Exact Low-rank Matrix Recovery via Nonconvex M_p-Minimization

Lingchen Kong and Naihua Xiu
Department of Applied Mathematics, Beijing Jiaotong University, Beijing, 100044, People's Republic of China

E-mail: lchkong@bjtu.edu.cn and nhxiu@bjtu.edu.cn

Abstract
The low-rank matrix recovery (LMR) arises in many fields such as signal and image processing, statistics, computer vision, system identification and control, and it is NP-hard. It is known that under some restricted isometry property (RIP) conditions we can obtain the exact low-rank matrix solution by solving its convex relaxation, the nuclear norm minimization. In this paper, we consider the nonconvex relaxations by introducing M_p-norm ($0 < p < 1$) of a matrix and establish RIP conditions for exact LMR via M_p-minimization. Specifically, letting $\mathcal{A}$ be a linear transformation from $\mathbb{R}^{m\times n}$ into $\mathbb{R}^s$ and r be the rank of recovered matrix $X \in \mathbb{R}^{m\times n}$, and if $\mathcal{A}$ satisfies the RIP condition $\sqrt{2}\delta_{\max\{r+\frac{3}{2}k,2k\}} + \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}}\delta_{2r+k} < \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}}$ for a given positive integer $k \leq m-r$, then r-rank matrix can be exactly recovered. In particular, we not only obtain a uniform bound on restricted isometry constant $\delta_{4r} < \sqrt{2}-1$ for any $p \in (0,1]$ for LMR via M_p-minimization, but also obtain the one $\delta_{2r} < \sqrt{2}-1$ for any $p \in (0,1]$ for sparse signal recovery via l_p-minimization.

AMS Subject Classification: 62B10, 90C26, 90C59

1 Introduction

The *low-rank matrix recovery* (LMR) is to find the lowest rank matrices based on fewer linear measurements. Mathematically, it is the *rank minimization problem* (RMP) defined as follows:

$$\min\ \operatorname{rank}(X) \quad \text{s.t.}\ \ \mathcal{A}X = b, \tag{1}$$

where $X \in \mathbb{R}^{m\times n}$ is the unknown matrix (information), and $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^s$ is a linear transformation (measurement ensemble) and $b \in \mathbb{R}^s$. It has many applications and appeared in the literature of a diverse set of fields including signal and image processing, statistics, computer vision, system identification and control. For more details, see the recent survey paper by Recht, Fazel and Parrilo [23]. Note that problem (1) is generally NP-hard and ill-posed. A well known heuristic introduced by Fazel, Hindi and Boyd [17] is the famous convex relaxation of LMR, which is called *nuclear norm minimization (NNM)*:

$$\min\ \|X\|_* \quad \text{s.t.}\ \ \mathcal{A}X = b, \tag{2}$$

where $\|X\|_*$ is the *nuclear norm* of X, i.e., the sum of its singular values. When $m = n$ and the matrix $X \triangleq Diag(x)(x \in \mathbb{R}^m)$ is diagonal, the LMR reduces to the *sparse signal recovery (SSR)*:

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

where $A : \mathbb{R}^m \to \mathbb{R}^s$ is a measurement matrix, $\|x\|_0$ is the l_0-*norm* of x, i.e., the number of its nonzero elements. (This is not a norm, as $\|\cdot\|_0$ is not positive homogeneous.) Similarly, the NNM reduces the l_1-*minimization*:

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

where $\|x\|_1$ is the l_1-*norm* of x, i.e., the sum of the absolute value of its entries. These are the problems of compressive sensing (CS, see, e.g., [9, 10, 16]) and there are hundreds of literature concerning them, see the survey paper by Bruckstein, Donoho and Elad [3] and reference therein.

It is well known that under a certain *restricted isometry property* (RIP) condition on the linear transformation one can obtain the exact LMR via NNM (respectively, exact SSR via l_0-minimization). The RIP notion was introduced by Candès and Tao [10] for SSR and generalized to LMR by Recht, Fazel and Parrilo [23]. Recall that the r-*restricted isometry constant (RIC)* δ_r of a linear transformation $\mathcal{A}$ is defined as the smallest constant such that the following holds for all r-*rank matrix* $X \in \mathbb{R}^{m\times n}$ (i.e., the rank of X is no more than r),

$$(1 - \delta_r)\|X\|_F^2 \le \|\mathcal{A}X\|^2 \le (1 + \delta_r)\|X\|_F^2,$$

where $\|X\|_F := \sqrt{\langle X, X\rangle} = \sqrt{trace(X^TX)}$ is the *Frobenius norm* of X, which is equal to the l_2-norm of the vector of singular values. Although the RIP is difficult to verify for a given linear transformation, it is one of the most important concepts in LMR via NNM (respectively, exact SSR via l_0-minimization). The research on RIC seems to be of independent interest, see, e.g., [4, 5, 6, 7, 8, 10, 21, 22] and reference therein.

Note that in the CS context, Chartrand [11] firstly show that fewer measurements are required for exact reconstruction if we replace l_1-norm with l_p-*norm* $(0 < p < 1)$, and Chartrand and Staneva [13] established p-RIP conditions for exact SSR via l_p-*minimization*, which is defined as

$$\min \ \|x\|_p^p \quad \text{s.t. } \ Ax = b, \tag{5}$$

where $\|x\|_p^p := \sum_i |x_i|^p$ and $\|x\|_p := (\sum_i |x_i|^p)^{1/p}$ is the l_p *norm* of x. As l_0-norm, $\|\cdot\|_p$ is not a norm when $p \in (0, 1)$, but $\|\cdot\|_p^p$ satisfies the triangle inequality and induces a metric. Moreover, the numerical experiments in magnetic resonance imaging (MRI) showed that this approach works very efficiently, see [12] for details. SSR and l_p-minimization have been the focus point of some recent research, see, e.g., [1, 2, 12, 13, 14, 15, 18, 19, 20, 25, 26, 27, 28]. Most of them deal with the performance of l_p-minimization and the random measurements with restricted p-isometry property introduced in [13]. For instance, Wang, Xu and Tang [28] studies the performance of l_p-minimization for strong recovery and *weak recovery* where we need to recover all the sparse vectors on one support with one sign pattern; Saab, Chartrand and Yilmaz [25] provides a sufficient condition for SSR via l_p-minimization and provides a lower bound of the support size up to which l_p-minimization can recover all such sparse vectors, and Foucart and Lai [18] improves this bound by considering a generalized version of RIP condition.

This paper deals with the nonconvex relaxation of LMR, *(matrix) M_p-minimization* $(0 < p < 1)$, which is defined as follows:

$$\min\ \|X\|_p^p \quad \text{s.t.}\ \ \mathcal{A}X = b, \tag{6}$$

where $\|X\|_p$ is the M_p *norm* of the matrix X, i.e., $\|X\|_p := (\sum_i \lambda_i^p(X))^{1/p}$ for its singular value decomposition (SVD) $X = UDiag(\lambda(X))V^T$ with $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$, and $\|X\|_p^p = \sum_i \lambda_i^p(X)$. Note that $\|\cdot\|_p$ is not a norm when $p \in (0,1)$ as in the vector case. Obviously, when X is diagonal, M_p-minimization reduces to l_p-minimization.

The main result of this paper is to establish RIP condition for exact LMR via M_p-minimization $(0 < p < 1)$. Based on a block decomposition of the matrix, we obtain a general RIP conditions as follows: if it holds for any positive integer $k \in \{1, 2, \cdots, m-r\}$

$$\sqrt{2}\delta_{\max\{r+\frac{3}{2}k,2k\}} + \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}} \delta_{2r+k} < \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}},$$

then r-rank matrix is guaranteed to be recovered exactly via M_p-minimization for any $p \in (0,1)$. Based on this result, we derive a uniform bound on RIC $\delta_{4r} < \sqrt{2}-1$ for LMR via M_p-minimization, which is independent with $p \in (0,1]$. To the best of our knowledge, these are the first such bounds on RIC for LMR via nonconvex minimizations. We also get a RIP condition $\delta_{2r} < \sqrt{2}-1$ for exact SSR via l_p-minimization, which is independent on $p \in (0,1]$.

The organization of this paper is as follows. In Section 2, by introducing a block decomposition of a matrix and giving some M_p-norm inequalities we prove our main result for LMR. We conclude this paper with the application of our approach to SSR in Section 3.

2 The main results

The main result in this paper is the following theorem.

Theorem 2.1 *Let $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^s$ be a linear transformation and $b \in \mathbb{R}^s$. Let W be a r-rank matrix with $\mathcal{A}W = b$, and $0 < p < 1$. For any positive integer $k \in \{1, 2, \cdots, m-r\}$, if $\mathcal{A}$ satisfies the RIP condition*

$$\sqrt{2}\delta_{\max\{r+\frac{3}{2}k,2k\}} + \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}} \delta_{2r+k} < \left(\frac{k}{2r}\right)^{\frac{1}{p}-\frac{1}{2}},$$

then the unique minimizer of problem (6) is exactly W.

In order to prove our main result, we begin with the following lemma with respect to M_p-norm.

Lemma 2.2 *Let $B, C \in \mathbb{R}^{m\times n}$ be matrices with $B^TC = 0$ and $C^TB = 0$. Let $0 < p < 1$. Then the following holds:*

i) $\|B + C\|_p^p = \|B\|_p^p + \|C\|_p^p$;

ii) $\|B + C\|_p \geq \|B\|_p + \|C\|_p$.

Proof. i) Let the SVDs of B and C as follows:

$$B = (U_B\ U_{B0}) \begin{pmatrix} Diag(\lambda(B)) & 0 \\ 0 & 0 \end{pmatrix} (V_B\ V_{B0})^T,$$

$$C = (U_C\ U_{C0}) \begin{pmatrix} Diag(\lambda(C)) & 0 \\ 0 & 0 \end{pmatrix} (V_C\ V_{C0})^T,$$

where block matrices $(U_B\ U_{B0}), (U_C\ U_{C0}) \in \mathbb{R}^{m\times m}$, $(V_B\ V_{B0}), (V_C\ V_{C0}) \in \mathbb{R}^{n\times n}$, submatrices U_B, V_B, U_C, V_C have the corresponding size with the singular values $\lambda(B), \lambda(C)$ of B and C, respectively. From the assumption, we obtain that $V_B^TU_C = 0$ and $U_B^TV_C = 0$. Therefore, there exists submatrices U_0 and V_0 such that $(U_B\ U_C\ U_0)$ and $(V_B\ V_C\ V_0)$ are orthogonal matrices. Then, we obtain the valid SVD of B and C as

$$B = (U_B\ U_C\ U_0) \begin{pmatrix} Diag(\lambda(B)) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} (V_B\ V_C\ V_0)^T,$$

$$C = (U_B\ U_C\ U_0) \begin{pmatrix} 0 & 0 & 0 \\ 0 & Diag(\lambda(C)) & 0 \\ 0 & 0 & 0 \end{pmatrix} (V_B\ V_C\ V_0)^T.$$

Clearly, $B + C$ has the SVD as

$$B + C = (U_B\ U_C\ U_0) \begin{pmatrix} Diag(\lambda(B)) & 0 & 0 \\ 0 & Diag(\lambda(C)) & 0 \\ 0 & 0 & 0 \end{pmatrix} (V_B\ V_C\ V_0)^T.$$

Thus,

$$\|B + C\|_p^p = \| \begin{pmatrix} Diag(\lambda(B)) & 0 & 0 \\ 0 & Diag(\lambda(C)) & 0 \\ 0 & 0 & 0 \end{pmatrix} \|_p^p = \|B\|_p^p + \|C\|_p^p.$$

This proves the part i).

ii) Note that $\|B + C\|_p = (\|B\|_p^p + \|C\|_p^p)^{1/p}$ by part i). The desired inequality holds immediately by $0 < p < 1$. □

We below introduce a new block decomposition of a matrix, which is basic in our subsequent analysis. For r-rank matrix $W \in \mathbb{R}^{m\times n}$, we denote the *singular value decomposition* (SVD) of W by

$$W = U \begin{pmatrix} Diag(\sigma^r(W)) & 0 \\ 0 & 0 \end{pmatrix} V^T,$$

where $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times m}$, and $\sigma^r(W) := (\sigma_1(W), \cdots, \sigma_r(W))^T$. For W given as above, we give a block decomposition of $Z \in \mathbb{R}^{m\times n}$ with respect to W as follows: let $U^T ZV$ have the block form as

$$U^T ZV = \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{pmatrix},$$

where $Z_{ij} \in \mathbb{R}^{m_i\times n_j}$ with $n_1 + n_2 = m_1 + m_2 = r$ and $m_3 = n_3 = m - r$. Thus, we decompose Z as

$$Z = Z_1 + Z_2 + Z_3 = Z^{(r)} + Z_c^{(r)}, \tag{7}$$

where $Z^{(r)} := Z_1 + Z_2$, $Z_c^{(r)} := Z_3$ and

$$\begin{aligned} Z_1 &= U \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & 0 & 0 \\ Z_{31} & 0 & 0 \end{pmatrix} V^T, \\ Z_2 &= U \begin{pmatrix} 0 & 0 & 0 \\ 0 & Z_{22} & Z_{23} \\ 0 & Z_{32} & 0 \end{pmatrix} V^T, \\ Z_3 &= U \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Z_{33} \end{pmatrix} V^T. \end{aligned}$$

Clearly, $\text{rank}(Z_1) \le m_1 + n_1$, $\text{rank}(Z_2) \le m_2 + n_2$ and $\text{rank}(Z_1 + Z_2) \le 2r$, and Z_1, Z_2, Z_3 are orthogonal one another. In terms of the above decomposition, we have the following property.

Lemma 2.3 *Let W be a r-rank matrix such that $\mathcal{A}W = b$ and let X^* be the optimal solution to the problem (6). Let $Z := X^* - W$ and $Z^{(r)}, Z_c^{(r)}$ defined as above. Then*

$$\|Z_c^{(r)}\|_p^p \le \|Z^{(r)}\|_p^p.$$

Proof. From the assumptions and the above decomposition, we have $W(Z_c^{(r)})^T = 0$ and $(Z_c^{(r)})W^T = 0$. By Lemma 2.2, we have $\|W + Z_c^{(r)}\|_p^p = \|W\|_p^p + \|Z_c^{(r)}\|_p^p$. Therefore, noting that W is the feasible solution to problem (6), we obtain

$$\begin{aligned} \|W\|_p^p \ge \|W + Z\|_p^p &\ge \|W + Z - Z^{(r)}\|_p^p - \|Z^{(r)}\|_p^p \\ &= \|W\|_p^p + \|Z_c^{(r)}\|_p^p - \|Z^{(r)}\|_p^p. \end{aligned}$$

The desired conclusion holds immediately. □

Lemma 2.4 *Let m_1, m_2, n_1, n_2 be positive integers such that $n_1 + n_2 = m_1 + m_2 = r$. Then it holds for given $k \in \{1, 2, \cdots, m - r\}$:*

i) $\min_{m_1,m_2,n_1,n_2} \max_k\{m_1 + n_1 + k, m_2 + n_2 + 2k\} = \max\{r + \frac{3}{2}k, 2k\}$;

ii) $\min_{n_1,n_2} \max\{n_1 + k, n_2 + 2k\} = \max\{\frac{r}{2} + \frac{3}{2}k, 2k\}$.

Proof. i) Note that $(n_1 + n_2 + k) + (m_1 + m_2 + 2k) = 2r + 3k$. Clearly, $\max\{m_1 + n_1 + k, m_2 + n_2 + 2k\} \geq r + \frac{3}{2}k$. If the equality holds, we must have $m_1 + n_1 = r + \frac{1}{2}k, m_2 + n_2 = r - \frac{1}{2}k$. This means that the necessity of the above equality is $k \leq 2r$. Also, when $k > 2r$, we easily obtain that $m_1 + n_1 + k < m_2 + n_2 + 2k$ since $2r \geq m_1 + n_1$. In this case, we may set $m_2 = n_2 = 0$ and get $\max\{m_1 + n_1 + k, m_2 + n_2 + 2k\} = 2k$. Combining the above arguments, we obtain the desired conclusion.

ii) Note that $\max\{n_1 + k, n_2 + 2k\} \geq \frac{r}{2} + \frac{3}{2}k$ because $(n_1 + n_2 + k) + 2k = r + 3k$. Following the similar arguments as in Part 1), we prove the desired result. □

We are ready to prove our main result for exact LMR via nonconvex M_p-minimization.

Proof of Theorem 2.1. Note that the function $\|\cdot\|_p^p$ is lower semi-continuous, level-bounded and proper. By Theorem 1.9 in [24], we obtain that the solution set of M_p-minimization problem (6) is nonempty and compact.

We remain to show that the solution set is a singleton $\{W\}$. Without loss of generality, let X^* be a optimal solution to problem (6). Take $Z = X^* - W$. Clearly, $\mathcal{A}Z = 0$ since W is a feasible solution to the problem (6). Let $W = UDiag(\sigma(W))V^T$, and let Z have the decompositions with respect to W as (7), i.e., $Z = Z_1 + Z_2 + Z_3$. In order to establish the RIP condition, we need further to decompose Z by a decomposition of $Z_c^{(r)} (= Z_3)$. Let SVD of Z_{33} in $\mathbb{R}^{(m-r)\times(m-r)}$ be specified by

$$Z_{33} = PDiag(\sigma(Z_{33}))Q^T$$

where $P, Q \in \mathbb{R}^{(m-r)\times(m-r)}$, and $\sigma(Z_{33}) = (\sigma_1(Z_{33}), \cdots, \sigma_{m-r}(Z_{33}))^T$ is the vector of the singular values of Z_{33} with $\sigma_1(Z_{33}) \geq \cdots \geq \sigma_{m-r}(Z_{33}) \geq 0$. We decompose $\sigma(Z_{33})$ into a sum of vectors $\sigma_{T_i}(Z_{33})(i = 1, 2, \cdots)$, each of sparsity at most k $(1 \leq k \leq m - r)$, where T_1 corresponds to the locations of the k largest entries of $\sigma(Z_{33})$, and T_2 to the locations of the next k largest entries, and so on. We define

$$Z_{T_i} := U \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & PDiag(\sigma_{T_i}(Z_{33})Q^T \end{pmatrix} V^T.$$

Then, Z_{T_1} is the part of Z_3 corresponding to the k largest singular values, Z_{T_2} is the part corresponding to the next k largest singular values, and so on. Clearly, Z_1, Z_2, Z_{T_i} are all orthogonal one another, and $\text{rank}(Z_{T_i}) \leq k$.

We proceed the proof in two steps: the first step shows that $\|Z_3 - Z_{T_1}\|_F$ is essentially bounded by $\|Z_1 + Z_2 + Z_{T_1}\|_F$; The second shows that $Z_1 + Z_2 + Z_{T_1} = 0$, and hence $Z = 0$.

Step 1: From the above decomposition, we easily obtain that for $j \geq 2$,

$$\|Z_{T_j}\|_F^2 \leq k\|Z_{T_j}\|^2 \leq k\left(\frac{\|Z_{T_{j-1}}\|_p^p}{k}\right)^{\frac{2}{p}},$$

where $\|Z_{T_j}\|$ is the spectral (operator) norm of a matrix $Z_{T_j} \in \mathbb{R}^{m\times n}$, i.e., the largest singular value of Z_{T_j}. Then

$$\|Z_{T_j}\|_F \leq k^{\frac{1}{2}-\frac{1}{p}}\|Z_{T_{j-1}}\|_p.$$

By Lemma 2.2, it follows

$$\sum_{j\geq 2}\|Z_{T_j}\|_F \leq k^{\frac{1}{2}-\frac{1}{p}}\sum_{j\geq 2}\|Z_{T_{j-1}}\|_p \leq k^{\frac{1}{2}-\frac{1}{p}}\|Z_3\|_p. \tag{8}$$

This yields

$$\|Z_3 - Z_{T_1}\|_F = \|\sum_{j\geq 2} Z_{T_j}\|_F \leq \sum_{j\geq 2}\|Z_{T_j}\|_F \leq k^{\frac{1}{2}-\frac{1}{p}}\|Z_3\|_p. \tag{9}$$

Noting that $\text{rank}(Z_1 + Z_2) \leq 2r$, we obtain that

$$\begin{aligned}
\|Z_1 + Z_2\|_p^p &= \sum_{1\leq l\leq 2r}\lambda_l^p(Z_1+Z_2)\\
&\leq \left[\sum_{1\leq l\leq 2r}(\lambda_l^p(Z_1+Z_2))^{\frac{2}{p}}\right]^{\frac{p}{2}}\left[\sum_{1\leq l\leq 2r}1\right]^{1-\frac{p}{2}}\\
&= (2r)^{1-\frac{p}{2}}\|Z_1+Z_2\|_F^p,
\end{aligned}$$

where the inequality holds from Hölder's inequality. By Lemma 2.3, it holds

$$\|Z_3\|_p^p \leq \|Z_1 + Z_2\|_p^p. \tag{10}$$

By (9) and (10), it holds that

$$\|Z_3 - Z_{T_1}\|_F \leq \left(\frac{k}{2r}\right)^{\frac{1}{2}-\frac{1}{p}}\|Z_1 + Z_2\|_F. \tag{11}$$

Step 2: Notice that $\mathcal{A}Z = 0$ and

$$\begin{aligned}
\|\mathcal{A}(Z_1 + Z_2 + Z_{T_1})\|^2 &= \langle\mathcal{A}(Z_1+Z_2+Z_{T_1}), \mathcal{A}(Z_1+Z_2+Z_{T_1})\rangle\\
&= \langle\mathcal{A}(Z_1+Z_2+Z_{T_1}), \mathcal{A}Z\rangle - \langle\mathcal{A}(Z_1+Z_2+Z_{T_1}), \mathcal{A}(Z_3 - Z_{T_1})\rangle\\
&= -\langle\mathcal{A}(Z_1+Z_2+Z_{T_1}), \mathcal{A}(Z_3 - Z_{T_1})\rangle.
\end{aligned}$$

Direct calculation yields

$$
\begin{aligned}
& |\langle \mathcal{A}(Z_1 + Z_2 + Z_{T_1}), \mathcal{A}(Z_3 - Z_{T_1})\rangle| \\
= \; & |\langle \mathcal{A}Z_1, \mathcal{A}(Z_3 - Z_{T_1})\rangle + \langle \mathcal{A}(Z_2 + Z_{T_1}), \mathcal{A}(Z_3 - Z_{T_1})\rangle| \\
\leq \; & |\langle \mathcal{A}Z_1, \mathcal{A}(Z_3 - Z_{T_1})\rangle| + |\langle \mathcal{A}(Z_2 + Z_{T_1}), \mathcal{A}(Z_3 - Z_{T_1})\rangle| \\
\leq \; & \delta_{m_1+n_1+k}\|Z_1\|_F \sum_{j\geq 2} \|Z_{T_j}\|_F + \delta_{m_2+n_2+2k}\|Z_2 + Z_{T_1}\|_F \sum_{j\geq 2} \|Z_{T_j}\|_F \\
\leq \; & \delta_{\max\{r+\frac{3}{2}k, 2k\}}(\|Z_1\|_F + \|Z_2 + Z_{T_1}\|_F) \sum_{j\geq 2} \|Z_{T_j}\|_F \\
\leq \; & \sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}} \|Z_1 + Z_2 + Z_{T_1}\|_F \sum_{j\geq 2} \|Z_{T_j}\|_F, \qquad (12)
\end{aligned}
$$

where the first inequality follows from the triangle inequality, the second follows from Lemma 3.3 in [8], and the third follows from Lemma 2.4 and the monotonicity of the RIP constant. It follows that

$$
\|\mathcal{A}(Z_1 + Z_2 + Z_{T_1})\|^2 \;\leq\; \sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}} \sum_{j\geq 2} \|Z_{T_j}\|_F \|Z_1 + Z_2 + Z_{T_1}\|_F.
$$

Combining with $\|\mathcal{A}(Z_1 + Z_2 + Z_{T_1})\|^2 \geq (1 - \delta_{2r+k})\|Z_1 + Z_2 + Z_{T_1}\|_F^2$, we obtain

$$
\|Z_1 + Z_2 + Z_{T_1}\|_F \;\leq\; \frac{\sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}}}{1 - \delta_{2r+k}} \sum_{j\geq 2} \|Z_{T_j}\|_F.
$$

This together with (9) and (10) yields

$$
\begin{aligned}
\|Z_1 + Z_2 + Z_{T_1}\|_F \;&\leq\; \frac{\sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}} k^{1/2-1/p}}{1 - \delta_{2r+k}} \|Z_3\|_p \qquad (13) \\
&\leq\; \beta\|Z_1 + Z_2\|_F \\
&\leq\; \beta\|Z_1 + Z_2 + Z_{T_1}\|_F,
\end{aligned}
$$

where $\beta := \frac{\sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}}}{1-\delta_{2r+k}} \left(\frac{k}{2r}\right)^{1/2-1/p}$. Therefore,

$$
(1 - \beta)\|Z_1 + Z_2 + Z_{T_1}\|_F \;\leq\; 0. \qquad (14)
$$

Since $1 - \beta > 0$ from the assumption, $\|Z_1 + Z_2 + Z_{T_1}\|_F = 0$. Then $Z_1 = Z_2 = Z_{T_1} = 0$ and hence $Z_{T_j} = 0$. Thus, $Z = 0$ and we complete the proof. □

Clearly, from the above proof we know that Theorem 1 is still true for $p = 1$. Moreover, by setting $k = 2r$ we get a RIP condition $\delta_{4r} < \sqrt{2} - 1$, which is independent on $p \in (0, 1]$. We below state this uniform bound result.

Theorem 2.5 *Let $\mathcal{A}:\mathbb{R}^{m\times n}\to\mathbb{R}^s$ be a linear transformation and $b\in\mathbb{R}^s$. Let W be a r-rank matrix with $\mathcal{A}W=b$. If $\mathcal{A}$ satisfies the RIP condition*

$$\delta_{4r}<\sqrt{2}-1,$$

then the unique minimizer of problem (6) for any $p\in(0,1]$ is exactly W.

Furthermore, by choosing different values of k, we easily obtain some bounds on restricted isometry constant for any $p\in(0,1]$ such as $\sqrt{2}\delta_{6r}+\left(\frac{3}{2}\right)^{\frac{1}{p}-\frac{1}{2}}\delta_{5r}<\left(\frac{3}{2}\right)^{\frac{1}{p}-\frac{1}{2}}$, $\delta_{\frac{7}{3}r}+6^{\frac{1}{p}-\frac{1}{2}}\sqrt{2}\delta_{1.5r}<1$, $\delta_{2.5r}+2^{\frac{2}{p}+\frac{1}{2}}\delta_{1.75r}<1$.

3 Uniform bound for SSR

We proceed to analyze the RIP conditions in SSR via l_p-minimization.

Theorem 3.1 *Let $A\in\mathbb{R}^{s\times m}$ be a matrix and $b\in\mathbb{R}^s$. Let $w\in\mathbb{R}^m$ be a r-rank matrix with $Aw=b$, and $0<p<1$. For any positive integer $k\in\{1,2,\cdots,m-r\}$, if A satisfies the RIP condition*

$$\sqrt{2}\delta_{\max\{\frac{r}{2}+\frac{3}{2}k,2k\}}+\left(\frac{k}{r}\right)^{\frac{1}{p}-\frac{1}{2}}\delta_{r+k}<\left(\frac{k}{r}\right)^{\frac{1}{p}-\frac{1}{2}},$$

then the unique minimizer of problem (5) is exactly w.

Proof Let $z=x^*-w$. Clearly, $Az=0$ since w is a feasible solution to the problem (5). Let Z,X^* and W be the diagonal matrix with $Z:=Diag(z),X^*:=Diag(x^*)$ and $W:=Diag(w)$, respectively. and Observe that in this case we only need $Z_1\in\mathbb{R}_{n_1\times n_1},Z_2\in\mathbb{R}_{n_2\times n_2}$. Thus, $rank(Z_1+Z_2)\le n_1+n_2=r$. Following the same analysis in the proof of Theorem 2.1, we then obtain the desired result. □

By setting $k=r$ in the above theorem, we obtain a RIP condition $\delta_{2r}<\sqrt{2}-1$ for SSR, which is clearly independent on $p\in(0,1]$.

Theorem 3.2 *Let $A\in\mathbb{R}^{s\times m}$ be a matrix and $b\in\mathbb{R}^s$. Let $w\in\mathbb{R}^m$ be a r-rank matrix with $Aw=b$. If A satisfies the RIP condition*

$$\delta_{2r}<\sqrt{2}-1,$$

then the unique minimizer of problem (5) for any $p\in(0,1]$ is exactly w.

Note that the RIP condition in Theorem 3.1 is different from those in [11, 13]. To the best of our knowledge, our result gives the first such bound on RIC for LMR via nonconvex minimizations which is independent of p .

Acknowledgments The work was supported in part by the National Natural Science Foundation of China (10831006), and the National Basic Research Program of China (2010CB732501).

References

[1] Berinde R, Gilbert A, Indyk P, Karloff H and Strauss M 2008 Combining geometry and combinatorics: a unified approach to sparse signal recovery *46th Annual Allerton Conference onCommunication, Control, and Computing*

[2] Blanchard J, Cartis C and Tanner J 2009 The restricted isometry property and l^q regularization: phase transitions for sparse approximation *Preprint*

[3] Bruckstein A M, Donoho D L, Elad M 2009 From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images *SIAM Review* **51** 34-81

[4] Cai T T, Wang L and Xu G 2010 Shifting inequality and recovery of sparse signals *IEEE Trans. Inf. Theory* **58** 1300-1308

[5] Cai T T, Wang L and Xu G 2010 New bounds for restricted isometry constants *IEEE Trans. Inf. Theory* **56** 4388-4394

[6] Cai T T, Xu G and Zhang J 2009 On recovery of sparse signals via l_1 minimization *IEEE Trans. Inf. Theory* **55** 3388-3397

[7] Candès E J 2008 The restricted isometry property and its implications for compressed sensing *Academie des Sciences*

[8] Candès E J and Plan Y 2009 Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements *IEEE Trans. Inf. Theory In Press*

[9] Candès E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489-509

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

[11] Chartrand R 2007 Exact reconstructions of sparse signals via nonconvex minimization *IEEE Signal Process. Lett.* **14** 707-710

[12] Chartrand R 2009 Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data *IEEE International Symposium on Biomedical Imaging (ISBI)*

[13] Chartrand R and Staneva V 2008 Restricted isometry properties and nonconvex compressive sensing *Inverse Problems* **24** 1-14

[14] Chen X, Xu F and Ye Y 2009 Lower Bound Theory of Nonzero Entries in Solutions of $l_2 - l_p$ Minimization *to appear in SIAM J. Scientific Computing* .

[15] Davies M E and Gribonval R 2009 Restricted isometry constants where l_p sparse recovery can fail for $0 < p \leq 1$ *IEEE Trans. Inf. Theory* **55** 2203-2214

[16] Donoho D L 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289-1306

[17] Fazel M, Hindi H and Boyd S 2001 A rank minimization heuristic with application to minimum order system approximation *In Proceedings American Control Conference*

[18] Foucart S, Lai M J 2009 Sparsest Solutions of Underdetermined Linear Systems via l_q-minimization for $0 < q \leq 1$ *Applied and Computational Harmonic Analysis* **26** 395-407

[19] Gasso G, Rakotomamonjy A and Canu S 2009 Recovering sparse signals with non-convex penalties and DC programming *IEEE Trans. Signal Processing*, **57** 4686-4698

[20] Ge D, Jiang X and Ye Y 2010 A Note on the Complexity of l_p-Minimization *submitted to Mathematical Programming*

[21] Lee K and Bresler Y 2009 Guaranteed minimum rank approximation from linear observations by nuclear norm minimization with an ellipsoidal constraint *arxiv0903.4742*

[22] Mohan K and Fazel M 2010 New restricted isometry results for noisy low-rank matrix recovery *Proc. Intl. Symp. Inf. Thoery (ISIT) (Austin TX)*

[23] Recht B, Fazel M and Parrilo P 2010 Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization *SIAM Review* **52** 471-501

[24] Rockafellar R T and Wets R J-B 2004 Variational Analysis *Second Edition (New York Springer)*

[25] Saab R, Chartrand R and Yilmaz O 2008 Stable sparse approximations via nonconvex optimization *in Proc. ICASSP*

[26] Saab R and Yilmaz O 2010 Sparse recovery by nonconvex optimization-instance optimality *Applied and Computational Harmonic Analysis Online*

[27] Shen Y and Li S 2011 Restricted p-isometry property and its application for nonconvex compressive sensing *arXiv:1007.4396*

[28] Wang M, Xu W Y and Tang A 2010 On the Performance of Sparse Recovery via l_p-minimization *arXiv:1011.5936v1*