Operations Research Letters 46 (2018) 76–80

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Minimizing the sum of linear fractional functions over the cone of positive semidefinite matrices: Approximation and applications

Yong Xia [a,*], Longfei Wang [a], Shu Wang [b]

[a] State Key Laboratory of Software Development Environment, LMIB of the Ministry of Education, School of Mathematics and System Sciences, Beihang University, Beijing, 100191, PR China

[b] College of Science, North China Institute of Science and Technology, Hebei, 065201, PR China

ARTICLE INFO

Article history:
Received 14 March 2017
Received in revised form 13 November 2017
Accepted 13 November 2017
Available online 22 November 2017

Keywords:
Fractional programming
Semidefinite programming
Rayleigh quotient
Total least squares
FPTAS

ABSTRACT

The problem of maximizing the sum of two generalized Rayleigh quotients and the total least squares problem with nonsingular Tikhonov regularization are reformulated as a class of sum-of-linear-ratios minimizing over the cone of symmetric positive semidefinite matrices, which is shown to have a Fully Polynomial Time Approximation Scheme.

1. Introduction

Consider the following linear fractional optimization over the cone of symmetric positive semidefinite matrices:

$$\text{(FSDP)} \ \min \ g(X) := \sum_{i=1}^{p} \frac{\operatorname{tr}(A_iX)}{\operatorname{tr}(B_iX)} \tag{1}$$

$$\text{s.t.} \ X \in \Omega := \{X \mid X \succeq 0, \ \operatorname{tr}(C_jX) \le d_j, j = 1, \ldots, m\}, \tag{2}$$

where $A_i \succ 0$, $B_i \succ 0$ for $i = 1, \ldots, p$, $C_j \in \mathbb{R}^{n\times n}$, $d_j \in \mathbb{R}$ for $j = 1, \ldots, m$, the notation $(\cdot) \succ (\succeq) 0$ indicates that $(\cdot)$ is symmetric positive (semi)definite, and $\operatorname{tr}(\cdot)$ denotes the trace of $(\cdot)$.

The well-known Rayleigh–Ritz theorem on the equivalence between the minimum of Rayleigh quotient and the smallest eigenvalue of a symmetric matrix can be regarded as a special case of (FSDP) with $p = 1$:

$$\lambda_{\min}(A) = \min_{x\neq 0} \frac{x^TAx}{x^Tx} = \min_{0\neq X\succeq 0} \frac{\operatorname{tr}(AX)}{\operatorname{tr}(X)}.$$

(FSDP) includes the classical semidefinite programming (SDP) [18,19] as special cases by setting $\Omega \subseteq \{X \mid \operatorname{tr}(B_iX) = 1, i = 1, \ldots, p\}$. We notice that (SDP) can be solved in polynomial time, see, for example, [5]. Moreover, when $p = 1$, applying the Charnes–Cooper transformation [4], (FSDP) can be equivalently reformulated as an SDP. For such an example of standard fractional quadratic problem, we refer to [1]. This is not surprise as the single linear ratio is pseudo-convexity and then any KKT solution of (FSDP) is a global minimizer.

When all A_i, B_i, C_j are diagonal matrices and $\Omega \subseteq \{X \mid X_{n,n} = 1\}$, (FSDP) reduces to the following linear fractional programming problem:

$$\text{(FLP)} \ \min_{x\ge 0, \operatorname{Diag}(x;1)\in\Omega} \sum_{i=1}^{p} \frac{a_i^Tx + a_{0i}}{b_i^Tx + b_{0i}}, \tag{3}$$

where $(a_i; a_{0i}) = \operatorname{diag}(A_i)$, $(b_i; b_{0i}) = \operatorname{diag}(B_i)$, the notation $\operatorname{Diag}(x)$ denotes the diagonal matrix with x_i being the (i, i)th main diagonal element, and $\operatorname{diag}(X)$ is defined as the column vector with its components being the diagonal elements of X. Problem (FLP) has many applications, see, for example, [3,16]. When $p \ge 2$, (FLP) is observed to have local non-global minimizers [15]. In general, it is shown to be NP-hard [8]. Actually, (FLP) with $p = 2$ is already NP-hard [10]. Recently, Depetrini and Locatelli [6] propose a Fully Polynomial Time Approximation Scheme (FPTAS) for (FLP) when p is fixed.

In this paper, we first present two new applications of (FSDP). The first is to maximize the sum of two generalized Rayleigh

* Corresponding author.
E-mail addresses: yxia@buaa.edu.cn (Y. Xia), lfking@buaa.edu.cn (L. Wang), wangshulxy@ncist.edu.cn (S. Wang).

https://doi.org/10.1016/j.orl.2017.11.010

quotients

$$(\text{RQ}) \max_{x\neq 0} \frac{x^T W_1 x}{x^T B_1 x} + \frac{x^T W_2 x}{x^T B_2 x}, \tag{4}$$

where $B_1 \succ 0$ and $B_2 \succ 0$. This problem is initially proposed by [23,24] with applications in the downlink of a multi-user MIMO system [13] and the sparse Fisher discriminant analysis in pattern recognition [9,20]. Recently, Nguyen et al. [11] proposed a parametric approach to solve (RQ) via the following univariate maximization:

$$\max_{\mu\in[\underline{\mu},\overline{\mu}]} q(\mu) := \mu + \max_{x^T(W_1-\mu B_1)x\geq 0, x^T B_2 x=1} x^T W_2 x,$$

where $\underline{\mu}$ and $\overline{\mu}$ are the smallest and the largest generalized eigenvalues of the matrix pencil (W_1, B_1), respectively. For each given μ, $q(\mu)$ can be evaluated by solving an SDP [11]. However, since $g(\mu)$ is not guaranteed to be unimodal, one cannot expect any nontrivial complexity analysis on this parametric approach.

The second application is the nonsingular Tikhonov regularization of the total least squares problem [2]:

$$(\text{RLS}) \min_{x\in\mathbb{R}^n} \frac{\|Ax-b\|^2}{\|x\|^2+1} + \rho\|Lx\|^2, \tag{5}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $L \in \mathbb{R}^{n\times n}$ is nonsingular and ρ is a positive penalty parameter. We show that both (RQ) and (RLS) can be equivalently reformulated as (FSDP). To our knowledge, it is still an open problem whether (RQ) and (RLS) are NP-hard or polynomial-time solvable problems.

Then, we extend Depetrini and Locatelli's ϵ-approximation algorithm [6] (which is proposed for (FLP)) to solve (FSDP) and then prove it is an FPTAS when $p+m$ is fixed. It follows that both (RQ) and (RLS) have an FPTAS.

The remainder of this paper is as follows. Section 2 reformulates the two applications, (RQ) and (RLS), as (FSDP). Section 3 establishes the FPTAS for a class of (FSDP). Conclusions are made in Section 4.

2. FSDP reformulation of the sum-of-quadratic-ratios applications

In this section, we reformulate the above sum-of-quadratic-ratios problems (RQ) (4) and (RLS) (5) as (FSDP).

First, (RQ) can be rewritten as the minimization problem (ignoring the constant term)

$$(\text{RQ}_{\min}) \min_{x^T x=1} r(x) := \frac{x^T(\tau B_1 - W_1)x}{x^T B_1 x} + \frac{x^T(\tau B_2 - W_2)x}{x^T B_2 x}, \tag{6}$$

where τ is a positive scalar such that $\tau B_1 - W_1 \succ 0$ and $\tau B_2 - W_2 \succ 0$, the constraint $x^T x = 1$ is added to keep away from $x = 0$ by noting that all the quadratic numerators and denominators are homogeneous. By lifting $x^T(\cdot)x$ to $\text{tr}((\cdot)X)$, $(\text{RQ}_{\min})$ (6) is further equivalent to

$$\min_{X\succeq 0, \text{tr}(X)=1, \text{rank}(X)=1} \frac{\text{tr}((\tau B_1 - W_1)X)}{\text{tr}(B_1 X)} + \frac{\text{tr}((\tau B_2 - W_2)X)}{\text{tr}(B_2 X)}.$$

Removing the rank constraint yields the FSDP relaxation:

$$\begin{aligned}(\text{RQ}_{\text{FSDP}}) \min\ R(X) := &\frac{\text{tr}((\tau B_1 - W_1)X)}{\text{tr}(B_1 X)} \\ &+ \frac{\text{tr}((\tau B_2 - W_2)X)}{\text{tr}(B_2 X)} \\ \text{s.t.}\ \ &\text{tr}(X) = 1,\ X \succeq 0.\end{aligned}$$

Moreover, we can show that the FSDP relaxation is tight.

Theorem 2.1. *The FSDP relaxation* $(\text{RQ}_{\text{FSDP}})$ *is equivalent to* $(\text{RQ}_{\min})$ *in the sense that both optimal values are equal and an optimal solution of* $(\text{RQ}_{\min})$ *can be recovered in polynomial time from that of* $(\text{RQ}_{\text{FSDP}})$.

Proof. Let $v(\cdot)$ be the optimal value of problem $(\cdot)$. According to the derivation of $(\text{RQ}_{\text{FSDP}})$, we have

$$v(\text{RQ}_{\text{FSDP}}) \leq v(\text{RQ}_{\min}). \tag{7}$$

Let $\widetilde{X}$ be any feasible solution of $(\text{RQ}_{\text{FSDP}})$. Define

$$\widetilde{t}_1 = \frac{\text{tr}((\tau B_1 - W_1)\widetilde{X})}{\text{tr}(B_1\widetilde{X})}, \widetilde{t}_2 = \frac{\text{tr}((\tau B_2 - W_2)\widetilde{X})}{\text{tr}(B_2\widetilde{X})}.$$

Then, it holds that $R(\widetilde{X}) = \widetilde{t}_1 + \widetilde{t}_2$ and

$$\text{tr}((\tau B_1 - W_1 - \widetilde{t}_1 B_1)\widetilde{X}) = 0, \tag{8}$$

$$\text{tr}((\tau B_2 - W_2 - \widetilde{t}_2 B_2)\widetilde{X}) = 0. \tag{9}$$

According to the rank-one decomposition theorem due to Sturm and Zhang [17], it follows from (8) that there exists nonzero vectors $p_1, \ldots, p_r$ (obtained in polynomial time) such that

$$\widetilde{X} = \sum_{i=1}^r p_i p_i^T \tag{10}$$

and

$$p_i^T(\tau B_1 - W_1 - \widetilde{t}_1 B_1)p_i = 0,\ i = 1, \ldots, r, \tag{11}$$

where r is the rank of $\widetilde{X}$. Substituting (10) into (9) yields

$$\sum_{i=1}^r p_i^T(\tau B_2 - W_2 - \widetilde{t}_2 B_2)p_i = 0.$$

Then, there is an index $k \in \{1, \ldots, r\}$ such that

$$p_k^T(\tau B_2 - W_2 - \widetilde{t}_2 B_2)p_k \leq 0. \tag{12}$$

Since $p_k \neq 0$, we can define $\widetilde{x} = p_k/\sqrt{p_k^T p_k}$. Thus, $\widetilde{x}^T\widetilde{x} = 1$. According to (11) and (12), we have

$$\widetilde{x}^T(\tau B_1 - W_1 - \widetilde{t}_1 B_1)\widetilde{x} = 0,\ \widetilde{x}^T(\tau B_2 - W_2 - \widetilde{t}_2 B_2)\widetilde{x} \leq 0.$$

It implies that

$$\begin{aligned}r(\widetilde{x}) &= \frac{\widetilde{x}^T(\tau B_1 - W_1)\widetilde{x}}{\widetilde{x}^T B_1 \widetilde{x}} + \frac{\widetilde{x}^T(\tau B_2 - W_2)\widetilde{x}}{\widetilde{x}^T B_2 \widetilde{x}} \\ &\leq \widetilde{t}_1 + \widetilde{t}_2 \\ &= R(\widetilde{X}).\end{aligned} \tag{13}$$

Therefore, we have

$$v(\text{RQ}_{\min}) \leq v(\text{RQ}_{\text{FSDP}}).$$

Combined with (7), it holds that $v(\text{RQ}_{\min}) = v(\text{RQ}_{\text{FSDP}})$. Consequently, $v(\text{RQ}_{\text{FSDP}})$ has an optimal solution, denoted by X^*, since $v(\text{RQ}_{\min})$ is attainable. Based on X^*, applying the above procedure for generating $\widetilde{x}$ from $\widetilde{X}$, we can obtain an optimal solution of $(\text{RQ}_{\min})$ in polynomial time. □

Now, we consider (RLS) (5), which has an equivalent FSDP-reformulation as follows:

$$\begin{aligned}(\text{RLS}_{\text{FSDP}}) \min\ &\frac{\text{tr}(A^T AX) - 2b^T Ax + \|b\|^2}{\text{tr}(X)+1} \\ &+ \rho\,\text{tr}(L^T LX) \\ \text{s.t.}\ &\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0.\end{aligned} \tag{14}$$