

# The Solution of Sparsity-Constrained Split Feasibility Problem

Hanxiao Chang, Jun Sun, Biao Qu

School of Management, Qufu Normal University, Rizhao Shandong  
Email: sunjun901203@163.com

Received: May 4<sup>th</sup>, 2016; accepted: May 23<sup>rd</sup>, 2016; published: May 26<sup>th</sup>, 2016

Copyright © 2016 by authors and Hans Publishers Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this paper, we mainly study the solution of sparsity-constrained split feasibility problem. Under some reasonable assumptions, we use IHT algorithm to get the stationary points of sparsity-constrained split feasibility problem and get a conclusion which plays an important role in local convergence analysis.

## Keywords

Split Feasible Problem, Sparsity-Constrained, IHT Algorithm, Stationery

---

# 带稀疏约束的分裂可行问题的算法

畅含笑, 孙 军, 屈 彪

曲阜师范大学管理学院, 山东 日照  
Email: sunjun901203@163.com

收稿日期: 2016年5月4日; 录用日期: 2016年5月23日; 发布日期: 2016年5月26日

---

## 摘 要

本文, 我们主要研究带稀疏约束的分裂可行问题。在某些合理的假设下用IHT算法, 得到了带稀疏约束的分裂可行问题的稳定点及给出在局部收敛性分析中起到了重要作用的结论。

## 关键词

分裂可行问题, 稀疏约束, IHT算法, 稳定点

## 1. 引言

近几年来, 带稀疏约束的最优化问题得到了学者们广泛的关注, 这是因为它在信号处理、图像恢复等领域有着重要的应用[1]-[4]。不论是在理论、算法还是在应用方面, 已经有了大量的研究。

大多数的理论研究都是关于带稀疏约束的一般连续可微函数, 问题形式如下:

$$\min f(x) \text{ 使得 } \|x\|_0 \leq s \quad (1.1)$$

其中,  $f: R^N \rightarrow R$  是一个连续可微的函数  $s > 0$ , 是一个小于  $N$  的整数。 $\|\cdot\|_0$  代表零范数(是向量中非零分量的个数)。目前, 问题(1.1)已经引起了大量学者的研究, 并提出了许多算法, 但大多数的算法产生的点列都只是收敛到问题(1.1)的稳定点。在[5]中, 作者引用了一种叫 IHT 的算法用来解问题(1.1), 并证明由该算法产生的点列的任意聚点都为问题(1.1)的  $L$ -稳定点。

关于问题(1.1)的特殊形式  $f(x) = \frac{1}{2} \|Ax - b\|^2$ , 其中  $A \in R^{M \times N}$ ,  $b \in R^M$ , 也得到了广大学者的广泛关注[6]-[8]。

本文, 我们主要研究的是带稀疏约束的分裂可行问题, 分裂可行问题就是下述形式的问题, 最早被 Censor 在[9]中提出:

找向量

$$x \in C, Ax \in Q \quad (1.2)$$

其中,  $C, Q$  分别为  $R^N$  和  $R^M$  中的非空闭凸集,  $A$  为一个  $M \times N$  的实矩阵。

本文研究的是带稀疏约束的分裂可行问题, 问题形式如下:

找向量

$$x \in C, Ax \in Q, \text{ 使得 } \|x\|_0 \leq s \quad (1.3)$$

受到[5]的启发, 我们通过求解下述最优问题得到(1.3)的近似解

$$\begin{aligned} \min \frac{1}{2} \|Ax - P_Q(Ax)\|^2 + \frac{1}{2} \|x - P_C(x)\|^2 \\ \|x\|_0 \leq s \end{aligned} \quad (1.4)$$

注 1: 在本文中我们取  $f(x) = \min \frac{1}{2} \|Ax - P_Q(Ax)\|^2 + \frac{1}{2} \|x - P_C(x)\|^2$ 。

## 2. 预备知识

### 2.1. 支撑映射

**定义:** 对任意的非空闭集  $\Omega \subset R^N$ , 称  $P_\Omega(x) = \arg \min_{y \in \Omega} \|x - y\|$  为  $x$  到  $\Omega$  的投影。称由  $R^N$  往稀疏集  $S$  上的投影为支撑映射。

显然, 往稀疏集  $S$  上的投影  $P_S(x)$  的分量包括  $s$  个绝对值最大的分量和  $N - s$  个零。即: 设  $I_s(x) = \{j_1, j_2, \dots, j_s\} \subseteq \{1, 2, \dots, N\}$  为  $x$  的指标集, 且

$$\min_{i \in I_s(x)} |x_i| \geq \max_{i \notin I_s(x)} |x_i|,$$

则

$$P_S(x) = \{y \in R^N \mid y_i = x_i, i \in I_S(x); y_i = 0, i \notin I_S(x)\}.$$

因为  $S$  非凸的, 所以  $P_S(x)$  一般不是一个单点集. 为使  $P_S(x)$  为单点集, 我们这样定义  $P_S(x)$ : 设  $\mu_i(x)$  代表  $x$  的第  $i$  大的分量, 则有  $\mu_1(x) \geq \mu_2(x) \geq \dots \geq \mu_N(x)$ . 记  $|x| = (|x_1|, |x_2|, \dots, |x_N|)^T$ . 当  $|x_i| \geq \mu_s(|x|)$  时, 取  $(P_S(x))_i = x_i$ , 如果有多于一个的  $x_i = \mu_s(|x|)$  时, 随机从中选一个  $(P_S(x))_i = x_i$ ; 否则, 取  $(P_S(x))_i = 0$ . 即:

$$(P_S(x))_i \begin{cases} x_i, & |x_i| > \mu_s(|x|) \\ x_i \text{ 或 } 0 & |x_i| = \mu_s(|x|) \\ 0 & \text{否则.} \end{cases}$$

### 2.2. $L$ -稳定点

定义: 如果对任意的  $L > 0$ ,  $x^* \in S$ , 满足如下:

$$x^* \in P_S\left(x^* - \frac{1}{L} \nabla f(x^*)\right) \quad (2.1)$$

则称  $x^*$  为(1.4)的稳定点. 其中,  $S = \{x \in R^N \mid \|x\|_0 \leq s\}$  为(1.4)的可行域,

$$\nabla f(x^*) = A^T(I - P_Q)Ax^* + (I - P_C)x^*.$$

### 2.3. 导集

定义: 记集合  $B$  的所有的聚点组成的集合为  $B'$ , 称  $B'$  为  $B$  的导集.

注 2: 导集的一些性质: 1) 若  $B$  为闭集, 则  $B' \subset B$ ; 2) 记  $\bar{B}$  为  $B$  的闭包, 则  $\bar{B} = B \cup B'$ .

引理 2.1: 对任意的  $L > 0$ ,  $x^*$  满足(2.1)当且仅当  $\|x^*\|_0 \leq 0$ , 且

$$\begin{cases} \left| \nabla_i f(x^*) \right| \leq L \mu_s(x^*), & i \notin I_S(x^*); \\ \left| \nabla_i f(x^*) \right| = 0, & i \in I_S(x^*). \end{cases}$$

引理 2.2: (1.4)式中, 目标函数的梯度函数  $\nabla f$  在  $R^N$  上是 Lipschitz 连续的, 且 Lipschitz 常数为  $L = \|A\|^2 + 1$ , 即:

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad x, y \in R^N$$

证明: 对任意的  $x, y \in R^N$ , 有

$$\begin{aligned} \|\nabla f(x) - \nabla f(y)\| &= \|A^T(I - P_Q)Ax + (I - P_C)x - A^T(I - P_Q)Ay - (I - P_C)y\| \\ &\leq \|A^T(I - P_Q)Ax - A^T(I - P_Q)Ay\| + \|(I - P_C)x - (I - P_C)y\| \\ &= \|A\| \|(I - P_Q)Ax - (I - P_Q)Ay\| + \|(I - P_C)x - (I - P_C)y\| \\ &\leq \|A\| \|Ax - Ay\| + \|x - y\| \\ &\leq \|A\|^2 \|x - y\| + \|x - y\| \\ &= (\|A\|^2 + 1) \|x - y\|. \end{aligned}$$

即:  $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$ 。

证毕。

**引理 2.3** (the descent lemma [10]): 假设  $f$  是一个连续可微的函数, 且其梯度函数是 Lipschitz 连续的, 并记其 Lipschitz 常数为  $L(f)$ , 则对任意  $L' \geq L(f)$ , 有

$$f(x) \leq h_{L'}(x, y), x, y \in \mathbb{R}^N$$

其中  $h_{L'}(x, y) \equiv f(y) + \langle \nabla f(y), x - y \rangle + \frac{L'}{2}\|x - y\|^2, x, y \in \mathbb{R}^N$ 。

**引理 2.4:** 对于任意的  $L' \geq L$ , 则对任意的  $x \in S, y \in \mathbb{R}^N$ , 满足于

$$y \in P_S \left( x - \frac{1}{L} \nabla f(x) \right) \quad (2.2)$$

则有  $f(x) - f(y) \geq \frac{L' - L}{2}\|x - y\|^2$ 。

其中,  $L = \|A\|^2 + 1, \nabla f(x) = A^T(I - P_Q)Ax + (I - P_C)x$ 。

证明: 我们知(2.2)式可以写成如下形式:

$$y \in \arg \min_{z \in S} \left\| z - \left( x - \frac{1}{L} \nabla f(x) \right) \right\|^2 \quad (2.3)$$

$$h_{L'}(z, x) = f(x) + \langle \nabla f(x), z - x \rangle + \frac{L'}{2}\|z - x\|^2$$

又因为

$$= \frac{L'}{2} \left\| z - \left( x - \frac{1}{L} \nabla f(x) \right) \right\|^2 + f(x) - \frac{1}{2L'} \|\nabla f(x)\|^2$$

又  $f(x) - \frac{1}{2L'} \|\nabla f(x)\|^2$  关于  $z$  为常数, 故(2.3)式等价于  $y \in \arg \min_{z \in S} h_{L'}(z, x)$ 。

即:

$$h_{L'}(y, x) \leq h_{L'}(x, x) = f(x) \quad (2.4)$$

由引理 2.4 可得  $f(x) - f(y) \geq f(x) - h_{L(f)}(y, x)$ 。

即:

$$f(y) \leq h_{L(f)}(y, x) \quad L(f) = L \quad (2.5)$$

联立(2.4), (2.5)可得  $f(x) - f(y) \geq \frac{L' - L}{2}\|x - y\|^2$ 。

证毕。

**定理 2.1:** 对于任意的  $L' \geq L$ , 设  $x^*$  为(1.4)式的最优解, 则

- (1)  $x^*$  是一个  $L$ -稳定点。
- (2)  $P_S \left( x^* - \frac{1}{L'} \nabla f(x^*) \right)$  为一个单点集。

证明: 我们同时证明(a)和(b), 假设存在向量  $y \neq x^*$ , 且有

$$y \in P_S \left( x^* - \frac{1}{L'} \nabla f(x^*) \right)$$

在引理 2.4 中取  $x = x^*$ , 有

$$f(x^*) - f(y) \geq \frac{L' - L}{2} \|x^* - y\|^2$$

这与  $x^*$  为最优解矛盾。

证毕。

### 3. IHT 算法

在本文我们引用求解带稀疏约束的一般连续可微函数的 IHT 算法来求解带稀疏约束的分裂可行问题。算法过程如下：

步骤 1: 任取  $L' \geq L, x^0 \in S$ ;

步骤 2:  $x^{k+1} = P_S \left( x^k - \frac{1}{L'} \nabla f(x^k) \right), k = 0, 1, 2, \dots$ ;

步骤 3: 若有  $\|x^{k+1} - x^k\| \leq \varepsilon$ , 则停止; 否则令  $k \leftarrow k+1$ , 转步骤 2。

其中,  $\nabla f(x^k) = A^T(I - P_Q)Ax^k + (I - P_C)x^k$ ,  $S$  稀疏集  $\varepsilon \geq 0$ 。

**引理 3.1:** 设  $\{x^k\}$  为由 IHT 算法产生的点列, 则

$$(1) f(x^k) - f(x^{k+1}) \geq \frac{L' - L}{2} \|x^k - x^{k+1}\|^2;$$

(2)  $\{f(x^k)\}$  是一个不增数列;

$$(3) \|x^k - x^{k+1}\| \rightarrow 0。$$

证明: 在引理 2.4 中, 取  $x = x^k, y = x^{k+1}$ , 从而由(1)成立。又由(1)可直接得出(2), (3)成立。

**定理 3.1:** 设  $\{x^k\}$  为由 IHT 算法产生的点列, 则  $\{x^k\}$  的任一聚点都为  $L$ -稳定点。

证明: 该证明与[5]中定理 3.1 类似, 证明过程如下:

设  $x^*$  为  $\{x^k\}$  的任意聚点, 则存在子列  $\{x^{k_n}\}$  收敛到  $x^*$ 。由引理 3.1 的(1)可得:

$$f(x^k) - f(x^{k+1}) \geq \frac{L' - L}{2} \|x^k - x^{k+1}\|^2 \quad (2.6)$$

因为  $\{f(x^k)\}$  为正数列, 结合引理 3.1 中(2), 可以得出  $\{f(x^{k_n})\}$  和  $\{f(x^{k_n+1})\}$  收敛到同一极限  $f^*$ 。

从而有  $n \rightarrow \infty$  时,  $f(x^{k_n}) - f(x^{k_n+1}) \rightarrow 0$ , 联立(2.6)式可以得出:

当  $n \rightarrow \infty$  时,  $x^{k_n+1} \rightarrow x^*$ 。

又

$$x^{k_n+1} = P_S \left( x^{k_n} - \frac{1}{L'} \nabla f(x^{k_n}) \right)。$$

取  $i \in I_s(x^*)$ 。因为  $x^{k_n}$  和  $x^{k_n+1}$  都收敛到  $x^*$ , 所以存在  $N$  使得当  $n > N$  时, 有:

$$x_i^{k_n}, x_i^{k_n+1} \neq 0$$

从而对所有的  $n > N$  :

$$x_i^{k_n+1} = x_i^{k_n} - \frac{1}{L'} \nabla_i f(x^{k_n})$$

取  $n \rightarrow \infty$  可得  $\nabla_i f(x^*) = 0$ 。

再取  $i \notin I_s(x^*)$ 。如果存在有限个  $k_n$  且  $x_i^{k_n+1} \neq 0$ 。则

$$x_i^{k_n+1} = x_i^{k_n} - \frac{1}{L} \nabla_i f(x^{k_n})$$

取极限可得  $\nabla_i f(x^*) = 0$ 。特别地,  $|\nabla_i f(x^*)| \leq L\mu_s(|x^*|)$ 。另一方面, 如果存在  $M > 0$ , 使得对所有的  $n > M$  都有  $x_i^{k_n+1} = 0$ , 则

$$\left| x_i^{k_n} - \frac{1}{L} \nabla_i f(x^{k_n}) \right| \leq \mu_s \left( \left| x_i^{k_n} - \frac{1}{L} \nabla f(x^{k_n}) \right| \right) = \mu_s(|x^{k_n+1}|)$$

再取  $n \rightarrow \infty$ , 利用函数  $\mu_s(|\cdot|)$  的连续性可得:

$$|\nabla_i f(x^*)| \leq L\mu_s(|x^*|)$$

从而结论得证。

证毕。

**引理 3.2:** 设  $\{x^k\}$  为由 IHT 算法产生的点列, 且  $B$  为包含  $\{x^k\}$  的所有聚点的集合,  $B \neq \emptyset$ , 则①  $\lim_{k \rightarrow \infty} x^k = x^*$ ; 或②  $B = B'$ , 其中  $B'$  为  $B$  的导集。

证明: 假设①不成立, 则证②成立。因为  $B$  为闭集, 所以我们只需要证明  $B \subset B'$ 。

设  $x^* \in B$ , 故存在  $\{x^k\}$  的子列  $\{x^{k_t}\}$  使得:

$$\lim_{t \rightarrow \infty} x^{k_t} = x^* \tag{2.7}$$

因为①不成立, 故存在  $\{x^k\}$  和  $x^*$  的一个  $\delta_0$  邻域, 记作  $N(x^*, \delta_0)$  使得:

$$\{x^k\} \cap N(x^*, \delta_0) \neq \emptyset \tag{2.8}$$

取  $\delta \in (0, \delta_0)$ , 由引理 3.1 的(3)和(2.7)式知  $\exists t_0$ , 使得:

$$x^{k_t} \in N\left(x^*, \frac{\delta}{2}\right), \forall t \geq t_0 \tag{2.9}$$

$$\|x^{k+1} - x^k\|^2 \leq \frac{\delta}{2}, \forall k \geq k_0 \tag{2.10}$$

不失一般性, 对于  $\{x^k\}$  我们可以假设对  $\forall t \geq t_0, \exists \bar{k}_t \in N$ , 使得  $k_t < \bar{k}_t < k_{t+1}$  且有对于  $\forall t \geq t_0$ , 由(2.8)式和(2.9)式可得  $x^{k_t} \in N\left(x^*, \frac{\delta}{2}\right)$ ,  $x^{\bar{k}_t} \notin N(x^*, \delta)$ , 所以, 存在  $k_{t'}: k_t < k_{t'} \leq \bar{k}_t$  使得:

$$x^{k_{t'-1}} \in N\left(x^*, \frac{\delta}{2}\right) \tag{2.11}$$

$$x^{k_{t'}} \notin N\left(x^*, \frac{\delta}{2}\right) \tag{2.12}$$

由(2.10)和(2.11)式可得, 当  $t \geq t_0$  时有:

$$\|x^{k_{t'}} - x^*\| \leq \|x^{k_{t'}} - x^{k_{t'-1}}\| + \|x^{k_{t'-1}} - x^*\| \leq \delta$$

即点列  $\{x^{k_{t'}}\} \subset N(x^*, \delta)$ , 故  $\{x^{k_{t'}}\}$  存在聚点, 设为  $x'$ , 则  $x' \in N(x^*, \delta)$ , 由(2.12)式可得  $x' \neq x^*$ , 由  $\delta > 0$  的任意性, 得  $x^* \in B$ 。

证毕。

**定理 3.2:** 设  $\{x^k\}$  为 IHT 算法产生的点列, 若  $x^* \in B$  为(1.4)的严格局部最小解, 则有  $\lim_{k \rightarrow \infty} x^k = x^*$ 。

证明: 由于  $x^*$  为  $\{x^k\}$  的一个聚点, 故(2.7)式成立, 又  $f(x) = \frac{1}{2} \|Ax - P_Q(Ax)\|^2 + \frac{1}{2} \|x - P_C(x)\|^2$ , 结合引理 3.2 的(2)知, 由 IHT 算法产生的点列  $\{x^k\}$  满足  $0 \leq f(x^{k+1}) \leq f(x^k)$ , 所以

$$\lim_{k \rightarrow \infty} f(x^k) = \lim_{t \rightarrow \infty} f(x^{k_t}) = f(x^*).$$

即:

$$\forall x \in B, f(x) = f(x^*) \tag{2.13}$$

由引理 3.1 的(3)知  $\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0$ 。

由引理 3.2 知, 若  $\lim_{k \rightarrow \infty} x^k \neq x^*$ , 则对于任意的  $\delta > 0$ , 存在  $x^{t'} \in B, x^{t'} \neq x^*$ , 都有  $x^{t'} \in U(x^*, \delta)$ , 而由(2.13)式知  $f(x^{t'}) = f(x^*)$  对所有  $t'$  成立。故与  $x^* \in B$  为(1.4)的严格局部最小解矛盾。

证毕。

#### 4. 总结

本文在带稀疏约束的一般连续可微函数的已有结论的基础上, 对带稀疏约束的分裂可行问题进行求解分析。由引理 3.2 我们可以得出, 由 IHT 算法产生的点列,  $\{x^k\}$  要么满足  $\lim_{k \rightarrow \infty} x^k = x^*$ , 此时, 有  $\{x^k\}$  收敛到(1.4)的一个  $L$ -稳定点; 要么有  $B = B'$  ( $B$  包含  $\{x^k\}$  的所有聚点)。并且我们给出了在局部收敛性分析中起到了重要作用的结论。

#### 参考文献 (References)

- [1] Bardsley, J. and Nagy, J. (2006) Covariance-Preconditioned Iterative Methods for Nonnegativity Constrained Astronomical. *SIAM Journal*, **27**, 1184-1198.
- [2] Bruckstein, A.M., Donoho, D.L. and Elad, M. (2009) From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images. *SIAM Review*, **51**, 34-81. <http://dx.doi.org/10.1137/060657704>
- [3] Bruckstein, A.M., Elad, M. and Zibulevsky, M. (2008) On the Uniqueness of Nonnegative Sparse Solutions to Underdetermined Systems of Equations. *IEEE Transactions on Information Theory*, **54**, 4813-4820. <http://dx.doi.org/10.1109/TIT.2008.929920>
- [4] Donoho, D.L. and Tanner, J. (2005) Sparse Nonnegative Solutions of Underdetermined Linear Equations by Linear Programming. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 9446-9951. <http://dx.doi.org/10.1073/pnas.0502269102>
- [5] Beck, A. and Eldar, Y. (2013) Sparsity Constrained Nonlinear Optimization: Optimality Conditions and Algorithms. *SIAM Journal*, **23**, 1480-1509. <http://dx.doi.org/10.1137/120869778>
- [6] Bahmani, S., Raj, B. and Boufounos, P.T. (2013) Greedy Sparsity-Constrained Optimization. *Journal of Machine Learning Research*, **14**, 807-841.
- [7] Cartis, C. and Thompson, A. (2013) A New and Improved Quantitative Recovery Analysis for Iterative Hard Thresholding Algorithms in Compressed Sensing. arXiv:1309.5406.pdf.
- [8] Tropp, J. (2004) Greed Is Good: Algorithmic Results for Sparse Approximation. *IEEE Transactions on Information Theory*, **50**, 2231-2242. <http://dx.doi.org/10.1109/TIT.2004.834793>
- [9] Censor, Y. and Elfving, T. (1994) A Multiprojection Algorithm Using Bregman Projection in a Product Space. *Numerical Algorithms*, **8**, 221-239. <http://dx.doi.org/10.1007/BF02142692>
- [10] Bertsekas, D.P. (1999) Nonlinear Programming. 2nd Edition, Athena Scientific, Belmont.
- [11] Wang, C., Liu, Q. and Ma, C. (2013) Smoothing SQP Algorithm for Semismooth Equations with Box Constraints. *Computational Optimization and Applications*, **55**, 399-425. <http://dx.doi.org/10.1007/s10589-012-9524-5>