ℓ_p Regularized Low-Rank Approximation via Iterative Reweighted Singular Value Minimization

Zhaosong Lu* Yong Zhang † Jian Lu ‡

December 31, 2016 (Revised: May 4, 2017; July 28, 2017)

Abstract

In this paper we study the ℓ_p (or Schatten-p quasi-norm) regularized low-rank approximation problems. In particular, we introduce a class of first-order stationary points for them and show that any local minimizer of these problems must be a first-order stationary point. In addition, we derive lower bounds for the nonzero singular values of the first-order stationary points and hence also of the local minimizers of these problems. The iterative reweighted singular value minimization (IRSVM) methods are then proposed to solve these problems, whose subproblems are shown to have a closed-form solution. Compared to the analogous methods for the ℓ_p regularized *vector* minimization problems, the convergence analysis of these methods is significantly more challenging. We develop a novel approach to establishing the convergence of the IRSVM methods, which makes use of the expression of a specific solution of their subproblems and avoids the intricate issue of finding the explicit expression for the Clarke subdifferential of the objective of their subproblems. In particular, we show that any accumulation point of the sequence generated by the IRSVM methods is a first-order stationary point of the problems. Our computational results demonstrate that the IRSVM methods generally outperform the recently developed iterative reweighted least squares methods in terms of solution quality and/or speed.

Key words: low-rank approximation, Schatten-p quasi-norm regularized matrix minimization, iterative reweighted singular value minimization, iterative reweighted least squares

AMS subject classifications: 15A18, 15A83, 65K05, 90C26, 90C30

1 Introduction

Over the last decade, finding a low-rank solution to a system or an optimization problem has attracted a great deal of attention in science and engineering. Numerous optimization models and methods have

*Department of Mathematics, Simon Fraser University, Canada. (email: `zhaosong@sfu.ca`). This author was supported in part by NSERC Discovery Grant.

†Istuary Innovation Group, Vancouver, BC, Canada (email: `michaelzhang917@gmail.com`).

‡College of Mathematics and Computational Science, Shenzhen University, Shenzhen, PR China (email: `jianlu@szu.edu.cn`). This author was supported in part by the National Natural Science Foundation of China under grants 61373087, 61472257, by the Natural Science Foundation of Guangdong, China under grants 2015A030313550, 2015A030313557.

been proposed for it (e.g., see [16, 37, 31, 39, 42, 33, 19, 26, 30]). In this paper we are interested in one of those models, particularly, the ℓ_p (or Schatten-p quasi-norm) regularized low-rank approximation problem:

$$\min_{X\in\Re^{m\times n}} \left\{F(X) := f(X) + \lambda\|X\|_p^p\right\} \tag{1}$$

for some $\lambda > 0$ and $p \in (0,1)$, which is also commonly referred to as the ℓ_p or Schatten-p quasi-norm regularized *matrix* minimization problem. Here, $\|X\|_p := (\sum_{i=1}^r \sigma_i(X)^p)^{1/p}$ for any $X \in \Re^{m\times n}$, where $r = \mathrm{rank}(X)$ and $\sigma_i(X)$ denotes the ith largest singular value of X. And f is assumed to be a smooth function with L_f-Lipschitz-continuous gradient in $\Re^{m\times n}$, that is,

$$\|\nabla f(X) - \nabla f(Y)\|_F \le L_f\|X - Y\|_F, \quad \forall X, Y \in \Re^{m\times n}, \tag{2}$$

and f is bounded below in $\Re^{m\times n}$, where $\|\cdot\|_F$ is the Frobenius norm. One can observe that as $p \downarrow 0$, problem (1) approaches the rank minimization problem

$$\min_{X\in\Re^{m\times n}} f(X) + \lambda\cdot\mathrm{rank}(X), \tag{3}$$

which is an exact formulation of finding a low-rank matrix to minimize f. In addition, as $p \uparrow 1$, problem (1) approaches the so-called nuclear (or trace) norm minimization problem

$$\min_{X\in\Re^{m\times n}} f(X) + \lambda\|X\|_*, \tag{4}$$

which is a widely used convex relaxation for (3). In the context of low-rank matrix completion, model (4) has been shown to be extremely effective in finding a low-rank matrix to minimize f. A variety of efficient methods were recently proposed for solving (4) (e.g., see [31, 39]). Since problem (1) is an intermediate model between (3) and (4), one can expect that it is also capable of seeking out a low-rank matrix to minimize f. This has indeed been demonstrated in the context of matrix completion and network localization by extensive computational studies in [26, 22]. In addition, under some restricted isometry property (RIP) conditions, it was shown that minimizing ℓ_p matrix quasi-norm over an affine matrix manifold can exactly recover a low-rank matrix from compressed linear measurements (e.g., see [23, 24, 45, 32]).

One can observe that if X is additionally required to be diagonal, problem (1) is reduced to an ℓ_p regularized *vector* minimization problem in the form of

$$\min_{x\in\Re^l} \left\{h(x) + \lambda\|x\|_p^p\right\}, \tag{5}$$

where $l := \min(m, n)$ and $\|x\|_p := (\sum_{i=1}^l |x_i|^p)^{1/p}$ for any $x \in \Re^l$. Problem (5) and its variants have been widely studied in the literature for recovering sparse vectors (e.g., see [8, 9, 10, 18, 44, 25, 13, 14, 15, 35, 11, 20, 38, 3, 12]). Efficient iterative reweighted l_1 (IRL$_1$) and l_2 (IRL$_2$) minimization algorithms were also proposed for finding an approximate solution to (5) or its variants (e.g., see [36, 10, 18, 15, 25, 14, 26, 29]).

Recently, Ji et al. [22] proposed an interior-point method for finding an approximate solution of a class of ℓ_p regularized *matrix* minimization problems arising in network localization. Lai et al. [26] extended the IRL$_2$ method that was proposed in the same paper for (5) to solve (1) with f being a convex quadratic function. In addition, Fornasier et al. [19], and Mohan and Fazel [33] extended the IRL$_2$ method, which was proposed in [15] for ℓ_p regularized *vector* minimization over an affine manifold, to minimize ℓ_p matrix quasi-norm over an affine manifold. In each iteration, these methods [26, 19, 33] solve a convex quadratic programming problem that has a closed-form solution. These methods are so-called iterative reweighted least squares (IRLS) methods in the literature. It is well known that the approximate solution obtained by IRL$_2$ methods when applied to (5) or its variants is generally dense. Analogously, the IRLS methods usually do not produce a low-rank solution. On the other hand, it is known that the IRL$_1$ methods tend to produce a sparse solution when applied to (5) or its variants. Thus one can expect that the extension of IRL$_1$ methods to (1) is capable of producing a low-rank solution of (1). Such an extension, however, has not yet been considered in the literature due to some technical difficulties.

To illustrate the aforementioned difficulties, we consider one of the IRL$_1$ methods proposed in [29] for solving (5), which generates a new iterate x^{k+1} by solving the weighted ℓ_1 subproblem:

$$x^{k+1} = \arg\min \left\{ \frac{1}{2} \left\| x - \left(x^k - \frac{1}{L_k} \nabla h(x^k) \right) \right\|^2 + \lambda p \sum_{i=1}^{l} s_i^k |x_i| \right\} \tag{6}$$

for some $L_k > 0$ and $s_i^k = (|x_i^k| + \epsilon_k)^{p-1}$ for all i, where $\epsilon_k > 0$. One can naturally extend this method to solve (1) by generating a sequence $\{X^k\} \subset \Re^{m \times n}$ according to

$$X^{k+1} \in \operatorname{Arg\,min} \left\{ \frac{1}{2} \left\| X - \left(X^k - \frac{1}{L_k} \nabla f(X^k) \right) \right\|_F^2 + \lambda p \sum_{i=1}^{l} s_i^k \sigma_i(X) \right\}^{1} \tag{7}$$

for $s_i^k = (\sigma_i(X^k) + \epsilon_k)^{p-1}$ for all i. Notice that the weights of the singular values of X are updated from one iteration to another. This method is thus referred to as an iterative reweighted singular value minimization (IRSVM) method. Though such an IRSVM method is an extension of the above IRL$_1$ method, there are some major differences between them. Firstly, the objective function of (7) is generally nonconvex while that of (6) is convex. One immediate question is whether (7) is solvable or not. Secondly, the convergence analysis of the IRSVM is significantly more challenging than that of IRL$_1$. Indeed, the convergence analysis of IRL$_1$ relies on the explicit expression of the (Clarke) subdifferential of the objective of (6), which is immediately available due to the simple structure of the objective of (6). Nevertheless, for problem (7), it is highly challenging to find an explicit expression for that due to the complication of $\Psi(X) = \sum_{i=1}^{l} s_i^k \sigma_i(X)$. To see this challenge, we follow [27, 28] to express Ψ as the composition of an *absolutely symmetric* function[2] ψ and all singular values of X, that is, $\Psi(X) = (\psi \circ \bar{\sigma})(X)$, where $\bar{\sigma}(X) = (\sigma_1(X), \sigma_2(X), \cdots, \sigma_l(X))$, $\psi(x) = \sum_{i=1}^{l} s_i^k \psi_i(x)$ and

$$\psi_i(x) = i\text{th largest element of } \{|x_1|, |x_2|, \ldots, |x_l|\}, \qquad \forall x \in \Re^l.$$

It then follows from [28, Theorem 3.7] that the Clarke subdifferential of Ψ at X is given by

$$\partial\Psi(X) = \{U \mathrm{Diag}(d) V^T : d \in \partial\psi(\bar{\sigma}(X)), (U, V) \in \overline{\mathcal{M}}(X)\},$$

where $\partial\psi$ is the Clarke subdifferential of ψ and

$$\overline{\mathcal{M}}(X) = \left\{ (U, V) \in \Re^{m \times l} \times \Re^{n \times l} : U^T U = V^T V = I, X = U \mathrm{Diag}(\bar{\sigma}(X)) V^T \right\}.$$

---

[1] By convention, the symbol "Arg" stands for the set of the solutions of the associated optimization problem. When this set is known to be a singleton, we use the symbol "arg" to stand for it instead.

[2] $\psi : \Re^l \to \Re$ is an absolutely symmetric function if $\psi(x_1, x_2 \ldots, x_l) = \psi(|x_{\pi(1)}|, |x_{\pi(2)}|, \ldots, |x_{\pi(l)}|)$ for any permutation π.

Due to the complication of ψ, it is rather challenging to find an explicit expression for $\partial\psi$ and hence for $\partial\Psi$. Therefore, the convergence analysis of IRL_1 in [29] is not applicable to IRSVM.

In this paper we propose two IRSVM methods for solving problem (1), which are the extension of some efficient IRL_1 methods studied in [29] for solving (5). As mentioned above, the convergence analysis of the IRSVM methods is significantly more challenging than that of the IRL_1 methods. Rather than following the approach for IRL_1 that involves the intricate issue of finding the explicit expression for the Clarke subdifferential of the objective of the subproblems, we develop a novel approach to establishing the convergence of the IRSVM methods, which makes use of the expression of a specific solution of the subproblems. In particular, we show that any accumulation point of the sequence generated by IRSVM is a first-order stationary point of problem (1). We also conduct numerical experiments to compare the IRSVM methods with the iterative reweighted least squares methods [26]. The computational results demonstrate that the IRSVM methods generally outperform those methods in terms of solution quality and/or speed.

The outline of this paper is as follows. In subsection 1.1, we introduce some notations that are used in the paper. In section 2, we introduce first-order stationary points for problem (1) and study some properties for them. In section 3, we extend two IRL_1 methods proposed in [29] to solve (1) and establish their convergence. We conduct numerical experiments in section 4 to compare the IRSVM methods with the iterative reweighted least squares methods [26]. Finally, in section 5 we present some concluding remarks.

1.1 Notation

The set of all n-dimensional nonnegative (resp., positive) vectors is denoted by $\Re^n_+$ (resp., $\Re^n_{++}$). $x \geq 0$ (resp., $x > 0$) means $x \in \Re^n_+$ (resp., $x \in \Re^n_{++}$). Given a vector x and a scalar α, $|x|^\alpha$ (resp., x^α) denotes the vector whose ith component is $|x_i|^\alpha$ (resp., x_i^α). In addition, $\text{Diag}(x)$ or $\text{Diag}(x_1, \ldots, x_n)$ denotes a square or non-square matrix with vector x on its main diagonal and zeros elsewhere, whose dimension shall be clear from the context. Given any $x, y \in \Re^n$, $x \circ y$ denotes the Hadamard product of x and y, namely, $(x \circ y)_i = x_i y_i$ for all i.

The space of $m \times n$ matrices is denoted by $\Re^{m\times n}$. Given any $X \in \Re^{m\times n}$, the Frobenius norm of X is denoted by $\|X\|_F$, namely, $\|X\|_F = \sqrt{\text{tr}(XX^T)}$, where $\text{tr}(\cdot)$ denotes the trace of a matrix. The entry-wise infinity norm of X is denoted by $\|X\|_{\max}$, that is, $\|X\|_{\max} = \max_{ij} |X_{ij}|$. For $X \in \Re^{m\times n}$, let $\sigma_i(X)$ denote the ith largest singular value of X for $i = 1, \ldots, \min(m, n)$, $\sigma(X) = (\sigma_1(X), \ldots, \sigma_r(X))^T$, and

$$\mathcal{M}(X) = \{(U, V) \in \Re^{m\times r} \times \Re^{n\times r} : U^T U = V^T V = I, X = U\text{Diag}(\sigma(X))V^T\}, \tag{8}$$

where $r = \text{rank}(X)$. Given any $X, Y \in \Re^{m\times n}$, the standard inner product of X and Y is denoted by $\langle X, Y\rangle$, that is, $\langle X, Y\rangle = \text{tr}(XY^T)$. If a symmetric matrix X is positive semidefinite (resp., definite), we write $X \succeq 0$ (resp., $X \succ 0$).

Finally, $|\Omega|$ denotes the cardinality of a finite set Ω. For any $\alpha < 0$, we define $0^\alpha = \infty$. The sign operator is denoted by sgn, that is,

$$\text{sgn}(t) = \begin{cases} 1 & \text{if } t > 0, \\ 0 & \text{if } t = 0, \\ -1 & \text{otherwise.} \end{cases}$$

Also, we define

$$\underline{f} \ = \inf_{X\in\Re^{m\times n}} f(X). \tag{9}$$

It follows from our early assumption on f that $-\infty < \underline{f} < \infty$.

2 Stationary points of (1) and lower bounds of their nonzero singular values

In this section we first introduce a class of first-order stationary points for problem (1). We then show that any local minimizer of (1) is a first-order stationary point. Finally, we derive lower bounds for nonzero singular values of the first-order stationary points and hence also of the local minimizers of problem (1).

Chen et al. [12] recently introduced a class of first-order stationary points for some non-Lipschitz optimization problems which include (5) as a special case. Specifically, $x^* \in \Re^l$ is a first-order stationary point of (5) if

$$X^* \nabla h(x^*) + \lambda p |x^*|^p = 0, \tag{10}$$

where $X^* = \mathrm{Diag}(x^*)$. It is not hard to observe that (10) is equivalent to

$$\mathrm{sgn}(x^*) \circ \nabla h(x^*) + \lambda p |x^*|^{p-1} = 0.$$

Inspired by this, we now introduce a class of first-order stationary points for problem (1).

Definition 1 *$X^* \in \Re^{m \times n}$ is a first-order stationary point of problem (1) if*

$$0 \in \left\{ U^T \nabla f(X^*) V + \lambda p \mathrm{Diag}(\sigma(X^*)^{p-1}) : (U, V) \in \mathcal{M}(X^*) \right\}. \tag{11}$$

We next show that any local minimizer of (1) is a first-order stationary point of the problem.

Theorem 2.1 *Suppose that X^* is a local minimizer of problem (1). Then X^* is a first-order stationary point of (1), that is, (11) holds at X^*.*

Proof. Let $X^* = \bar{U} \mathrm{Diag}(\sigma(X^*)) \bar{V}^T$ for some $(\bar{U}, \bar{V}) \in \mathcal{M}(X^*)$ and $r = \mathrm{rank}(X^*)$. By the assumption that X^* is a local minimizer of (1), one can see that 0 is a local minimizer of the problem

$$\min_{Z \in \Re^{r \times r}} f(X^* + \bar{U} Z \bar{V}^T) + \lambda \|X^* + \bar{U} Z \bar{V}^T\|_p^p.$$

This, together with the relation $X^* = \bar{U} \mathrm{Diag}(\sigma(X^*)) \bar{V}^T$, implies that 0 is a local minimizer of the problem

$$\min_{Z \in \Re^{r \times r}} \underbrace{f(X^* + \bar{U} Z \bar{V}^T) + \lambda \|\mathrm{Diag}(\sigma(X^*)) + Z\|_p^p}_{w(Z)}. \tag{12}$$

By [28, Theorem 3.7] and the definition of $\mathcal{M}(\cdot)$, the Clarke subdifferential of w at $Z = 0$ is given by

$$\partial w(0) = \left\{ \bar{U}^T \nabla f(X^*) \bar{V} + \lambda p U_\sigma \mathrm{Diag}(\sigma(X^*)^{p-1}) V_\sigma^T : (U_\sigma, V_\sigma) \in \mathcal{M}(\mathrm{Diag}(\sigma(X^*))) \right\}.$$

Since 0 is a local minimizer of problem (12), the first-order optimality condition of (12) yields $0 \in \partial w(0)$. Hence, there exists some $(U_\sigma, V_\sigma) \in \mathcal{M}(\mathrm{Diag}(\sigma(X^*)))$ such that

$$\bar{U}^T \nabla f(X^*) \bar{V} + \lambda p U_\sigma \mathrm{Diag}(\sigma(X^*)^{p-1}) V_\sigma^T = 0. \tag{13}$$

Due to $(U_\sigma, V_\sigma) \in \mathcal{M}(\mathrm{Diag}(\sigma(X^*)))$, one has $U_\sigma^T U_\sigma = V_\sigma^T V_\sigma = I$. Using this relation and upon pre- and post-multiplying (13), we obtain that

$$U^T \nabla f(X^*) V + \lambda p \mathrm{Diag}(\sigma(X^*)^{p-1}) = 0, \tag{14}$$

where $U = \bar{U} U_\sigma$ and $V = \bar{V} V_\sigma$. Since $(U_\sigma, V_\sigma) \in \mathcal{M}(\mathrm{Diag}(\sigma(X^*)))$ and $(\bar{U}, \bar{V}) \in \mathcal{M}(X^*)$, we have

$$U \mathrm{Diag}(\sigma(X^*)) V^T = \bar{U}(U_\sigma \mathrm{Diag}(\sigma(X^*)) V_\sigma^T) \bar{V}^T = \bar{U} \mathrm{Diag}(\sigma(X^*)) \bar{V}^T = X^*,$$

which together with (14) implies that (11) holds at X^*. ■

Before ending this section we derive lower bounds for the nonzero singular values of the first-order stationary points and hence also of the local minimizers of problem (1).

Theorem 2.2 *Let X^* be a first-order stationary point of (1) satisfying $F(X^*) \le M$ for some $M \in \Re$, $\mathcal{B} = \{i : \sigma_i(X^*) \ne 0\}$, L_f and $\underline{f}$ be defined in (2) and (9) , respectively. Then the following property holds:*

$$\sigma_i(X^*) \ \ge \ \left(\frac{\lambda p}{\sqrt{2L_f(M - \underline{f})}}, \right)^{\frac{1}{1-p}}, \qquad \forall i \in \mathcal{B}. \tag{15}$$

Proof. Since f satisfies (2), it is well-known that

$$f(Y) \ \le \ f(X) + \langle \nabla f(X), Y - X \rangle + \frac{L_f}{2} \|Y - X\|_F^2, \qquad \forall X, Y \in \Re^{m \times n}.$$

Substituting $X = X^*$ and $Y = X^* - \nabla f(X^*)/L_f$ into the above inequality, we obtain that

$$f(X^* - \nabla f(X^*)/L_f) \ \le \ f(X^*) - \frac{1}{2L_f} \|\nabla f(X^*)\|_F^2. \tag{16}$$

Note that

$$f(X^* - \nabla f(X^*)/L_f) \ge \inf_{X \in \Re^{m \times n}} f(X) = \underline{f}, \qquad f(X^*) \le F(X^*) \le M.$$

Using these relations and (16), we have

$$\|\nabla f(X^*)\|_F \ \le \ \sqrt{2L_f[f(X^*) - f(X^* - \nabla f(X^*)/L_f)]} \ \le \ \sqrt{2L_f(M - \underline{f})}. \tag{17}$$

Since X^* is a first-order stationary point of (1), we know that X^* satisfies (11) for some $(U, V) \in \mathcal{M}(X^*)$. Using (11) and $p \in (0,1)$, we obtain that for every $i \in \mathcal{B}$,

$$\sigma_i(X^*) = \left(\frac{[-U^T \nabla f(X^*) V]_{ii}}{\lambda p} \right)^{\frac{1}{p-1}} \ \ge \ \left(\frac{\|U^T \nabla f(X^*) V\|_F}{\lambda p} \right)^{\frac{1}{p-1}}. \tag{18}$$

Since $(U, V) \in \mathcal{M}(X^*)$, we know that $U^T U = V^T V = I$. By these relations, it is not hard to see that $\|U^T \nabla f(X^*) V\|_F \ \le \ \|\nabla f(X^*)\|_F$, which together with (18) yields

$$\sigma_i(X^*) \ge \left(\frac{\|\nabla f(X^*)\|_F}{\lambda p} \right)^{\frac{1}{p-1}}, \qquad i \in \mathcal{B}.$$

Using this relation, (17), and $p \in (0,1)$, one can see that (15) holds. ■

3 Iterative reweighted singular value minimization methods for (1)

In this section we propose iterative reweighted singular value minimization (IRSVM) methods for problem (1) by extending two IRL_1 methods proposed by Lu [29] for (5) to solve (1) and establish their convergence. The Computational study in [29] demonstrates that these two are most efficient among all existing IRL_1 methods for (5).

Throughout this section, let $l := \min(m, n)$. As seen in sequel, the key component of our IRSVM methods is to solve a weighted singular value minimization problem each iteration in the form of

$$\min_{X \in \Re^{m \times n}} \left\{ \langle C, X - B \rangle + \frac{L}{2} \|X - B\|_F^2 + \sum_{i=1}^{l} s_i \sigma_i(X) \right\} \tag{19}$$

for some $B, C \in \Re^{m \times n}$, $L > 0$ and $s \in \Re^l_+$ satisfying $s_1 \le s_2 \le \cdots \le s_l$. We now show that problem (19) has a closed-form solution, which resembles the weighted ℓ_1 minimization problem (6).

Theorem 3.1 *Given $B, C \in \Re^{m \times n}$, $L > 0$ and $s \in \Re^l_+$ satisfying $s_1 \le s_2 \le \cdots \le s_l$, let $U\mathrm{Diag}(d)V^T$ be the singular value decomposition of $B - C/L$, and $x^* = \max(d - s/L, 0)$. Then $X^* = U\mathrm{Diag}(x^*)V^T$ is an optimal solution to problem* (19).

Proof. First, one can verify $(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 \le (\alpha_2 - \beta_1)^2 + (\alpha_1 - \beta_2)^2$ for any $\alpha_1 \ge \alpha_2$ and $\beta_1 \ge \beta_2$. By this relation, the monotonicity and nonnegativity of d, and the permutation invariance of $\sum_{i=1}^l s_i \sigma_i(\mathrm{Diag}(x))$ with respect to $x \in \Re^l$, one can observe that the problem

$$\min_{x \in \Re^l} \left\{ \frac{L}{2} \|x - d\|^2 + \sum_{i=1}^{l} s_i \sigma_i(\mathrm{Diag}(x)) \right\} \tag{20}$$

has an optimal solution $\tilde{x}^* \in \Re^l_+$ with $\tilde{x}^*_1 \ge \tilde{x}^*_2 \ge \cdots \ge \tilde{x}^*_l$. It follows from this observation and the fact $\sum_{i=1}^l s_i \sigma_i(\mathrm{Diag}(x)) = \sum_{i=1}^l s_i |x_i|$ for any $x \in \Re^l_+$ satisfying $x_1 \ge x_2 \ge \cdots \ge x_l$ that $\tilde{x}^*$ is also the optimal solution of the problem

$$\min_{x \in \Re^l_+} \left\{ \frac{L}{2} \|x - d\|^2 + \sum_{i=1}^{l} s_i x_i : x_1 \ge x_2 \ge \cdots \ge x_l \right\}. \tag{21}$$

In addition, by $L > 0$, $d, s \in \Re^l_+$ and $x^* = \max(d - s/L, 0)$, we can observe that

$$x^* = \arg\min_{x \in \Re^l} \left\{ \frac{L}{2} \|x - d\|^2 + \sum_{i=1}^{l} s_i |x_i| \right\}.$$

Also, by $d_1 \ge d_2 \ge \cdots \ge d_l$, $s_1 \le s_2 \le \cdots \le s_l$ and $x^* = \max(d - s/L, 0)$, we also see that $x^*_1 \ge x^*_2 \ge \cdots \ge x^*_l \ge 0$. It follows from these facts that x^* is also an optimal solution of (21). Since (21) has a unique optimal solution, we conclude that $\tilde{x}^* = x^*$. Hence, x^* is an optimal solution of (20). Finally, one can observe that problem (19) is equivalent to

$$\min_{X \in \Re^{m \times n}} \left\{ \frac{L}{2} \left\| X - \left(B - \frac{C}{L} \right) \right\|_F^2 + \sum_{i=1}^{l} s_i \sigma_i(X) \right\}. \tag{22}$$

Invoking [30, Proposition 2.1] with $F(X) = \sum_{i=1}^{l} s_i\sigma_i(X)$, $\phi(t) = Lt^2/2$ and $\|\cdot\| = \|\cdot\|_F$ and using the fact that x^* is an optimal solution of (20), we can conclude that $X^* = U\mathrm{Diag}(x^*)V^T$ is an optimal solution of (22). Due to the equivalence of (19) and (22), X^* is also an optimal solution of (19). ■

3.1 The first IRSVM method for (1)

In this subsection we present the first IRSVM method for (1), which is an extension of an IRL$_1$ method (namely, [29, Algorithm 5]) that was proposed for solving the ℓ_p regularized *vector* minimization in the form of (5) and study its convergence. To proceed, we define

$$\bar{F}_\epsilon(X) := f(X) + \lambda \sum_{i=1}^{l} (\sigma_i(X) + \epsilon)^p,$$

where $\epsilon \geq 0$ and $l = \min(m, n)$.

Algorithm 1: The first IRSVM method for (1)

Let $0 < L_{\min} < L_{\max}$, $\tau > 1$, $c > 0$, integer $N \geq 0$, and $\{\epsilon_k\}$ be a sequence of non-increasing positive scalars converging to 0. Choose an arbitrary $X^0 \in \Re^{m\times n}$ and set $k = 0$.

1) Choose $L_k^0 \in [L_{\min}, L_{\max}]$ arbitrarily. Set $L_k = L_k^0$.

   1a) Apply Theorem 3.1 to find a solution to the subproblem

$$X^{k+1} \in \mathrm{Arg}\min_{X\in\Re^{m\times n}} \left\{ \langle \nabla f(X^k), X - X^k \rangle + \frac{L_k}{2}\|X - X^k\|_F^2 + \lambda p \sum_{i=1}^{l} s_i^k \sigma_i(X) \right\}, \tag{23}$$

   where $s_i^k = (\sigma_i(X^k) + \epsilon_k)^{p-1}$ for all i.

   1b) If

$$\bar{F}_{\epsilon_{k+1}}(X^{k+1}) \leq \max_{[k-N]^+ \leq i \leq k} \bar{F}_{\epsilon_i}(X^i) - \frac{c}{2}\|X^{k+1} - X^k\|_F^2 \tag{24}$$

   is satisfied, then go to step 2).

   1c) Set $L_k \leftarrow \tau L_k$ and go to step 1a).

2) Set $\bar{L}_k = L_k$, $k \leftarrow k + 1$ and go to step 1).

end

The following theorem states that for each outer iteration of the above method, the number of its inner iterations is uniformly bounded. Its proof follows from the fact that $\{\epsilon_k\}$ is non-increasing and a similar argument as used in the proof of [29, Theorem 3.7].

Theorem 3.2 *For each $k \geq 0$, the inner termination criterion* (24) *is satisfied after at most*

$$\max\left\{ \left\lfloor \frac{\log(L_f + c) - \log(L_{\min})}{\log\tau} + 1 \right\rfloor, 1 \right\}$$

inner iterations.

We next establish some convergence result for the outer iterations of our first IRSVM method. As mentioned in section 1, the convergence analysis of the IRL$_1$ method ([29, Algorithm 5]) relies on the explicit expression of the (Clarke) subdifferential of the objective function of its subproblems, which is immediately available due to their simple structure. It is, however, extremely difficult to find the explicit expression of the Clarke subdifferential of the objective of (23). Therefore, the approach for analyzing the IRL$_1$ method is not applicable to this IRSVM method. Our approach below is novel and substantially different from that one. The key idea is to use the expression of a specific solution of (23) to establish a first-order optimality condition of (23) (see (31)), which avoids the intricate issue of finding the explicit expression of the Clarke subdifferential of the objective of (23).

Theorem 3.3 *Suppose that $\{\epsilon_k\}$ is a sequence of non-increasing positive scalars and $\epsilon_k \to 0$ as $k \to \infty$. Let $\{X^k\}$ be generated by the first IRSVM method. Then the following statements hold:*

(i) The sequence $\{X^k\}$ is bounded.

(ii) Let X^* be any accumulation point of $\{X^k\}$. Then X^* is a first-order stationary point of (1)*, i.e.,* (11) *holds at X^*. Moreover, the nonzero entries of X^* satisfy the bound* (15) *with $M = \bar{F}_{\epsilon_0}(X^0)$.*

Proof. (i) Using (24) and an inductive argument, we can conclude that $\bar{F}_{\epsilon_k}(X^k) \le \bar{F}_{\epsilon_0}(X^0)$ for all k. This together with (9) , $\epsilon_k > 0$, and the definition of $\bar{F}_\epsilon$ implies that

$$\underline{f} + \lambda\|X^k\|_p^p \;\le\; f(X^k) + \lambda\sum_{i=1}^{l}(\sigma_i(X^k)+\epsilon_k)^p \;=\; \bar{F}_{\epsilon_k}(X^k) \;\le\; \bar{F}_{\epsilon_0}(X^0).$$

It follows that $\|X^k\|_p^p \;\le\; (\bar{F}_{\epsilon_0}(X^0) - \underline{f})/\lambda$ and hence $\{X^k\}$ is bounded.

(ii) From the proof of statement of (i) and the assumption that $\{\epsilon_k\}$ is non-increasing, we know that

$$\{(X^k,\epsilon_k)\} \subset \Omega = \{(X,\epsilon) \in \Re^{m\times n}\times\Re : \|X\|_p^p \le (\bar{F}_{\epsilon_0}(X^0)-\underline{f})/\lambda, 0 \le \epsilon \le \epsilon_0\}.$$

Observe that $\bar{F}_\epsilon(X)$, viewed as a function of (X,ϵ), is uniformly continuous in Ω. Using this fact, (24), $\{\epsilon_k\} \to 0$ and a similar argument as used in the proof of [43, Lemma 4], one can show that $\|X^{k+1}-X^k\| \to 0$. By the updating rule of Algorithm 1 and Theorem 3.2, we can observe that $\{\bar{L}_k\}$ is bounded, and moreover, X^{k+1} is the solution of subproblem (23) with $L_k = \bar{L}_k$ that is obtained by applying Theorem 3.1. Let $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$ and $U^k\mathrm{Diag}(d^k)(V^k)^T$ the singular value decomposition of Z^k, where $U^k \in \Re^{m\times l}$, $V^k \in \Re^{n\times l}$ satisfy $(U^k)^T U^k = I$ and $(V^k)^T V^k = I$ and $d^k \in \Re^l_+$ consists of all singular values of Z^k arranged in descending order. It then follows from (23) and Theorem 3.1 that

$$X^{k+1} = U^k\mathrm{Diag}(x^{k+1})(V^k)^T, \tag{25}$$

where

$$x^{k+1} = \max(d^k - \lambda p s^k/\bar{L}_k, 0). \tag{26}$$

By the definitions of d^k and s^k, we see that $d_1^k \ge \cdots \ge d_l^k$ and $s_1^k \le \cdots \le s_l^k$. These together with (26) imply that $x_1^{k+1} \ge \cdots \ge x_l^{k+1} \ge 0$. This relation and (25) yield

$$x_i^{k+1} = \sigma_i(X^{k+1}), \quad \forall i. \tag{27}$$

Since X^* is an accumulation point of $\{X^k\}$, there exists a subsequence $\mathcal{K}$ such that $\{X^k\}_{\mathcal{K}} \to X^*$. This together with $\epsilon_k \to 0$ and $s_i^k = (\sigma_i(X^k) + \epsilon_k)^{p-1}$ implies that

$$\{s_i^k\}_{\mathcal{K}} \to \sigma_i(X^*)^{p-1}, \quad i = 1, \ldots, r, \tag{28}$$

where $r = \mathrm{rank}(X^*)$. Due to $\|X^{k+1} - X^k\| \to 0$ and $\{X^k\}_{\mathcal{K}} \to X^*$, we also have $\{X^{k+1}\}_{\mathcal{K}} \to X^*$. This along with (27) leads to

$$\{x_i^{k+1}\}_{\mathcal{K}} \to \sigma_i(X^*), \quad \forall i. \tag{29}$$

One can observe from (29) that there exists some $k_0 > 0$ such that $x_i^{k+1} > 0$ for all $1 \le i \le r$ and $k \in \mathcal{K}_0 = \{j \in \mathcal{K} : j > k_0\}$. It then follows from (26) that

$$x_i^{k+1} = d_i^k - \lambda p s_i^k / \bar{L}_k, \quad 1 \le i \le r, \; k \in \mathcal{K}_0.$$

Hence,

$$\bar{L}_k (x_i^{k+1} - d_i^k) + \lambda p s_i^k = 0, \quad 1 \le i \le r, \; k \in \mathcal{K}_0,$$

which implies that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k \sum_{i=1}^{r} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T + \lambda p \sum_{i=1}^{r} s_i^k U_i^k (V_i^k)^T \;=\; 0. \tag{30}$$

Using (25) and $Z^k = U^k \mathrm{Diag}(d^k)(V^k)^T$, one can see that

$$\sum_{i=1}^{r} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T = X^{k+1} - Z^k - \sum_{i=r+1}^{l} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T,$$

which together with $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$ yields

$$\sum_{i=1}^{r} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T = X^{k+1} - X^k + \frac{1}{\bar{L}_k} \nabla f(X^k) - \sum_{i=r+1}^{l} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T.$$

Substituting it into (30), we obtain that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k (X^{k+1} - X^k) + \nabla f(X^k) - \bar{L}_k \sum_{i=r+1}^{l} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T + \lambda p \sum_{i=1}^{r} s_i^k U_i^k (V_i^k)^T \;=\; 0. \tag{31}$$

Let $\bar{U}^k = [U_1^k \cdots U_r^k]$ and $\bar{V}^k = [V_1^k \cdots V_r^k]$. Upon pre- and post-multiplying (31) by $(\bar{U}^k)^T$ and $\bar{V}^k$, and using $(U^k)^T U^k = (V^k)^T V^k = I$, we see that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k (\bar{U}^k)^T (X^{k+1} - X^k) \bar{V}^k + (\bar{U}^k)^T \nabla f(X^k) \bar{V}^k + \lambda p \cdot \mathrm{Diag}(s_1^k, \ldots, s_r^k) \;=\; 0. \tag{32}$$

Notice that $\{\bar{U}^k\}_{\mathcal{K}}$ and $\{\bar{V}^k\}_{\mathcal{K}}$ are bounded. Considering a convergent subsequence if necessary, assume without loss of generality that $\{\bar{U}^k\}_{\mathcal{K}} \to \bar{U}^*$ and $\{\bar{V}^k\}_{\mathcal{K}} \to \bar{V}^*$. Using (28), the boundedness of $\{\bar{L}_k\}$, $\|X^{k+1} - X^k\| \to 0$, $\{X^k\}_{\mathcal{K}} \to X^*$, and taking limits on both sides of (32) as $k \in \mathcal{K}_0 \to \infty$, we have

$$(\bar{U}^*)^T \nabla f(X^*) \bar{V}^* + \lambda p \cdot \mathrm{Diag}(\sigma(X^*)^{p-1}) \;=\; 0. \tag{33}$$

Observe $(\bar{U}^k)^T\bar{U}^k = (\bar{V}^k)^T\bar{V}^k = I$. Hence, we have $(\bar{U}^*)^T\bar{U}^* = (\bar{V}^*)^T\bar{V}^* = I$. Using (25), (29), $r = \mathrm{rank}(X^*)$, $\{X^{k+1}\}_\mathcal{K} \to X^*$, $\{\bar{U}^k\}_\mathcal{K} \to \bar{U}^*$, and $\{\bar{V}^k\}_\mathcal{K} \to \bar{V}^*$, one can obtain that $X^* = \bar{U}^*\mathrm{Diag}(\sigma(X^*))(\bar{V}^*)^T$. Hence, $(\bar{U}^*, \bar{V}^*) \in \mathcal{M}(X^*)$. Using this relation and (33), we can conclude that (11) holds at X^* with $U = \bar{U}^*$, $V = \bar{V}^*$. Finally, recall that $\bar{F}_{\epsilon_k}(X^k) \le \bar{F}_{\epsilon_0}(X^0)$ for all k. This along with the definition of $\bar{F}_\epsilon(\cdot)$ implies $F(X^k) \le \bar{F}_{\epsilon_k}(X^k) \le \bar{F}_{\epsilon_0}(X^0)$. Taking limits on both sides as $\mathcal{K} \ni k \to \infty$, one has $F(X^*) \le \bar{F}_{\epsilon_0}(X^0)$. The second part of statement (ii) then follows from this relation and Theorem 2.2. ∎

Remark: Let U^k, $\bar{U}^k$, V^k, $\bar{V}^k$, X^*, $\bar{U}^*$, $\bar{V}^*$ and $\mathcal{K}$ be defined as in the proof of Theorem 3.3. We know that $\{\bar{U}^k\}_\mathcal{K} \to \bar{U}^*$, $\{\bar{V}^k\}_\mathcal{K} \to \bar{V}^*$ and $\{X^k\}_\mathcal{K} \to X^*$. It follows from (33) that

$$\lim_{\mathcal{K} \ni k \to \infty} (\bar{U}^k)^T \nabla f(X^k) \bar{V}^k + \lambda p \cdot \mathrm{Diag}((v^k)^{p-1}) \;=\; 0, \tag{34}$$

where $v^k = (\sigma_1(X^k), \ldots, \sigma_r(X^k))^T$ with $r = \mathrm{rank}(X^*)$. Notice that $\{\sigma_i(X^k)\}_\mathcal{K} \to 0$ for all $i \ge r+1$ due to $\{X^k\}_\mathcal{K} \to X^*$ and $r = \mathrm{rank}(X^*)$. It follows that (34) is equivalent to

$$\lim_{\mathcal{K} \ni k \to \infty} \mathrm{Diag}(\sigma(X^k)^{\frac{1}{2}})(U^k)^T \nabla f(X^k) V^k \mathrm{Diag}(\sigma(X^k)^{\frac{1}{2}}) + \lambda p \cdot \mathrm{Diag}(\sigma(X^k)^p) \;=\; 0. \tag{35}$$

Therefore, one suitable termination criterion for the first IRSVM method is

$$\left\| \mathrm{Diag}(\sigma(X^k)^{\frac{1}{2}})(U^k)^T \nabla f(X^k) V^k \mathrm{Diag}(\sigma(X^k)^{\frac{1}{2}}) + \lambda p \cdot \mathrm{Diag}(\sigma(X^k)^p) \right\|_{\max} \le \bar{\varepsilon} \tag{36}$$

for some prescribed accuracy parameter $\bar{\varepsilon}$.

3.2 The second IRSVM method for (1)

In this subsection we extend another efficient IRL$_1$ method (namely, [29, Algorithm 7]) to solve problem (1) and establish a global convergence for the resulting IRSVM method. To proceed, we introduce some notations as follows.

Let $q < 0$ be such that

$$1/p + 1/q = 1. \tag{37}$$

For any $u > 0$, let

$$h_u(t) := \min_{0 \le s \le u} p\left(|t|s - s^q/q\right), \quad \forall t \in \Re.$$

Given any $\epsilon > 0$, define

$$F_\epsilon(X) := f(X) + \lambda \sum_{i=1}^{l} h_{u_\epsilon}(\sigma_i(X)),$$

where $l = \min(m, n)$, and

$$u_\epsilon := \sqrt[q]{\epsilon/(\lambda l)}. \tag{38}$$

By a similar argument as in the proof of [29, Proposition 2.6], it can be shown that F_ϵ is locally Lipschitz continuous, and moreover,

$$0 \;\le\; F_\epsilon(X) - F(X) \;\le\; \epsilon, \qquad \forall X \in \Re^{m \times n}. \tag{39}$$

Given any $X^0 \in \Re^{m\times n}$, we define

$$\varepsilon(X^0) = \left\{ \epsilon : 0 < \epsilon < l\lambda \left[\frac{\sqrt{2L_f[F(X^0)+\epsilon-\underline{f}]}}{\lambda p} \right]^q \right\}. \tag{40}$$

We are now ready to present the second IRSVM method for solving (1), which is an extension of [29, Algorithm 7] proposed for solving the ℓ_p regularized *vector* minimization in the form of (5).

Algorithm 2: The second IRSVM method for (1)

Let $l = \min(m,n)$, $0 < L_{\min} < L_{\max}$, $\tau > 1$, $c > 0$, and integer $N \geq 0$ be given. Let q be defined in (37). Choose an arbitrary $X^0 \in \Re^{m\times n}$ and $\epsilon \in \varepsilon(X^0)$. Set $k = 0$.

1) Choose $L_k^0 \in [L_{\min}, L_{\max}]$ arbitrarily. Set $L_k = L_k^0$.

   1a) Apply Theorem 3.1 to find a solution to the subproblem

   $$X^{k+1} \in \operatorname{Arg}\min_{X\in\Re^{m\times n}} \left\{ \langle \nabla f(X^k), X - X^k\rangle + \frac{L_k}{2}\|X-X^k\|_F^2 + \lambda p \sum_{i=1}^{l} s_i^k \sigma_i(X) \right\}, \tag{41}$$

   where $s_i^k = \min\left\{ (\frac{\epsilon}{\lambda l})^{\frac{1}{q}}, \sigma_i(X^k)^{\frac{1}{q-1}} \right\}$ for all i.

   1b) If

   $$F_\epsilon(X^{k+1}) \leq \max_{[k-N]^+\leq i\leq k} F_\epsilon(X^i) - \frac{c}{2}\|X^{k+1}-X^k\|_F^2 \tag{42}$$

   is satisfied, then go to step 2).

   1c) Set $L_k \leftarrow \tau L_k$ and go to step 1a).

2) Set $\bar{L}_k = L_k$, $k \leftarrow k+1$ and go to step 1).

end

The following result states that for each outer iteration of the above method, the number of its inner iterations is uniformly bounded. Its proof is similar to that of [29, Theorem 4.2].

Theorem 3.4 *For each $k \geq 0$, the inner termination criterion* (42) *is satisfied after at most*

$$\max\left\{ \left\lfloor \frac{\log(L_f + c) - \log(L_{\min})}{\log \tau} + 1 \right\rfloor, 1 \right\}$$

inner iterations.

We next establish some convergence result for the outer iterations of the above IRSVM method. For a similar reason as mentioned in section 1, the approach for analyzing the IRL$_1$ method ([29, Algorithm 7]) is not applicable to this IRSVM method. Our approach below makes use of the expression of a specific solution of (41) to establish a first-order optimality condition of (41) (see (47)). This novel approach avoids the intricate issue of finding the explicit expression of the Clarke subdifferential of the objective of (41).

Theorem 3.5 *Let $\{X^k\}$ be generated by the second IRSVM method. Assume that $\epsilon \in \varepsilon(X^0)$, where $\varepsilon(X^0)$ is defined in* (40). *Then the following statements hold:*

(i) The sequence $\{X^k\}$ is bounded.

(ii) Let X^* be any accumulation point of $\{X^k\}$. Then X^* is a first-order stationary point of (1), *i.e.,* (11) *holds at X^*. Moreover, the nonzero entries of X^* satisfy the bound* (15).

Proof. (i) Using (42) and an inductive argument, we can conclude that $F_\epsilon(X^k) \le F_\epsilon(X^0)$ for all k. This together with (39) implies that $F(x^k) \le F_\epsilon(X^0)$. Using this relation, (1) and (9), we see that $\|X^k\|_p^p \le (F_\epsilon(X^0) - \underline{f})/\lambda$ and hence $\{X^k\}$ is bounded.

(ii) From the proof of statement (i), we know that

$$\{X^k\} \subset \Omega = \{X \in \Re^{m\times n} : \|X\|_p^p \le (F_\epsilon(X^0) - \underline{f})/\lambda\}.$$

Observe that F_ϵ is uniformly continuous in Ω. Using this fact, (24) and a similar argument as used in the proof of [43, Lemma 4], one can show that $\|X^{k+1} - X^k\| \to 0$. By the updating rule of Algorithm 2 and Theorem 3.4, we can observe that $\{\bar{L}_k\}$ is bounded, and moreover, X^{k+1} is the solution of subproblem (41) with $L_k = \bar{L}_k$ that is obtained by applying Theorem 3.1. Let $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$ and $U^k \mathrm{Diag}(d^k)(V^k)^T$ the singular value decomposition of Z^k, where $U^k \in \Re^{m\times l}$, $V^k \in \Re^{n\times l}$ satisfy $(U^k)^T U^k = I$ and $(V^k)^T V^k = I$ and $d^k \in \Re^l_+$ consists of all singular values of Z^k arranged in descending order. It then follows from (41) and Theorem 3.1 that

$$X^{k+1} = U^k \mathrm{Diag}(x^{k+1})(V^k)^T, \tag{43}$$

where

$$x^{k+1} = \max(d^k - \lambda p s^k/\bar{L}_k, 0). \tag{44}$$

By the definitions of d^k and s^k, we see that $d_1^k \ge \cdots \ge d_l^k$ and $s_1^k \le \cdots \le s_l^k$. These together with (44) imply that $x_1^{k+1} \ge \cdots \ge x_l^{k+1} \ge 0$. This relation and (43) yield

$$x_i^{k+1} = \sigma_i(X^{k+1}), \quad \forall i. \tag{45}$$

Since X^* is an accumulation point of $\{X^k\}$, there exists a subsequence $\mathcal{K}$ such that $\{X^k\}_\mathcal{K} \to X^*$. Due to $\|X^{k+1} - X^k\| \to 0$ and $\{X^k\}_\mathcal{K} \to X^*$, we also have $\{X^{k+1}\}_\mathcal{K} \to X^*$. This along with (45) leads to

$$\{x_i^{k+1}\}_\mathcal{K} \to \sigma_i(X^*), \quad \forall i. \tag{46}$$

Let $r = \mathrm{rank}(X^*)$. One can observe from (46) that there exists some $k_0 > 0$ such that $x_i^{k+1} > 0$ for all $1 \le i \le r$ and $k \in \mathcal{K}_0 = \{j \in \mathcal{K} : j > k_0\}$. Using this fact, (44) and a similar argument as used in the proof of Theorem 3.3, we can show that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k(\bar{U}^k)^T(X^{k+1} - X^k)\bar{V}^k + (\bar{U}^k)^T \nabla f(X^k)\bar{V}^k + \lambda p \cdot \mathrm{Diag}(s_1^k, \ldots, s_r^k) \ = \ 0, \tag{47}$$

where

$$\bar{U}^k = [U_1^k \cdots U_r^k], \qquad \bar{V}^k = [V_1^k \cdots V_r^k].$$

Notice that $\{\bar{U}^k\}_\mathcal{K}$ and $\{\bar{V}^k\}_\mathcal{K}$ are bounded. Considering a convergent subsequence if necessary, assume without loss of generality that $\{\bar{U}^k\}_\mathcal{K} \to \bar{U}^*$ and $\{\bar{V}^k\}_\mathcal{K} \to \bar{V}^*$. Using (46), the boundedness of $\{\bar{L}_k\}$, $\|X^{k+1}-X^k\| \to 0$, and $\{X^k\}_{\mathcal{K}} \to X^*$, and taking limits on both sides of (47) as $k \in \mathcal{K}_0 \to \infty$, we have

$$(\bar{U}^*)^T \nabla f(X^*) \bar{V}^* + \lambda p \cdot \mathrm{Diag}(s_1^*, \ldots, s_r^*) \;=\; 0, \tag{48}$$

where

$$s_i^* = \min\left\{ \left(\frac{\epsilon}{\lambda l}\right)^{\frac{1}{q}}, \sigma_i(X^*)^{\frac{1}{q-1}} \right\}, \quad i = 1, \ldots, r. \tag{49}$$

Similar to the proof of Theorem 3.3, one can show that $(\bar{U}^*, \bar{V}^*) \in \mathcal{M}(X^*)$. Recall that $F_\epsilon(X^k) \leq F_\epsilon(X^0)$ for all k. Using this relation, the continuity of F_ϵ and $\{X^k\}_{\mathcal{K}} \to X^*$, one has $F_\epsilon(X^*) \leq F_\epsilon(X^0)$, which together with (39) yields

$$f(X^*) \leq F(X^*) \leq F_\epsilon(X^*) \leq F_\epsilon(X^0) \leq F(X^0) + \epsilon. \tag{50}$$

Using this relation and the similar arguments as for deriving (17), we see that (17) also holds for such X^*. It then follows from (17), (48), $(\bar{U}^*)^T\bar{U}^* = (\bar{V}^*)^T\bar{V}^* = I$ and $s_i^* > 0$ for every i that for $1 \leq i \leq r$,

$$\begin{aligned} s_i^* &= \frac{1}{\lambda p}\left|[(\bar{U}^*)^T \nabla f(X^*)\bar{V}^*]_{ii}\right| \;\leq\; \frac{1}{\lambda p}\|(\bar{U}^*)^T \nabla f(X^*)\bar{V}^*]\|_F \\ &\leq \frac{1}{\lambda p}\|\nabla f(X^*)\|_F \;\leq\; \frac{\sqrt{2L_f[F(x^0)+\epsilon-\underline{f}]}}{\lambda p}. \end{aligned} \tag{51}$$

We now claim that $\sigma_i(X^*) > {u_\epsilon}^{q-1}$ for all $1 \leq i \leq r$, where u_ϵ is defined in (38). Suppose for contradiction that there exists some $1 \leq i \leq r$ such that $0 < \sigma_i(X^*) \leq {u_\epsilon}^{q-1}$. It then follows from (49) that $s_i^* = u_\epsilon = \sqrt[q]{\frac{\epsilon}{\lambda l}}$. Using this and (51), we have

$$\left(\frac{\epsilon}{\lambda l}\right)^{\frac{1}{q}} \;\leq\; \frac{\sqrt{2L_f[F(x^0)+\epsilon-\underline{f}]}}{\lambda p},$$

which contradicts with the assumption $\epsilon \in \varepsilon(X^0)$. Therefore, $\sigma_i(X^*) > {u_\epsilon}^{q-1}$ for all $1 \leq i \leq r$. Using this relation, (37) and (49), we see that

$$s_i^* = \sigma_i(X^*)^{p-1}, \quad i = 1, \ldots, r.$$

Substituting it into (48), we obtain that

$$(\bar{U}^*)^T \nabla f(X^*)\bar{V}^* + \lambda p \; \mathrm{Diag}(\sigma(X^*)^{p-1}) \;=\; 0.$$

Using this relation and $(\bar{U}^*, \bar{V}^*) \in \mathcal{M}(X^*)$, we can conclude that (11) holds at X^* with $U = \bar{U}^*$, $V = \bar{V}^*$. Finally, recall from (50) that $F(X^*) \leq F(X^0) + \epsilon$. Using this relation and Theorem 2.2, we immediately see that the second part of statement (ii) also holds. ■

Remark: Let U^k, V^k and $\mathcal{K}$ be defined as in the proof of Theorem 3.5. By a similar argument as for the first IRSVM method, one can show that (35) also holds for the second IRSVM method with the above U^k, V^k and $\mathcal{K}$. Therefore, (36) with these U^k and V^k can also be used as a termination criterion for the second IRSVM method.

4 Numerical results

In this section we conduct numerical experiments to test the performance of the IRSVM methods (Algorithms 1 and 2) proposed in Section 3. In particular, we apply them to solve the problem (1) with $f(X) = \|\mathcal{P}_\Omega(X - M)\|_F^2$, that is

$$\min_{X \in \Re^{m\times n}} \left\{ \|\mathcal{P}_\Omega(X - M)\|_F^2 + \lambda \|X\|_p^p \right\}, \tag{52}$$

where $M \in \Re^{m\times n}$, Ω is a subset of index pairs (i, j), and $\mathcal{P}_\Omega(\cdot)$ is the projection onto the subspace of sparse matrices with nonzeros restricted to the index subset Ω. Problem (52) and its variant have applications in matrix completion, which aims to fill in the missing entries of a partially observed low-rank matrix M with the known entries $\{M_{ij} : (i, j) \in \Omega\}$ (see, for example, [5, 33, 19, 26, 41]).

Notice that the IRSVM methods are an extension of the iterative reweighted l_1 methods [29] from problem (5) to problem (1), while the iterative reweighted least squares methods, namely, IRucLq-M and tIRucLq-M [26], are an extension of the iterative reweighted l_2 methods from problem (5) to problem (1) with h and f being a convex quadratic function. It is thus plausible to compare our methods with IRucLq-M and tIRucLq-M. For convenience of presentation, we name the IRSVM methods as IRSVM-1 (Algorithm 1) and IRSVM-2 (Algorithm 2), respectively. In addition, the codes of all methods tested in this section are written in MATLAB and all experiments are performed in MATLAB 7.14.0 (2012a) on a desktop with an Intel Core i7-3770 CPU (3.40 GHz) and 16GB RAM running 64-bit Windows 7 Enterprise (Service Pack 1).

We terminate IRSVM-1 and IRSVM-2 according to the criterion (36). In addition, we apply a continuation technique to IRSVM-1 and IRSVM-2, which is similar to the one used in APGL [39]. In detail, set λ to be the target parameter, and let $\{\lambda_0, \lambda_1, \ldots, \lambda_\ell = \lambda\}$ be a set of parameters in descending order. We start with $X^0 = \mathcal{P}_\Omega(M)$ and apply a method to problem (1) with λ replaced by λ_0 to find an approximate solution, denoted by $X^{(0)}$. Then we use $X^{(0)}$ as the initial point and apply the same method to (1) with λ replaced by λ_1 to obtain an approximate solution, denoted by $X^{(1)}$. This process is repeated until the target parameter λ is reached and its approximate solution is found.

For IRSVM-1 and IRSVM-2, we set $L_{\min} = 10^{-2}$, $L_{\max} = 1$, $c = 10^{-4}$, $\tau = 2$, $N = 10$, and $L_0^0 = 1$. And we update L_k^0 by the same strategy as used in [1, 4, 43], that is,

$$L_k^0 = \max\left\{ L_{\min}, \min\left\{ L_{\max}, \frac{\mathrm{tr}(\Delta X \Delta G^T)}{\|\Delta X\|_F^2} \right\} \right\},$$

where $\Delta X = X^k - X^{k-1}$ and $\Delta G = \nabla f(X^k) - \nabla f(X^{k-1})$. In addition, we set $\epsilon_k = 0.5^k$ for IRSVM-1. For IRSVM-2, ϵ is chosen to be the one within 10^{-6} to the supremum of $\varepsilon(X^0)$ that is defined in (40) with $\underline{f}$ being replaced by 0.

Given an approximate recovery X^* for M, the relative error is defined as

$$\text{rel\_err} := \frac{\|X^* - M\|_F}{\|M\|_F}.$$

We adopt the same criterion as used in [37, 6], and say a matrix M is *successfully recovered* by X^* if the corresponding relative error is less than 10^{-3}.

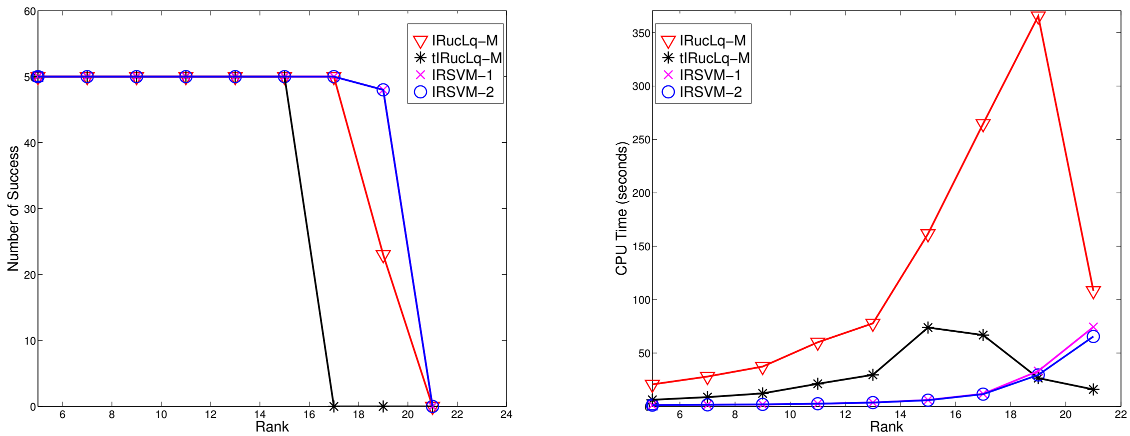


Figure 1: Comparison on random data with $SR = 0.2$.

4.1 Numerical results on random data

In this subsection we conduct numerical experiments to test the performance of IRSVM-1 and IRSVM-2 for solving (52) on random data. We also compare our methods with IRucLq-M and tIRucLq-M [26].

We aim to recover a random matrix $M \in \Re^{m\times n}$ with rank r based on a subset of entries $\{M_{ij}\}_{(i,j)\in\Omega}$. For this purpose, we randomly generate M and Ω by a similar procedure as described in [31]. In detail, we first generate random matrices $M_L \in \Re^{m\times r}$ and $M_R \in \Re^{n\times r}$ with i.i.d. standard Gaussian entries and let $M = M_L M_R^T$. We then sample a subset Ω with sampling ratio SR uniformly at random, where $SR = |\Omega|/(mn)$. In our experiment, we set $m = n = 200$ and generate Ω with three different values of SR, which are 0.2, 0.5 and 0.8.

For each sample ratio SR and rank r, we apply IRSVM-1, IRSVM-2, IRucLq-M and tIRucLq-M to solve (52) on 50 instances that are randomly generated above. In particular, we set a limit of 5000 on maximum number of iterations for all methods. In addition, we set tol $= 10^{-6}$, $K = \lfloor 1.5r \rfloor$, $\lambda = 10^{-6}$, $q = 0.5$ for IRucLq-M and tIRucLq-M. All other parameters for these two methods are chosen by default. For IRSVM-1 and IRSVM-2, we choose $p = 0.5$, $\bar{\varepsilon} = 10^{-3}$, $\lambda_0 = 10$, and $\lambda_\ell = \max(0.1\lambda_{\ell-1}, 10^{-6})$ for $\ell \geq 1$. The computational results are presented in Figures 1-3. In detail, the left and right plots in each figure show the number of successfully recovered matrices and the CPU time (in seconds) of each method, respectively. One can see that the recoverability of IRSVM-1 and IRSVM-2 is generally better than that of IRucLq-M and tIRucLq-M. For example, when $SR = 0.2$ and $r = 19$, IRSVM-1 and IRSVM-2 are capable of recovering almost all instances while tIRucLq-M recovers none of the instances and IRucLq-M only recovers about half of the instances. For the instances that are successfully recovered by all four methods, the CPU time of IRSVM-1 and IRSVM-2 is less than that of the other two methods. Nevertheless, for the instances that fail to be recovered by IRucLq-M or tIRucLq-M, the CPU time of these methods is less than that of our two methods. In addition, we observe that IRSVM-2 is slightly faster than IRSVM-1 for these instances.

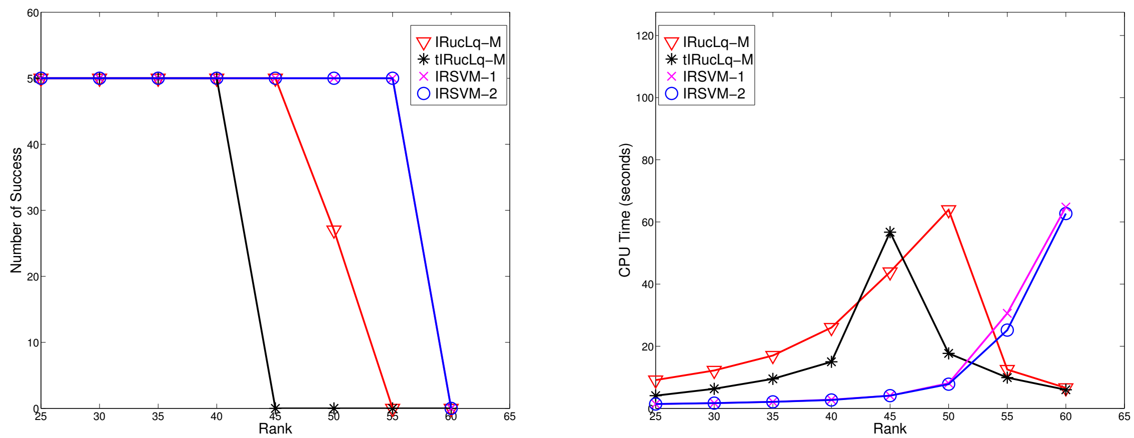


Figure 2: Comparison on random data with $SR = 0.5$.

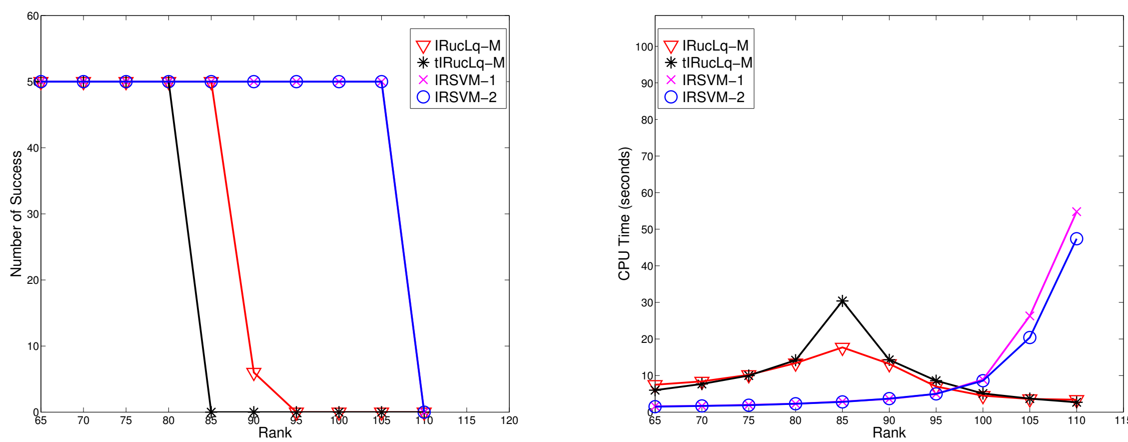
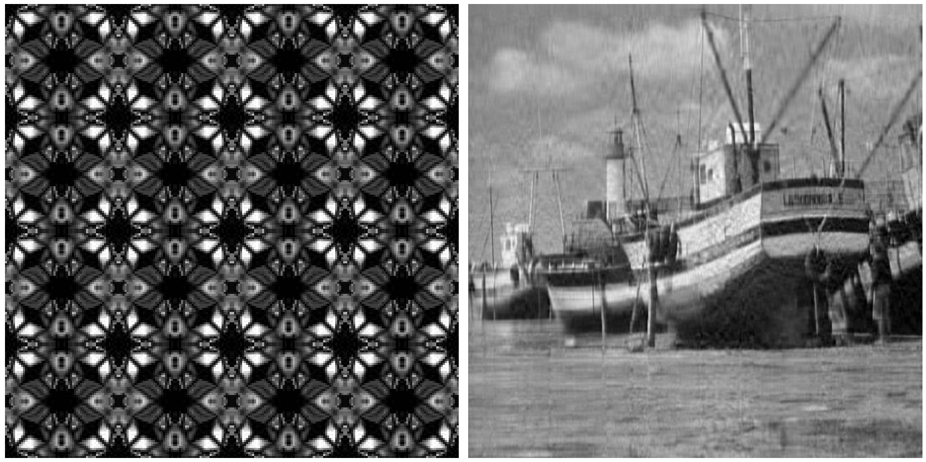


Figure 3: Comparison on random data with $SR = 0.8$.

4.2 Numerical results on image data

In this subsection we compare the performance of IRSVM-1 and IRSVM-2 with IRucLq-M and tIRucLq-M for solving a grayscale image inpainting problem [2], which was used in [26] for testing IRucLq-M and tIRcuLq-M. For an image inpainting problem, the goal is to fill the missing pixel values of the image at given pixel locations. As shown in [40, 34], this problem can be solved as a matrix completion problem if the image is of low-rank. In our test, we consider two different grayscale images which are "pattern" and "boat" as shown in Figure 4. In detail, "pattern" is a texture image with 224×224 pixels and rank 28. The image "boat" is obtained by first applying the singular value decomposition to the original image with 512×512 pixels and then truncating the decomposition so that the resulting image has rank 40.

We first apply all four methods to solve the image inpainting problem with three different sample



(a) pattern (b) boat

Figure 4: Testing images. “pattern”: grayscale image of 224×224 pixels with rank = 28. “boat”: grayscale image of 512×512 pixels with rank = 40.

Table 1: Results of image recovery (best rel_err in boldface).

		IRucLq-M		tIRucLq-M		IRSVM-1		IRSVM-2	
Image	SR	rel_err	Time	rel_err	Time	rel_err	Time	rel_err	Time
“pattern”	0.1	4.02e−1	5.6	5.25e−1	1.4	3.19e−1	6.9	**3.18e−1**	5.1
	0.2	1.27e−1	7.6	2.29e−1	2.9	**9.08e−2**	2.7	9.19e−2	2.2
	0.3	4.53e−2	5.1	7.03e−2	2.2	**3.13e−2**	1.5	3.23e−2	1.4
“boat”	0.1	1.66e−1	158.4	2.34e−1	9.3	**1.58e−1**	45.1	**1.58e−1**	42.3
	0.2	3.72e−2	148.0	5.52e−2	10.5	1.39e−2	29.7	**1.16e−2**	33.0
	0.3	4.61e−3	97.8	5.84e−3	8.6	2.64e−3	9.4	**1.21e−3**	10.3

ratios ($SR = 0.1$, 0.2, and 0.3). In particular, we set maxrank and K equal to the rank of the testing image, and tol=10^{-3} for IRucLq-M and tIRucLq-M. In addition, we set $\bar{\varepsilon} = 5 \times 10^{-3}$ for IRSVM-1 and IRSVM-2. The other parameter settings for all methods are the same as those used in the random data experiment. We present the results of this experiment in Table 1 and Figures 5 and 6. In Table 1, the name of the images and the sample ratio SR are given in the first two columns. The results of all the methods in terms of relative error and CPU time (in seconds) are reported in columns three to ten. In Figures 5 and 6, we display the sample images in the first column and the recovered images by different methods in the rest columns. One can observe that IRSVM-1 and IRSVM-2 achieve smaller rel_err than the other two methods. The CPU time of our two methods is generally less than that of IRucLq-M. Though tIRucLq-M outperforms our methods in terms of CPU time, their rel_err is much higher than that of our methods.

5 Concluding remarks

In this paper we considered the ℓ_p (or Schatten-p quasi-norm) regularized low-rank approximation problems. In particular, we studied the first-order optimality condition for them and also derived lower bounds for nonzero singular values of their first-order stationary points. We also proposed the iterative reweighted singular value minimization (IRSVM) methods for solving them and established

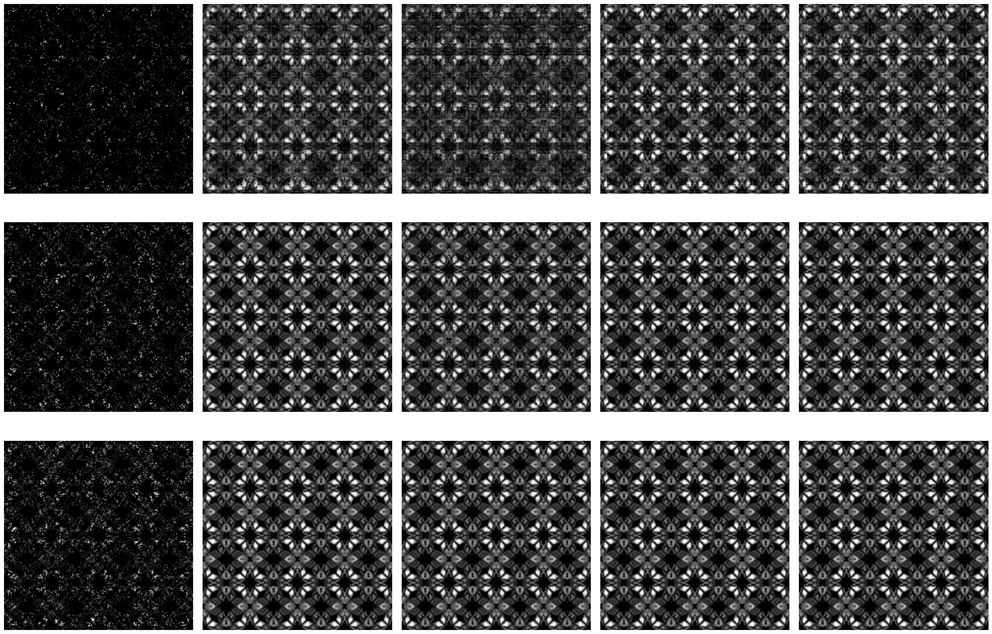


Figure 5: Results of image recovery. First column: the sample images with $SR = 0.1$, 0.2 and 0.3. The rest columns: the images recovered by IRucLq-M, tIRucLq-M, IRSVM-1 and IRSVM-2, respectively.

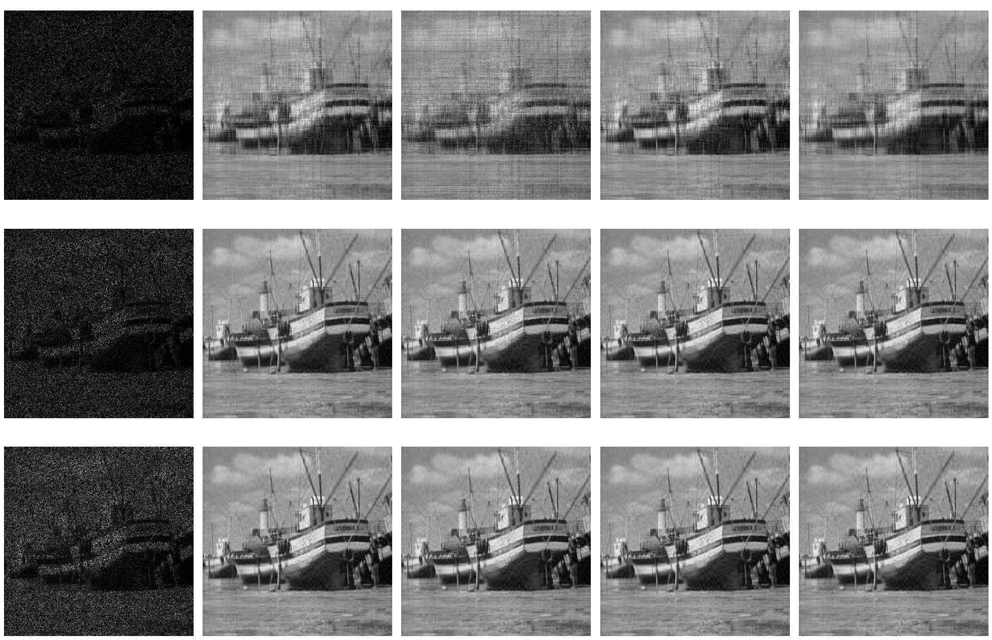


Figure 6: Results of image recovery. First column: the sample images with $SR = 0.1$, 0.2 and 0.3. The rest columns: the images recovered by IRucLq-M, tIRucLq-M, IRSVM-1 and IRSVM-2, respectively.

their convergence by a novel approach that makes use of the expression of a specific solution of their subproblems and avoids the intricate issue of finding the explicit expression for the Clarke subdifferential of the objective of their subproblems. The computational results demonstrate that the IRSVM methods generally outperform the iterative reweighted least squares methods [26] in terms of solution quality and/or speed.

Besides the ℓ_p quasi-norm regularizer, there are some other popular nonconvex regularizers for producing a sparse solution of a system or a vector optimization problem (e.g., see [17, 7, 46, 47, 21] and references therein). They can be extended to find a low-rank solution of a matrix optimization problem. Though we only studied the ℓ_p regularized matrix minimization problems, most of the results and ideas presented in this paper can be moderately modified for the matrix minimization problems with other regularizers by using the similar techniques developed in this paper.

References

[1] J. Barzilai and J. M. Borwein. Two point step size gradient methods. *IMA J. Numer. Anal.*, 8:141–148, 1988.

[2] M. Bertalmío, G. Sapiro, V. Caselles and C. Ballester. Image inpainting. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, 2000.

[3] W. Bian and X. Chen. Smoothing SQP algorithm for non-Lipschitz optimization with complexity analysis. Preprint, February 2012.

[4] E. G. Birgin, J. M. Martínez, and M. Raydan. Nonmonotone spectral projected gradient methods on convex sets. *SIAM J. Optimiz*, 4:1196–1211, 2000.

[5] J.-F. Cai, E. J. Candés, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20, pp. 1956–1982, 2010.

[6] E. J. Candés and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 2009.

[7] E. J. Candés, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.*, 14(5–6): 877–905, 2008.

[8] R. Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Process. Lett.*, 14:707–710, 2007.

[9] R. Chartrand and V. Staneva. Restricted isometry properties and nonconvex compressive sensing. *Inverse Probl.*, 24:1–14, 2008.

[10] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressive sensing. *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

[11] X. Chen, D. Ge, Z. Wang, and Y. Ye. Complexity of unconstrained L_2-L_p minimization. *Math. Program.*, 143(1):371–383, 2014.

[12] X. Chen, L. Niu, and Y. Yuan. Optimality conditions and a smoothing trust region Newton method for non-Lipschitz optimization. *SIAM J. Optim.*, 23(3):1528–1552, 2013.

[13] X. Chen, F. Xu, and Y. Ye. Lower bound theory of nonzero entries in solutions of l_2-l_p minimization. *SIAM J. Sci. Comput.*, 32: 2832–2852, 2010.

[14] X. Chen and W. Zhou. Convergence of reweighted l_1 minimization algorithm for l_2-l_p minimization. *Comp. Optim. Appl.*, 59:47–61, 2014.

[15] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Gunturk. Iteratively reweighted least squares minimization for sparse recovery. *Commun. Pure. Appl. Math.*, 63:1–38, 2010.

[16] M. Fazel, H. Hindi, and S. P. Boyd. A rank minimization heuristic with application to minimum order system approximation. *P. Amer. Contr. Conf.*, 6:4734-4739, 2001.

[17] J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.*, 96(456):1348–1360, 2001.

[18] S. Foucart and M. Lai. Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$. *Appl. Comput. Harmon. Anal.*, 26:395–407, 2009.

[19] M. Fornasier, H. Rauhut, and R. Ward. Low-rank matrix recovery via iteratively reweighted least squares minimization. *SIAM J. Optim.*, 21:1614–1640, 2011.

[20] D. Ge, X. Jiang, and Y. Ye. A note on the complexity of L_p minimization. *Math. Program.*, 21: 1721–1739, 2011.

[21] P. Gong, C. Zhang, Z. Lu, J. Huang, and J. Ye. A general iterative shrinkage and thresholding algorithm for non-convex regularized optimization problems. *The 30th International Conference on Machine Learning (ICML)*, 2013.

[22] S. Ji, K.-F. Sze, Z. Zhou, A. So, and Y. Ye. Beyond convex relaxation: A polynomialtime nonconvex optimization approach to network localization. INFOCOM 2013.

[23] L. Kong and N. Xiu. Exact low-rank matrix recovery via nonconvex Schatten p-minimization. *Asia-Pacific Journal of Operational Research*, 30(3):1340010, 2013.

[24] M. Lai, S. Li, L. Y. Liu and H. Wang. Two results on the schatten p-quasi-norm minimization for low-rank matrix recovery. Preprint, 2012.

[25] M. Lai and J. Wang. An unconstrained l_q minimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems. *SIAM J. Optim.*, 21:82–101, 2010.

[26] M. Lai, Y. Xu, and W. Yin. Improved iteratively reweighted least squares for unconstrained smoothed l_q minimization. *SIAM J. Numer. Anal.*, 5(2):927–957, 2013.

[27] A. S. Lewis and H. S. Sendov. Nonsmooth analysis of singular values. Part I: Theory. *Set-Valued Anal.*, 13: 213-241, 2005.

[28] A. S. Lewis and H. S. Sendov. Nonsmooth analysis of singular values. Part II: Applications. *Set-Valued Anal.*, 13: 243-264, 2005.

[29] Z. Lu. Iterative reweighted minimization methods for l_p regularized unconstrained nonlinear programming. *Math. Program.*, 147(1-2): 277-307, 2014.

[30] Z. Lu, Y. Zhang, and X. Li. Penalty decomposition methods for rank minimization. *Optim. Method Softw.*, 30(3): 531–558, 2015.

[31] S. Ma, D. Goldfarb, and L. Chen. Fixed point and Bregman iterative methods for matrix rank minimization. *Math. Program.*, 128(1):321-353, 2011.

[32] M. Malek-Mohammadi, M. Babaie-Zadeh, and M. Skoglund. Performance guarantees for Schatten-p quasi-norm minimization in recovery of low-rank matrices. *Signal Processing*, 114:225–230, 2015.

[33] K. Mohan and M. Fazel. Iterative reweighted least squares for matrix rank minimization. *48th Annual Allerton Conference on Communication, Control, and Computing*, 653-661, 2010.

[34] T. Mrita and T. Kanade. A sequential factorization method for recovering shape and motion from image streams. *IEEE T. Pattern Anal.*, 19:858-867, 1997.

[35] N. Mourad and J. P. Reilly. Minimizing nonconvex functions for sparse vector reconstruction. *IEEE Trans. Signal Process.*, 58:3485–3496 , 2010.

[36] B. D. Rao and K. Kreutz-Delgado. An affine scaling methodology for best basis selection. *IEEE Trans. Signal Process.*, 47:187–200, 1999.

[37] B. Recht, M. Fazel, and P. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[38] Q. Sun. Recovery of sparsest signals via l_q minimization. *Appl. Comput. Harmon. Anal.*, 32:329–341, 2012.

[39] K. Toh and S. Yun. An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems. *Pac. J. Optim.*, 6:615–640, 2010.

[40] C. Tpmasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *Int. J. Comput. Vision*, 9:137–154, 1992.

[41] Z. Wang, M.-J. Lai, Z. Lu, W. Fan, H. Davulcu, and J. Ye. Orthogonal rank-one matrix pursuit for low rank matrix completion. *SIAM J. Sci. Comput.*, 37: A488–A514, 2015.

[42] Z. Wen, W. Yin, and Y. Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Math. Prog. Comp.*, 4:333-361, 2012.

[43] S. J. Wright, R. Nowak, and M. Figueiredo. Sparse reconstruction by separable approximation. *IEEE T. Image Process.*, 57:2479–2493, 2009.

[44] Z. Xu, H. Zhang, Y. Wang, and X. Chang. $L_{1/2}$ regularizer. *Sci. China*, 52:1–9, 2009.

[45] M. Yue and A.M. So. A perturbation inequality for the schatten-p quasi-norm and its applications to low-rank matrix recovery. Preprint, 2014.

[46] C. H. Zhang. Nearly unbiased variable selection under minimax concave penalty *Ann. Stat.*, 38(2): 894–942, 2010.

[47] T. Zhang. Analysis of multi-stage convex relaxation for sparse regularization. *J. Mach. Learn. Res.*, 11:1081–1107, 2010.