Ann. of Appl. Math.
32:4(2016), 380-395

EQUIVALENCE BETWEEN NONNEGATIVE SOLUTIONS TO PARTIAL SPARSE AND WEIGHTED l_1-NORM MINIMIZATIONS[*][†]

Xiuqin Tian, Zhengshan Dong, Wenxing Zhu[‡]

(*Center for Discrete Math. and Theoretical Computer Science, Fuzhou University, Fuzhou 350116, Fujian, PR China*)

Abstract

Based on the range space property (RSP), the equivalent conditions between nonnegative solutions to the partial sparse and the corresponding weighted l_1-norm minimization problem are studied in this paper. Different from other conditions based on the spark property, the mutual coherence, the null space property (NSP) and the restricted isometry property (RIP), the RSP-based conditions are easier to be verified. Moreover, the proposed conditions guarantee not only the strong equivalence, but also the equivalence between the two problems. First, according to the foundation of the strict complementarity theorem of linear programming, a sufficient and necessary condition, satisfying the RSP of the sensing matrix and the full column rank property of the corresponding sub-matrix, is presented for the unique nonnegative solution to the weighted l_1-norm minimization problem. Then, based on this condition, the equivalence conditions between the two problems are proposed. Finally, this paper shows that the matrix with the RSP of order k can guarantee the strong equivalence of the two problems.

Keywords compressed sensing; sparse optimization; range space property; equivalent condition; l_0-norm minimization; weighted l_1-norm minimization

2000 Mathematics Subject Classification 94A12; 15A29; 90C25

1 Introduction

In this paper, we consider the following partial sparse minimization problem

$$\min_{x,y} \sum_{i=1}^{n_1} w_i |x_i|_0 + a^{\mathrm{T}} y \quad s.t.\ Ax + By = b,\ x \geq 0,\ y \geq 0, \tag{1}$$

where $x = (x_1, x_2, \cdots, x_{n_1})^{\mathrm{T}} \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$. $|x_i|_0 = 0$ if $x_i = 0$; otherwise $|x_i|_0 = 1$.

[*]Research supported by the National Natural Science Foundation of China under Grant 61672005.

[†]Manuscript received May 23, 2016; Revised September 18, 2016

[‡]Corresponding author. E-mail: wxzhu@fzu.edu.cn

$w_i \in R$ is the weight on $|x_i|_0$, $i = 1, 2, \cdots, n_1$. $a \in \mathbb{R}^{n_2}$, $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{m\times n_2}$, and $b \in \mathbb{R}^m$ $(m < n_1 + n_2)$ are the problem datas. Let $\|x\|_0$ be the number of nonzero components of x, that is, $\|x\|_0 = \sum_{i=1}^{n_1} |x_i|_0$. Although $\|x\|_0$ is not a norm, we still call it l_0-norm for simplicity.

By relaxing $|x_i|_0$ as $|x_i|$, and taking into account $x \geq 0$, $\sum_{i=1}^{n_1} w_i|x_i| = w^{\mathrm{T}}x$, we get the following linear program

$$\min_{x,y} \ w^{\mathrm{T}}x + a^{\mathrm{T}}y \quad s.t.\ Ax + By = b,\ (x, y) \geq 0, \tag{2}$$

where $w = (w_1, w_2, \cdots, w_{n_1})^{\mathrm{T}}$. We are interested in what conditions can ensure the equivalence of problems (1) and (2).

In recent years, l_0-norm minimization problems have been widely researched, and have been successfully applied to signal processing [1], pattern recognition [2], machine learning [3], computational biology [4], medical imaging [5], and other fields [6-9]. Recent research indicates that l_1-norm relaxation can promote sparsity [11]. This is based on equivalence between the l_0-norm and l_1-norm minimization problems.

Up to now, the study of the equivalence between l_0-norm and l_1-norm minimization problems is mainly for the following two problems:

$$\min_{x} \ \|x\|_0 \quad s.t.\ Ax = b, \tag{3}$$

and

$$\min_{x} \ \|x\|_1 \quad s.t.\ Ax = b. \tag{4}$$

It has been proved that, all k-sparse solutions to problem (3) can be found by solving problem (4), if the spark of the sensing matrix A is greater than $2k$ [19], or the order of the null space property (NSP) [21] or the restricted isometry property (RIP) of A is $2k$ or above [25]. However, these conditions remain restrictive and are hard to be verified. From geometric perspective, Donoho and Tanner [30] showed that the outward k-neighborliness of A could also guarantee the equivalence of problems (3) and (4). Donoho and Romberg [11] also analysed the equivalence from probabilistic perspective.

Based on the RSP, Zhao [26,27] presented equivalent conditions for problems (3) and (4), no matter whether the nonnegative constraints exist or not. The conditions guarantee not only the strong equivalence but also the equivalence between the l_0-norm and l_1-norm minimization problems. Moreover, Zhao presented an RSP which could be verified easily, by solving a linear programming problem [26].

In this paper, we consider the equivalence between problems (1) and (2). The definition of equivalence is similar to that in [26], which is as follows.

Definition 1.1[26] (i) Problems (1) and (2) are said to be equivalent if there exists a solution to problem (1) that coincides with the unique solution to problem (2);

(ii) problems (1) and (2) are said to be strongly equivalent if the unique solution to problem (1) coincides with the unique solution to problem (2).

We consider only nonnegative solution throughout this paper. The case without nonnegative constrains can be reduced to that one, by replacing (x, y) with $(x^+ - x^-, y^+ - y^-)$, where $x^+, x^- \in \mathbb{R}^{n_1}_+$, and $y^+, y^- \in \mathbb{R}^{n_2}_+$. Moreover, when there exist inequality constraints, by introducing slack variables, they can be written as the form of (1). Some more details can refer to Section 5.

The rest of this paper is organized as follows. Section 2 shows the existence of the unique nonnegative solution to the weighted l_1-norm minimization problem (2). The equivalent and strongly equivalent conditions are described in Sections 3 and 4 respectively. In Section 5, we present several variants of problem (1). Conclusions are given in Section 6.

2 Uniqueness of Weighted l_1-norm Minimization

From Definition 1.1, the existence of a unique solution to the weighted l_1-norm minimization problem (2) is necessary for the equivalence between the partial sparse and the weighted l_1-norm minimizations. Hence, we first consider some properties of the unique solution to problem (2) in this section.

2.1 Range space property

In this subsection, we first reformulate problem (2) as a linear program. Then based on the duality theory, we present a necessary condition for a unique solution to problem (2).

Suppose that problem (2) admits a unique solution (x^*, y^*). Then for its any feasible solution $(x, y) \neq (x^*, y^*)$, there is $w^{\mathrm{T}}x + a^{\mathrm{T}}y > w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*$. In other words,

$$\{(x, y) : Ax + By = b, w^{\mathrm{T}}x + a^{\mathrm{T}}y \leq w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*, (x, y) \geq 0\} = \{(x^*, y^*)\}.$$

Now consider the following linear optimization problem:

$$\begin{aligned} \min_{x,y} \quad & 0^{\mathrm{T}}x + 0^{\mathrm{T}}y, \\ s.t. \quad & Ax + By = Ax^* + By^*, \\ & w^{\mathrm{T}}x + a^{\mathrm{T}}y \leq w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*, \\ & (x, y) \geq 0. \end{aligned} \tag{5}$$

It is easy to verify that problem (5) has a feasible solution (x^*, y^*), and its optimal value is finite (always equals zero). By introducing slack variable $t \geq 0$, the above problem (5) is equivalent to:

$$\begin{aligned}
\min_{x,y,t} \quad & 0^{\mathrm{T}}x + 0^{\mathrm{T}}y, \\
s.t. \quad & Ax + By = Ax^* + By^*, \\
& w^{\mathrm{T}}x + a^{\mathrm{T}}y + t = w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*, \\
& (x, y, t) \geq 0.
\end{aligned} \tag{6}$$

Before proceeding, we can obviously obtain the following lemma, which will be used below.

Lemma 2.1 *The following three statements are equivalent:*

(i) (x^*, y^*) *is the unique solution to the weighted l_1-norm minimization problem* (2).

(ii) (x^*, y^*) *is the unique solution to problem* (5).

(iii) $(x, y, t) = (x^*, y^*, 0)$ *is the unique solution to problem* (6).

The dual of problem (6) is given by:

$$\begin{aligned}
\max_{z,c} \quad & (Ax^* + By^*)^{\mathrm{T}}z + (w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*)c, \\
s.t. \quad & A^{\mathrm{T}}z + cw \leq 0, \\
& B^{\mathrm{T}}z + ca \leq 0, \\
& c \leq 0,
\end{aligned} \tag{7}$$

where $z \in \mathbb{R}^m$, $c \in \mathbb{R}$ are the dual variables. Assume that $\alpha, \beta \in \mathbb{R}^m$, $\gamma \in \mathbb{R}_+$ are slack variables of problem (7), then we have

$$\alpha = -A^{\mathrm{T}}z - cw, \quad \beta = -B^{\mathrm{T}}z - ca, \quad \gamma = -c.$$

Next, if problem (2) admits a unique solution, then $(A, B)^{\mathrm{T}}$ satisfies the range space property at this point, which is given by the following lemma.

Lemma 2.2 *If* (x^*, y^*) *is the unique solution to problem* (2), *then there exists a* $z \in \mathbb{R}^m$ *satisfying*

$$\begin{cases} (A^{\mathrm{T}}z)_i = w_i, & i \in J_+, \\ (A^{\mathrm{T}}z)_i < w_i, & i \in J_0, \end{cases} \qquad \begin{cases} (B^{\mathrm{T}}z)_i = a_i, & i \in S_+, \\ (B^{\mathrm{T}}z)_i < a_i, & i \in S_0, \end{cases} \tag{8}$$

where $J_+ = \{i : x_i^* > 0\}$, $J_0 = \{i : x_i^* = 0\}$, $S_+ = \{i : y_i^* > 0\}$, $S_0 = \{i : y_i^* = 0\}$.

Proof First, it is easy to check that problem (6) and its duality (7) have feasible solutions. Then by the complementary slackness theory [31], there exists a pair of strictly complementary optimal solution $((x, y, t), (z^1, c^1))$ to problems (6) and (7). Let $(\alpha, \beta, \gamma) = (-A^{\mathrm{T}}z^1 - c^1w, -B^{\mathrm{T}}z^1 - c^1a, -c^1)$. Then we have

$$x^{\mathrm{T}}\alpha = 0, \quad y^{\mathrm{T}}\beta = 0, \quad t\gamma = 0, \tag{9}$$

and
$$x+\alpha>0, \quad y+\beta>0, \quad t+\gamma>0. \tag{10}$$
Since (x^*,y^*) is the unique solution, by Lemma 2.1, $(x^*,y^*,0)$ is the unique solution to problem (6). Then we obtain
$$(x,y,t)=(x^*,y^*,0),$$
which together with $x^*\geq 0,\ y^*\geq 0$ implies that
$$\begin{cases} x_i>0, & i\in J_+,\\ x_i=0, & i\in J_0,\end{cases} \quad \begin{cases} y_i>0, & i\in S_+,\\ y_i=0, & i\in S_0,\end{cases} \quad t=0. \tag{11}$$
From (9)-(11), we have
$$\begin{cases} \alpha_i=0, & i\in J_+,\\ \alpha_i>0, & i\in J_0,\end{cases} \quad \begin{cases} \beta_i=0, & i\in S_+,\\ \beta_i>0, & i\in S_0,\end{cases} \quad \gamma>0.$$
By the definitions of these slack variables, we obtain
$$\begin{cases} (A^{\mathrm{T}}z^1+c^1w)_i=0, & i\in J_+,\\ (A^{\mathrm{T}}z^1+c^1w)_i<0, & i\in J_0,\end{cases} \quad \begin{cases} (B^{\mathrm{T}}z^1+c^1a)_i=0, & i\in S_+,\\ (B^{\mathrm{T}}z^1+c^1a)_i<0, & i\in S_0,\end{cases} \quad c^1<0,$$
which is
$$\begin{cases} (A^{\mathrm{T}}(\frac{z^1}{-c^1})-w)_i=0, & i\in J_+,\\ (A^{\mathrm{T}}(\frac{z^1}{-c^1})-w)_i<0, & i\in J_0,\end{cases} \quad \begin{cases} (B^{\mathrm{T}}(\frac{z^1}{-c^1})-a)_i=0, & i\in S_+,\\ (B^{\mathrm{T}}(\frac{z^1}{-c^1})-a)_i<0, & i\in S_0.\end{cases}$$
If denote $z=\frac{z^1}{-c^1}$, the above formula can be rewritten as
$$\begin{cases} (A^{\mathrm{T}}z)_i=w_i, & i\in J_+,\\ (A^{\mathrm{T}}z)_i<w_i, & i\in J_0,\end{cases} \quad \begin{cases} (B^{\mathrm{T}}z)_i=a_i, & i\in S_+,\\ (B^{\mathrm{T}}z)_i<a_i, & i\in S_0.\end{cases}$$
This completes the proof of Lemma 2.2.

Hence, the inequality system (8) is necessary for (x^*,y^*) to be the unique solution to the weighted l_1-norm minimization problem (2). And throughout this paper, the inequality system (8) is called the range space property (RSP) of $(A,B)^{\mathrm{T}}$ at (x^*,y^*).

2.2 Full column rank condition

In this subsection, we present another necessary condition for the existence of a unique solution to the weighted l_1-norm minimization problem (2), which is given by the following lemma.

Lemma 2.3 *Let (x^*, y^*) be the unique solution to problem* (2)*, then the matrix*

$$M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w_{J_+}^{\mathrm{T}} & a_{S_+}^{\mathrm{T}} \end{pmatrix} \tag{12}$$

is of full column rank, where $J_+ = \{i : x_i^* > 0\}$, $S_+ = \{i : y_i^* > 0\}$.

Proof Suppose that the columns of M are linear dependent. Then there exists a vector $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \neq 0$ satisfying

$$M\mu = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w_{J+}^{\mathrm{T}} & a_{S_+}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = 0. \tag{13}$$

Let us define a vector (x, y, t) as

$$x = (x_{J_+}, x_{J_0}) = (x_{J_+}^*, 0), \quad y = (y_{S_+}, y_{S_0}) = (y_{S_+}^*, 0), \quad t = 0,$$

where $J_0 = \{i : x_i^* = 0\}$, $S_0 = \{i : y_i^* = 0\}$. By Lemma 2.1, (x, y, t) is an optimal solution to problem (6). Next, we define another vector $(\tilde{x}, \tilde{y}, \tilde{t})$ as:

$$\tilde{x} = (\tilde{x}_{J_+}, \tilde{x}_{J_0}) = (x_{J_+}^* + \lambda\mu_1, 0), \quad \tilde{y} = (\tilde{y}_{S_+}, \tilde{y}_{S_0}) = (y_{S_+}^* + \lambda\mu_2, 0), \quad \tilde{t} = 0.$$

Since $x_{J_+}^* > 0$ and $y_{S_+}^* > 0$, there exists a $\lambda \neq 0$ such that

$$\tilde{x}_{J_+} = x_{J_+}^* + \lambda\mu_1 \geq 0, \quad \tilde{y}_{S_+} = y_{S_+}^* + \lambda\mu_2 \geq 0.$$

Therefore, $(\tilde{x}, \tilde{y}, \tilde{t}) \geq 0$. In addition, taking (13) into consideration, it is easy to verify that $(\tilde{x}, \tilde{y}, \tilde{t})$ is a feasible solution to problem (6). And it is also an optimal solution. Moreover, combing $\lambda\mu \neq 0$ with the definitions of (x, y, t) and $(\tilde{x}, \tilde{y}, \tilde{t})$, we have

$$(x, y, t) \neq (\tilde{x}, \tilde{y}, \tilde{t}).$$

Hence, problem (6) has at least two solutions, which contradicts the assumption of Lemma 2.3. We have thus proved the lemma.

2.3 Sufficient condition

Combing the above two necessary conditions, we obtain the following sufficient condition for the existence of a unique solution to problem (2).

Lemma 2.4 *Let (x^*, y^*) be a feasible solution to problem* (2)*. If the RSP of $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and the matrix $M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w_{J_+}^{\mathrm{T}} & a_{S_+}^{\mathrm{T}} \end{pmatrix}$ is of full column rank, then (x^*, y^*) is the unique solution to problem* (2)*, where $J_+ = \{i : x_i^* > 0\}$, $S_+ = \{i : y_i^* > 0\}$.*

Proof According to Lemma 2.1, it suffices to prove that $(x^*, y^*, 0)$ is the unique solution to problem (6). Since $(A, B)^{\mathrm{T}}$ satisfies the RSP at (x^*, y^*), there exists a $z \in \mathbb{R}^m$ satisfying

$$\begin{cases} (A^{\mathrm{T}}z)_i = w_i, & i \in J_+, \\ (A^{\mathrm{T}}y)_i < w_i, & i \in J_0, \end{cases} \quad \begin{cases} (B^{\mathrm{T}}z)_i = a_i, & i \in S_+, \\ (B^{\mathrm{T}}z)_i < a_i, & i \in S_0, \end{cases}$$

where $J_0 = \{i : x_i^* = 0\}$, $S_0 = \{i : y_i^* = 0\}$. Let $c = -1$. The above formula can be rewritten as

$$\begin{cases} (A^{\mathrm{T}}z)_i + cw_i = 0, & i \in J_+, \\ (A^{\mathrm{T}}z)_i + cw_i < 0, & i \in J_0, \end{cases} \quad \begin{cases} (B^{\mathrm{T}}z)_i + ca_i = 0, & i \in S_+, \\ (B^{\mathrm{T}}z)_i + ca_i < 0, & i \in S_0. \end{cases} \tag{14}$$

Combing (14) with $c = -1$, it is easy to verify that, if (z, c) satisfies (14) then it is a feasible solution to problem (7). Now we prove that (z, c) is also an optimal solution to problem (7). Substituting such (z, c) into the objective function of (7), we have

$$\begin{aligned} &(Ax^* + By^*)^{\mathrm{T}}z + (w^{\mathrm{T}}x^* + a^{\mathrm{T}}y^*)c \\ =\ & x^{*T}(A^{\mathrm{T}}z) + y^{*T}(B^{\mathrm{T}}z) + c(w^{\mathrm{T}}x^*) + ca^{\mathrm{T}}y^* \\ =\ & \sum_{i \in J_+} x_i^*(A^{\mathrm{T}}z)_i + \sum_{i \in S_+} y_i^*(B^{\mathrm{T}}z)_i + c\sum_{i \in J_+} w_i x_i^* + c\sum_{i \in S_+} a_i y_i^* \\ =\ & -c\sum_{i \in J_+} w_i x_i^* - c\sum_{i \in S_+} a_i y_i^* + c\sum_{i \in J_+} w_i x_i^* + c\sum_{i \in S_+} a_i y_i^* \\ =\ & 0. \end{aligned}$$

On the other hand, the optimal value of its primal problem (6) is 0, which together with the strong duality theorem implies that problem (7) has an optimal value 0 and (z, c) is its optimal solution.

Next, we prove that such $(x^*, y^*, 0)$ is the unique solution to problem (6). Let (x', y', t') be its any optimal solution. Since (z, c) satisfying (14) is an optimal solution to problem (7), $((x', y', t'), (z, c))$ is a pair of optimal solutions to problems (6) and (7). Assume that α, β, γ are slack variables of (7) and defined by

$$\begin{cases} \alpha = -(A^{\mathrm{T}}z + cw), \\ \beta = -(B^{\mathrm{T}}z + ca), \\ \gamma = -c. \end{cases}$$

According to (14), we have

$$\begin{cases} \alpha_i = -(A^{\mathrm{T}}z + cw)_i > 0, & i \in J_0, \\ \beta_i = -(B^{\mathrm{T}}z + ca)_i > 0, & i \in S_0, \\ \gamma = -c = 1 > 0. \end{cases} \tag{15}$$

On the other hand, by the complementary slackness conditions, (x', y', t') and (α, β, γ) satisfy

$$\alpha^{\mathrm{T}} x' = 0, \quad \beta^{\mathrm{T}} y' = 0, \quad \gamma t' = 0, \tag{16}$$

which together with (15) implies that

$$\begin{cases} x'_i = 0, & i \in J_0, \\ y'_i = 0, & i \in S_0, \\ t' = 0. \end{cases} \tag{17}$$

Substituting (17) into the constraints of (6), we get

$$\begin{aligned} &Ax' + By' = A_{J_+} x'_{J_+} + B_{S_+} y'_{S_+} = A_{J_+} x^*_{J_+} + B_{S_+} y^*_{S_+}, \\ &w^{\mathrm{T}} x' + a^{\mathrm{T}} y' + t' = w^{\mathrm{T}}_{J_+} x'_{J_+} + a^{\mathrm{T}}_{S_+} y'_{S_+} = w^{\mathrm{T}}_{J_+} x^*_{J_+} + a^{\mathrm{T}}_{S_+} y^*_{S_+}, \end{aligned}$$

which is

$$\begin{pmatrix} A_{J_+} & B_{S_+} \\ w^{\mathrm{T}}_{J+} & a^{\mathrm{T}}_{S_+} \end{pmatrix} \begin{pmatrix} x'_{J_+} \\ y'_{S_+} \end{pmatrix} = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w^{\mathrm{T}}_{J+} & a^{\mathrm{T}}_{S_+} \end{pmatrix} \begin{pmatrix} x^*_{J_+} \\ y^*_{S_+} \end{pmatrix}.$$

Since $M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w^{\mathrm{T}}_{J+} & a^{\mathrm{T}}_{S_+} \end{pmatrix}$ is of full column rank, $(x'_{J_+}, y'_{S_+}) = (x^*_{J_+}, y^*_{S_+})$. Taking (17) into consideration, we get $(x', y', t') = (x^*, y^*, 0)$. This completes the proof.

2.4 Sufficient and necessary condition

In this subsection, we present a sufficient and necessary condition for the unique solution to problem (2). This condition is the foundation of the equivalence between the partial sparse and the weighted l_1-norm minimization problems with nonnegative constrains in this paper.

Theorem 2.1 *Let (x^*, y^*) be a feasible solution to problem (2). Then (x^*, y^*) is the unique solution if and only if the RSP of $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and the matrix $M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w^{\mathrm{T}}_{J+} & a^{\mathrm{T}}_{S_+} \end{pmatrix}$ is of full column rank, where $J_+ = \{i : x^*_i > 0\}$, $S_+ = \{i : y^*_i > 0\}$.*

The proof of Theorem 2.1 is evident from Lemmas 2.2, 2.3 and 2.4. Furthermore, if (A_{J_+}, B_{S_+}) is of full column rank, then so is the matrix

$$M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w^{\mathrm{T}}_{J_+} & a^{\mathrm{T}}_{S_+} \end{pmatrix}.$$

Unfortunately, its converse does not hold. For example, for

$$A_{J_+} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_{S_+} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad w^{\mathrm{T}}_{J_+} = (1, 1), \quad a^{\mathrm{T}}_{S_+} = 0,$$

it is obvious that

$$M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w_{J_+}^{\mathrm{T}} & a_{S_+}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{pmatrix}$$

is of full column rank, but the columns of

$$(A_{J_+}, B_{S_+}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

are linear dependent.

However, if the RSP of $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), then (A_{J_+}, B_{S_+}) is of full column rank if and only if $M = \begin{pmatrix} A_{J_+} & B_{S_+} \\ w_{J+}^{\mathrm{T}} & a_{S_+}^{\mathrm{T}} \end{pmatrix}$ is of full column rank, which directly produces the following theorem from Theorem 2.1.

Theorem 2.2 *Let (x^*, y^*) be a feasible solution to problem* (2). *Then (x^*, y^*) is the unique solution if and only if the RSP of $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and the submatrix (A_{J_+}, B_{S_+}) is of full column rank, where $J_+ = \{i : x_i > 0\}$, $S_+ = \{i : y_i > 0\}$.*

Theorems 2.1 and 2.2 adequately characterize the condition for the existence of a unique solution to problem (2). Next we present an example to show that it is easy to verify the sufficient and necessary condition.

Example 2.1 Consider the following linear system $Ax + By = b$, where

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ 1 & -6 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

If let $w = (1, 1, 1)^{\mathrm{T}}$, $a = (0, 0)^{\mathrm{T}}$, it is easy to verify that $(x^*, y^*) = \big((0, 0, \frac{1}{2}), (1, 0)\big)^{\mathrm{T}}$ is a feasible solution to problem (2). Denote

$$J_+ = \{i : x_i^* > 0\} = \{3\}, \quad S_+ = \{i : y_i^* > 0\} = \{1\}.$$

Then the submatrix

$$(A_{J_+}, B_{S_+}) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}$$

corresponding to (x^*, y^*) is of full column rank. Moreover, there exists a $z = (\frac{2}{7}, \frac{5}{7}, 1)^{\mathrm{T}}$ such that $A^{\mathrm{T}} z = (\frac{2}{7}, -\frac{5}{7}, 1)^{\mathrm{T}}$, $B^{\mathrm{T}} z = (0, -\frac{25}{7})^{\mathrm{T}}$. Then the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*). According to Theorem 2.2, we know that (x^*, y^*) is the unique solution to the weighted l_1-norm minimization problem.

If $a = (\frac{1}{2}, 1)$, then there exists a $z = (\frac{1}{2}, 1, 1)^{\mathrm{T}}$ such that $A^{\mathrm{T}} z = (\frac{1}{2}, -1, 1)^{\mathrm{T}}$, $B^{\mathrm{T}} z = (\frac{1}{2}, -\frac{11}{2})^{\mathrm{T}}$. In this case we also find out that the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and (x^*, y^*) is the unique solution to the weighted l_1-norm minimization problem.

3 Equivalent Condition

In this section, we study the equivalence between problems (1) and (2). From the sufficient and necessary conditions given by Theorems 2.1 and 2.2, if (x^*, y^*) is the unique solution to problem (2), then (A_{J_+}, B_{S_+}) must be of full column rank. Hence,

$$\|(x^*, y^*)\|_0 = \operatorname{rank}(A_{J_+}, B_{S_+}) = |J_+| + |S_+| \leq m,$$

which shows that, if (x^*, y^*) is the unique solution, it must be m-sparse, and that any nonnegative vector (x, y), whose sparsity is greater than m, must not be an optimal solution to problem (2).

As we know, Gaussian elimination can easily obtain an m-sparse solution. However, it cannot guarantee that the obtained solution is the sparsest. In fact, problem (1) is an NP-hard combinatorial optimization problem. And many relaxation methods have been proposed, such as relaxing the problem to the weighted l_1-norm minimization problem. Hence, we must find out that, under what condition problems (1) and (2) have the same sparse solutions. The following theorem provides somewhat an answer according to the sufficient and necessary conditions in Theorems 2.1 and 2.2.

Theorem 3.1 *Let (x^*, y^*) be an optimal solution to the l_0-norm minimization problem* (1)*. Then it also is a solution to the weighted l_1-norm minimization problem* (2) *if and only if the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and (A_{J_+}, B_{S_+}) is of full column rank, where $J_+ = \{x_i^* > 0\}$, $S_+ = \{y_i^* > 0\}$.*

Proof If problems (1) and (2) are equivalent, then there exists an optimal solution (x^*, y^*) to problem (1), which is also the unique optimal solution to problem (2). Let $J_+ = \{x_i^* > 0\}$, $S_+ = \{y_i^* > 0\}$. Due to Theorem 2.2, the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and (A_{J_+}, B_{S_+}) is of full column rank.

On the other hand, if the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at (x^*, y^*), and (A_{J_+}, B_{S_+}) is of full column rank, then by Theorem 2.2, (x^*, y^*) is also a solution to problem (2). The proof is completed.

Next, we present an example to show that, when the l_0-norm minimization problem has several sparsest solutions, the weighted l_1-minimization problem could still has one of the sparsest solutions.

Example 3.1 Consider the linear system $Ax + By = b$, where

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Let $w = (\frac{3}{4}, 1)^{\mathrm{T}}$, $a = (0, 0)^{\mathrm{T}}$. It is easy to verify that $(x^1, y^1) = \big((0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\big)^{\mathrm{T}}$, $(x^2, y^2) = \big((1, 0), (0, 0)\big)^{\mathrm{T}}$ are the sparsest solutions to problem (1). Since $w^{\mathrm{T}} x^2 +$

$a^{\mathrm{T}}y^2 > w^{\mathrm{T}}x^1 + a^{\mathrm{T}}y^1$, (x^2, y^2) is not the optimal solution to problem (2), and the RSP of $(A, B)^{\mathrm{T}}$ does not hold at (x^2, y^2). Next, we consider (x^1, y^1). Denote $J_+ = \{2\}$, $S_+ = \{1, 2\}$. Then $(A_{J_+}, B_{S_+}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ is of full column rank. There exists a $z = (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^{\mathrm{T}}$ such that $A^{\mathrm{T}}z = (\frac{1}{2}, 1)^{\mathrm{T}}$, $B^{\mathrm{T}}z = (0, 0)^{\mathrm{T}}$. Hence, the RSP of $(A, B)^{\mathrm{T}}$ holds at (x^1, y^1). Due to Theorem 2.2, (x^1, y^1) is the unique solution to the weighted l_1-norm minimization problem.

If let $w = (\frac{3}{4}, -1)^{\mathrm{T}}$, $a = (\frac{1}{4}, 0)^{\mathrm{T}}$, then there exists a $z = (\frac{5}{8}, -\frac{3}{8}, -\frac{5}{8})^{\mathrm{T}}$ satisfying the RSP at (x^1, y^1), since $A^{\mathrm{T}}z = (-\frac{3}{8}, -1)^{\mathrm{T}}$, $B^{\mathrm{T}}z = (\frac{1}{4}, 0)^{\mathrm{T}}$. Hence, in this case, (x^1, y^1) is also the unique solution to the weighted l_1-norm minimization problem.

It is easy to verify that if $w = (-\frac{3}{4}, 1)^{\mathrm{T}}$, then $a = (\frac{1}{4}, 0)^{\mathrm{T}}$, (x^2, y^2) is the unique solution to the weighted l_1-norm minimization problem.

Theorem 2.1 and Example 3.1 show that the existence of a unique solution to l_0-norm minimization is not necessary for the equivalence between the l_0-norm and weighted l_1-norm minimization problems.

4 Strongly Equivalent Condition

In practice, the matrices A and B should be suitable such that we can find all k-sparse solutions. In this section, we will discuss the RSP of order k, which guarantees finding all k-sparse vectors. First, we present the definition of the RSP of order k as follows.

Definition 4.1 Let $A \in \mathbb{R}^{m\times n_1}$, $B \in \mathbb{R}^{m\times n_2}$ with $m < n_1 + n_2$. The matrix $(A, B)^{\mathrm{T}}$ is said to satisfy the RSP of order k, if for any subset S_1 of $\{1, 2, \cdots, n_1\}$ and subset S_2 of $\{1, 2, \cdots, n_2\}$ with $|S_1| \le k$ and $|S_1| + |S_2| \le m$, (A_{S_1}, A_{S_2}) is of full column rank, and there exists a $z \in \mathbb{R}^m$ satisfying

$$\begin{cases} (A^{\mathrm{T}}z)_i = w_i, & i \in S_1, \\ (A^{\mathrm{T}}z)_i < w_i, & i \in S_1^c, \end{cases} \quad \begin{cases} (B^{\mathrm{T}}z)_i = a_i, & i \in S_2, \\ (B^{\mathrm{T}}z)_i < a_i, & i \in S_2^c, \end{cases} \tag{18}$$

where $w \in \mathbb{R}^{n_1}$, $a \in \mathbb{R}^{n_2}$, $S_1^c = \{1, 2, \cdots, n_1\} \setminus S_1$, $S_2^c = \{1, 2, \cdots, n_2\} \setminus S_2$.

Now we give an example to show the existence of the matrix satisfying the RSP of order k. Let $A = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $w = 1$, $a = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. It is easy to verify that:

If $S_1 = \emptyset$, $S_2 = \{1\}$, then there exists a $z = (0, -\frac{1}{4})^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = \frac{1}{2}, B^{\mathrm{T}}z = (0, -\frac{1}{4})$;

if $S_1 = \emptyset$, $S_2 = \{2\}$, then there exists a $z = (-\frac{1}{2}, 0)^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = \frac{1}{2}$, $B^{\mathrm{T}}z = (-\frac{1}{2}, 0)$;

if $S_1 = \emptyset$, $S_2 = \{1, 2\}$, then there exists a $z = (0, 0)^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = 0$, $B^{\mathrm{T}}z = (0, 0)$;

if $S_1 = \{1\}$, $S_2 = \emptyset$, then there exists a $z = (-\frac{1}{2}, -\frac{1}{4})^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = 1$, $B^{\mathrm{T}}z = (-\frac{1}{2}, -\frac{1}{4})$;

if $S_1 = \{1\}$, $S_2 = \{1\}$, then there exists a $z = (0, -\frac{1}{2})^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = 1$, $B^{\mathrm{T}}z = (0, -\frac{1}{2})$;

if $S_1 = \{1\}$, $S_2 = \{2\}$, then there exists a $z = (-1, 0)^{\mathrm{T}}$ satisfying $A^{\mathrm{T}}z = 1$, $B^{\mathrm{T}}z = (-1, 0)$.

Hence, for any S_1, S_2 with $|S_1| \leq 1$, $|S_1| + |S_2| \leq 2$, there exists a $z \in R^2$ such that $(A, B)^{\mathrm{T}}$ satisfies the RSP, that is, the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k.

Next, we claim that if the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k, then we can find all k-sparse solutions, which is presented as follows.

Theorem 4.1 *Weighted l_1-norm minimization could find any nonnegative solution (x, y) with $\|x\|_0 \leq k$, $\|(x, y)\|_0 \leq m$ if and only if the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k.*

Proof Assume that any vector (x, y) with $\|x\|_0 \leq k$, $\|(x, y)\|_0 \leq m$ can be found by weighted l_1-norm minimization. Then (x, y) is the unique solution to

$$\min_{(\tilde{x}, \tilde{y})} \; w^{\mathrm{T}}\tilde{x} + a^{\mathrm{T}}\tilde{y} \quad A\tilde{x} + B\tilde{y} = b = Ax + By, \; (\tilde{x}, \tilde{y}) \geq 0.$$

Denote $S_1 = \{i : x_i > 0\}$, $S_2 = \{i : y_i > 0\}$. By Theorem 2.2, the matrix (A_{S_1}, B_{S_2}) is of full column rank, and there exists a $z \in R^m$ satisfying

$$\begin{cases} (A^{\mathrm{T}}z)_i = w_i, & i \in S_1, \\ (A^{\mathrm{T}}z)_i < w_i, & i \in S_1^c, \end{cases} \qquad \begin{cases} (B^{\mathrm{T}}z)_i = a_i, & i \in S_2, \\ (B^{\mathrm{T}}z)_i < a_i, & i \in S_2^c. \end{cases}$$

Since (x, y) is arbitrary and satisfies $\|x\|_0 \leq k$, $\|(x, y)\|_0 \leq m$, the corresponding sets S_1, S_2 are also arbitrary subsets of $\{1, 2, \cdots, n_1\}$, $\{1, 2, \cdots, n_2\}$ respectively and satisfy $|S_1| \leq k$, $|S_1| + |S_2| \leq m$, and there exists a $z \in R^m$ satisfying (18). Hence, the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k.

On the contrary, assume the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k. Then for any vector (x, y) with $\|x\|_0 \leq k$, $\|(x, y)\|_0 \leq m$, the corresponding submatrix (A_{S_1}, B_{S_2}) is of full column rank satisfying the RSP at (x, y), where $S_1 = \{i : x_i > 0\}$, $S_2 = \{i : y_i > 0\}$. According to Theorem 2.2, we claim that (x, y) is the unique solution to the weighted l_1-norm minimization problem. The proof is completed.

Theorem 4.2 *Assume that $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k. Then any nonnegative vector (x, y) with $\|x\|_0 \leq k$ and $\|(x, y)\|_0 \leq m$ is the unique solution to both the l_0-norm and weighted l_1-norm minimization problems.*

Proof By Theorem 4.1, any nonnegative vector (x, y) with $\|x\|_0 \leq k$, $\|(x, y)\|_0 \leq m$ can be found by weighted l_1-norm minimization. Hence, such (x, y) is the unique solution to the weighted l_1-norm minimization problem. Now we prove that it is also the unique solution to the l_0-norm minimization problem. Suppose (x^1, y^1) is an optimal solution to the l_0-norm minimization problem and satisfies $\|x^1\|_0 \le \|x\|_0$. Denote $S_1 = \{i : x_i^1 > 0\}$, $S_2 = \{i : y_i^1 > 0\}$. Then $|S_1| = \|x^1\|_0 \le \|x\|_0 \le k$, $|S_1| + |S_2| \le m$, which together with the RSP of order k of $(A, B)^{\mathrm{T}}$ imply that (A_{S_1}, B_{S_2}) is of full column rank and satisfies the RSP at (x^1, y^1). By Theorem 2.4, (x^1, y^1) is the unique solution to the weighted l_1-norm minimization problem. Hence, $(x, y) = (x^1, y^1)$. Then by arbitrariness of (x^1, y^1), (x, y) is the unique solution to the l_0-norm minimization problem. The proof is completed.

Theorem 4.2 shows that if the matrix $(A, B)^{\mathrm{T}}$ satisfies the RSP of order k, then the l_0-norm and weighted l_1-norm minimization problems are strongly equivalent. However, Theorem 3.1 states that, when the RSP of the matrix $(A, B)^{\mathrm{T}}$ holds at one sparsest solution to the l_0-norm minimization problem, the l_0-norm and weighted l_1-norm minimization problems are equivalent. Hence, the RSP of order k is stronger then the RSP at a solution. This may be also explained that the RSP is a local property, which only requires to be satisfied at one special solution, while the RSP of order k is a global property, which requires to be satisfied for all (x, y) with $\|x\|_0 \le k$, $\|(x, y)\|_0 \le m$.

5 Several Variants

In this section, we discuss several variants of problem (1). Fortunately, all these variants can be rewritten as the form of (1).

No nonnegative constraints In this case, the variables $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$ are not nonnegative. The problem is

$$\min_{x,y} \ \|x\|_0 + a^{\mathrm{T}} y \quad s.t.\ Ax + By = b. \tag{19}$$

To obtain the form of (1), for a given (x, y), we define

$$\begin{cases} x_i^+ = x_i, & \text{if } x_i > 0, \\ x_i^+ = 0, & \text{if } x_i \le 0, \end{cases} \qquad \begin{cases} x_i^- = 0, & \text{if } x_i > 0, \\ x_i^- = -x_i, & \text{if } x_i \le 0, \end{cases} \tag{20}$$

and

$$\begin{cases} y_i^+ = y_i, & \text{if } y_i > 0, \\ y_i^+ = 0, & \text{if } y_i \le 0, \end{cases} \qquad \begin{cases} y_i^- = 0, & \text{if } y_i > 0, \\ y_i^- = -y_i, & \text{if } y_i \le 0. \end{cases} \tag{21}$$

Then $x = x^+ - x^-$, $y = y^+ - y^-$. And the equality constraint can be rewritten as

$$(A, -A) \begin{pmatrix} x^+ \\ x^- \end{pmatrix} + (B, -B) \begin{pmatrix} y^+ \\ y^- \end{pmatrix} = b,$$

and $\|x\|_0 = \|x^+ - x^-\|_0 = \|x^+\|_0 + \|x^-\|_0$. Therefore, problem (19) can be reformulated as

$$\min_{\tilde{x},\tilde{y}} \ \|\tilde{x}\|_0 + \tilde{a}^{\mathrm{T}}\tilde{y} \quad s.t.\ \tilde{A}\tilde{x} + \tilde{B}\tilde{y} = b,\ (\tilde{x},\tilde{y}) \geq 0, \tag{22}$$

where $\tilde{a} = \begin{pmatrix} a \\ -a \end{pmatrix}$, $\tilde{A} = (A, -A)$, $\tilde{B} = (B, -B)$, $\tilde{x} = \begin{pmatrix} x^+ \\ x^- \end{pmatrix}$, $\tilde{y} = \begin{pmatrix} y^+ \\ y^- \end{pmatrix}$. It is easy to verify that problems (19) and (22) have the same solutions.

Inequality constraints When the constraints are inequalities, the problem is

$$\min_{x,y} \ \|x\|_0 + a^{\mathrm{T}}y \quad s.t.\ Ax + By \geq b,\ (x,y) \geq 0.$$

By introducing slack variables $z \in \mathbb{R}^m_+$, the above problem can be rewritten as

$$\min_{x,y,z} \ \|x\|_0 + \begin{pmatrix} a \\ 0 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} y \\ z \end{pmatrix} \quad s.t.\ Ax + (B, -I)\begin{pmatrix} y \\ z \end{pmatrix} = b,\ (x,y,z) \geq 0,$$

where I is the identity matrix of m-dimensions.

Noisy case In many practical problems, the measurement data b always is noisy. Then the optimization problem is

$$\min_{x,y} \ \|x\|_0 + a^{\mathrm{T}}y \quad s.t.\ \|Ax + By - b\|_\infty \leq \sigma,\ (x,y) \geq 0,$$

where $\sigma \geq 0$ controls the error between $Ax + By$ and b. In this case, the constraint can be written as

$$Ax + By \geq b - \sigma, \quad -Ax - By \geq -b - \sigma,$$

which is

$$\begin{pmatrix} A \\ -A \end{pmatrix} x + \begin{pmatrix} B \\ -B \end{pmatrix} \geq \begin{pmatrix} b - \sigma \\ -b - \sigma \end{pmatrix}.$$

Then this case can be reduced to the second case.

6 Conclusions

In this paper, we have considered the equivalence conditions between the partial sparse optimization problem and its relaxation, that is, the weighted l_1-norm minimization problem. The proposed conditions are based on the RSP, and guarantee not only the strong equivalence, but also the equivalence between the two problems. The considered problem is more general than the problems considered in literatures.

References

[1] M. Davenport, P. Boufounos, and M. Wakin. Signal processing with compressive measurements, *IEEE Journal of Selected Topics in Signal Processing*, **4**:2(2010),445-460.

[2] J. Wright, A.Y. Yang, A. Ganesh, and S.S. Sastryand Y. Ma, Robust face recognition via sparse representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**:2(2009),210-227.

[3] R. He, W. Zheng, B. Hu, and X. Kong, Nonnegative sparse coding for discriminative semi-supervised learning, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2849-2856, 2011.

[4] M. Sheikh, S. Sarvotham, and O. Milenkovic, DNA array decodin from nonlinear measurements by belief propagation, IEEE Workshop on Statistical Signal Processing (SSP), Madison, Wisconsin, 2007.

[5] M. Lustig, D. Donoho, and J.M. Pauly, Sparse MRI: The application of compressed sensing for rapid mr imaging, *Magnetic Resonance in Medicine*, **58**:6(2007),1182-1195.

[6] W.U. Bajwa, J.D. Haupt, A.M. Sayeed, and R.D. Nowak, Joint source-channel communication for distributed estimation in sensor networks, *IEEE Transactions on Information Theory*, **53**:10(2007),3629-3653.

[7] S.D. Cabrera and T.W. Parks, Extrapolation and spectral estimation with iterative weighted norm modification, *IEEE Transactions on Signal Processing*, **39**:4(1991),842-851.

[8] M. Elad, M.A.T. Figueiredo, and Y. Ma, On the role of sparse and redundant representations in image processing, *Proceedings of the IEEE*, **98**:6(2010),972-982.

[9] X. Zhan, R. Zhang, D. Yin, and C. Huo, SAR image compression using multiscale dictionary learning and sparse representation, *IEEE Geoscience and Remote Sensing Letters*, **10**:5(2013),1090-1094.

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

[11] E. Candès, J. Romberg, and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory*, **52**:2(2006),489-509.

[12] X. Chen and W. Zhou, Convergence of reweighted l_1 minimization algorithms for l_2-l_p minimization, *Computational Optimization and Applications*, **59**:2(2014),47-61.

[13] I. Daubechies, R. DeVore, M. Fornasier, and C.S. Güntürk, Iteratively reweighted least squares minimization for sparse recovery, *Communications on Pure and Applied Mathematics*, **63**:1(2010),1-38.

[14] S. Foucart and M. Lai, Sparsest solutions of undertermined linear systems via l_p-minimization for $0 < p \leq 1$, *Applied and Computational Harmonic Analysis*, **26**(2009), 395-407.

[15] M. Lai and J. Wang, An unconstrainted lpminimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems, *SIAM Journal on Optimination*, **21**(2010),82-101.

[16] P. Combettes and J. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM Journal on Optimination*, **18**:4(2007),1351-1376.

[17] R. Chartrand, Nonconvex compressed sensing and error correction, International Conference on Acoustics Speech and Signal Processing (ICASSP), pages 889-892, 2007.

[18] J. Darbon and M. Sigelle, Image restoration with discrete constrained total variation part i: Fast and exact optimization, *Journal of Mathematical Imaging and Vision*, **26**:3(2006),261-276.

[19] D.L. Donoho and M. Elad Optimally sparse representation in general (nonorthogonal) dictionaries vis l_1 minimization, *Proceedings of the National Academy of Sciences*, **100**:5(2003),2197-2202.

[20] R. Gribonval and M. Nielsen, Sparse representations in unions of bases, *IEEE Transactions on Information Theory*, **49**:12(2003),3320-3325.

[21] A. Cohen, W. Dahmen, and R. DeVore, Compressed sensing and best k-term approximation, *J. Amer. Math. Soc.*, **22**(2009),211-231.

[22] Y. Zhang, A simple proof for recoverability of l_1-minimization (ii): the nonnegative case, *Rice University, Technical Report*, 2005.

[23] E. Candès, Compressive sampling, International Congress of Mathematicians, Madrid, Spain, pages 1433-1452, 2006.

[24] E. Candès and T. Tao, Decoding by linear programming, *IEEE Transactions on Information Theory*, **51**:12(2005),4203C4215.

[25] E. Candès, The restricted isometry property and its implications for compressed sensing, *Comptes Rendus Mathematique*, **346**:10(2008),589-592.

[26] Y.B. Zhao, RSP-based analysis for sparsest and least l_1-norm solutions to underdetermined linear systems, *IEEE Transactions on Signal Processing*, **61**:22(2013),5777-5788.

[27] Y.B. Zhao. Equivalence and strong equivalence between sparsest and least l_1-norm nonnegative solutions to linear systems and their applications, *Journal of the Operations Research Society of China*, **2**:2(2014),171-193.

[28] S. Foucart and H. Rauhut, A Mathematical Introduction to Compressive Sensing, NY: Birkhäuser, New York, 2013.

[29] M. Elad and A. Bruckstein, A generalized uncertainty principle and sparse representation in pairs of bases, *IEEE Transactions on Information Theory*, **48**:9(2002),2558-2567.

[30] D.L. Donoho and J. Tanner, Sparse nonnegative solutions of underdetermined linear equations by linear programming, *Proceeding of the National Academy of sciences USA*, **102**:27(2005),9446-9451.

[31] A. Schrijver, Theory of Linear and Integer Programming, Wiley, New York, 1986.

(*edited by Liangwei Huang*)