示整个区间上的信息

图 2 无噪情况下各种算法所需迭代数

Fig. 2 Average numbers of iterations of different algorithms in ideal situation

### 3.2 加噪情况下的表现

本小节给出各算法在加噪情况下的表现,所有结果均为 500 次实验后的平均结果。在本小节实验中随机生成  $m$  维单位高斯误差向量  $e$ ,令观测值  $f = Ax + \tau e$ ,其中  $\tau = 0.01 \|x\|_2$ 。此时各算法的停机准则为  $\|f - Ax^n\|_2 \leq \tau$  或达到最大迭代步数 500,一次重建的相对误差  $\|x - x^n\|_2 \leq 10^{-2} \|x\|_2$  则为成功。

当测量值包含噪声时,图(3)说明 APGHTP、GHTP 的表现极大地优于其它算法,APHTP 与 CoSaMP 的表现相近,IHT 表现最差,原因可能是所选择的停机准则使得 IHT 过早的停止。从迭代数来看,图(4)再次说明在恢复成功率为 100% 的区间上,APGHTP、GHTP 恢复  $k$  稀疏向量时所需的迭代数为  $k$ ,同时在整个区间上 APGHTP 与 GHTP 所需迭代数几乎相同且最大迭代数不超过 150。线性化 Bregman 算法在区间  $[0, 50]$  所需迭代数与 APHTP 相似,当稀疏度  $k > 50$  时其所需迭代数逐渐降低到 150 附近,但此时该算法的恢复成功率为 0。HTP 在该情况下所需的迭代数在整个区间上是最小的而 CoSaMP 是最大的。

### 3.3 运行时间

本小节给出各算法在无噪情况下的平均运行时间,该时间为 100 次实验后的平均结果。各算法停机准则与节 3.1 中的停机准则相同。

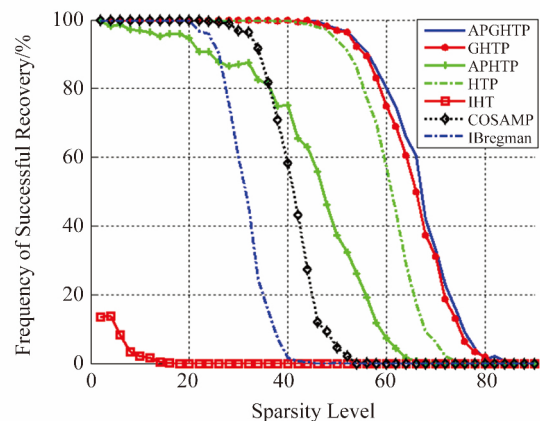


图 3 加噪情况下各种算法恢复成功率

Fig. 3 Numerical test of different algorithms on frequency of success in realistic situation

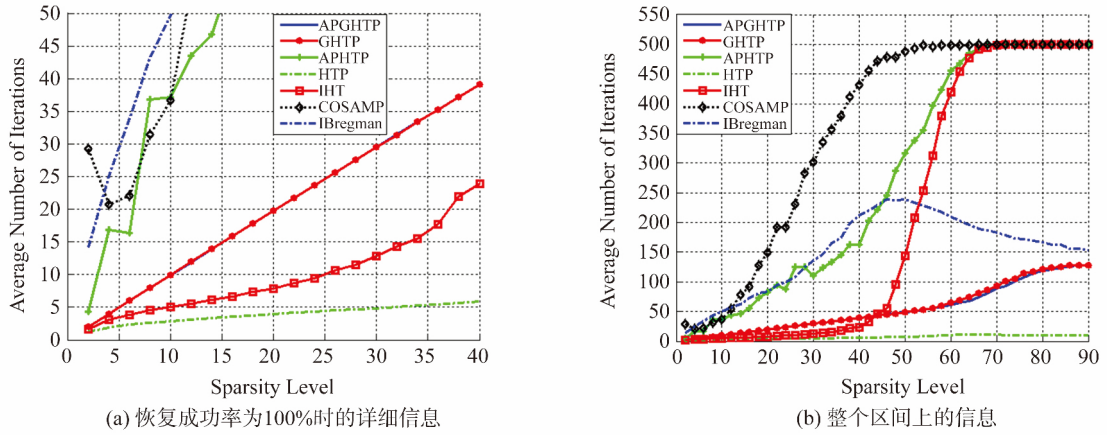


图 4 加噪情况下各种算法所需迭代数  
Fig. 4 Average numbers of iterations of different algorithms in realistic situation

图 5 表明,当信号高度稀疏时,HTP 所需的运行时间是最少的,而 APHTP 与 APGHTP 所需时间相当且略多,这与计算  $A^\dagger$  有关(如有必要,可以采用高度稀疏时的  $A^\dagger$  近似算法来减少运行时间)。随着稀疏度的不断增加各算法所需时间也不断增加。由图 1、图 2 可知,当稀疏度  $k > 50$  时,所有算法的恢复成功率开始降低且所需迭代数迅速增加到所限定的最大迭代数 500,故运行时间也迅速增加。特别地,尽管此时线性 Bregman 算法、IHT 所需时间保持稳定,但恢复成功率相对较低。而 CoSaMP、GHTP 所需的时间迅速的赶上了其它算法所需的时间、且远远大于其它算法。

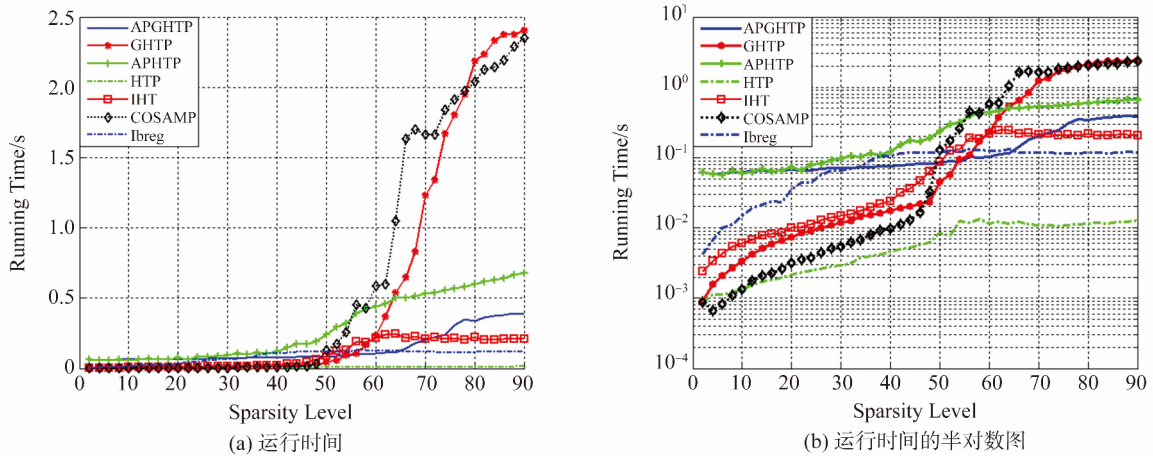


图 5 无噪情况下各算法平均运行时间及其半对数图  
Fig. 5 Running time of different algorithms and its semilog plot

### 4 结论

本文提出了  $A^\dagger$  硬阈值追踪算法 (APHTP),该算法与硬阈值追踪算法 (HTP) 的主要区别是引入 Moore-Penrose 逆  $A^\dagger$  代替  $A'$ 。 $A^\dagger$  分级硬阈值追踪算法 (APGHTP) 是 APHTP 的变体,其最大优点在于避免了稀疏度的估计。数值实验说明 APGHTP 表现优于 HTP, GHTP, CoSaMP, IHT 及线性化 Bregman 算法。在恢复  $k$  稀疏向量时, APGHTP 所需的迭代数与稀疏向量的稀疏度相同。尽管计算 Moore-Penrose 逆的代价较大,但在两种情况下 APGHTP 良好的表现、需要较少的迭代数以及算法的简单性使得  $A^\dagger$  分级硬阈值追踪算法 (APGHTP) 在存在噪声且稀疏度未知的情况下可以作为一个较好的选择。

#### 参考文献:

[1] NATARAJAN B K. Sparse approximate solutions to linear systems [J]. SIAM Journal on Computing, 1995, 24(2): 227-234.

(下转第 69 页)

- [5] HU Shuhe, WANG Xuejun, YANG Wenzhi, et al. The Hájek-Rényi-type inequality for associated random variables[J]. *Statistics and Probability Letters*, 2009, 79(7): 884-888.
- [6] WANG Xuejun, HU Shuhe. Maximal inequalities for demimartingales and their applications[J]. *Science in China Series A: Mathematics*, 2009, 52(10): 2207-2217.
- [7] WANG Xuejun, PRAKASA RAO B L S, HU Shuhe. On some maximal inequalities for demisubmartingales and N-demisupermartingales based on concave Young functions[J]. *Journal of Mathematical Analysis and Applications*, 2012, 396(2): 434-440.
- [8] PRAKASA RAO B L S. On some maximal inequalities for demisubmartingales and N-demisupermartingales[J]. *Journal of Inequalities in Pure and Applied Mathematics*, 2007, 8(4): 112.
- [9] PRAKASA RAO B L S. Remarks on maximal inequalities for nonnegative demisubmartingales[J]. *Statistics and Probability Letters*, 2012, 82(7): 1388-1390.
- [10] PRAKASA RAO B L S. Conditional independence, conditional mixing and conditional association[J]. *Annals of the Institute of Statistical Mathematics*, 2009, 61(2): 441-460.
- [11] 胡舒合, 杨文志, 王学军 等. 关于 N-弱鞅和弱鞅不等式的一个注记[J]. *系统科学与数学*, 2010, 30(8): 1052-1058.  
HU Shuhe, YANG Wenzhi, WANG Xuejun, et al. A note on the inequalities for N-demimartingales and demimartingales[J]. *Journal of Systems Science and Mathematical Sciences*, 2010, 30(8): 1052-1058.
- [12] CHRISTOFIDES T C. Maximal inequalities for N-demimartingales[J]. *Archives of Inequalities and Applications*, 2003, 1: 387-398.
- [13] WANG Xinghui, WANG Xuejun. Some inequalities for conditional demisubmartingales and conditional N-demisupermartingales[J]. *Statistics and Probability Letters*, 2013, 83(7): 700-709.
- [14] DAI Pingping, SHEN Yan, HU Shuhe, et al. Some results for demimartingales and N-demimartingales[J]. *Journal of Inequalities and Applications*, 2014(1): 489-501.
- [15] 冯德成, 蒋文君, 陈彩龙. 弱鞅的极小值不等式(英)[J]. *兰州文理学院学报: 自然科学版*, 2015, 29(2): 1-4.  
FENG Decheng, JIANG Wenjun, CHEN Cailong. On minimal inequalities of demimartingales[J]. *Journal of Lanzhou University of Arts and Science: Natural Sciences*, 2015, 29(2): 1-4.

(编辑: 许力琴)

(上接第 64 页)

- [2] CANDÈS E J, ROMBERG J K, TAO TERENCE. Stable signal recovery from incomplete and inaccurate measurements[J]. *Communications on Pure & Applied Mathematics*, 2005, 59(8): 410-412.
- [3] DONOHO D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [4] DONOHO D L. Neighborly polytopes and sparse solution of underdetermined linear equations[EB/OL]. [2016-04-06]. [http://statweb.stanford.edu/~donoho/Reports/2005/NPa\\_Ssule-01-28-05.pdf](http://statweb.stanford.edu/~donoho/Reports/2005/NPa_Ssule-01-28-05.pdf)
- [5] DONOHO D L. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension[J]. *Discrete & Computational Geometry*, 2006, 35(4): 617-652.
- [6] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666.
- [7] NEEDELL D, TROPP J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples[J]. *Applied & Computational Harmonic Analysis*, 2008, 26(3): 301-321.
- [8] BLUMENSATH T, DAVIES M E. Iterative hard thresholding for compressed sensing[J]. *Applied & Computational Harmonic Analysis*, 2009, 27(3): 265-274.
- [9] DONOHO D L. De-noising by soft-thresholding[J]. *IEEE Transactions on Information Theory*, 1995, 41(3): 613-627.
- [10] YIN Wotao, OSHER S, GOLDFARB D, et al. Bregman iterative algorithms for  $\ell_1$ -minimization with applications to compressed sensing[J]. *Siam Journal on Imaging Sciences*, 2008, 1(1): 143-168.
- [11] YIN Wotao. Analysis and generalizations of the linearized bregman method[J]. *Siam Journal on Imaging Sciences*, 2010, 3(4): 856-877.
- [12] FOUCAIT S. Hard thresholding pursuit: an algorithm for compressive sensing[J]. *Siam Journal on Numerical Analysis*, 2011, 49(6): 2543-2563.
- [13] BOUCHOT J L, FOUCAIT S, HITCZENKO P. Hard thresholding pursuit algorithms: number of iterations[J]. *Applied & Computational Harmonic Analysis*, 2016, 41(2): 412-435.
- [14] BEN-ISRAEL A, GREVILLE T N E. Generalized inverses[M]. New York: Springer-Verlag, 2003.

(编辑: 许力琴)