北京交通大學
Beijing Jiaotong University

Exact Recovery for Sparse Signal via Weighted ℓ_1 Minimization

SL Zhou YN Wang LC Kong **NH Xiu**

Department of Applied Mathematics, Beijing Jiaotong University

PKU Workshop On Optimization and Data Sciences, Dec 22, 2013, Peking University, Beijing

Outline

1 Introduction
1.1 Background

The concept of compressed sensing was first introduced by Donoho [D], Candès, Romberg and Tao [CRT] and Candès and Tao [CT]. Since then myriads of researchers have been lured to this area owing to its wide applications in signal processing, communications, astronomy, biology, medicine and so forth, see, e.g., [EK].

[D] D.L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289-1306, 2006.
[CRT] E.J. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans.Inf. Theory*, vol. 52, pp. 489-509, 2006.
[CT] E.J. Candès and T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203-4215, 2005.
[EK] Y.C. Eldar and G.Kutyniok, Compressed Sensing: Theory and Applications, Cambridge University Press, 2012.

The fundamental problem in compressed sensing is to recover a sparse solution $x \in \mathbb{R}^n$ of the underdetermined system of the form

$$\Phi x = y,$$

where $y \in \mathbb{R}^m$ is the available measurement and $\Phi \in \mathbb{R}^{m \times n}$ is a known measurement matrix.

To recover a sparse solution $x \in \mathbb{R}^n$ of the form $\Phi x = y$, the underlying model is the following ℓ_0 *minimization*:

$$\min \|x\|_0, \quad \text{s.t. } \Phi x = y, \tag{1}$$

where $\|x\|_0$ is ℓ_0-norm of the vector $x \in \mathbb{R}^n$. However (1) is NP-Hard.

One common approach is to solve (1) via convex ℓ_1 *minimization*:

$$\min \|x\|_1, \quad \text{s.t. } \Phi x = y. \tag{2}$$

The use of ℓ_1 minimization has become so extensively that it could arguably be considered *the modern least squares*, see, e.g., [BDE],[CWX] and [CZ].

[BDE] A.M. Bruckstein, D.L. Donoho, and A. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.*, vol. 51, pp. 34-81, 2009.
[CWX] T. Cai, L. Wang and G. Xu, New bounds for restricted isometry constants, *IEEE Trans. Inform. Theory*, vol. 56, pp. 4388-4394, 2010.
[CZ] T. Cai, and A. Zhang, Sparse Representation of a Polytope and Recovery of Sparse Signals and Low-rank Matrices, to appear in *IEEE Trans.Inf. Theory*, 2013.

Inspired by the efficiency of ℓ_1 minimization, it is natural to ask, for example, whether a different (but perhaps again convex) alternative to ℓ_0 minimization might also find the correct solution, but with a lower measurement requirement than ℓ_1 minimization.

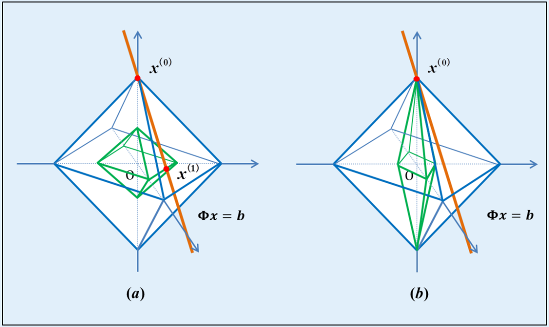
Numerical experiments indicate that the reweighted ℓ_1 minimization does outperform unweighted ℓ_1 minimization in many situations [CWB],[DDFG] and [ZL].

[CWB] E.J. Candès, M.B. Wakin and S. P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization, *J. Fourier Anal. Appl.*, vol. 14, pp. 877-905, 2008.
[DDFG] I. Daubechies, R. DeVore, M. Fornasier and C.S. Güntürk, Iteratively reweighted least squares minimization for sparse recovery, *Commun. Pure. Appl. Math.*, vol. 63, pp.1-38, 2010.
[ZL] Y.B. Zhao and D. Li, Reweighted ℓ_1-Minimization for Sparse Solutions to Underdetermined Linear Systems. *SIAM Journal on Optimization*, vol. 22, pp. 1065-1088, 2012.

In this talk, as a sequence, we consider the theoretical properties of the *weighted ℓ_1 minimization:*

$$\min \|\omega \circ x\|_1, \quad \text{s.t. } \Phi x = b, \tag{3}$$

where $\circ$ denotes the Hadamard product, that is $\|w \circ x\|_1 = \sum \omega_i |x_i|$, and $0 < \omega_i \leq 1, \; i = 1, 2, \cdots, n$.



Some cases that ℓ_1 minimization will fail to recover the sparse signal while exact recovery can be succeeded via weighted ℓ_1 minimization. (a) Sparse signal $x^{(0)} = (0, 0, 2)^T$, feasible set $\Phi x = b$, and in ℓ_1 ball there exists an $x^{(1)} = (\frac{3}{4}, \frac{3}{4}, 0)^T$ but $\|x^{(1)}\|_0 > \|x^{(0)}\|_0$. (b) In weighted ℓ_1 ball, there does not exist an $x \neq x^{(0)}$ such that $\|x\|_0 \leq \|x^{(0)}\|_0$.

The *null space property* (NSP) is the necessary and sufficient condition for (2) to reconstruct the system $b = \Phi x$ exactly, see, e.g., [Z].

Definition I.1 (NSP)

A matrix $\Phi \in \mathbb{R}^{m \times n}$ satisfies the null space property of order k if for all subsets $S \in \mathcal{C}_n^k$ it holds

$$\|h_S\|_1 < \|h_{S^c}\|_1 \tag{4}$$

for any $h \in \mathcal{N}(\Phi) \setminus \{0\}$, where $\mathcal{N}(\Phi) = \{h \in \mathbb{R}^n | \ \Phi h = 0\}$ and $\mathcal{C}_n^k = \{S \subset \{1, 2, \cdots, n\} \mid |S| = k\}$.

[Z] Y. Zhang, Theory of compressive sensing via ℓ-mimimization: A Non-RIP analysis and extensions, Technical Report, Rice Univ., 2008.

Another most popular sufficient condition for exact sparse recovery is *Restricted Isometry Property* (RIP) introduced by Candès and Tao [CT].

Definition I.2 (RIP)

For $k \in \{1, 2, \cdots, n\}$, the restricted isometry constant is the smallest positive number δ_k such that

$$(1 - \delta_k)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_k)\|x\|_2^2 \tag{5}$$

hold for all k-sparse vector $x \in \mathbb{R}^n$, i.e., $\|x\|_0 \leq k$.

[CT] E.J. Candès and T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203-4215, 2005.

	δ_k	δ_{2k}
Candès	- -	0.4142
Foucart and Lai	- -	0.4531
Foucart	- -	0.4652
Cai, Wang and Xu	- -	0.4721
Mo and Li	- -	0.4931
Cai and Zhang	1/3	0.5000
Zhou, Kong and Xiu	- -	0.5746 with $\delta_{8k} < 1$
Andersson and Strömberg	- -	0.6246

Table: Different bounds on δ_k and δ_{2k}.

Recently, Cai and Zhang [CZ] got a sharp bound

$$\delta_{tk} < \sqrt{\frac{t-1}{t}}. \tag{6}$$

★ Particularly, $\delta_{2k} < \frac{\sqrt{2}}{2}$. It is worth mentioning that (6) is the sharp bound for ℓ_1 minimization which has been proved in [CZ].

[CZ] T. Cai, and A. Zhang, Sparse Representation of a Polytope and Recovery of Sparse Signals and Low-rank Matrices, to appear in *IEEE Trans.Inf. Theory*, 2013.

1.7 Current Results for Weighted ℓ_1 Minimization

As for the weighted ℓ_1 minimization, literature [FMSY] presented us the upper bound on δ_k might be $\delta_k < 0.4343$ under some cases.

[FMSY] M.P. Friedlander, H. Mansour, R. Saab, and Ö. Yilmaz, Recovering Compressively Sampled Signals Using Partial Support Information, *IEEE Trans. Inf. Theory*, vol. 58, pp. 1122-1134, 2012.

Definition II.1

F A matrix $\Phi \in \mathbb{R}^{m\times n}$ satisfies the null space property of order k if for all subsets $S \in \mathcal{C}_n^k$ it holds

$$\|h_S\|_1 < \|h_{S^c}\|_1 \tag{7}$$

for any $h \in \mathcal{N}_1 := \{h \in \mathbb{R}^n |\ h \in \mathcal{N}(\Phi), \|h\|_1 = 1\}$.

Lemma II.2

Definition **I.1** is equivalent to Definition **II.1**.

Definition II.2 (WNSP)

For a given weight $\omega \in \mathbb{R}^n$, a matrix $\Phi \in \mathbb{R}^{m\times n}$ satisfies the weighted null space property of order k if for all subsets $S \in \mathcal{C}_n^k$ it holds

$$\|\omega \circ h_S\|_1 < \|\omega \circ h_{S^c}\|_1 \tag{8}$$

for any $h \in \mathcal{N}_1$.

Theorem II.2

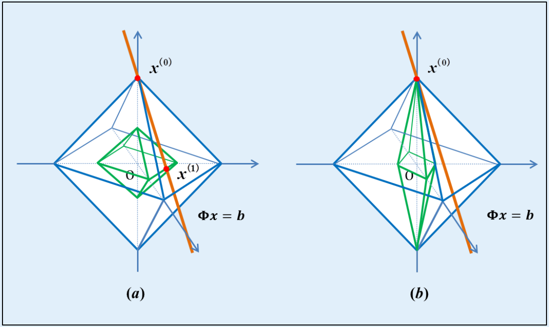
Every k-sparse vector $\hat{x} \in \mathbb{R}^n$ is the unique solution of the weighted minimization (3) with $b = \Phi\hat{x}$ iff Φ satisfies the WNSP of order k.

$$\Phi = \begin{pmatrix} 4/5 & 0 & 3/10 \\ 0 & 4/5 & 3/10 \end{pmatrix}, \quad b = \begin{pmatrix} 3/5 \\ 3/5 \end{pmatrix}.$$

Clearly, the unique solution of ℓ_0 and ℓ_1 models are $x^{(0)} = (0,0,2)^T$ and $x^{(1)} = (\frac{3}{4}, \frac{3}{4}, 0)^T$. If setting $\omega_2 = \omega_1, \omega_3 < \frac{3}{4}\omega_1$, $x^{(0)}$ is also the unique solution of the weighted ℓ_1 model.

For any $h \in \mathcal{N}_1$, we have $h = (\frac{3}{8}h_3, \frac{3}{8}h_3, -h_3)^T$ with $h_3 = 4/7$. Then for all subset $S \in \mathcal{C}_3^1$ and the given ω it holds $\|\omega \circ h_S\|_1 < \|\omega \circ h_{S^C}\|_1$, which means Φ satisfies WNSP. It is worth mentioning that this Φ does not satisfy the NSP due to $|h_3| \not< |\frac{3}{4}h_3| = |h_1| + |h_2|$.

2.3 Two Examples



Some cases that ℓ_1 minimization will fail to recover the sparse signal while exact recovery can be succeeded via weighted ℓ_1 minimization. (a) Sparse signal $x^{(0)} = (0, 0, 2)^T$, feasible set $\Phi x = b$, and in ℓ_1 ball there exists an $x^{(1)} = (\frac{3}{4}, \frac{3}{4}, 0)^T$ but $\|x^{(1)}\|_0 > \|x^{(0)}\|_0$. (b) In weighted ℓ_1 ball, there does not exist an $x \neq x^{(0)}$ such that $\|x\|_0 \leq \|x^{(0)}\|_0$.

$$\Phi = \begin{pmatrix} 3/4 & -1/2 & 3/8 & 1/2 & -1/4 \\ 3/4 & -1/2 & -1/8 & 1/2 & 0 \\ 0 & 1/4 & 3/8 & -1/8 & -3/8 \end{pmatrix}, b = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/8 \end{pmatrix}.$$

$$x^{(0)} = (0,0,0,1,0)^T, \qquad x^{(1)} = (\frac{1}{3}, -\frac{1}{2}, 0, 0, 0)^T,$$

$$\omega_2 = \frac{2}{3}\omega_1, \omega_4 = \frac{1}{2}\omega_1, \omega_3 = \omega_5 = \omega_1,$$

$$h = \left(\frac{-8h_2 + 13h_5}{12}, h_2, \frac{h_5}{2}, \frac{4h_2 - 3h_5}{2}, h_5 \right)^T.$$

Likely, Φ satisfies the WNSP we defined while does not content the NSP.

We first design a way of weighing and introduce some notations. Let T_0 and $\widehat{h}$ be the optimal solution of the following model

$$(T_0, \widehat{h}) := \underset{T \in \mathcal{C}_n^k, h \in \mathcal{N}_1}{\operatorname{argmax}} \|h_T\|_1. \tag{9}$$

For a constant $0 < \gamma \leq 1$, we define ω based on T_0 as

$$\omega_i = \begin{cases} \gamma, & i \in T_0, \\ 1, & i \in T_0^C, \end{cases} \tag{10}$$

where T_0^C is the complementary set of T_0 in $\{1, 2, \cdots, n\}$.

Lemma III.1

Let T_0 and $\widehat{h}$ be defined as (9). If T_0 uniquely exists, then there exists ω defined as (10) with $0 < \gamma < 1$ such that

$$\|\omega \circ \widehat{h}_{T_0}\|_1 = \max_{T \in \mathcal{C}_n^k, h \in \mathcal{N}_1} \|\omega \circ h_T\|_1. \tag{11}$$

If T_0 exists but not uniquely, then ω defined as (10) with $\gamma = 1$ that satisfies (11).

Theorem III.2

For the given γ and ω as (9) and (10), if

$$\delta_{ak} < \sqrt{\frac{a-1}{a-1+\gamma^2}} \tag{12}$$

holds for some $a > 1$, then each k sparse minimizer $\hat{x}$ of the weighted ℓ_1 minimization (3) is the solution of (1).

γ	δ_{2k}	δ_{3k}	δ_{4k}
1	$\sqrt{2}/2$	$\sqrt{6}/3$	$\sqrt{3}/2$
3/4	0.800	0.883	0.917
1/2	0.894	0.942	0.960
1/4	0.970	0.984	0.989

Table: Bounds on δ_{2k}, δ_{3k} and δ_{4k} with different cases.

Theorem III.3

For the given γ and ω as (9) and (10), if

$$\delta_k < \begin{cases} \dfrac{1}{1 + 2\lceil \gamma k \rceil / k}, & \text{for even number } k \geq 2, \quad (13) \\ \dfrac{1}{1 + 2\lceil \gamma k \rceil / \sqrt{k^2 - 1}}, & \text{for odd number } k \geq 3, \quad (14) \end{cases}$$

holds, where $\lceil a \rceil$ denotes the smallest integer that is no less than a, then each k sparse minimizer $\hat{x}$ of the weighted ℓ_1 minimization (3) is the solution of (1).

γ	$k \geq 2$ is even	$k \geq 3$ is odd
1	$1/3$	0.3203
$3/4$	$3/8$ ($k \geq 4$)	0.3797 ($k \geq 5$)
$1/2$	$1/2$ ($k \geq 2$)	$\sqrt{6} - 2$ ($k \geq 5$)
$1/4$	$2/3$ ($k \geq 4$)	$3 - \sqrt{6}$ ($k \geq 5$)
$1/6$	$3/4$ ($k \geq 6$)	0.7101 ($k \geq 5$)

Table: Bounds on δ_k with different cases. From the table one cannot difficultly find that under some mild situation, the upper bounds are greater than 0.4343.

$$\Phi = \begin{pmatrix} 4/5 & 0 & 3/10 \\ 0 & 4/5 & 3/10 \end{pmatrix}, \quad b = \begin{pmatrix} 3/5 \\ 3/5 \end{pmatrix}.$$

From $h = (\frac{3}{8}h_3, \frac{3}{8}h_3, -h_3)^T \in \mathcal{N}_1$ with $h_3 = \frac{4}{7}$, $|h_3|$ is the largest entry of h, i.e. $T_0 = \{3\}$ uniquely exists. Therefore by setting $\frac{3}{8} < \omega_3 = \gamma < 0.418, \omega_1 = \omega_2 = 1$, we have $\gamma\|h_{\{3\}}\|_1 < \|h_{\{1,2\}}\|_1$, which means that $x^{(0)}$ is the unique solution of weighted ℓ_1 model. We directly calculate that $\delta_2 = 0.9224$ with $n = 3, k = 2$ by the following formula

$$\delta_k = \max_{S \in \mathcal{C}_n^k} \|\Phi_S^T \Phi_S - I_k\|, \tag{15}$$

where $\|\cdot\|$ denotes the spectral norm of a matrix. Since T_0 uniquely exists and $\gamma < 0.418$, it yields $\delta_2 < 0.9226$ from (12) by taking $a = 2, k = 1$. Hence the ℓ_0 minimization can be exactly reconstructed by the weighted ℓ_1 minimization from our Theorem III.2

$$\Phi = \begin{pmatrix} 3/4 & -1/2 & 3/8 & 1/2 & -1/4 \\ 3/4 & -1/2 & -1/8 & 1/2 & 0 \\ 0 & 1/4 & 3/8 & -1/8 & -3/8 \end{pmatrix}, b = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/8 \end{pmatrix}.$$

From $h = \left(\frac{-8h_2+13h_5}{12}, h_2, \frac{h_5}{2}, \frac{4h_2-3h_5}{2}, h_5 \right)^T$, it follows that

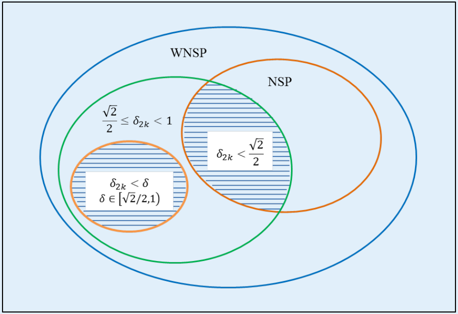
$$T_0 = \{4\}, \ \widehat{h} = (-2h_2/3, h_2, 0, 2h_2, 0)^T, \ h_2 = 6/11,$$

which manifests that T_0 uniquely exists. By setting $\omega_4 = \gamma = 0.3, \omega_1 = \omega_2 = \omega_3 = \omega_5 = 1$, we have $\gamma \|h_{\{4\}}\|_1 < \|h_{\{1,2,3,5\}}\|_1$, which means that $x^{(0)}$ is the unique solution of weighted ℓ_1 minimization. We compute $\delta_2 = 0.9572$ by (15) with $n = 5, k = 2$. Since T_0 uniquely exists and $\gamma = 0.3$, it yields $\delta_2 < 0.9578$ from (12) by taking $a = 2, k = 1$. And thus the ℓ_0 minimization can be exactly recovered via the weighted ℓ_1 minimization from Theorem III.2.

4 Discussion

Although T_0 defined by (9) always exists but not uniquely sometimes. However, from Examples above, we can see the assumption that T_0 uniquely exists is actually not a strong assumption to a certain extent.

4 Discussion



The relationship between WNSP, NSP and RIP, the dashed area denotes the scale of matrices that satisfy the RIP via weighted ℓ_1 minimization.

Thank you !