On the Schatten p-Quasi-Norm Minimization for Low-Rank Matrix Recovery

Ming-Jun Lai [*], Yang Liu [†], Song Li [‡], Huimin Wang[§]

November 4, 2020

Abstract

The first part of the paper proves the conjectures on an inequalities in the Schatten p-quasi-norm of matrices. The second part of the paper uses the inequalities for proving a sufficient condition when the Schatten p-quasi-norm minimization can be used for low rank matrix recovery. More precisely, when the restricted isometry constant $\delta_{2s} < 1$, there exists a real number $p_0 < 1$ such that any solution of the p minimization is the minimal rank solution for any $p \le p_0$. In addition, in the noisy setting, the estimate of the difference $\|X - X_0\|$ is also given which is useful for applications.

Keywords and Phrases: singular values, schatten p norm, restricted isometry property, low rank matrix recovery
Subject Classifications:[AMS 2000] 15A04; 46B09; 15A60

1 Introduction

In this paper, we consider matrices of size $m \times n$ for integer m, n. Let Ω be a given subset of the index set $\{(i,j), i = 1, \cdots, m, j = 1, \cdots, n\}$. Suppose the entries of a matrix M with indices in Ω are given. We look for this matrix M. This is the so-called matrix completion problem. Of course, there are many such matrices. We thus find a matrix X^* of size $m \times n$

[*] mjlai@math.edu. This author is associated with Department of Mathematics, The University of Georgia, Athens, GA 30602

[†] yliu9t@gmail.com. This author is associated with Shenzhen Tech. Univ., Shenzhen, 518118, P. R. China

[‡] songli@math.zju.edu.cn. This author is associated with Department of Mathematics, Zhejiang University, Hangzhou, 310027, P. R. China

[§] glanme@126.com. This author is associated with Department of Mathematics, Shaoxing University, Shaoxing, Zhejiang, P. R. China

with smallest rank such that the entries of X^* in Ω are the same as the entries of M in Ω. In general, we solve the following rank minimization problem:

$$\min_{X\in\mathbb{R}^{m\times n}} \quad \text{rank } (X): \quad \text{such that} \quad \mathcal{A}(X)=\mathcal{A}(M), \tag{1}$$

where $\mathcal{A}:\mathbb{R}^{m\times n}\mapsto\mathbb{R}^{\ell}$ is a linear mapping. For example, in the matrix completion setting, $\mathcal{A}(X)=[x_{ij},\forall(i,j)\in\Omega]$ for any matrix $X=[x_{ij}]_{1\leq i\leq m,1\leq j\leq n}\in\mathbb{R}^{m\times n}$. For a general $\mathcal{A}$, the problem (1) is the so-called low-rank matrix recovery problem which has been an active research area in last 15 years. In this topic of research, there have been significant works on extending the conditions from recovering sparse vectors to those for recovering low-rank matrices, including the research on developing various algorithms for recovering low-rank matrices. We refer the interested reader to [11], [25], [10], [22], [23], [19], [24] for theoretical understanding and to [12], [20], and the references therein for computational approaches. Many more studies on matrix completion have been published although the problem in this paper has not been completely settled to the best of the authors' knowledge.

In [24], Oymak, Mohan, Fazel and Hassibi extended recovery conditions from sparse vectors to low-rank matrices in a simple and transparent way. That is, they observed a key equivalence between the sparse vector recovery and low rank matrices recovery using the ℓ_1 minimization in terms of the well-known restricted isometry property and the null space conditions. However, when extending the necessary and sufficient condition in terms of null space property for sparse vector recovery in p quasi-norm with $p\in(0,1)$ (cf. e.g., [15] and [14]) to the setting of the recovery of low-rank matrices by Schatten p quasi-norm minimization:

$$\begin{array}{rl} \text{minimize} & \|X\|_p^p \\ \text{subject to} & \mathcal{A}(X)=y, \end{array} \tag{2}$$

where $\|X\|_p^p=\sum_{i=1}^k\sigma_i^p(X)$ with $k=\min\{m,n\}$, they could not get a condition which is both necessary and sufficient for recovery low-rank matrices (cf. [24]), where $p\in(0,1)$, $y=\mathcal{A}(X_0)$ with X_0 being the low-rank solution to be recovered. Their sufficient condition for the recovery of a low-rank matrix of rank s is that

$$\sum_{i=1}^{2s}\sigma_i^p(W)\leq\sum_{i=2s+1}^{n}\sigma_i^p(W) \tag{3}$$

for all $W\in\mathcal{N}(\mathcal{A})$ with $W\neq 0$, where $\sigma_i(W)$ is the i-th singular value of W and $\mathcal{N}(\mathcal{A})$ is the null space of $\mathcal{A}$. Their necessary condition for the recovery is that

$$\sum_{i=1}^{s}\sigma_i^p(W)\leq\sum_{i=s+1}^{n}\sigma_i^p(W) \tag{4}$$

for all $W\in\mathcal{N}(\mathcal{A})\backslash\{0\}$. To bridge the gap between these two conditions, they will need to have the following inequality in (5). As pointed out by these researchers, the inequalities are true for $p = 1$ and $p = 0$ and for general $p \in (0, 1)$ in their numerical experiments. Thus they conjectured that this inequality holds for all matrices A and B of the same size for general integer $1 \leq k \leq min\{m, n\}$. This is a well-known conjecture in community of compressed sensing and low-ranked matrix recovery. See a recent study in [13], where the conjecture was settled for one case when $k = \min\{m, n\}$.

One of the goals of the study in this paper is to explain how to prove the following conjecture (cf. [[24], Section VI]) for any positive integer $1 \leq k \leq \min\{m, n\}$.

Theorem 1.1 *Let $p \in [0, 1]$. For all real matrices A and B of size m by n,*

$$\sum_{i=1}^{k} |\sigma_i^p(A) - \sigma_i^p(B)| \leq \sum_{i=1}^{k} \sigma_i^p(A - B). \tag{5}$$

for $1 \leq k \leq \ell = \min\{m, n\}$, where $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_\ell(A) \geq 0$, $\sigma_1(B) \geq \sigma_2(B) \geq \cdots \geq \sigma_\ell(B) \geq 0$ and $\sigma_1(A - B) \geq \sigma_2(A - B) \geq \cdots \geq \sigma_\ell(A - B) \geq 0$ are singular values of A, B and $A - B$, respectively.

This motivates a more general conjecture (cf. [[3], Conjecture 6]) given below.

Theorem 1.2 *Let A and B be $m \times n$ matrices. Let f be a nonnegative concave function: $\mathbb{R}_+ \to \mathbb{R}_+$ with $f(0) = 0$. Then*

$$\sum_{i=1}^{k} |f(\sigma_i(A)) - f(\sigma_i(B))| \leq \sum_{i=1}^{k} f(\sigma_i(A - B)) \tag{6}$$

for any $1 \leq k \leq \min(n, m)$.

In addition to establish Theorems 1.1 and 1.2, we use the results to answer the original question on matrix completion problem: under what condition on $p \in (0, 1)$ the minimizer of (2) is the solution of (1)? Although Yue and So in [32] and Foucart in [13] mentioned that the result would answer the question on matrix recovery problem, they did not write down any detail. In particular, the study of matrix recovery in the noisy setting (7) has not been discussed in [32] and [13] which is most useful for applications. Recently we found out that such estimates were given in [34] for the Schatten p-quasi-norm minimization by using a complicated approach. We feel that it is important to use Theorems 1.1 to present these results explicitly. This motivates us to present Theorems 1.3 and 1.4 below which are similar to the results in [34] with a much simpler proof.

To describe when the minimizer of (2) is the solution of (1) when $p = 1$, the researchers in [25] introduced the following concept similar to the restricted isometry property for the study of compressed sensing.

Definition 1.1 *Let $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^l$ be a linear map . Let δ_s be the smallest number such that*

$$(1-\delta_s)\|X\|_F^2 \le \|\mathcal{A}(X)\|_2^2 \le (1+\delta_s)\|X\|_F^2$$

holds for all matrices X of rank at most s, where $\|\cdot\|_F$ stands for the well-known Frobenius norm of matrices. δ_s is called matrix restricted isometry constant (RIC).

Note that our definition is slightly different from the one in [25]. The researchers in [25] showed that there is an overwhelming probability that the entries of M in Ω satisfies the RIP for any fixed RIC $\delta_s < 1$. Many results have been obtained based on the RIP. For example, in [11], Candès and Plan showed that when $\delta_{4s} < \sqrt{2} - 1$, the solution of the minimization (2) with $p = 1$ is unique and is the low-rank solution (1). The matrix RIC has been improved, e.g., $\delta_{5s} < 0.607, \delta_{4s} < 0.558, \delta_{3s} < 0.4721$ in Mohan and Fazel's work [22], $\delta_{2s} < 0.4931, \delta_s < 0.307$ in Wang and Li's work in [31] and $\delta_{2s} \le 1/2$ and $\delta_s < 1/3$ in Cai and Zhang's work [6] and [7]. In fact, $\delta_s < 1/3$ seems the best estimate. Furthermore, in [8] and [33], the researchers are able to show that the sparse recovery for $\delta_{2s} < 1/\sqrt{2}$ and $\delta_{ts} < t/(4-t)$ for $0 < t < 4/3$, respectively. The result in [24] is also worthy mentioning. The researchers showed that any sufficient condition for recovery of the sparse solution in the compressed setting can be translated into a sufficient condition to recover the minimal rank solution in the matrix completion setting. That is, the results in [8] and [33] have corresponding version in matrix recovery. Recently, in [34], Zhang and Li also study on the RIP condition for Schatten p quasi-norm minimization (2). They obtained the results similar to our Theorems 1.3 and 1.4 with $\delta_{2s} < \delta(p) = \eta/(2-p-\eta)$ with $\eta \in (1-p, 1-p/2)$. As seen from our results below, we are able to establish the desired results under the condition $\delta_{2s} < 1$ which is the best possible.

Now let us use the Schatten p-quasi-norm to describe a sufficient condition when a minimizer of (2) is the solution of (1). Another one of the main results in this paper is the following

Theorem 1.3 *Suppose that the linear mapping $\mathcal{A}$ whose matrix RIP constant $\delta_{2s} < 1$. There exists a number $p_0 \in (0,1)$ such that for any $p \le p_0$, each minimizer X^* of the Schatten p quasi-norm minimization (2) is the lowest rank solution of (1).*

This shows that using Schatten p quasi-norm minimization (2) with $p < 1$ has a better chance to recovery a rank r solution than using the standard ℓ_1 norm minimization (2) with $p = 1$ for matrix recovery. Similarly, we shall discuss the noisy recovery case. Let X^* be the optimal solution of the following problem

$$\min_{X\in\mathbb{R}^{m\times n}} \sum_{i=1}^m \sigma_i^p(X), \quad \text{such that} \quad \|\mathcal{A}(X) - \mathcal{A}(M)\|_2 \le \eta, \tag{7}$$

where $\eta > 0$ is a noise level. We can establish the following result.

Theorem 1.4 *If the linear mapping $\mathcal{A}$ whose matrix RIP constant $\delta_{2s} < 1$, there exists some $p_0 < 1$, such that for any $p \leq p_0$ we have*

$$\|X^* - X^0\|_p \leq \frac{2^{1/p} s^{1/p-1/2} \eta}{\sqrt{1-\delta_{2s}}(1 - \frac{1}{\sqrt{s}})} \tag{8}$$

Let us review briefly some known results related to Theorems 1.1 and 1.2. The conjecture is true for $p = 1$. That is, we have the well-known Mirsky's inequality:

Theorem 1.5 *For all real matrices A and B of size m by n,*

$$\sum_{i=1}^{k} |\sigma_i(A) - \sigma_i(B)| \leq \sum_{i=1}^{k} \sigma_i(A - B) \tag{9}$$

for $1 \leq k \leq \min(m, n)$.

In other words, Theorem 1.1 holds for $p = 1$. See a simple proof in Appendix 5.1. Clearly, Theorem 1.1 holds for $p = 0$. The inequality in (5) is stronger than the following one

Theorem 1.6 *For any two matrices A, B which have the same size, say $m \times n$ with $m \leq n$, it holds that*

$$\sum_{i=1}^{k} (\sigma_i^p(A) - \sigma_i^p(B)) \leq \sum_{i=1}^{k} \sigma_i^p(A - B) \tag{10}$$

for all $k = 1, ..., m$, where $0 \leq p \leq 1$.

The above inequality in (10) is known to be true as early as in [26] (for an English version, see [27]). See [29] for a general version with a short proof. It is also a consequence of the more general result in [30].

Recently, Yue and So in [32] and Foucart in [13] established the Miao conjecture for the following case $k = \min\{m, n\}$.

Theorem 1.7 *Let A and B be $m \times n$ matrices. Let f be a nonnegative concave function: $\mathbb{R}_+ \to \mathbb{R}_+$ with $f(0) = 0$. Then*

$$\sum_{i=1}^{\min\{m,n\}} |f(\sigma_i(A)) - f(\sigma_i(B))| \leq \sum_{i=1}^{\min\{m,n\}} f(\sigma_i(A - B)). \tag{11}$$

However, the result in Theorem 1.2 is stronger and hence, it is worthwhile to give a proof. Although the proof in [32] is elementary and the proof in [13] is simple, our proof of Theorem 1.2 is also simple and elementary.

The paper is organized as follows. We first establish Theorems 1.1 and 1.2) in §2. Then we show Theorems 1.3 and 1.4 in §3. Finally, we present some remarks in §4.

2 Proof of Theorems 1.1 and 1.2

First of all, we have

Proposition 2.1 *Suppose that the inequalities in (5) hold for any symmetric matrices A and B of the same size. Then the inequalities hold for any matrices A and B of size $m \times n$.*

Proof. For any matrix A and B of size $m \times n$, we consider the $(m+n)\times(m+n)$ symmetric matrices

$$\tilde{A} = \begin{bmatrix} 0 & A \\ A^\top & 0 \end{bmatrix} \text{ and } \tilde{B} = \begin{bmatrix} 0 & B \\ B^\top & 0 \end{bmatrix}. \tag{12}$$

Then it is known that the eigenvalues of $\tilde{A}$ are $\pm\sigma_1(A), \cdots, \pm\sigma_\ell(A)$ together with $m+n-2\ell$ zeros, where $\ell = \min\{m,n\}$. Similarly for $\tilde{B}$. For $k \le \ell$, we use the assumption to have

$$\sum_{i=1}^k |\sigma_i^p(A) - \sigma_i^p(B)| = \sum_{i=1}^k |\sigma_i^p(\tilde{A}) - \sigma_i^p(\tilde{B})| \le \sum_{i=1}^k \sigma_i^p(\tilde{A} - \tilde{B}) = \sum_{i=1}^k \sigma_i^p(A-B).$$

This completes the proof. ■

Let us focus on A and B which are symmetric or Hermitian.

Theorem 2.1 *When A and B are symmetric matrices of the same size, Theorem 1.1 is true. That is, (5) holds for all symmetric matrices A and B of the same size.*

Proof of Theorem 1.1. Combining the results in the above Proposition 2.1 and Theorem 2.1 establishes the inequality in Theorem 1.1 for any rectangular matrices A and B of the same size $m \times n$. ■

Our major task in this section is to prove Theorem 2.1. To do so, we begin with the definition of $|A|$ for any square matrix A: $|A| = \sqrt{A^\top A}$. If $A = U\Sigma_A V^\top$ in SVD format, we have $|A| = V\sqrt{\Sigma_A^\top \Sigma_A} V^\top$. When A is symmetric, $A = C^\top \Lambda C$ for an orthonormal matrix C and diagonal matrix Λ which contains all eigenvalues of A, we have $|A| = C^\top |\Lambda| C$, where $|\Lambda|$ is the diagonal matrix obtained from Λ by replacing the diagonal by the absolute value of the diagonal. For any nonnegative function $f(t)$ defined on $[0,\infty)$, we can define $f(|A|) = C^\top f(|\Lambda|)C$. In [29] Thompson proved the following

Lemma 2.1 (Thompson, 1976 [29]) *For any real square matrices A and B of the same size, there exist orthonormal matrices U and V such that*

$$|A+B| \preceq U|A|U^T + V|B|V^\top, \tag{13}$$

namely $U|A|U^T + V|B|V^\top - |A+B|$ is positive semidefinite.

We remark that we do not have $|A+B| \le |A| + |B|$ in general. Furthermore, we have a more general version which is called Bourin–Uchiyama triangle inequality [[4], Corollary 2.6]. For self-containedness, we include a proof. We thank Simon Foucart for helpful discussion on the proof of Lemma 2.2 when he was associated with University of Georgia.

Lemma 2.2 *Let f be a nonnegative function defined on $[0, \infty)$ which is increasing and concave with $f(0) = 0$. For any matrices A and B of same size, there exists unitary matrices U and V of appropriate size such that*

$$f(|A|) \preceq Uf(|A - B|)U^\top + Vf(|B|)V^\top. \tag{14}$$

In particular, for $f(t) = t^p$ for $p \in (0, 1)$ and $t \geq 0$, we have

$$|A|^p \preceq U|A - B|^pU^\top + V|B|^pV^\top. \tag{15}$$

That is, we have

$$|A|^p - V|B|^pV^\top \preceq U|A - B|^pU^\top \tag{16}$$

for any symmetric matrices A, B of size $m \times m$.

Proof. For any matrices A, B of the same size, we use Lemma 2.1 to find unitary matrices U and V such that

$$|A| \preceq U|A - B|U^\top + V|B|V^\top.$$

Since $f(t)$ is nonnegative and nondecreasing, we have $f(|A|) \preceq f(U|A-B|U^\top + V|B|V^\top)$. Next we claim that

$$f(U|A - B|U^\top + V|B|V^\top) \preceq U_1f(U|A - B|U^\top)U_1^\top + V_1f(V|B|V^*)V_1^\top \tag{17}$$

for another two unitary matrices U_1 and V_1.

If the claim is true, as $f(U|A-B|U^\top) = Uf(|A-B|)U^\top$ and $f(V|B|V^\top) = Vf(|B|)V^\top$ as U and V are unitary, we obtain the desired inequality $f(|A|) \preceq U_1Uf(|A-B|)(U_1U)^\top + V_1Vf(|B|)(V_1V)^\top$.

We now prove the claim (17). For convenience, let us prove the following simplified version of the claim: if A and B are positive definite, then there exist unitary matrices U and V such that

$$f(A + B) \preceq Uf(A)U^\top + Vf(B)V^\top. \tag{18}$$

Without loss of generality, we may assume that $A + B > 0$. Let $X = (A + B)^{-1/2}A^{1/2}$ and $Y = (A + B)^{-1/2}B^{1/2}$. It is easy to see that $X^\top X = A^{1/2}(A + B)^{-1}A^{1/2} \leq I$ and similarly, $Y^\top Y \leq I$. Also, we note that $A = X^\top(A + B)X$ and $B = Y^\top(A + B)Y$. Next we need a few more facts:

- (1) For a concave function f with $f(0) = 0$, we have $f(\langle A\mathbf{x}, \mathbf{x}\rangle) \geq \langle f(A)\mathbf{x}, \mathbf{x}\rangle$ for nonnegative definite matrix A.
- (2) For a nonnegative definite matrix A, $f(C^\top AC) \geq C^\top f(A)C$ for any matrix C with $C^\top C \leq I$, where I is the identity matrix.
- (3) $\lambda_i(A^\top A) = \lambda_i(AA^\top)$ for any square matrix A.

Then we use the facts (2) and (3) to have

$$\begin{aligned}\lambda_i(f(A)) &= \lambda_i(f(X^\top(A+B)X)) \geq \lambda_i(X^\top f(A+B)X) \\ &= \lambda_i(X^\top f(A+B)^{1/2} f(A+B)^{1/2} X) = \lambda_i(f(A+B)^{1/2} X\, X^\top f(A+B)^{1/2})\end{aligned}$$

for all $i = 1, \cdots, n$ since A, B are symmetric.

By Lemma 2.3 below, there exists a unitary matrix U such that

$$Uf(A)U^* \geq f(A+B)^{1/2} X\, X^* f(A+B)^{1/2}.$$

Similarly, there exists a unitary matrix V such that

$$Vf(B)V^* \geq f(A+B)^{1/2} Y\, Y^* f(A+B)^{1/2}$$

Adding the two inequalities above, we have

$$Uf(A)U^* + Vf(B)V^* \geq f(A+B)^{1/2}(XX^* + YY^*)f(A+B)^{1/2} = f(A+B).$$

since $X\,X^* + Y\,Y^* = I$. This finishes a proof of the claim (17) and hence, we complete the proof. ■

Lemma 2.3 *If A and B are symmetric positive definite matrices and eigenvalues of A and B satisfy $\lambda_i(A) \leq \lambda_i(B)$ for all i, then there exists a matrix W such that $A \preceq WBW^\top$.*

Proof. See Thompson, 1976[29] for a proof. ■

Proof of Theorem 2.1. By Lemma 2.2, i.e. (16)

$$\sum_{i=1}^k \sigma_i(|A|^p - V|B|^p V^\top) \leq \sum_{i=1}^k \sigma_i(U|A-B|^p U^\top) = \sum_{i=1}^k \sigma_i(|A-B|^p) = \sum_{i=1}^k \sigma_i^p(A-B). \quad (19)$$

On the other hand, when both A and B are symmetric, the left-hand side of the inequality (5) can be calculated as follows:

$$\begin{aligned}\sum_{i=1}^k |\sigma_i^p(A) - \sigma_i^p(B)| &= \sum_{i=1}^k |\sigma_i^p(|A|) - \sigma_i^p(|B|)| = \sum_{i=1}^k |\sigma_i(|A|^p) - \sigma_i(|B|^p)| \\ &= \sum_{i=1}^k |\sigma_i(|A|^p) - \sigma_i(V|B|^p V^\top)| \leq \sum_{i=1}^k \sigma_i(|A|^p - V|B|^p V^\top), \quad (20)\end{aligned}$$

where we have used Theorem 1.5 in the last inequality. We now combine (19) and (20) to obtain the desired inequality (5). ■

Next we prove Theorem 1.2.

Proof of Theorem 1.2. For symmetric matrices A and B, we have

$$\begin{aligned}
&\sum_{i=1}^{k}|f(\sigma_i(A))-f(\sigma_i(B))| = \sum_{i=1}^{k}|f(\sigma_i(|A|))-f(\sigma_i(|B|))| \\
&= \sum_{i=1}^{k}|\sigma_i(f(|A|))-\sigma_i(f(|B|))| = \sum_{i=1}^{k}|\sigma_i(f(|A|))-\sigma_i(Vf(|B|)V^\top)| \\
&\le \sum_{i=1}^{k}\sigma_i(f(|A|)-Vf(|B|)V^\top) \le \sum_{i=1}^{k}\sigma_i(Uf(|A-B|)U^\top) \\
&= \sum_{i=1}^{k}\sigma_i(f(|A-B|)) = \sum_{i=1}^{k}f(\sigma_i(|A-B|)) = \sum_{i=1}^{k}f(\sigma_i(A-B)), \qquad (21)
\end{aligned}$$

where we have used Theorem 1.5 and Lemma 2.2. This completes the proof of Theorem 1.2 for the case when A and B are symmetric.

Next for any general matrices A and B of the same size, e.g. $m \times n$, let $\tilde{A}$ and $\tilde{B}$ be the symmetrization of A and B as in (12). Then for $k \le \min\{m, n\}$, we use (21) to have

$$\begin{aligned}
&\sum_{i=1}^{k}|f(\sigma_i(A))-f(\sigma_i(B))| = \sum_{i=1}^{k}|f(\sigma_i(\tilde{A}))-f(\sigma_i(\tilde{B}))| \le \sum_{i=1}^{k}f(\sigma_i(\tilde{A}-\tilde{B})) \\
&= \sum_{i=1}^{k}f(\sigma_i(A-B)).
\end{aligned}$$

Therefore, we have completed the proof of Theorem 1.2. ■

3 Proof of Theorems 1.3 and 1.4

To prove Theorem 1.3, we start with the following Lemma.

Lemma 3.1 *Let X^0 be the minimizer of (1) with rank s and X^* be a global minimizer of (2). Recall that both of them satisfy $\mathcal{A}(X^0) = \mathcal{A}(X^*)$. Let $H = X^0 - X^*$ which is in the null space of $\mathcal{A}$. Then*

$$\sum_{i=s+1}^{m}\sigma_i^p(H) \le \sum_{i=1}^{s}\sigma_i^p(H). \qquad (22)$$

Proof. We use (5) to have

$$\|X^0\|_p^p \ge \|X^*\|_p^p \;=\; \|X^0 - H\|_p^p = \sum_{i=1}^{m}\sigma_i^p(X^0-H) \ge \sum_{i=1}^{m}|\sigma_i^p(X^0)-\sigma_i^p(H)|$$

$$\geq \sum_{i=1}^{s}(\sigma_i^p(X^0) - \sigma_i^p(H)) + \sum_{i=s+1}^{m} \sigma_i^p(H)$$

since $\sigma_i(X^0) = 0$ for $i \geq s+1$. After rearranging the above inequality, since $\|X^0\|_p^p = \sum_{i=1}^s \sigma_i^p(X^0)$, we obtain the result. ■

Next we need an inequality which can be found in [11].

Lemma 3.2 *For any matrices X and Y with $\langle X, Y\rangle = trace(X^\top Y) = 0$ with rank$(X) \leq r$ and rank$(Y) \leq s$,*

$$\langle \mathcal{A}(X), \mathcal{A}(Y)\rangle \leq \delta_{r+s}\|X\|_F\|Y\|_F, \tag{23}$$

where $\|X\|_F$ stands for the Frobenius norm of X and similar for $\|Y\|_F$.

Proof of Theorem 1.3. Let $H = X^0 - X^*$ and write $H = U\Sigma_H V^*$ in SVD format. We decompose $\Sigma_H = \Sigma_0 + \Sigma_1 + \Sigma_2 + \cdots$ such that Σ_0 contains the first s singular values of Σ_H, Σ_1 the second s singular values of Σ_H, and so on. Similarly, write $U = [U_0\, U_1\, U_2\, \cdots]$ and $V = [V_0\, V_1\, V_2\, \cdots]$ accordingly. Thus, $H = H_0 + H_1 + H_2 + \cdots$ with $H_i = U_i\Sigma_i V_i^\top$. Clearly, $\langle H_i, H_j\rangle = 0$ if $i \neq j$. Thus, we can use the definition of matrix version RIP to have

$$\begin{aligned}
& (1-\delta_{2s})(\|H_0\|_F^2 + \|H_1\|_F^2) \\
\leq\ & \|\mathcal{A}(H_0+H_1)\|_2^2 = \|\mathcal{A}(H - \sum_{j\geq 2} H_j)\|_2^2 \\
=\ & \sum_{i,j\geq 2}\langle \mathcal{A}(H_j), \mathcal{A}(H_i)\rangle \leq \sum_{i\geq 2}(1+\delta_{2s})\|H_i\|_F^2 + \sum_{\substack{i,j\geq 2\\ i\neq j}}\langle \mathcal{A}(H_i), \mathcal{A}(H_j)\rangle \\
\leq\ & \sum_{i\geq 2}(1+\delta_{2s})\|H_i\|_F^2 + \delta_{2s}\sum_{\substack{i,j\geq 2\\ i\neq j}}\|H_i\|_F\|H_j\|_F \\
\leq\ & \sum_{i\geq 2}\|H_i\|_F^2 + \delta_{2s}(\sum_{j\geq 2}\|H_j\|_F)^2 \\
\leq\ & (1+\delta_{2s})\left(\sum_{j\geq 2}\|H_j\|_F\right)^2.
\end{aligned}$$

If $\|H_2\|_F = 0$, the right-hand side of the inequality above is zero and hence $H = 0$. Hence $X^* = X^0$. That is, the minimizer is the low rank solution of (1).

Otherwise, the above inequality can be rewritten as follows:

$$\|H_0\|_F \leq \left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2}\left(\sum_{j\geq 2}\|H_j\|_F\right) \tag{24}$$

under the assumption that $\|H_2\|_F \neq 0$. In this situation, $\sigma_{2s+1}(H) \neq 0$. We have

$$(2s)^{-1/p}\|H_0 + H_1 + H_2\|_p = (2s)^{-1/p}\sigma_{2s+1}(H)\left(\sum_{i=1}^{3s}\left(\frac{\sigma_i(H)}{\sigma_{2s+1}(H)}\right)^p\right)^{1/p}$$

$$
\begin{aligned}
&\geq \sigma_{2s+1}(H)(2s)^{-1/p}\left(\sum_{i=1}^{2s+1}\left(\frac{\sigma_i(H)}{\sigma_{2s+1}(H)}\right)^p\right)^{1/p} \\
&\geq \sigma_{2s+1}(H)(2s)^{-1/p}(2s+1)^{1/p} \\
&= \sigma_{2s+1}(H)(1+\frac{1}{2s})^{1/p}. \qquad (25)
\end{aligned}
$$

On the other hand, we have

$$\sum_{j\geq 2}\|H_j\|_F \leq \frac{n}{s}\|H_2\|_F \leq \frac{n}{s}\sqrt{s}\;\sigma_{2s+1}(H) = \frac{n}{\sqrt{s}}\;\sigma_{2s+1}(H). \qquad (26)$$

Now we claim that for any $\epsilon \in (0,1)$, there exists a real number $p_\epsilon \leq 1$ such that when $p \leq p_\epsilon$, we have

$$(2s)^{1/p}\left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2}\sum_{j\geq 2}\|H_j\|_F \leq \epsilon\|H_0+H_1+H_2\|_p. \qquad (27)$$

For convenience, let

$$C_s = \left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2}.$$

We consider the following function $f(p,H)$, using (26) and (25),

$$
\begin{aligned}
f(p,H) &= \left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2}\frac{(2s)^{1/p}\sum_{j\geq 2}\|H_j\|_F}{\|H_0+H_1+H_2\|_p} \\
&\leq C_s\frac{(2s)^{1/p}\frac{n}{\sqrt{s}}\sigma_{2s+1}(H)}{\sigma_{2s+1}(H)(2s+1)^{1/p}} \leq \frac{C_s\frac{n}{\sqrt{s}}}{(1+\frac{1}{2s})^{1/p}}.
\end{aligned}
$$

For $0<\epsilon<1$, since $(1+\dfrac{1}{2s})^{1/p}\to\infty$ when $p\to 0_+$, for any $H\in\mathbb{R}^{m\times n}$,

$$f(p,H)\leq\epsilon.$$

That is, we have (27).

Finally, we use the Hölder inequality, the inequality in (24), and then the inequality in (27) to have

$$
\begin{aligned}
\|H_0\|_p^p &\leq s^{1-p/2}\|H_0\|_F^p \leq s^{1-p/2}\left(\epsilon(2s)^{-1/p}\|H_0+H_1+H_2\|_p\right)^p \\
&\leq \epsilon^p s^{-p/2}2^{-1}\|H_0+H_1+H_2\|_p^p \leq \frac{\epsilon^p}{2s^{p/2}}\left(\|H_0\|_p^p+\sum_{i\geq s+1}\sigma_i^p(H)\right) \\
&\leq \frac{\epsilon^p}{2s^{p/2}}2\|H_0\|_p^p = \frac{\epsilon^p}{s^{p/2}}\|H_0\|_p^p
\end{aligned}
$$

by using Lemma 3.1.

Under the assumption that we used p minimization (2) with $p \leq p_\epsilon$ and $\epsilon < \sqrt{s}$, we get a contradiction if $\|H_0\|_F \neq 0$ since $t = \dfrac{\epsilon^p}{s^{p/2}} < 1$. Thus, $\|H_0\|_F = 0$. Therefore, $H = X^0 - X^* = 0$ and hence, the minimizer of (2) is the minimal rank solution of (1). ■

We now study the ℓ_q minimization in the noisy matrix recovery setting.

Proof of Theorem 1.4. Let $H = X^0 - X^*$. Suppose that $\|\mathcal{A}(H)\|_2 \leq 2\eta$. We decompose $H = H_0 + H_1 + H_2 + ...$ similarly as in the proof of Theorem 1.3. Using the definition of matrix version RIP we have

$$\begin{aligned} & (1-\delta_{2s})(\|H_0\|_F^2 + \|H_1\|_F^2) & (28)\\ \leq\ & \|\mathcal{A}(H_0+H_1)\|_2^2 = \|\mathcal{A}(H - \sum_{j\geq 2} H_j)\|_2^2 \leq (2\eta + \|\sum_{j\geq 2}\mathcal{A}(H_j)\|_2)^2. & (29)\end{aligned}$$

It follows that

$$\|H_0\|_F \leq \frac{2\eta + \|\sum_{j\geq 2}\mathcal{A}(H_j)\|_2}{\sqrt{1-\delta_{2s}}}. \tag{30}$$

Then by the same argument as (24), we have

$$\|\sum_{j\geq 2}\mathcal{A}(H_j)\|_2^2 \leq (1+\delta_{2s})(\sum_{j\geq 2}\|H_j\|_F)^2.$$

Thus, we have

$$\|H_0\|_F \leq \frac{2\eta + \sqrt{1+\delta_{2s}}\sum_{j\geq 2}\|H_j\|_F}{\sqrt{1-\delta_{2s}}}. \tag{31}$$

If $\sigma_{2s+1}(H) \neq 0$, then by (27), for any $0 < \epsilon < 1$, there exists some $0 < p_0 < 1$, such that for any $0 < p \leq p_0$,

$$\left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2} \frac{(2s)^{1/p}\sum_{j\geq 2}\|H_j\|_F}{\|H_0+H_1+H_2\|_p} \leq \epsilon.$$

Using Hölder inequality, we have

$$\|H_0\|_p^p \leq s^{1-p/2}\|H_0\|_F^p \leq s^{1-p/2}\left(\frac{2\eta + \sqrt{1+\delta_{2s}}\sum_{j\geq 2}\|H_j\|_F}{\sqrt{1-\delta_{2s}}}\right)^p,$$

that is,

$$\begin{aligned} \|H_0\|_p \ &\leq\ s^{1/p-1/2}\left(\frac{2\eta}{\sqrt{1-\delta_{2s}}} + \left(\frac{1+\delta_{2s}}{1-\delta_{2s}}\right)^{1/2}\sum_{j\geq 2}\|H_j\|_F\right) & (32)\\ &\leq\ s^{1/p-1/2}\left(\frac{2\eta}{\sqrt{1-\delta_{2s}}} + \epsilon(2s)^{-1/p}\|H_0+H_1+H_2\|_p\right). & (33)\end{aligned}$$

By Lemma 3.1, we have

$$\|H_0 + H_1 + H_2\|_p^p \leq \|H\|_p^p = \|H_0\|_p^p + \sum_{i \geq s+1} \sigma_i^p(H) \leq 2\|H_0\|_p^p. \tag{34}$$

Substitute the above inequality into (32), we have

$$\|H_0\|_p \leq \frac{s^{1/p-1/2}2\eta}{\sqrt{1-\delta_{2s}}} + \frac{\epsilon\|H_0\|_p}{s^{1/2}},$$

Then by direct computation,

$$\|H_0\|_p \leq \frac{s^{1/p-1/2}2\eta}{\sqrt{1-\delta_{2s}}(1-\frac{\epsilon}{s^{1/2}})}$$

Now we use (34) to evaluate the Schatten p quasi norm of H,

$$\|H\|_p \leq \frac{2^{1/p}s^{1/p-1/2}2\eta}{\sqrt{1-\delta_{2s}}(1-\frac{\epsilon}{s^{1/p}})}.$$

Letting $\epsilon = 1$, we obtain the result (8) of Theorem 1.4.

If $\delta_{2s+1} = 0$, $H = H_0 + H_1$, by (31), we have

$$\|H_0\|_F \leq \frac{2\eta + \sqrt{1+\delta_{2s}}\sum_{j\geq 2}\|H_j\|_F}{\sqrt{1-\delta_{2s}}} = \frac{2\eta}{\sqrt{1-\delta_{2s}}}. \tag{35}$$

In terms of Schatten p-norm, we have

$$\|H_0\|_p^p \leq s^{1-p/2}\|H_0\|_F^p \leq s^{1-p/2}(\frac{2\eta}{\sqrt{1-\delta_{2s}}})^p. \tag{36}$$

Hence, by (34)

$$\|H\|_p^p \leq 2\|H_0\|_p^p \leq 2s^{1-p/2}(\frac{2\eta}{\sqrt{1-\delta_{2s}}})^p \text{ or } \|H\|_p \leq 2^{1/p}s^{1/p-1/2}(\frac{2\eta}{\sqrt{1-\delta_{2s}}}).$$

These complete the proof. ■

4 Remarks

We have a few remarks in order.

Remark 4.1 *In the vector case, we have the following inequality*

$$\sum_{i=1}^{n}|x_i - y_i|^p \geq \sum_{i=1}^{n}(|x_i|^p - |y_i|^p).$$

So Theorem 1.3 is true in the compressed sensing setting. Thus, we recover a known result in [28], i.e. Theorem 1.2. Our proof is much simpler.

Remark 4.2 (Counter Example to a Result in [2]) *Audenaert posted a paper on arxiv, which has not been published in any journal though, which uses the Thompson–Freede inequality in matrix analysis to show the inequality in Theorem 1 of [2]. However, a counter example to Theorem 1 in [2] is the following: Let*

$$A = \begin{pmatrix} 0.2068 & 0.5829 & 0.6446 & 0.2887 & 0.8949 \\ 0.6631 & 0.8282 & 0.0823 & 0.3543 & 0.8262 \\ 0.3367 & 0.7249 & 0.7107 & 0.2332 & 0.7953 \\ 0.1886 & 0.3429 & 0.7456 & 0.3608 & 0.8450 \\ 0.9713 & 0.8871 & 0.3066 & 0.2953 & 0.7165 \end{pmatrix} \tag{37}$$

and

$$B = \begin{pmatrix} 0.0476 & 0.2111 & 0.8173 & 0.9058 & 0.8816 \\ 0.8424 & 0.2926 & 0.4545 & 0.3807 & 0.0241 \\ 0.1914 & 0.5593 & 0.9373 & 0.9929 & 0.0159 \\ 0.1988 & 0.5522 & 0.7141 & 0.2116 & 0.1705 \\ 0.6147 & 0.6757 & 0.1920 & 0.6496 & 0.8566 \end{pmatrix}. \tag{38}$$

Let $i_1 = 3$ and $i_2 = 5$, which form an increasing sequence of integers, and let $f(x) = \sqrt{x}$, which is concave with $f(0) = 0$. They satisfy all the conditions of Theorem 1 in [2], but the inequality in the conclusion of Theorem 1 in [2] does not hold for theses matrices.

References

[1] T. Ando. Comparison of norms $||f(a) - f(b)||$ and $||f(|a-b|)||$. Mathematische Zeitschrift, 197(3):403–409, 1998.

[2] K. Audenaert. A generalisation of Mirsky's singular value inequalities. arXiv preprint arXiv:1410.4941, 2014.

[3] K. Audenaert and F. Kittaneh. Problems and conjectures in matrix and operator inequalities. Operator theory, (Banach Center Publications), 2012.

[4] J.-C. Bourin and M. Uchiyama. A matrix subadditivity inequality for $f(A+B)$ and $f(A)+f(B)$. Linear Algebra and its Applications 423.2, 512–518, 2007.

[5] R. Bhatia. Matrix analysis, Springer Verlag, 1997.

[6] T. T. Cai and Zhang, A., Sharp RIP bound for sparse signal and low-rank matrix recovery. Appl. Comput. Harmon. Anal. 35(2013), 74–93.

[7] T. T. Cai and Zhang, A., Compressed sensing and affine rank minimization under restricted isometry. IEEE Transactions on Signal Processing 61(2013), 3279–3290.

[8] T. T. Cai and Zhang, A., Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. IEEE Transactions on Information Theory 60(2014), 122–132.

[9] E. J. Candès, The restricted isometry property and its implications for compressed sensing, comptes Rendus de Academie des Sciences, Series I, 346(2008), 589–592.

[10] K. Dvijotham and M. Fazel. A nullspace analysis of the nuclear norm heuristic for rank minimization. In Proc. of ICASSP, 2010.

[11] E.J. Candès and Y. Plan, Tight Oracle Inequalities for Low-Rank Matrix Recovery From a Minimal Number of Noisy Random Measurements, IEEE Transactions on Information Theory, 57(2011), 2342 - 2359.

[12] M. Fornasier, H. Rauhut, and R. Ward, Low-rank matrix recovery via iteratively reweighted least squares minimization, SIAM J. Optim. 21 (2011), no. 4, 1614–1640.

[13] S. Foucart, Concave Mirsky inequality and low-rank recovery. SIAM Journal on Matrix Analysis and Applications, 39(2018), 99–103.

[14] S. Foucart, M.J. Lai, Sparsest solutions of under-determined linear systems via ℓ_q-minimization for $0 < q < 1$, Appl. Comput. Harmon. Anal. 26 (2009) 395–407.

[15] R. Gribonval and M. Nielsen, Sparse decompositions in unions of bases. IEEE Trans. Info. Theory, 49(2003), pp. 3320–3325.

[16] R. Horn and C. Johnson, Matrix Analysis, Cambridge Univ. Press, 1985.

[17] R. Horn and C. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, 1991.

[18] T. Kato. Perturbation theory for linear operators, Springer Verlag, 1995.

[19] L. Kong, and N. Xiu, New Bounds for Restricted Isometry Constants in Low-rank Matrix Recovery, Optimization on-line Digest, 2011.

[20] Lai, M. -J., Xu, Y. Y. and Yin, W. T., Improved Iteratively Reweighted Least Squares for Unconstrained Smoothed p Minimization, SIAM Journal on Numerical Analysis, vol. 51 (2013) pp. 927–957.

[21] V. B. Lidskii, C.D. Benster, and G.E. Forsythe, The proper values of the sum and product of symmetric matrices, United States Office of Naval Research, 1953.

[22] K. Mohan, M. Fazel, New Restricted Isometry results for noisy low-rank recovery. ISIT, Austin, 2010.

[23] S. Oymak and B. Hassibi. New null space results and recovery thresholds for matrix rank minimization. 2010. Available on-line.

[24] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi. A simplified approach to recovery conditions for low rank matrices. Proc. Intl. Sympo. Information Theory (ISIT), 2011.

[25] B. Recht, M. Fazel, and P. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Review, 52(2010), 471–501.

[26] S. Yu. Rotfel'd, Remarks on the singular numbers of a sum of completely continuous operators, Functional Anal. Appl., 1 (1967), 252–253.

[27] S. Yu. Rotfel'd, The singular numbers of the sum of completely continuous operators, Topics in Mathematical Physics, vol. 3, Spectral Theory, edited by M. S. Berman, (1969), 73–78. Consultants Bureau, New York.

[28] Q. Sun, Recovery of sparsest signals via ℓ^q minimization, Appl. Comput. Harmon. Anal., 32(2012), 329–341.

[29] R. C. Thompson, Convex and concave functions of singular values of matrix sums. Pacific Journal of Mathematics, 66(1):285–290, 1976.

[30] M. Uchiyama, Subadditivy of eigenvalue sums. Proc. American Mathematical Society, (2005) 1405–1412.

[31] H. Wang and S. Li, The bounds of restricted isometry constants for low rank matrices recovery, Science China Mathematics, 56(2013) 1117 –1127.

[32] Man-Chung Yue, Anthony Man-Cho So, A perturbation inequality for concave functions of singular values and its applications in low-rank matrix recovery, Appl. Comput. Harmon. Anal., Vol. 40(2016), 396–416.

[33] R. Zhang and S. Li, A Proof of Conjecture on Restricted Isometry Property Constants δ_{tk} $(0 < t < 4/3)$, IEEE Transactions on Information Theory, 64(2018), 1699–1705.

[34] R. Zhang and S. Li, Optimal RIP bounds for sparse signals recovery via ℓ_p minimization, Appl. Comput. Harmon. Anal. 47 (2019), 566–584.

5 Appendix

5.1 Proof of Theorem 1.5

Let us begin with

Theorem 5.1 *Let A and B be matrices of the same size $m \times n$. Fix $p \in [0, 1]$. Write $\sigma_j(A)$ and $\sigma_j(B), j = 1, \cdots, n$ be singular values of A and B, respectively. For each k, we define*

$$\alpha = \max_{k=1,\cdots,n} \frac{\max_{j=1,\cdots,k} |\sigma_j(A - B)|}{\min_{j=1,\cdots,k} |\sigma_j(A) - \sigma_j(B)| \neq 0} \geq 1. \tag{39}$$

Then we have

$$\sum_{i=1}^{k} |\sigma_i^p(A) - \sigma_i^p(B)| \leq \alpha^{1-p} \sum_{i=1}^{k} \sigma_i^p(A-B) \tag{40}$$

for $k = 1, 2, \cdots, m$.

Proof of Theorem 1.5. In particular, when $p = 1$, we have Theorem 1.5 from (40). ■

Proof of Theorem 5.1. Let $\beta_k = \min_{j=1,\cdots,k} |\sigma_j(A) - \sigma_j(B)| \neq 0$. ALso, let $\gamma_k = \max_{j=1,\cdots,k} |\sigma_j(A-B)|$. Then for $p \in (0,1)$, we have

$$\begin{aligned}
&\sum_{j=1}^{k} |\sigma_j^p(A) - \sigma_j^p(B)| \leq \sum_{j=1}^{k} |\sigma_j(A) - \sigma_j(B)|^p = \beta_k^p \sum_{j=1}^{k} \left| \frac{|\sigma_j(A) - \sigma_j(B)|}{\beta_k} \right|^p \\
&\leq \beta_k^p \sum_{j=1}^{k} \left| \frac{|\sigma_j(A) - \sigma_j(B)|}{\beta_k} \right| \leq \beta_k^{p-1} \sum_{j=1}^{k} \sigma_j(A-B) = \beta_k^{p-1} \gamma_k \sum_{j=1}^{k} \frac{\sigma_j(A-B)}{\gamma_k} \\
&\leq \beta_k^{p-1} \gamma_k \sum_{j=1}^{k} \left| \frac{\sigma_j(A-B)}{\gamma_k} \right|^p \leq \beta_k^{p-1} \gamma_k^{1-p} \sum_{j=1}^{k} \sigma_j^p(A-B) \leq \alpha^{1-p} \sum_{j=1}^{k} \sigma_j^p(A-B).
\end{aligned}$$

This completes the proof. ■