

A Perturbation Analysis of Low-Rank Matrix Recovery by Schatten p-Minimization

Zhaoying Sun, Huimin Wang*, Zhihui Zhu

Department of Mathematics, Shaoxing University, Shaoxing, China

Email: *gfanme@126.com

How to cite this paper: Sun, Z.Y., Wang, H.M. and Zhu, Z.H. (2024) A Perturbation Analysis of Low-Rank Matrix Recovery by Schatten p-Minimization. *Journal of Applied Mathematics and Physics*, 12, 475-487.
<https://doi.org/10.4236/jamp.2024.122032>

Received: January 2, 2024

Accepted: February 26, 2024

Published: February 29, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A number of previous papers have studied the problem of recovering low-rank matrices with noise, further combining the noisy and perturbed cases, we propose a nonconvex Schatten p-norm minimization method to deal with the recovery of fully perturbed low-rank matrices. By utilizing the p-null space property (p-NSP) and the p-restricted isometry property (p-RIP) of the matrix, sufficient conditions to ensure that the stable and accurate reconstruction for low-rank matrix in the case of full perturbation are derived, and two upper bound recovery error estimation ns are given. These estimations are characterized by two vital aspects, one involving the best r-approximation error and the other concerning the overall noise. Specifically, this paper obtains two new error upper bounds based on the fact that p-RIP and p-NSP are able to recover accurately and stably low-rank matrix, and to some extent improve the conditions corresponding to RIP.

Keywords

Nonconvex Schatten p-Norm, Low-Rank Matrix Recovery, p-Null Space Property, The Restricted Isometry Property

1. Introduction

Low-rank matrix recovery (LMR) is a fast-growing field nowadays, attracting much attention in numerous applications such as quantum state tomography scanners [1], deep learning [2], nonlinear system identification [3], computer visualization [4], and medical imaging [5]. The mathematical expression of the low-rank matrix recovery issue is described as

$$b = \mathcal{A}(X) \quad (1)$$

where $X \in \mathbb{R}^{m_1 \times m_2}$ is an unknown low-rank matrix or approximate low-rank

matrix, which need be recovered and $b \in \mathbb{R}^N$ is a known observation vector, $\mathcal{A} : \mathbb{R}^{m_1 \times m_2} \rightarrow \mathbb{R}^N$ is a given measurement operator or linear mapping which is defined by the following formula

$$\mathcal{A}(X) = \left[\text{tr}\left(X^T B^{(1)}\right), \text{tr}\left(X^T B^{(2)}\right), \dots, \text{tr}\left(X^T B^{(N)}\right) \right]^T \tag{2}$$

where $B^{(1)}, B^{(2)}, \dots, B^{(N)}$ are named the measurement matrices, X^T is the transpose matrix of X and $\text{tr}(\cdot)$ is the trace function. The main goal of LMR is to recover the low-rank matrix X on the basis of observation vector b and operator \mathcal{A} .

Actually, the linear measurement b is affected by the noisy vector y . The noisy LMR model is shown by

$$\hat{b} = \mathcal{A}(X) + y, \tag{3}$$

\hat{b} is an observed measurement disturbed by the noise vector y , y is an additive noise that does not depend on X .

Moreover, more LMR models may involve case in which the observed vector b is perturbed noise vector, meanwhile, the linear mapping \mathcal{A} is hampered by Φ , i.e., \mathcal{A} is replaced by $\hat{\mathcal{A}} = \mathcal{A} + \Phi$ to bring about multiplicative noise $\Phi(X)$ associated with X . In a large number of applications of complete separation problems such as remote sensing [6], source separation [7], and telecommunication [8], complete perturbation problems usually arise. In order to require the optimal solution from this fully perturbed problems, a common approach is to solve a class of nuclear norm minimization issues (NNM) as described below:

$$\min_{X \in \mathbb{R}^{m_1 \times m_2}} \|X\|_* \quad \text{s.t.} \quad \|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{4}$$

where $\hat{\epsilon}_{\mathcal{A},r,b} \geq 0$ represents the overall noise, $\|X\|_*$ is the trace norm of the matrix X , namely the sum of its singular values. While $m_1 = m_2$ and $X = \text{diag}(x) (x \in \mathbb{R}^m)$ is diagonal matrix, issue (4) relegates into a compressed sensing issue

$$\min_{x \in \mathbb{R}^m} \|x\|_1 \quad \text{s.t.} \quad \|\hat{\mathcal{A}}x - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{5}$$

here $\|x\|_1$ is l_1 the norm of the vector x , in a word, the sum of the absolute values of the elements of x .

Chartrand's study [9] revealed that nonconvex variants of (5) can produce accurate reconstruction with fewer measurements. Specifically, the l_1 norm minimization is substituted with the l_p norm minimization:

$$\min_{x \in \mathbb{R}^m} \|x\|_p^p \quad \text{s.t.} \quad \|\hat{\mathcal{A}}x - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{6}$$

in which $\|x\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$ is l_p -quasi-norm of vector x . Even though the l_p -quasi-norm is not a norm, but it satisfies the triangle inequality.

Numerous studies have focused on the recovery of vector x by l_p minimization ($0 < p \leq 1$) [10] [11] [12]. Chartrand [9] conducts numerical experiments using random and non-random Fourier measurements, which acquired fewer

measurements are required for accurate restoration than when $p = 1$. Chartrand [13] extended the result of Restricted Isometry Property (RIP) in Candès and Tao [14] to the case of l_p minimization. Kong and Xiu [15] explored nonconvex relaxation methods for recovering the vector x . In summary, the case of (6) where x is free of the noise and perturbation for $(\hat{\epsilon}_{\mathcal{A},r,b} = 0)$ extends to the matrix, referred to as M_p -minimization:

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \|X\|_p^p \quad \text{s.t. } \mathcal{A}(X) = b. \tag{7}$$

The related work [13] considers the scenario of matrix recovery with noise but without perturbation, *i.e.*, where the linear mapping \mathcal{A} is not interfered with by Φ . From an applied and practical perspective, it is crucial to investigate the problem of rank- r matrix recovery in the fully perturbed case.

Therefore, we present the fully perturbed model of LMR through nonconvex Schatten p -norm minimization ($0 < p \leq 1$):

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \|X\|_p^p \quad \text{s.t. } \|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{8}$$

a p -norm of matrix X , and $\|X\|_p = \left(\sum_i \sigma_i^p(X)\right)^{\frac{1}{p}}$ ($0 < p \leq 1$), its singular value decomposition (SVD) is $X = U \text{diag}(\sigma(X)) V^T$, $\sigma_i(X)$ being the i th singular value of X . This model characterizes the problem of rank- r matrix recovery in a fully perturbed scenario, with $\epsilon_{\mathcal{A}}^{(r)}$, $\eta_{\mathcal{A}}^{(r)}$, $\kappa_{\mathcal{A}}$, α_r , β_r , $\epsilon_{\mathcal{A}}$, and ϵ_y as parameters in the model.

Now, unlike the previous concept of restricted isometry constant, this paper follows the notion of restricted isometry property (p-RIP) given by Zhang in the article [13], which is viewed below:

Definition 1.1 (p-RIP of the measurement mapping (or operator) \mathcal{A}) For the measurement operator $\mathcal{A}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^M$, a positive integer r and $0 < p \leq 1$, the restricted p -isometry constant (p-RIC) of order r denoted by δ_r and for any matrix $X \in \mathbb{R}^{m \times n}$ and $R(X) \leq r$, which has

$$(1 - \delta_r) \|X\|_F^p \leq \|\mathcal{A}(X)\|_p^p \leq (1 + \delta_r) \|X\|_F^p. \tag{9}$$

If $\delta \in (0, 1)$, then \mathcal{A} meets the restricted p -isometry property of order r .

The restricted isometry property (RIP) of matrix is an essential tool for LMR theory analysis. For the accurate LMR (*i.e.*, $y = 0$, $\Phi = 0$ in (1)) or a noisy/partly perturbed LMR (*i.e.*, $\Phi = 0$ in (7)), there are some sufficient conditions based on RIP, which include $\delta_k + \theta \delta_{2r+k} < \theta - 1$, when $\theta > 1$ and

$$k = \left\lceil 2r\theta^{\frac{2}{2-p}} \right\rceil \left(0 < p \leq 1\right) \text{ [16]}, \quad \sqrt{2}\delta \max \left\{ r + \left\lceil \frac{3k}{2} \right\rceil, 2k \right\} \frac{k^{\frac{1}{p}-\frac{1}{2}}}{2r} \delta_{2r+k} < \frac{k^{\frac{1}{p}-\frac{1}{2}}}{2r} \text{ [17]}.$$

Moreover, another crucial tool for analyzing low-rank matrix recovery is the null space property (NSP) of the linear mapping \mathcal{A} . Gao and Peng *et al.* [18] extended the general null space property to obtain the sparse vector p -null space property (p-NSP). Furthermore, we further extend sparse vector x recovery to the low-rank matrix, the notion is defined like this:

Definition 1.2 (p-NSP of measurement operator \mathcal{A}) For the measured operator $\mathcal{A}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^M$, with a constant $0 < s < 1$ and $\nu > 0$, for any matrix $X \in \mathbb{R}^{m \times n}$ and $\text{rank}(X) \leq r$, there exist

$$\|X_r\|_p^p \leq s \|X_r^c\|_p^p + \nu \|\mathcal{A}(X)\|_p^p, \tag{10}$$

so the operator \mathcal{A} fulfills the p-null space property of order r .

2. Symbols and Main Results

Before presenting the key results of this paper, some notations similar to those of the article [19] which quantize disturbances Φ and γ with different upper bounds are given like such:

$$\frac{\|\Phi\|_p}{\|\mathcal{A}\|_p} \leq \epsilon_{\mathcal{A}}, \frac{\|\Phi\|_p^{(r)}}{\|\mathcal{A}\|_p^{(r)}} \leq \epsilon_{\mathcal{A}}^r, \frac{\|\Phi\|_2}{\|\mathcal{A}\|_2} \leq \epsilon_b, \tag{11}$$

Here $\|\mathcal{A}\|_p = \sup \left\{ \frac{\|\mathcal{A}(X)\|_p}{\|X\|_F}, X \in \mathbb{R}^{m_1 \times m_2} \text{ with } X \neq 0 \right\}$ is the Schatten p-norm

of the linear mapping \mathcal{A} and

$\|\mathcal{A}\|_p^{(r)} = \sup \left\{ \frac{\|\mathcal{A}(X)\|_p}{\|X\|_F}, X \in \mathbb{R}^{m_1 \times m_2}, X \neq 0, R(X) = r \right\}$, $\|\mathcal{A}\|_p^{(r)}$ is the norm of its

initial image set of \mathcal{A} consisting of nonzero matrix of order r , furthermore

$$\alpha_r = \frac{\|X_r^c\|_F}{\|X_r\|_F}, \beta_r = \frac{\|X_r^c\|_p}{r^{\frac{1}{p}-\frac{1}{2}} \|X_r\|_F}, \eta_{\mathcal{A}^{(r)}} = \sqrt[p]{\frac{1+\delta_r}{1-\delta_r}}, \kappa_{\mathcal{A}} = \frac{\|\mathcal{A}\|_p}{\sqrt[p]{1-\delta_r}}, \tag{12}$$

$X_r^c = X - X_r$, X_r stands for the best r -rank approximation of X whose singular values consist of the r -largest singular values of X .

Secondly, two theorems are obtained based on the matrix of restricted p-isometry property (p-RIP) (see Definition 1.1) and the p-null space property (p-NSP) (see Definition 1.2) defined in section 1. Two theorems derive sufficient conditions guaranteeing stable and accurate recovery of rank r matrices, and offer recovery error upper bounds. The content of two theorems is described as below:

Theorem 2.1: Let there be a linear operator $\mathcal{A}: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^M$, $\hat{y} \in \mathbb{R}^M$ and $0 < p \leq 1$. Let $t > 1$ and $k = 2tr$, if the restricted p-isometry constant of the linear operator \mathcal{A} fulfills

$$\left(1 + \delta_{2r(t+1)}\right) \left(1 + \epsilon_{\mathcal{A}}^{2r(t+1)}\right)^p + t^{\frac{p}{2}-1} \delta_{2tr} \left(1 + \epsilon_{\mathcal{A}}^{2tr}\right)^p < 2 - t^{\frac{p}{2}-1} \left(1 + \epsilon_{\mathcal{A}}^{2tr}\right)^p, \tag{13}$$

a general matrix X with r -rank meets

$$\alpha_r + \beta_r < \frac{1}{\eta_{\mathcal{A}}^{(r)}}, \tag{14}$$

and the overall noise is

$$\hat{\epsilon}_{\mathcal{A},r,b} = \left(\frac{\epsilon_{\mathcal{A}}^{(r)} \eta_{\mathcal{A}}^{(r)} + \epsilon_{\mathcal{A}} \kappa_{\mathcal{A}} \alpha_r}{1 - \eta_{\mathcal{A}}^{(r)} (\alpha_r + \beta_r)} + \epsilon_y \right) \|b\|_2. \tag{15}$$

$X \in \mathbb{R}^{m \times n}$, $rank(X) \leq r$, $\hat{\mathcal{A}} = \mathcal{A} + \Phi$ and $\|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}$. X^* is the feasible solution of the fully perturbed Schatten p -norm minimization problem

$$\min_{X \in \mathbb{R}^{m_1 \times m_2}} \|X\|_p^p \quad \text{s.t.} \quad \|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{16}$$

The error estimation of X and X^* fulfills

$$\|X - X^*\|_p^p \leq C (\hat{\epsilon}_{\mathcal{A},r,b})^p + D \|X_r^c\|_p^p, \tag{17}$$

where

$$C = \frac{2^{p+1} (2rM)^{1-\frac{p}{2}}}{2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}}},$$

$$D = 2 + \frac{4(1 + \delta_{2tr}) t^{\frac{p-1}{2}} (1 + \epsilon_{\mathcal{A}}^{2tr})^p}{2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}}}.$$

Theorem 2.2: For a given $\epsilon_{\mathcal{A}}$, supposing that the linear measured operator \mathcal{A} fulfills the p -null space property (p -NSP) with constants $0 < s < 1 - 2\tau \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p$ and $0 < v < \frac{1}{\epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}$ and conditions of (13) and (14) hold. Suppose $X \in \mathbb{R}^{m \times n}$,

$rank(X) \leq r$, $\hat{\mathcal{A}} = \mathcal{A} + \Phi$, $\|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}$ holds. X^* is the feasible solution to the completely perturbed Schatten p -norm minimization problem

$$\min_{X \in \mathbb{R}^{m_1 \times m_2}} \|X\|_p^p \quad \text{s.t.} \quad \|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}, \tag{18}$$

then error estimation of X and X^* satisfies

$$\|X - X^*\|_p^p \leq C' (\hat{\epsilon}_{\mathcal{A},r,b})^p + D' \|X_r^c\|_p^p, \tag{19}$$

where

$$C' = \frac{vM^{1-\frac{p}{2}} 2^{p+1}}{1 - s - 2v\epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}, \quad D' = \frac{2(1+s)}{1 - s - 2v\epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}.$$

3. Proof of Key Results

Proofs of our key results are presented in this section. To prove theorem 2.1 and theorem 2.2 such that we need the support of the following five lemmas and their proofs. We start this section with a lemma with respect to Schatten p norm.

Lemma 3.1: If $0 < p \leq 1$. Presume that $D, E \in \mathbb{R}^{m \times n}$ are matrices with $D^T E = 0$ and $DE^T = 0$. Then

$$1) \|D + E\|_p^p = \|D\|_p^p + \|E\|_p^p; \quad 2) \|D + E\|_p \geq \|D\|_p + \|E\|_p.$$

When $p = 1$, $\|D\|_p^p$ and $\|D\|_p$ are the trace(or nuclear) norm of matrix D .

Lemma 3.2: For any $p \in (0, 1]$, there is

$$\|Z_r^c\|_p^p \leq \|Z_r\|_p^p + 2\|X_r^c\|_p^p. \tag{20}$$

Proof of the lemma 3.2 suppose $Z = X - X^*$ and X^* is the optimal solution to the problem (8), which yields

$$\|X\|_p^p \geq \|X^*\|_p^p = \|X - Z\|_p^p. \tag{21}$$

Applying the inverse triangle inequality to the above Equation (21), we get

$$\|X - Z\|_p^p = \|X_r - Z_r^c + X_r^c - Z_r\|_p^p \geq \|X_r - Z_r^c\|_p^p - \|X_r^c - Z_r\|_p^p. \tag{22}$$

Again, by Lemma 3.1 and inequality (22), we obtain

$$\begin{aligned} \|X_r - Z_r^c\|_p^p - \|X_r^c - Z_r\|_p^p &= \|X_r\|_p^p + \|Z_r^c\|_p^p - \|X_r^c - Z_r\|_p^p \\ &\geq \|X_r\|_p^p + \|Z_r^c\|_p^p - \|X_r^c\|_p^p - \|Z_r\|_p^p. \end{aligned} \tag{23}$$

Combining (21), (22), (23) and integrating their shifted terms, it is easy to show that

$$\begin{aligned} \|Z_r^c\|_p^p &\leq \|X\|_p^p - \|X_r\|_p^p + \|X_r^c\|_p^p + \|Z_r\|_p^p \\ &\leq \|X - X_r\|_p^p + \|X_r^c\|_p^p + \|Z_r\|_p^p \\ &\leq \|Z_r\|_p^p + 2\|X_r^c\|_p^p \end{aligned} \tag{24}$$

which finishes the proof of the above lemma.

Lemma 3.3: For any vector $y \in \mathbb{R}^b$, there is

$$\|y\|_p^p \leq b^{1-\frac{p}{2}} \|y\|_2^p. \tag{25}$$

Proof of the lemma 3.3 By definition of the l_p norm, there is

$\|y\|_p = \left(\sum_{i=1}^b |y_i|^p\right)^{\frac{1}{p}}$, and exploiting the Holder's inequality, we suffice to obtain

$$\|y\|_p^p = \sum_{i=1}^b |y_i|^p \cdot 1 \leq \left(\sum_{i=1}^b (|y_i|^p)^{\frac{2}{p}}\right)^{\frac{p}{2}} \cdot \left(\sum_{i=1}^b 1\right)^{1-\frac{p}{2}} = b^{1-\frac{p}{2}} \|y\|_2^p.$$

Therefore, the lemma 3.3 is proved.

Next, the restricted isometry constant δ_r and the relative disturbance upper bound $\epsilon_{\mathcal{A}}^{(r)}$ of the measured operator \mathcal{A} which does not have perturbations are already present, and Lemma 3.4 gives p-RIP condition of the perturbed measurement operator $\hat{\mathcal{A}}$ and $\hat{\delta}_r$.

Lemma 3.4: (The perturbed measurement operator $\hat{\mathcal{A}}$ of the P-RIP) Assume that the r-order RIC of the \mathcal{A} is denoted as δ_r , and the upper bound on the relative perturbation corresponding to the operator Φ is $\epsilon_{\mathcal{A}}^{(r)}$ and fix constant $\hat{\delta}_{r,\max} = (1 + \delta_r)(1 + \epsilon_{\mathcal{A}}^{(r)})^p - 1$, then $\hat{\mathcal{A}} = \mathcal{A} + \Phi$ of the RIC $\hat{\delta}_r \leq \hat{\delta}_{r,\max}$, $\hat{\delta}_r$ as the smallest and positive number that obeys

$$(1 - \delta_r) X_F^p \leq \|\hat{\mathcal{A}}(X)\|_p^p \leq (1 + \delta_r) \|X\|_F^p \tag{26}$$

for any matrices X which are r -rank.

Proof of the lemma 3.4 inspired by [20], first we define t_r and μ_r are the smallest positive or zero numbers that satisfy

$$(1 - t_r) \|X\|_F^p \leq \|\hat{\mathcal{A}}(X)\|_p^p \leq (1 + \mu_r) \|X\|_F^p \tag{27}$$

for any matrix X with rank at most r .

Using the triangle inequality, (9) and (11), we acquire

$$\begin{aligned} \|\hat{\mathcal{A}}(X)\|_p^p &= (\|\mathcal{A}(X)\|_p + \|\Phi(X)\|_p)^p \\ &\leq \left(\sqrt[p]{1 + \delta_r} + \|\Phi\|_p^{(r)}\right)^p \|X\|_F^p \\ &\leq \left(\sqrt[p]{1 + \delta_r} + \epsilon_{\mathcal{A}}^{(r)}\|\mathcal{A}\|_p^{(r)}\right)^p \|X\|_F^p \\ &\leq \left(\sqrt[p]{1 + \delta_r} + \epsilon_{\mathcal{A}}^{(r)}\sqrt[p]{1 + \delta_r}\right)^p \|X\|_F^p \\ &\leq (1 + \delta_r)(1 + \epsilon_{\mathcal{A}}^{(r)})^p \|X\|_F^p \end{aligned} \tag{28}$$

Because of the concept of μ_r , it means that

$$1 + \mu_r \leq (1 + \delta_r)(1 + \epsilon_{\mathcal{A}}^{(r)})^p, \tag{29}$$

by applying the above inequality (29), we obtain a minimum upper bound

$$\mu_r = (1 + \delta_r)(1 + \epsilon_{\mathcal{A}}^{(r)})^p - 1. \tag{30}$$

Similarly, taking advantage of the inverse triangle inequality, combined with the concept of RIC and (11) yields

$$t_r = 1 - (1 - \delta_r)(1 - \epsilon_{\mathcal{A}}^{(r)})^p. \tag{31}$$

Observe carefully that $1 - \mu_r \leq 1 - t_r$ and $1 + t_r \leq 1 + \mu_r$. Based on the given δ_r and $\epsilon_{\mathcal{A}}^{(r)}$, we choose $\mu_r = \hat{\delta}_{r, \max}$ the smallest nonnegative constant that makes (27) symmetric. Clearly, the true RIC of $\hat{\mathcal{A}}_{\hat{\delta}_r}$ satisfies $\hat{\delta}_r \leq \hat{\delta}_{r, \max}$. The proof of Lemma 3.4 is completed.

Finally, the following lemma 3.5 clarifies that the perturbed measurement operator $\hat{\mathcal{A}}$ can also comply with p-NSP, provided that the constants s and τ satisfy specific conditions and the measurement operator $\hat{\mathcal{A}}$ satisfies p-NSP.

Lemma 3.5: For the given $\epsilon_{\mathcal{A}}$, the measured operator $\hat{\mathcal{A}}$ satisfies the p-null space property with a constant $0 < s < 1 - 2v\epsilon_{\mathcal{A}}^p\|\mathcal{A}\|_p^p$ and a constant

$$0 < v < \frac{1}{\epsilon_{\mathcal{A}}^p\|\mathcal{A}\|_p^p}, \text{ which satisfies the condition of (10) for any } X \in \mathbb{R}^{m \times n}, \text{ and fix}$$

$$\text{two constants } \hat{v} = \frac{v}{1 - v\epsilon_{\mathcal{A}}^p\|\mathcal{A}\|_p^p} \text{ and } \hat{s} = s + \frac{v\epsilon_{\mathcal{A}}^p(\rho + 1)}{1 - v\epsilon_{\mathcal{A}}^p\|\mathcal{A}\|_p^p}. \text{ Then two constants}$$

$0 < \hat{s} < 1$ and $\hat{v} > 0$, and the perturbed measurement operator $\hat{\mathcal{A}}$ satisfies p-NSP.

Proof of the lemma 3.5 Utilizing (11) and the inequality, there are

$$\begin{aligned} \|\hat{\mathcal{A}}(X)\|_p^p &\leq \|\mathcal{A}(X)\|_p^p + \|\Phi(X)\|_p^p \\ &\leq \|\mathcal{A}(X)\|_p^p + \|\Phi\|_p^p \|X\|_p^p \\ &\leq \|\mathcal{A}(X)\|_p^p + \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p \|X\|_p^p. \end{aligned} \tag{32}$$

Since \mathcal{A} satisfies the p-null space property and $\|X\|_p^p \leq \|X_r\|_p^p + \|X_r^c\|_p^p$ holds, we achieve

$$\begin{aligned} \|X_r\|_p^p &\leq s \|X_r^c\|_p^p + v \|\mathcal{A}(X)\|_p^p \\ &\leq s \|X_r^c\|_p^p + v \|\hat{\mathcal{A}}(X)\|_p^p + v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p \|X\|_p^p \\ &\leq s \|X_r^c\|_p^p + v \|\hat{\mathcal{A}}(X)\|_p^p + v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p (\|X_r\|_p^p + \|X_r^c\|_p^p). \end{aligned} \tag{33}$$

Then we collapse the inequality (33) to conclude that

$$\|X_r\|_p^p \leq \frac{s + v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p} \|X_r^c\|_p^p + \frac{v}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p} \|\hat{\mathcal{A}}(X)\|_p^p.$$

Let $\hat{v} = \frac{v}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}$ and $\hat{s} = \frac{s + v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}$, basing the fact that $v > 0$ to

make $\hat{v} > 0$, we need to solve the following inequality $\frac{v}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p} > 0$, which

$$\text{imitates } 0 < v < \frac{1}{\epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}.$$

In view of $\epsilon_{\mathcal{A}}$ in (11), and it is also known that $s > 0$ to make $\hat{s} > 0$, we must solve

$$0 < \frac{s + v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}{1 - v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p} < 1,$$

after sorting out the above inequality, we obtain

$$v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p < s < 1 - 2v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p.$$

Combining the above, while two constants $0 < v < \frac{1}{\epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p}$ and

$0 < s < 1 - 2v \epsilon_{\mathcal{A}}^p \|\mathcal{A}\|_p^p$, we are able to come true $\hat{v} > 0$ and $0 < \hat{s} < 1$, hence the measurement operator $\hat{\mathcal{A}}$ obeys p-RNSP. In summary, we prove the Lemma 3.5.

After the previous preparations, we now prove two theorems. The upper bound estimation of the error of a real matrix X to be recovered with the optimal solution X^* of problem (8) is derived from the restricted isometry property, *i.e.*, theorem 2.1.

Proof of the theorem 2.1 Let $Z = X - X^*$, X be the real matrix that we expect to recover, X^* be the optimal solution to the problem (8). We apply the block decomposition of the SVD of the matrix Z given by

$$U^T ZV = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix},$$

where $Z_{ij} \in \mathbb{R}^{m_i \times n_j}$ with $m_1 = n_1 = r$ and $m_2 = n_2 = m - r$, $Z = Z_r + Z_r^c$,

$$Z_r = U \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & 0 \end{pmatrix} V^T, Z_r^c = U \begin{pmatrix} 0 & 0 \\ 0 & Z_{22} \end{pmatrix} V^T,$$

it is clear that $\text{rank}(Z_r) \leq \text{rank}[(Z_{11} \ Z_{12})] + \text{rank}[(Z_{21} \ 0)] \leq 2r$, and $X_r (Z_r^c)^T = 0$ and $(X_r)^T Z_r^c = 0$.

Let SVD of $Z_{22} \in \mathbb{R}^{(m-r) \times (m-r)}$ be described by

$$Z_{22} = H \text{diag}(\sigma(Z_{22})) G^T,$$

where matrices $H, G \in \mathbb{R}^{(m-r) \times (m-r)}$ are orthogonal,

$\sigma(Z_{22}) = (\sigma_1(Z_{22}), \sigma_2(Z_{22}), \dots, \sigma_{m-r}(Z_{22}))^T$ denotes a vector consisting of the singular values of Z_{22} and $\sigma_1(Z_{22}) \geq \sigma_2(Z_{22}) \geq \dots \geq \sigma_{m-r}(Z_{22}) \geq 0$. We divide $\sigma(Z_{22})$ into the sum of the vectors $\sigma_{T_i}(Z_{22}) (i = 1, 2, \dots)$,

$T_i = 2r(i-1) + 1, \dots, 2ri$, $T_1 \cup T_2 \cup \dots \cup T_{j=1,2,\dots,m-r}$, each T_i has sparsity $2r$ (except possibly T_j). Then, Z_{T_i} is the portion of Z_r^c that corresponds to the $k (k = 2tr)$ largest singular values, Z_{T_2} is the part that corresponds to the next k largest singular values, then so forth. Obviously, for all $i \neq j$, $Z_{T_i}^T Z_{T_j} = 0$ and $Z_{T_i} Z_{T_j}^T = 0$, and $\text{rank}(Z_{T_i}) \leq k$.

The following results can be derived from [17]

$$\|Z_{T_j}\|_F^p \leq (2tr)^{\frac{p}{2}-1} \|Z_{T_{j-1}}\|_F^p, \tag{34}$$

$$\|Z_r\|_p^p \leq (2r)^{\frac{p}{2}-1} \|Z_r\|_F^p \leq (2r)^{\frac{p}{2}-1} \|Z_r + Z_{T_1}\|_F^p. \tag{35}$$

According to (34), we get

$$\sum_{j \geq 2} \|Z_{T_j}\|_F^p \leq k^{\frac{p}{2}-1} \sum_{j \geq 2} \|Z_{T_{j-1}}\|_F^p = k^{\frac{p}{2}-1} \|Z_r^c\|_F^p. \tag{36}$$

Applying Lemma 3.2 to (36), we can have

$$\begin{aligned} \sum_{j \geq 2} \|Z_{T_j}\|_F^p &\leq k^{\frac{p}{2}-1} (\|Z_r\|_F^p + 2\|X_r^c\|_F^p) \\ &\leq k^{\frac{p}{2}-1} (2r)^{\frac{p}{2}-1} \|Z_r + Z_{T_1}\|_F^p + 2k^{\frac{p}{2}-1} \|X_r^c\|_F^p \\ &= t^{\frac{p}{2}-1} \|Z_r + Z_{T_1}\|_F^p + 2(2tr)^{\frac{p}{2}-1} \|X_r^c\|_F^p. \end{aligned} \tag{37}$$

Since $\|\hat{\mathcal{A}}(X) - \hat{b}\|_2 \leq \hat{\epsilon}_{\mathcal{A},r,b}$ and the triangular inequality, it suffices to

$$\|\hat{\mathcal{A}}(Z)\|_2 \leq \|\hat{\mathcal{A}}(X) - \hat{b}\|_2 + \|\hat{\mathcal{A}}(X^*) - \hat{b}\|_2 \leq 2\hat{\epsilon}_{\mathcal{A},r,b}.$$

Therefore, by combining the above equation with lemma 3.3, we have

$$\|\hat{\mathcal{A}}(Z)\|_p^p \leq M^{1-\frac{p}{2}} \|\hat{\mathcal{A}}(Z)\|_2^p \leq M^{1-\frac{p}{2}} (2\hat{\epsilon}_{\mathcal{A},r,b})^p. \tag{38}$$

Moreover, by Lemma 3.4 and (37), we get

$$\begin{aligned}
 \|\hat{\mathcal{A}}(Z)\|_p^p &= \|\hat{\mathcal{A}}(Z_r + Z_{T_1}) + \hat{\mathcal{A}}(Z_r^c - Z_{T_1})\|_p^p \\
 &\geq \|\hat{\mathcal{A}}(Z_r + Z_{T_1})\|_p^p - \sum_{j \geq 2} \|\hat{\mathcal{A}}(Z_{T_j})\|_p^p \\
 &\geq (1 - \delta_{2r+2tr}) \|Z_r + Z_{T_1}\|_F^p - (1 + \delta_{2tr}) \sum_{j \geq 2} \|Z_{T_j}\|_F^p \\
 &\geq \left[(1 - \delta_{2r(t+1)}) - (1 + \delta_{2tr}) t^{\frac{p-1}{2}} \right] \|Z_r + Z_{T_1}\|_F^p - 2(1 + \delta_{2tr}) (2tr)^{\frac{p-1}{2}} \|X_r^c\|_p^p \\
 &\geq \left[2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}} \right] \|Z_r + Z_{T_1}\|_F^p \\
 &\quad - 2(1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p (2tr)^{\frac{p-1}{2}} \|X_r^c\|_p^p,
 \end{aligned} \tag{39}$$

in that case

$$\|Z_r + Z_{T_1}\|_F^p \leq \frac{M^{1-\frac{p}{2}} (2\hat{\epsilon}_{\mathcal{A},r,b})^p + 2(2tr)^{\frac{p-1}{2}} (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p \|X_r^c\|_p^p}{2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}}} \tag{40}$$

From (40) and Lemma 3.2, we have that

$$\begin{aligned}
 \|Z\|_p^p &\leq \|Z_r + Z_r^c\|_p^p \leq \|Z_r\|_p^p + \|Z_r^c\|_p^p \leq 2\|Z_r\|_p^p + 2\|Z_r^c\|_p^p \\
 &\leq 2(2r)^{1-\frac{p}{2}} \|Z_r + Z_{T_1}\|_F^p + 2\|X_r^c\|_p^p \\
 &\leq \frac{2^{p+1} (2rM)^{1-\frac{p}{2}}}{2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}}} (\hat{\epsilon}_{\mathcal{A},r,b})^p \\
 &\quad + \left(2 + \frac{4(1 + \delta_{2tr}) t^{\frac{p-1}{2}} (1 + \epsilon_{\mathcal{A}}^{2tr})^p}{2 - (1 + \delta_{2r(t+1)}) (1 + \epsilon_{\mathcal{A}}^{2r(t+1)})^p - (1 + \delta_{2tr}) (1 + \epsilon_{\mathcal{A}}^{2tr})^p t^{\frac{p-1}{2}}} \right) \|X_r^c\|_p^p
 \end{aligned} \tag{41}$$

which finishes the proof of theorem 2.1.

When constants $0 < p \leq 1$ and $t > 1$, setting $k = 2tr$, a new RIP condition that can robustly and accurately recover low-rank matrices is obtained *i.e.*, (2 - 3) in Theorem 2.1, and it is slightly weaker than the sufficient condition of Zhang M [16] $\delta_k + \theta\delta_{2r+k} < \theta - 1$.

Theorem 2.2 is based on this element of the NSP of matrix, and provides an error-bound estimation of the actual r-rank matrix X to be found and the optimal solution X^* of problem (8). Its proof steps as below:

Proof According to Lemma 3.4, while $\hat{\mathcal{A}}$ meets the p-null space property (pNSP), there exist

$$\|X_r\|_p^p \leq \hat{s} \|X_r^c\|_p^p + \hat{v} \|\hat{\mathcal{A}}(X)\|_p^p. \tag{42}$$

Let $Z = X - X^*$, it is obvious that we get that by the lemma 3.2 and (42), we

$$\left\| (X - X^*)^c \right\|_r^p \leq \left\| (X - X^*) \right\|_r^p + 2 \left\| X_r^c \right\|_p^p \leq \hat{s} \left\| X_r^c \right\|_p^p + \hat{v} \left\| \hat{\mathcal{A}}(X) \right\|_p^p + 2 \left\| X_r^c \right\|_p^p, \quad (43)$$

holds through the lemma 3.2 and (42). After simplifying, we further obtain

$$\left\| (X - X^*)^c \right\|_r^p \leq \frac{1}{1 - \hat{s}} \left(\hat{v} \left\| \hat{\mathcal{A}}(X) \right\|_p^p + 2 \left\| X_r^c \right\|_p^p \right). \quad (44)$$

Finally, we conclude from (38), (43) and (44) that

$$\begin{aligned} \left\| X - X^* \right\|_p^p &\leq \left\| (X - X^*) \right\|_r^p + \left\| (X - X^*)^c \right\|_r^p \\ &\leq \hat{s} \left\| X_r^c \right\|_p^p + \hat{v} \left\| \hat{\mathcal{A}}(X) \right\|_p^p + \left\| (X - X^*)^c \right\|_r^p \\ &\leq (1 + \hat{s}) \left\| X_r^c \right\|_p^p + \hat{v} \left\| \hat{\mathcal{A}}(X) \right\|_p^p \\ &\leq \left(\frac{1 + \hat{s}}{1 - \hat{s}} + 1 \right) \hat{v} \left\| \hat{\mathcal{A}}(X) \right\|_p^p + \frac{2(1 + \hat{s})}{1 - \hat{s}} \left\| X_r^c \right\|_p^p \\ &\leq \frac{2\hat{v}}{1 - \hat{s}} M^{1-\frac{p}{2}} (2\hat{\epsilon}_{\mathcal{A},r,b})^p + \frac{2(1 + \hat{s})}{1 - \hat{s}} \left\| X_r^c \right\|_p^p \\ &\leq \frac{vM^{1-\frac{p}{2}} 2^{p+1}}{1 - s - 2v\epsilon_{\mathcal{A}}^p} (\hat{\epsilon}_{\mathcal{A},r,b})^p + \frac{2(1 + s)}{1 - s - 2v\epsilon_{\mathcal{A}}^p} \left\| X_r^c \right\|_p^p. \end{aligned}$$

holds for $0 < p \leq 1$. Meanwhile, we finish the proof of the theorem 2.2.

In brief, the above completes all proofs of the two theorems.

4. Conclusion

We primarily investigate mainly study fully perturbed problem of reconstructing a low-rank matrix through nonconvex Schatten p-norm minimization and give sufficient conditions for recovering the error along with corresponding upper bound estimations. These results show that nonconvex Schatten p-minimization provides a stable and accurate guarantee for reconstructing low-rank matrix in the existence of overall noise. The obtained results involve two implications, firstly, it suffices to guide the selection of measurement operators for low-rank matrix reconstruction, *i.e.*, operators that satisfy weaker RIC sufficient conditions can also better promote the recovery capacity; secondly, it provides theoretical support for boundary error by taking advantage of the following two properties, respectively p-RIP and p-NSP.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Recht, B., Fazel, M. and Parrilo, P.A. (2010) Guaranteed Minimum Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization. *SIAM Review*, **52**, 471-501. <https://doi.org/10.1137/070697835>

- [2] Chang, X., Zhong, Y., Wang, Y. and Lin, S. (2019) Unified Low-Rank Matrix Estimate via Penalized Matrix Least Squares Approximation. *IEEE Transactions on Neural Networks and Learning Systems*, **30**, 474-485. <https://doi.org/10.1109/TNNLS.2018.2844242>
- [3] Batselier, K. (2022) Low-Rank Tensor Decompositions for Nonlinear System Identification: A Tutorial with Examples. *IEEE Control Systems Magazine*, **42**, 54-74. <https://doi.org/10.1109/MCS.2021.3122268>
- [4] Dass, J., Wu, S., Shi, H., et al. (2023) Vitality: Unifying Low-Rank and Sparse Approximation for Vision Transformer Acceleration with a Linear Taylor Attention. *2023 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*. Montreal, 25 February-01 March 2023, 415-428. <https://doi.org/10.1109/HPCA56546.2023.10071081>
- [5] Wang, Z., Qian, C., Guo, D., et al. (2022) One-Dimensional Deep Low-Rank and Sparse Network for Accelerated MRI. *IEEE Transactions on Medical Imaging*, **42**, 79-90. <https://doi.org/10.1109/TMI.2022.3203312>
- [6] Wang, Y., Lin, L., Zhao, Q., et al. (2017) Compressive Sensing of Hyperspectral Images via Joint Tensor Tucker Decomposition and Weighted Total Variation Regularization. *IEEE Geoscience and Remote Sensing Letters*, **14**, 2457-2461. <https://doi.org/10.1109/LGRS.2017.2771212>
- [7] Gundupalli, S.P., Hait, S. and Thakur, A. (2017) A Review on Automated Sorting of Source-Separated Municipal Solid Waste for Recycling. *Waste Management*, **60**, 56-74. <https://doi.org/10.1016/j.wasman.2016.09.015>
- [8] Ye, G., Pan, C., Dong, Y., et al. (2021) A Novel Multi-Image Visually Meaningful Encryption Algorithm Based on Compressive Sensing and Schur Decomposition. *Transactions on Emerging Telecommunications Technologies*, **32**, e4071. <https://doi.org/10.1002/ett.4071>
- [9] Chartrand, R. (2007) Exact Reconstruction of Sparse Signals via Nonconvex Minimization. *IEEE Signal Processing Letters*, **14**, 707-710. <https://doi.org/10.1109/LSP.2007.898300>
- [10] Lai, M.J., Xu, Y. and Yin, W. (2013) Improved Iteratively Reweighted Least Squares for Unconstrained Smoothed l_q Minimization. *SIAM Journal on Numerical Analysis*, **51**, 927-957. <https://doi.org/10.1137/110840364>
- [11] Wang, Y., Wang, J. and Xu, Z. (2014) Restricted P-Isometry Properties of Nonconvex Block-Sparse Compressed Sensing. *Signal Processing*, **104**, 188-196. <https://doi.org/10.1016/j.sigpro.2014.03.040>
- [12] Wang, J., Zhang, J., Wang, W., et al. (2015) A Perturbation Analysis of Nonconvex Block-Sparse Compressed Sensing. *Communications in Nonlinear Science and Numerical Simulation*, **29**, 416-426. <https://doi.org/10.1016/j.cnsns.2015.05.022>
- [13] Chartrand, R. (2007) Nonconvex Compressed Sensing and Error Correction. *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP07*. Honolulu, 15-20 April 2007, III-889-III-892. <https://doi.org/10.1109/ICASSP.2007.366823>
- [14] Candes, E.J. and Tao, T. (2005) Decoding by Linear Programming. *IEEE Transactions on Information Theory*, **51**, 4203-4215. <https://doi.org/10.1109/TIT.2005.858979>
- [15] Kong, L.C. and Xiu, N.H. (2011) Exact Low-Rank Matrix Recovery via Nonconvex l_p -Minimization. *Optimization*.
- [16] Zhang, M., Huang, Z.H. and Zhang, Y. (2013) Restricted P-Isometry Properties of

- Nonconvex Matrix Recovery. *IEEE Transactions on Information Theory*, **59**, 4316-4323. <https://doi.org/10.1109/TIT.2013.2250577>
- [17] Kong, L. and Xiu, N. (2013) Exact Low-Rank Matrix Recovery via Nonconvex Schatten P-Minimization. *Asia-Pacific Journal of Operational Research*, **30**, Article 1340010. <https://doi.org/10.1142/S0217595913400101>
- [18] Gao, Y., Peng, J., Yue, S., *et al.* (2015) On the Null Space Property of l_q -Minimization for $0 < q \leq 1$ in Compressed Sensing. *Journal of Function Spaces*, **2015**, Article ID 579853. <https://doi.org/10.1155/2015/579853>
- [19] Herman, M.A. and Strohmer, T. (2010) General Deviants: An Analysis of Perturbations in Compressed Sensing. *IEEE Journal of Selected Topics in Signal Processing*, **4**, 342-349. <https://doi.org/10.1109/JSTSP.2009.2039170>
- [20] Huang, J., Wang, J., Zhang, F., *et al.* (2021) Perturbation Analysis of Low-Rank Matrix Stable Recovery. *International Journal of Wavelets, Multiresolution and Information Processing*, **19**, Article 2050091. <https://doi.org/10.1142/S0219691320500915>