Hindawi
Journal of Function Spaces
Volume 2017, Article ID 5309698, 6 pages
https://doi.org/10.1155/2017/5309698

Research Article

A $\mathcal{UV}$-Method for a Class of Constrained Minimized Problems of Maximum Eigenvalue Functions

Wei Wang, Ming Jin, Shanghua Li, and Xinyu Cao

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

Correspondence should be addressed to Wei Wang; wangwei@lnnu.edu.cn

Received 31 August 2016; Revised 7 November 2016; Accepted 4 December 2016; Published 2 January 2017

Academic Editor: Stojan Radenovic

In this paper, we apply the $\mathcal{UV}$-algorithm to solve the constrained minimization problem of a maximum eigenvalue function which is the composite function of an affine matrix-valued mapping and its maximum eigenvalue. Here, we convert the constrained problem into its equivalent unconstrained problem by the exact penalty function. However, the equivalent problem involves the sum of two nonsmooth functions, which makes it difficult to apply $\mathcal{UV}$-algorithm to get the solution of the problem. Hence, our strategy first applies the smooth convex approximation of maximum eigenvalue function to get the approximate problem of the equivalent problem. Then the approximate problem, the space decomposition, and the $\mathcal{U}$-Lagrangian of the object function at a given point will be addressed particularly. Finally, the $\mathcal{UV}$-algorithm will be presented to get the approximate solution of the primal problem by solving the approximate problem.

1. Introduction

The eigenvalue optimization problems have attracted wide attention to the nonsmooth optimization. Such problems arise from many applications such as signal recovery [1], shape optimization [2], and robotics [3]. Therefore, the research on methods for solving such problems plays an important role in enriching the blend of classical mathematical techniques and contemporary optimization theory. Various methods have been proposed to deal with such problems; for example, the bundle method was used by Helmberg and Oustry to solve a class of unconstrained maximum eigenvalue optimization problems [4]. Recently, Oustry applied $\mathcal{U}$-Newton algorithm to solve the maximum eigenvalue optimization problem [5]. However, this method must satisfy the transversality condition. In this paper, we design a $\mathcal{UV}$-algorithm which does not satisfy the strict condition above to solve the constrained maximum eigenvalue optimization problem approximately. Here, we focus our attention on the following mode problem particularly:

$$\begin{aligned} \min \quad & \lambda_1 (A(x)) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad i = 1, 2, \ldots, m, \end{aligned} \tag{P}$$

where $\lambda_1(A(x))$ is the maximum eigenvalue function and the mapping $A(x) := A_0 + \mathscr{A}(x)$ is affine, $A_0 \in S_n$ is given, $\mathscr{A} : R^n \to S_n$ is a linear operator, and S_n is the space of $n \times n$ symmetric matrices. Consider an exact penalty function associated with (P) as follows:

$$P(x) := \lambda_1 (A(x)) + \pi \max \{ f_0(x), f_1(x), \ldots, f_m(x) \}, \tag{1}$$

where $f_0(x) \equiv 0$, $\nabla f_0(x) \equiv 0$, and $\pi > 0$ is a penalty parameter. For π large enough, it is well known that the problem (P) is equivalent to the following form:

$$\min_{x \in R^n} P(x). \tag{P_1}$$

It is known that the $\mathcal{UV}$-decomposition theory must be applied on the condition that the dimension of the $\mathcal{V}$-space is not full dimensional. Since $P(x)$ inherits the nondifferentiability of $\lambda_1(A(x))$ and the function $\max\{f_i\}$, it is difficult to apply $\mathcal{UV}$-decomposition theory to (P_1) in that the $\mathcal{V}$-space of $P(x)$ at a given point is full dimensional. Hence, it is imperative to consider the smooth approximation function

$\theta_\varepsilon(x)$ [6] to the function $\lambda_1(A(x))$. Then the approximate problem of (P_1) is given as follows:

$$\min_{x\in R^n} P_\varepsilon(x), \qquad (P_2)$$

where $P_\varepsilon(x) := \theta_\varepsilon(x) + \pi \max\{f_0(x), f_1(x), \ldots, f_m(x)\}$. Thus the problem (P_2) can be solved by $\mathscr{U}\mathscr{V}$-algorithm and we can get the approximate solution of the problem (P) at the same time.

The rest of the paper is organized as follows. Section 2 introduces three equivalent $\mathscr{U}\mathscr{V}$-space decomposition definitions of $P_\varepsilon(x)$, associated with a given $\overline{x} \in R^n$. The $\mathscr{U}$-Lagrangian of $P_\varepsilon(x)$ and relevant property will be addressed more detailedly in Section 3. Section 4 is devoted to the $\mathscr{U}\mathscr{V}$-algorithm for solving the approximate problem and the convergence analysis of the method. Finally, Section 5 gives some conclusive comments.

To be convenient for explanation, we give the set of the act indicators throughout the paper

$$J(x) := \left\{ j \in \{0, 1, \ldots, m\} : \theta_\varepsilon(x) + \pi f_j(x) = P_\varepsilon(x) \right\} \qquad (2)$$

and set

$$\varphi(x) := \pi \max\{f_0(x), f_1(x), \ldots, f_m(x)\}. \qquad (3)$$

The solution of the problem (P) depends on the study of the objective function of problem (P_2). The $\mathscr{U}\mathscr{V}$-space decomposition theory of $P_\varepsilon(x)$ will be shown firstly.

2. $\mathscr{U}\mathscr{V}$-Space Decomposition for $P_\varepsilon(x)$

Firstly, we can easily obtain the description of the subdifferential about $P_\varepsilon(x)$ as follows:

$$\begin{aligned} \partial P_\varepsilon(x) &= \nabla\theta_\varepsilon(x) + \partial\varphi(x) \\ &= \nabla\theta_\varepsilon(x) + \pi \operatorname{conv}\left\{\nabla f_j(x)\right\}_{j\in J(x)} \end{aligned} \qquad (4)$$

and the relative interior of $\partial P_\varepsilon(x)$

$$\operatorname{ri}\partial P_\varepsilon(x) = \nabla\theta_\varepsilon(x) + \operatorname{ri}\partial\varphi(x). \qquad (5)$$

We start by defining a decomposition of space $R^n = \mathscr{U} \oplus \mathscr{V}$, associated with a given $\overline{x} \in R^n$. We give three definitions for the subspaces $\mathscr{U}$ and $\mathscr{V}$ as follows.

Definition 1. (i) Define $\mathscr{U}_1$ as the subspace where $P'_\varepsilon(\overline{x}, \cdot)$ is linear and take $\mathscr{V}_1 := \mathscr{U}_1^\perp$, and since $P'_\varepsilon(\overline{x}, \cdot)$ is sublinear, we have

$$\mathscr{U}_1 := \left\{ d \in R^n : P'_\varepsilon(\overline{x}, d) = -P'_\varepsilon(\overline{x}, -d) \right\}. \qquad (6)$$

(ii) Define $\mathscr{V}_2$ as the subspace parallel to the affine hull of $\partial P_\varepsilon(x)$; in other words,

$$\mathscr{V}_2 := \operatorname{lin}\left(\partial P_\varepsilon(\overline{x}) - \overline{g}\right), \qquad (7)$$

where $\overline{g} \in \partial P_\varepsilon(\overline{x})$ is arbitrary and take $\mathscr{U}_2 := \mathscr{V}_2^\perp$.

(iii) Define $\mathscr{U}_3$ and $\mathscr{V}_3$ as the normal and tangent cones to $\partial P_\varepsilon(\overline{x})$ at a given point g°; that is,

$$\begin{aligned} \mathscr{U}_3 &:= N_{\partial P_\varepsilon(\overline{x})}(g^\circ), \\ \mathscr{V}_3 &:= T_{\partial P_\varepsilon(\overline{x})}(g^\circ). \end{aligned} \qquad (8)$$

In the meantime, $\mathscr{U}_3$ and $\mathscr{V}_3$ are subspaces.

Theorem 2. *In Definition 1, we have the following:*

(i) The subspace $\mathscr{U}_3$ is actually given by

$$\begin{aligned} \mathscr{U}_3 = \Big\{ d \in R^n : \left\langle g' + \nabla\theta_\varepsilon(\overline{x}) - g^\circ, d \right\rangle = 0,\ \forall g' \\ \in \partial\varphi(\overline{x}) \Big\} = N_{\partial\varphi(\overline{x})}\left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right) \end{aligned} \qquad (9)$$

and is independent of the particular $g^\circ \in ri\partial\varphi(\overline{x})$.

(ii) Subspace $\mathscr{V}_2$ is parallel to the affine hull of $\partial\varphi(\overline{x})$; that is,

$$\mathscr{V}_2 = \operatorname{lin}\left(\partial\varphi(\overline{x}) - g'\right), \quad \forall g' \in \partial\varphi(\overline{x}). \qquad (10)$$

More specifically, $\mathscr{V}_2 = \operatorname{lin}\{\nabla f_j(\overline{x}) - \nabla f_0(\overline{x})\}_{j\in J(\overline{x})}$.

(iii) $\mathscr{U}_1 = \mathscr{U}_2 = \mathscr{U}_3 =: \mathscr{U}$.

Proof. (i) On one hand, by Definition 1 and a normal cone, we have

$$\begin{aligned} & d \in N_{\partial P_\varepsilon(\overline{x})}(g^\circ) \Longleftrightarrow \\ & \left\langle d, g' + \nabla\theta_\varepsilon(\overline{x}) - g^\circ \right\rangle \le 0 \Longleftrightarrow \\ & \left\langle d, g' - \left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right) \right\rangle \le 0 \Longleftrightarrow \\ & d \in N_{\partial\varphi(\overline{x})}\left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right), \end{aligned} \qquad (11)$$

where $g' + \nabla\theta_\varepsilon(\overline{x}) \in \partial P_\varepsilon(\overline{x})$, $g' \in \partial\varphi(\overline{x})$, and $g^\circ \in \operatorname{ri}\partial P_\varepsilon(\overline{x})$.

On the other hand, let

$$\begin{aligned} N := \Big\{ d \in R^n : \left\langle d, g' + \nabla\theta_\varepsilon(\overline{x}) - g^\circ \right\rangle = 0,\ \forall g' \\ \in \partial\varphi(\overline{x}) \Big\}. \end{aligned} \qquad (12)$$

By the definition of a normal cone, $N \subseteq N_{\partial\varphi(\overline{x})}(g^\circ - \nabla\theta_\varepsilon(\overline{x}))$. Next, we only need to establish the converse case. Let $d \in N_{\partial\varphi(\overline{x})}(g^\circ - \nabla\theta_\varepsilon(\overline{x}))$ and $g' \in \partial\varphi(\overline{x})$ and it suffices to prove $\langle d, g' - (g^\circ - \nabla\theta_\varepsilon(\overline{x}))\rangle \ge 0$. Indeed, let

$$\beta = -\frac{g' - \left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right)}{\left\| g' - \left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right)\right\|} \in \mathscr{V}_2. \qquad (13)$$

By the definition of relative interior, there exists a positive constant η such that $g^\circ - \nabla\theta_\varepsilon(\overline{x}) + \eta\beta \in \partial\varphi(\overline{x})$ and $d \in N_{\partial\varphi(\overline{x})}(g^\circ - \nabla\theta_\varepsilon(\overline{x}))$ implies that

$$\begin{aligned} & \left\langle d, \left(g^\circ - \nabla\theta_\varepsilon(\overline{x}) + \eta\beta\right) - g^\circ + \nabla\theta_\varepsilon(\overline{x}) \right\rangle \\ & = -\frac{\eta}{\left\| g' - \left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right)\right\|} \left\langle d, g' - \left(g^\circ - \nabla\theta_\varepsilon(\overline{x})\right) \right\rangle \\ & \le 0. \end{aligned} \qquad (14)$$

Then the result (i) is proved.

(ii) Taking the affine hull of $\partial P_\varepsilon(\bar{x}) = \nabla\theta_\varepsilon(\bar{x}) + \partial\varphi(\bar{x})$, we obtain

$$\text{aff } \partial P_\varepsilon(\bar{x}) = \nabla\theta_\varepsilon(\bar{x}) + \text{aff } \partial\varphi(\bar{x}). \quad (15)$$

Hence, the subspace which is parallel to the affine hull of $\partial P_\varepsilon(\bar{x})$ is also parallel to the affine hull of $\partial\varphi(\bar{x})$; that is, $\mathcal{V}_2 = \text{lin}(\partial\varphi(\bar{x}) - g')$, where $g' \in \partial\varphi(\bar{x})$ is arbitrary. Moreover, by the definition of $\partial\varphi(\bar{x})$, we have

$$\partial\varphi(\bar{x}) = \pi\text{conv}\left\{\nabla f_j(\bar{x})\right\}_{j\in J(\bar{x})} = \pi\sum_{j=0}^{m}\alpha_j\nabla f_j(\bar{x}), \quad (16)$$

where $\alpha_j > 0$ and $\sum_{j=0}^{m}\alpha_j = 1$. Let $\alpha_0 = 1$ and $\alpha_j = 0$, $j \neq 0$, then $\pi\nabla f_0(\bar{x}) \in \partial\varphi(\bar{x})$, and we can obtain

$$\begin{aligned}\mathcal{V}_2 &= \text{lin}\left\{\pi\sum_{j=0}^{m}\alpha_j\nabla f_j(\bar{x}) - \pi\nabla f_0(\bar{x})\right\}_{j\in J(\bar{x})} \\ &= \text{lin}\left\{\nabla f_j(\bar{x}) - \nabla f_0(\bar{x})\right\}_{j\in J(\bar{x})}.\end{aligned} \quad (17)$$

Then result (ii) is proved.

(iii) By the property of $P'_\varepsilon(\bar{x},\cdot)$ and the definition of $\mathcal{U}_1$, we have $\varphi'(\bar{x},d) = -\varphi'(\bar{x},-d)$. By the convexity of $\varphi(x)$, we have

$$\begin{aligned}\mathcal{U}_1 &= \left\{d \in R^n : \varphi'(\bar{x},d) = -\varphi'(\bar{x},-d)\right\} \\ &= \left\{d \in R^n : \max_{g'\in\partial\varphi(\bar{x})}\left\langle g',d\right\rangle = \min_{g'\in\partial\varphi(\bar{x})}\left\langle g',d\right\rangle\right\}.\end{aligned} \quad (18)$$

Let $d \in \mathcal{U}_1$, $\forall g' \in \partial\varphi(\bar{x})$, take $g'' \in \partial\varphi(\bar{x})$, and, by (18), we have $\langle d, g' - g''\rangle = \langle d, g'\rangle - \langle d, g''\rangle = 0$. It implies that, $\forall\beta' \in \mathcal{V}_2$, $\langle d,\beta'\rangle = 0$; that is, $d \in \mathcal{U}_2$. Hence, $\mathcal{U}_1 \subseteq \mathcal{U}_2$.

Let $d \in \mathcal{U}_2$, we have $\langle d, g'\rangle = \langle d, g''\rangle$ for all $g' \in \partial\varphi(\bar{x})$ and $g'' \in \partial\varphi(\bar{x})$. By the assumption $g^\circ \in \text{ri}\partial P_\varepsilon(\bar{x})$, we have $g^\circ - \nabla\theta_\varepsilon(\bar{x}) \in \text{ri}\partial\varphi(\bar{x}) \subseteq \partial\varphi(\bar{x})$. Then $\langle d, g'\rangle = \langle d, g''\rangle = \langle d, g^\circ - \nabla\theta_\varepsilon(\bar{x})\rangle$. By (i), we have $d \in \mathcal{U}_3$. Hence, $\mathcal{U}_2 \subseteq \mathcal{U}_3$.

By (i) and (18), we obtain $\mathcal{U}_3 \subseteq \mathcal{U}_1$. The proof of (iii) is completed. □

The solution of problem (P) is not only based on the $\mathcal{U}\mathcal{V}$-space decomposition of $P_\varepsilon(x)$ but also based on the study of the $\mathcal{U}$-Lagrangian of $P_\varepsilon(x)$, which will be shown next.

3. The $\mathcal{U}$-Lagrangian of $P_\varepsilon(x)$

Let $\nabla\theta_\varepsilon(\bar{x}) \coloneqq \tilde{g} = \tilde{g}_{\mathcal{U}} \oplus \tilde{g}_{\mathcal{V}}$, let $H_2 \coloneqq \nabla^2\theta_\varepsilon(\bar{x})$ be a positive semidefinite matrix, and let $\widehat{U}$ be a basis matrix for $\mathcal{U}$. $\forall g' \in \partial\varphi(\bar{x})$, we define the $\mathcal{U}$-Lagrange function of $P_\varepsilon(\bar{x})$ as follows:

$$\begin{aligned}L_{\mathcal{U},P_\varepsilon}(u,\bar{g}) &= \inf\left\{\varphi(\bar{x} + u \oplus v) - \left\langle g'_{\mathcal{V}}, v\right\rangle_{\mathcal{V}}\right\} \\ &\quad + \left(\theta_\varepsilon(\bar{x}) + \langle\tilde{g}_{\mathcal{U}}, u\rangle_{\mathcal{U}} + \frac{1}{2}\left\langle\widehat{U}^\top H_2\widehat{U}u, u\right\rangle_{\mathcal{U}}\right).\end{aligned} \quad (19)$$

Associated with (19) we have the set of minimizers

$$W(u) \coloneqq \text{Arg}\min_{v\in\mathcal{V}}\left\{\varphi(\bar{x} + u \oplus v) - \left\langle g'_{\mathcal{V}}, v\right\rangle_{\mathcal{V}}\right\}. \quad (20)$$

In the following paragraphs, a series of theorems and corollaries will be given to specify the property of $L_{\mathcal{U},P_\varepsilon}$ and the expansions of $P_\varepsilon(x)$.

Theorem 3. *By the definition of $L_{\mathcal{U},P_\varepsilon}$, we have the following conclusions:*

(i) *$L_{\mathcal{U},P_\varepsilon}(u,\bar{g})$ is a proper convex function.*

(ii) *A minimum point $w \in W(u)$ in (20) is characterized by the existence of some $g \in \partial P_\varepsilon(\bar{x})$ such that $\bar{g}_{\mathcal{V}} = g'_{\mathcal{V}} + \tilde{g}_{\mathcal{V}}$, where $g' \in \partial\varphi(\bar{x})$, and $g' = g'_{\mathcal{U}} \oplus g'_{\mathcal{V}}$.*

(iii) *In particular, $0 \in W(0)$ and $L_{\mathcal{U},P_\varepsilon}(0,\bar{g}) = P_\varepsilon(\bar{x})$.*

(iv) *If $g' \in ri\partial\varphi(\bar{x})$, then $W(u)$ is nonempty for each $u \in \mathcal{U}$ and $W(0) = \{0\}$.*

Theorem 4. *Let u satisfy $W(u) \neq \emptyset$. Then, $\forall w \in W(u)$, the subdifferential of $L_{\mathcal{U},P_\varepsilon}$ at this u has the expression*

$$\begin{aligned}&\partial L_{\mathcal{U},P_\varepsilon}(u,\bar{g}) \\ &\quad = \left\{g_{\mathcal{U}} : g_{\mathcal{U}} \oplus \left(g'_{\mathcal{V}} + \tilde{g}_{\mathcal{V}}\right) \in \partial P_\varepsilon(\bar{x} + u \oplus w)\right\}.\end{aligned} \quad (21)$$

In particular, $L_{\mathcal{U},P_\varepsilon}$ is differentiable at 0, with $\nabla L_{\mathcal{U},P_\varepsilon}(0,\bar{g}) = g'_{\mathcal{U}} + \tilde{g}_{\mathcal{U}}$.

Corollary 5. *If $g' \in ri\partial\varphi(\bar{x})$, then $W(u) = o(\|u\|_{\mathcal{U}})$.*

Theorem 6. *Let u satisfy $W(u) \neq \emptyset$, then, $\forall w \in W(u)$, $g' \in \partial\varphi(\bar{x})$, $\tilde{g} = \nabla\theta_\varepsilon(\bar{x})$, and we have*

$$\begin{aligned}P_\varepsilon(\bar{x} + u \oplus w) &= P_\varepsilon(\bar{x}) + \left\langle g' + \tilde{g}, u \oplus w\right\rangle_{\mathcal{U}} \\ &\quad + o\left(\|u\|_{\mathcal{U}}\right).\end{aligned} \quad (22)$$

Theorem 7. *Assume the function $\bar{L}_{\mathcal{U},P_\varepsilon}(u,\bar{g}) = \inf_{v\in\mathcal{V}}\{\varphi(\bar{x} + u \oplus v) - \langle g'_{\mathcal{V}}, v\rangle_{\mathcal{V}}\}$ has a generalized Hessian H_1 at $u = 0$ and $g' \in ri\partial\varphi(\bar{x})$. For $u \in \mathcal{U}$ and $x \in u \oplus W(u)$, it holds that*

$$\begin{aligned}P_\varepsilon(\bar{x} + x) &= P_\varepsilon(\bar{x}) + \left\langle g' + \tilde{g}, x\right\rangle \\ &\quad + \frac{1}{2}\left\langle\left(H_1 + \widehat{U}^\top H_2\widehat{U}\right)u, u\right\rangle_{\mathcal{U}} \\ &\quad + o\left(\|u\|^2_{\mathcal{U}}\right).\end{aligned} \quad (23)$$

The proofs of the above theorems and corollary are similar to [7] and here we ignore the details of them.

Based on the study of $\mathcal{U}\mathcal{V}$-space decomposition theory and the $\mathcal{U}$-Lagrangian of $P_\varepsilon(x)$, the $\mathcal{U}\mathcal{V}$-algorithm which can solve the problem (P_2) will be addressed in the next section.

4. The $\mathcal{U}\mathcal{V}$-Method

Depending on the $\mathcal{U}\mathcal{V}$-theory mentioned above, the constrained minimization problem of maximum eigenvalue function has been converted into the convex minimization problem which can be solved by the $\mathcal{U}\mathcal{V}$-algorithm in [8]. Hence, we apply the $\mathcal{U}\mathcal{V}$-algorithm in [8] and do some appropriate modifications for solving the problem (P_2).

In this section, some definitions and two quadratic programming problems will be denoted for easy understanding.

Given a tolerance $\sigma \in (0, 1/2]$, a prox-parameter $\mu > 0$, and a prox-center $x \in R^n$, to find σ-approximation of $P_\mu(x)$, our bundle subroutine accumulates information from the candidates $\{y_i\}_{i\in\mathbf{B}}$, where $\mathbf{B} := \{j : y_j = x\}$.

Definition 8. Let $x, y_i \in R^n$, $i \in \mathbf{B}$, $g_i \in \partial P_\varepsilon(x)$, and the linearization error is defined by

$$e_i := e(x, y_i) = P_\varepsilon(x) - P_\varepsilon(y_i) - g_i^\top (x - y_i). \tag{24}$$

Definition 9. Given a positive scalar parameter μ, the proximal point function depending on $P_\varepsilon(x)$ is defined by

$$p_\mu(x) := \operatorname*{argmin}_{p\in R^n} \left\{ P_\varepsilon(p) + \frac{1}{2}\mu \|p - x\|^2 \right\} \quad \text{for } x \in R^n. \tag{25}$$

The first quadratic programming subproblem $(\chi - QP)$ has the following form and properties; see [9]. The problem

$$(\chi - QP)\ \min\left\{ r + \frac{1}{2}\mu\|p - x\|^2 : (r, p) \in R^{1+n}, r \geq P_\varepsilon(x) - e_i + g_i^\top (p - x),\ \forall i \in \mathbf{B},\ \mu > 0 \right\} \tag{26}$$

has a dual problem

$$\min\left\{ \frac{1}{2\mu}\left\|\sum_{i\in\mathbf{B}} \alpha_i g_i\right\|^2 + \sum_{i\in\mathbf{B}} \alpha_i e_i : \alpha_i \geq 0,\ i \in \mathbf{B},\ \sum_{i\in\mathbf{B}} \alpha_i = 1 \right\}. \tag{27}$$

Their respective solutions, denoted by $(\widetilde{r}, \widetilde{p})$ and $\widetilde{\alpha} = (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_{|\mathbf{B}|})$, satisfy

$$\begin{aligned} \widetilde{r} &= \phi(\widetilde{p}), \\ \widetilde{p} &= x - \frac{1}{\mu}\widetilde{g}, \\ \widetilde{g} &:= \sum_{i\in\mathbf{B}} \widetilde{\alpha}_i g_i, \end{aligned} \tag{28}$$

where $\phi(x) := P_\varepsilon(x) - \sum_{i\in\mathbf{B}} \widetilde{\alpha}_i e_i - 1/\mu|\widetilde{g}|^2$ and $\widetilde{\alpha}_i = 0$, for all $i \in \mathbf{B}$, satisfies $\widetilde{r} > P_\varepsilon(x) - e_i + g_i^\top(\widetilde{p} - x)$. For convenience, in the sequel we denote the output of these calculations by

$$(\widetilde{r}, \widetilde{p}) = \chi - QP\left(\mu, x, \{(e_i, g_i)\}_{i\in\mathbf{B}}\right). \tag{29}$$

The second quadratic programming subproblem is

$$(\gamma - QP)\ \min\left\{ r + \frac{1}{2}\|p - x\|^2 : (r, p) \in R^{1+n}, r \geq g_i^\top (p - x),\ \forall i \in \widetilde{\mathbf{B}} \right\}, \tag{30}$$

where $\widetilde{\mathbf{B}} := \{i \in \mathbf{B} : \widetilde{r} = P_\varepsilon(x) - e_i + g_i^\top(\widetilde{p} - x)\} \cup \{i_+\}$, $y_{i_+} = \widetilde{p}$, and $g_{i_+} \in \partial P_\varepsilon(\widetilde{p})$. The above problem has a dual problem without linearization error terms:

$$\min\left\{ \frac{1}{2}\left\|\sum_{i\in\widetilde{\mathbf{B}}} \alpha_i g_i\right\|^2 : \alpha_i > 0,\ i \in \widetilde{\mathbf{B}},\ \sum_{i\in\widetilde{\mathbf{B}}} \alpha_i = 1 \right\}. \tag{31}$$

Similar to (28), the respective solutions, denoted by $(\overline{r}, \overline{p})$ and $\overline{\alpha}$, satisfy

$$\begin{aligned} \overline{p} - x &= -\widetilde{s}, \\ \widetilde{s} &:= \sum_{i\in\widetilde{\mathbf{B}}} \overline{\alpha}_i g_i. \end{aligned} \tag{32}$$

Since the need of the algorithm, the solution of the problem $(\gamma - QP)$ will be applied to get the matrix $\widetilde{U}$. Firstly, define an active index by $\widetilde{\mathbf{B}}_{\text{act}} := \{i \in \widetilde{\mathbf{B}} : \overline{r} = \overline{g}_i(\overline{p} - x)\}$. Then, from (32), $\overline{r} = -g_i^\top \widetilde{s}$ for all $i \in \widetilde{\mathbf{B}}_{\text{act}}$, so

$$(g_i - g_l)^\top \widetilde{s} = 0 \tag{33}$$

for all such i and for a fixed $l \in \widetilde{\mathbf{B}}_{\text{act}}$. Define a full-column rank matrix $\widetilde{V}$ by choosing the largest number of indices i satisfying (33) such that the corresponding vectors $g_i - g_l$ are linearly independent and by letting these vectors be the columns of $\widetilde{V}$. Then let $\widetilde{U}$ be a matrix where columns form an orthogonal basis for the null-space of $\widetilde{V}^\top$ with $\widetilde{U} = I$ if $\widetilde{V}$ is vacuous.

For convenience, in the sequel we denote the output from these calculations by

$$(\widetilde{s}, \widetilde{U}) = \gamma - QP\left(\{g_i\}_{i\in\mathbf{B}}\right). \tag{34}$$

The algorithm depending on the above quadratic programming problems is given as follows.

Algorithm 10.

Step 0. Choose positive parameters ε, μ, and m with $m < 1$. Let $p_0 \in R^n$ and $g_0 \in \partial P_\varepsilon(p_0)$, respectively, be an initial point and subgradient. Also, let U_0 be a matrix with orthogonal n-dimensional columns estimating an optimal $\mathcal{U}$-basis. Set $s_0 = g_0$ and $k := 0$.

Step 1. Stop if $\|s_k\|^2 \le \varepsilon$.

Step 2. Choose an $n_k \times n_k$ positive definite matrix H_k, where n_k is the number of columns of U_k and U_k approximating a basis for $\mathcal{V}(\chi(u))^\top$. For $\overline{x}$ which is a minimizer of $P_\varepsilon(x)$, $\chi(u) = \overline{x} + u \oplus v(u)$, where $v := R^{\dim \mathcal{U}} \mapsto R^{\dim \mathcal{V}}$ is a C^2-function satisfying $\overline{V}v(u) \in W(u)$ for all $\overline{g} \in \mathrm{ri}\partial P_\varepsilon(\overline{x})$. $\overline{V}$ is a basis matrix of $\mathcal{V}$ and H_k is the approximation of $\nabla^2 L_{\mathcal{U},P_\varepsilon}(u,0)$.

Step 3. Compute an approximate $\mathcal{U}$-*Newton* step by solving the linear system

$$H_k \Delta u = -U_k^\top s_k \quad \text{for } \Delta u = \Delta u_k \in R^{n_k} \tag{35}$$

and set $x_{k+1}^c := p_k + U_k \Delta u_k = p_k - U_k H_k^{-1} U_k^\top s_k$.

Step 4. Choose $\mu_{k+1} \ge \underline{\mu}$, $\sigma_{k+1} \in (0, 1/2]$, initialize $\mathbf{B}$, and run the following bundle subprocedure with $x = x_{k+1}^c$:

Compute recursively

$$\begin{aligned}(\widetilde{r}, \widetilde{p}) &= \chi - QP\left(\mu_{k+1}, x, \left\{(e_i, g_i)_{i\in\mathbf{B}}\right\}\right)\\ \widetilde{\varepsilon} &= P_\varepsilon\left(\widetilde{p}\right) - \widetilde{r}\\ \left(\widetilde{s}, \widetilde{U}\right) &= \gamma - QP\left(\{g_i\}_{i\in\widetilde{\mathbf{B}}}\right)\end{aligned} \tag{36}$$

until satisfying

$$\widetilde{\varepsilon} \le \frac{\sigma_{k+1}}{\mu_{k+1}} |\widetilde{s}|^2. \tag{37}$$

Then set

$$\left(\varepsilon_{k+1}^c, p_{k+1}^c, s_{k+1}^c, U_{k+1}^c\right) := \left(\widetilde{\varepsilon}, \widetilde{p}, \widetilde{s}, \widetilde{U}\right) \tag{38}$$

Step 5. If

$$P_\varepsilon\left(p_{k+1}^c\right) - P_\varepsilon\left(p_{k+1}\right) \le \frac{m}{2\mu_{k+1}} \left|s_{k+1}^c\right|^2, \tag{39}$$

and then declare a successful candidate and set

$$\begin{aligned}&\left(x_{k+1}, \varepsilon_{k+1}, p_{k+1}, s_{k+1}, U_{k+1}\right)\\ &\quad := \left(x_{k+1}^c, \varepsilon_{k+1}^c, p_{k+1}^c, s_{k+1}^c, U_{k+1}^c\right).\end{aligned} \tag{40}$$

Otherwise, execute a line search on the line determined by p_k and p_{k+1}^c to find x_{k+1} thereon satisfying $P_\varepsilon(x_{k+1}) \le P_\varepsilon(p_k)$; reinitialize $\mathbf{B}$ and return the above bundle subroutine, but with $x = x_{k+1}$, to find new values for $(\widetilde{\varepsilon}, \widetilde{p}, \widetilde{s}, \widetilde{U})$; then set $(\varepsilon_{k+1}, p_{k+1}, s_{k+1}, U_{k+1}) := (\widetilde{\varepsilon}, \widetilde{p}, \widetilde{s}, \widetilde{U})$.

Step 6. Replace k by $k + 1$ and go to stopping test.

Next, we will show the convergence of Algorithm 10. From here on, we assume that $\varepsilon = 0$ and that Algorithm 10 does not terminate. When the primal track at the initial point exists, firstly, it shows that if some execution of the bundle procedure in Algorithm 10 continues indefinitely, there is convergence to a minimizer of P_ε.

Theorem 11. *If the bundle procedure does not terminate, that is, if (37) never holds, then the sequence of $\widetilde{p}$-values converges to $p_\mu(x)$ and $p_\mu(x)$ minimizes $P_\varepsilon(x)$. If the procedure terminates with $\widetilde{s} = 0$, the corresponding $\widetilde{p}$ equals $p_\mu(x)$ and minimizes $P_\varepsilon(x)$. In both of these cases $p_\mu(x) - x \in \mathcal{V}(p_\mu(x))$.*

Proof. The recursion in the bundle subprocedure replacing $\mathbf{B}$ by $\widetilde{\mathbf{B}}$ satisfies conditions (4.7) to (4.9) in [9]. By Proposition 4.3 in [9], if this procedure does not terminate it generates an infinite sequence of $\widetilde{\varepsilon}$-values converging to zero. Since (37) does not hold, the sequence of $\|\widetilde{s}\|$-values also converges to 0. Thus, by lemma 5 in [8] and continuity of P_ε, we can get $P_\varepsilon(z) \ge P_\varepsilon(p_\mu(x))$ for all $z \in R^n$. The termination case with $\widetilde{\varepsilon} = 0$ follows in a similar manner, since (37) implies $\widetilde{\varepsilon} = 0$ in this case. In either case, by the minimality of $p_\mu(x)$, $0 \in \partial P_\varepsilon(p_\mu(x))$. From (3) in [8], $0 - \mu(x - p_\mu(x)) \in \mathcal{V}(p_\mu(x))$, and the final result follows, since $\mu \ne 0$. □

Next theorem shows minimizing convergence from any initial point without assuming the existence of a primal track. Here we assume that all executions of bundle procedure terminate.

Theorem 12. *Suppose that the algorithm sequence $\{\mu_k\}$ is bounded above by $\overline{\mu}$. Then the following hold:*

(i) *The sequence $\{P_\varepsilon(p_k)\}$ is decreasing and either $\{P_\varepsilon(p_k)\} \to -\infty$ or $\{\|s_k\|\}$ and $\{\varepsilon_k\}$ both converge to 0.*

(ii) *If P_ε is bounded from below, then any accumulation point of $\{p_k\}$ minimizes P_ε.*

Proof. (i) Since $\|s_k\| \ne 0$, whether or not p_{k+1}^c is successful candidate, the inequality

$$P_\varepsilon\left(p_{k+1}\right) - P_\varepsilon\left(p_k\right) \le \frac{m}{2\mu_{k+1}} \left\|s_{k+1}\right\|^2 \tag{41}$$

holds. Equation (41) implies that $\{P_\varepsilon(p_k)\}$ is decreasing. Suppose $\{P_\varepsilon(p_k)\} \nrightarrow -\infty$. Then summing (41) over k and using the fact that $m/(2\mu_k) \ge m/(2\overline{\mu})$ for all k imply that $\{\|s_k\|\} \to 0$. From Lemma 5 in [8] and (37) with $(\mu, \sigma, \widetilde{\varepsilon}, \widetilde{p}, \widetilde{s}) = (\mu_k, \sigma_k, \varepsilon_k, s_k)$, we have

$$P_\varepsilon\left(p_k\right) + s_k^\top\left(z - p_k\right) - \varepsilon_k \le P_\varepsilon(z) \quad \forall z \in R^n, \tag{42}$$

$$\varepsilon_k \le \frac{\sigma_k}{\mu_k} \left\|s_k\right\|^2. \tag{43}$$

Then (43) with $\sigma_k \le 1/2$ and $\mu_k \ge \underline{\mu} > 0$ implies that $\varepsilon_k \to 0$ which establishes (i).

(ii) Now suppose P_ε is bounded below and $\overline{p}$ is any accumulation point of $\{P_k\}$. Then, because $\|s_k\|$ and $\{\varepsilon_k\}$ converge to 0 by item (i), (42) together with the continuity of P_ε implies that $P_\varepsilon(\overline{p}) \le P_\varepsilon(z)$ for all $z \in R^n$ and (ii) is proved. □

In order to obtain convergence of the whole sequence p_k, we need the concept of a strong minimizer.

Definition 13. We say that $\overline{x}$ is a strong minimizer of P_ε if $0 \in \mathrm{ri}\partial P_\varepsilon(\overline{x})$ and the corresponding $\mathcal{U}$-Lagrangian $L_{\mathcal{U},P_\varepsilon}(u,0)$ has a Hessian at $u = 0$ that is positive definite.

Corollary 14. *Suppose that* $\overline{x}$ *is a strong minimizer of* P_ε*, as in Definition 13, and that the algorithm sequence* $\{\mu_k\}$ *is bounded above by* $\overline{\mu}$*. Then* $\{p_k\}$ *converges to* $\overline{x}$*. If, in addition, the sequence* $\{H_k^{-1}\}$ *is bounded, then* $\{x_{k+1}^c\}$ *and* $\{x_k\}$ *converge to* $\overline{x}$ *and* $\{s_{k+1}^c\}$ *converges to* $0 \in R^n$*.*

Proof. The proofs will be finished when $\widetilde{p}$, $\widetilde{s}$, and P_ε take the place of $\widehat{p}$, $\widehat{s}$, and f, respectively. □

5. Conclusions

The principal result is that we present the $\mathscr{U}\mathscr{V}$-algorithm for solving the constrained minimization problem of maximum eigenvalue functions. The innovative point is converting the constrained problem into the approximate unconstrained problem. By using the smooth convex approximation of maximum eigenvalue function, the latter problem can be solved by $\mathscr{U}\mathscr{V}$-algorithm. Although this method is based on some assumptions, it enriches the ways to deal with the constrained minimization problem of maximum eigenvalue functions.

Additional Points

Wei Wang (1960–) is a Professor in School of Mathematics of Liaoning Normal University.

Competing Interests

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 11671184, no. 11171138, and no. 11671183).

References

[1] P. Apkarian, D. Noll, J.-B. Thevenet, and H. D. Tuan, "A spectral quadraric-SDP method with applications to fixed-order H_2 and H_∞ synthesis," *European Journal of Control*, vol. 10, no. 6, pp. 527–538, 2004.

[2] A. R. Diaaz and N. Kikuchi, "Solutions to shape and topology eigenvalue optimization problems using a homogenization method," *International Journal for Numerical Methods in Engineering*, vol. 35, no. 7, pp. 1487–1502, 1992.

[3] K. Mombaur, *Stability optimization of open-loop controlled walking robots [Ph.D. thesis]*, Universität Heidelberg, Heidelberg, Germany, 2001.

[4] C. Helmberg and F. Oustry, "Bundle methods to minimize the maximum eigenvalue function," in *Handbook of semidefinite programming*, vol. 27 of *Internat. Ser. Oper. Res. Management Sci.*, pp. 307–337, Kluwer Acad. Publ., Boston, Mass, USA, 2000.

[5] F. Oustry, "The U-Lagrangian of the maximum eigenvalue function," *SIAM Journal on Optimization*, vol. 9, no. 2, pp. 526–549, 1999.

[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, Article ID PIPS5118271, pp. 253–270, 2004.

[7] W. Wang, L.-P. Pang, and Z.-Q. Xia, "The UV-decomposition on a class of D.C. functions and optimality conditions," *Acta Mathematicas Applicatas Sinica*, vol. 23, no. 1, pp. 29–38, 2007.

[8] R. Mifflin and C. Sagstizábal, "A UV-algorithm for convex minimization," *Mathematical Programming, Series B*, vol. 104, pp. 583–608, 2005.

[9] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagstizábal, *Numerical Optimization. Theoretical and Practical Aspects*, Universitext, Springer-Verlag, Berlin, Germany, 2003.