

Research Article

An Efficient Method for Convex Constrained Rank Minimization Problems Based on DC Programming

Wanping Yang, Jinkai Zhao, and Fengmin Xu

School of Economics and Finance, Xi'an Jiaotong University, Xi'an 710061, China

Correspondence should be addressed to Fengmin Xu; fengminxu@mail.xjtu.edu.cn

Received 26 January 2016; Revised 19 May 2016; Accepted 2 June 2016

Academic Editor: Srdjan Stankovic

Copyright © 2016 Wanping Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The constrained rank minimization problem has various applications in many fields including machine learning, control, and signal processing. In this paper, we consider the convex constrained rank minimization problem. By introducing a new variable and penalizing an equality constraint to objective function, we reformulate the convex objective function with a rank constraint as a difference of convex functions based on the closed-form solutions, which can be reformulated as DC programming. A stepwise linear approximative algorithm is provided for solving the reformulated model. The performance of our method is tested by applying it to affine rank minimization problems and max-cut problems. Numerical results demonstrate that the method is effective and of high recoverability and results on max-cut show that the method is feasible, which provides better lower bounds and lower rank solutions compared with improved approximation algorithm using semidefinite programming, and they are close to the results of the latest researches.

1. Introduction

In recent decades, with the increase of the system acquisition and data services, the explosion of data poses challenges to storage, transmission, and processing, as well as device design. The burgeoning theory of sparse recovery reveals the outstanding sensing ability on the large scales of high-dimensional data. Recently, a new theory called compressed sensing (or compressive sampling (CS)) has appeared. And this theory is praised as a big deal in signal processing area. This approach can acquire the signals while compressing data properly. Its sampling frequency is lower than that of Nyquist, which makes the collections of high resolution signal possible. One noticeable merit of this approach is its ability to combine traditional data collection and data compression into a union when facing sparse representation signals. This merit means that sampling frequency of signals, time and computational cost consuming on data processing, and expense for data storage and transmission are all greatly reduced. Thus this approach leads signal processing to a new age. However, same with CS, Low Rank Matrix theory can also solve signal reconstruction problems that are related to theory and practice. It is also carried out through exerting

dimension reduction treatments on unknown multidimensional signals validly. And then use sparse concepts in a more broad sense to the problems. That is to say, reconstruct the original high-dimensional signals array completely or approximately by a small number of values measured by sparse (or reducing dimensions) sampling. Considering that both of these two theories are closely related to intrinsic sparsity of signals, we generally view them as sparse recovery problems. The main goal of sparse recovery problems is using less linear measurement data to recovery high-dimensional sparse signal. Its core mainly includes three steps: signal sparse representation, linear dimension reduction measure, and the nonlinear reconstruction of signal. The three parts for the sparse recovery theory are closely related; each part will have a significant impact on the quality of signal reconstruction. However, the nonlinear reconstruction of signal can be seen as a special case of rank minimization problem to be solved.

Constrained rank minimization problems have attracted a great deal of attention over recent years, which aim to find sparse solutions of a system. The problems are widely applicable in a variety of fields including machine learning [1, 2], control [3, 4], Euclidean embedding [5], image processing

[6], and finance [7, 8], to name but a few. There are some typical examples of constrained rank minimization problems; for example, one is affine matrix rank minimization problems [5, 9, 10]

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathcal{A}(X) - b\|_F^2 \\ \text{s.t.} \quad & \text{rank}(X) \leq k, \end{aligned} \quad (1)$$

where $X \in R^{m \times n}$ is the decision variable and the linear map $\mathcal{A} : R^{m \times n} \rightarrow R^p$ and vector $b \in R^p$ are known. Equation (1) includes matrix completion, based on affine rank minimization problem and compressed sensing. Matrix completion is widely used in collaborative filtering [11, 12], system identification [13, 14], sensor network [15, 16], image processing [17, 18], sparse channel estimation [19, 20], spectrum sensing [21, 22], and multimedia coding [23, 24]. Another is max-cut problems [25] in combinatorial optimization

$$\begin{aligned} \min \quad & \text{Tr}(LX) \\ \text{s.t.} \quad & \text{diag}(X) = e \\ & \text{rank}(X) = 1 \\ & X \geq 0, \end{aligned} \quad (2)$$

where $X = xx^T$, $x = (x_1, \dots, x_n)^T$, $L = (1/4)(\text{diag}(We) - W)$, $e = (1, 1, \dots, 1)^T$, W is the weight matrix, and so on. As a matter of fact, these problems can be written as a common convex constrained rank minimization model; that is,

$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & \text{rank}(X) \leq k \\ & X \in \Gamma \cap \Omega, \end{aligned} \quad (3)$$

where $f : R^{m \times n} \rightarrow R$ is a continuously differentiable convex function, $\text{rank}(X)$ is the rank of a matrix X , $k \geq 0$ is a given integer, Γ is a closed convex set, and Ω is a closed unitarily invariant convex set. According to [26], a set $C \subseteq R^{n \times n}$ is a unitarily invariant set if

$$\{UXV : U \in U^n, V \in U^n, X \in C\} = C, \quad (4)$$

where U^n denotes the set of all unitary matrices in $R^{n \times n}$. Various types of methods have been proposed to solve (3) or its special cases with different types of $\Gamma \cap \Omega$.

When $\Gamma \cap \Omega$ is $R^{m \times n}$, (3) is the unconstrained rank minimization problem. One approach is to solve this case by convex relaxation of $\text{rank}(\cdot)$. Cai et al. [9] proposed singular value thresholding (SVT) algorithm; Ma et al. [10] proposed fixed point continuation (FPC) algorithm. Another approach is to solve rank constraint directly. Jain et al. [27] proposed simple and efficient algorithm Singular Value Projection (SVP) based on the projected gradient algorithm. Haldar and Hernando [28] proposed an alternating least squares approach. Keshavan et al. [29] proposed an algorithm based on optimization over a Grassmann manifold.

When $\Gamma \cap \Omega$ is a convex set, (3) is a convex constrained rank minimization problem. One approach is to solve it by replacing rank constraint with other constraints, such as trace constraint and norm constraint. Based on this, several heuristic methods have been proposed in literature; for example, Nesterov et al. [30] proposed interior-point polynomial algorithms; Weimin [31] proposed an adaptive seminuclear norm regularization approach; in addition, semidefinite programming (SDP) has also been applied; see [32]. Another approach is to solve rank constraint directly. Grigoriadis and Beran [33] proposed alternating projection algorithms for linear matrix inequalities problems. Mohseni et al. [24] proposed penalty decomposition (PD) method based on penalty function. Gao and Sun [34] recently proposed the majorized penalty approach. Burer et al. [35] proposed nonlinear programming (NLP) reformulation approach.

In this paper, we consider problem (3). We introduce a new variable and penalize equality constraints to objective function. By this technique, we can obtain closed-form solution and use it to reformulate the objective function in problem (3) as a difference of convex functions; then (3) is reformulated as DC programming which can be solved via linear approximation method. A function is called DC if it can be represented as the difference of two convex functions. Mathematical programming problems dealing with DC functions are called DC programming problems. Our method is different from original PD [26]; for one thing we penalize equality constraints to objective function and keep other constraints, while PD penalizes all except rank constraint to objective function, including equality and inequality constraints; for another each subproblem of PD is approximately solved by a block coordinate descent method, while we approximate problem (3) to a convex optimization to be solved finally. Compared with PD method, our method uses the closed-form solutions to remove rank constraint, while PD needs to consider rank constraint in each iteration. It is known that rank constraint is a nonconvex and discontinuous constraint, which is not easy to deal. And due to the remaining convex constraints, the formulated programming can be solved by solving a series of convex programming. In our method, we use linear function instead of the last convex function to formulate the programming to convex programming. We test the performance of our method by applying it to affine rank minimization problems and max-cut problems. Numerical results show that the algorithm is feasible and effective.

The rest of this paper is organized as follows. In Section 1.1, we introduce the notation that is used throughout the paper. Our method is shown in Section 2. In Section 3, we apply it to some practical problems to test its performance.

1.1. Notations. In this paper, the symbol R^n denotes the n -dimensional Euclidean space, and the set of all $m \times n$ matrices with real entries is denoted by $R^{m \times n}$. Given matrices X and Y in $R^{m \times n}$, the standard inner product is defined by $\langle X, Y \rangle := \text{Tr}(XY^T)$, where $\text{Tr}(\cdot)$ denotes the trace of a matrix. The Frobenius norm of a real matrix X is defined as $\|X\|_F := \sqrt{\text{Tr}(XX^T)}$. $\|X\|_*$ denotes the nuclear norm of X , that is, the sum of singular values of X . The rank of a matrix X is denoted

by $\text{rank}(X)$. We use I_q to denote the identity matrix, whose dimension is q . Throughout this paper, we always assume that the singular values are arranged in nonincreasing order, that is, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 = \sigma_{r+1} = \dots = \sigma_{\min\{m,n\}}$. ∂f denotes the subdifferential of the function f . We denote the adjoint of A by A^* . $\|AA^T\|_{(k)}$ is used to denote Ky Fan k -norms. Let P_Ω be the projection onto the closed convex set Ω .

2. An Efficient Algorithm Based on DC Programming

In this section, we first reformulated problem (3) as a DC programming problem with the convex constrained set, which can be solved via stepwise linear approximation method. Then, the convergence is provided.

2.1. Model Transformation. Let $X = Y$. Problem (3) is rewritten as

$$\begin{aligned} \min \quad & f(X) \\ \text{s.t.} \quad & X = Y, \\ & \text{rank}(Y) \leq k, \\ & X \in \Gamma \cap \Omega. \end{aligned} \quad (5)$$

Adopting penalty function method to penalize $X = Y$ to objective function and choosing proper penalty parameter ρ , then (5) can be reformulated as

$$\begin{aligned} \min \quad & f(X) + \frac{\rho}{2} \|X - Y\|_F^2 \\ \text{s.t.} \quad & \text{rank}(Y) \leq k, \\ & X \in \Gamma \cap \Omega. \end{aligned} \quad (6)$$

Solving Y with fixed $X \in \Gamma \cap \Omega$, (6) can be treated as

$$\begin{aligned} \min_X \quad & \left\{ f(X) + \min_{\text{rank}(Y) \leq k} \frac{\rho}{2} \|X - Y\|_F^2 \right\} \\ \text{s.t.} \quad & X \in \Gamma \cap \Omega. \end{aligned} \quad (7)$$

In fact, given the singular value decomposition, $X = USV^T$, and singular values of X : $\sigma_j(X)$, $j = 1, \dots, \min\{m, n\}$, where $U \in R^{m \times m}$ and $V \in R^{n \times n}$ are orthogonal matrices, the minimal optimal problem with respect to Y with X fixed in (7), that is, $\min_{\text{rank}(Y) \leq k} (\rho/2) \|X - Y\|_F^2$, has closed-form solution $Y = UCV^T$ with

$$C = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \quad (8)$$

where $D = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k) \in R^{k \times k}$. Thus, by the closed-form solution and the unitary invariance of F -norm, we have

$$\begin{aligned} \min_{\text{rank}(Y) \leq k} \frac{\rho}{2} \|X - Y\|_F^2 &= \min_{\text{rank}(Y) \leq k} \frac{\rho}{2} \|U^T X V - U^T Y V\|_F^2 \\ &= \frac{\rho}{2} \|S - C\|_F^2 \\ &= \frac{\rho}{2} \|X\|_F^2 - \frac{\rho}{2} \sum_{i=1}^k \sigma_i^2(X). \end{aligned} \quad (9)$$

Substituting the above into problem (7), problem (3) is reformulated as

$$\begin{aligned} \min \quad & f(X) + \frac{\rho}{2} \|X\|_F^2 - \frac{\rho}{2} \sum_{i=1}^k \sigma_i^2(X) \\ \text{s.t.} \quad & X \in \Gamma \cap \Omega. \end{aligned} \quad (10)$$

Clearly, $f(X) + (\rho/2)\|X\|_F^2$ is a convex function about X . Define function $g : R^{m \times n} \rightarrow R$, $g(X) = \sum_{i=1}^k \sigma_i^2(X)$, $1 \leq k \leq \min\{m, n\}$. Next, we will study the property of $g(X)$. Problem (3) would be reformulated as a DC programming problem with convex constraints if $g(X)$ is a convex function, which can be solved via stepwise linear approximation method. In fact, it does hold, and $g(X)$ equals a trace function.

Theorem 1. *Let*

$$W = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix} \in R^{m \times m}, \quad (11)$$

where O denotes zero matrix whose dimension is determined depending on context; then, for any $X = USV^T \in R^{m \times n}$, the following hold:

(1) $g(X)$ is a convex function;

(2) $g(X) = \text{Tr}(X^T U W U^T X)$.

Proof. (1) Since $\{\sigma_j^2(X) \mid j = 1, \dots, \min\{m, n\}\}$ are eigenvalues of XX^T , then $g(X) = \sum_{i=1}^k \sigma_i^2(X) = \|XX^T\|_{(k)}$, where $\|\cdot\|_{(k)}$ is Ky Fan k -norms (defined in [36]). Clearly, $\|\cdot\|_{(k)}$ is a convex function. Let $\varphi(\cdot) = \|\cdot\|_{(k)}$, $1 \leq k \leq \min\{m, n\}$. For $\forall A, B \in R^{m \times n}$ and $\alpha \in [0, 1]$ we have

$$\begin{aligned} g(\alpha A + (1 - \alpha) B) &= \varphi\left((\alpha A + (1 - \alpha) B)(\alpha A + (1 - \alpha) B)^T\right) \\ &= \varphi\left(\alpha^2 AA^T + \alpha(1 - \alpha) AB^T + \alpha(1 - \alpha) BA^T \right. \\ &\quad \left. + (1 - \alpha)^2 BB^T\right) \leq \alpha^2 \varphi(AA^T) + \alpha(1 - \alpha) \\ &\quad \cdot \varphi(AB^T) + \alpha(1 - \alpha) \varphi(BA^T) + (1 - \alpha)^2 \varphi(BB^T). \end{aligned} \quad (12)$$

Note that $BA^T = (AB^T)^T$; thus they have the same nonzero singular values, and then $\varphi(BA^T) = \varphi(AB^T)$ holds. Together with (12), we have

$$\begin{aligned} & \varphi\left((\alpha A + (1 - \alpha)B)(\alpha A + (1 - \alpha)B)^T\right) \\ & \leq \alpha^2 \varphi(AA^T) + 2\alpha(1 - \alpha)\varphi(AB^T) \\ & \quad + (1 - \alpha)^2 \varphi(BB^T). \end{aligned} \quad (13)$$

It is known that for arbitrary matrices $A, B \in R^{m \times m}$,

$$\begin{aligned} \sigma_j(A^*B) & \leq \frac{1}{2}\sigma_j(A^*A + B^*B), \\ 1 \leq j \leq m, & \text{ (see Theorem IX.4.2 in [36]).} \end{aligned} \quad (14)$$

Further,

$$\sum_{j=1}^k \sigma_j(A^T B) \leq \sum_{j=1}^k \frac{1}{2} \sigma_j(A^*A + B^*B), \quad 1 \leq k \leq n. \quad (15)$$

That is,

$$\varphi(AB^T) \leq \frac{1}{2}\varphi(AA^T + BB^T), \quad (16)$$

where AA^T and BB^T are symmetric matrices. It is also known that, for arbitrary Hermitian matrices $AA^T, BB^T \in R^{m \times m}$, $k = 1, 2, \dots, m$, the following holds (see Exercise II.1.15 in [24])

$$\sum_{j=1}^k \sigma_j(AA^T + BB^T) \leq \sum_{j=1}^k \sigma_j(AA^T) + \sum_{j=1}^k \sigma_j(BB^T). \quad (17)$$

Thus

$$\begin{aligned} \varphi(A^T A + B^T B) & = \varphi(AA^T + BB^T) \\ & \leq \varphi(AA^T) + \varphi(BB^T). \end{aligned} \quad (18)$$

Together with (13) and (16), (18) yields

$$\begin{aligned} g(\alpha A + (1 - \alpha)B) & \leq \alpha g(AA^T) + (1 - \alpha)g(BB^T) \\ & = \alpha g(A) + (1 - \alpha)g(B). \end{aligned} \quad (19)$$

Hence, $g(X) = \sum_{i=1}^k \sigma_i^2(X)$ is a convex function.

(2) Let $\Theta = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_r)$ and r is the rank of X , $k \leq r$

$$\bar{S} = \begin{pmatrix} \Theta & O \\ O & O \end{pmatrix} \quad (20)$$

and we have

$$\begin{aligned} \text{Tr}(X^T U W U^T X) & = \text{Tr}(V \bar{S}^T U^T U W U^T U \bar{S} V^T) \\ & = \text{Tr}(V \bar{S}^T W \bar{S} V^T) = \text{Tr}(\bar{S}^T W \bar{S}). \end{aligned} \quad (21)$$

Moreover

$$\bar{S}^T W \bar{S} = \begin{pmatrix} \Theta & O \\ O & O \end{pmatrix}^T \begin{pmatrix} I_k & O \\ O & O \end{pmatrix} \begin{pmatrix} \Theta & O \\ O & O \end{pmatrix} = \begin{pmatrix} \bar{\Theta} & O \\ O & O \end{pmatrix}, \quad (22)$$

where $\bar{\Theta} = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_k^2)$.

Thus, $\text{Tr}(\bar{S}^T W \bar{S}) = \sum_{i=1}^k \sigma_i^2(X) = g(X)$ holds. \square

According to above theorem, (10) can be rewritten as

$$\begin{aligned} \min \quad & f(X) + \frac{\rho}{2} \|X\|_F^2 - \frac{\rho}{2} \text{Tr}(X^T U W U^T X) \\ \text{s.t.} \quad & X \in \Gamma \cap \Omega. \end{aligned} \quad (23)$$

Since the objective function of (23) is a difference of convex functions, the constrained set $\Gamma \cap \Omega$ is a closed convex constraint, so (23) is DC programming. Next, we will use stepwise linear approximate method to solve it.

2.2. Stepwise Linear Approximative Algorithm. In this subsection, we use stepwise linear approximation method [37] to solve reformulated model (23). Let $F(X)$ represent objective function of model (23); that is,

$$F(X) = f(X) + \frac{\rho}{2} \|X\|_F^2 - \frac{\rho}{2} \text{Tr}(X^T U W U^T X), \quad (24)$$

and the linear approximation function of $F(X)$ can be defined as

$$\begin{aligned} H(X, X^i) & = f(X) + \frac{\rho}{2} \|X\|_F^2 \\ & \quad - \frac{\rho}{2} \{g(X^i) + \langle \partial g(X^i), X - X^i \rangle\}, \end{aligned} \quad (25)$$

where $(\partial g(X)/\partial X)(\partial \text{Tr}(X^T U W U^T X)/\partial X) = 2U W U^T X$ (see 5.2.2 in [36]).

Stepwise linear approximative algorithm for solving model (23) can be described as follows.

Given the initial matrix $X^0 \in \Gamma \cap \Omega$, penalty parameter $\rho > 0$, rank $k > 0$, $\tau > 1$, set $i = 0$.

Step 1. Compute the single value decomposition of X^i , $U^i \Sigma^i V^i = X^i$.

Step 2. Update

$$X^{i+1} \in \arg \min_{X \in \Gamma \cap \Omega} \{H(X, X^i)\} \quad (26)$$

Step 3. If stopping criterion is satisfied, stop; else set $i \leftarrow i + 1$ and $\rho \leftarrow (1 + \tau)\rho$ and go to Step 1.

Remark

(1) Subproblem (26) can be done by solving a series of convex subproblems. The efficient method for convex programming is multiple and it has perfect theory guarantees.

- (a) When $\Gamma \cap \Omega$ is $R^{m \times n}$, closed-form solutions can be obtained; for example, when solving model (1), $X^{i+1} = (A^*A + \rho I)^{-1}(\rho U^i W (U^i)^T X^i)$ are obtained by the first-order optimality conditions

$$0 \in \partial f(X) + \rho X - \rho U^i W (U^i)^T X^i. \quad (27)$$

- (b) When $\Gamma \cap \Omega$ is a general convex set, the closed-form solution is hard to get, but subproblem (26) can be solved by some convex optimization software programs (such as CVX and CPLEX). According to different closed convex constraint sets, we can pick and choose suitable convex programming algorithm to combine with our method. The toolbox listed in this paper is just an alternative choice.

- (2) The stop rule that we adopted here is $\|X^{i+1} - X^i\|_F \leq \varepsilon$ with some ε small enough.

2.3. Convergence Analysis. In this subsection, we analyze the convergence properties of the sequence $\{X^i\}$ generated by above algorithm. Before we prove the main convergence result, we give the following lemma which is analogous to Lemma 1 in [10].

Lemma 2. For any Y_1 and $Y_2 \in R^{m \times n}$

$$\|\bar{Y}_1 - \bar{Y}_2\|_F \leq \|Y_1 - Y_2\|_F, \quad (28)$$

where \bar{Y}_i is the pre- k singular values approximate of Y_i ($i = 1, 2$).

Proof. Without loss of generality, we assume $m \leq n$ and the singular value decomposition $Y_1 = U_1 S_1 V_1^T Y_2 = U_2 S_2 V_2^T$ with

$$S_1 = \begin{pmatrix} \text{diag}(\sigma) & 0 \\ 0 & 0 \end{pmatrix} \in R^{m \times n} \quad (29)$$

$$S_2 = \begin{pmatrix} \text{diag}(\gamma) & 0 \\ 0 & 0 \end{pmatrix} \in R^{m \times n},$$

where $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_s)$, $\sigma_1 \geq \dots \geq \sigma_s > 0$ and $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_t)$, $\gamma_1 \geq \dots \geq \gamma_t > 0$. \bar{Y}_i can be written as $\bar{Y}_i = U_i W S_i V_i^T$, ($i = 1, 2$). Note that U_1, V_1, U_2, V_2 are orthogonal matrices; we have

$$\begin{aligned} & \|Y_1 - Y_2\|_F^2 - \|\bar{Y}_1 - \bar{Y}_2\|_F^2 \\ &= \text{Tr}((Y_1 - Y_2)^T (Y_1 - Y_2)) \\ & \quad - \text{Tr}((\bar{Y}_1 - \bar{Y}_2)^T (\bar{Y}_1 - \bar{Y}_2)) \\ &= \text{Tr}(Y_1^T Y_1 - \bar{Y}_1^T \bar{Y}_1 + Y_2^T Y_2 - \bar{Y}_2^T \bar{Y}_2) \\ & \quad - 2 \text{Tr}(Y_1^T Y_2 - \bar{Y}_1^T \bar{Y}_2) \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^s \sigma_i^2 - \sum_{i=1}^k \sigma_i^2 + \sum_{i=1}^t \gamma_i^2 - \sum_{i=1}^k \gamma_i^2 \\ & \quad - 2 \text{Tr}(Y_1^T Y_2 - \bar{Y}_1^T \bar{Y}_2). \end{aligned} \quad (30)$$

And we note that

$$\begin{aligned} \text{Tr}(Y_1^T Y_2 - \bar{Y}_1^T \bar{Y}_2) &= \text{Tr}((Y_1 - \bar{Y}_1)^T (Y_2 - \bar{Y}_2)) \\ & \quad + (Y_1 - \bar{Y}_1)^T \bar{Y}_2 + \bar{Y}_1^T (Y_2 - \bar{Y}_2) \\ &= \text{Tr}(V_1 (S_1 - WS_1)^T U_1^T U_2 (S_2 - WS_2) V_2^T \\ & \quad + V_1 (S_1 - WS_1)^T U_1^T U_2 WS_2 V_2^T \\ & \quad + V_1 (WS)^T U_1^T U_2 (S_2 - WS_2) V_2^T) \\ &= \text{Tr}((S_1 - WS_1)^T U (S_2 - WS_2) V^T \\ & \quad + (S_1 - WS_1)^T U WS_2 V^T \\ & \quad + (WS_1)^T U (S_2 - WS_2) V^T), \end{aligned} \quad (31)$$

where $U = U_1^T U_2, V = V_1^T V_2$ are also orthogonal matrices.

Next we estimate an upper bound for $\text{Tr}(Y_1^T Y_2 - \bar{Y}_1^T \bar{Y}_2)$. According to Theorem 7.4.9 in [38], we get that an orthogonal matrix U is a maximizing matrix for the problem $\max\{\text{Tr}(AU): U \text{ is orthogonal}\}$, if and only if AU is positive semidefinite matrix. Thus, $\text{Tr}((S_1 - WS_1)^T U (S_2 - WS_2) V^T)$, $\text{Tr}((S_1 - WS_1)^T U WS_2 V^T)$, and $\text{Tr}((WS)^T U (S_2 - WS_2) V^T)$ achieve their maximum, if and only if $(S_1 - WS_1)^T U (S_2 - WS_2) V^T$, $(S_1 - WS_1)^T U WS_2 V^T$, and $(WS)^T U (S_2 - WS_2) V^T$ are all positive semidefinite. It is known that when AB is positive semidefinite,

$$\text{Tr}(AB) = \sum_i \sigma_i(AB) \leq \sum_i \sigma_i(A) \sigma_i(B). \quad (32)$$

Applying (32) to above three terms, we have

$$\begin{aligned} & \text{Tr}((S_1 - WS_1)^T U (S_2 - WS_2) V^T) \\ & \leq \sum_i \sigma_i(S_1 - WS_1) \sigma_i(S_2 - WS_2) \\ & \quad \text{Tr}((S_1 - WS_1)^T U WS_2 V^T) \\ & \leq \sum_i \sigma_i(S_1 - WS_1) \sigma_i(WS_2) \\ & \quad \text{Tr}((WS)^T U (S_2 - WS_2) V^T) \\ & \leq \sum_i \sigma_i(WS_1) \sigma_i(S_2 - WS_2). \end{aligned} \quad (33)$$

Hence, without loss of generality, assuming $k \leq s \leq t$, we have

$$\begin{aligned} & \|Y_1 - Y_2\|_F^2 - \|\bar{Y}_1 - \bar{Y}_2\|_F^2 \\ & \geq \sum_{i=1}^s \sigma_i^2 - \sum_{i=1}^k \sigma_i^2 + \sum_{i=1}^t \gamma_i^2 - \sum_{i=1}^k \gamma_i^2 - 2 \left(\sum_{i=k+1}^s \sigma_i \gamma_i \right) \\ & = \sum_{i=k+1}^s (\sigma_i - \gamma_i)^2 + \sum_{i=s+1}^t \gamma_i^2 \geq 0 \end{aligned} \quad (34)$$

which implies that $\|\bar{Y}_1 - \bar{Y}_2\|_F \leq \|Y_1 - Y_2\|_F$ holds. \square

Next, we claim the value of iterations of stepwise linear approximative algorithm is nonincreasing.

Theorem 3. *Let $\{X^i\}$ be the iterative sequence generated by stepwise linear approximative algorithm. Then $\{F(X^i)\}$ is monotonically nonincreasing.*

Proof. Since $X^{i+1} \in \arg \min\{H(X, X^i) \mid X \in \Gamma \cap \Omega\}$, we have

$$\begin{aligned} H(X^{i+1}, X^i) & \leq H(X^i, X^i) \\ & = f(X^i) + \frac{\rho}{2} \|X^i\|_F^2 - \frac{\rho}{2} g(X^i) \\ & = F(X^i), \end{aligned} \quad (35)$$

and $\partial g(X)$ is the subdifferential of $g(X)$, which implies

$$\forall XY, \quad g(X) \geq g(Y) + \langle \partial g(Y), X - Y \rangle. \quad (36)$$

Let $X = X^{i+1}$, $Y = X^i$ in above inequality; we have

$$g(X^{i+1}) \geq g(X^i) + \langle \partial g(X^i), X^{i+1} - X^i \rangle. \quad (37)$$

Therefore,

$$\begin{aligned} H(X^{i+1}, X^i) & = f(X^{i+1}) + \frac{\rho}{2} \|X^{i+1}\|_F^2 \\ & \quad - \frac{\rho}{2} \{g(X^i) + \langle \partial g(X^i), X - X^i \rangle\} \\ & \geq f(X^{i+1}) + \frac{\rho}{2} \|X^{i+1}\|_F^2 - \frac{\rho}{2} g(X^{i+1}) \\ & = F(X^{i+1}). \end{aligned} \quad (38)$$

Consequently,

$$F(X^{i+1}) \leq H(X^{i+1}, X^i) \leq H(X^i, X^i) = F(X^i). \quad (39)$$

We immediately obtain the conclusion as desired. \square

Next, we show the convergence of iterative sequence generated by stepwise linear approximative algorithm when solving (23). Theorem 4 shows the convergence when solving (23) with $\Gamma \cap \Omega = R^{m \times n}$.

Theorem 4. *Let Y be the set of stable points of problem (23) with no constraints; then the iterative sequence $\{X^i\}$ generated by linear approximative algorithm converges to some $X^{\text{opt}} \in Y$ when $\rho \rightarrow \infty$.*

Proof. Assuming that $X^{\text{opt}} \in Y$ is any optimal solution of problem (23) with $\Gamma \cap \Omega = R^{m \times n}$, according to the first-order optimality conditions, we have

$$0 = \partial f(X^{\text{opt}}) + \rho X^{\text{opt}} - \rho U^{\text{opt}T} W (U^{\text{opt}})^T X^{\text{opt}}. \quad (40)$$

Since $X^{i+1} \in \arg \min f(X) + (\rho/2)\|X\|_F^2 - (\rho/2)\{g(X^i) + \langle \partial g(X^i), X - X^i \rangle\}$, then

$$0 = \partial f(X^{i+1}) + \rho X^{i+1} - \rho U^i W (U^i)^T X^i. \quad (41)$$

Equation (41) subtracts (40):

$$\begin{aligned} X^{i+1} - X^{\text{opt}} & = \frac{1}{\rho} (\partial f(X^{i+1}) - \partial f(X^{\text{opt}})) \\ & \quad + U^{\text{opt}T} W (U^{\text{opt}})^T X^{\text{opt}} \\ & \quad - U^i W (U^i)^T X^i \end{aligned} \quad (42)$$

in which $U^{\text{opt}T} W (U^{\text{opt}})^T X^{\text{opt}} - U^i W (U^i)^T X^i = U^{\text{opt}T} W \sum^{\text{opt}} (V^{\text{opt}})^T - U^i W \sum^i (V^{\text{opt}})^T$.

By Lemma 2, we have

$$\begin{aligned} & \left\| U^{\text{opt}T} W (U^{\text{opt}})^T X^{\text{opt}} - U^i W (U^i)^T X^i \right\|_F \\ & = \left\| U^{\text{opt}T} W \sum^{\text{opt}} (V^{\text{opt}})^T - U^i W \sum^i (V^i)^T \right\|_F \\ & \leq \|X^i - X^{\text{opt}}\|_F. \end{aligned} \quad (43)$$

Norming both sides of (42), by the triangle inequalities property of norms, we have

$$\begin{aligned} \|X^{i+1} - X^{\text{opt}}\|_F & = \lim_{\rho \rightarrow \infty} \|X^{i+1} - X^{\text{opt}}\|_F \\ & \leq \lim_{\rho \rightarrow \infty} \left(\frac{1}{\rho} \|\partial f(X^{i+1}) - \partial f(X^{\text{opt}})\|_F \right. \\ & \quad \left. + \|X^i - X^{\text{opt}}\|_F \right) = \lim_{\rho \rightarrow \infty} \frac{1}{\rho} \|\partial f(X^{i+1}) \\ & \quad - \partial f(X^{\text{opt}})\|_F + \lim_{\rho \rightarrow \infty} \|X^i - X^{\text{opt}}\|_F = \|X^i \\ & \quad - X^{\text{opt}}\|_F, \end{aligned} \quad (44)$$

which implies that the sequence $\{\|X^i - X^{\text{opt}}\|_F\}$ is monotonically nonincreasing when $\rho \rightarrow \infty$. We then observe that $\{\|X^i - X^{\text{opt}}\|_F\}$ is bounded, and hence $\lim_{i \rightarrow \infty} \|X^i - X^{\text{opt}}\|_F$ exists, denoted by $\|\bar{X} - X^{\text{opt}}\|_F$, where \bar{X} is the limited point of $\{X^i\}$. Let $i \rightarrow \infty$ in (41); by the continuity, we get

$0 = \partial f(\bar{X}) + \rho \bar{X} - \rho \bar{U} W \bar{U}^T \bar{X}$, which implies that $\bar{X} \in Y$. By setting $\bar{X} = X^{\text{opt}}$, we have

$$\lim_{i \rightarrow \infty} \|X^i - X^{\text{opt}}\|_F = \|\bar{X} - X^{\text{opt}}\|_F = 0, \quad (45)$$

which completes the proof. \square

Now, we are ready to give the convergence of iterative sequence when solving (23) with general closed convex $\Gamma \cap \Omega$.

Theorem 5. *Let ξ be the set of optimal solutions of problem (23); then the iterative sequence $\{Y^i\}$ generated by stepwise linear approximative algorithm converges to some $Y^{\text{opt}} \in \xi$ when $\rho \rightarrow \infty$.*

Proof. Generally speaking, subproblem (26) can be solved by projecting a solution obtained via solving a problem with no constraints onto a closed convex set; that is, $Y^{i+1} = P_{\Gamma \cap \Omega}(X^{i+1})$, where $X^{i+1} \in \arg \min\{H(X, X^i)\}$. According to Theorem 4, the sequence $\{\|X^i - X^{\text{opt}}\|_F\}$ is monotonically nonincreasing and convergent when $\rho \rightarrow \infty$; together with the well known fact that P is a nonexpansive operator, we have

$$\begin{aligned} \|Y^i - Y^{\text{opt}}\|_F &= \|P_{\Gamma \cap \Omega}(X^i) - P_{\Gamma \cap \Omega}(X^{\text{opt}})\|_F \\ &\leq \|X^i - X^{\text{opt}}\|_F, \end{aligned} \quad (46)$$

where Y^{opt} is the projection of X^{opt} onto $\Gamma \cap \Omega$. Clearly, $Y^{\text{opt}} \in \xi$.

Hence, $0 \leq \lim_{i \rightarrow \infty} \|Y^i - Y^{\text{opt}}\|_F \leq \lim_{i \rightarrow \infty} \|X^i - X^{\text{opt}}\|_F = 0$, which implies that Y^i converge to Y^{opt} . We immediately obtain the conclusion. \square

3. Numerical Results

In this section, we demonstrate the performance of our method proposed in Section 2 by applying it to solve affine rank minimization problems (image reconstruction and matrix completion) and max-cut problems. The codes implemented in this section are written in MATLAB and all experiments are performed in MATLAB R2013a running Windows 7.

3.1. Image Reconstruction Problems. In this subsection, we apply our method to solve one case of affine rank minimization problems. It can be formulated as

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathcal{A}(X) - b\|_F^2 \\ \text{s.t.} \quad & \text{rank}(X) \leq k \\ & X \in R^{m \times n}, \end{aligned} \quad (47)$$

where $X \in R^{m \times n}$ is the decision variable and the linear map $\mathcal{A} : R^{m \times n} \rightarrow R^p$ and vector $b \in R^p$ are known. Clearly, (47) is a special case of problem (3). Hence, our method can be suitably applied to (47).

Recht et al. [39] have demonstrated that when we sample linear maps from a class of probability distributions, then they

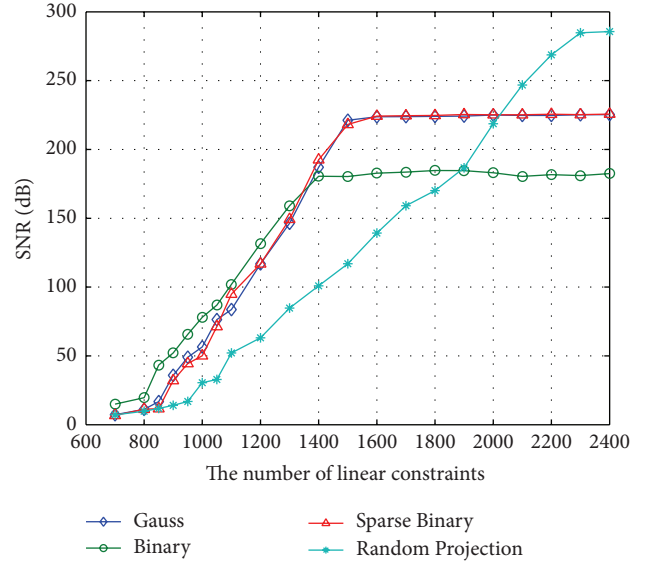


FIGURE 1: Performance curves under four different measurement matrices: SNR versus the number of linear constraints.

would obey the Restricted Isometry Property. There are two ingredients for a random linear map to be nearly isometric. First, it must be isometric in expectation. Second, the probability of large distortions of length must be exponentially small. In our numerical experiments, we use four nearly isometric measurements. The first one is the ensemble with independent, identically distributed (i.i.d.) Gaussian entries. The second one has entries sampled from an i.i.d. symmetric Bernoulli distribution, which we call Bernoulli. The third one has zeros in two-thirds of the entries and others sampled from an i.i.d. symmetric Bernoulli distribution, which we call Sparse Bernoulli. The last one has an orthonormal basis, which we call Random Projection.

Now we conduct numerical experiments to test the performance of our method for solving (47). To illustrate the scaling of low rank recovery for a particular matrix X , consider the MIT logo presented in [39]. The image has 46 rows and 81 columns (the total number of pixels is $46 \times 81 = 3726$). Since the logo only has 5 distinct rows, it has rank 5. We sample it using Gaussian i.i.d., Bernoulli i.i.d., Sparse Bernoulli i.i.d., and Random Projection measurements matrix with the number of linear constraints ranging between 700 and 2400. We declared MIT logo X to be recovered if $\|X - X^{\text{opt}}\|_F / \|X\|_F \leq 10^{-3}$; that is, $\text{SNR} \geq 60$ dB ($\text{SNR} = -20 \log(\|X - X^{\text{opt}}\|_F / \|X\|_F)$). Figures 1 and 2 show the results of our numerical experiments.

Figure 1 plots the curves of SNR under four different measurement matrices mentioned above with different numbers of linear constraints. The numerical values on the x -axis represent the number of linear constraints p , that is, number of measurements, while those on y -axis represent the Signal to Noise Ratio between the true original X and X^{opt} recovered by our method. Noticeably, Gauss and Sparse Binary can successfully reconstruct MIT logo images from about 1000 constraints, while Random Projection requires about

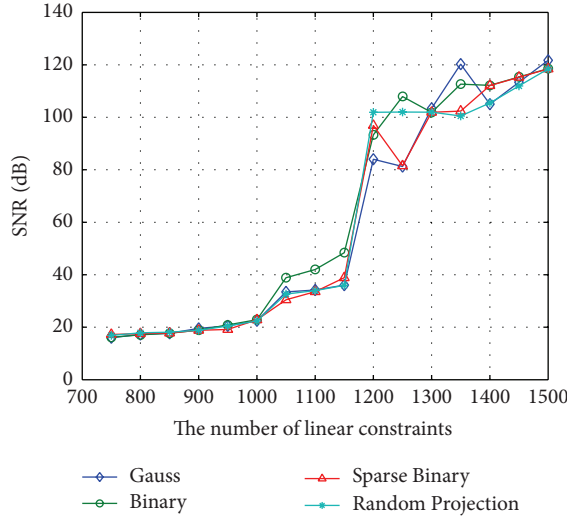


FIGURE 2: Performance curves of nuclear norm heuristic method under four different measurement matrices: SNR versus the number of linear constraints.

1200 constraints. Binary requires the minimal constraints about 950. The Gauss and Sparse Binary measurements reach their maximal SNR of 220 dB at around 1500 constraints; the Binary measurements can reach its maximal SNR of 180 dB at around 1400 constraints. Random Projection has more large phase transition point but has the best recovery error. We observe a sharp transition to perfect recovery near 1000 measurements lower than $2k(m + n - k)$. Compared with Figure 2, which shows results of nuclear norm heuristic method in [39] and the method that is a kind of convex relaxations, there is a sharp transition to perfect recovery near 1200 measurements which is approximately equal to $2k(m + n - k)$. The comparison shows our method has better performance that it can recover an image with less constraints; namely, it needs less samples. In addition, we observe that SNR of nuclear norm heuristic method achieves 120 dB when the number of measurements approximately equals 1500, while our method just needs 1200 measurements, which decrease by 20% compared with 1500. If the number of measurements is 1500, accuracy of our method would be higher. Among precisions of four different measurement matrixes, the smallest one is precision of Random Projection. However, the precision value is also higher than 120 dB. The performance illustrates that our method is more accurate. Observing the curve of between 1200 and 1500, we can find that our curve rises smoothly compared with the curve in the same interval in Figure 2; this shows that our method is stable for reconstructing image.

Figure 3 shows recovered images under four different measurement matrices with 800, 900, and 1000 linear constraints. And Figure 4 shows recovered images by our method and that by SeDuMi. We chose to use this interior-point method because it yields the highest accuracy. Distinctly, the recovered images by SeDuMi have some shadow, while ours is clear. Moreover, as far as time is concerned, our method is more fast and the computational time is not extremely

TABLE 1: DC versus SeDuMi in time (second).

p	800	900	1000	1100	1200	1500
DC	46.82	50.49	50.78	51.38	51.96	54.09
SeDuMi	93.26	120.16	138.24	159.44	161.76	245.38

sensitive to p compared to SeDuMi. The reason is that closed-form solutions can be obtained when we solve this problem. Comparisons are shown in Table 1.

3.2. *Matrix Completion Problems.* In this subsection, we apply our method to matrix completion problems. It can be reformulated as

$$\begin{aligned} \min \quad & \frac{1}{2} \|P_{\Omega}(X) - P_{\Omega}(M)\|_F^2 \\ \text{s.t.} \quad & \text{rank}(X) \leq k \\ & X \in R^{m \times n}, \end{aligned} \quad (48)$$

where X and M are both $m \times n$ matrices and Ω is a subset of index pairs (i, j) . Up to now, numerous methods were proposed to solve the matrix completion problem (see, e.g., [9, 10, 29, 31]). It is easy to see that problem (48) is a special case of general rank minimization problem (3) with $f(X) = (1/2)\|P_{\Omega}(X) - P_{\Omega}(M)\|_F^2$.

In our experiments, we first created random matrices $M_L \in R^{m \times k}$ and $M_R \in R^{m \times k}$ with i.i.d. Gaussian entries and then set $M = M_L M_R^T$. We sampled a subset Ω of p entries uniformly at random. We compare against singular value thresholding (SVT) method [9], FPC, FPCA [10], and the OptSpace (OPT) method [29] using

$$\frac{\|P_{\Omega}(X) - P_{\Omega}(M)\|_F}{\|P_{\Omega}(M)\|_F} \leq 10^{-3} \quad (49)$$

as a stop criterion. The maximum number of iterations is no more than 1500 in all methods; other parameters are default. We report result averaged over 10 runs.

We use $SR = p/(mn)$ to denote sampling ratio and $FR = k(m + n - k)/p$ to represent the dimension of the set of rank k matrices divided by the number of measurements. NS and AT denote number of matrices that are recovered successfully and average times (seconds) that are successfully solved, respectively. RE represents the average relative error of the successful recovered matrices. We use the relative error

$$\|M - X^*\|_F / \|M\|_F \quad (50)$$

to estimate the closeness of X^* to M , where X^* is the optimal solution to (48) produced by our method. We declared M to be recovered if the relative error was less than 10^{-3} .

Table 2 shows that our method has the close performance with OptSpace in column of NS. These performances are better than those of SVT, FPC, and FPCA. However, SVT, FPC, and FPCA cannot obtain the optimal matrix within 1500 iterations which shows that our method has advantage in finding the neighborhood of optimal solution. The results do not mean that these algorithms are invalid, but the maximum

TABLE 2: Numerical results for problems with $m = n = 40$ and $SR = 0.5$.

Alg	k	FR	NS	AT	RE	k	FR	NS	AT	RA
DC			10	0.02	$1.23E - 04$			9	0.12	$4.64E - 04$
SVT			10	0.11	$1.59E - 04$			0	—	—
FPC	1	0.0988	9	0.13	$6.52E - 04$	6	0.5550	0	—	—
FPCA			10	0.03	$3.98E - 05$			0	—	—
OptSpace			10	0.01	$1.31E - 04$			10	0.21	$4.05E - 04$
DC			10	0.03	$1.50E - 04$			9	0.17	$6.11E - 04$
SVT			9	0.18	$1.83E - 04$			0	—	—
FPC	2	0.1950	8	0.14	$6.52E - 04$	7	0.6388	0	—	—
FPCA			10	0.02	$5.83E - 05$			0	—	—
OptSpace			10	0.01	$1.51E - 04$			9	0.37	$5.37E - 04$
DC			10	0.05	$2.68E - 04$			7	0.28	$7.90E - 04$
SVT			2	0.36	$2.47E - 04$			0	—	—
FPC	3	0.2888	5	0.25	$8.58E - 04$	8	0.7200	0	—	—
FPCA			10	0.06	$1.38E - 04$			0	—	—
OptSpace			10	0.04	$1.75E - 04$			9	0.82	$6.68E - 04$
DC			10	0.06	$2.84E - 04$			4	0.40	$8.93E - 04$
SVT			1	0.62	$1.38E - 04$			0	—	—
FPC	4	0.3800	1	0.25	$8.72E - 04$	9	0.7987	0	—	—
FPCA			0	—	—			0	—	—
OptSpace			10	0.05	$2.68E - 04$			4	1.48	$9.28E - 04$
DC			10	0.08	$3.82E - 04$					
SVT			0	—	—					
FPC	5	0.4688	0	—	—					
FPCA			0	—	—					
OptSpace			10	0.09	$3.16E - 04$					

TABLE 3: Numerical results for problems with $m = n = 100$ and $SR = 0.2$.

Alg	k	FR	NS	AT	RE	k	FR	NS	AT	RA
DC			10	0.30	$1.91E - 04$			10	1.50	$5.05E - 04$
SVT			7	1.82	$2.04E - 04$			0	—	—
FPC	1	0.0995	10	3.14	$8.06E - 04$	5	0.4875	0	—	—
FPCA			10	0.12	$1.73E - 04$			0	—	—
OptSpace			10	0.03	$1.71E - 04$			10	0.58	$5.57E - 04$
DC			10	0.77	$3.34E - 04$			6	2.82	$7.45E - 04$
SVT			0	—	—			0	—	—
FPC	2	0.1980	0	—	—	6	0.5820	0	—	—
FPCA			8	0.21	$2.97E - 04$			0	—	—
OptSpace			10	0.05	$2.27E - 04$			9	0.97	$6.57E - 04$
DC			10	0.64	$3.03E - 04$			5	3.24	$7.66E - 04$
SVT			0	—	—			0	—	—
FPC	3	0.2955	0	—	—	7	0.6755	0	—	—
FPCA			6	0.09	$2.20E - 04$			0	—	—
OptSpace			10	0.12	$3.76E - 04$			5	1.89	$7.95E - 04$
DC			9	1.16	$4.40E - 04$					
SVT			0	—	—					
FPC	4	0.3920	0	—	—					
FPCA			0	—	—					
OptSpace			10	0.22	$3.85E - 04$					



FIGURE 3: Recovered images under four measurement matrices with different numbers of linear constraints. From left to right: the original, results under Gauss, Binary, Sparse Binary, and Random Projection measurement. From top to bottom: $p = 800$, $p = 900$, and $p = 1000$.

TABLE 4: Numerical results for problems with $m = n = 100$ and $SR = 0.3$.

Alg	k	FR	NS	AT	RE	k	FR	NS	AT	RA
DC			10	0.19	$1.49E - 04$			10	0.61	$3.64E - 04$
SVT			10	0.68	$1.66E - 04$			0	—	—
FPC	1	0.0663	10	1.33	$4.47E - 04$	7	0.4530	0	—	—
FPCA			10	0.17	$6.12E - 05$			0	—	—
OptSpace			10	0.05	$1.29E - 04$			10	0.47	$3.39E - 04$
DC			10	0.20	$1.69E - 04$			10	0.87	$4.70E - 04$
SVT			8	1.39	$1.66E - 04$			0	—	—
FPC	2	0.1320	10	2.18	$4.84E - 04$	8	0.5120	0	—	—
FPCA			10	0.16	$7.68E - 05$			0	—	—
OptSpace			10	0.04	$1.48E - 04$			10	0.63	$3.67E - 04$
DC			10	0.25	$1.92E - 04$			9	1.03	$4.85E - 04$
SVT			4	1.77	$1.23E - 04$			0	—	—
FPC	3	0.1970	9	4.32	$5.82E - 04$	9	0.5730	0	—	—
FPCA			9	0.19	$1.21E - 04$			0	—	—
OptSpace			10	0.07	$1.98E - 04$			10	1.40	$4.87E - 04$
DC			10	0.32	$2.30E - 04$			8	1.36	$6.21E - 04$
SVT			4	9.82	$4.67E - 04$			0	—	—
FPC	4	0.2613	7	5.02	$6.59E - 04$	10	0.6333	0	—	—
FPCA			10	0.18	$6.63E - 04$			0	—	—
OptSpace			10	0.09	$2.08E - 04$			10	1.80	$5.97E - 04$
DC			10	0.37	$2.72E - 04$			9	2.19	$7.29E - 04$
SVT			0	—	—			0	—	—
FPC	5	0.3250	1	8.18	$7.34E - 04$	11	0.6930	0	—	—
FPCA			10	0.10	$2.20E - 04$			9	—	—
OptSpace			10	0.19	$2.67E - 04$			5	4.11	$7.05E - 04$
DC			10	0.47	$3.00E - 04$			6	2.83	$8.79E - 04$
SVT			0	—	—			0	—	—
FPC	6	0.3880	0	—	—	12	0.7520	0	—	—
FPCA			0	0.09	$2.20E - 04$			0	—	—
OptSpace			10	0.25	$3.25E - 04$			6	6.44	$8.47E - 04$

number of iterations is limited to 1500. When the rank of matrix is large (i.e., $k \geq 5$), the average time of our algorithm is less than that of OptSpace. In some cases, the relative errors of DC are minimal. When the dimension of matrix becomes larger, DC method has a better performance of NS and similar relative errors with that of OptSpace. In terms of time, our time is longer than that of OptSpace, but shorter than that of FPC and SVT. Tables 3 and 4 show the above results. In

conclusion, our method has better performance in “hard” problems; that is, $FR > 0.34$. The problems involved matrices that are not of very low rank, according to [10].

Table 5 gives the numerical results for problems with large dimensions. In these numerical experiments, the maximum number of iterations is chosen by default of each compared method. We observe that our method can recover problems successfully as other four methods, while our method is

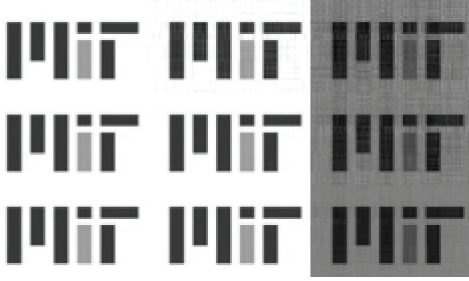


FIGURE 4: Recovered images under different numbers of linear constraints. From left to right: the original, results under our method, and results by SeDuMi. From top to bottom: $p = 800$, $p = 900$, and $p = 1000$.

TABLE 5: Numerical results for problems with different dimension.

$m(n)$	k	p	SR	FR	Alg	NS	AT	RA
100	10	5666	0.57	0.34	DC	10	0.14	$1.91E-04$
					SVT	10	2.81	$1.85E-04$
					FPC	10	5.23	$1.83E-04$
					FPCA	10	0.28	$4.94E-05$
					OptSpace	10	0.9	$1.41E-04$
200	10	15665	0.39	0.25	DC	10	0.80	$1.65E-04$
					SVT	10	2.93	$2.53E-04$
					FPC	10	4.18	$2.73E-04$
					FPCA	10	0.08	$6.88E-05$
					OptSpace	10	0.41	$1.72E-04$
300	10	27160	0.30	0.22	DC	10	2.72	$1.67E-04$
					SVT	10	2.57	$1.79E-04$
					FPC	10	9.8	$1.63E-04$
					FPCA	10	0.39	$6.71E-05$
					OptSpace	10	1.85	$1.50E-04$
400	10	41269	0.26	0.19	DC	10	5.73	$1.56E-04$
					SVT	10	2.78	$1.66E-04$
					FPC	10	13.28	$1.30E-04$
					FPCA	10	0.73	$6.65E-05$
					OptSpace	10	3.47	$1.46E-04$
500	10	49471	0.20	0.2	DC	10	13.63	$1.73E-04$
					SVT	10	2.94	$1.74E-04$
					FPC	10	25.75	$1.44E-04$
					FPCA	10	1.12	$1.07E-04$
					OptSpace	10	5.05	$1.60E-04$
1000	10	119406	0.12	0.17	DC	10	356.41	$1.70E-04$
					SVT	10	4.56	$1.66E-04$
					FPC	10	110.53	$1.29E-04$
					FPCA	10	5.16	$1.64E-04$
					OptSpace	10	14.61	$1.50E-04$

slower with the increase of dimension. However, our method is faster than FPC. Though SVT, FPCA, and OptSpace outperform our method in terms of speed, they are specifically

developed method for matrix completion problems while our method can be applied to more class of problems. To prove this, we will apply our method to max-cut problems.

3.3. *Max-Cut Problems.* In this subsection, we apply our method to max-cut problems. It can be formulated as follows:

$$\begin{aligned}
 \min \quad & \text{Tr}(LX) \\
 \text{s.t.} \quad & \text{diag}(X) = e \\
 & \text{rank}(X) = 1 \\
 & X \geq 0,
 \end{aligned} \tag{51}$$

where $X = xx^T$, $x = (x_1, \dots, x_n)^T$, $L = (1/4)(\text{diag}(We) - W)$, $e = (1, 1, \dots, 1)^T$, and W is the weight matrix. Equation (51) is one general case of problem (3) with $\Gamma \cap \Omega = \{X \mid \text{diag}(X) = eX \in S_+\}$. Hence, our method can be suitably applied to (51).

We now address the initialization and termination criteria for our method when applied to (51). We set $X^0 = ee^T$; in addition, we choose the initial penalty parameter ρ to be 0.01 and set the parameter τ to be 1.1. And we use $\|X^{i+1} - X^i\|_F \leq \epsilon$ as the termination criteria with $\epsilon = 10^{-5}$.

In our numerical experiments, we first test several small scale max-cut problems; then we perform our method on the standard “G” set problems written by Professor Ginaldi [40]. In all, we consider 10 problems ranging in size from 800 variables to 1000 variables. The results are shown in Tables 6 and 7, in which “ m ” and “ k ” represent corresponding dimension of graph and the rank of solution, respectively. The “obv” denotes objective values of optimal matrix solutions operated by our method.

In Table 6, “KV” denotes known optimal objective values and “UB” presents upper bounds for max-cut by the semidefinite relaxation. It is easy to observe that our method is effective for solving these problems and it has computed the global optimal solutions of all problems. Moreover, we obtain lower rank solution compared with semidefinite relaxation method.

Table 7 shows the numerical results for standard “G” set problems. “SUB” denotes upper bounds by spectral bundle method proposed by Helmberg and Rendl [41]. “GW-cut” denotes the approximative algorithm proposed by Goemans and Williamson [42]. “DC” denotes results generated by our method, DC^+ is the results of DC with neighborhood approximation algorithm, and the right most column, titled BK, gives best results as we have known according to [43, 44]. We use “—” to indicate that we do not obtain the rank. From Table 3 we can note that our optimal values are close to the results of latest researches, better than values given by GW-cut, and they are all less than the corresponding upper bounds. For the problems, we all get the lower rank solution compared with GW-cut; moreover, most solutions satisfy the constraints $\text{rank} = 1$. The results prove that our algorithm is effective when used to solve max-cut problems.

4. Concluding Remarks

In this paper we used closed-form of penalty function to reformulate convex constrained rank minimization problems

TABLE 6: Numerical results for famous max-cut problems.

Graph	m	KV	GW-cut		DC		Optimal solutions
			UB	k	obv	k	
C5	5	4	4.525	5	4	1	(-1; 1; -1; 1; -1)
K5	5	6	6.25	5	6	1	(-1; 1; -1; 1; -1)
AG	5	9.28	9.6040	5	9.28	1	(-1; 1; 1; -1; -1)
BG	12	88	90.3919	12	88	1	(-1; 1; 1; -1; -1; -1; -1; -1; -1; 1; 1; 1)

TABLE 7: Numerical results for G problems.

Graph	m	SUB	GW-cut		DC		DC ⁺	BK
			obv	k	obv	k		
G01	800	12078	10612	14	11281	1	11518	11624
G02	800	12084	10617	31	11261	1	11550	11620
G03	800	12077	10610	14	11275	1	11540	11622
G11	800	627	551	30	494	2	528	564
G14	800	3187	2800	13	2806	1	2980	3063
G15	800	3169	2785	70	2767	1	2969	3050
G16	800	3172	2782	14	2800	1	2974	3052
G43	1000	7027	6174	—	6321	1	6519	6660
G45	1000	7020	6168	—	6348	1	6564	6654
G52	1000	4009	3520	—	3383	1	3713	3849

as DC programming, which could be solved by stepwise linear approximative method. Under some suitable assumptions, we showed that accumulation point of the sequence generated by the stepwise linear approximative method was a stable point. The numerical results on affine rank minimization and max-cut problems demonstrate that our method is generally comparable or superior to some existing methods in terms of solution quality, and it can be applied to a class of problems. However, the speed of our method needs to be improved.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors need to acknowledge the fund projects. Wanping Yang and Jinkai Zhao are supported by Key Projects of Philosophy and Social Sciences Research (Ministry of Education, China, 15JZD012), Soft Science Fund of Shaanxi Province of China (2015KRM114), and the Fundamental Research Funds for the Central Universities. Fengmin Xu is supported by the Fundamental Research Funds for the Central Universities and China NSFC Projects (no. 11101325).

References

- [1] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific Journal of Optimization*, vol. 6, no. 3, pp. 615–640, 2010.
- [2] Y. Amit, M. Fink, N. Srebro, and S. Ullman, "Uncovering shared structures in multiclass classification," in *Proceedings of the 24th International Conference on Machine Learning (ICML '07)*, pp. 17–24, ACM, Corvallis, Ore, USA, June 2007.
- [3] L. El Ghaoui and P. Gahinet, "Rank minimization under LMI constraints: a framework for output feedback problems," in *Proceedings of the European Control Conference (ECC '93)*, Gahinet P. Rank minimization under LMI constraints, Groningen, The Netherlands, June–July 1993.
- [4] M. Fazel, H. Hindi, and S. P. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proceedings of the IEEE American Control Conference (ACC '01)*, vol. 6, pp. 4734–4739, Arlington, Va, USA, June 2001.
- [5] M. W. Trosset, "Distance matrix completion by numerical optimization," *Computational Optimization and Applications*, vol. 17, no. 1, pp. 11–22, 2000.
- [6] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.
- [7] L. Wu, "Fast at-the-money calibration of the labor market model through lagrange multipliers," *Journal of Computational Finance*, vol. 6, no. 2, pp. 39–77, 2003.
- [8] Z. Zhang and L. Wu, "Optimal low-rank approximation to a correlation matrix," *Linear Algebra and its Applications*, vol. 364, pp. 161–187, 2003.
- [9] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [10] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, no. 1-2, pp. 321–353, 2011.
- [11] K. Barman and O. Dabeer, "Analysis of a collaborative filter based on popularity amongst neighbors," *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7110–7134, 2012.
- [12] S. T. Aditya, O. Dabeer, and B. K. Dey, "A channel coding perspective of collaborative filtering," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2327–2341, 2011.
- [13] N. E. Zubov, E. A. Mikrin, M. S. Misrikhanov, V. N. Ryabchenko, S. N. Timakov, and E. A. Cheremnykh, "Identification of the position of an equilibrium attitude of the international space station as a problem of stable matrix completion," *Journal of Computer and Systems Sciences International*, vol. 51, no. 2, pp. 291–305, 2012.
- [14] T. M. Guerra and L. Vermeiren, "LMI-based relaxed non-quadratic stabilization conditions for nonlinear systems in the Takagi-Sugeno's form," *Automatica*, vol. 40, no. 5, pp. 823–829, 2004.
- [15] J. J. Meng, W. Yin, H. Li, E. Hossain, and Z. Han, "Collaborative spectrum sensing from sparse observations in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 327–337, 2011.

- [16] Z. Q. Dong, S. Anand, and R. Chandramouli, "Estimation of missing RTTs in computer networks: matrix completion vs compressed sensing," *Computer Networks*, vol. 55, no. 15, pp. 3364–3375, 2011.
- [17] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
- [18] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1791–1798, San Francisco, Calif, USA, June 2010.
- [19] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.
- [20] G. Tauböck, F. Hlawatsch, D. Eiwien, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, 2010.
- [21] Y. Tachwali, W. J. Barnes, F. Basma, and H. H. Refai, "The feasibility of a fast fourier sampling technique for wireless microphone detection in IEEE 802.22 air interface," in *Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM '02)*, pp. 1–5, IEEE, San Diego, Calif, USA, March 2010.
- [22] J. Meng, W. Yin, H. Li, E. Houssain, and Z. Han, "Collaborative spectrum sensing from sparse observations using matrix completion for cognitive radio networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '10)*, pp. 3114–3117, Dallas, Tex, USA, March 2010.
- [23] S. Pudlewski and T. Melodia, "On the performance of compressive video streaming for wireless multimedia sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '10)*, pp. 1–5, IEEE, Cape Town, South Africa, May 2010.
- [24] A. H. Mohseni, M. Babaie-Zadeh, and C. Jutten, "Inflating compressed samples: a joint source-channel coding approach for noise-resistant compressed sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 2957–2960, IEEE, Taipei, Taiwan, April 2009.
- [25] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," *SIAM Journal on Optimization*, vol. 10, no. 2, pp. 443–461, 2000.
- [26] Z. Lu, Y. Zhang, and X. Li, "Penalty decomposition methods for rank minimization," *Optimization Methods and Software*, vol. 30, no. 3, pp. 531–558, 2015.
- [27] P. Jain, R. Meka, and I. S. Dhillon, "Guaranteed rank minimization via singular value projection," in *Advances in Neural Information Processing Systems*, pp. 937–945, 2010.
- [28] J. P. Haldar and D. Hernando, "Rank-constrained solutions to linear matrix equations using powerfactorization," *IEEE Signal Processing Letters*, vol. 16, no. 7, pp. 584–587, 2009.
- [29] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2980–2998, 2010.
- [30] Y. Nesterov, A. Nemirovskii, and Y. Ye, *Interior-Point Polynomial Algorithms in Convex Programming*, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 1994.
- [31] M. Weimin, *Matrix completion models with fixed basis coefficients and rank regularized problems with hard constraints [Ph.D. thesis]*, National University of Singapore, 2013.
- [32] M. Fazel, *Matrix rank minimization with applications [Ph.D. thesis]*, Stanford University, 2002.
- [33] K. M. Grigoriadis and E. B. Beran, "Alternating projection algorithms for linear matrix inequalities problems with rank constraints," in *Advances in Linear Matrix Inequality Methods in Control*, pp. 512–267, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 1999.
- [34] Y. Gao and D. Sun, "A majorized penalty approach for calibrating rank constrained correlation matrix problems," 2010, <http://www.math.nus.edu.sg/~matsundf/MajorPen.pdf>.
- [35] S. Burer, R. D. C. Monteiro, and Y. Zhang, "Maximum stable set formulations and heuristics based on continuous optimization," *Mathematical Programming*, vol. 94, no. 1, pp. 137–166, 2002.
- [36] R. Bhatia, *Matrix Analysis*, Springer, Berlin, Germany, 1997.
- [37] Z. Lu, "Sequential convex programming methods for a class of structured nonlinear programming," <https://arxiv.org/abs/1210.3039>.
- [38] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [39] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.
- [40] <http://www.stanford.edu/~yyye/yyye/Gset/>.
- [41] C. Helmberg and F. Rendl, "A spectral bundle method for semidefinite programming," *SIAM Journal on Optimization*, vol. 10, no. 3, pp. 673–696, 2000.
- [42] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of ACM*, vol. 42, no. 6, pp. 1115–1145, 1995.
- [43] G. A. Kochenberger, J.-K. Hao, Z. Lü, H. Wang, and F. Glover, "Solving large scale max cut problems via tabu search," *Journal of Heuristics*, vol. 19, no. 4, pp. 565–571, 2013.
- [44] G. Palubeckis and V. Krivickiene, "Application of multistart tabu search to the max-cut problem," *Information Technology and Control*, vol. 31, no. 2, pp. 29–35, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

