New Bounds for Restricted Isometry Constants in Low-rank Matrix Recovery *

Lingchen Kong†, and Naihua Xiu‡

(January 25, 2011)

Abstract

In this paper, we establish new bounds for restricted isometry constants (RIC) in low-rank matrix recovery. Let $\mathcal{A}$ be a linear transformation from $\mathbb{R}^{m\times n}$ into $\mathbb{R}^p$, and r the rank of recovered matrix $X \in \mathbb{R}^{m\times n}$. Our main result is that if the condition on RIC satisfies $\delta_{2r+k} + 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}} < 1$ for a given positive integer $k \leq m-r$, then r-rank matrix can be exactly recovered via nuclear norm minimization problem in noiseless case, and estimated stably in the noise case. Taking different k, we obtain some improved and new RIC bounds such as $\delta_{\frac{7}{3}r} + 2\sqrt{3}\delta_{1.5r} < 1$, $\delta_{2.5r} + 2\sqrt{2}\delta_{1.75r} < 1$, $\delta_{2r+1} + 2\sqrt{r}\delta_{r+2} < 1$, $\delta_{2r+2} + \sqrt{2r}\delta_{r+3} < 1$, or $\delta_{2r+4} + \sqrt{r}\delta_{r+7} < 1$. To the best of our knowledge, these are the first such conditions on RIC.

Keywords: low-rank matrix recovery, restricted isometry constant, bound, nuclear norm minimization.

1 Introduction

The *noisy low-rank matrix recovery* (noisy LMR) aims to find the lowest rank matrices based on fewer linear measurements, which is a *rank minimization problem* (RMP) with noisy linear measurements:

$$\min\ \mathrm{rank}(X) \quad \text{s.t.}\ \ \mathcal{A}X + \zeta = b, \quad \|\zeta\| \leq \varepsilon, \tag{1}$$

where $X \in \mathbb{R}^{m\times n}$ is the unknown information, and $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^p$ is a linear transformation and $\zeta, b \in \mathbb{R}^p$, $\|\zeta\|$ is the l_2-norm of $\zeta \in \mathbb{R}^p$, an unknown noise term, and ε is an upper bound on the size of noisy contribution. When $\varepsilon = 0$, we call it the *low-rank matrix recovery* (LMR). Problem (1) has many applications and appeared in the literature of

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

† Lingchen Kong, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail: konglchen@126.com)

‡ Naihua Xiu, Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, P. R. China (e-mail: nhxiu@bjtu.edu.cn)

a diverse set of fields including signal and image processing, statistics, computer vision, system identification and control, Euclidean embedding, and collaborative filtering. For more details, see the recent survey paper by Recht, Fazel and Parrilo [13]. Note that problem (1) is generally NP-hard and ill-posed. A popular heuristic introduced by Fazel, Hindi and Boyd [10] is the famous convex relaxation of RMP in noisy case, which is called *nuclear norm minimization (NNM)*:

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

where $\|X\|_*$ is the *nuclear norm* of X, i.e., the sum of its singular values. Similarly, we have the corresponding relaxation for the noiseless case as follows:

$$\min \|X\|_* \quad \text{s.t.} \quad \mathcal{A}X = b. \tag{3}$$

When $m = n$ and the matrix $X := \mathrm{Diag}(x)(x \in \mathbb{R}^n)$ is diagonal, the LMR and NNM reduce respectively to *sparse signal recovery (SSR)* and l_1 *norm minimization*, which are the problems of compressed sensing (CS, see, e.g., [6, 8, 9]).

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. Without loss of generality, let $m \le n$ in this paper. For $r + r' \le m$, the r, r'*- restricted orthogonality constant* $\theta_{r,r'}$ of $\mathcal{A}$ is the smallest constant which satisfies

$$|\langle \mathcal{A}X, \mathcal{A}X'\rangle| \le \theta_{r,r'}\|X\|_F\|X'\|_F,$$

for all r-rank matrices X and r'-rank matrices X', where X, X' are such that if the singular value decomposition (SVD) of X is $X = (U_rU_{m-r})\begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} V_r^T \\ V_{n-r}^T \end{pmatrix}$ then $X' = U_{m-r}Z_1 + Z_2V_{n-r}^T$ for some $Z_1 \in \mathbb{R}^{(m-r)\times n}, Z_2 \in \mathbb{R}^{m\times(n-r)}$.

Although the RIP is difficult to verify for a given linear transformation $\mathcal{A}$, it is one of the most important concepts in low-rank matrix recovery. Therefore, the research on RIC is of independent interest. Recht et al [13] showed that if $\delta_{2r} < 1$, then the r-rank matrix solution to RMP in the noiseless case is unique, and if $\delta_{5r} < 1/10$, then this unique solution can be exactly recovered via NNM. Lee and Bresler [11] gave $\delta_{3r} < 1/(1 + 4/\sqrt{3})$ by employing an analogue of the approach for SSR in [4]; Candès and Plan [5] obtained $\delta_{4r} < \sqrt{2} - 1$ based on the work [4, 8]; Mohan and Fazel [12] provided $\delta_{2r+\alpha} + \sqrt{1/\beta}\theta_{2r+\alpha,\beta r} < 1$[§] $(2\alpha \le \beta \le 4\alpha, \alpha > 0)$ by combining a condition on $\theta_{2r+\alpha,\beta r}$, which extended the recent work in CS [1, 2, 3]. Moreover, they derives RIP condition on RIC δ_{cr} with $c > 2$ such as $\delta_{3r} < 2\sqrt{5} - 4$, and $\delta_{4r} < (8 - \sqrt{40})/3$.

The aim of this paper is to establish new RIP condition for LMR via NNM. Based on a new block decomposition of the matrix, we obtain a general RIP conditions as follows: if it holds for $k \in \{1, 2, \cdots, m - r\}$

$$\delta_{2r+k} + 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3k}{2},2k\}} < 1,$$

[§]For a positive real number β, βr is understood as $\lceil \beta r\rceil$.

then r-rank matrices are guaranteed to be recovered exactly via NNM in noiseless case and this recovery is stable in general case, which improves the existing RIP results. For instants, letting $k = r$, we get $\delta_{3r} + 2\delta_{2.5r} < 1$; and letting $k = 2r$, we get $\delta_{4r} < \sqrt{2} - 1$. In particular, we give new bounds on RICs such as $\delta_{\frac{7}{3}r} + 2\sqrt{3}\delta_{1.5r} < 1$, $\delta_{2.5r} + 2\sqrt{2}\delta_{1.75r} < 1$, $\delta_{2r+1} + 2\sqrt{r}\delta_{r+2} < 1$, $\delta_{2r+2} + \sqrt{2r}\delta_{r+3} < 1$, or $\delta_{2r+4} + \sqrt{r}\delta_{r+7} < 1$; see Theorem 3.1 for more details. To the best of our knowledge, these are the first such conditions on RIC for LMR via NNM.

The organization of this paper is as follows. In Section 2, we introduce a new block decomposition of a matrix and then show our main result. We conclude this paper with new bounds for RICs in Section 3.

2 The main results

The main results in this paper include two theorems. One is Theorem 2.1 for noisy LMR via NNM, and the other is Theorem 2.5 for noiseless case. We first consider the noisy recovery.

Theorem 2.1 (Noisy recovery) *Let W be a matrix such that $\|\mathcal{A}W - b\| \leq \varepsilon$ with $\varepsilon \geq 0$, and let X^* be the optimal solution to the problem (2). For $1 \leq k \leq m - r$, assume that $\delta_{2r+k} + 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}} < 1$. Then it holds that*

$$\|X^* - W\|_F \quad \leq \quad C_1\|W - W^r\|_* + C_0\varepsilon, \tag{4}$$

where W^r is the best r-rank approximation *of W specified as*

$$W^r := \text{argmin}_Y\{\|Y - W\|_F : \text{rank}(Y) \leq r, Y \in \mathbb{R}^{m\times n}\},$$

and the constants C_0, C_1 are given explicitly as

$$\begin{aligned} C_0 &:= \frac{2(1 + \sqrt{2}(\frac{r}{k})^{1/2})\sqrt{1 + \delta_{2r+k}}}{1 - \delta_{2r+k} - 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}, \\ C_1 &:= 2k^{-1/2} + \frac{4(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}{1 - \delta_{2r+k} - 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}. \end{aligned}$$

In order to prove the above result, we begin with a new block decomposition of a matrix. For $W \in \mathbb{R}^{m\times n}$, we denote the *singular value decomposition* (SVD) of W by

$$W = U Diag(\sigma(W))V^T,$$

where $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times m}$, and $Diag(\sigma(W))$ is the diagonal matrix of the vector $\sigma(W) := (\sigma_1(W), \cdots, \sigma_m(W))^T$ of the singular values of W. Without loss of generality, let $\sigma_1(W) \geq \cdots \geq \sigma_m(W) \geq 0$. In this case, we specify r-rank matrix W^r as

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

where $\sigma^r(W) := (\sigma_1(W), \cdots, \sigma_r(W))^T$. It is easy to see that W^r is the best r-rank approximation of W. We denote $W_c^r := W - W^r$. 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

$$Z = Z_1 + Z_2 + Z_3 = Z^{(r)} + Z_c^{(r)}, \tag{5}$$

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

Lemma 2.2 *Let W be a matrix such that $\|\mathcal{A}W - b\| \le \varepsilon$ with $\varepsilon \ge 0$, and let X^* be the optimal solution to the problem (2). Let $Z := X^* - W$ and $Z^{(r)}, Z_c^{(r)}, W_c^r$ defined as above. Then*

$$\|Z_c^{(r)}\|_* \le \|Z^{(r)}\|_* + 2\|W_c^r\|_*.$$

Proof. From the assumptions and the above decomposition, we have $W^r(Z_c^{(r)})^T = 0$ and $(Z_c^{(r)})(W^r)^T = 0$. By Lemma 2.3 in [13], we have $\|W^r + Z_c^{(r)}\|_* = \|W^r\|_* + \|Z_c^{(r)}\|_*$. Therefore, noting that W is the feasible solution to problem (2), we obtain

$$\begin{aligned} \|W\|_* \ge \|W + Z\|_* &\ge \|W^s + Z - Z^{(r)}\|_* - \|Z^{(r)} + W - W^r\|_* \\ &= \|W^r\|_* + \|Z - Z^{(r)}\|_* - \|Z^{(r)}\|_* - \|W - W^r\|_*. \end{aligned}$$

Together with the fact that $\|W\|_* - \|W^r\|_* = \|W_c^r\|_*$, we yield the desired conclusion. □

Before proving our main result, we need the following lemma.

Lemma 2.3 *Let m_1, m_2, n_1, n_2 be integers such that $n_1 + n_2 = m_1 + m_2 = r$. Then for given $k \in \{1, 2, \cdots, m - r\}$ one may set m_1, m_2, n_1, n_2 to satisfy the following system*

$$\max\{m_1 + n_1 + k, m_2 + n_2 + 2k\} = \max\{r + \frac{3}{2}k, 2k\}.$$

Proof. 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. □

We are ready to prove our main result for noisy LMR via NNM.

Proof of Theorem 4. The proof is inspired by similar work for CS [4, 6]. Let $Z = X^* - W$. We begin with the simple system

$$\|\mathcal{A}Z\|_F = \|\mathcal{A}(X^* - W)\|_F \leq \|\mathcal{A}X^* - b\|_F + \|\mathcal{A}W - b\|_F \leq 2\varepsilon, \tag{6}$$

which is immediate from the triangle inequality and the fact that W is a feasible solution to the problem (2). Let $W = UDiag(\sigma(W))V^T$, and let Z have the decompositions with respect to W as (5), i.e., $Z = Z_1 + Z_2 + Z_3$. In order to prove (4), 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 $\operatorname{rank}(Z_{T_i}) \leq k$.

As in [4, 7], 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 $\|Z_3 - Z_{T_1}\|_F$ is appropriately small.

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

$$\|Z_{T_j}\|_F \leq k^{1/2}\|Z_{T_j}\| \leq k^{-1/2}\|Z_{T_{j-1}}\|_*,$$

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 it follows

$$\sum_{j\geq 2} \|Z_{T_j}\|_F \leq k^{-1/2} \sum_{j\geq 2} \|Z_{T_{j-1}}\|_* \leq k^{-1/2}\|Z_3\|_*. \tag{7}$$

This yields

$$\|Z_3 - Z_{T_1}\|_F = \|\sum_{j\geq 2} Z_{T_j}\|_F \leq \sum_{j\geq 2} \|Z_{T_j}\|_F \leq k^{-1/2}\|Z_3\|_*. \tag{8}$$

By Lemma 2.2, it holds

$$\|Z_3\|_* \le \|Z_1 + Z_2\|_* + 2\|W_c^r\|_*. \tag{9}$$

Note that $\|Z_1 + Z_2\|_* \le \sqrt{2r}\|Z_1 + Z_2\|_F$ by the Cauchy-Schwarz inequality and the fact that $\text{rank}(Z_1 + Z_2) \le 2r$. By (8) and (9), it holds that

$$\|Z_3 - Z_{T_1}\|_F \le k^{-1/2}\sqrt{2r}\|Z_1 + Z_2\|_F + 2k^{-1/2}\|W_c^r\|_*. \tag{10}$$

Step 2: Notice that

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

From (6) and the RIP definition, we obtain that

$$\begin{aligned}|\langle \mathcal{A}(Z_1 + Z_2 + Z_{T_1}), \mathcal{A}Z\rangle| &\le \|\mathcal{A}(Z_1 + Z_2 + Z_{T_1})\|\|\mathcal{A}Z\| \\ &\le 2\varepsilon\sqrt{1 + \delta_{2r+k}}\|Z_1 + Z_2 + Z_{T_1}\|_F.\end{aligned} \tag{11}$$

On the other hand, direct calculation yields

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

where the first inequality follows from the triangle inequality, the second follows from Lemma 3.3 in [5], and the third follows from Lemma 2.3 and the monotonicity of the RIP constant. It holds from (11) and (12) that

$$\|\mathcal{A}(Z_1 + Z_2 + Z_{T_1})\|_F^2 \le [2\varepsilon\sqrt{1 + \delta_{2r+k}} + \sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}}\sum_{j\ge 2}\|Z_{T_j}\|_F]\|Z_1 + Z_2 + Z_{T_1}\|_F.$$

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

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

This together with (8) and (9) yields

$$\begin{aligned}\|Z_1 + Z_2 + Z_{T_1}\|_F &\le \frac{2\varepsilon\sqrt{1 + \delta_{2r+k}}}{1 - \delta_{2r+k}} + \frac{\sqrt{2}\delta_{\max\{r+\frac{3}{2}k, 2k\}}k^{-1/2}}{1 - \delta_{2r+k}}\|Z_3\|_* \\ &\le \alpha\varepsilon + \beta[\|Z_1 + Z_2\|_* + 2\|W_c^r\|_*] \\ &\le \alpha\varepsilon + \sqrt{2r}\beta[\|Z_1 + Z_2 + Z_{T_1}\|_F + 2\Delta],\end{aligned} \tag{13}$$

where $\alpha := \frac{2\sqrt{1+\delta_{2r+k}}}{1-\delta_{2r+k}}$, $\beta := \frac{\sqrt{2}k^{-1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}{1-\delta_{2r+k}}$ and $\Delta := \|W_c^r\|_*$. Therefore,

$$\|Z_1 + Z_2 + Z_{T_1}\|_F \quad \leq \quad (1 - \sqrt{2r}\beta)^{-1}(\alpha\varepsilon + 2\sqrt{2r}\beta\Delta). \tag{14}$$

From the above two steps, combining (10) with (14), we obtain that

$$\begin{aligned} \|Z\|_F &\leq \|Z_1 + Z_2 + Z_{T_1}\|_F + \|Z_3 - Z_{T_1}\|_F \\ &\leq (1 + k^{-1/2}\sqrt{2r})\|Z_1 + Z_2 + Z_{T_1}\|_F + 2k^{-1/2}\Delta \\ &\leq \alpha(1 + k^{-1/2}\sqrt{2r})(1 - \sqrt{2r}\beta)^{-1}\varepsilon + (2k^{-1/2} + 2\sqrt{2r}\beta(1 - \sqrt{2r}\beta)^{-1})\Delta \\ &= C_1\Delta + C_0\varepsilon, \end{aligned}$$

where $C_1 := 2k^{-1/2} + 2\sqrt{2r}\beta(1 - \sqrt{2r}\beta)^{-1}$ and $C_0 := \alpha(1 + k^{-1/2}\sqrt{2r})(1 - \sqrt{2r}\beta)^{-1}$. Noting that $\sqrt{2r}\beta < 1$ is equivalent to the RIP condition $\delta_{2r+k} + 2\sqrt{\frac{r}{k}}\delta_{\max\{r+\frac{3}{2}k,2k\}} < 1$, we complete the proof. □

As a byproduct of the proof of Theorem 2.1, we can obtain a useful inequality.

Proposition 2.4 *One has*

$$\|Z_1 + Z_2\|_* \quad \leq \quad \gamma\|Z_3\|_*, \quad \gamma := \frac{2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}{1 - \delta_{2r+k}}.$$

Proof. From the proof of Theorem 2.1, setting $\varepsilon = 0$ in (13), we obtain

$$\|Z_1 + Z_2 + Z_{T_1}\|_F \leq \frac{\sqrt{2}\delta_{\max\{r+\frac{3}{2}k,2k\}}k^{-1/2}}{1 - \delta_{2r+k}}\|Z_3\|_*.$$

It is easy to see $\|Z_1 + Z_2\|_* \leq \sqrt{2r}\|Z_1 + Z_2\|_F \leq \sqrt{2r}\|Z_1 + Z_2 + Z_{T_1}\|_F$, and hence we conclude the proof. □

As an application of Proposition 2.4, in the noiseless case, we obtain from (9) that

$$\|Z_3\|_* \leq \gamma\|Z_3\|_* + 2\|W_c^r\|_*.$$

Then, it holds $\|Z_3\|_* \leq 2(1 - \gamma)^{-1}\|W_c^r\|_*$. Therefore, we have

$$\|X^* - W\|_* = \|Z\|_* \leq \|Z_1 + Z_2\|_* + \|Z_3\|_* \leq C_2\|W_c^r\|_*,$$

where $C_2 := 2(1 + \gamma)(1 - \gamma)^{-1}$. Summarizing the above analysis we give the following result for the noiseless recovery.

Theorem 2.5 (Noiseless recovery) *Let W be a matrix such that $\mathcal{A}W = b$, and let X^* be the optimal solution to the problem (3). For $1 \leq k \leq m - r$, assume that $\delta_{2r+k} + 2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}} < 1$. Then it holds that*

$$\|X^* - W\|_F \quad \leq \quad C_1\|W - W^r\|_*, \tag{15}$$

and

$$\|X^* - W\|_* \quad \leq \quad C_2\|W - W^r\|_*, \tag{16}$$

for the same constant C_1 given in Theorem 2.1, and C_2 given explicitly as

$$C_2 := \frac{2(1-\delta_{2r+k}+2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}})}{1-\delta_{2r+k}-2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}}}.$$

In particular, if W is r-rank, the recovery is exact.

The above theorems say that we can stably recovery low-rank matrix via NNM under the assumption on RIP constant of $\delta_{2r+k}+2(\frac{r}{k})^{1/2}\delta_{\max\{r+\frac{3}{2}k,2k\}} < 1$ for $1 \leq k \leq m-r$, which compares well with the recent RIP results. Moreover, r-rank matrices are guaranteed to be recovered exactly via NNM in noiseless case. Our results generalizes Theorems 1.1 and 1.2 for SSR in [4, 7].

3 New RIC bounds

We proceed to analyze the RIP conditions in Theorems 2.1 and 2.5. By choosing different values of k, we obtain that each of the following conditions is sufficient for the exact recovery of low-rank matrix in the noiseless case and stable recovery in the noisy case.

$k = cr$	RIP conditions
$k = \frac{r}{3}$	$\delta_{\frac{7}{3}r} + 2\sqrt{3}\delta_{1.5r} < 1$
$k = \frac{r}{2}$	$\delta_{2.5r} + 2\sqrt{2}\delta_{1.75r} < 1$
$k = r$	$\delta_{3r} + 2\delta_{2.5r} < 1$
$k = 2r$	$\delta_{4r} < \sqrt{2} - 1$
$k = 3r$	$\delta_{5r} + \frac{2}{\sqrt{3}}\delta_{6r} < 1$

Table 1

It is easy to see from Table 1 that our results improve the previous results of [13, 11, 5], where recovery is shown via NNM if $\delta_{5r} < 1/10$ in [13], $\delta_{3r} < 1/(1+4/\sqrt{3})$ in [11], or $\delta_{4r} < \sqrt{2}-1$ in [5]. Moreover, when $r > 20$, $2r+4 \leq 2.2r$ and when $r > 30$, $2r+3 \leq 2.1r$. Thus, taking $k \in \{1, 2, 3, 4\}$, we get the following conditions which are of certain interest.

k	RIP conditions
$k = 1$	$\delta_{2r+1} + 2\sqrt{r}\delta_{r+2} < 1$
$k = 2$	$\delta_{2r+2} + \sqrt{2r}\delta_{r+3} < 1$
$k = 3$	$\delta_{2r+3} + 2\sqrt{r/3}\delta_{r+5} < 1$
$k = 4$	$\delta_{2r+4} + \sqrt{r}\delta_{r+7} < 1$

Table 2

Summarizing Theorems 2.1 and 2.5, Tables 1 and 2, we obtain the following new bounds on RIC in LMR via NNM:

Theorem 3.1 *Under the above assumptions as in Theorems 2.1 and 2.5. Each of the following conditions is sufficient for the exact recovery of r-rank matrix in the noiseless case and stable recovery in the noisy case:*

i) $\delta_{\frac{7}{3}r} + 2\sqrt{3}\delta_{1.5r} < 1$;

ii) $\delta_{2.5r} + 2\sqrt{2}\delta_{1.75r} < 1$;

iii) $\delta_{2r+1} + 2\sqrt{r}\delta_{r+2} < 1$;

iv) $\delta_{2r+2} + \sqrt{2r}\delta_{r+3} < 1$;

v) $\delta_{2r+3} + 2\sqrt{r/3}\delta_{r+5} < 1$;

vi) $\delta_{2r+4} + \sqrt{r}\delta_{r+7} < 1$.

An interesting fact in above theorem is that we do not need additional condition such as r, r'-restricted orthogonality property of $\mathcal{A}$. However, we do not completely obtain the RIC bounds of form $\delta_{cr} < \gamma$ with $c = 1$ or $c < 2$ as done in the CS context [2], which is desirable and important among the RIP conditions for r-rank matrix recovery. We will leave this as future research topic.

References

[1] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals", IEEE Trans. Inf. Theory, vol. 58, no. 3, pp. 1300-1308. 2010.

[2] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants", IEEE Trans. Inf. Theory, vol. 56, no. 9, pp. 4388-4394, 2010.

[3] T. T. Cai, G. Xu, and J. Zhang, "On recovery of sparse signals via l_1 minimization", IEEE Trans. Inf. Theory, vol. 55, no. 7, pp. 3388-3397, 2009.

[4] E.J. Candès, "The restricted isometry property and its implications for compressed sensing", Academie des Sciences, 2008.

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

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information", IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489-509, 2006.

[7] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements", Comm. Pure Appl. Math., vol. 59, no. 8, pp. 1207-1223, 2006.

[8] E. J. Candès and T. Tao, "Decoding by linear programming", IEEE Trans. Inf. Theory, vol. 51, no. 12, pp. 4203-4215, 2005.

[9] D. L. Donoho, "Compressed sensing", IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289-1306, 2006.

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

[11] K. Lee, and Y. Bresler, "Guaranteed minimum rank approximation from linear observations by nuclear norm minimization with an ellipsoidal constraint", Available online at http://arxiv.org/abs/0903.4742 Submitted on 27 Mar 2009.

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

[13] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization", SIAM Review, vol. 52, no. 3, pp. 471-501, 2010.