# Iterative Reweighted Singular Value Minimization Methods for $l_p$ Regularized Unconstrained Matrix Minimization*

Zhaosong Lu† Yong Zhang ‡

January 4, 2014

**Abstract**

In this paper we study general $l_p$ regularized unconstrained matrix minimization problems. In particular, we first introduce a class of first-order stationary points for them. And we show that the first-order stationary points introduced in [11] for an $l_p$ regularized *vector* minimization problem are equivalent to those of an $l_p$ regularized *matrix* minimization reformulation. We also establish that any local minimizer of the $l_p$ regularized matrix minimization problems must be a first-order stationary point. Moreover, we derive lower bounds for nonzero singular values of the first-order stationary points and hence also of the local minimizers for the $l_p$ matrix minimization problems. The iterative reweighted singular value minimization (IRSVM) approaches are then proposed to solve these problems in which each subproblem has a closed-form solution. We show that any accumulation point of the sequence generated by these methods is a first-order stationary point of the problems. In addition, we study a nonmontone proximal gradient (NPG) method for solving the $l_p$ matrix minimization problems and establish its global convergence. Our computational results demonstrate that the IRSVM and NPG methods generally outperform some existing state-of-the-art methods in terms of solution quality and/or speed. Moreover, the IRSVM methods are slightly faster than the NPG method.

**Key words:** $l_p$ regularized matrix minimization, iterative reweighted singular value minimization, iterative reweighted least squares, nonmonotone proximal gradient method

*This work was supported in part by NSERC Discovery Grant.

†Department of Mathematics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. (email: zhaosong@sfu.ca).

‡Department of Mathematics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. (email: yza30@sfu.ca).

# 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 been proposed for it (e.g., see [15, 32, 27, 34, 36, 28, 18, 23, 26]). In this paper we are interested in one of those models, namely, the $l_p$ regularized unconstrained nonlinear matrix optimization 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 $p \in (0,1)$, where $f$ is 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 $\Re^{m \times n}$. 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 $i$th largest singular value of $X$. 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 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 [27, 34]). Since problem (1) is intermediate between problems (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 [23, 21].

One can observe that if $f$ only depends on the diagonal entries of $X$, problem (1) is equivalent to an $l_p$ regularized unconstrained vector minimization problem in the form of

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

where $\|x\|_p := (\sum_{i=1}^l |x_i|^p)^{1/p}$ 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 [7, 8, 9, 17, 38, 22, 12, 13, 14, 30, 10, 19, 33, 3, 11]). Efficient iterative reweighted $l_1$ ($\mathrm{IRL}_1$) and $l_2$ ($\mathrm{IRL}_2$) minimization algorithms were also proposed for finding an approximate solution to (5) or its variants (e.g., see [31, 9, 17, 14, 22, 13, 23, 25]).

Recently, Ji et al. [21] proposed an interior-point method for finding an approximate solution to a class of $l_p$ regularized matrix optimization problems arising in network localization. Lai et al. [23] proposed an iterative reweighted least squares (IRLS) method for solving (1) with a convex quadratic function $f$, which is an extension of the IRL$_2$ method that was proposed in the same paper for (5). In addition, Fornasier et al. [18], and Mohan and Fazel [28] extended the IRL$_2$ method, which was proposed in [14] for solving the linear constrained $l_p$ regularized vector optimization problems, to solve the corresponding matrix optimization problems, that is,

$$\min_{X \in \Re^{m\times n}} \{\|X\|_p^p : \mathcal{A}(X) = b\},$$

where $\mathcal{A} : \Re^{m\times n} \to \Re^l$ is a linear operator, and $b \in \Re^l$. In each iteration, these methods solve a certain convex quadratic programming problem that has a closed-form solution. They are so called iterative reweighted least squares (IRLS) methods in the literature. It is well-known that the IRL$_2$ methods when applied to (5) or its variants generally do not produce a sparse solution. Similarly, the IRLS methods usually does not generate a low-rank solution either, and thus a post-processing scheme based on singular value thresholding is often used to obtain a low-rank solution. Unlike the IRL$_2$ methods, the IRL$_1$ methods tend to produce a sparse solution when applied to (5) or its variants. One can naturally expect that the iterative reweighted singular value minimization (IRSVM) methods, namely, the counterpart of the IRL$_1$ methods for (1), are also capable of yielding a low-rank solution of (1) without the aid of a post-processing. Nevertheless, the extension of the IRL$_1$ methods to solve (1) or its variants have not yet been studied in the literature.

In this paper we first introduce a class of first-order stationary points to problem (1). We show that the first-order stationary points introduced in [11] for (5) are equivalent to those of an $l_p$ regularized matrix minimization problem. Also, we establish that any local minimizer of problem (1) must be a stationary point. Moreover, we derive lower bounds for nonzero singular values of the first-order stationary points and hence also of the local minimizers of (1). Then we extend two IRL$_1$ methods proposed in [25] to solve (1), and the resulting IRSVM methods per iteration solve a weighted singular value minimization subproblem which has a closed-form solution. And we show that any accumulation point of the sequence generated by these methods is a first-order stationary point of problem (1). In addition, we study a nonmontone proximal gradient (NPG) method for solving (1) and establish its global convergence. Finally we conduct numerical experiments to compare the IRSVM and NPG methods with some state-of-the-art methods in the literature. The computational results demonstrate that the IRSVM and NPG methods generally outperform those methods in terms of solution quality and/or speed. Moreover, the IRSVM methods are slightly faster than the NPG method.

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 [25] to solve (1), and establish their convergence. In addition, we study a nonmontone proximal gradient method for solving (1) in section 4 and establish its convergence. We conduct numerical experiments in section 5 to compare the IRSVM and NPG methods with some state-of-the-art methods proposed in the literature. Finally, in section 6 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 that $x \in \Re^n_+$ (resp., $x \in \Re^n_{++}$). Given any $x \in \Re^n$ and a scalar $\alpha$, $|x|^\alpha$ (resp., $x^\alpha$) denotes an $n$-dimensional vector whose $i$th component is $|x_i|^\alpha$ (resp., $x_i^\alpha$). Given an index set $\mathcal{B} \subseteq \{1,\ldots,n\}$, $x_\mathcal{B}$ denotes the sub-vector of $x$ indexed by $\mathcal{B}$. $\mathrm{Diag}(x)$ or $\mathrm{Diag}(x_1,\ldots,x_n)$ denotes an $n \times n$ diagonal matrix whose diagonal is formed by the vector $x$. 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, $\|X\|_F = \sqrt{\mathrm{tr}(XX^T)}$, where $\mathrm{tr}(\cdot)$ 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 $i$th largest singular value of $X$ for $i = 1,\ldots,\min(m,n)$, $\sigma(X) = (\sigma_1(X),\ldots,\sigma_r(X))^T$, and

$$\mathcal{M}(X) = \{(U,V) \in \Re^{m\times r} \times \Re^{n\times r} : U^TU = V^TV = I, X = U\mathrm{Diag}(\sigma(X))V^T\}, \tag{6}$$

where $r = \mathrm{rank}(X)$. 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, $\langle X, Y\rangle = \mathrm{tr}(XY^T)$. If a symmetric matrix $X$ is positive semidefinite (resp., definite), we write $X \succeq 0$ (resp., $X \succ 0$). Given a positive definite diagonal matrix $D$ and a scalar $\alpha$, we define $D^\alpha$ as a diagonal matrix whose $i$th diagonal entry is $(D_{ii})^\alpha$, that is, $[D^\alpha]_{ii} = (D_{ii})^\alpha$ for all $i$. In addition, for a square matrix $X$, $\mathrm{diag}(X)$ denotes the vector extracted from its diagonal.

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

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

Also, we define

$$\underline{f} \;=\; \inf_{X \in \Re^{m\times n}} f(X). \tag{7}$$

It follows from the 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). Then we show that the first-order stationary points of (5) as defined in [12] are equivalent to those of an $l_p$ regularized matrix minimization reformulation. Also, we 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).

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

$$0 \in \left\{ U^T \nabla f(X^*) V + \lambda p \mathrm{Diag}(\sigma(X^*)^{p-1}) : (U,V) \in \mathcal{M}(X^*) \right\}. \tag{8}$$

Chen et al. [11] 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

$$X^* \nabla h(x^*) + \lambda p |x^*|^p = 0, \tag{9}$$

where $X^* = \mathrm{Diag}(x^*)$. It is not hard to observe that (5) is equivalent to the matrix minimization problem

$$\min_{X\in\Re^{l\times l}} \left\{ \widehat{h}(X) + \lambda \|X\|_p^p \right\}, \tag{10}$$

where $\widehat{h}(X) = h(\mathrm{diag}(X))$ for all $X \in \Re^{l\times l}$. As we later show, the aforementioned first-order stationary points of (5) are equivalent to those given in Definition 1 for the matrix minimization problem (10). Therefore, our definition of the first-order stationary points for (1) is a natural extension of that introduced in [11] for (5).

Before proceeding, we establish two technical lemmas that will be used subsequently.

**Lemma 2.1** *Suppose that $Y \in \Re^{n_1\times l_1}$, $Z \in \Re^{r\times r}$, $U \in \Re^{n_1\times r}$, $\bar{U} \in \Re^{n_2\times r}$, $V \in \Re^{l_1\times r}$, $\bar{V} \in \Re^{l_2\times r}$ satisfy*

$$\begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} U \\ \bar{U} \end{bmatrix} Z \begin{bmatrix} V^T & \bar{V}^T \end{bmatrix}. \tag{11}$$

*Assume that $Z$ is nonsingular and $E^T E = G^T G = I$, where*

$$E = \begin{bmatrix} U \\ \bar{U} \end{bmatrix}, \quad G = \begin{bmatrix} V \\ \bar{V} \end{bmatrix}. \tag{12}$$

*Then $\bar{U} = 0$, $\bar{V} = 0$ and $U^T U = V^T V = I$.*

*Proof.* Post-multiplying $G$ on both sides of (11), and using (12) and $G^T G = I$, one has

$$\begin{bmatrix} YV \\ 0 \end{bmatrix} = \begin{bmatrix} UZ \\ \bar{U}Z \end{bmatrix}.$$

It follows that $\bar{U}Z = 0$, which together with the nonsingularity of $Z$ implies that $\bar{U} = 0$. In addition, pre-multiplying $E^T$ on both sides of (11) and using a similar argument as above, one can show that $\bar{V} = 0$. Finally, $U^T U = V^T V = I$ follows from (12), $\bar{U} = 0$, $\bar{V} = 0$ and $E^T E = G^T G = I$. ∎

**Lemma 2.2** *Suppose that $D$ is a positive definite $r \times r$ diagonal matrix. Assume that $U, V \in \Re^{r\times r}$ are such that $U^TU = V^TV = I$ and*

$$UDV^T = D. \tag{13}$$

*Then there hold:*

*(i) $U = V$;*

*(ii) $UD^\alpha U^T = U^TD^\alpha U = D^\alpha$ for every scalar $\alpha$.*

*Proof.* (i) We first show that $DU = UD$. Indeed, it follows from (13) and $U^TU = V^TV = I$ that

$$D^2U = DD^TU = (UDV^T)(VDU^T)U = UD^2.$$

This relation implies that $(D^2 - D^2_{ii}I)U_i = 0$ for all $i$, where $U_i$ denotes the $i$th column of $U$. Hence, we have

$$(D + D_{ii}I)(D - D_{ii}I)U_i = (D^2 - D^2_{ii}I)U_i = 0. \tag{14}$$

Since $D \succ 0$, one has $D + D^2_{ii}I \succ 0$. Using this and (14), we see that $(D - D_{ii}I)U_i = 0$ for all $i$ and hence $DU = UD$. This together with (13) implies that

$$DUV^T = UDV^T = D.$$

Using this relation and the nonsingularity of $D$, we obtain $UV^T = I$, which along with $V^TV = I$ yields $U = V$ as desired.

(ii) Let $1 \le i \le r$ be arbitrarily chosen, and let

$$\mathcal{J} = \{j : D_{jj} \ne D_{ii}\}.$$

Recall from the proof of statement (i) that $(D-D_{ii})U_i = 0$. Hence, $U_{ij} = 0$ for all $j \in \mathcal{J}$. Using this fact, one can observe that $(D^\alpha - D^\alpha_{ii})U_i = 0$, which together with the arbitrarity of $i$ yields $D^\alpha U = UD^\alpha$. In view of this and $U^TU = UU^T = I$, we obtain that $UD^\alpha U^T = U^TD^\alpha U = D^\alpha$ for every scalar $\alpha$. ■

We are now ready to establish that the first-order stationary points of (5) are equivalent to those as given in Definition 1 for problem (10).

**Theorem 2.3** *$x^* \in \Re^l$ is a first-order stationary point of problem (5), i.e., $x^*$ satisfies (9) if and only if $X^* = \mathrm{Diag}(x^*)$ is a first-order stationary point of problem (10), i.e., $X^*$ satisfies (8) with $f$ replaced by $\widehat{h}$.*

*Proof.* ($\Leftarrow$) Suppose that $X^* = \mathrm{Diag}(x^*)$ is a first-order stationary point of problem (10), namely, $X^*$ satisfies (8) with $f$ replaced by $\widehat{h}$. Without loss of generality, assume that $x^*$ satisfies

$$|x^*_1| \ge |x^*_2| \ge \cdots \ge |x^*_r| > |x^*_{r+1}| = \cdots = |x^*_l| = 0$$

for some $r$ (otherwise, one can re-index $x^*$ and $x$ simultaneously). Let $\mathcal{B} = \{1, \ldots, r\}$. Then we have

$$X^* = \left[ \begin{array}{cc} \mathrm{Diag}\,(x^*_{\mathcal{B}}) & 0 \\ 0 & 0 \end{array} \right], \qquad |x^*_{\mathcal{B}}| = \sigma(X^*), \qquad r = \mathrm{rank}(X^*). \tag{15}$$

Since $X^*$ satisfies (8) with $f$ replaced by $\widehat{h}$, there exists some $(U, V) \in \mathcal{M}(X^*)$ such that

$$U^T \nabla \widehat{h}(X^*) V + \lambda p \mathrm{Diag}\left(\sigma(X^*)^{p-1}\right) = 0. \tag{16}$$

Note that $\widehat{h}(\cdot) = h(\mathrm{diag}(\cdot))$. Hence, $\nabla \widehat{h}(X^*) = \mathrm{Diag}(\nabla h(x^*))$. Using this and (16), one has

$$U^T \mathrm{Diag}\,(\nabla h(x^*))\, V + \lambda p \mathrm{Diag}\left(\sigma(X^*)^{p-1}\right) = 0. \tag{17}$$

Let $U_{\mathcal{B}}, V_{\mathcal{B}} \in \Re^{r \times r}$ and $U_{\bar{\mathcal{B}}}, V_{\bar{\mathcal{B}}} \in \Re^{(l-r)} \times r$ such that

$$U = \left[ \begin{array}{c} U_{\mathcal{B}} \\ U_{\bar{\mathcal{B}}} \end{array} \right], \qquad V = \left[ \begin{array}{c} V_{\mathcal{B}} \\ V_{\bar{\mathcal{B}}} \end{array} \right].$$

This, together with $(U, V) \in \mathcal{M}(X^*)$, (6) and (15), yields

$$U^T U = V^T V = I, \qquad \left[ \begin{array}{cc} \mathrm{Diag}\,(x^*_{\mathcal{B}}) & 0 \\ 0 & 0 \end{array} \right] = \left[ \begin{array}{c} U_{\mathcal{B}} \\ U_{\bar{\mathcal{B}}} \end{array} \right] \mathrm{Diag}\,(|x^*_{\mathcal{B}}|) \left[ V^T_{\mathcal{B}} \;\; V^T_{\bar{\mathcal{B}}} \right] \tag{18}$$

It then follows from Lemma 2.1 that $U_{\bar{\mathcal{B}}} = V_{\bar{\mathcal{B}}} = 0$ and $U^T_{\mathcal{B}} U_{\mathcal{B}} = V^T_{\mathcal{B}} V_{\mathcal{B}} = I$, which along with the second relation of (18) implies that

$$\mathrm{Diag}\,(x^*_{\mathcal{B}}) = U_{\mathcal{B}} \mathrm{Diag}\,(|x^*_{\mathcal{B}}|)\, V^T_{\mathcal{B}}.$$

Post-multiplying it by $\mathrm{Diag}(\mathrm{sgn}(x^*_{\mathcal{B}}))$, one can obtain that

$$\mathrm{Diag}\,(|x^*_{\mathcal{B}}|) = U_{\mathcal{B}} \mathrm{Diag}\,(|x^*_{\mathcal{B}}|)\, \bar{V}^T_{\mathcal{B}}, \tag{19}$$

where $\bar{V}_{\mathcal{B}} = \mathrm{Diag}(\mathrm{sgn}(x^*_{\mathcal{B}})) V_{\mathcal{B}}$. Since $V^T_{\mathcal{B}} V_{\mathcal{B}} = I$ and $[\mathrm{Diag}(\mathrm{sgn}(x^*_{\mathcal{B}}))]^2 = I$, it is clear to see that $\bar{V}^T_{\mathcal{B}} \bar{V}_{\mathcal{B}} = I$. Using this relation, $U^T_{\mathcal{B}} U_{\mathcal{B}} = I$, $\mathrm{Diag}(|x^*_{\mathcal{B}}|) \succ 0$, (19), and Lemma 2.2, we have $U_{\mathcal{B}} = \bar{V}_{\mathcal{B}}$ and

$$\mathrm{Diag}\left(|x^*_{\mathcal{B}}|^{p-1}\right) = U^T_{\mathcal{B}} \mathrm{Diag}\left(|x^*_{\mathcal{B}}|^{p-1}\right) U_{\mathcal{B}}. \tag{20}$$

In view of (17), $|x^*_{\mathcal{B}}| = \sigma(X^*)$ and $U_{\bar{\mathcal{B}}} = V_{\bar{\mathcal{B}}} = 0$, one has

$$U^T_{\mathcal{B}} \mathrm{Diag}\,([\nabla h(x^*)]_{\mathcal{B}})\, V_{\mathcal{B}} + \lambda p \mathrm{Diag}(|x^*_{\mathcal{B}}|^{p-1}) = 0.$$

Using this relation, $\bar{V}_{\mathcal{B}} = \mathrm{Diag}\,(\mathrm{sgn}(x^*_{\mathcal{B}}))\, V_{\mathcal{B}}$ and $U_{\mathcal{B}} = \bar{V}_{\mathcal{B}}$, we obtain that

$$U^T_{\mathcal{B}} \mathrm{Diag}\,([\nabla h(x^*)]_{\mathcal{B}} \circ \mathrm{sgn}(x^*_{\mathcal{B}}))\, U_{\mathcal{B}} + \lambda p \mathrm{Diag}\left(|x^*_{\mathcal{B}}|^{p-1}\right) = 0.$$

In view of this relation, (20), and the invertibility of $U_{\mathcal{B}}$, one can conclude that

$$[\nabla h(x^*)]_{\mathcal{B}} \circ \mathrm{sgn}(x^*_{\mathcal{B}}) + \lambda p |x^*_{\mathcal{B}}|^{p-1} = 0,$$

which, along with the definitions of sgn and $X^*$ and the fact that $x_i^* = 0$ for $i \notin \mathcal{B}$, implies that $x^*$ satisfies (9).

($\Rightarrow$) Suppose that $x^*$ is a first-order stationary point of (5), that is, $x^*$ satisfies (9). Let $X^* = \mathrm{Diag}(x^*)$, $\mathcal{B} = \{i : x_i^* \neq 0\}$, and $I_{\mathcal{B}}$ the submatrix of the $l \times l$ identity matrix $I$, consisting of the rows indexed by $\mathcal{B}$. Clearly, there exists a permutation matrix $Q \in \Re^{r\times r}$ such that

$$
\begin{aligned}
\mathrm{Diag}\left(|x^*_{\mathcal{B}}|\right) &= Q\mathrm{Diag}\left(\sigma(X^*)\right)Q^T, &\quad (21)\\
\mathrm{Diag}\left(|x^*_{\mathcal{B}}|^{p-1}\right) &= Q\mathrm{Diag}\left(\sigma(X^*)^{p-1}\right)Q^T. &\quad (22)
\end{aligned}
$$

Let $U = I_{\mathcal{B}}^T Q$ and $V = I_{\mathcal{B}}^T \mathrm{Diag}\left(\mathrm{sgn}(x^*_{\mathcal{B}})\right)Q$. Using these and (21), one can observe that $U^TU = V^TV = I$, and moreover,

$$
\begin{aligned}
X^* &= I_{\mathcal{B}}^T \mathrm{Diag}(x^*_B) I_{\mathcal{B}} = I_{\mathcal{B}}^T \mathrm{Diag}\left(|x^*_B|\right)\mathrm{Diag}\left(\mathrm{sgn}(x^*_{\mathcal{B}})\right) I_{\mathcal{B}} \\
&= I_{\mathcal{B}}^T Q \mathrm{Diag}\left(\sigma(X^*)\right) Q^T \mathrm{Diag}\left(\mathrm{sgn}(x^*_{\mathcal{B}})\right) I_{\mathcal{B}} = U\mathrm{Diag}\left(\sigma(X^*)\right)V^T.
\end{aligned}
$$

Hence, $(U, V) \in \mathcal{M}(X^*)$. In addition, it follows from (9) that

$$\mathrm{Diag}\left([\nabla h(x^*)]_{\mathcal{B}}\right)\mathrm{Diag}\left(\mathrm{sgn}(x^*_{\mathcal{B}})\right) + \lambda p \mathrm{Diag}\left(|x^*_{\mathcal{B}}|^{p-1}\right) = 0,$$

which together with (22) and the definition of $I_{\mathcal{B}}$ leads to

$$I_{\mathcal{B}}\mathrm{Diag}\left(\nabla h(x^*)\right) I_{\mathcal{B}}^T \mathrm{Diag}\left(\mathrm{sgn}(x^*_{\mathcal{B}})\right) + \lambda p Q \mathrm{Diag}\left(\sigma(X^*)^{p-1}\right)Q^T = 0.$$

Pre- and post-multiplying this equality by $Q^T$ and $Q$, respectively, and using the definitions of $U$ and $V$, one can obtain that

$$U^T\mathrm{Diag}\left(\nabla h(x^*)\right)V + \lambda p \mathrm{Diag}\left(\sigma(X^*)^{p-1}\right) = 0.$$

The conclusion follows from this equality, $\nabla\widehat{h}(X^*) = \mathrm{Diag}(\nabla h(x^*))$, and $(U, V) \in \mathcal{M}(X^*)$. ∎

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

**Theorem 2.4** *Let $X^*$ be a local minimizer of problem* (1)*. Then $X^*$ is a stationary point of* (1)*, that is,* (8) *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

$$\min_{Z\in\Re^{r\times r}} \underbrace{f(X^* + \bar{U}Z\bar{V}^T) + \lambda\|\mathrm{Diag}(\sigma(X^*)) + Z\|_p^p}_{w(Z)}. \quad (23)$$

By [24, 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 (23), the first-order optimality condition of (23) yields $0 \in \partial w(0)$. Hence, there exists some $(U_\sigma, V_\sigma) \in \mathcal{M}(\mathrm{Diag}(\sigma(X^*)))$ such that

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

Pre- and post-multiplying it by $U_\sigma^T$ and $V_\sigma$, respectively, and using $U_\sigma^T U_\sigma = V_\sigma^T V_\sigma = I$, we obtain that

$$U^T \nabla f(X^*) V + \lambda p \mathrm{Diag}(\sigma(X^*)^{p-1}) = 0, \qquad (24)$$

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 (24) implies that (8) 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 2.5** *Let $X^*$ be a first-order stationary point of* (1) *satisfying $F(X^*) \le F(X^0) + \epsilon$ for some $X^0 \in \Re^{m \times n}$ and $\epsilon \ge 0$, $\mathcal{B} = \{i : \sigma_i(X^*) \neq 0\}$, $L_f$ and $\underline{f}$ be defined in* (2) *and* (7) *, respectively. Then there holds:*

$$\sigma_i(X^*) \ \ge \ \left( \frac{\lambda p}{\sqrt{2L_f[F(X^0) + \epsilon - \underline{f}]}} \right)^{\frac{1}{1-p}} \qquad \forall i \in \mathcal{B}. \qquad (25)$$

*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

$$f(X^* - \nabla f(X^*)/L_f) \ \le \ f(X^*) - \frac{1}{2L_f} \|\nabla f(X^*)\|_F^2. \qquad (26)$$

Note that

$$f(X^* - \nabla f(X^*)/L_f) \ge \inf_{X \in \Re^{m \times n}} f(X) = \underline{f}, \qquad f(X^*) \le F(X^*) \le F(X^0) + \epsilon.$$

Using these relations and (26), we have

$$\|\nabla f(X^*)\|_F \ \le \ \sqrt{2L_f[f(X^*) - f(X^* - \nabla f(X^*)/L_f)]} \ \le \ \sqrt{2L_f[F(X^0) + \epsilon - \underline{f}]}. \qquad (27)$$

Since $X^*$ is a first-order stationary point of (1), we know that $X^*$ satisfies (8) for some $(U,V) \in \mathcal{M}(X^*)$. Using (8) and $p \in (0,1)$, we obtain that for every $i \in \mathcal{B}$,

$$\sigma_i(X^*) = \left(\frac{[U^T \nabla f(X^*) V]_{ii}}{\lambda p}\right)^{\frac{1}{p-1}} \ \geq \ \left(\frac{\|U^T \nabla f(X^*) V\|_F}{\lambda p}\right)^{\frac{1}{p-1}}. \tag{28}$$

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

$$\|U^T \nabla f(X^*) V\|_F \ \leq \ \|\nabla f(X^*)\|_F,$$

which together with (28) yields

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

Using this relation, (27), and $p \in (0,1)$, one can see that (25) holds. ■

# 3 Iterative reweighted singular value minimization methods for (1)

In this section we extend two IRL$_1$ methods proposed by Lu [25] for solving $l_p$ regularized vector minimization problem to solve problem (1). The resulting methods are referred to as iterative reweighted singular value minimization (IRSVM) methods. Throughout this section, we set

$$l := \min(m,n).$$

We start by reviewing a technical result that was established by Lu and Zhang [26] and will play a crucial role in our subsequent convergence analysis of the IRSVM methods.

Let $\mathcal{U}^n$ denote the set of all unitary matrices in $\Re^{n\times n}$. A norm $\|\cdot\|$ is a *unitarily invariant norm* on $\Re^{m\times n}$ if $\|UXV\| = \|X\|$ for all $U \in \mathcal{U}^m$, $V \in \mathcal{U}^n$, $X \in \Re^{m\times n}$. More generally, a function $F : \Re^{m\times n} \to \Re$ is a *unitarily invariant function* if $F(UXV) = F(X)$ for all $U \in \mathcal{U}^m$, $V \in \mathcal{U}^n$, $X \in \Re^{m\times n}$. A set $\mathcal{X} \subseteq \Re^{m\times n}$ is a *unitarily invariant set* if

$$\{UXV : U \in \mathcal{U}^m, V \in \mathcal{U}^n, X \in \mathcal{X}\} = \mathcal{X}.$$

The following result shows that a class of matrix optimization problems over a certain subset of $\Re^{m\times n}$ can be solved as lower dimensional vector optimization problems. Its proof is given in [26, Proposition 2.1] .

**Lemma 3.1** *Let $A \in \Re^{m\times n}$ be given, $l = \min(m,n)$, and let the operator $\mathcal{D} : \Re^l \to \Re^{m\times n}$ be defined as follows:*

$$[\mathcal{D}(x)]_{ij} = \begin{cases} x_i & \text{if } i = j; \\ 0 & \text{otherwise} \end{cases} \qquad \forall x \in \Re^l.$$

*Suppose that $\|\cdot\|$ is a unitarily invariant norm on $\Re^{m\times n}$, $\Theta : \Re^{m\times n} \to \Re$ is a unitarily invariant function, and that $\mathcal{X} \subseteq \Re^{m\times n}$ is a unitarily invariant set. Assume that $\phi$ is a non-decreasing function on $[0, \infty)$. Let $U\Sigma V^T$ be the singular value decomposition of $A$. Then $X^* = U\mathcal{D}(x^*)V^T$ is an optimal solution of the problem*

$$\begin{array}{ll} \min & \Theta(X) + \phi(\|X - A\|) \\ \text{s.t.} & X \in \mathcal{X}, \end{array}$$

*where $x^* \in \Re^l$ is an optimal solution of the problem*

$$\begin{array}{ll} \min & \Theta(\mathcal{D}(x)) + \phi(\|\mathcal{D}(x) - \Sigma\|) \\ \text{s.t.} & \mathcal{D}(x) \in \mathcal{X}. \end{array}$$

As a consequence of Lemma 3.1, we can show that the following weighted singular value minimization (WSVM) problem, which is in the same form as the subproblems of the IRSVM methods presented subsequently, has a closed-form solution. It thus offers a tool for solving the subproblems of the IRSVM methods.

**Corollary 3.2** *Given any $B, C \in \Re^{m\times n}$, $L > 0$ and $s \in \Re^l_+$, where $l = \min(m, n)$, consider the WSVM problem*

$$\min_{X\in\Re^{m\times n}} \left\{ \langle C, X - B\rangle + \frac{L}{2}\|X - B\|_F^2 + \sum_{i=1}^{l} s_i\sigma_i(X) \right\}. \tag{29}$$

*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 (29).*

*Proof.* One can observe that problem (29) is equivalent to

$$\min_{X\in\Re^{m\times n}} \left\{ \frac{L}{2} \left\| X - \left(B - \frac{C}{L}\right) \right\|_F^2 + \sum_{i=1}^{l} s_i\sigma_i(X) \right\}. \tag{30}$$

Let $x^*$ be defined above. Note that $L > 0$, and $d, s \in \Re^l_+$. Using this fact, it is not hard to see that

$$x^* = \arg\min_{x\in\Re^l} \left\{ \frac{L}{2}\|x - d\|_2^2 + \sum_{i=1}^{l} s_i|x_i| \right\}.$$

Invoking Lemma 3.1 with $\Theta(X) = \sum_{i=1}^{l} s_i\sigma_i(X)$, $\phi(t) = Lt^2/2$ and $\|\cdot\| = \|\cdot\|_F$, one can conclude that $X^* = U\mathrm{Diag}(x^*)V^T$ is an optimal solution of (30). Due to the equivalence of (29) and (30), $X^*$ is also an optimal solution of (29). ∎

## 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, [25, Algorithm 5]) that was proposed for solving $l_p$ regularized *vector* minimization in the form of (5), and study its convergence.

Define

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

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

**Algorithm 1: The first IRSVM method for** (1)

Let $0 < L_{\min} < L_{\max}$, $\tau > 1$, $c > 0$, integer $N \geq 0$, and $\{\epsilon^k\} \subset \Re^n$ be a sequence of non-increasing positive vectors 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 Corollary 3.2 to find a solution to the WSVM subproblem

$$X^{k+1} \in \operatorname{Arg} \min_{X\in\Re^{m\times n}} \left\{ \langle \nabla f(X^k), X - X^k \rangle + \frac{L_k}{2}\|X - X^k\|_F^2 + \lambda p \sum_{i=1}^{l} s_i^k \sigma_i(X) \right\}, \tag{31}$$

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

   1b) If

$$\bar{F}_{\epsilon^{k+1}}(X^{k+1}) \leq \max_{[k-N]^+ \leq i \leq k} \bar{F}_{\epsilon^i}(X^i) - \frac{c}{2}\|X^{k+1} - X^k\|_F^2 \tag{32}$$

   is satisfied, then go to step 2).

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

2) Set $k \leftarrow k + 1$ 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 [25, Theorem 3.7].

**Theorem 3.3** *For each $k \geq 0$, the inner termination criterion* (32) *is satisfied after at most* $\left\lceil \frac{\log(L_f + c) - \log(L_{\min})}{\log \tau} + 2 \right\rceil$ *inner iterations.*

We next show that the sequence $\{X^k\}$ generated by the first IRSVM method is bounded, and moreover, any accumulation point of $\{X^k\}$ is a first-order stationary point of (1).

**Theorem 3.4** *Suppose that* $\{\epsilon^k\}$ *is a sequence of non-increasing positive vectors in* $\Re^n$ *and* $\epsilon^k \to 0$ *as* $k \to \infty$. *Let the sequence* $\{X^k\}$ *be generated by the first IRSVM method. There 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.,* (8) *holds at* $X^*$. *Moreover, the nonzero entries of* $X^*$ *satisfy the bound* (25) *with* $\epsilon = \bar{F}_{\epsilon^0}(X^0) - F(X^0)$.

*Proof.* (i) Using (32) 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 (7) , $\epsilon^k > 0$, and the definition of $\bar{F}_\epsilon$ implies that

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

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

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

Observe that $\bar{F}_\epsilon(X)$, viewed as a function of $(X, \epsilon)$, is uniformly continuous in $\Omega$. Using this fact, (32), $\{\epsilon^k\} \to 0$, and a similar argument as used in the proof of [37, Lemma 4], one can show that $\|X^{k+1} - X^k\| \to 0$. Let $\bar{L}_k$ denote the final value of $L_k$ at the $k$th outer iteration. Clearly, $\bar{L}_k \ge L_{\min} > 0$. Moreover, it follows from Theorem 3.3 that $\{\bar{L}_k\}$ is bounded. As mentioned in Algorithm 1, $X^{k+1}$ is the solution of the subproblem (31) with $L_k = \bar{L}_k$ found by applying Corollary 3.2. Let $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$, and $U^k \mathrm{Diag}(d^k)(V^k)^T$ be 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 (31) and Corollary 3.2 that

$$X^{k+1} = U^k \mathrm{Diag}(x^{k+1})(V^k)^T, \tag{33}$$

where

$$x^{k+1} = \max(d^k - \lambda p s^k/\bar{L}_k, 0).^1 \tag{34}$$

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

$$\{s_i^k\}_\mathcal{K} \to \sigma_i(X^*)^{p-1} =: s_i^* \qquad \forall i. \tag{35}$$

[1]Notice that $\{s_1^k, \ldots, s_l^k\}$ may not be in increasing order. Therefore, $\{x_1^{k+1}, \ldots, x_l^{k+1}\}$ is generally not in descending order either.

By the boundedness of $\{X^k\}$, $\bar{L}_k > L_{\min}$, and $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$, we can see that $\{Z^k\}$ is bounded and so is $\{d^k\}$. Without loss of generality, assume that $\{d^k\}_{\mathcal{K}} \to d^*$ and $\{\bar{L}_k\}_{\mathcal{K}} \to \bar{L}^* \geq L_{\min} > 0$ (otherwise, one can consider their convergent subsequence). It then follows from (34) that

$$\{x^{k+1}\}_{\mathcal{K}} \to \max(d^* - \lambda p s^*/\bar{L}^*, 0) =: x^*. \tag{36}$$

Since $d_1^k \geq \cdots \geq d_l^k$, we know that $d_1^* \geq \cdots \geq d_l^*$. Also, one can observe from (35) that $s_1^* \leq \cdots \leq s_l^*$. Using these inequalities and (36), we can conclude that $x_1^* \geq \cdots \geq x_l^* \geq 0$. Since $\|X^{k+1} - X^k\| \to 0$ and $\{X^k\}_{\mathcal{K}} \to X^*$, we also have $\{X^{k+1}\}_{\mathcal{K}} \to X^*$. Using this relation, (33), (36), and $x_1^* \geq \cdots \geq x_l^* \geq 0$, it is not hard to show that

$$\{x_i^{k+1}\}_{\mathcal{K}} \to x_i^* = \sigma_i(X^*) \qquad \forall i. \tag{37}$$

Let $r = \mathrm{rank}(X^*)$. One can observe from (37) that there exists some $k_0 > 0$ such that $x_i^{k+1} > 0$ for all $1 \leq i \leq r$ and $k \in \mathcal{K}_0 = \{i \in \mathcal{K} : i > k_0\}$. It then follows from (34) that

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

Hence,

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

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

$$\bar{L}_k \sum_{i=1}^{r} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T + \lambda p \sum_{i=1}^{r} s_i^k U_i^k (V_i^k)^T \ = \ 0. \tag{38}$$

Using (33) 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 $Z^k = X^k - \frac{1}{\bar{L}_k}\nabla f(X^k)$ 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 (38), we obtain that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k(X^{k+1} - X^k) + \nabla f(X^k) - \bar{L}_k \sum_{i=r+1}^{l} (x_i^{k+1} - d_i^k) U_i^k (V_i^k)^T + \lambda p \sum_{i=1}^{r} s_i^k U_i^k (V_i^k)^T \ = \ 0. \tag{39}$$

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 (39) 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 $k \in \mathcal{K}_0$,

$$\bar L_k (\bar U^k)^T (X^{k+1} - X^k) \bar V^k + (\bar U^k)^T \nabla f(X^k) \bar V^k + \lambda p \cdot \mathrm{Diag}(s_1^k, \ldots, s_r^k) \ = \ 0.^2 \tag{40}$$

Notice that $\{\bar U^k\}_{\mathcal{K}}$ and $\{\bar V^k\}_{\mathcal{K}}$ are bounded. Without loss of generality, assume that $\{\bar U^k\}_{\mathcal{K}} \to \bar U^*$ and $\{\bar V^k\}_{\mathcal{K}} \to \bar V^*$ (one can consider their convergent subsequence if necessary). Using (35), the boundedness of $\{\bar L_k\}$, $\|X^{k+1} - X^k\| \to 0$ and $X^k \to X^*$ as $k \in \mathcal{K} \to \infty$, and taking limits on both sides of (40) as $k \in \mathcal{K}_0 \to \infty$, we have

$$(\bar U^*)^T \nabla f(X^*) \bar V^* + \lambda p \cdot \mathrm{Diag}(\sigma(X^*)^{p-1}) \ = \ 0. \tag{41}$$

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 (33), (37), $r = \mathrm{rank}(X^*)$, $\{X^{k+1}\}_{\mathcal{K}} \to X^*$, $\{\bar U^k\}_{\mathcal{K}} \to \bar U^*$, 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 (41), we can conlcude that (8) 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 together with the definition of $\bar F_\epsilon(\cdot)$ implies that

$$F(X^k) \ \le \ \bar F_{\epsilon^k}(X^k) \ \le \ \bar F_{\epsilon^0}(X^0).$$

Taking limits on both sides as $k \in \mathcal{K} \to \infty$, one has $F(X^*) \le \bar F_{\epsilon^0}(X^0)$. The second part of statement (ii) then follows from this relation and Theorem 2.5. ■

*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 3.4. 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 (41) that

$$\lim_{k \in \mathcal{K} \to \infty} (\bar U^k)^T \nabla f(X^k) \bar V^k + \lambda p \cdot \mathrm{Diag}((v^k)^{p-1}) \ = \ 0, \tag{42}$$

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 $\{X^k\}_{\mathcal{K}} \to X^*$. Using this fact and (42), one can observe that

$$\lim_{k \in \mathcal{K} \to \infty} \mathrm{Diag}(\sigma(X^k)^{\frac12})(U^k)^T \nabla f(X^k) V^k \mathrm{Diag}(\sigma(X^k)^{\frac12}) + \lambda p \cdot \mathrm{Diag}(\sigma(X^k)^p) \ = \ 0. \tag{43}$$

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

$$\left\| \mathrm{Diag}(\sigma(X^k)^{\frac12})(U^k)^T \nabla f(X^k) V^k \mathrm{Diag}(\sigma(X^k)^{\frac12}) + \lambda p \cdot \mathrm{Diag}(\sigma(X^k)^p) \right\|_{\max} \le \bar\varepsilon \tag{44}$$

for some prescribed accuracy parameter $\bar\varepsilon$. ■

[2]In the context of $l_p$ regularized vector minimization, a result similar to (40) was derived in the proof of [25, Theorem 3.8] by using the first-order optimality condition of a subproblem similar to (31). It is not hard to observe that under the assumption $\sigma_1(X^{k+1}) > \cdots > \sigma_{r+1}(X^{k+1})$, (40) can be derived analogously by using the Clarke subdifferential of singular values (see [24, Corollary 6.4]) and the first-order optimality condition of (31). The assumption, however, generally does not hold and thus this traditional approach is not applicable here. Instead, our approach is based on exploiting the structure of (31), which is substantially different from the traditional approach.

### 3.2 The second IRSVM method for (1)

In this subsection we extend another IRL$_1$ method (namely, [25, Algorithm 7]) to solve problem (1) and establish a global convergence for the resulting IRSVM method.

Let $q$ be such that

$$\frac{1}{p}+\frac{1}{q}=1. \tag{45}$$

For any $u>0$, let

$$h_u(t) := \min_{0\le s\le u} p\left(|t|s-\frac{s^q}{q}\right) \quad \forall t\in\Re. \tag{46}$$

Given any $\epsilon>0$, define

$$F_\epsilon(X) := f(X)+\lambda\sum_{i=1}^{l} h_{u_\epsilon}(\sigma_i(X)), \tag{47}$$

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

$$u_\epsilon := \left(\frac{\epsilon}{\lambda l}\right)^{\frac{1}{q}}. \tag{48}$$

By a similar argument as in the proof of [25, Proposition 2.6], it can be shown that $F_\epsilon$ is locally Lipschitz continuous, and moreover,

$$0 \ \le\ F_\epsilon(X)-F(X) \ \le\ \epsilon \qquad \forall X\in\Re^{m\times n}. \tag{49}$$

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

$$\varepsilon(X^0)=\left\{\epsilon : 0<\epsilon<l\lambda\left[\frac{\sqrt{2L_f[F(X^0)+\epsilon-\underline{f}]}}{\lambda p}\right]^q\right\}. \tag{50}$$

We are now ready to present the second IRSVM method for solving (1), which is an extension of [25, Algorithm 7] proposed for solving $l_p$ 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 $N\ge 0$ be given. Let $q$ be defined in (45). Choose an arbitrary $X^0$ 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 Corollary 3.2 to find a solution to the WSVM subproblem

$$X^{k+1}\in \operatorname{Arg}\min_{x\in\Re^{m\times n}}\left\{\langle\nabla f(X^k),X-X^k\rangle+\frac{L_k}{2}\|X-X^k\|_F^2+\lambda p\sum_{i=1}^{l}s_i^k\sigma_i(X)\right\}, \tag{51}$$

   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

$$F_\epsilon(X^{k+1}) \leq \max_{[k-N]^+ \leq i \leq k} F_\epsilon(X^i) - \frac{c}{2}\|X^{k+1} - X^k\|_F^2 \tag{52}$$

is satisfied, where $F_\epsilon$ is defined in (47), then go to step 2).

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

2) Set $k \leftarrow k+1$ 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 [25, Theorem 4.2].

**Theorem 3.5** *For each $k \geq 0$, the inner termination criterion* (32) *is satisfied after at most* $\left\lceil \frac{\log(L_f+c)-\log(L_{\min})}{\log \tau} + 2 \right\rceil$ *inner iterations.*

We next establish that the sequence $\{X^k\}$ generated by the second IRSVM method is bounded and moreover any accumulation point of $\{X^k\}$ is a stationary point of (1).

**Theorem 3.6** *Let the sequence $\{X^k\}$ be generated by the second IRSVM method. Assume that $\epsilon \in \varepsilon(X^0)$, where $\varepsilon(X^0)$ is defined in* (50)*. There hold:*

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

*(ii) Let $X^*$ be any accumulation point of $\{X^k\}$. Then $X^*$ is a stationary point of* (1)*, i.e.,* (8) *holds at $X^*$. Moreover, the nonzero entries of $X^*$ satisfy the bound* (25).

*Proof.* (i) Using (52) and an inductive argument, we can conclude that $F_\epsilon(X^k) \leq F_\epsilon(X^0)$ for all $k$. This together with (49) implies that $F(x^k) \leq F_\epsilon(X^0)$. Using this relation, (1) and (7), we see that $\|X^k\|_p^p \leq (F_\epsilon(X^0) - \underline{f})/\lambda$ and hence $\{X^k\}$ is bounded.

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

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

Observe that $F_\epsilon$ is uniformly continuous in $\Omega$. Using this fact, (32), and a similar argument as used in the proof of [37, Lemma 4], one can show that $\|X^{k+1} - X^k\| \to 0$. Let $\bar{L}_k$ denote the final value of $L_k$ at the $k$th outer iteration. It follows from Theorem 3.5 that $\{\bar{L}_k\}$ is bounded. As observed from Algorithm 2, $X^{k+1}$ is the solution of the subproblem (51) with $L_k = \bar{L}_k$ found by applying Corollary 3.2. Let $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$, and $U^k \mathrm{Diag}(d^k)(V^k)^T$ be 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 (51) and Corollary 3.2 that

$$X^{k+1} = U^k \mathrm{Diag}(x^{k+1})(V^k)^T, \tag{53}$$

where

$$x^{k+1} = \max(d^k - \lambda p s^k/\bar{L}_k, 0). \tag{54}$$

By the definitions of $d^k$ and $s^k$, we know that $d_1^k \geq \cdots \geq d_l^k$ and $s_1^k \leq \cdots \leq s_l^k$. These together with (54) imply that $x_1^{k+1} \geq \cdots \geq x_l^{k+1}$. This relation and (53) yields

$$x_i^{k+1} = \sigma_i(X^{k+1}) \quad \forall i. \tag{55}$$

Suppose that $X^*$ is an accumulation point of $\{X^k\}$. Then 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 (55) leads to

$$\{x_i^{k+1}\}_{\mathcal{K}} \to \sigma_i(X^*) \quad \forall i. \tag{56}$$

Let $r = \mathrm{rank}(X^*)$. One can observe from (56) that there exists some $k_0 > 0$ such that $x_i^{k+1} > 0$ for all $1 \leq i \leq r$ and $k \in \mathcal{K}_0 = \{i \in \mathcal{K} : i > k_0\}$. Using this fact, (54), and a similar argument as used in the proof of Theorem 3.4, we can show that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k(\bar{U}^k)^T(X^{k+1} - X^k)\bar{V}^k + (\bar{U}^k)^T\nabla f(X^k)\bar{V}^k + \lambda p \cdot \mathrm{Diag}(s_1^k, \ldots, s_r^k) \ = \ 0,^3 \tag{57}$$

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. Without loss of generality, assume that $\{\bar{U}^k\}_{\mathcal{K}} \to \bar{U}^*$ and $\{\bar{V}^k\}_{\mathcal{K}} \to \bar{V}^*$ (one can consider their convergent subsequence if necessary). Using (56), the boundedness of $\{\bar{L}_k\}$, $\|X^{k+1} - X^k\| \to 0$, and $X^k \to X^*$ as $k \in \mathcal{K} \to \infty$, and taking limits on both sides of (57) as $k \in \mathcal{K}_0 \to \infty$, we have

$$(\bar{U}^*)^T\nabla f(X^*)\bar{V}^* + \lambda p \cdot \mathrm{Diag}(s_1^*, \ldots, s_r^*) \ = \ 0, \tag{58}$$

where

$$s_i^* = \min\left\{\left(\frac{\epsilon}{\lambda l}\right)^{\frac{1}{q}}, \sigma_i(X^*)^{\frac{1}{q-1}}\right\}, \quad i = 1, \ldots, r. \tag{59}$$

Similar to the proof of Theorem 3.4, 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 $F_\epsilon$ and $\{X^k\}_{\mathcal{K}} \to X^*$, one has $F_\epsilon(X^*) \leq F_\epsilon(X^0)$, which together with (49) yields

$$f(X^*) \ \leq \ F(X^*) \ \leq \ F_\epsilon(X^*) \ \leq \ F_\epsilon(X^0) \ \leq \ F(X^0) + \epsilon. \tag{60}$$

Using this relation and similar arguments as for deriving (27), we see that (27) also holds for such $X^*$. It then follows from (27), (58), and $(\bar{U}^*)^T\bar{U}^* = (\bar{V}^*)^T\bar{V}^* = I$ that, for $1 \leq i \leq r$,

$$\begin{aligned} s_i^* \quad &= \quad \frac{1}{\lambda p}\left|[(\bar{U}^*)^T\nabla f(X^*)\bar{V}^*]_{ii}\right| \ \leq \ \frac{1}{\lambda p}\|(\bar{U}^*)^T\nabla f(X^*)\bar{V}^*]\|_F \\ &\leq \ \frac{1}{\lambda p}\|\nabla f(X^*)\|_F \ \leq \ \frac{\sqrt{2L_f[F(x^0) + \epsilon - \underline{f}]}}{\lambda p}. \end{aligned} \tag{61}$$

---

[3]We shall mention that under the assumption $\sigma_1(X^{k+1}) > \cdots > \sigma_{r+1}(X^{k+1})$, (57) can be derived directly by using the Clarke subdifferential of singular values (see [24, Corollary 6.4]) and the first-order optimality condition of (51). This assumption, however, generally does not hold and thus this traditional approach is not applicable.

We now claim that $\sigma_i(X^*) > u_\epsilon{}^{q-1}$ for all $1 \le i \le r$, where $u_\epsilon$ is defined in (48). Suppose for contradiction that there exists some $1 \le i \le r$ such that $0 < \sigma_i(X^*) < u_\epsilon{}^{q-1}$. It then follows from (59) that

$$s_i^* = u_\epsilon = \left(\frac{\epsilon}{\lambda l}\right)^{\frac{1}{q}}.$$

Using this relation and (61), we have

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

which contradicts with the assumption $\epsilon \in \varepsilon(X^0)$. Therefore, $\sigma_i(X^*) > u_\epsilon{}^{q-1}$ for all $1 \le i \le r$. Using this relation, (59) and (45), we see that

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

Substituting it into (58), we obtain that

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

Using this relation and $(\bar{U}^*, \bar{V}^*) \in \mathcal{M}(X^*)$, we can conlcude that (8) holds at $X^*$ with $U = \bar{U}^*$, $V = \bar{V}^*$. Finally, recall from (60) that $F(X^*) \ \le \ F(X^0)+\epsilon$. Using this relation and Theorem 2.5, 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 3.6. By a similar argument as for the first IRSVM method, one can show that (43) also holds for the second IRSVM method with the above $U^k$, $V^k$ and $\mathcal{K}$. Therefore, (44) with these $U^k$ and $V^k$ can also be used as a termination criterion for the second IRSVM method. ∎

# 4 Nonmonotone proximal gradient method for (1)

In this section we study a nonmonotone proximal gradient (NPG) method for solving problem (1), which is an extension of the method proposed by Wright et al. [37] for minimizing sum of a Lipschitz continuously differentiable function and a possibly nonsmooth function.

Before proceeding, we establish that the following special $l_p$ regularized matrix minimization problem, which is in the same form as the subproblems of the NPG method presented below, can be solved as a lower-dimensional vector minimization problem. Since the latter problem is separable and can be suitably solved, this provides an efficient tool for solving the subproblems of the NPG method.

**Lemma 4.1** *Given any $B, C \in \Re^{m\times n}$, and $L, \lambda > 0$, consider the proximal $l_p$ regularized matrix minimization problem*

$$\min_{X\in\Re^{m\times n}} \left\{ \langle C, X-B\rangle + \frac{L}{2}\|X-B\|_F^2 + \lambda\|X\|_p^p \right\}. \tag{62}$$

*Let $U\mathrm{Diag}(d)V^T$ be the singular value decomposition of $B - C/L$, $l = \min(m,n)$, and*

$$x^* = \arg\min_{x\in\Re^l}\left\{\frac{L}{2}\|x-d\|_2^2 + \lambda\sum_{i=1}^{l}|x_i|^p\right\}.$$

*Then $X^* = U\mathrm{Diag}(x^*)V^T$ is an optimal solution to problem* (62).

*Proof.* One can observe that problem (62) is equivalent to

$$\min_{X\in\Re^{m\times n}}\left\{\frac{L}{2}\left\|X-\left(B-\frac{C}{L}\right)\right\|_F^2 + \lambda\|X\|_p^p\right\}.$$

The conclusion immediately follows from Lemma 3.1 with $\Theta(X) = \lambda\|X\|_p^p$, $\phi(t) = Lt^2/2$ and $\|\cdot\| = \|\cdot\|_F$. ∎

We are now ready to present an NPG method for solving problem (1).

**Algorithm 3: A nonmonotone proximal gradient (NPG) method for** (1)

Let $0 < L_{\min} < L_{\max}$, $\tau > 1$, $c > 0$ and integer $N \geq 0$ be given. 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 Lemma 4.1 to solve the subproblem

$$X^{k+1} \in \mathrm{Arg}\min_{X\in\Re^{m\times n}}\left\{\langle\nabla f(X^k), X - X^k\rangle + \frac{L_k}{2}\|X-X^k\|_F^2 + \lambda\|X\|_p^p\right\}. \tag{63}$$

   1b) If

$$F(X^{k+1}) \leq \max_{[k-N]^+\leq i\leq k} F(X^i) - \frac{c}{2}\|X^{k+1}-X^k\|_F^2 \tag{64}$$

   is satisfied, then go to step 2).

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

2) Set $k \leftarrow k+1$ and go to step 1).

**end**

We first state 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 [25, Theorem 4.2].

**Theorem 4.2** *For each $k \geq 0$, the inner termination criterion* (64) *is satisfied after at most $\left\lceil \frac{\log(L_f+c)-\log(L_{\min})}{\log\tau} + 2\right\rceil$ inner iterations.*

We next establish that the sequence $\{X^k\}$ generated above is bounded and moreover any accumulation point of $\{X^k\}$ is a stationary point of problem (1).

**Theorem 4.3** *Let $\{X^k\}$ be the sequence generated by the above NPG method. There hold:*

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

*(ii) Let $X^*$ be any accumulation point of $\{X^k\}$. Then $X^*$ is a stationary point of (1). Moreover, the nonzero entries of $X^*$ satisfy the bound (25) with $\epsilon = 0$.*

*Proof.* (i) Using (64) and an inductive argument, one can see that $F(X^k) \leq F(X^0)$ for all $k$. Using this fact and (7), we have $\underline{f} + \lambda \|X^k\|_p^p \leq F(X^0)$. It then follows that $\|X^k\|_p^p \leq (F(X^0) - \underline{f})/\lambda$ and hence $\{X^k\}$ is bounded.

(ii) Using (64) and a similar proof as in [37], one can show that $\|X^{k+1} - X^k\| \to 0$. Let $\bar{L}_k$ denote the final value of $L_k$ at the $k$th outer iteration. It follows from Theorem 4.2 that $\{\bar{L}_k\}$ is bounded. As observed from Algorithm 3, $X^{k+1}$ is the solution of the subproblem (63) with $L_k = \bar{L}_k$ found by applying Lemma 4.1. Let $l = \min(m, n)$, $Z^k = X^k - \nabla f(X^k)/\bar{L}_k$, and $U^k \mathrm{Diag}(d^k)(V^k)^T$ be 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 (63) and Lemma 4.1 that

$$X^{k+1} = U^k \mathrm{Diag}(x^{k+1})(V^k)^T, \tag{65}$$

where

$$x^{k+1} = \arg\min_{x \in \Re^l} \left\{ \frac{\bar{L}_k}{2} \left\|x - d^k\right\|_2^2 + \lambda \left\|x\right\|_p^p \right\}. \tag{66}$$

Suppose that $X^*$ is an accumulation point of $\{X^k\}$. Then there exists a subsequence $\mathcal{K}$ such that $\{X^k\}_{\mathcal{K}} \to X^*$, which together with $\|X^{k+1} - X^k\| \to 0$ implies that $\{X^{k+1}\}_{\mathcal{K}} \to X^*$. Without loss of generality, assume that $\{x_1^{k+1}, \ldots, x_l^{k+1}\}$ is in descending order (one can rearrange the components of $x^{k+1}$ and $d^k$ and the columns of $U^k$ and $V^k$ if necessary). It then follows from (65) and $\{X^k\}_{\mathcal{K}} \to X^*$ that for all $i$,

$$x_i^{k+1} = \sigma_i(X^{k+1}) \to \sigma_i(X^*) \text{ as } k \in \mathcal{K} \to \infty. \tag{67}$$

Let $r = \mathrm{rank}(X^*)$. One can observe from (67) that there exists some $k_0 > 0$ such that $x_i^{k+1} > 0$ for all $1 \leq i \leq r$ and $k \in \mathcal{K}_0 = \{i \in \mathcal{K} : i > k_0\}$. By the first-order optimality conditions of (66), we have

$$\bar{L}_k(x_i^{k+1} - d_i^k) + \lambda p (x_i^{k+1})^{p-1} = 0, \quad 1 \leq i \leq r.$$

Using this relation and a similar argument as used in the proof of Theorem 3.4, one can show that for all $k \in \mathcal{K}_0$,

$$\bar{L}_k (\bar{U}^k)^T (X^{k+1} - X^k) \bar{V}^k + (\bar{U}^k)^T \nabla f(X^k) \bar{V}^k + \lambda p \cdot \mathrm{Diag}((x_1^{k+1})^{p-1}, \ldots, (x_r^{k+1})^{p-1}) \ = \ 0, \tag{68}$$

where

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

The rest of the proof follows from (67), (68), and a similar argument as used in the proof of Theorem 3.4. ■

*Remark.* By a similar argument as for the IRSVM methods, (44) with $U^k$, $V^k$ given in the proof of Theorem 4.3 can be used as a termination criterion for the NPG method. ■

# 5 Numerical results

In this section we conduct numerical experiments to test the performance of the IRSVM methods (Algorithms 1 and 2) and the NPG method (Algorithm 3) proposed in sections 3 and 4 by applying them to solve the matrix completion problem. In particular, we apply these methods to problem (1) with $f(X) = \|\mathcal{P}_\Omega(X - M)\|_F^2$, that is

$$\min_{X\in\Re^{m\times n}} \left\{\|\mathcal{P}_\Omega(X - M)\|_F^2 + \lambda\|X\|_p^p\right\}, \tag{69}$$

where $M \in \Re^{m\times n}$, $\Omega$ 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 $\Omega$. 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, IRSVM-2 and NPG according to the criterion (44). In addition, we apply a continuation technique to IRSVM-1, IRSVM-2, and NPG, which is similar to the one used in APGL [34]. In detail, set $\lambda$ 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 $\lambda$ replaced by $\lambda_0$ 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 $\lambda$ replaced by $\lambda_1$ to obtain an approximate solution, denoted by $X^{(1)}$. This process is repeated until the target parameter $\lambda$ is reached and its approximate solution is found.

For IRSVM-1, IRSVM-2, and NPG, 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, 37], 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 $\epsilon^k = 0.5^k e$ for IRSVM-1, where $e$ is the all-ones vector. For IRSVM-2, $\epsilon$ is chosen to be the one within $10^{-6}$ to the supremum of $\varepsilon(X^0)$ that is defined in (50) 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 [32, 5], and say a matrix $M$ is *successfully recovered* by $X^*$ if the corresponding relative error is less than $10^{-3}$.

## 5.1 Matrix completion with random data

In this subsection we conduct numerical experiments to test the performance of IRSVM-1, IRSVM-2, and NPG for solving (69) on random data. We also compare our methods with three other related methods, that is, APGL [34], IRucLq-M [23], and tIRucLq-M [23]. APGL solves a (convex) nuclear norm relaxation of (69) while IRucLq-M is an iterative reweighted least squares method and tIRucLq-M is a variant of IRucLq-M.

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 $\Omega$ by a similar procedure as described in [27]. 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 $\Omega$ with sampling ratio $SR$ uniformly at random, where $SR = |\Omega|/(mn)$. In our experiment, we set $m = n = 200$ and generate $\Omega$ 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, NPG, APGL, IRucLq-M, and tIRucLq-M to solve (69) 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, and set truncation $= 1$, truncation_gap $= 10$, maxrank $= \lfloor 1.5r \rfloor$, $\mu_0 = 10^{-2}\|\mathcal{P}_\Omega(M)\|_F$, $\mu_k = \max(0.7\mu_{k-1}, 10^{-6}\|\mathcal{P}_\Omega(M)\|_F)$ for APGL. All other parameters for these three methods are chosen by default. For IRSVM-1, IRSVM-2 and NPG, 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, IRSVM-2, and NPG are generally better than the other three methods. For example, when $SR = 0.2$ and $r = 19$, our three methods are capable of recovering almost all instances while APGL and tIRucLq-M recover none of the instances and IRucLq-M only recovers about half of the instances. For the instances that are successfully recovered by all six methods, the CPU time of IRSVM-1, IRSVM-2 and NPG is less than that of other three methods. Nevertheless, for the instances that fail to be recovered by APGL, IRucLq-M or tIRucLq-M, the CPU time of these methods is less than that of our three methods. In addition, we observe that IRSVM-2 is slightly faster than IRSVM-1 and NPG for these instances.

## 5.2 Matrix completion with image data

In this subsection we compare the performance of IRSVM-1, IRSVM-2 and NPG with APGL, IRucLq-M and tIRucLq-M for solving a grayscale image inpainting problem [2], which was used in [23] for testing APGL 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 [35, 29], this

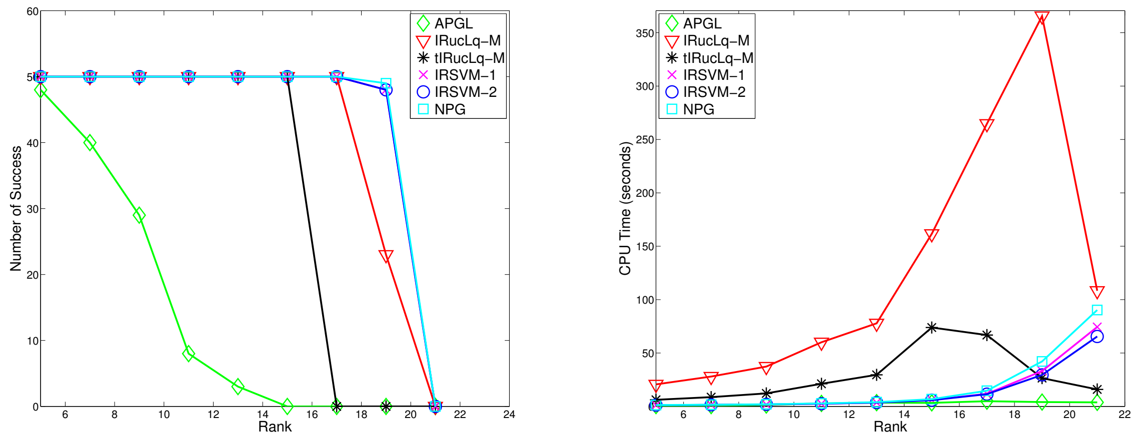



Figure 1: Comparison on random data with $SR = 0.2$.

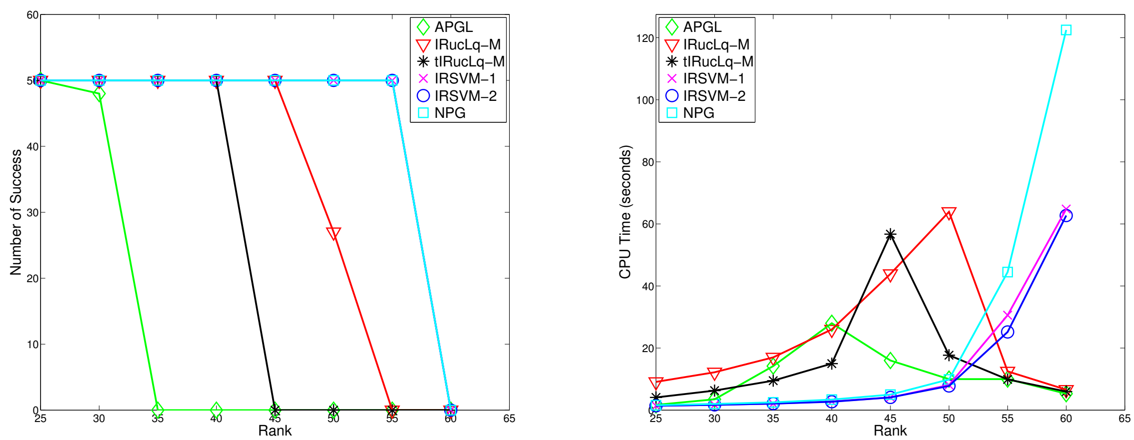



Figure 2: Comparison on random data with $SR = 0.5$.

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 $224 \times 224$ pixels and rank 28. The image "boat" is obtained by first applying the singular value decomposition to the original image with $512 \times 512$ pixels and then truncating the decomposition so that the resulting image has rank 40.

We first apply all six methods to solve the image inpainting problem with three different sample 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 APGL, IRucLq-M and tIRucLq-M. In addition, we set $\bar{\varepsilon} = 5 \times 10^{-3}$ for IRSVM-1, IRSVM-2 and NPG. The other parameter settings for all methods are the same as those used in the random data experiment. We present the results of this

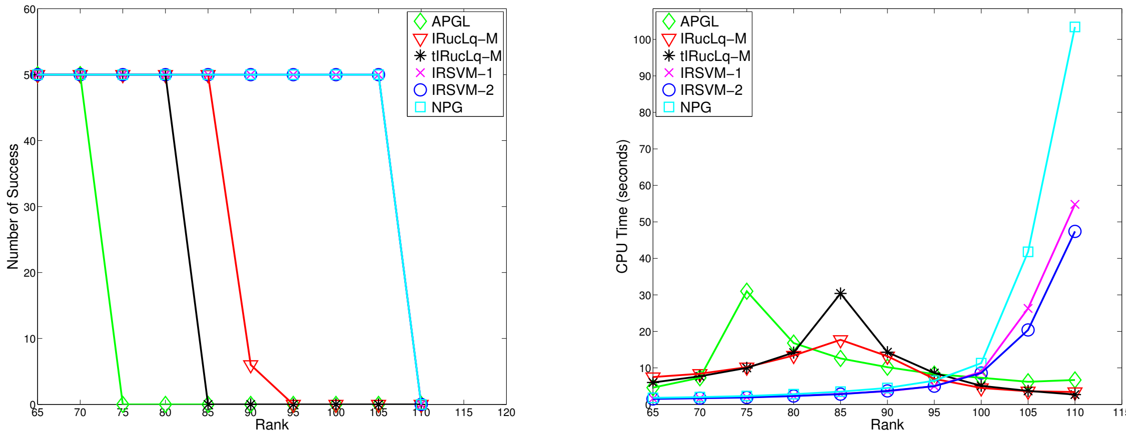
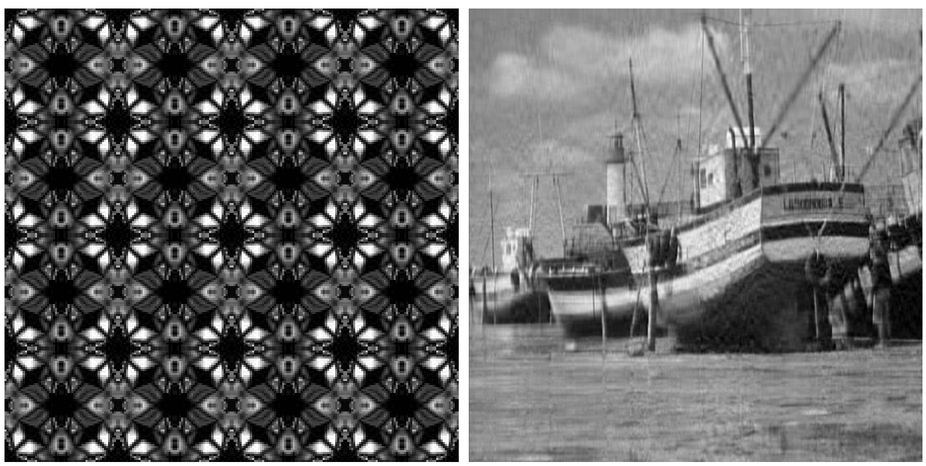



Figure 3: Comparison on random data with $SR = 0.8$.



(a) pattern (b) boat

Figure 4: Testing images. "pattern": grayscale image of $224 \times 224$ pixels with rank $= 28$. "boat": grayscale image of $512 \times 512$ pixels with rank $= 40$.

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 fourteen. 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, I RSVM-2 and NPG achieve smaller rel_err than the other three methods. The CPU time of these methods is generally less than that of IRucLq-M. Though APGL and tIRucLq-M outperform the other methods in terms of CPU time, their rel_err is much higher than that of the other four methods. In addition, we observe that IRSVM-1 and IRSVM-2 are slightly faster than NPG for these instances.

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

| Image | SR | APGL rel_err | APGL Time | IRucLq-M rel_err | IRucLq-M Time | tIRucLq-M rel_err | tIRucLq-M Time | IRSVM-1 rel_err | IRSVM-1 Time | IRSVM-2 rel_err | IRSVM-2 Time | NPG rel_err | NPG Time |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| "pattern" | 0.1 | 5.73e−1 | 1.5 | 4.02e−1 | 5.6 | 5.25e−1 | 1.4 | 3.19e−1 | 6.9 | 3.18e−1 | 5.1 | **3.14e−1** | 9.5 |
| | 0.2 | 2.15e−1 | 1.1 | 1.27e−1 | 7.6 | 2.29e−1 | 2.9 | **9.08e−2** | 2.7 | 9.19e−2 | 2.2 | 9.22e−2 | 5.7 |
| | 0.3 | 2.15e−1 | 1.2 | 4.53e−2 | 5.1 | 7.03e−2 | 2.2 | 3.13e−2 | 1.5 | 3.23e−2 | 1.4 | **2.99e−2** | 2.6 |
| "boat" | 0.1 | 2.53e−1 | 3.7 | 1.66e−1 | 158.4 | 2.34e−1 | 9.3 | **1.58e−1** | 45.1 | **1.58e−1** | 42.3 | 1.60e−1 | 53.5 |
| | 0.2 | 4.03e−2 | 6.4 | 3.72e−2 | 148.0 | 5.52e−2 | 10.5 | 1.39e−2 | 29.7 | **1.16e−2** | 33.0 | 1.36e−2 | 50.4 |
| | 0.3 | 1.64e−2 | 4.2 | 4.61e−3 | 97.8 | 5.84e−3 | 8.6 | 2.64e−3 | 9.4 | **1.21e−3** | 10.3 | 1.48e−3 | 26.6 |

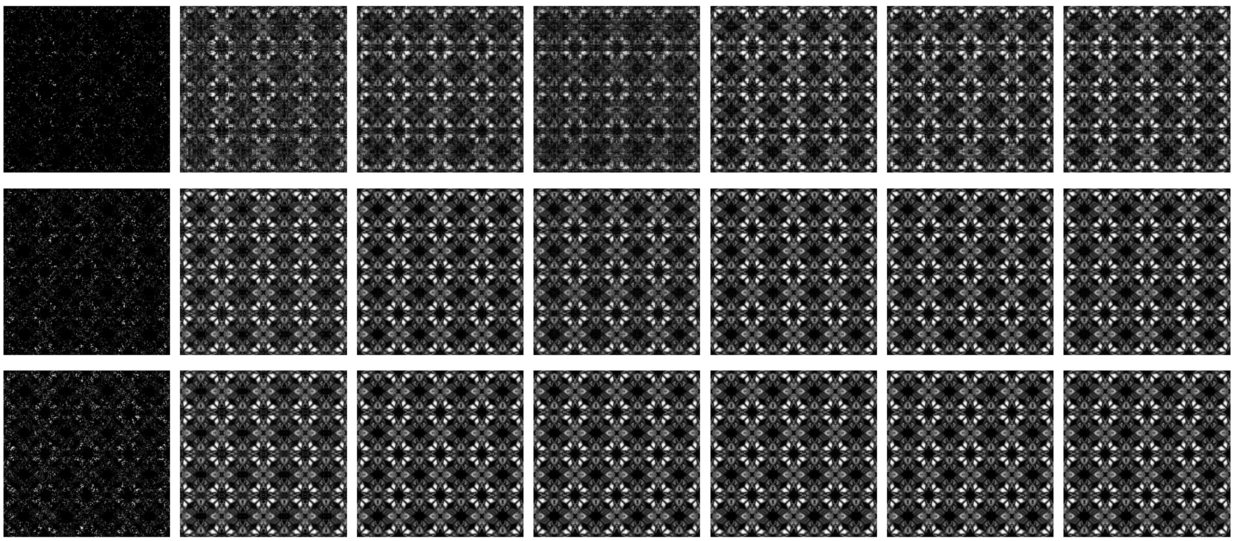


Figure 5: Results of image recovery. First column: the sample images with $SR = 0.1$, 0.2 and 0.3. The rest columns: the images recovered by APGL, IRucLq-M, tIRucLq-M, IRSVM-1, IRSVM-2 and NPG, respectively.

# 6 Concluding remarks

In this paper we studied general $l_p$ regularized unconstrained matrix minimization problems (1). In particular, we introduced a class of first-order stationary points for them. And we showed that the first-order stationary points introduced in [11] for an $l_p$ regularized *vector* minimization problem are equivalent to those of an $l_p$ regularized *matrix* minimization reformulation. Also, we established that any local minimizer of problem (1) must be a first-order stationary point. Moreover, we derived lower bounds for nonzero singular values of the first-order stationary points and hence also of the local minimizers for problem (1). The iterative reweighted singular value minimization (IRSVM) approaches were also proposed to solve these problems in which each subproblem has a closed-form solution. We showed that any accumulation point of the sequence generated by these methods is a first-order stationary point of the problems. In addition, we studied a nonmontone proximal gradient method for solving the $l_p$ matrix minimization problems and established its global convergence. Our computational results demonstrate that the IRSVM and NPG methods generally outperform some existing state-of-the-art methods in terms of solution quality and/or speed. Moreover, the IRSVM

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 APGL, IRucLq-M, tIRucLq-M, IRSVM-1, IRSVM-2 and NPG, respectively.

methods are slighly faster than the NPG method.

Besides $l_p$ regularizer, there are some other popular nonconvex regularizers for producing a sparse solution of a system or a vector optimization problem (e.g., see [16, 6, 39, 40, 20] 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 methods 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 V. Ballester. Image inpainting. *SIGGRAPH 2000*, New Orleans, USA, 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] E. J. Candés and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 2009.

[6] E. J. Candés, M. B. Wakin, S. P. Boyd. Enhancing sparsity by reweighted $\ell_1$ minimization. *J. Fourier Anal. Appl.*, 14(5–6): 877–905, 2008.

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

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

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

[10] X. Chen, D. Ge, Z. Wang, and Y. Ye. Complexity of unconstrained $L_2$-$L_p$ minimization. Accepted by *Math. Program.*.

[11] 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.

[12] 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.

[13] X. Chen and W. Zhou. Convergence of reweighted $l_1$ minimization algorithm for $l_2$-$l_p$ minimization. Accepted by *Comp. Optim. Appl.*.

[14] 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.

[15] 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.

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

[17] 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.

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

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

[20] 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*, 2013.

[21] 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.

[22] 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.

[23] 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.

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

[25] Z. Lu. Iterative Reweighted Minimization Methods for $l_p$ Regularized Unconstrained Nonlinear Programming. To appear in *Math. Program.*.

[26] Z. Lu and Y. Zhang. Penalty decomposition methods for rank minimization. Technical report, Department of Mathematics, Simon Fraser University, Canada, 2013.

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

[28] 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.

[29] 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.

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

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

[32] 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.

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

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

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

[36] 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.

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

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

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

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