On Finding a Generalized Lowest Rank Solution to a Linear Semi-definite Feasibility Problem

Chee-Khian Sim[1]

Department of Decision Sciences, NUS Business School

National University of Singapore

Mochtar Riady Building, BIZ 1 #8-69

15 Kent Ridge Drive

Singapore 119245

March 6, 2013

[1]Email address: bizsimc@nus.edu.sg

Abstract

In this note, we generalize the affine rank minimization problem and the vector cardinality minimization problem and show that the resulting generalized problem can be solved by solving a sequence of continuous concave minimization problems. In the case of the vector cardinality minimization problem, we show that it can be solved by solving the continuous concave minimization problem.

Keywords. Affine rank minimization problem; Vector cardinality minimization problem; Concave minimization problem.

1 Introduction

The affine rank minimization problem is to find a matrix of the lowest rank that satisfies a given system of linear equality constraints. Such problem arises when solving problems in diverse fields including system identification and control, collaborative filtering and Euclidean embedding. It is well-known that solving the affine rank minimization problem is NP-hard. In [10], it is shown that if a certain restricted isometry property holds for the linear constraints, the affine rank minimization problem can be solved by solving a convex optimization problem, namely, the minimization of the nuclear norm over the given affine space. This work has its basis on the work [2], which considers the vector cardinality minimization problem. Previous attempts to solve the affine rank minimization problem include [3, 4] in which a heuristic is used to solve the problem.

In this note, we generalize the above problems and consider the problem of finding a generalized lowest rank solution to a linear semi-definite feasibility problem (LSDFP). We observe that the affine rank minimization problem and the vector cardinality minimization problem are special cases of the problem to find a generalized lowest rank solution to an LSDFP. We define this problem in Section 2. The main result of this paper is also presented in the section. We show that by solving a sequence of continuous concave minimization problems, we can find a generalized lowest rank solution to an LSDFP. This result also explains the good numerical results using the log-det heuristic in [3]. Similar result as this paper has been obtained in [13] for a different sequence of minimization problems. In Section 3, we show that the vector cardinality minimization problem can be solved by solving the continuous concave minimization problem. Although, similar results have been shown in [5, 11], in this note, we provide a different and completely new proof of these results. We conclude the note with Section 4.

1.1 Notations and Definitions

The space of symmetric $n \times n$ matrices is denoted by S^n. The cone of positive semi-definite (resp., positive definite) symmetric matrices is denoted by S^n_+ (resp., S^n_{++}).

Given a symmetric matrix $X \in S^n$, denote its real eigenvalues by $\lambda_i(X), i = 1, \ldots, n$, with $\lambda_1(X) \leq \ldots \leq \lambda_n(X)$.

Also, given $X \in S^n$, we denote its component at the intersection of the i^{th} row and the j^{th} column by X_{ij}. In case X is partitioned into blocks of submatrices, then X_{ij} refers to the submatrix in the corresponding (i, j) position.

For $X \in S^n$, $\text{diag}(X)$ stands for a vector in $\Re^n$ whose entries are the corresponding main diagonal elements of X, while given $x \in \Re^n$, $\text{Diag}(x)$ is a matrix in S^n with main diagonal entries equal to the corresponding component entries in x, with the rest of entries in the matrix equal to zero.

Given $Y \in \Re^{k_1 \times k_2}$, $\text{rank}(Y)$ refers to the dimension of the column space of Y, which is the same as the dimension of the row space of Y. In case $Y \in S^n$, then $\text{rank}(Y) =$ number of nonzero eigenvalues of Y, including multiplicities. $\|Y\|_F$ stands for the Frobenius norm of Y.

2 Main Results

The affine rank minimization problem is to find an $Y^* \in \Re^{k_1 \times k_2}$ which satisfies the following minimization problem:

$$\begin{array}{ll} \min & \text{rank}(Y) \\ \text{subject to} & \mathcal{A}(Y) = c, \\ & Y \in \Re^{k_1 \times k_2}. \end{array} \tag{1}$$

Here, $\mathcal{A} : \Re^{k_1 \times k_2} \to \Re^p$ is a linear map, and $c \in \Re^p$.

It is known [3] that the affine rank minimization problem can be written as

$$\begin{array}{ll} \min & \text{rank}(X) + \text{rank}(Z) \\ \text{subject to} & \mathcal{A}(Y) = c, \\ & \begin{pmatrix} X & Y \\ Y^T & Z \end{pmatrix} \in S_+^{k_1+k_2} \end{array} \tag{2}$$

The problem we consider in this paper, which we called the problem of finding a generalized lowest rank solution to an LSDFP, is a generalization of (2), and is to find an $X^* \in S^n$ that solves

$$\begin{array}{ll} \min & \sum_{k=1}^{N} \text{rank}(X_{kk}) \\ \text{subject to} & \text{Tr}(A_i X) = b_i, \ i = 1, \ldots, m, \\ & X \in S_+^n. \end{array} \tag{3}$$

Here, $A_i \in S^n, i = 1, \ldots, m$. Note that $X_{kk} \in S_+^{n_k}, k = 1, \ldots, N$, formed the main block diagonal submatrices of $X \in S_+^n$. Hence, $\sum_{k=1}^{N} n_k = n$.

From now onwards, whenever we consider $X \in S^n$ a feasible point of (3), it is partitioned into block submatrices with respective sizes defined from the set $\{n_k \mid k =$

$1, \ldots, N\}$. For example, the block submatrix at the (i, j) position has size $n_i \times n_j$.

Let $\mathcal{C}$ = the feasible set of (3). We assume that $\mathcal{C} \neq \emptyset$.

We have the following proposition which states the existence of an optimal solution to (3).

Proposition 2.1 *There exists an optimal solution to (3).*

Proof: Suppose the minimum in (3) is not attained. Then for any given $X \in \mathcal{C}$, there exists an $X' \in \mathcal{C}$ such that

$$\sum_{k=1}^{N} \text{rank}((X')_{kk}) < \sum_{k=1}^{N} \text{rank}(X_{kk}).$$

Since the rank of a matrix is a nonnegative integer, the above implies that there exists an $X^* \in \mathcal{C}$ with

$$\sum_{k=1}^{N} \text{rank}((X^*)_{kk}) = 0.$$

This must mean that $X^* = 0$, and is also a contradiction to minimum in (3) not attained.

QED

To make the problem meaningful, assume that the zero matrix is not a feasible solution to (3). If not, then it is the optimal solution to the problem.

Let X^* be an optimal solution to (3). We have a property satisfied by X^* as follows:

Proposition 2.2 *If X^* is an optimal solution to (3), then X^* is an extreme point of $\mathcal{C}$.*

Proof: Suppose to the contrary that X^* is not an extreme point of $\mathcal{C}$, then there exist $X_1, X_2 \in \mathcal{C}$, $X_1, X_2 \neq X^*$ with $X^* = \beta X_1 + (1 - \beta) X_2$ for some $\beta \in (0, 1)$. Therefore, we have

$$(X^*)_{kk} = \beta(X_1)_{kk} + (1-\beta)(X_2)_{kk}, \tag{4}$$

where $1 \le k \le N$.

Since $(X_1)_{kk}, (X_2)_{kk}$ are symmetric, positive semidefinite matrices, we have by (4) that $\mathrm{Ker}((X^*)_{kk}) \subseteq \mathrm{Ker}((X_1)_{kk})$ and $\mathrm{Ker}((X^*)_{kk}) \subseteq \mathrm{Ker}((X_2)_{kk})$, $k = 1, \ldots, N$. Furthermore, since X^* is an optimal solution to (3), we must have $\mathrm{Ker}((X_1)_{kk}) = \mathrm{Ker}((X_2)_{kk}) = \mathrm{Ker}((X^*)_{kk}) = d_k$, $k = 1, \ldots, N$.

Hence, there exists an orthogonal matrix $Q_k \in \Re^{n_k \times n_k}$ such that each of these matrices is of the form

$$Q_k \begin{pmatrix} B_k & 0 \\ 0 & 0 \end{pmatrix} Q_k^T,$$

where $B_k \in S_{++}^{n_k - d_k}$ different for each $(X^*)_{kk}$, $(X_1)_{kk}$ and $(X_2)_{kk}$, $k = 1, \ldots, N$.

By extending the line containing X^*, X_1, X_2 in S^n in one of the two directions, we can find an $\hat{X} \in \mathcal{C}$ such that

$$(\hat{X})_{kk} = Q_k \begin{pmatrix} \hat{B}_k & 0 \\ 0 & 0 \end{pmatrix} Q_k^T,$$

where $\hat{B}_k$ is a symmetric, positive semi-definite matrix in $S^{n_k - d_k}$, but non-invertible for some $k = 1, \ldots, N$, while the rest of $\hat{B}_k$ are symmetric, positive semi-definite matrices in $S^{n_k - d_k}$ which may or may not be invertible. This implies that

$$\sum_{k=1}^{N} \mathrm{rank}((\hat{X})_{kk}) < \sum_{k=1}^{N} \mathrm{rank}((X^*)_{kk}).$$

But this is a contradiction to X^* being an optimal solution of (3). Hence, X^* is an extreme point of $\mathcal{C}$. **QED**

Let us now consider a general framework in which to solve (3).

Let $f : (0,1] \times \Re_+ \to \Re$ be a continuous function which is strictly concave on $\Re_+$ in the second variable for fixed value of the first variable on $(0,1]$. For ease of presentation, let us denote $f(\alpha, x)$ by $f^\alpha(x)$ for $\alpha \in (0,1]$ and $x \in \Re_+$.

Let f^α satisfies the following properties:

(a) For each $0 < \alpha \leq 1$, $f^\alpha(x) \geq C_\alpha$ for all $x \in \Re_+$, where C_α is a constant, and $f^\alpha(x) \to \infty$, as $x \to \infty$.

(b) $f^\alpha(0) \to -\infty$ as $\alpha \to 0^+$, and $f^\alpha(x) \not\to -\infty$, for each $x > 0$, as $\alpha \to 0^+$.

(c) if $g : (0,1] \to \Re_+$ is any function such that $g(\alpha) \to \infty$ as $\alpha \to 0^+$, we have $f^\alpha(g(\alpha)) \to \infty$ as $\alpha \to 0^+$.

(d) if $h : (0,1] \to \Re_+$ is any function such that $h(\alpha) \to 0$ as $\alpha \to 0^+$, we have $f^\alpha(h(\alpha)) - f^\alpha(0) \not\to -\infty$ as $\alpha \to 0^+$.

Property (a) of f^α ensures existence of optimal solutions to (5) as shown in Proposition 2.3 below, while properties (b) to (d) of f^α relates (5) to (3), the problem under consideration in this note.

Hence $F_k^\alpha : \Re_+^{n_k} \to \Re$ defined by

$$F_k^\alpha(x_1, \ldots, x_{n_k}) := \sum_{j=1}^{n_k} f^\alpha(x_j)$$

is a strictly concave separable function on $\Re_+^{n_k}$, for each $k = 1, \ldots, N$, $0 < \alpha \leq 1$.

F_k^α is also a permutation-invariant function[1].

We have an associated eigenvalue function $G_k^\alpha = F_k^\alpha \circ \lambda$ defined on $S_+^{n_k}$ by

$$G_k^\alpha(Y) = (F_k^\alpha \circ \lambda)(Y) := \sum_{j=1}^{n_k} f^\alpha(\lambda_j(Y)),$$

where $Y \in S_+^{n_k}$, and $0 \leq \lambda_1(Y) \leq \ldots \leq \lambda_{n_k}(Y)$ are the eigenvalues of Y. Note that G_k^α is a continuous function on $S_+^{n_k}$ since F_k^α is continuous on $\Re_+^{n_k}$. See [7] for further properties of an eigenvalue function. We have G_k^α is also strictly concave on $S_+^{n_k}$ since F_k^α is strictly concave on $\Re_+^{n_k}$ (see for example, [8, 7]).

Consider the following continuous concave minimization problem:

$$\begin{array}{ll} \min & \sum_{k=1}^N G_k^\alpha(X_{kk}) \\ \text{subject to} & \text{Tr}(A_i X) = b_i, \ i = 1, \ldots, m, \\ & X \in S_+^n. \end{array} \tag{5}$$

We have the following proposition on (5):

Proposition 2.3 *There exists an optimal solution to (5) for each α, $0 < \alpha \leq 1$.*

Proof: The objective function of (5) is bounded from below on $\mathcal{C}$ by property (a) of f^α. Since $\mathcal{C} \neq \emptyset$, this implies that there exists a sequence $\{X_l\} \subset \mathcal{C}$ such that $\sum_{k=1}^N G_k^\alpha((X_l)_{kk})$ converges to the finite optimal value of (5), as $l \to \infty$.

We have $\{X_l\}$ is bounded. If not, then given that $X_l \in S_+^n$ for each l, we can assume without loss of generality that $\sum_{k=1}^N \sum_{j=1}^{n_k} \lambda_j((X_l)_{kk}) \to \infty$, as $l \to \infty$. This implies that $\sum_{k=1}^N G_k^\alpha((X_l)_{kk}) \to \infty$, as $l \to \infty$, by property (a) of f^α, which is impossible.

Hence, there exists a cluster point $\hat{X}$ of $\{X_l\}$, as $l \to \infty$. Since $\mathcal{C}$ is closed, $\hat{X} \in \mathcal{C}$ and is in fact an optimal solution to (5). **QED**

[1]The definition of a permutation-invariant function can be found in [7].

Let X_α^* be an optimal solution to (5).

Since $\sum_{k=1}^N G_k^\alpha(X_{kk})$ is a concave function on S_+^n, we have by Corollary 32.3.1 of [12], an optimal solution to (5) is an extreme point of $\mathcal{C}$. In fact, due to the strict concavity of $\sum_{k=1}^N G_k^\alpha(X_{kk})$, all optimal solutions of (5) are extreme points of $\mathcal{C}$.

We have the following theorem:

Theorem 2.1 *$\{X_\alpha^* \mid 0 < \alpha \leq 1\}$ is bounded, and every cluster point of $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$, as $\alpha \to 0^+$, is an optimal solution to (3).*

Proof: Let X^* be an optimal solution to (3).

We have, by definition of X_α^*, that

$$\sum_{k=1}^N G_k^\alpha((X^*)_{kk}) \geq \sum_{k=1}^N G_k^\alpha((X_\alpha^*)_{kk}). \tag{6}$$

We first show that $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$ is bounded by contradiction.

Suppose $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$ is unbounded.

Then, there exists a sequence $\{\alpha_l\} \subseteq (0,1]$ with $\alpha_l \to 0^+$, as $l \to \infty$, such that $\|X_{\alpha_l}^*\|_F \to \infty$, as $l \to \infty$.

Since $X_{\alpha_l}^* \in S_+^n$, we have $\sum_{k=1}^N \sum_{j=1}^{n_k} \lambda_j((X_{\alpha_l}^*)_{kk}) \to \infty$, as $l \to \infty$.

By property (c) of f^{α_l}, we have $\sum_{k=1}^N G_k^{\alpha_l}((X_{\alpha_l}^*)_{kk}) \to \infty$, as $l \to \infty$, which is impossible from (6).

Hence, $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$ is bounded.

We now show that every cluster point of $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$, as $\alpha \to 0^+$, is an optimal solution to (3).

Let the sequence $\{\alpha_l\} \subseteq (0,1]$ with $\alpha_l \to 0^+$, as $l \to \infty$, be such that $X_{\alpha_l}^*$ converges to $\hat{X}^* \in \mathcal{C}$, as $l \to \infty$. That is, $\hat{X}^*$ is a cluster point of $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$ as $\alpha \to 0^+$.

Let

$$\begin{aligned}
N_k^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) = 0, \lambda_j((X^*_{\alpha_l})_{kk}) \to 0 \text{ as } l \to \infty, \text{ with infinitely} \\
&\qquad \text{many } l \text{ such that } \lambda_j((X^*_{\alpha_l})_{kk}) > 0\}, \\
N_{k'}^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) = 0, \lambda_j((X^*_{\alpha_l})_{kk}) = 0 \text{ for all } l \text{ large enough}\}, \\
B_{k,1}^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) \neq 0, \lambda_j((X^*_{\alpha_l})_{kk}) \to 0 \text{ as } l \to \infty, \text{ with infinitely} \\
&\qquad \text{many } l \text{ such that } \lambda_j((X^*_{\alpha_l})_{kk}) > 0\}, \\
B_{k,1'}^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) \neq 0, \lambda_j((X^*_{\alpha_l})_{kk}) = 0, \text{ for all } l \text{ large enough}\}, \\
B_{k,2}^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) = 0, \lambda_j((X^*_{\alpha_l})_{kk}) \not\to 0 \text{ as } l \to \infty\}, \\
B_{k,3}^* &= \{j \in \{1,\ldots,n_k\} \mid \lambda_j((X^*)_{kk}) \neq 0, \lambda_j((X^*_{\alpha_l})_{kk}) \not\to 0 \text{ as } l \to \infty\}.
\end{aligned}$$

It is clear that the above sets are disjoint, and $|N_k^*| + |N_{k'}^*| + |B_{k,1'}^*| + \sum_{i=1}^3 |B_{k,i}^*| = n_k$, for $k = 1,\ldots,N$.

Note that $\operatorname{rank}((\hat{X}^*)_{kk}) = |B_{k,2}^*| + |B_{k,3}^*|$ and $\operatorname{rank}((X^*)_{kk}) = |B_{k,1}^*| + |B_{k,1'}^*| + |B_{k,3}^*|$.

We therefore have

$$\sum_{k=1}^{N} \left[|B_{k,2}^*| - |B_{k,1'}^*| - |B_{k,1}^*|\right] \geq 0, \tag{7}$$

since X^* is an optimal solution to (3).

In the following argument, we consider a subsequence of $\{\alpha_l\}$ is necessary.

Rearranging terms in (6) using the definitions of G_k^α and the above sets, the following inequality holds:

$$\sum_{k=1}^{N} \left[(|B_{k,2}^*| - |B_{k,1'}^*| - |B_{k,1}^*|) f^{\alpha_l}(0) + \sum_{j \in B_{k,1}^* \cup B_{k,1'}^* \cup B_{k,3}^*} f^{\alpha_l}(\lambda_j((X^*)_{kk})) \right]$$

$$\geq \sum_{k=1}^{N}\left[\sum_{j\in B_{k,2}^*} f^{\alpha_l}(\lambda_j((X_{\alpha_l}^*)_{kk})) + \sum_{j\in N_k^*\cup B_{k,1}^*}(f^{\alpha_l}(\lambda_j((X_{\alpha_l}^*)_{kk})) - f^{\alpha_l}(0))\right.$$

$$\left. + \sum_{j\in B_{k,3}^*} f^{\alpha_l}(\lambda_j((X_{\alpha_l}^*)_{kk}))\right]. \tag{8}$$

We claim that (7) is actually an equality.

If not, then

$$\sum_{k=1}^{N}\left[|B_{k,2}^*| - |B_{k,1'}^*| - |B_{k,1}^*|\right] \geq 1. \tag{9}$$

We have in (8), its right-hand side does not tend to $-\infty$ as $l \to \infty$, using properties (b) and (d) of f^{α_l}. However, since $\lim_{l\to\infty} f^{\alpha_l}(0) = -\infty$ and from (9), we have the left-hand side of (8) tends to $-\infty$ as $l \to \infty$. But this is a contradiction. Hence, we have

$$\sum_{k=1}^{N}\left[|B_{k,2}^*| - |B_{k,1'}^*| - |B_{k,1}^*|\right] = 0. \tag{10}$$

Thus, $\hat{X}^*$ is an optimal solution to (3). Therefore, every cluster point of $\{X_\alpha^* \mid 0 < \alpha \leq 1\}$, as $\alpha \to 0^+$, is an optimal solution to (3). **QED**

The following corollary follows immediately from the above theorem.

Corollary 2.1 *Suppose (3) has an unique solution. Then X_α^* converges to the unique solution, as $\alpha \to 0^+$.*

2.1 An Example of f^α

Consider $f^\alpha(x) = \log(\beta x + \alpha)$ for $0 < \alpha \leq 1$, where $\beta > 0$ is fixed.

Then $f^\alpha(x)$ is a continuous function of $(\alpha, x) \in (0,1] \times \Re_+$ and a strictly concave function of $x \in \Re_+$ for fixed $\alpha \in (0,1]$. f^α is bounded below on $\Re_+$ by $\log\alpha$ for each $\alpha \in (0,1]$, with $f^\alpha(x) = \log(\beta x + \alpha) \to \infty$, as $x \to \infty$. Property (b) of f^α is satisfied since $f^\alpha(0) = \log \alpha \to -\infty$ as $\alpha \to 0^+$, with $f^\alpha(x) \to \log(\beta x) > -\infty$ for each $x > 0$ as $\alpha \to 0^+$. Also, if $g : (0,1] \to \Re_+$ is any function such that $g(\alpha) \to \infty$ as $\alpha \to 0^+$, then $f^\alpha(g(\alpha)) = \log(\beta g(\alpha) + \alpha) \to \infty$ as $\alpha \to 0^+$. We can also verify easily that if $h : (0,1] \to \Re_+$ is any function such that $h(\alpha) \to 0$ as $\alpha \to 0^+$, then $\log(\beta h(\alpha) + \alpha) - \log(\alpha) \not\to -\infty$ as $\alpha \to 0^+$, since $h(\alpha) \geq 0$ for all $0 < \alpha \leq 1$.

In case $f^\alpha(x) = \log(\beta x + \alpha)$, we can write (5) in a neat way as

$$\begin{array}{ll} \min & \sum_{k=1}^N \log\det(\beta X_{kk} + \alpha I) \\ \text{subject to} & \mathrm{Tr}(A_i X) = b_i,\ i = 1, \ldots, m, \\ & X \in S^n_+. \end{array} \tag{11}$$

In [3], this is used as a heuristic to solve the affine rank minimization problem, with $\beta = 1$ and $N = 2$. In Theorem 2.1, we show that, indeed, using this heuristic is well justified, in theory, to solve the problem.

Another example of $f^\alpha(x) = -\frac{1}{x+\alpha} + x$, for $0 < \alpha \leq 1$ and $x \geq 0$.

3 A Related Problem

A closely related problem to (3) is the following problem:

$$\begin{array}{ll} \min & \|\mathrm{diag}(X)\|_0 \\ \text{subject to} & \mathrm{Tr}(A_i X) = b_i,\ i = 1, \ldots, m, \\ & X \in S^n_+. \end{array} \tag{12}$$

Here, $\|x\|_0$ is the L_0 norm of $x \in \Re^n$, and is defined as the number of nonzero components of x.

Note that (12) is a special case of (3) with $N = n$, and $X_{kk}, 1 \leq k \leq N$, are the main diagonal entries of X in (3).

We have (12) is a generalization of its linear counterpart

$$\begin{array}{ll} \min & \|x\|_0 \\ \text{subject to} & Bx = b, \\ & x \geq 0 \end{array} \tag{13}$$

to the space of symmetric matrices.

Consider the following problem, which is called the vector cardinality minimization problem, studied in [2] (see also [6]):

$$\begin{array}{ll} \min & \|x\|_0 \\ \text{subject to} & Bx = b. \end{array} \tag{14}$$

We have the following proposition relating (13) to (14):

Proposition 3.1 $\begin{pmatrix} (x^+)^* \\ (x^-)^* \end{pmatrix} \in \Re^{2n}$ *is an optimal solution to*

$$\begin{array}{ll} \min & \|x^+\|_0 + \|x^-\|_0 \\ \textit{subject to} & (B \ \ -B) \begin{pmatrix} x^+ \\ x^- \end{pmatrix} = b, \\ & x^+, x^- \geq 0 \end{array} \tag{15}$$

if and only if $x^* = (x^+)^* - (x^-)^*$ *is an optimal solution to (14). In this case,*

$$\begin{aligned} (x^+)^*_i &= \max\{x^*_i, 0\}, \\ (x^-)^*_i &= -\min\{x^*_i, 0\}, \end{aligned} \tag{16}$$

$i = 1, \ldots, n.$

Proof: Suppose $\begin{pmatrix} (x^+)^* \\ (x^-)^* \end{pmatrix}$ is an optimal solution to (15). This means that $(x^+)^*_i > 0$ implies that $(x^-)^*_i = 0$, and $(x^-)^*_i > 0$ implies that $(x^+)^*_i = 0$. Hence, (16) holds and $\|x^*\|_0 = \|(x^+)^*\|_0 + \|(x^-)^*\|_0$.

Suppose x^{**} is an optimal solution to (14) with $\|x^{**}\|_0 < \|x^*\|_0$. Define

$$\begin{aligned} (x^+)^{**}_i &= \max\{x^{**}_i, 0\} \\ (x^-)^{**}_i &= -\min\{x^{**}_i, 0\} \end{aligned}$$

for $i = 1, \ldots, n$. Then it is clear that $\|x^{**}\|_0 = \|(x^+)^{**}\|_0 + \|(x^-)^{**}\|_0$ and also $\begin{pmatrix} (x^+)^{**} \\ (x^-)^{**} \end{pmatrix}$ is feasible to (15). But this is a contradiction to $\begin{pmatrix} (x^+)^* \\ (x^-)^* \end{pmatrix} \in \Re^{2n}$ being an optimal solution to (15), since we have $\|(x^+)^{**}\|_0 + \|(x^-)^{**}\|_0 = \|x^{**}\|_0 < \|x^*\|_0 = \|(x^+)^*\|_0 + \|(x^-)^*\|_0$. Hence, $x^* = (x^+)^* - (x^-)^*$ is an optimal solution to (14).

The reverse direction can be shown to be true in a similar manner. **QED**

The above proposition shows that solving (13) is as hard as solving (14).

We have the following theorem on solving (13), which ends this section:

Theorem 3.1 *In case $A_i, 1 \leq i \leq m$, in (12) are diagonal matrices, then X^*_α, which is an optimal solution to (5), where in (5), $N = n$, and X_{kk} are the main diagonal entries of X, is also an optimal solution to (12), for all $\alpha > 0$ small enough.*

Proof: We prove the theorem by contradiction, that is, assuming that there exists a sequence $\{\alpha_l\}$ with $\alpha_l \to 0^+$, as $l \to \infty$, and such that $X^*_{\alpha_l}$ is not an optimal solution to

(12) for each l. We have $\{X^*_{\alpha_l}\}$ is bounded, by Theorem 2.1. Let X^* be a cluster point of $\{X^*_{\alpha_l}\}$, as $l \to \infty$. By Theorem 2.1 again, X^* is an optimal solution to (12).

Without loss of generality, let $X^*_{\alpha_l} \to X^*$, as $l \to \infty$.

Observe that if X is an extreme point of $\mathcal{C}$, then $\text{diag}(X) \in \Re^n$ is an extreme point of the following convex polyhedron in $\Re^n$:

$$\mathcal{C}' = \{x \in \Re^n \mid \text{Tr}(A_i \text{Diag}(x)) = b_i,\ i = 1, \ldots, m, x \geq 0\}.$$

By Proposition 2.2 applied to (12), X^* is an extreme point of $\mathcal{C}$. Thus $\text{diag}(X^*)$ is an extreme point of $\mathcal{C}'$.

Now, $X^*_{\alpha_l}$ is also an extreme point of $\mathcal{C}$ for each l, since it is an optimal solution to (5). This implies that $\text{diag}(X^*_{\alpha_l})$ is an extreme point of $\mathcal{C}'$.

Since $X^*_{\alpha_l} \to X^*$ as $l \to \infty$, we have $\text{diag}(X^*_{\alpha_l}) \to \text{diag}(X^*)$ as $l \to \infty$. But since the set of extreme points of $\mathcal{C}'$ is finite (by Corollary 19.1.1 of [12]), we must have $\text{diag}(X^*_{\alpha_l}) = \text{diag}(X^*)$ for all l large enough. Now, given $X \in \mathcal{C}$, only $\text{diag}(X)$ is needed to determine whether X solves (12), we then have $X^*_{\alpha_l}$ solves (12) for l large enough. This is a contradiction. **QED**

Similar results as Theorem 3.1 have been obtained in [5, 11]. Above, we provide a different proof of these results.

4 Conclusion

This note shows that a minimization problem with discrete objective function (3) can be solved by solving a continuous concave minimization problem (5) under certain conditions (Theorem 3.1), and approximated by the continuous concave minimization problem in general (Theorem 2.1 and Corollary 2.1). There are known methods used to solve the class of continuous concave minimization problems, see for example, [1, 9]. As a future work, we would like to explore ways to solve (5) effectively and efficiently, by studying the special structure of the problem. This is crucial if the problem is large-scale. This can be done for example by further exploring ways to solve the corresponding linear semi-definite program with special structure, which was given in [3]. This in turns allows us to solve (3) approximately, if not exactly.

References

[1] H. P. Benson, *Deterministic algorithms for constrained concave minimization: a unified critical survey,* Naval Research Logistics, Vol. 43(1996), pp. 765-795.

[2] E. J. Candes & T. Tao, *Decoding by linear programming,* IEEE Transactions on Information Theory, Vol. 51(2005), pp. 4203-4215.

[3] M. Fazel, H. Hindi & S. P. Boyd, *Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices,* Proceedings of the American Control Conference, Denver, Colorado, 2003, pp. 2156-2162.

[4] M. Fazel, H. Hindi & S. P. Boyd, *A rank minimization heuristic with application to minimium order system approximation,* Proceedings of the American Control Conference, Arlington, VA, 2001, pp. 4734 - 4739.

[5] G. M. Fung & O. L. Mangasarian, *Equivalence of minimal* l_0- *and* l_p- *norm solutions of linear equalities, inequalities and linear programs for sufficiently small p,* Journal of Optimization Theory and Applications, Vol. 151(2011), pp. 1-10.

[6] D. Ge, X. Jiang & Y. Ye, *A note on the complexity of* L_p *minimization,* Mathematical Programming, published online: 2011.

[7] A. S. Lewis, *Nonsmooth analysis of eigenvalues,* Mathematical Programming, Series A, Vol. 84(1999), pp. 1-24.

[8] A. S. Lewis, *Group invariance and convex matrix analysis,* SIAM Journal on Matrix Analysis and Applications, Vol. 17(1996), pp. 927-949.

[9] P. M. Pardalos & J. B. Rosen, *Methods for global concave minimization: a bibliographic survey,* SIAM Review, Vol. 28(1986), pp. 367-379.

[10] B. Recht, M. Fazel & P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,* SIAM Review, Vol. 52(2010), pp. 471-501.

[11] F. Rinaldi, F. Schoen & M. Sciandrone, *Concave programming for minimizing the zero-norm over polyhedral sets,* Computational Optimization and Applications, Vol. 46(2010), pp. 467-486.

[12] R. T. Rockafellar, *Convex Analysis,* Princeton Landmarks in Mathematics, Princeton University Press, 1970.

[13] Y.-B. Zhao, *An approximation theory of matrix rank minimization and its application to quadratic equations,* Linear Algebra and its Applications, Vol. 437(2012), pp. 77-93.