Research Article

Recovery of High-Dimensional Sparse Signals via ℓ_1-Minimization

Shiqing Wang and Limin Su

College of Mathematics and Information Sciences, North China University of Water Resources and Electric Power, Zhengzhou 450011, China

Correspondence should be addressed to Shiqing Wang; wangshiqing@ncwu.edu.cn

We consider the recovery of high-dimensional sparse signals via ℓ_1-minimization under mutual incoherence condition, which is shown to be sufficient for sparse signals recovery in the noiseless and noise cases. We study both ℓ_1-minimization under the ℓ_2 constraint and the Dantzig selector. Using the two ℓ_1-minimization methods and a technical inequality, some results are obtained. They improve the results of the error bounds in the literature and are extended to the general case of reconstructing an arbitrary signal.

1. Introduction

The problem of recovering a high-dimensional sparse signal based on a small number of measurements, possibly corrupted by noise, has attracted much recent attention. In the existing literature on sparse signals recovery and compressed sensing (see [1–13] and references therein), the emphasis is on assessing sparse signal $w \in R^n$ from an observation $y \in R^m$:

$$y = Aw + z, \tag{1}$$

where the matrix $A \in R^{m \times n}$ with $m \ll n$ is given and $z \in R^m$ is a vector of measurement errors. The goal is to reconstruct the unknown vector $w \in R^n$ based on y and A. Throughout the paper, we will assume that the columns of A are standardized to have unit ℓ_2-norm.

When noise is present, there are two well-known ℓ_1-minimization methods, which are well suited for recovering sparse signals. One is ℓ_1-minimization under the ℓ_2-constraint on the residuals:

$$(P_1) \quad \widehat{w} = \arg \min_w \left\{ \|w\|_1 : \|Aw - y\|_2 \le \varepsilon \right\}. \tag{2}$$

Another method, called Dantzig selector, is proposed by Candes and Tao [4]. The Dantzig selector solves the sparse recovery problem through ℓ_1-minimization with a constraint

on the correlation between the residuals and the column vectors of A:

$$(P_2) \quad \widehat{w} = \arg \min_w \left\{ \|w\|_1 : \left\| A^T (Aw - y) \right\|_\infty \le \varepsilon \right\}. \tag{3}$$

It is clear that regularity conditions are needed in order for these problems to be well behaved. Over the last few years, many interesting results for recovering sparse signals have been obtained in the framework of the mutual incoherence property (MIP) introduced by Donoho and Huo [14]. The MIP requires the pairwise correlations among the column vectors of A to be small. Let

$$\mu = \max_{i \neq j} \left| \left\langle A_i, A_j \right\rangle \right|. \tag{4}$$

See, for example, [8, 10, 14, 15].

It was first shown by Donoho and Huo [14], in the noiseless case for the setting where A is a concatenation of two square orthogonal matrices, that

$$k < \frac{1}{2} \left(\frac{1}{\mu} + 1 \right), \tag{5}$$

ensuring that the exact recovery of w when w has at most k nonzero entries (such a signal is called k-sparse). This result was then extended in the noiseless case in [11, 16] to a general dictionary A.

Stronger MIP conditions have been used in the literature to guarantee stable recovery of sparse signals in the noisy case. When noise is assumed to be bounded in ℓ_2-norm, Donoho et al. [15] showed that sparse signals can be recovered approximately through ℓ_1-minimization, with the error being at worst proportional to the input noise level, when

$$k < \frac{1}{4}\left(\frac{1}{\mu} + 1\right). \qquad (6)$$

The results in [17] imply that

$$k < \frac{2}{3 + \sqrt{6}}\left(\frac{1}{\mu} + 1\right) \approx 0.368\left(\frac{1}{\mu} + 1\right) \qquad (7)$$

is sufficient for stable recovery. And Tseng [18] used

$$k < \left(\frac{1}{2} - O(\mu)\right)\frac{1}{\mu} + 1. \qquad (8)$$

Cai et al. [19] have showed that the condition

$$k < \frac{1}{2}\left(\frac{1}{\mu} + 1\right) \qquad (9)$$

is not only sufficient but in fact sharp for stable recovery with bounded noise as well as Gaussian noise.

In this paper, we consider the problem of recovering a high-dimensional sparse signal via two well ℓ_1-minimization methods under the condition $k < (1/2)(1/\mu + 1)$. We study both ℓ_1-minimization under the ℓ_2 constraint (P_1) and the Dantzig selector (P_2). Using the two methods and a technical inequality, we give some results, which slightly improve those in [19]. Moreover, we obtain some results when unknown vector is not k-sparse in the noise case and noiseless case.

The rest of the paper is organized as follows. In Section 2, some basic notation and definitions are reviewed; an elementary inequality, which allow us to make finer analysis of the sparse recovery problem, is introduced. We begin the analysis of ℓ_1-minimization methods for sparse signals recovery by considering the exact recovery in the noise case in Section 3; our results are similar to those in [19] and to some extent, we provide tighter error bounds than the existing results in the literature. In Section 4, we consider the case of unknown vector which is not k-sparse under the condition $k < (1/2)(1/\mu + 1)$. We give some facts and the proofs of the theorems in Section 5.

2. Preliminaries

We begin by introducing basic notation and definitions and then develop an important inequality which will be used in proving our main results.

For a vector $v = (v_i) \in \mathbb{R}^n$, we will denote by $\mathrm{supp}(v) = \{i : v_i \neq 0\}$ the support of a vector v. We use the standard notation $\|v\|_q = \left(\sum_{i=1}^{p} |v_i|^q\right)^{1/q}$ to denote the ℓ_q-norm of the vector of v. Moreover, a vector v is said to be k-sparse if $|\mathrm{supp}(v)| \leq k$. We also treat a vector $v = (v_i) \in \mathbb{R}^n$ as a function $v : \{1, 2, \ldots, n\} \to \mathbb{R}$ by assigning $v(i) = v_i$.

We now introduce a useful elementary inequality, which is used in the proofs of the theorems.

Proposition 1. *Let $k < n$ be positive integers. For any descending chain of real numbers*

$$x_1 \geq x_2 \geq \cdots \geq x_k \geq x_{k+1} \geq \cdots \geq x_{n-1} \geq x_n \geq 0, \qquad (10)$$

One has

$$x_{k+1}^2 + x_{k+2}^2 + \cdots + x_n^2$$
$$\leq \frac{(x_{k+1} + x_{k+2} + \cdots + x_n)(x_1 + x_2 + \cdots + x_k)}{k} \qquad (11)$$

Proof. Since $x_i \geq x_j$ for $i < j$, we have

$$x_{k+1}^2 + x_{k+2}^2 + \cdots + x_n^2$$
$$\leq x_k x_{k+1} + x_{k+1}x_{k+2} + \cdots + x_{n-1}x_n$$
$$= \frac{k(x_k x_{k+1} + x_{k+1}x_{k+2} + \cdots + x_{n-1}x_n)}{k} \qquad (12)$$
$$\leq \frac{(x_{k+1} + x_{k+2} + \cdots + x_n)(x_1 + x_2 + \cdots + x_k)}{k}.$$
□

3. Recovery of k-Sparse Signals

As previously above, the condition (5) has been proved to guarantee the recovery of k-sparse signal in noiseless case. Cai et al. [19] have shown that this condition is also sufficient for stable reconstruction of k-sparse signals in the noisy case when the error is in a bounded set. We will also give the results for reconstruction of k-sparse signals both in the noiseless and noisy cases with error bounded, which are proved using different methods from [19].

Theorem 2. *Consider the model (1) with $\|z\|_2 \leq \varepsilon$. Suppose that w is k-sparse and $\widehat{w}$ is the solution of ℓ_1-minimization problem (P_1). Then, under the condition $k < (1/2)(1/\mu + 1)$,*

$$\|\widehat{w} - w\|_2 \leq \frac{2\sqrt{2(1 + (k-1)\mu)}}{1 - (2k-1)\mu}\varepsilon. \qquad (13)$$

We now consider sparse recovery of k-sparse signals with error in a different bounded set. Candes and Tao [4] treated the sparse signal recovery in the Gaussian noise case by solving minimization

$$\min \|w\|_1 \quad \text{subject to} \quad \left\|A^T(y - Aw)\right\|_\infty \leq \eta, \qquad (14)$$

with bounded set $\{z : \|A^T z\|_\infty \leq \eta\}$ and referred the solution as the Dantzig Selector. The following result shows that the condition $k < (1/2)(1/\mu + 1)$ is also sufficient when the error is in the bounded set $\{z : \|A^T z\|_\infty \leq \varepsilon\}$.

Theorem 3. *Consider the model (1) with $\|A^T z\|_\infty \leq \varepsilon$. Suppose that w is k-sparse and $\widehat{w}$ is the solution of ℓ_1-minimization problem (P_2). Then, under the condition $k < (1/2)(1/\mu + 1)$,*

$$\|\widehat{w} - w\|_2 \leq \frac{2\sqrt{2k}}{1 - (2k-1)\mu}\varepsilon. \qquad (15)$$

Remark 4. We consider the stable recovery of sparse signals with error in the ℓ_2-ball; for example, $B \subset \mathbb{R}^n$ is a bounded set. B is taken to be $\{0\}$ in the noiseless case and can be $B^{\ell_2} = \{z : \|z\|_2 \leq \eta\}$ or $B^{DS} = \{z : \|A^T z\|_\infty \leq \eta\}$ in the noisy case.

Note that $2k\mu < \mu+1$, under the condition $k < (1/2)(1/\mu+1)$, the result of Theorem 2 is equivalent to

$$\|\widehat{w} - w\|_2 \leq \frac{\sqrt{3-\mu}}{1-(2k-1)\mu}(\varepsilon + \eta), \qquad (16)$$

and Theorem 3 is equivalent to

$$\|\widehat{w} - w\|_2 \leq \frac{\sqrt{2k}}{1-(2k-1)\mu}(\varepsilon + \eta). \qquad (17)$$

To some extent, Theorem 2 improves Theorem 2.1 in [19], while Theorem 3 is improved using different method from [19] and gets the same results as Theorem 2.2 in [19].

4. Recovery of Approximately k-Sparse Signals

In the previous section, the focus was on recovering k-sparse signals. As discussed in [17, 19, 20], our results can also be stated in the general setting of reconstructing an arbitrary signal under the condition $k < (1/2)(1/\mu + 1)$.

We begin in this section by considering the problem of exact recovery of spares signals when no noise is present. This is an interesting problem in itself and has been considered in a number of papers; see, for example, [9, 11, 17, 21]. More importantly, the solutions to this "clean" problem shed light on the noisy case.

When w is not k-sparse, ℓ_1-minimization can also recover w with accuracy if w has good k-term approximation. For a general vector $w \in R^n$, denote by $w_{\max(k)}$ with all but the k-largest entries (in absolute value) set to zero and $w_{-\max(k)}$ the vector w with the k-largest entries (in absolute value) set to zero.

Theorem 5. *Let $A \in R^{m \times n}$. Suppose that $k \geq 1$ satisfies $k < (1/2)(1/\mu + 1)$ and $\widehat{w}$ is the solution of the following ℓ_1-minimization problem*

$$(P) \quad \min \|w\|_1 \quad subject\ to \quad y = Aw. \qquad (18)$$

Then

$$\|\widehat{w} - w\|_2 \leq \frac{2\mu\sqrt{2k}\|w_{-\max(k)}\|_1}{1-(2k-1)\mu}. \qquad (19)$$

We now turn to the noisy case. Suppose that $A \in R^{m \times n}$ and $y = Aw + z$, where the noise z is bounded. We will specifically consider two cases: $\|z\|_2 \leq \varepsilon$ and $\|A^T z\|_\infty \leq \varepsilon$. We will first consider the case $\|z\|_2 \leq \varepsilon$.

Theorem 6. *Consider the model (1) with z satisfying $\|z\|_2 \leq \varepsilon$. Suppose that $k \geq 1$ satisfies $k < (1/2)(1/\mu + 1)$ and $\widehat{w}$ is the solution of the ℓ_1-minimization problem (P_1). Then*

$$\|\widehat{w} - w\|_2 \leq \frac{2\varepsilon\sqrt{2(1+(k-1)\mu)}+2\mu\sqrt{2k}\|w_{-\max(k)}\|_1}{1-(2k-1)\mu}. \qquad (20)$$

We next turn to the case $\|A^T z\|_\infty \leq \varepsilon$, which is called Dantzig selector.

Theorem 7. *Consider the model (1) with $\|A^T z\|_\infty \leq \varepsilon$. Suppose that $k \geq 1$ satisfies $k < (1/2)(1/\mu+1)$ and $\widehat{w}$ is the solution of ℓ_1-minimization problem (P_2). Then*

$$\|\widehat{w} - w\|_2 \leq \frac{2\varepsilon\sqrt{2k} + 2\mu\sqrt{2k}\|w_{-\max(k)}\|_1}{1-(2k-1)\mu}. \qquad (21)$$

We have so far focused on stable recovery with bounded error. The results can be extended directly to the Gaussian noise case. This is due to the fact that Gaussian noise is "essentially bounded." See, for example, [17, 19, 20].

5. The Proofs of the Theorems

Before giving the proofs of the theorems, we introduce three widely used facts, which are useful for the proofs.

(A) The following fact is well known in the recovery of sparse signals. Let $x \in \mathbb{R}^n$ be any k-sparse signal and $A \in R^{m \times n}$; then

$$(1-(k-1)\mu)\|x\|_2^2$$
$$\leq \|Ax\|_2^2 \leq (1+(k-1)\mu)\|x\|_2^2, \qquad (22)$$

where μ is defined by (4); see, for example, [17, 18, 22, 23].

Let $\widehat{w}$ be a solution to the minimization problem; then by definition $\|\widehat{w}\|_1 \leq \|w\|_1$. Let $\delta = \widehat{w} - w$, and $\delta_{S_0} = \delta I \operatorname{supp}(w)$. Here, I_J denotes the indicator function of a set J; that is, $I_J(j) = 1$ if $j \in J$ and 0 if $j \notin J$.

(B) The following is a widely used fact (see, e.g., [4, 7, 14, 17]):

$$\left\|\delta - \delta_{S_0}\right\|_1 \leq \left\|\delta_{S_0}\right\|_1. \qquad (23)$$

This follows directly from the fact that

$$\|w\|_1 \geq \|\widehat{w}\|_1 = \|w + \delta\|_1$$
$$= \left\|w + \delta_{S_0}\right\|_1 + \left\|\delta - \delta_{S_0}\right\|_1 \qquad (24)$$
$$\geq \|w\|_1 - \left\|\delta_{S_0}\right\|_1 + \left\|\delta - \delta_{S_0}\right\|_1.$$

(C) The following fact, which is based on the minimality of $\widehat{w}$, has been widely used; see for example, [14, 19, 20]:

$$\left\|\delta_{T_0}\right\|_1 \geq \left\|\delta_{T_0^c}\right\|_1, \qquad (25)$$

where T_0 is the support of w. This follows directly from the fact that

$$\|w\|_1 \geq \|\widehat{w}\|_1 = \|w + \delta\|_1$$
$$= \left\|w + \delta_{T_0}\right\|_1 + \left\|\delta_{T_0}^c\right\|_1 \qquad (26)$$
$$\geq \|w\|_1 - \left\|\delta_{T_0}\right\|_1 + \left\|\delta_{T_0^c}\right\|_1.$$

Proof of Theorem 2. The proof makes use of the ideas from [17, 19, 22].

Let $\delta = \widehat{w} - w$. Rearranging the indices if necessary, we assume that

$$\left|\delta_{(1)}\right| \geq \left|\delta_{(2)}\right| \geq \cdots \geq \left|\delta_{(k+1)}\right| \geq \left|\delta_{(k+2)}\right| \geq \cdots. \tag{27}$$

Let $T_1 = \{1, 2, \ldots, k\}$ and T_2 be the support of w; then, from the fact (C),

$$\left\|\delta_{T_2}\right\|_1 \geq \left\|\delta_{T_2^c}\right\|_1. \tag{28}$$

Note that $T_2^c \cap T_1$ and $T_2 \cap T_1^c$ both have $k - |T_1 \cap T_2|$ elements, so we have

$$\left\|\delta_{T_2^c \cap T_1}\right\|_1 \geq \left\|\delta_{T_2 \cap T_1^c}\right\|_1. \tag{29}$$

We will show that this implies that

$$\left\|\delta_{T_1}\right\|_1 \geq \left\|\delta_{T_1^c}\right\|_1. \tag{30}$$

In fact

$$
\begin{aligned}
\left\|\delta_{T_1}\right\|_1 &= \left\|\delta_{T_1 \cap T_2}\right\|_1 + \left\|\delta_{T_1 \cap T_2^c}\right\|_1 \\
&= \left\|\delta_{T_2}\right\|_1 - \left\|\delta_{T_2 \cap T_1^c}\right\|_1 + \left\|\delta_{T_1 \cap T_2^c}\right\|_1 \\
&\geq \left\|\delta_{T_2^c}\right\|_1 - \left\|\delta_{T_2 \cap T_1^c}\right\|_1 + \left\|\delta_{T_1 \cap T_2^c}\right\|_1 \\
&= \left\|\delta_{T_1^c \cap T_2^c}\right\|_1 - \left\|\delta_{T_2 \cap T_1^c}\right\|_1 + 2\left\|\delta_{T_1 \cap T_2^c}\right\|_1 \\
&= \left\|\delta_{T_1^c}\right\|_1 - 2\left\|\delta_{T_2 \cap T_1^c}\right\|_1 + 2\left\|\delta_{T_1 \cap T_2^c}\right\|_1 \geq \left\|\delta_{T_1^c}\right\|_1.
\end{aligned}
\tag{31}
$$

For simplicity, partition $\{1, 2, \ldots, n\}$ into the following sets:

$$
\begin{aligned}
S_0 &= \{1, 2, \ldots, k\}, \qquad S_1 = \{k+1, k+2, \ldots, k+k_1\}, \\
S_2 &= \{k+k_1+1, k+k_1+2, \ldots, k+k_1+k_2\}, \ldots,
\end{aligned}
\tag{32}
$$

where $k_j (j > 0)$ is an positive integer.

Note that $k < (1/2)(1/\mu + 1)$ is equivalent to $(2k-1)\mu < 1$ or $1 - (2k-1)\mu > 0$. Now

$$
\begin{aligned}
\left|\left\langle A\delta, A\delta_{S_0}\right\rangle\right| &= \left|\left\langle A\delta_{S_0}, A\delta_{S_0}\right\rangle + \left\langle A\left(\delta - \delta_{S_0}\right), A\delta_{S_0}\right\rangle\right| \\
&\geq (1 - (k-1)\mu)\left\|\delta_{S_0}\right\|_2^2 - \mu\left\|\delta - \delta_{S_0}\right\|_1 \left\|\delta_{S_0}\right\|_1 \\
&\geq (1 - (k-1)\mu)\left\|\delta_{S_0}\right\|_2^2 - \mu\left\|\delta_{S_0}\right\|_1^2 \\
&\geq (1 - (k-1)\mu)\left\|\delta_{S_0}\right\|_2^2 - \mu k\left\|\delta_{S_0}\right\|_2^2 \\
&\geq (1 - (2k-1)\mu)\left\|\delta_{S_0}\right\|_2^2,
\end{aligned}
\tag{33}
$$

where the second inequality applies the facts (A) and (B).

On the other hand, it follows from the fact (A) that

$$\left\|A\delta_{S_0}\right\|_2^2 \leq (1 + (k-1)\mu)\left\|\delta_{S_0}\right\|_2^2. \tag{34}$$

Note that

$$\|A\delta\|_2 = \|A(\widehat{w} - w)\|_2 \leq \|A\widehat{w} - y\|_2 + \|Aw - y\|_2 \leq 2\varepsilon. \tag{35}$$

From Proposition 1 and the fact that $|\delta_{(1)}| \geq |\delta_{(2)}| \geq \cdots \geq |\delta_{(k+1)}| \geq |\delta_{(k+2)}| \geq \cdots$, we get

$$\left\|\delta_{S_0^c}\right\|_2^2 \leq \left\|\delta_{S_0^c}\right\|_1 \frac{\left\|\delta_{S_0}\right\|_1}{k} \leq \frac{\left\|\delta_{S_0}\right\|_1^2}{k} \leq \left\|\delta_{S_0}\right\|_2^2, \tag{36}$$

which implies

$$\|\delta\|_2^2 = \left\|\delta_{S_0}\right\|_2^2 + \left\|\delta_{S_0^c}\right\|_2^2 \leq 2\left\|\delta_{S_0}\right\|_2^2. \tag{37}$$

Putting them together, we get

$$
\begin{aligned}
\|\delta\|_2 &\leq \sqrt{2}\left\|\delta_{S_0}\right\|_2 \leq \frac{\sqrt{2}\left|\left\langle A\delta, A\delta_{S_0}\right\rangle\right|}{(1 - (2k-1)\mu)\left\|\delta_{S_0}\right\|_2} \\
&\leq \frac{\sqrt{2}\|A\delta\|_2 \left\|A\delta_{S_0}\right\|_2}{(1 - (2k-1)\mu)\left\|\delta_{S_0}\right\|_2} \\
&\leq \frac{2\sqrt{2(1 + (k-1)\mu)}\left\|\delta_{S_0}\right\|_2}{(1 - (2k-1)\mu)\left\|\delta_{S_0}\right\|_2}\varepsilon \\
&= \frac{2\sqrt{2(1 + (k-1)\mu)}}{1 - (2k-1)\mu}\varepsilon.
\end{aligned}
\tag{38}
$$

$\square$

Proof of Theorem 3. Note that from the fact (A) and the first part of the proof of Theorem 2, we have

$$\left|\left\langle A\delta, A\delta_{S_0}\right\rangle\right| \geq (1 - (2k-1)\mu)\left\|\delta_{S_0}\right\|_2^2. \tag{39}$$

On the other hand, we also obtain the following relation:

$$
\begin{aligned}
\left|\left\langle A\delta, A\delta_{S_0}\right\rangle\right| &= \left(A\delta_{S_0}\right)^T A\delta \\
&\leq \left\|A^T A\delta\right\|_\infty \left\|\delta_{S_0}\right\|_1 \leq 2\varepsilon\sqrt{k}\left\|\delta_{S_0}\right\|_2,
\end{aligned}
\tag{40}
$$

We get, together with them, that

$$\left\|\delta_{S_0}\right\|_2 \leq \frac{2\varepsilon\sqrt{k}}{1 - (2k-1)\mu}. \tag{41}$$

Then

$$\|\delta\|_2 \leq \sqrt{2}\left\|\delta_{S_0}\right\|_2 \leq \frac{2\sqrt{2k}}{1 - (2k-1)\mu}\varepsilon, \tag{42}$$

where the last second inequality uses (37). $\square$

Proof of Theorem 5. Let $\delta = \widehat{w} - w$ and $T = \{1, 2, \ldots, k\} \subset \{1, 2, \ldots, n\}$ be the support of $w_{\max(k)}$. Following the notation

and the first part in the proof of Theorem 2, we first give the following relation:

$$\left\| \delta - \delta_{S_0} \right\|_1 \leq \left\| \delta_{S_0} \right\|_1 + 2 \left\| w_{-\max(k)} \right\|_1. \tag{43}$$

In fact, since $\|w\|_1 \geq \|\widehat{w}\|_1$, we have

$$\begin{aligned}
\|w\|_1 &\geq \|\widehat{w}\|_1 = \|w + \delta\|_1 \\
&= \left\| w_{\max(k)} + w_{-\max(k)} + \delta - \delta_{S_0} + \delta_{S_0} \right\|_1 \\
&= \left\| w_{\max(k)} + \delta_{S_0} \right\|_1 + \left\| w_{-\max(k)} + \delta - \delta_{S_0} \right\|_1 \\
&\geq \left\| w_{\max(k)} \right\|_1 - \left\| \delta_{S_0} \right\|_1 \\
&\quad + \left\| \delta - \delta_{S_0} \right\|_1 - \left\| w_{-\max(k)} \right\|_1.
\end{aligned} \tag{44}$$

Since $w = w_{\max(k)} + w_{-\max(k)}$, this yields

$$\left\| \delta - \delta_{S_0} \right\|_1 \leq \left\| \delta_{S_0} \right\|_1 + 2 \left\| w_{-\max(k)} \right\|_1. \tag{45}$$

Note that

$$A\delta = A\widehat{w} - Aw = 0. \tag{46}$$

So

$$\begin{aligned}
0 &= \left| \left\langle A\delta, A\delta_{S_0} \right\rangle \right| = \left| \left\langle A\delta_{S_0}, A\delta_{S_0} \right\rangle + \left\langle A\left(\delta - \delta_{S_0} \right), A\delta_{S_0} \right\rangle \right| \\
&\geq \left(1 - (k-1)\mu \right) \left\| \delta_{S_0} \right\|_2^2 - \mu \left\| \delta - \delta_{S_0} \right\|_1 \left\| \delta_{S_0} \right\|_1 \\
&\geq \left(1 - (k-1)\mu \right) \left\| \delta_{S_0} \right\|_2^2 - \mu \left(\left\| \delta_{S_0} \right\|_1 + 2 \left\| w_{-\max(k)} \right\|_1 \right) \left\| \delta_{S_0} \right\|_1 \\
&\geq \left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2^2 - 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2,
\end{aligned} \tag{47}$$

where the last second inequality holds from (43).

Then

$$\left\| \delta_0 \right\|_2 \leq \frac{2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}. \tag{48}$$

From (37), we have that

$$\|\delta\|_2 \leq \sqrt{2} \left\| \delta_0 \right\|_2 \leq \frac{2\mu\sqrt{2k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}. \tag{49}$$

The proof is completed. □

Proof of Theorem 6. From the proof of Theorem 5, we have

$$\begin{aligned}
\left| \left\langle A\delta, A\delta_{S_0} \right\rangle \right| &\geq \left(1 - (2k-1)\mu \right) \\
&\quad \times \left\| \delta_{S_0} \right\|_2^2 - 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2.
\end{aligned} \tag{50}$$

It follows from the fact (A) that

$$\left\| A\delta_{S_0} \right\|_2^2 \leq \left(1 + (k-1)\mu \right) \left\| \delta_{S_0} \right\|_2^2. \tag{51}$$

Note that

$$\|A\delta\|_2 = \|A(\widehat{w} - w)\|_2 \leq \|A\widehat{w} - y\|_2 + \|Aw - y\|_2 \leq 2\varepsilon. \tag{52}$$

Together with them, we get

$$\begin{aligned}
\left\| \delta_{S_0} \right\|_2 &\leq \frac{\left| \left\langle A\delta, A\delta_{S_0} \right\rangle \right| + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2}{\left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2} \\
&\leq \frac{\|A\delta\|_2 \left\| A\delta_{S_0} \right\|_2 + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2}{\left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2} \\
&\leq \frac{2\varepsilon\sqrt{1 + (k-1)\mu} \left\| \delta_{S_0} \right\|_2 + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2}{\left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2} \\
&\leq \frac{2\varepsilon\sqrt{1 + (k-1)\mu} + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}.
\end{aligned} \tag{53}$$

Then, from (37),

$$\begin{aligned}
\|\delta\|_2 &\leq \sqrt{2} \|\delta_0\|_2 \\
&\leq \frac{2\varepsilon\sqrt{2(1 + (k-1)\mu)} + 2\mu\sqrt{2k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}.
\end{aligned} \tag{54}$$

□

Proof of Theorem 7. From the proof of Theorems 3 and 5, we get

$$\begin{aligned}
\left\| \delta_{S_0} \right\|_2 &\leq \frac{\left| \left\langle A\delta, A\delta_{S_0} \right\rangle \right| + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2}{\left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2} \\
&\leq \frac{2\varepsilon\sqrt{k} \left\| \delta_{S_0} \right\|_2 + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1 \left\| \delta_{S_0} \right\|_2}{\left(1 - (2k-1)\mu \right) \left\| \delta_{S_0} \right\|_2} \\
&= \frac{2\varepsilon\sqrt{k} + 2\mu\sqrt{k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}.
\end{aligned} \tag{55}$$

Then

$$\|\delta\|_2 \leq \sqrt{2} \|\delta_0\|_2 \leq \frac{2\varepsilon\sqrt{2k} + 2\mu\sqrt{2k} \left\| w_{-\max(k)} \right\|_1}{1 - (2k-1)\mu}. \tag{56}$$

□

References

[1] S. Q. Wang and L. M. Su, "The oracle inequalities on simultaneous Lasso and Dantzig selector in high-dimensional nonparametric regression," *Mathematical Problems in Engineering*, vol. 2013, Article ID 571361, 6 pages, 2013.

[2] P. J. Bickel, Y. Ritov, and A. B. Tsybakov, "Simultaneous analysis of lasso and Dantzig selector," *The Annals of Statistics*, vol. 37, no. 4, pp. 1705–1732, 2009.

[3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[4] E. Candes and T. Tao, "The Dantzig selector: statistical estimation when p is much larger than n," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

[5] E. Candes, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians*, pp. 1433–1452, Madrid, Spain, August 2006.

[6] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[7] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[8] J. A. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1030–1051, 2006.

[9] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[10] J. J. Fuchs, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3601–3608, 2005.

[11] J.-J. Fuchs, "On sparse representations in arbitrary redundant bases," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1341–1344, 2004.

[12] A. Nemirovski, "Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 229–251, 2004.

[13] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society B*, vol. 58, no. 1, pp. 267–288, 1996.

[14] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[15] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[16] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.

[17] T. T. Cai, G. Xu, and J. Zhang, "On recovery of sparse signals via ℓ_1 minimization," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3388–3397, 2009.

[18] P. Tseng, "Further results on stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 888–899, 2009.

[19] T. T. Cai, L. Wang, and G. Xu, "Stable recovery of sparse signals and an oracle inequality," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3516–3522, 2010.

[20] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1300–1308, 2010.

[21] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006.

[22] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4388–4394, 2010.

[23] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

Advances in
Operations Research

Advances in
Decision Sciences

Mathematical Problems
in Engineering

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Journal of
Complex Analysis

Journal of
Mathematics

International Journal of
Stochastic Analysis

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

Journal of
Applied Mathematics

Journal of
Optimization