Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2014, Article ID 749475, 7 pages
http://dx.doi.org/10.1155/2014/749475

Research Article

A Relax Inexact Accelerated Proximal Gradient Method for the Constrained Minimization Problem of Maximum Eigenvalue Functions

Wei Wang, Shanghua Li, and Jingjing Gao

School of Mathematics, Liaoning Normal University, Liaoning, Dalian 116029, China

Correspondence should be addressed to Shanghua Li; lishanghua1989@163.com

Received 26 February 2014; Accepted 21 June 2014; Published 9 July 2014

Academic Editor: Yuesheng Xu

For constrained minimization problem of maximum eigenvalue functions, since the objective function is nonsmooth, we can use the approximate inexact accelerated proximal gradient (AIAPG) method (Wang et al., 2013) to solve its smooth approximation minimization problem. When we take the function $g(X) = \delta_{\Omega}(X)$ ($\Omega := \{X \in S^n : \mathcal{F}(X) = b, X \succeq 0\}$) in the problem $\min\{\lambda_{\max}(X) + g(X) : X \in S^n\}$, where $\lambda_{\max}(X)$ is the maximum eigenvalue function, $g(X)$ is a proper lower semicontinuous convex function (possibly nonsmooth) and $\delta_{\Omega}(X)$ denotes the indicator function. But the approximate minimizer generated by AIAPG method must be contained in Ω otherwise the method will be invalid. In this paper, we will consider the case where the approximate minimizer cannot be guaranteed in Ω. Thus we will propose two different strategies, respectively, constructing the feasible solution and designing a new method named relax inexact accelerated proximal gradient (RIAPG) method. It is worth mentioning that one advantage when compared to the former is that the latter strategy can overcome the drawback. The drawback is that the required conditions are too strict. Furthermore, the RIAPG method inherits the global iteration complexity and attractive computational advantage of AIAPG method.

1. Introduction

The minimization problem of maximum eigenvalue functions in nonsmooth optimization presents a fascinating mathematical challenge. Such problems arise in many different areas of applied mathematics, especially in engineering design [1], and also play important roles in enriching blend of classical mathematical techniques and contemporary optimization theory. The constrained minimization problem of maximum eigenvalue functions can be transformed into the minimization problem of the sum of two convex functions. Various methods have been proposed to deal with such problems, such as in [2], a forward-backward splitting algorithm was used to solve the minimization problem of two proper lower semicontinuous convex functions. Besides, several fixed point algorithms based on proximity operator were introduced in [3] for ROF denoising model which is actually the minimization problem of the sum of two convex functions. More recently, the AIAPG method which is based on accelerated proximal gradient (APG) method [4] was introduced in [5] for solving the minimization problem of the sum of maximum eigenvalue function and proper lower semicontinuous convex function. If the approximate minimizer is infeasible, that is, the approximate minimizer is not strictly contained in the feasible set Ω, the AIAPG method will not be available. Hence, we design the RIAPG method which is based on AIAPG method to solve the smooth approximation problem of constrained minimization problem of maximum eigenvalue functions.

We consider the following constrained minimization problem of the maximum eigenvalue function:

$$\begin{aligned} \min \quad & \lambda_{\max}(X) \\ \text{s.t.} \quad & \mathcal{F}(X) = b \\ & X \succeq 0, \end{aligned} \tag{P}$$

where $\lambda_{\max}(X)$ is the maximum eigenvalue function, $\mathcal{F} : S^n \to R^m$ is a linear map, $b \in R^m$, and $X \succeq 0$ means X is a positive semidefinite matrix. S^n is the space of $n \times n$

real symmetric matrices. The problem (P) is equivalent to the following form:

$$\min \{\lambda_{\max}(X) + \delta_\Omega(X) : X \in S^n\}, \qquad (P_1)$$

where $\Omega = \{X \in S^n : \mathcal{F}(X) = b, X \succeq 0\}$, $\delta_\Omega(X)$ denotes the indicator function. Then we consider the smooth approximation $(h_\varepsilon \circ \lambda)(X)$ [6] to the maximum eigenvalue function $\lambda_{\max}(X)$ which is a proper, lower semicontinuous, convex function and $\nabla(h_\varepsilon \circ \lambda)(X)$ is Lipschitz continuous. This thought for dealing with the problem resembles the technique used in [7]. Hence, the approximation form of (P_1) is given by

$$\min \{(h_\varepsilon \circ \lambda)(X) : X \in \Omega\}. \qquad (P_2)$$

Problem (P_2) can be solved by AIAPG method in feasible case. In infeasible case, we will propose two strategies. On the one hand, we use infeasible approximate minimizer to construct feasible solution which satisfies the conditions required by AIAPG method. On the other hand, we enlarge the feasible set Ω suitably and present RIAPG method to solve problem (P_2).

The rest of paper is organized as follows. Section 2 introduces the constructive technique of feasible approximate minimizer that satisfies the requirement of AIAPG method. Due to the drawback of AIAPG method is that the required conditions are strict. It makes challenge to the efficiency of practical performance and the accuracy of calculation. Hence, the relax inexact accelerated proximal gradient method will be addressed more formally in Section 3. Section 4 is devoted to a series of lemmas and theorems to show the convergence analysis of the method. Finally, we have a conclusion section.

Notation. For any X, Y in S^n, $\langle X, Y\rangle$ denotes stand trace inner product, $\|\cdot\|$ and $\|\cdot\|_2$, respectively, stand for Frobenius norm and spectral norm. $\mathcal{F}^* : R^m \to S^n$ is the adjoint operator of linear operator $\mathcal{F}$ such that $\langle \mathcal{F}^* X, Y\rangle = \langle X, \mathcal{F} Y\rangle$ for all $(X, Y) \in R^m \times S^n$. To facilitate the latter specification, we have also given the following notations. Let $\mathcal{N}_k$ be a self-adjoint positive definite operator that is chosen by the user. In addition, $\{\zeta_k\}, \{\rho_k\}, \{\theta_k\}$ are all the given convergent sequences of nonnegative numbers such that $\sum_{k=1}^{\infty} \zeta_k < \infty$, $\sum_{k=1}^{\infty} \rho_k < \infty$, $\sum_{k=1}^{\infty} \theta_k < \infty$.

2. Construction of Feasible Solution

Problem (P_2) can be solved by AIAPG method [5], but note that the approximate minimizer X_k generated by above method must be feasible; that is, $\mathcal{F}(X_k) = b$ and $X_k \succeq 0$. At the same time, given Y_k in [5], the approximate solution (X_k, P_k, Z_k) should satisfy the KKT optimality conditions. More precisely

$$\begin{gathered} \nabla (h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k (X_k - Y_k) - \mathcal{F}^* P_k - Z_k =: \delta_k \approx 0, \\ \mathcal{F}(X_k) = b, \\ \langle X_k, Z_k\rangle =: \varepsilon_k \approx 0, \\ X_k \succeq 0, \quad Z_k \succeq 0. \end{gathered} \qquad (1)$$

In practice, the positive semidefiniteness of approximate solution X_k is easy to stipulate by performing projection onto S_+^n, but the vector $R_k := \mathcal{F}(X_k) - b$ is usually not exactly equal to 0. Hence, we present a strategy that uses infeasible solution X_k to construct $\widehat{X}_k$ which is a feasible solution such that $(\widehat{X}_k, P_k, Z_k)$ satisfies corresponding KKT optimality conditions.

Suppose (X_k, P_k, Z_k) satisfies the conditions $X_k \succeq 0$, $Z_k \succeq 0$, $\|\mathcal{N}_k^{-1/2}\delta_k\| \le \rho_k/\sqrt{2}t_k$, and $\varepsilon_k \le \zeta_k/(2t_k^2)$, and there exists $\overline{X} \succ 0$ such that $\mathcal{F}(\overline{X}) = b$, where $t_1 = 1$, $t_{k+1} = (1/2)(1 + \sqrt{1 + 4t_k^2})$, $\mathcal{F}$ is surjective. Then the constructive form of feasible solution is given as follows:

$$\widehat{X}_k = \lambda (X_k + W_k) + (1 - \lambda)\overline{X}, \qquad (2)$$

where $\lambda \in [0, 1]$ and $W_k = -\mathcal{F}^*(\mathcal{F}\mathcal{F}^*)^{-1}(R_k)$.

In the following paragraph, we will show that $\widehat{X}_k$ is feasible and satisfies corresponding KKT optimality conditions for above construction. By the definition of $\widehat{X}_k$, W_k, and R_k, we have

$$\begin{aligned} \mathcal{F}(\widehat{X}_k) &= \mathcal{F}[\lambda (X_k + W_k) + (1 - \lambda)\overline{X}] \\ &= \lambda \mathcal{F}(X_k + W_k) + (1 - \lambda)\mathcal{F}(\overline{X}) \\ &= \lambda [\mathcal{F}(X_k) - R_k] + (1 - \lambda) b = b. \end{aligned} \qquad (3)$$

It is easy to get $\|W_k\|_2 \le \|R_k\|/\sigma_{\min}(\mathcal{F})$ and $\widehat{X}_k$ will be positive semidefinite whenever $\lambda = 1 - (\|W_k\|_2/(\|W_k\|_2 + \lambda_{\min}(\overline{X})))$. In addition, for $\widehat{X}_k$ the following results are also valid

$$\begin{gathered} 0 \le \langle \widehat{X}_k, Z_k\rangle \le 2\varepsilon_k, \qquad \left\|\mathcal{N}_k^{-1/2}\widehat{\delta}_k\right\| \le \frac{\rho_k}{\sqrt{2}t_k}, \\ \nabla (h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k (\widehat{X}_k - Y_k) - \mathcal{F}^* P_k - Z_k \\ = \delta_k + \mathcal{N}_k (\widehat{X}_k - X_k) =: \widehat{\delta}_k. \end{gathered} \qquad (4)$$

But above results were established on the condition of the requirement of W_k, that is,

$$\begin{aligned} &\|W_k\|_2 \\ &\le \min \left\{ \frac{\zeta_k}{4t_k^2\sqrt{n}\|Z_k\|}\left(1 + \frac{\lambda_{\max}(\overline{X})}{\lambda_{\min}(\overline{X})}\right)^{-1}, \right. \\ &\qquad \left. \frac{\rho_k}{2\sqrt{2n}t_k}(\lambda_{\max}(\mathcal{N}_1))^{-1/2}\left(1 + \frac{\|\overline{X} - X_k\|_2}{\lambda_{\min}(\overline{X})}\right)^{-1} \right\}. \end{aligned} \qquad (5)$$

The proof of above conclusions is similar as in [8] and we omit it here.

Though we have succeeded in constructing a feasible solution $\widehat{X}_k \in \Omega$, the requirement of W_k is too stringent to be difficult for computational efficiency. To overcome the drawbacks above we propose RIAPG method to solve problem (P_2) for which the iterate points X_k generated by method need not be strictly contained in Ω.

3. A Relax Inexact Accelerated Proximal Gradient Method

The RIAPG algorithm for solving the problem (P_2) is described as follows.

Given a tolerance $\varepsilon > 0$. Input $Y_1 = X_0 \in \mathrm{dom}(\delta_\Omega(X))$, $t_1 = 1$. Set $k = 1$. Iterate the following steps.

Step 1. Find an approximate minimizer

$$X_k \approx \operatorname{argmin}\left\{ (h_\varepsilon \circ \lambda)(Y_k) + \langle \nabla (h_\varepsilon \circ \lambda)(Y_k), X - Y_k \rangle + \frac{1}{2}\langle X - Y_k, \mathcal{N}_k (X - Y_k)\rangle : X \in \Omega \right\}, \tag{6}$$

where X_k is allowed to be contained in a suitable enlargement $\Omega_k := \{X \in S^n : \|\mathcal{F}(X) - b\| \le \theta_k/t_k^2, X \succeq 0\}$ of Ω, and the sequence $\{\theta_k/t_k^2\}$ is monotonically decreasing. Consider

$$\nabla (h_\varepsilon \circ \lambda)(Y_k) = Q\mathrm{Diag}[\nabla h_\varepsilon(\lambda)]Q^\top = Q\mathrm{Diag}[\alpha(\varepsilon,\lambda)]Q^\top,$$
$$\alpha_i(\varepsilon,\lambda) = \frac{e^{\lambda_i/\varepsilon}}{\sum_{j=1}^n e^{\lambda_j/\varepsilon}} = \frac{e^{(\lambda_i-\lambda_1)/\varepsilon}}{\sum_{j=1}^n e^{(\lambda_j-\lambda_1)/\varepsilon}}, \quad \text{where } i = 1,\ldots,n. \tag{7}$$

Step 2. Compute $t_{k+1} = (1/2)(1 + \sqrt{1 + 4t_k^2})$.

Step 3. Compute $Y_{k+1} = X_k + ((t_k - 1)/t_{k+1})(X_k - X_{k-1})$.

Let $l_k(X) = (h_\varepsilon\circ\lambda)(Y_k) + \langle\nabla(h_\varepsilon\circ\lambda)(Y_k), X - Y_k\rangle + (1/2)\langle X - Y_k, \mathcal{N}_k(X - Y_k)\rangle$. When $\Omega_k = \Omega$ the dual of (6) is given by

$$\max\{l_k(X) - \langle \nabla l_k(X), X\rangle + \langle b, P\rangle : \nabla l_k(X) - \mathcal{F}^* P - Z = 0, Z \succeq 0, X \succeq 0\}. \tag{8}$$

We assume that the approximate minimizer X_k in (6) and its corresponding dual variables (P_k, Z_k) satisfy the following conditions:

$$\begin{aligned}
&(h_\varepsilon \circ \lambda)(X_k) \le l_k(X_k) + \frac{\zeta_k}{(2t_k^2)} \\
&\left|\langle l_k(X_k), X_k\rangle + \langle b, P_k\rangle\right| \le \Delta, \\
&\nabla l_k(X_k) - \mathcal{F}^* P_k - Z_k = \delta_k \quad \text{with } \left\|\mathcal{N}_k^{-1/2}\delta_k\right\| \le \frac{\rho_k}{\sqrt{2}t_k}, \\
&\langle X_k, Z_k\rangle \le \frac{\zeta_k}{(2t_k^2)}, \\
&\|R_k\| \le \frac{\theta_k}{t_k^2}, \quad X_k \succeq 0, \quad Z_k \succeq 0,
\end{aligned} \tag{9}$$

where Δ is a given positive number and we also assume that the sequence $\{\rho_k/t_k\}$ is monotonically decreasing.

Let X_* be the optimal solution of (P_2), and the dual of (P_2) is given as follows:

$$\max\{(h_\varepsilon\circ\lambda)(X) - \langle\nabla(h_\varepsilon\circ\lambda)(X), X\rangle + \langle b, P\rangle : \nabla(h_\varepsilon\circ\lambda)(X) - \mathcal{F}^* P - Z = 0, Z \succeq 0, X \succeq 0\}. \tag{10}$$

Let (X_*, P_*, Z_*) be the optimal solution of above dual problem.

To facilitate the later proof, we define the following quantities:

$$\begin{aligned}
&v_k = (h_\varepsilon \circ \lambda)(X_k) - (h_\varepsilon\circ\lambda)(X_*), \\
&u_k = t_k X_k - (t_k - 1)X_{k-1} - X_*, \\
&a_k = t_k^2 v_k, \qquad b_k = \frac{1}{2}\langle u_k, \mathcal{N}_k(u_k)\rangle \ge 0, \\
&e_k = t_k\langle\delta_k, u_k\rangle, \\
&\eta_k = \langle P_k, t_k^2 R_k - t_{k-1}^2 R_{k-1}\rangle, \qquad \eta_1 = \langle P_1, R_1\rangle, \\
&\chi_k = \|P_{k-1} - P_k\|\theta_k, \qquad \chi_1 = 0, \\
&\tau = \frac{1}{2}\langle X_0 - X_*, \mathcal{N}_1(X_0 - X_*)\rangle, \qquad \overline{\rho}_k = \sum_{j=1}^k \rho_j, \\
&\overline{\zeta}_k = \sum_{j=1}^k (\zeta_j + \rho_j^2).
\end{aligned} \tag{11}$$

It should be noted that comparing to the quantities of AIAPG method, a_k and v_k here may be negative since the lack of the feasibility of X_k.

4. Convergence Analysis

In the following paragraphs, a series of lemmas and theorems will be given to specify the convergence analysis of the RIAPG method. We should mention that the lack of the feasibility of X_k introduces nontrivial technical difficulties in the proof of the convergence.

Lemma 1. *Given $Y_k \in S^n$ and a positive definite linear operator $\mathcal{N}_k$ on S^n such that the conditions in* (9) *hold, then for all $X \in S_+^n$, we have*

$$\begin{aligned}
(h_\varepsilon\circ\lambda)(X) - (h_\varepsilon\circ\lambda)(X_k) &\ge \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle \\
&\quad + \langle Y_k - X, \mathcal{N}_k(X_k - Y_k)\rangle \\
&\quad + \langle \delta_k + \mathcal{F}^* P_k, X - X_k\rangle - \frac{\zeta_k}{t_k^2}.
\end{aligned} \tag{12}$$

Proof. Noting the first inequality of (9), we have

$$\begin{aligned}(h_\varepsilon \circ \lambda)(X) - (h_\varepsilon \circ \lambda)(X_k) &\geq (h_\varepsilon \circ \lambda)(X) - l_k(X_k) - \frac{\zeta_k}{2t_k^2} \\ &= (h_\varepsilon \circ \lambda)(X) - (h_\varepsilon \circ \lambda)(Y_k) \\ &\quad - \langle \nabla (h_\varepsilon \circ \lambda)(Y_k), X_k - Y_k \rangle \\ &\quad - \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle - \frac{\zeta_k}{2t_k^2}.\end{aligned} \tag{13}$$

By the convexity of $(h_\varepsilon \circ \lambda)(X)$, we have

$$(h_\varepsilon \circ \lambda)(X) - (h_\varepsilon \circ \lambda)(Y_k) \geq \langle \nabla (h_\varepsilon \circ \lambda)(Y_k), X - Y_k \rangle. \tag{14}$$

Then,

$$\begin{aligned}(h_\varepsilon \circ \lambda)(X) - (h_\varepsilon \circ \lambda)(X_k) &\geq \langle \nabla (h_\varepsilon \circ \lambda)(Y_k), X - X_k \rangle \\ &\quad - \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle - \frac{\zeta_k}{2t_k^2}.\end{aligned} \tag{15}$$

Since $\nabla l_k(X_k) = \nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)$ and the third inequality of (9), we have

$$\begin{aligned}&(h_\varepsilon \circ \lambda)(X) - (h_\varepsilon \circ \lambda)(X_k) \\ &\quad\geq \langle \delta_k + \mathcal{F}^* P_k + Z_k - \mathcal{N}_k(X_k - Y_k), X - X_k \rangle \\ &\qquad - \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle - \frac{\zeta_k}{2t_k^2} \\ &\quad= \langle \delta_k + \mathcal{F}^* P_k, X - X_k \rangle + \langle Z_k, X \rangle \\ &\qquad - \langle \mathcal{N}_k(X_k - Y_k), X - X_k \rangle \\ &\qquad - \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle - \frac{\zeta_k}{2t_k^2} - \langle Z_k, X_k \rangle.\end{aligned} \tag{16}$$

Then the required result is proved by considering the fact that $\langle Z_k, X \rangle \geq 0$ and $\langle X_k, Z_k \rangle \leq \zeta_k/(2t_k^2)$. □

Lemma 2. *Suppose that $\mathcal{N}_{k-1} \succeq \mathcal{N}_k \succ 0$, for all k. Then*

(i) $a_{k-1} + b_{k-1} \geq a_k + b_k - e_k - \zeta_k - \eta_k$;

(ii) *in addition, the conditions in* (9) *are satisfied for all k. Then*

$$a_k + b_k \leq (\sqrt{\tau} + \bar{\rho}_k)^2 + \|P_k\| \theta_k + 2(\bar{\zeta}_k + \bar{\chi}_k + J_k), \tag{17}$$

where $J_k = \sum_{j=1}^{k} \rho_j \sqrt{A_j}$, $A_j = \|P_j\|\theta_j + a_j^*$, $a_j^* = \max\{0, -a_j\}$.

Proof. (i) According to Lemma 1, taking $X = X_{k-1}$ in (12), we have

$$\begin{aligned}v_{k-1} - v_k &\geq \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\quad + \langle Y_k - X_{k-1}, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\quad + \langle \delta_k + \mathcal{F}^* P_k, X_{k-1} - X_k \rangle - \frac{\zeta_k}{t_k^2}.\end{aligned} \tag{18}$$

Similarly, taking $X = X_*$ in (12), we have

$$\begin{aligned}-v_k &\geq \frac{1}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\quad + \langle Y_k - X_*, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\quad + \langle \delta_k + \mathcal{F}^* P_k, X_* - X_k \rangle - \frac{\zeta_k}{t_k^2}.\end{aligned} \tag{19}$$

By multiplying (18) throughout by $t_k - 1$ and adding that to (19), we have

$$\begin{aligned}&(t_k - 1) v_{k-1} - t_k v_k \\ &\quad\geq \frac{t_k}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\qquad + \langle t_k Y_k - (t_k - 1) X_{k-1} - X_*, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\qquad - \langle \delta_k + \mathcal{F}^* P_k, t_k X_k - (t_k - 1) X_{k-1} - X_* \rangle - \frac{\zeta_k}{t_k}.\end{aligned} \tag{20}$$

In addition, by multiplying (20) throughout by t_k, and using $t_{k-1}^2 = t_k^2 - t_k$, we have

$$\begin{aligned}&a_{k-1} - a_k \\ &\quad\geq \frac{t_k^2}{2}\langle X_k - Y_k, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\qquad + t_k \langle t_k Y_k - (t_k - 1) X_{k-1} - X_*, \mathcal{N}_k(X_k - Y_k)\rangle \\ &\qquad - \left\langle \delta_k + \mathcal{F}^* P_k, t_k^2 X_k - t_{k-1}^2 X_{k-1} - t_k X_* \right\rangle - \zeta_k \\ &\quad\geq \frac{1}{2}\langle u_k, \mathcal{N}_k u_k \rangle - \frac{1}{2}\langle u_{k-1}, \mathcal{N}_k u_{k-1} \rangle \\ &\qquad - \langle \delta_k + \mathcal{F}^* P_k, t_k u_k \rangle - \zeta_k \\ &\quad\geq b_k - b_{k-1} - e_k - \langle \mathcal{F}^* P_k, t_k u_k \rangle - \zeta_k.\end{aligned} \tag{21}$$

Note that the second inequality above follows the fact that the definition of Y_k and $t_k^2 X_k - t_{k-1}^2 X_{k-1} - t_k X_* = t_k u_k$. Since $\mathcal{N}_{k-1} \succeq \mathcal{N}_k \succ 0$, $t_{k-1}^2 = t_k^2 - t_k$ and (11), we have

$$\begin{aligned}&\langle \mathcal{F}^* P_k, t_k u_k \rangle \\ &\quad= \langle P_k, \mathcal{F}(t_k u_k) \rangle \\ &\quad= \left\langle P_k, t_k^2 (\mathcal{F}(X_k) - b) - t_{k-1}^2 (\mathcal{F}(X_{k-1}) - b) \right\rangle \\ &\quad= \left\langle P_k, t_k^2 R_k - t_{k-1}^2 R_{k-1} \right\rangle = \eta_k.\end{aligned} \tag{22}$$

Then the result (i) is proved.

(ii) We have $|e_k| = |t_k\langle \delta_k, u_k\rangle| \le \|\mathcal{N}_k^{-1/2}\delta_k\|\|\mathcal{N}_k^{1/2}u_k\|t_k \le \rho_k\|H_k^{1/2}u_k\|/\sqrt{2} = \rho_k\sqrt{b_k}$.

First, we show that $a_1 + b_1 \le \tau + |\langle P_1, R_1\rangle| + \rho_1\sqrt{b_1} + \zeta_1$. As $a_1 = (h_\varepsilon \circ \lambda)(X_1) - (h_\varepsilon \circ \lambda)(X_*)$, $b_1 = (1/2)\langle X_1 - X_*, \mathcal{N}_1(X_1 - X_*)\rangle$, $Y_1 = X_0$, $t_1 = 1$. By applying the Lemma 1, taking $X = X_*$ we have

$$\begin{aligned} -a_1 &= (h_\varepsilon \circ \lambda)(X_*) - (h_\varepsilon \circ \lambda)(X_1) \\ &\ge \frac{1}{2}\langle X_1 - Y_1, \mathcal{N}_1(X_1 - Y_1)\rangle \\ &\quad + \langle Y_1 - X_*, \mathcal{N}_1(X_1 - Y_1)\rangle \\ &\quad + \langle \delta_1 + \mathcal{F}^* P_1, X_* - X_1\rangle - \frac{\zeta_1}{t_1^2} \\ &= \frac{1}{2}\langle X_1 - X_*, \mathcal{N}_1(X_1 - X_*)\rangle \\ &\quad - \frac{1}{2}\langle Y_1 - X_*, \mathcal{N}_1(Y_1 - X_*)\rangle \\ &\quad + \langle \delta_1 + \mathcal{F}^* P_1, X_* - X_1\rangle - \zeta_1 \\ &= b_1 - \tau + \langle \delta_1, X_* - X_1\rangle + \langle \mathcal{F}^* P_1, X_* - X_1\rangle - \zeta_1. \end{aligned} \tag{23}$$

Since $\|\mathcal{N}_1^{-1/2}\delta_1\| \le \rho_1/\sqrt{2}$ and $e_1 = \langle \delta_1, X_1 - X_*\rangle$, we have

$$\begin{aligned} a_1 + b_1 &\le \tau - e_1 + \langle P_1, \mathcal{F}(X_* - X_1)\rangle + \zeta_1 \\ &\le \tau + |\langle P_1, R_1\rangle| + \rho_1\sqrt{b_1} + \zeta_1. \end{aligned} \tag{24}$$

Next we will show

$$a_k + b_k \le \tau + \left|\left\langle P_k, t_k^2 R_k\right\rangle\right| + s_k, \tag{25}$$

where $s_k = \sum_{j=1}^k \rho_j\sqrt{b_j} + \sum_{j=1}^k \zeta_j + \sum_{j=1}^k \chi_j$. By (i), we can get

$$\begin{aligned} \tau &\ge a_1 + b_1 - \rho_1\sqrt{b_1} - \zeta_1 - \eta_1 \\ &\ge a_2 + b_2 - e_2 - \zeta_2 - \eta_2 - \rho_1\sqrt{b_1} - \zeta_1 - \eta_1 \\ &\ge a_2 + b_2 - \rho_2\sqrt{b_2} - \rho_1\sqrt{b_1} - \zeta_1 - \zeta_2 - \eta_1 - \eta_2 \\ &\ge a_k + b_k - \sum_{j=1}^k \rho_j\sqrt{b_j} - \sum_{j=1}^k \zeta_j - \sum_{j=1}^k \eta_j. \end{aligned} \tag{26}$$

We use the fact that

$$\begin{aligned} \sum_{j=1}^k \eta_j &= \left\langle P_k, t_k^2 R_k\right\rangle + \sum_{j=1}^{k-1}\left\langle P_j - P_{j+1}, t_j^2 R_j\right\rangle \\ &\le \left|\left\langle P_k, t_k^2 R_k\right\rangle\right| + \sum_{j=1}^{k-1}\left\|P_j - P_{j+1}\right\| t_j^2 \\ &\quad \cdot \frac{\theta_k}{t_j^2} = \left|\left\langle P_k, t_k^2 R_k\right\rangle\right| + \sum_{j=1}^k \chi_j; \end{aligned} \tag{27}$$

consequently, (25) holds. Hence

$$\begin{aligned} & b_k \le \tau_k + s_k, \\ & \text{where } \tau_k := \tau + \left|\left\langle P_k, t_k^2 R_k\right\rangle\right| - a_k \le \tau + A_k. \end{aligned} \tag{28}$$

Then we can get

$$\begin{aligned} s_k &= s_{k-1} + \rho_k\sqrt{b_k} + \zeta_k + \chi_k \\ &\le s_{k-1} + \rho_k\sqrt{\tau_k + s_k} + \zeta_k + \chi_k. \end{aligned} \tag{29}$$

According to (28), we have $\tau_1 \ge b_1 - \rho_1\sqrt{b_1} - \zeta_1$; this implies

$$\sqrt{b_1} \le \frac{1}{2}\left(\rho_1 + \sqrt{\rho_1^2 + 4(\tau_1 + \zeta_1)}\right) \le \rho_1 + \sqrt{\tau_1 + \zeta_1}, \tag{30}$$

and then

$$\begin{aligned} s_1 &= \rho_1\sqrt{b_1} + \zeta_1 \le \rho_1\left(\rho_1 + \sqrt{\tau_1 + \zeta_1}\right) + \zeta_1 \\ &\le \rho_1^2 + \zeta_1 + \rho_1\left(\sqrt{\tau_1} + \sqrt{\zeta_1}\right). \end{aligned} \tag{31}$$

Adding τ_k to both sides of (29) and moving the terms, we get

$$(\tau_k + s_k) - \rho_k\sqrt{\tau_k + s_k} - (\tau_k + s_{k-1} + \zeta_k + \chi_k) \le 0; \tag{32}$$

this implies

$$\sqrt{\tau_k + s_k} \le \frac{1}{2}\left[\rho_k + \sqrt{\rho_k^2 + 4(\tau_k + s_{k-1} + \zeta_k + \chi_k)}\right]; \tag{33}$$

thus,

$$\begin{aligned} s_k &\le s_{k-1} + \frac{1}{2}\rho_k\left[\rho_k + \sqrt{\rho_k^2 + 4(\tau_k + s_{k-1} + \zeta_k + \chi_k)}\right] \\ &\quad + \zeta_k + \chi_k \\ &\le s_{k-1} + \frac{1}{2}\rho_k^2 + \frac{1}{2}\rho_k\sqrt{\rho_k^2 + 4(\tau + A_k + s_{k-1} + \zeta_k + \chi_k)} \\ &\quad + \zeta_k + \chi_k \\ &\le s_{k-1} + \left(\rho_k^2 + \zeta_k\right) + \chi_k + \rho_k\sqrt{\tau + A_k} \\ &\quad + \rho_k\sqrt{s_{k-1} + \zeta_k + \chi_k} \\ &\le s_1 + \sum_{j=2}^k\left(\rho_j^2 + \zeta_j\right) + \sum_{j=2}^k \chi_j \\ &\quad + \sum_{j=2}^k \rho_j\sqrt{\tau + A_j} + \sum_{j=2}^k \rho_j\sqrt{s_{j-1} + \zeta_j + \chi_j} \\ &\le \sum_{j=1}^k\left(\rho_j^2 + \zeta_j\right) + \sum_{j=1}^k \chi_j + \sum_{j=1}^k \rho_j\sqrt{\tau} \\ &\quad + \sum_{j=1}^k \rho_j\sqrt{A_j} + \sum_{j=1}^k \rho_j\sqrt{s_j} \\ &\le \bar{\zeta}_k + \bar{\chi}_k + \bar{\rho}_k\sqrt{\tau} + \bar{J}_k + \bar{\rho}_k\sqrt{s_k}. \end{aligned} \tag{34}$$

In the last two inequalities, we use the fact that $s_{j-1}+\zeta_j+\chi_j \le s_j$ and $0 \le s_1 \le s_2 \le \cdots \le s_k$.

Let $\omega_k := \overline{\zeta}_k + \overline{\chi}_k + \overline{J}_k + \overline{\rho}_k\sqrt{\tau}$; then we have $\sqrt{s_k} \le (1/2)(\overline{\rho}_k + \sqrt{\overline{\rho}_k^2 + 4\omega_k})$ which implies

$$s_k \le \overline{\rho}_k^2 + \omega_k. \tag{35}$$

The result (ii) follows from (35), (25), and the fact that $|\langle P_k, t_k^2 R_k\rangle| \le \|P_k\|\theta_k$. □

Lemma 3. *(i) Suppose that there exists (X', P', Z') such that*

$$\begin{gathered}\mathcal{F}\left(X'\right) = b, \quad X' \ge 0, \\ \nabla\left(h_\varepsilon \circ \lambda\right)\left(X'\right) = \mathcal{F}^* P' + Z', \quad Z' \succ 0.\end{gathered} \tag{36}$$

If the sequence $\{(h_\varepsilon \circ \lambda)(X_k)\}$ is bounded from above, then the sequence $\{X_k\}$ is bounded.

(ii) Suppose that $\mathcal{N}_{k-1} \succeq \mathcal{N}_k \succ 0$ for all k, $\{X_k\}$ is bounded, and there exists $\widetilde{X}$ such that

$$\mathcal{F}\left(\widetilde{X}\right) = b, \quad \widetilde{X} \ge 0. \tag{37}$$

Then the sequence $\{Z_k\}$ is bounded. In addition, the sequence $\{P_k\}$ is also bounded.

Proof. (i) By using the convexity of $(h_\varepsilon \circ \lambda)(\cdot)$, $X_k \in \Omega_k$, and monotonicity of the sequence of $\{\theta_k/t_k^2\}$, we have

$$\begin{aligned}&(h_\varepsilon \circ \lambda)\left(X'\right) - (h_\varepsilon \circ \lambda)(X_k) \\ &\quad\le \left\langle \nabla(h_\varepsilon \circ \lambda)\left(X'\right), X' - X_k\right\rangle = \left\langle \mathcal{F}^* P' + Z', X' - X_k\right\rangle \\ &\quad= \left\langle P', \mathcal{F}\left(X'\right) - \mathcal{F}(X_k)\right\rangle + \left\langle Z', X'\right\rangle - \left\langle Z', X_k\right\rangle \\ &\quad\le \left\|P'\right\| \|b - \mathcal{F}(X_k)\| + \left\langle Z', X'\right\rangle - \left\langle Z', X_k\right\rangle \\ &\quad\le \left\|P'\right\| \theta_1 + \left\langle Z', X'\right\rangle - \left\langle Z', X_k\right\rangle.\end{aligned} \tag{38}$$

Thus

$$\begin{aligned}&\lambda_{\min}\left(Z'\right) T_r(X_k) \\ &\quad\le \left\langle X_k, Z'\right\rangle \\ &\quad\le \left\|P'\right\| \theta_1 + \left\langle Z', X'\right\rangle - (h_\varepsilon \circ \lambda)\left(X'\right) + (h_\varepsilon \circ \lambda)(X_k).\end{aligned} \tag{39}$$

Then the required result is proved.

(ii) Noting (9) and monotonicity of the sequence of $\{\rho_k/t_k\}$, we have

$$\begin{aligned}&\lambda_{\min}\left(\widetilde{X}\right) T_r(Z_k) \\ &\quad\le \langle \widetilde{X}, Z_k\rangle \\ &\quad= \left\langle \widetilde{X}, \nabla l_k(X_k) - \mathcal{F}^* P_k - \delta_k\right\rangle \\ &\quad= \left(\left\langle X_k, \nabla l_k(X_k)\right\rangle - \langle b, P_k\rangle\right) \\ &\qquad+ \left\langle \widetilde{X} - X_k, \nabla l_k(X_k)\right\rangle - \left\langle \widetilde{X}, \delta_k\right\rangle \\ &\quad\le \Delta + \left\langle \widetilde{X} - X_k, \nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)\right\rangle \\ &\qquad+ \left\langle \mathcal{N}_k^{1/2}\widetilde{X}, \mathcal{N}_k^{-1/2}\delta_k\right\rangle \\ &\quad\le \Delta + \left\|\widetilde{X} - X_k\right\| \|\nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)\| \\ &\qquad+ \frac{\left\|\mathcal{N}_k^{1/2}\widetilde{X}\right\| \rho_1}{\sqrt{2}}.\end{aligned} \tag{40}$$

Then that the sequence $\{Y_k\}$ is bounded follows from the fact that $\{X_k\}$ is bounded. By the continuity of $\nabla((h_\varepsilon \circ \lambda))$ and the fact that $\mathcal{N}_1 \succeq \mathcal{N}_k \succ 0$, the sequence $\{\|\nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)\|\}$ is also bounded.

Next, we will show that $\{P_k\}$ is bounded. Take $\mathcal{F}^\circ = (\mathcal{F}\mathcal{F}^*)^{-1}\mathcal{F}$ and we can get $P_k = \mathcal{F}^\circ(\nabla l_k(X_k) - Z_k - \delta_k)$ from (9). Hence

$$\begin{aligned}&\|P_k\| \\ &\quad\le \left\|\mathcal{F}^\circ\right\| \|\nabla l_k(X_k) - Z_k - \delta_k\| \\ &\quad\le \left\|\mathcal{F}^\circ\right\| \left(\|\nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)\| + \|Z_k\| + \|\delta_k\|\right).\end{aligned} \tag{41}$$

So we have $\|\delta_k\| \le \sqrt{\lambda_{\max}(\mathcal{N}_1)}\|\mathcal{N}_k^{-1/2}\delta_k\| \le \sqrt{\lambda_{\max}(\mathcal{N})_1}\rho_1/\sqrt{2}$ directly from $\lambda_{\max}(\mathcal{N}_1)I \succeq \mathcal{N}_1 \succeq \mathcal{N}_k$. Then the boundedness of $\{P_k\}$ is proved by using the fact that the sequence $\{\|\nabla(h_\varepsilon \circ \lambda)(Y_k) + \mathcal{N}_k(X_k - Y_k)\|\}$ and $\{Z_k\}$ are bounded. □

Lemma 4. *For all $k \ge 1$, we have*

$$0 \le (h_\varepsilon \circ \lambda)(X_*) - (h_\varepsilon \circ \lambda)\left(X_*^k\right) \le \frac{\|P_*\|\theta_k}{t_k^2}, \tag{42}$$

where X_*^k is an optimal solution of the problem $\min\{(h_\varepsilon \circ \lambda)(X) : X \in \Omega_k\}$.

Proof. By the convexity of $(h_\varepsilon \circ \lambda)(X)$ and the fact that $\mathcal{F}(X_*) = b$, $\langle X_*, Z_*\rangle = 0$, $\langle Z_*, X_*^k\rangle \ge 0$, we have

$$\begin{aligned}&0 \le (h_\varepsilon \circ \lambda)(X_*) - (h_\varepsilon \circ \lambda)\left(X_*^k\right) \\ &\quad\le \left\langle \nabla(h_\varepsilon \circ \lambda)(X_*), X_* - X_*^k\right\rangle = \left\langle \mathcal{F}^* P_* + Z_*, X_* - X_*^k\right\rangle\end{aligned}$$

$$= \left\langle P_*, \mathcal{F}(X_*) - \mathcal{F}\left(X_*^k\right)\right\rangle + \langle Z_*, X_* \rangle - \left\langle Z_*, X_*^k \right\rangle$$
$$\leq \|P_*\| \left\|b - \mathcal{F}\left(X_*^k\right)\right\| \leq \frac{\|P_*\| \theta_k}{t_k^2}. \tag{43}$$

□

Theorem 5. *Suppose that* $\mathcal{N}_{k-1} \succeq \mathcal{N}_k \succ 0$ *for all k. Taking* $M_k := \max_{1\leq j\leq k}\{\sqrt{(\|P_*\| + \|P_j\|)\theta_j}\}$, *we have*

$$\begin{aligned} &-\frac{4\|P_*\|\theta_k}{(k+1)^2} \\ &\quad \leq (h_\varepsilon \circ \lambda)(X_k) - (h_\varepsilon \circ \lambda)(X_*) \\ &\quad \leq \frac{4}{(k+1)^2} \\ &\qquad \times \left(\left(\sqrt{\tau} + \overline{\rho_k}\right)^2 + 2\overline{\rho_k} M_k + 2\left(\zeta_k + \chi_k\right) + \|P_k\|\theta_k\right). \end{aligned} \tag{44}$$

Proof. Taking the problem $\min\{(h_\varepsilon \circ \lambda)(X) : X \in \Omega_k\}$ into account, X_*^k is an optimal solution of it. Since $X_*, X_k \in \Omega_k$, we have

$$\begin{aligned} &(h_\varepsilon \circ \lambda)\left(X_*^k\right) - (h_\varepsilon \circ \lambda)(X_*) \\ &\qquad \leq (h_\varepsilon \circ \lambda)(X_k) - (h_\varepsilon \circ \lambda)(X_*). \end{aligned} \tag{45}$$

Then the inequality on the left side of (44) follows from Lemma 4, (45) and the fact that $t_k \geq (k+1)/2$.

Next, we will show the inequality on the right side of (44). By Lemma 2(ii) and using $b_k \geq 0$, we have

$$\begin{aligned} t_k^2 v_k &= t_k^2 \left((h_\varepsilon \circ \lambda)(X_k) - (h_\varepsilon \circ \lambda)(X_*)\right) \\ &= a_k \leq \left(\sqrt{\tau} + \overline{\rho}_k\right)^2 + \|P_k\|\theta_k + 2\left(\overline{\zeta}_k + \overline{\chi}_k + J_k\right). \end{aligned} \tag{46}$$

Since $-a_j = t_j^2((h_\varepsilon \circ \lambda)(X_*) - (h_\varepsilon \circ \lambda)(X_j)) \leq t_j^2((h_\varepsilon \circ \lambda)(X_*) - (h_\varepsilon \circ \lambda)(X_*^j)) \leq \|P_*\|\theta_j$. Then we have

$$\begin{aligned} a_j^* &\leq \|P_*\|\theta_j, \\ J_k &\leq \sum_{j=1}^{k} \rho_j \sqrt{\left(\|P_j\| + \|P_*\|\right)\theta_j} \leq M_k \overline{\rho}_k. \end{aligned} \tag{47}$$

Using $t_k \geq (k+1)/2$ again, the required result is proved. □

From the assumption on the sequences of $\{\zeta_k\}$, $\{\theta_k\}$, $\{\rho_k\}$, we can get the result that the sequences $\{\overline{\rho}_k\}$ and $\{\overline{\theta}_k\}$ are bounded. Moreover, by using Lemma 3, we note the sequence $\{\|P_k\|\}$ is also bounded; at the same time, we can also get the boundedness of $\{M_k\}$ and $\{\overline{\chi}_k\}$. Then the convergence of the RIAPG method with the convergent rate $O(1/k^2)$ is proved.

5. Conclusion

The principal result given here is that we have presented the implementable and globally convergent method (RIAPG method) for solving the constrained minimization problem of maximum eigenvalue functions. RIAPG method, being an extension of AIAPG method, is especially suited for the case where the approximate minimizer generated by AIAPG method may not be in the feasible set. Though this method is based on some assumptions, it enriches the way to deal with the constrained minimization problem of maximum eigenvalue functions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank the referees for their beneficial suggestions for the improvement of this paper. This paper is supported by the National Natural Science Foundation of China under Project no. 11171138.

References

[1] E. S. Mistakidis and G. E. Stavroulakis, *Nonconvex Optimization in Mechanics. Smooth and Nonsmooth Algotithms , Heuristics and Engineering Applications*, Kluwer Academic Publisher, Dordrechet, The Netherlands, 1998.

[2] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[3] C. A. Micchelli, L. Shen, and Y. Xu, "Proximity algorithms for image models: denoising," *Inverse Problems*, vol. 27, no. 4, Article ID 045009, 30 pages, 2011.

[4] Y. Nesterov, "Gradient methods for minimizing composite functions," *Mathematical Programming*, vol. 140, no. 1, pp. 125–161, 2013.

[5] W. Wang, J. J. Gao, and S. H. Li, "Approximate inexact accelerated proximal gradient method for the minimization problem of a class of maximum eigenvalue functions," *Journal of Liaoning Normal University*, vol. 36, no. 3, pp. 314–317, 2013.

[6] X. Chen, H. Qi, L. Qi, and K.-L. Teo, "Smooth convex approximation to the maximum eigenvalue function," *Journal of Global Optimization*, vol. 30, no. 2-3, pp. 253–270, 2004.

[7] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical Programming*, vol. 103, no. 1, pp. 127–152, 2005.

[8] K. Jiang, D. Sun, and K. Toh, "An inexact accelerated proximal gradient method for large scale linearly constrained convex SDP," *SIAM Journal on Optimization*, vol. 22, no. 3, pp. 1042–1064, 2012.