An Iterative Approach to Rank Minimization Problems

Chuangchuang Sun and Ran Dai

Abstract—This paper investigates an iterative approach to solve the Rank Minimization Problems (RMPs) constrained in a convex set. The matrix rank function is discontinuous and nonconvex and the general RMP is classified as NP-hard. A continuous function is firstly introduced to approximately represent the matrix rank function with prescribed accuracy by selecting appropriate parameters. The RMPs are then converted to rank constrained optimization problems. An Iterative Rank Minimization (IRM) method is proposed to gradually approach the constrained rank. Convergence proof of the IRM method using the duality theory and Karush-Kuhn-Tucker conditions is provided. Two representative applications of RMP, matrix completion and output feedback stabilization problems, are presented to verify the feasibility and improved performance of the proposed IRM method.

Index Terms—Matrix Rank Minimization, Convex Relaxation, Semidefinite Programming, Matrix Completion, Output Feedback Stabilization Problem

I. INTRODUCTION

Applications of Rank Minimization Problems (RMPs) can be found in a variety of areas, such as matrix completion [1]–[3], control system analysis and design [4]–[8], machine learning [9], [10] and quadratically constrained quadratic program [11], [12], just to name a few. Particularly, when a design matrix is restricted to be diagonal, the RMP is reduced to a cardinality minimization problem which is general NP-hard as well. Due to its wide applications, RMP has attracted extensive attention.

However, due to the discontinuous and nonconvex nature of the matrix rank function, most of the existing methods focus on solving relaxed or simplified RMPs. Examples include the alternating projections and its variations [13], [14], linearization [15], [16], augmented lagrangian method [17], and log-det heuristic method [18]. In general, convergence of these methods depends on the initial guess and optimality is not guaranteed.

Recently, nuclear norm heuristic method, which minimizes the nuclear norm of a matrix instead of the rank over a convex set, has been introduced in the work of [19]. This heuristic method is computationally favorable, as nuclear norm is considered to be the best convex hull of the rank function for a set of matrices with spectral norm no greater than one [19]. In addition, the relaxed formulation with nuclear norm objective does not depend on the initial guess and global optimality is guaranteed for the convex relaxed problem. When the unknown matrix is constrained to be positive semidefinite, the relaxation of RMP with trace objective function is equivalent to the formulation with nuclear norm objective based on the fact that the trace of a positive semidefinite matrix equals to its nuclear norm [8]. For cases when the unknown matrix is not positive semidefinite, work in [19] introduces a semidefinite embedding lemma to extend the trace heuristic method to general cases. The nuclear norm heuristic method reduces the rank of an unknown matrix to a lower one, even to minimum rank in special cases. For example, it can guarantee the minimum rank solution when the constraints are equality affine ones satisfying some strong assumptions, such as the restricted isometry property [1].

However, in general cases, the nuclear norm, as a relaxed function, cannot represent the exact rank function and performance of the heuristic method is not guaranteed. In the worst case, when the unknown matrix is positive semidefinite and has equality trace constraint, the nuclear norm heuristic method is not applicable to such type of problems. Other heuristic methods, i.e., the Iterative Reweighted Least Square (IRLS) algorithm [3] which iteratively minimizes the reweighted frobenius norm of the matrix, cannot lead to the minimum rank solution either. The uncertainty of the performances in heuristic methods stems from the fact that these methods are minimizing a relaxed function and generally there is a gap between the relaxed objective and the real one. An iterative method is proposed in [20] to gradually minimize the rank to a prescribed one. However, this iterative method focuses on finding a feasible solution to meet specified rank constraint, not on searching the minimum rank. After reviewing the literature, we come to a conclusion that a more efficient approach is required to solve RMPs.

In this paper, a continuous function is introduced to approximate the rank function with prescribed accuracy by selecting appropriate parameters in the continuous function. We then reformulate the RMPs as rank constrained optimization problems. An Iterative Rank Minimization (IRM) method, with subproblem at each step formulated as a convex optimization problem, is proposed to solve the rank constrained optimization problems. Convergence of IRM is proved via the duality theory and the Krush-Kuhn-Tucker conditions. To verify the effectiveness and improved performance of IRM, two representative applications of RMPs, the matrix completion and output feedback stabilization problems, are presented with comparative results obtained from nuclear norm heuristic method, log-det heuristic method, IRLS algorithm, and Trace Penalty Method (TPM) [21].

The rest of the paper is organized as follows. In §II, the problem formulation of RMP and its conversion to rank constrained optimization problems are described. The IRM approach and its convergence proof is addressed in §III. Two application examples of RMPs and their comparative results...
are presented in §IV. We conclude the paper with a few remarks in §V.

II. PROBLEM FORMULATION

A. Rank Minimization Problem

A RMP constrained in a convex set is formulated as

\[
J = \min_X \text{rank}(X) \\
\text{s.t.} \quad X \in C, \quad (2.1)
\]

where \( X \in \mathbb{R}^{m \times n} \) is an unknown matrix and \( C \) is a convex set. Without loss of generality, we assume \( m \leq n \) in the above formulation. The matrix rank function is discontinuous and highly nonlinear. Thus, existing approaches usually build a relaxed optimization by replacing the rank function with an approximate one to achieve computational efficiency while sacrificing optimality. In the following, we introduce an alternative approach to reformulate RMPs as rank constrained optimization problems.

B. Conversion to Rank Constrained Optimization Problem

Based on the fact that the trace of a projection matrix is the dimension of the target space, the mathematical expression of this statement is in the form of

\[
P(X) = X(X^T X)^{-1}X^T \\
\text{trace}(P(X)) = \text{rank}(X), \quad (2.2)
\]

where \( P(X) \in \mathbb{R}^{m \times m} \) is the projection matrix and its trace is equivalent to the rank of \( X \) if \( X^T X \) is nonsingular. To consider singular cases, a small regularization parameter, \( \epsilon \), is introduced to reformulate \( P(X) \) as \( P_\epsilon(X) = X(X^T X + \epsilon I_n)^{-1}X^T \). It has been verified that trace \( (P_\epsilon(X)) \) can approximately represent rank \( (X) \) at any prescribed accuracy as long as \( \epsilon \) is properly given \cite{21}. Since trace \( (P_\epsilon(X)) \) is continuous and differentiable with respect to \( X \), the rank function can be replaced by trace \( (P_\epsilon(X)) \) and RMP in (2.1) is rewritten as

\[
J = \min_{X,Y} \text{trace}(Y) \\
\text{s.t.} \quad X \in C \\
Y \succeq X(X^T X + \epsilon I_n)^{-1}X^T, \quad (2.3)
\]

where \( Y \in \mathbb{S}^m \) is a slack symmetric matrix variable here. The new formulation in (2.3) is equivalent to (2.1) based on the fact that if \((Y^*, X^*)\) is an optimal solution pair to (2.3), its matrix equality constraint will be active such that \( Y^* = X^*((X^*)^T X^* + \epsilon I_n)^{-1}(X^*)^T \). When \( Y \) is the upper bound of \( X(X^T X + \epsilon I_n)^{-1}X^T \), minimizing trace \( (Y) \) is equivalent to minimizing \( X(X^T X + \epsilon I_n)^{-1}X^T \). Hence, \( X^* \) is an optimum to (2.1).

In addition, by using Schur complement to convert the nonlinear matrix inequality in (2.3) to a linear matrix equality and introducing a new matrix \( Z \in \mathbb{S}^n \), (2.3) can be equivalently transformed to the following form,

\[
J = \min_{X,Y} \text{trace}(Y) \\
\text{s.t.} \quad X \in C \\
\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\
Z = X^T X. \quad (2.4)
\]

The new formulation does not require handling matrix inverse operation and eliminates the regularization parameter, \( \epsilon \), in the semidefinite constraint of (2.3). However, the quadratic equality constraint, \( Z = X^T X \), is nonconvex. Work in [21] relaxes the quadratic equality constraint with semidefinite constraint, denoted as \( Z \succeq X^T X \), and adds a penalty in the objective function. The semidefinite relaxation method finds the minimum rank solution for a subset of problems with additional conditions, i.e., when the optimal solution with the minimum rank also has the least Frobenius norm. Therefore, an alternative approach for general RMPs is required. The first step is to transform the RMPs into rank constrained optimization problems.

**Proposition 2.1:** \( Z = X^T X \) is equivalent to \( \text{rank}\left( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \right) \leq m \) and \( \text{rank}\left( \begin{bmatrix} I_m & X^T \\ Z & X \end{bmatrix} \right) \geq 0 \), where \( Z \in \mathbb{S}^m \) and \( X \in \mathbb{R}^{m \times n} \). As noted before, \( m \leq n \).

**Proof:**

1. If \( Z = X^T X \), we get

\[
\begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \succeq 0, \quad \text{rank}\left( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \right) \leq m.
\]

2. If \( \text{rank}\left( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \right) \leq m \), and \( \text{rank}\left( \begin{bmatrix} I_m & X^T \\ Z & X \end{bmatrix} \right) \geq 0 \), we get

\[
\begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} = LL^T, \quad L \in \mathbb{R}^{(m+n) \times m} \quad \text{and here we write} \\
L = [A,B]^T \quad \text{with} \quad A \in \mathbb{R}^{m \times m} \quad \text{and} \quad B \in \mathbb{R}^{m \times n}.
\]

Since the block of the first \( m \) rows and \( m \) columns of \( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \) is an identity matrix of dimension \( m \), it indicates that each column of \( A^T \) must be orthogonal to each other, namely, \( A^T A = I_m \). It is obvious that \( X = A^T B \), \( X^T = B^T A \) and \( Z = B^T B \). Thus \( X^T X = (B^T A)(A^T B) = B^T B = Z. \)

From the above proposition, \( Z = X^T X \) could be transformed equivalently to \( \text{rank}\left( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \right) \leq m \). Combining the formulation in (2.4), RMPs are transformed as rank constrained optimization problems in the form of

\[
J = \min_{X,Y} \text{trace}(Y) \\
\text{s.t.} \quad X \in C \\
\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0 \\
\text{rank}\left( \begin{bmatrix} I_m & X \\ X^T & Z \end{bmatrix} \right) \leq m. \quad (2.5)
\]
III. An Iterative Rank Minimization Approach

A. IRM Approach to Rank Constrained Problems

To make it general, we consider the following formulation of rank constrained optimization problem,

\[
J = \min_X f(X)
\]
\[\text{s.t. } X \in C \quad (3.6)\]

\[\text{rank}(X) \leq r,
\]

where \(X \in \mathbb{R}^{m \times n}\) is an unknown matrix, \(f(X)\) is a convex function and \(C\) is a convex set. Satisfying the rank constraint for an unknown matrix is computationally complicated. Existing methodologies for rank constrained problems are mainly focused on matrix factorization and/or linearization [13], [15], which have slow convergence rate and are sensitive to the initial guess.

In fact, the number of nonzero eigenvalues of a matrix is identical to its rank. For an unknown square matrix \(U \in \mathbb{R}^{n \times n}\), it is not feasible to examine its eigenvalues before it is determined. We focus on the fact that when a matrix rank is \(r\), it has \(r\) nonzero eigenvalues. Therefore, instead of making constraint on the rank, we focus on constraining the eigenvalues of \(U\) such that the rank constraint for an unknown matrix is computationally efficient. The eigenvalue constraints on matrices have been used for graph design [22] and are applied here for rank constrained problems. Before addressing the detailed approach for rank constrained problems, we first provide necessary observations that will be used subsequently in the approach.

Proposition 3.1: The \((r + 1)\)th largest eigenvalue \(\lambda_{n-r}\) of matrix \(U \in \mathbb{R}^{n \times n}\) is less equal than \(e\) if and only if \(eI_{n-r} - V^T U V \succeq 0\), where \(I_{n-r}\) is the identity matrix with a dimension of \(n - r\), \(V \in \mathbb{R}^{n \times (n-r)}\) are the eigenvectors corresponding to the \(n - r\) smallest eigenvalues of \(U\).

Proof: Assume the eigenvalues of \(U\) is sorted in descending orders in the form of \([\lambda_n, \lambda_{n-1}, \ldots, \lambda_1]\). Since the Rayleigh quotient of an eigenvector is its associated eigenvalue, then

\[
V^T U V = \begin{bmatrix}
\lambda_{n-r} & 0 & \ldots & 0 \\
0 & \lambda_{n-r-1} & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda_1
\end{bmatrix},
\]

Hence,

\[
eI_{n-r} - V^T U V = \begin{bmatrix}
e - \lambda_{n-r} & 0 & \ldots & 0 \\
0 & e - \lambda_{n-r-1} & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & e - \lambda_1
\end{bmatrix}.
\]

Therefore \(e \geq \lambda_{n-r}\) if and only if \(eI_{n-r} - V^T U V \succeq 0\).

Corollary 3.2: When \(e = 0\) and \(U\) is a positive semidefinite matrix, the \(\text{rank}(U) \leq r\) if and only if \(eI_{n-r} - V^T U V \succeq 0\) where \(V \in \mathbb{R}^{n \times (n-r)}\) are the eigenvectors corresponding to the \(n - r\) smallest eigenvalues of \(U\).

However, the prerequisite of corollary (3.2) is that \(U\) is a positive semidefinite matrix. For a general rectangular matrix, \(X \in \mathbb{R}^{m \times n}\), in (3.6), a semidefinite embedding lemma [7], [19] is introduced to extend corollary (3.2) to general cases.

Lemma 3.3: (Lemma 1. in [7]) Let \(X \in \mathbb{R}^{m \times n}\) be a given matrix. Then \(\text{rank}(X) \leq r\) if and only if there exists matrices \(Y = Y^T \in \mathbb{R}^{r \times m}\) and \(Z = Z^T \in \mathbb{R}^{n \times r}\) such that \(\text{rank}(Y) + \text{rank}(Z) \leq 2r\) and \(Y^T X / X^T Z \succeq 0\).

Given Lemma (3.3), corollary (3.2) can be extended to general cases below.

Corollary 3.4: When \(e = 0\) and \(X \in \mathbb{R}^{m \times n}\), the rank of \(X\) is no great than \(r\) if and only if there exist matrices \(Y = Y^T \in \mathbb{R}^{r \times m}\) and \(Z = Z^T \in \mathbb{R}^{n \times r}\) such that

\[
Y^T X / X^T Z \succeq 0, eI_{m+n-2r} - V^T \begin{bmatrix} Y & 0_{m \times n} \\
0_{n \times m} & Z \end{bmatrix} V \succeq 0
\]

where \(V \in \mathbb{R}^{(m+n) \times (m+n-2r)}\) are the eigenvectors corresponding to the \(m + n - 2r\) smallest eigenvalues of \(Y^T 0_{m \times n} Z\).

From the above discussion, the rank constraint \(\text{rank}(X) \leq r\) in (3.6) can be substituted by two semidefinite constraints,

\[
\begin{bmatrix} Y & 0_{m \times n} \\
0_{n \times m} & Z \end{bmatrix} \succeq 0 \quad \text{and} \quad eI_{m+n-2r} - V^T \begin{bmatrix} Y & 0_{m \times n} \\
0_{n \times m} & Z \end{bmatrix} V \succeq 0, \quad (3.7)
\]

where \(e = 0\) and \(V \in \mathbb{R}^{(m+n) \times (m+n-2r)}\) are the eigenvectors corresponding to the \(m + n - 2r\) smallest eigenvalues of \(Y^T 0_{m \times n} Z\).

However, before we solve \(Y^T 0_{m \times n} Z\), we cannot obtain the exact \(V\) matrix, therefore an iterative method is proposed to solve (3.6) by gradually satisfying (3.7). At each step \(k\), we will solve the following semidefinite programming subproblem formulated as

\[
J = \min_{X_k, Y_k, Z_k, e_k} f(X_k) + w_k e_k
\]
\[\text{s.t. } X_k \in C \quad (3.8)\]
\[e_k I_{m+n-2r} - V_k^T \begin{bmatrix} Y_k & 0_{m \times n} \\
0_{n \times m} & Z_k \end{bmatrix} V_k \succeq 0,
\]

where \(w > 1\) is a weighting factor, \(V_{k-1} \in \mathbb{R}^{(m+n) \times (m+n-2r)}\) are the eigenvectors corresponding to the \(m + n - 2r\) smallest eigenvalues of \(Y_{k-1} 0_{m \times n} Z_{k-1}\) solved at previous iteration \(k - 1\). At each step, we are trying to optimize the original objective function and at the same time minimize parameter \(e\) such that when \(e = 0\), the rank constraint on \(X\) is satisfied. The above approach is repeated until \(e \leq \epsilon\), where \(\epsilon\) is a small threshold for the stopping criteria.
**B. IRM Approach to RMPs**

Combining the algorithm to rank constrained optimization problem in the above subsection, problem in (2.5) can be solved iteratively and the subproblem at iteration \( k \) is formulated as

\[
J = \min_{X_k, Y_k, Z_k, e_k} \text{trace}(Y_k) + w^k e_k
\]

s.t.
\[
X_k \in \mathcal{C},
\begin{bmatrix}
Y_k & X_k \\
X_k^T & Z_k
\end{bmatrix} \succeq 0
\]

\[
e_k I_n - V_{k-1}^T \begin{bmatrix}
I_m & X_{k-1}^T \\
X_{k-1} & Z_{k-1}
\end{bmatrix} V_{k-1} \succeq 0,
\]

where \( V_{k-1} \in \mathbb{R}^{(m+n) \times n} \) are the eigenvectors corresponding to the \( n \) smallest eigenvalues of \( \begin{bmatrix}
I_m & X_{k-1}^T \\
X_{k-1} & Z_{k-1}
\end{bmatrix} \) solved at previous iteration \( k-1 \).

In addition, an initial guess of \( V_0 \) is required at the first iteration \( k = 1 \). Since the trace heuristic method is often considered as a good candidate for initial guess due to its simplicity in implementation and thus is adopted here to obtain initial input of \( V_0 \). The trace heuristic method for RMPs formulated in (2.1) is listed below by applying the aforementioned semidefinite embedding lemma for general unknown matrix \( X \in \mathbb{R}^{m \times n} \) [5],

\[
J = \min_{X, Y, Z} \text{trace}(Y) + \text{trace}(Z)
\]

s.t.
\[
X \in \mathcal{C},
\begin{bmatrix}
Y & X \\
X^T & Z
\end{bmatrix} \succeq 0.
\]

The IRM approach is summarized below.

---

**Algorithm: Iterative Rank Minimization**

**Input:** Problem information \( \mathcal{C}, w, \epsilon \)

**Output:** \( X^* \) with minimum rank

**begin**

1) **initialize** Set \( k = 0 \), solve the trace heuristic problem in (3.10) to obtain \( V_0 \) from \( \begin{bmatrix}
I_m & X_k \\
X_k^T & Z_k
\end{bmatrix} \)

set \( k = k + 1 \)

2) while \( r_k > \epsilon \)

3) Solve subproblem (3.9) and obtain \( X_k, Y_k, Z_k, e_k \)

4) Update \( V_k \) from \( \begin{bmatrix}
I_m & X_k \\
X_k^T & Z_k
\end{bmatrix} \)

5) \( k = k + 1 \)

6) **end while**

7) Find \( X^* \)

**end**

---

**C. Convergence Analysis of IRM**

In the following, we provide the convergence analysis of the proposed IRM method. The convex set \( \mathcal{C} \) is defined as

\[
\mathcal{C} := \{ X \mid A(X) + b \leq 0, \ A(X) + B \preceq 0 \},
\]

where \( A \) and \( \tilde{A} \) are affine functions for linear vector and linear matrix inequalities, respectively, \( b \in \mathbb{R}^p \) and \( B \in \mathbb{R}^{s \times s} \) are given vector and matrix sets. Integrating the definition of \( \mathcal{C} \), the subproblems in (3.9) can be generalized as

\[
J = \min_{X_k, Y_k, Z_k, e_k} \text{trace}(Y_k) + w^k e_k
\]

s.t.
\[
\begin{aligned}
& A(X_k) + b \leq 0 \\
& \tilde{A}(X_k, Y_k, Z_k) + B \preceq 0 \\
& V_{k-1}^T W_k V_{k-1} - e_k I_n \preceq 0,
\end{aligned}
\]

where \( W_k = \begin{bmatrix}
I_m & X_k \\
X_k^T & Z_k
\end{bmatrix} \in \mathbb{S}^{m+n} \) and \( \tilde{A} \) represents the affine functions in convex set \( \mathcal{C} \) and transformations of \( X_k, Y_k, \) and \( Z_k \) defined in (3.9).

**Proposition 3.5:** \( \lim_{k \to +\infty} e_k = 0 \) in the IRM algorithm for a feasible optimization problem formulated in (3.12).

**Proof:** A Lagrange dual function of (3.12) is constructed as

\[
\mathcal{L} = \text{trace}(Y_k) + w^k e_k + \lambda^T (A(X_k) + b) + \text{trace}(S_1 (\tilde{A}(X_k, Y_k, Z_k) + B)) + \text{trace}(S_2 (V_{k-1}^T W_k V_{k-1} - e_k I_n))
\]

where \( \lambda_i \in \mathbb{R} \geq 0, i = 1, \ldots, p, S_1 \succeq 0, S_2 \preceq 0 \) are the Lagrange dual multipliers. The dual function is then expressed as

\[
g(\lambda, S_1, S_2) = \inf_{X_k, Y_k, Z_k, e_k} \mathcal{L}(X_k, Y_k, Z_k, e_k, \lambda, S_1, S_2).
\]

Based on the Lagrange dual function in (13.3), the first order condition is derived as

\[
\frac{\partial \mathcal{L}}{\partial X_k} = h_1 (\lambda, C_A) + h_2 (S_1, C_{\tilde{A}}) + h_3 (S_2, C_e) = 0
\]

\[
\frac{\partial \mathcal{L}}{\partial e_k} = w^k - \text{trace}(S_2) = 0,
\]

where \( C_A, C_{\tilde{A}} \) and \( C_e \) are constants derived from the first order derivatives of linear vector and linear matrix constraints in (3.12), respectively, and \( h_1, h_2, h_3 \) are affine functions. Consequently, we build the dual problem as follows,

\[
J = \max_{\lambda, S_1, S_2} \lambda^T b + \text{trace}(S_1 \tilde{B})
\]

s.t.
\[
\begin{aligned}
& h_1 (\lambda, C_A) + h_2 (S_1, C_{\tilde{A}}) + h_3 (S_2, C_e) = 0 \\
& w^k - \text{trace}(S_2) = 0
\end{aligned}
\]

\[
\lambda_i \geq 0, S_1 \succeq 0, S_2 \preceq 0.
\]

It is obvious that the problem described in (3.12) is convex. Moreover, it can be verified that the Slater’s conditions are satisfied as all constraints are affine. We come to the conclusion that the strong duality holds. As a result, at the optimal point \( \{X_k^*, Y_k^*, Z_k^*, e_k^*, \lambda^*, S_1^*, S_2^* \} \), the objective function value of the primal problem is equal to the corresponding one in the dual problem. Hence, at iteration \( k \) and \( k + 1 \),

\[
g_k = \text{trace}(Y_k^*) + w^k e_k^* = (\lambda_k^*)^T b + \text{trace}(S_k^* \tilde{B})
\]

\[
g_{k+1} = \text{trace}(Y_{k+1}^*) + w^{k+1} e_{k+1}^* = (\lambda_{k+1}^*)^T b + \text{trace}(S_{k+1}^* \tilde{B}).
\]

For a linear feasible optimization problem of (3.12), the dual problem in (3.15) is bounded, i.e., \( g_k \) and \( g_{k+1} \) are finite.
Since the weighting factor $w > 1$ and $X^*$ and $Y^*$ are finite matrices, it can be verified that
\[
\lim_{k \to +\infty} \left( w e^*_{k+1} - e^*_k \right) =
\lim_{k \to +\infty} \frac{(g_{k+1} - g_k) - \left( \text{trace}(Y^*_{k+1}) - \text{trace}(Y^*_k) \right)}{w^k} = 0.
\]
Hence,
\[
\lim_{k \to +\infty} e^*_{k+1} = \frac{1}{w} \lim_{k \to +\infty} e^*_k.
\]
Since $e^* = 0$, $e^*_{k+1} > 0$, $e^*_k > 0$, and $0 < \frac{1}{w} < 1$, then
\[
\lim_{k \to +\infty} \frac{|e^*_{k+1} - e^*_k|}{|e^*_k - e^*_m|} = \frac{1}{w}.
\]
The above equation indicates that $e^*_k$ converges to $e^*$ linearly and $w^k e^*_k$ is non-increasing.

**Proposition 3.6:** $W_k$ converges to a local optimal solution in the IRM approach.

**Proof:** When $\lim_{k \to \infty} e_k = 0$, the last constraint in (3.12) will become
\[
\lim_{k \to \infty} e_k I_n - V_k^T W_{k+1} V_k = \lim_{k \to \infty} -V_k^T W_{k+1} V_k \geq 0.
\]
Combining the above relationship with $W_{k+1} \geq 0$, it leads to
\[
\lim_{k \to \infty} V_k^T W_{k+1} V_k = 0. \tag{3.17}
\]
As the columns of $V_k \in \mathbb{R}^{(m+n) \times n}$ are orthogonal to each other, the rank of $W_{k+1}$ is no more than $m$. It indicates that when $k$ approaches infinity, the rank of $W_k$ is no more than $m$. Thus, when $k \to \infty$, $W_k \in \mathbb{S}^{m+n}$ satisfies that
\[
\lim_{k \to \infty} V_k^T W_k V_k = 0. \tag{3.18}
\]
Subtracting (3.18) from (3.17) yields
\[
\lim_{k \to \infty} V_k^T (W_{k+1} - W_k) V_k = 0. \tag{3.19}
\]
Equations in (3.17)-(3.20) indicate that when $k$ approaches infinity, the rank of matrices $V_k$, $W_{k+1}$, and $W_{k+1} - W_k$ is less than $m$. It leads to
\[
\lim_{k \to \infty} W_{k+1} = \alpha W_k. \tag{3.20}
\]
Consequently, when $k$ approaches infinity, $W_{k+1}$ and $W_k$ will have identical eigenvectors, that is, $V_k$ remains constant. As a result, the subproblem will converge with fixed $V_k$. Hence, the RMP will converge to a local minimum, which is the minimum of the convergent subproblem.

**IV. Applications and Simulation Results**

In this section, two representative applications of RMPs are solved via the proposed IRM method. They are matrix completion problems and output feedback stabilization problems. To verify the effectiveness and improved performance of IRM, results obtained from IRM are compared with those obtained from existing approaches, including the nuclear norm heuristic method [5], log-det heuristic method [7], IRLS [3] and TPM [21]. All simulation is run on a desktop computer with a 3.50 GHz processor and a 16.0 RAM.

**A. Matrix Completion Problem**

Matrix completion problems require to reconstruct a matrix given a set of entries to improve a predefined performance index. One commonly used performance index is the rank of the reconstructed matrix and lower rank is generally preferred. Methodologies in matrix completion have been applied in machine learning [9], i.e., low-rank covariance matrix recovery with partial of the entries obtained from observation. Another application of matrix completion is in the recommendation system to recommend preferred items to customers according to their shopping/rating records. In general, information in these systems is incomplete and matrix completion is required to find the missing information. Assume matrix $X$ has a set of given entries, denoted as $X_{I_k,J_k} = v_k$, the matrix completion problem is formulated as
\[
J = \min_{X} \text{rank}(X) \quad \text{s.t.} \quad X_{I_k,J_k} = v_k, \quad k = 1, \ldots, p, \tag{4.21}
\]
where $p$ is the cardinality of the given entries.

A particular example, image reconstruction, is used here to test IRM and compare it with the existing methods. Three red letters ‘ISU’ on a blue background, shown in Fig. 1, has 21 rows and 36 columns (756 elements) and two different nonzero numbers are used to represent the red and blue color. The matrix representing the real image has 5 distinct rows and thus its rank is 5. Intuitively, the cardinality of the given entries, denoted as $p$ in (4.21), will affect the recovery result. Therefore, the image reconstruction is simulated for different number of given entries, varying from $p = 100$ to $p = 700$. In the simulation, we use the MATLAB function ‘randperm’ to randomly generate the index of the given entries. IRM shows its advantages in three aspects. Firstly, Fig. 2 demonstrates the relative error of recovered matrices from different methods where IRM has the smallest relative error. The relative error is defined by the relative Frobenius norm, denoted as, $\|X - X_0\|_F / \|X_0\|_F$. Secondly, it is verified in Fig. 3 that IRM can always generate a matrix with lower or equivalent rank compared to the best solution obtained from other methods, especially in the cases when less number of entries are given. Thirdly, although trace heuristic method and log-det method can recover the matrix when $p$ is relatively large, i.e., $p = 500$, the quality of reconstructed image is poor when $p$ is relatively small. However, IRM can reconstruct higher quality images even with fewer number of given entries, as demonstrated in Fig. 4.

![Fig. 1. Test image. The corresponding matrix has 21 × 36 elements and the rank is 5.](image-url)
constraints [4].

B. Output Feedback Stabilization Problem

Consider the following linear time-invariant continuous dynamical system

\[ \dot{x} = Ax + Bu \\
\]  
\[ y = Cx, \]  
(4.22)

where \( x \in \mathbb{R}^n \), \( u \in \mathbb{R}^m \), \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \) and \( C \in \mathbb{R}^{p \times n} \). For system (4.22) and a given number \( k \leq n \), the stabilizing control law is defined as

\[ \dot{z} = A_K z + B_K u \]
\[ u = C_K z + D_K y, \]  
(4.23)

where \( z \in \mathbb{R}^k \), \( A_K \in \mathbb{R}^{k \times k} \), \( B_K \in \mathbb{R}^{k \times m} \), \( C_K \in \mathbb{R}^{m \times k} \), and \( D_K \in \mathbb{R}^{m \times p} \). Existence of the \( k \)th order control law can be determined by solving a RMP with linear matrix constraints [4].

Lemma 4.1: (Corollary 1. in [4]) There exists a stabilizing output feedback law of order \( k \) if and only if the global minimum objective of the program

\[ J = \min_{W_1, W_2, \gamma} \ \text{rank}(\begin{bmatrix} W_1 & I_n \\
I_n & W_2 \end{bmatrix}) \]
\[ s.t. \quad AW_1 + W_2 A^T \preceq \gamma BB^T \]
\[ A^T W_2 + W_2 A \preceq \gamma C^T C \]  
(4.25)

where constraints are linear functions with respect to \( W_1, W_2 \) and \( \gamma \). Using the above converted formulation, the existence of the \( k \)th order stabilizing output feedback law can be determined via the proposed IRM approach. In the simulation, we assume \( n = 10, m = 4, p = 5 \) and randomly generate the matrices \( A, B \) and \( C \) for 100 simulation cases. Similar to matrix completion example, we compare the results obtained from five different methods. Among the 100 cases, IRM obtains lower rank solution than the other methods for 79 cases, IRM finds the same lowest rank solution for 18 cases, and IRM cannot find the best solution for 3 cases. The results indicate that IRM has benchmark performance in solving most of the RMPs. Figure 5 provides comparison results of 50 cases. In addition, for one of the 50 cases, Table I shows the comparison of the computation time and the number of iterations for the five algorithms. The performance indices from the relative algorithms demonstrate the trade-
off between computation time and optimal objective value. Furthermore, figure (6) shows the convergence history of the IRM, which verifies that the algorithm reaches convergence within a few iterations under the specified stopping threshold \( \epsilon = 5e - 5 \).

![Fig. 5. Results Comparison for the output feedback stabilization problem](image)

**TABLE I**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Total Time(s)</th>
<th>Iteration Number</th>
<th>Time per Iteration(s)</th>
<th>Optimal Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trace</td>
<td>0.429267</td>
<td>1</td>
<td>0.429267</td>
<td>14</td>
</tr>
<tr>
<td>log-det</td>
<td>3.194775</td>
<td>10</td>
<td>0.319477</td>
<td>13</td>
</tr>
<tr>
<td>TPM</td>
<td>1.788778</td>
<td>1</td>
<td>1.788778</td>
<td>18</td>
</tr>
<tr>
<td>IRLS</td>
<td>7.163360</td>
<td>20</td>
<td>0.358168</td>
<td>17</td>
</tr>
<tr>
<td>IRM</td>
<td>57.861792</td>
<td>16</td>
<td>3.616362</td>
<td>12</td>
</tr>
</tbody>
</table>

![Fig. 6. IRM convergence history of one case for the output feedback stabilization problem](image)

**V. CONCLUSIONS**

In this paper, an Iterative Rank Minimization (IRM) method is proposed to solve rank minimization problems (RMPs). The convergence of the IRM to a local optimum is proved through the duality theory as well as the Karush-Kuhn-Tucker conditions. We apply IRM to two representative applications, matrix completion problems and output feedback stabilization control problems. Comparative results are demonstrated to verify the effectiveness and improved performance of IRM. Future work will focus on seeking the global optimal solution of RMPs.

**REFERENCES**