

SOLVING A MINIMIZATION PROBLEM FOR A CLASS OF CONSTRAINED MAXIMUM EIGENVALUE FUNCTION

Wei Wang^{1 §}, Miao Chen², Lingling Zhang³

¹School of Mathematics Liaoning Normal University
Liaoning, Dalian, 116029, P.R. CHINA

²School of Mathematics Liaoning Normal University
Liaoning, Dalian, 116029, P.R. CHINA

Abstract: Nonsmooth convex optimization problem is a class of important problems in operational research. Bundle methods are considered as one of the most efficient methods for solving nonsmooth optimization problems. The methods have already been applied to many practical problems. In this paper, using bundle method, the optimization problem that the sum of maximum eigenvalue function and general non-smooth convex function can be solved. Through approximation to the objective function, the proximal bundle method based on approximate model is given. We prove that the sequences generated by the algorithm converge to the optimal solution of the original problem. Finally, the algorithm is used to solve a class of constrained maximum eigenvalue function.

AMS Subject Classification: 15A18, 49J52, 52A41

Key Words: nonsmooth optimization, bundle method, maximum eigenvalue function

1. Introduction

Nonsmooth optimization problems arise in many fields of applications, for example, in economics (see [1]), mechanics (see [2]), engineering (see [3]) and optimal control (see [4]). They are generally difficult to solve. The methods

Received: August 28, 2013

© 2014 Academic Publications, Ltd.
url: www.acadpubl.eu

[§]Correspondence author

for nonsmooth optimization can be divided into two main classes: subgradient methods and bundle methods. In this paper, we focus on bundle methods, specifically on their approximate model.

Consider a class of problem as follows:

$$(P) \quad \min_{y \in R^m} \lambda_{\max}(A(y)) + g(y),$$

where $\lambda_{\max}(A(y))$ is a maximum eigenvalue function, A is a linear operator from R^m to S^n , and $g(y)$ is a nonsmooth convex function. C. Helmbert and F. Oustry (see [5]) used bundle method to solve a class of unconstrained maximum eigenvalue function. C. Sagastizabal and M. Solodov (see [6]) adopted bundle-filter method to deal with nonsmooth convex constrained optimization. Here, the optimization problem that a class of the sum of maximum eigenvalue function and general non-smooth convex function can be solved by using proximal bundle method. Furthermore, we extend the algorithm to a class of constrained maximum eigenvalue function.

In order to obtain a minimizer of (P) , we should solve the following subproblems. We consider a class of proximal bundle method for (P) . The sequence of stability centers $\{x^k\}$ is a subsequence of $\{y^k\}$, where $\{y^k\}$ is a sequence of sample points which is used to define an approximate model of the objective function

$$F(y) := \lambda_{\max}(A(y)) + g(y).$$

In order to generate candidate point y^k , we should construct an approximate model $\hat{F}_k(y)$ of the objective function $F(y)$.

The structure of this paper is as follows: Section 2 shows the approximate model of the objective function, which is under the condition of

$$\text{ri}(\text{dom}\lambda_{\max}(A(x^i))) \cap \text{ri}(\text{dom}g(x^i)) \neq \phi.$$

The proximal bundle method is derived in Section 3 and its convergence is studied in Section 4. In Section 5, a class of constrained maximum eigenvalue function optimization problem is solved by proximal bundle method.

Throughout the paper, $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ are the standard norm and inner product in Hilbert space.

2. The Approximate Model of the Objective Function

In this section, we will mainly study the approximate model of the objective function. Then, we introduce the subdifferential of the maximum eigenvalue function: $\lambda_{\max}(X)$, $X \in S^n$.

For the convenience of calculation, we will deal with the maximum eigenvalue function. Convexity is the important property enjoyed by maximum eigenvalue function. It is the support function of the compact convex set $C_n := \{V \in S^n : V \succeq 0, \text{tr}V = 1\}$ (see [7]), Then $\lambda_{\max}(X) = \max_{v \in R^n, |v|=1} v^T X v = \max_{V \in C_n} V \cdot X$, where “ \cdot ” is the standard scalar product in S^n . $\partial\lambda_{\max}(X)$ is the face of C_n exposed by X . Let r be the multiplicity of $\lambda_{\max}(X)$ and let Q be an $n \times r$ matrix whose columns form an orthonormal basis of the corresponding eigenspace. Then

$$\partial\lambda_{\max}(X) = \{QZQ^T : Z \in C_r\}.$$

we consider the maximum eigenvalue function form as $\lambda_{\max}(A(x))$, where A is a linear operator from R^m to $S^n : A(x) = A_0 + \mathcal{A}x$, Then

$$\begin{aligned} \partial[\lambda_{\max}(A(x))] &= \mathcal{A}^* \partial\lambda_{\max}(A(x)) \\ &= \mathcal{A}^* \{Q(A(x))ZQ(A(x))^T, Z \in C_r\}. \end{aligned}$$

It follows from this that

$$\partial\lambda_{\max}(A(x^i)) = \{\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T), Z \in C_r\}.$$

Suppose $\text{ri}(\text{dom}\lambda_{\max}(A(y))) \cap \text{ri}(\text{dom}g(y)) \neq \phi$. Then we construct the following approximate model for $F(y)$

$$\hat{F}_k(y) := \max_{i=1,2,\dots,k} \{F(x^i) + \langle m^i, y - x^i \rangle\},$$

where $m^i \in \partial F(x^i) = \partial[\lambda_{\max}(A(x^i)) + g(x^i)]$.

Choose $m^i = \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i$, where

$$\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) \in \partial\lambda_{\max}(A(x^i)), s^i \in \partial g(x^i).$$

By the condition of $\text{ri}(\text{dom}\lambda_{\max}(A(y))) \cap \text{ri}(\text{dom}g(y)) \neq \phi$, we have the result from the convex analysis in [8] that

$$\begin{aligned} \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i &\in \partial\lambda_{\max}(A(x^i)) + \partial g(x^i) \\ &= \partial[\lambda_{\max}(A(x^i)) + g(x^i)]. \end{aligned}$$

Set the terms e^i be the linearization errors at x^k ,

$$e^i := F(x^k) - F(x^i) - \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i, x^k - x^i \rangle.$$

With the notation, the approximate models have the form

$$\hat{F}_k(y) = F(x^k) + \max_{i=1,2,\dots,k} \{-e^i + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i, y - x^k \rangle\}.$$

In next section, the proximal bundle method algorithm based on approximate model will be given.

3. The Proximal Bundle Method Algorithm

Algorithm (the Proximal Bundle Method)

Step 0. Let $\varepsilon \geq 0$, $m \in (0, 1)$ be given parameters. Choose x^1 , call black box with $y = x^1$, we can obtain $F(x^1)$ and

$$\mathcal{A}^*(Q(A(x^1))ZQ(A(x^1))^T) + s^1 \in \partial(\lambda_{\max}(A(x^1)) + g(x^1)),$$

construct the model \hat{F}_1 , and let $k = 1$, $\delta_1 = \infty$.

Step 1. If $\delta_k \leq \varepsilon$, stop.

Step 2. Solve the quadratic program as follows:

$$(P_1) \quad \min_{y \in R^n} \hat{F}_k(y) + \frac{1}{2}\eta_k \|y - x^k\|^2,$$

the nominal decrease $\delta_{k+1} := F(x^k) - \hat{F}_k(y^{k+1}) - \frac{1}{2}\eta_k \|y^{k+1} - x^k\|^2$.

Step 3. Call black box with $y = y^{k+1}$, if $F(x^k) - F(y^{k+1}) \geq m\delta_{k+1}$, let $x^{k+1} = y^{k+1}$, we call it serious-step. Otherwise, let $x^{k+1} = x^k$, we call it null-step.

Step 4. Append y^{k+1} to bundle model, construct \hat{F}_{k+1} . Change $k = k + 1$, go to Step 1.

At Step 2, we can get the candidate point y^{k+1} by the dual problem of (P_1) . This can be guaranteed by the following theorem.

Theorem 1. *If y^{k+1} be the unique solution to (P_1) and assume $\eta_k > 0$. Then*

$$y^{k+1} = x^k - \frac{1}{\eta_k} \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i),$$

where $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_{np_k})$ is a solution to

$$(D_1) \quad \begin{cases} \min \frac{1}{2\eta_k} \left\| \sum_{i=1}^{np_k} \alpha_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i) \right\|^2 + \sum_{i=1}^{np_k} \alpha_i e^i \\ \text{s.t. } \alpha \in \Delta_k = \{ \alpha_i \in [0, 1], \sum_{i=1}^{np_k} \alpha_i = 1, i = 1, 2, \dots, np_k \} \end{cases}$$

In addition, the following relations hold:

$$(1) \quad \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i) \in \partial \hat{F}_k(y^{k+1});$$

$$(2) \delta_{k+1} = \varepsilon_k + \frac{1}{2\eta_k} \left\| \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i) \right\|^2,$$

$$\text{where } \varepsilon_k = \sum_{i=1}^{np_k} \bar{\alpha}_i e^i;$$

$$(3) \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i) \in \partial_{\varepsilon_k} F(x^k).$$

Proof. Write (P_1) as a QP with an extra scalar variable r . (P_1) is equivalent to

$$(P_2) \begin{cases} \min_{(y,r) \in R^m \times R} r + \frac{1}{2}\eta_k \|y - x^k\|^2 \\ \text{s.t. } F(x^k) - e^i + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i, y - x^k \rangle \leq r \\ i = 1, 2, \dots, np_k. \end{cases}$$

In view of strong convexity, the dual problem of (P_1) is equivalent to the following problem:

$$(D_1) \begin{cases} \min \frac{1}{2\eta_k} \left\| \sum_{i=1}^{np_k} \alpha_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i) \right\|^2 + \sum_{i=1}^{np_k} \alpha_i e^i \\ \text{s.t. } \alpha \in \Delta_k = \{ \alpha_i \in [0, 1], \sum_{i=1}^{np_k} \alpha_i = 1, i = 1, 2, \dots, np_k \} \end{cases}$$

and $y^{k+1} = x^k - \frac{1}{\eta_k} \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i)$ is the solution of (P_1) , then (1) holds.

Because there is no duality gap, the optimal value in (P_1) is equal to the dual optimal value in (P_1) , hence (2) holds.

The relation $F(y) \geq \hat{F}_k(y) \geq \hat{F}_k(y^{k+1})$ gives the desired result of (3). \square

As iterations go along, the number of elements in the bundle increases. When the size of the bundle becomes too big, it is necessary to compress it and clean the model. Let np_{\max} be the maximal size of the bundle, and np_k be its current size. The compression sub-algorithm to be appended at Step 4 is the following:

Step 4'. Let $n_a = \{i \leq np_k, \bar{\alpha}_i > 0\}$ be the cardinality of active indices.

If $n_a \leq np_{\max} - 1$, then delete all inactive couples from the bundle, set $n_{left} = n_a$, and define $np_{k+1} = n_{left} + 1$. Otherwise, delete all inactive couples from the bundle, and discard two or more couples $(\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i, e^i)$,

then compress the discarded couples into a single couple. $n_{left} = np_{max} - 2$ or $n_{left} < np_{max} - 2$, define $np_{k+1} = n_{left} + 2$.

If $n_a \neq np_{max} - 1$, append

$$(\mathcal{A}^*(Q(A(x^{np_{k+1}}))ZQ(A(x^{np_{k+1}}))^T) + s^{np_{k+1}}, e_{np_{k+1}})$$

to the bundle, with

$$e_{np_{k+1}} = \begin{cases} 0, & \text{if serious-step,} \\ F(x^k) - F(y^{k+1}) - \langle \mathcal{A}^*(Q(A(x^{np_{k+1}}))ZQ(A(x^{np_{k+1}}))^T) \\ + s^{np_{k+1}}, x^k - y^{k+1} \rangle, & \text{if null-step.} \end{cases}$$

Construct \hat{F}_{k+1} , let $k = k + 1$, go to Step 1.

Remark 1. When the algorithm reaches an iteration where the number np_k becomes too big, delete all inactive couples from the bundle. If the remaining couples are still too many, then synthesizes indispensable information of active bundle elements. Using the information of $(\sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i), \varepsilon_k)$, which is defined by Theorem 4, construct aggregate linearization

$$F_\alpha(y) := F(x^k) - \varepsilon_k + \langle \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i), y - x^k \rangle.$$

For $F_\alpha(y)$, it holds that:

- (1) $F_\alpha(y) = \hat{F}_k(y^{k+1}) + \langle \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i), y - y^{k+1} \rangle,$
- (2) For all $y \in R^n, \hat{F}_k(y) \geq F_\alpha(y).$

Remark 2. When the maximum capacity is reached, for instance, when $k = np_{max}$, suppose we decide to discard the elements x^1, x^2, \dots, x^t ($t < k$) from the bundle, and to append the aggregate couple. The resulting model will be

$$\hat{F}_{k+1}(y) = \max_{t+1 \leq i \leq k+1} \{ \max \{ F(x^i) + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i, y - y^i \rangle \}, F_\alpha(y) \},$$

Then

$$F_\alpha(y) \leq \hat{F}_{k+1}(y) \leq F(y);$$

$$\hat{F}_{k+1}(y) \geq F(x^{k+1}) + \langle \mathcal{A}^*(Q(A(x^{k+1}))ZQ(A(x^{k+1}))^T) + s^{k+1}, y - y^{k+1} \rangle.$$

4. Convergence Analysis

Next, we discuss the convergence result in two cases: (1) $\varepsilon > 0$; (2) $\varepsilon = 0$.

Case (1): When $\varepsilon > 0$, by the following Theorem 5 there is an index k_{last} for which $\delta_{k_{last}} \leq \varepsilon$ if (P) has minimizers. Therefore, $x^{k_{last}}$ is the minimizer.

Theorem 2. *Consider the algorithm and suppose it loops forever. Use the notation $F^* := \lim_{k \in K_s} F(x^k)$ and $F^* > -\infty$ (K_s is the set of indices k for which a new serious-step is done). Then*

$$0 \leq \sum_{k \in K_s} \delta_k \leq \frac{F_1 - F^* + \varepsilon}{m}$$

Proof. Note first that, since $\varepsilon \geq 0$, for the algorithm to loop forever the nominal decrease must satisfy $\delta_k > 0$ for all $k \in K_s$. Since the descent test is satisfied: $x^{k+1} = y^{k+1}$, then $F(x^k) - F(x^{k+1}) \geq 0$. Let k' be the index following k in K_s . Between k and k' the algorithm makes null-steps only: $x^{k+1} = x^{k+j}$, for all $j = 2, 3, \dots, k' - k$. The descent test at k' gives $F(x^{k+1}) - F(x^{k'+1}) \geq m\delta_{k'+1}$. Hence, for any $k' \in K_s$, $\varepsilon > 0$,

$$m \sum_{k \in K_s}^{k''} \delta_{k+1} \leq \sum_{k \in K_s}^{k''} F(x^k) - F(x^{k+1}) = F_{x^1} - F_{x^{k''+1}} \leq F_{x^1} - F^* + \varepsilon,$$

Now letting $k'' \rightarrow \infty$ gives the desired result. □

Case(2) : When $\varepsilon = 0$, if $\delta_{k+1} = 0$, by the result of Theorem 4, the algorithm will find a solution $x^{\tilde{k}}$ to (P). if $\delta_{k+1} > 0$, the algorithm loops indefinitely. In this case, there are two possibilities for the sequence of descent steps $\{x^k\}_{k \in K_s}$. Either it has infinitely many elements, or there is an iteration k_{last} where a last serious-step is done, i.e., $x^k = x^{k_{last}}$ for all $k \geq k_{last}$. We consider these two situations separately.

Theorem 3. *Suppose the algorithm generates infinitely many descent-steps x^k . Then either (P) has an empty solution set and $\{F(x^k)\} \rightarrow -\infty$, or (P) has minimizers, In this case, the following holds:*

- (1) Both $\{\delta_k\} \rightarrow 0$ and $\{\varepsilon_k\} \rightarrow 0$ as $k \in K_s, k \rightarrow -\infty$.
- (2) If for all $k \in K_s$, $0 < \eta_{k+1} \leq \eta_k$, then the sequence $\{x^k\}$ is bounded and converges to a minimizer of (P).

Proof. Note first that, since $\varepsilon = 0$ and the algorithm does not stop. We have that $\delta_{k+1} > 0$ for all $k \in K_s$. If (P) has no solution, $\{F(x^k)\}$ goes to $-\infty$.

To see item (2), first, we shall show that the sequence x^k is minimizing for (P). Since $0 < \eta_{k+1} \leq \eta_k$ and $y^{k+1} = x^{k+1}$, we can obtain $\langle x^{k+1} - x, \eta_{k+1}(x^{k+1} - x) \rangle \leq \langle x^{k+1} - x, \eta_k(x^{k+1} - x) \rangle$, we notice from

$$x^{k+1} = x^k - \frac{1}{\eta_k} \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i)))^T + s^i) \tag{1}$$

that

$$\begin{aligned} & \eta_k \langle x^{k+1} - x, x^{k+1} - x \rangle \\ = & \eta_k \|x^k - x\|^2 + \frac{1}{\eta_k} \left\| \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i)))^T + s^i) \right\|^2 \\ & - 2 \langle x^k - x, \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i)))^T + s^i) \rangle \end{aligned}$$

We bound the right hand side terms by using Theorem 4(2) and 4(3). Then we obtain the relation

$$\eta_{k+1} \langle x^{k+1} - x, x^{k+1} - x \rangle \leq \eta_k \langle x^k - x, x^k - x \rangle + 2(F(x) - F(x^k) + \delta_{k+1}). \tag{2}$$

It follows from this that the sequence x^k is minimizing for (P).

To see the sequence $\{x^k\}$ is bounded, we suppose \bar{x} is a solution to (P). Take in (2) $x = \bar{x}$, and sum over $k \in K_s$, we can obtain the desire result. \square

Theorem 4. *Suppose the algorithm generates a last serious-step $x^{k_{last}}$, followed by infinitely many null-steps, if $0 < \eta_{k+1} \leq \eta_k$, then the sequence $\{y^k\}$ converges to $x^{k_{last}}$ and $x^{k_{last}}$ is the minimizer of $F(x)$.*

Proof. For any $y \in R^m$, consider the function

$$M_k(y) = \hat{F}_k(y^{k+1}) + \frac{1}{2}\eta_k \|y - x^{k_{last}}\|^2 + \frac{1}{2}\eta_k \|y^{k+1} - y\|^2.$$

Since y^{k+1} is the solution of (P₁), So

$$M_k(y^{k+1}) \leq F(x^{k_{last}}), \forall k \geq k_{last}. \tag{3}$$

Furthermore, the equality in

$$\begin{aligned} \hat{F}_{k+1}(y) = & \max \left\{ \max_{t+1 \leq i \leq k+1} \{F(x^i) + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i)))^T \right. \\ & \left. + s^i, y - y^i \rangle\}, F_\alpha(y) \right\}, \end{aligned}$$

and the identity about $F_\alpha(y)$, give the relations

$$\hat{F}_{k+1}(y) \geq \hat{F}_k(y^{k+1}) + \left\langle \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i), y - y^{k+1} \right\rangle. \quad (4)$$

Using inequality (4) written for $y = y^{k+2}$, and we obtain that

$$\begin{aligned} & M_{k+1}(y^{k+2}) \\ \geq & M_k(y^{k+1}) - \frac{1}{2}\eta_k \|y^{k+1} - x^{k_{last}}\|^2 + \frac{1}{2}\eta_k \|y^{k+2} - x^{k_{last}}\|^2 \\ & + \left\langle \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i), y^{k+2} - y^{k+1} \right\rangle. \end{aligned}$$

By expanding the difference of squares we see that

$$\begin{aligned} & \|y^{k+2} - x^{k_{last}}\|^2 - \|y^{k+1} - x^{k_{last}}\|^2 \\ = & \|y^{k+2} - y^{k+1}\|^2 + 2\langle y^{k+2} - y^{k+1}, y^{k+1} - x^{k_{last}} \rangle, \end{aligned}$$

it follows from this that

$$M_{k+1}(y^{k+2}) \geq M_k(y^{k+1}) + \frac{1}{2}\eta_k \|y^{k+2} - y^{k+1}\|^2. \quad (5)$$

Since the increasing sequence $M_k(y^{k+1})$ is bounded from above by (3), it must converge.

We now show that the sequence $\{y^{k+1}\}$ is bounded, with $\{y^{k+1} - y^k\} \rightarrow 0$. Using the identity

$$\eta_k(x^{k_{last}} - y^{k+1}) = \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + s^i)$$

and the relation $M_k(y^{k+1}) = \hat{F}_k(y^{k+1}) + \frac{1}{2}\eta_k \|y^{k+1} - x^{k_{last}}\|^2$, we see that

$$\begin{aligned} M_k(y^{k+1}) + \frac{1}{2}\eta_k \|y^{k+1} - x^{k_{last}}\|^2 &= \hat{F}_k(y^{k+1}) + \eta_k \|y^{k+1} - x^{k_{last}}\|^2 \\ &= F_\alpha(x^{k_{last}}) \leq F(x^{k_{last}}). \end{aligned}$$

It follows from this that the sequence $\{y^{k+1}\}$ must be bounded. In addition, by (5) and passing to the limit, we conclude that $\{y^{k+1} - y^k\} \rightarrow 0$.

It is obvious from the definition of convexity that

$$\hat{F}_k(y^{k+1}) - F(y^k) \rightarrow 0, \quad (6)$$

From the bounded sequence $\{y^k\}$ extract a subsequence $\{y^{k_i}\}$, where $\{y^{k_i}\} \rightarrow \bar{y}$ as $i \rightarrow \infty$. Since $\{y^{k+1} - y^k\} \rightarrow 0$, the sequence $\{y^{k_i+1}\} \rightarrow \bar{y}$. Therefore

$$F(y^{k_i+1}) - \hat{F}_{k_i}(y^{k_i+1}) = F(y^{k_i+1}) - F(y^{k_i}) + F(y^{k_i}) - \hat{F}_{k_i}(y^{k_i+1}) \rightarrow 0$$

as $i \rightarrow \infty$ this implies that $\hat{F}_{k_i}(y^{k_i+1}) \rightarrow F(\bar{y})$ as $i \rightarrow \infty$. To show that $x^{k_{last}}$ minimizes (P) , recall that for all $k > k_{last}$, the serious-step is never satisfied. This means that $F(y^{k_i+1}) - F(x^{k_{last}}) \geq -m\delta_{k_i+1}$, so, we obtain

$$0 \leq (1 - m)\delta_{k_i+1} \leq F(y^{k_i+1}) - \hat{F}_{k_i}(y^{k_i+1}).$$

Passing to the limit as $i \rightarrow \infty$ and using (6) we conclude that $\delta_{k_i+1} \rightarrow 0$. Hence, it is obvious from the Theorem 4(2), (3) that $x^{k_{last}}$ minimizes (P) .

Finally, we show that \bar{y} is equal to $x^{k_{last}}$. Use the facts that $F(y) \geq F(x^{k_{last}})$ and that $\hat{F}_{k_i}(y) \leq F(y)$, then we have

$$F(\bar{y}) \geq \hat{F}_{k_i}(y^{k_i+1}) + \frac{1}{2}\eta_{k_{last}} \|y^{k_i+1} - x^{k_{last}}\|^2.$$

By (6), we obtain in the limit that

$$\begin{aligned} F(\bar{y}) &\geq \lim_{i \rightarrow \infty} (\hat{F}_{k_i}(y^{k_i+1}) + \frac{1}{2}\eta_{k_{last}} \|y^{k_i+1} - x^{k_{last}}\|^2) \\ &= F(\bar{y}) + \frac{1}{2}\eta_{k_{last}} \|\bar{y} - x^{k_{last}}\|^2, \end{aligned}$$

an inequality that is possible only if $\bar{y} = x^{k_{last}}$, and the proof is complete. \square

5. Bundle Method for Constrained Maximum Eigenvalue Function

Consider the problem as follows :

$$(\bar{P}) \quad \begin{cases} \min \lambda_{\max}(A(y)) \\ \text{s.t. } By = c \\ y \geq 0, \end{cases}$$

where $\lambda_{\max}(A(y))$ is the maximum eigenvalue function, A is a linear operator from R^m to S^n , $B \in R^{m \times m}$, Then (\bar{P}) is equivalent to

$$(\bar{P}_1) \quad \min_{y \in R^m} \lambda_{\max}(A(y)) + \delta_{\Omega}(y),$$

where $\delta_\Omega(y)$ denotes the indicator function on the set $\Omega = \{y \in R_+^m : By = c\}$. $\delta_\Omega(y)$ is a nonsmooth convex function, which satisfies the condition of the proximal bundle method. Therefore, we can apply the proximal bundle method to \bar{P} .

Let $G(y) = \lambda_{\max}(A(y)) + \delta_\Omega(y)$, the sub-problem that the approximate model of $G(y)$ as follows:

$$\hat{G}(y) = G(x^k) + \max_{i=1,2,\dots,k} \{-e^i + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i, y - x^k \rangle\}.$$

where

$$\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) \in \partial\lambda_{\max}(A(x^i)), c^i \in \partial\delta_\Omega(x^i) = N_\Omega(x^i).$$

Suppose $\text{ri}(\text{dom}\lambda_{\max}(A(y))) \cap \text{ri}(\text{dom}\delta_\Omega(y)) \neq \emptyset$, then

$$\hat{G}(y) = G(x^k) + \max_{i=1,2,\dots,k} \{-e^i + \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i, y - x^k \rangle\},$$

which let the terms e^i are the linearization errors at x^k ,

$$e^i := G(x^k) - G(x^i) - \langle \mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i, x^k - x^i \rangle.$$

Note that $N_\Omega(x^i) = \{v = v_1 + v_2 : v_1 \in R^m, v_2 \in B^T y, y \in R^m\}$, Indeed, as we shall show next:

$$\begin{aligned} \Omega &= \{y \in R_+^m : By = c\} \\ &= \{y \in R_+^m\} \cap \{y \in R^m : Bx - c = 0\}. \end{aligned}$$

Let $\Omega_1 = \{y \in R_+^m\}$, $\Omega_2 = \{y \in R^m : By - c = 0\}$, we have

$$\begin{aligned} N_\Omega(x^i) &= N_{\Omega_1 \cap \Omega_2}(x^i) = N_{\Omega_1}(x^i) + N_{\Omega_2}(x^i) \\ &= \{v \in R_-^m\} + \{v = B^T y, y \in R^m\}. \end{aligned}$$

Just as in the previous section, we can solve the problem (\bar{P}_1) .

Theorem 5. *If y^{k+1} be the unique solution to (\bar{P}_1) and assume $\eta_k > 0$. Then*

$$y^{k+1} = x^k - \frac{1}{\eta_k} \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i),$$

where $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \dots, \bar{\alpha}_{np_k})$ is a solution to

$$(D_1) \begin{cases} \min \frac{1}{2\eta_k} \left\| \sum_{i=1}^{np_k} \alpha_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i) \right\|^2 + \sum_{i=1}^{np_k} \alpha_i e^i \\ \alpha \in \Delta_k = \{\alpha_i \in [0, 1], \sum_{i=1}^{np_k} \alpha_i = 1, i = 1, 2, \dots, np_k\}. \end{cases}$$

In addition, the following relations hold:

- (1) $\sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i) \in \partial \hat{G}_k(y^{k+1}).$
- (2) $\delta_{k+1} = \varepsilon_k + \frac{1}{2\eta_k} \left\| \sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i) \right\|^2,$
 where we defined $\varepsilon_k = \sum_{i=1}^{np_k} \bar{\alpha}_i e^i.$
- (3) $\sum_{i=1}^{np_k} \bar{\alpha}_i (\mathcal{A}^*(Q(A(x^i))ZQ(A(x^i))^T) + c^i) \in \partial_{\varepsilon_k} G(x^k).$

Then, by algorithm it is easy to obtain the desired result.

References

- [1] J. Outrata, M. Kočvara and J. Zowe, Nonsmooth Approach to Optimization Problems With Equilibrium Constraints, *Theory, Applications and Numerical Results*, Kluwer Academic Publishers, Dordrecht (1998).
- [2] J.J. Moreau, P.D. Panagiotopoulos and G. Strang, Eds., *Topics in Nonsmooth Mechanics*, Birkhäuser Verlag, Basel (1988).
- [3] E.S. Mistakidis and G.E. Stavroulakis, Nonconvex Optimization in Mechanics, *Smooth and Nonsmooth Algorithms, Heuristics and Engineering Applications by the F.E.M.*, Kluwer Academic Publisher, Dordrecht (1998).
- [4] F.H. Clarke, Yu. S. Ledyaev, R.J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York(1998).
- [5] C. Helmberg, F. Oustry, *Bundle methods to minimize the maximum eigenvalue function*, Handbook of Semidefinite Programming, **27** (2000), 307-337.
- [6] C. Sagastizábal, M. Solodov, An infeasible bundle method for nonsmooth convex constrained optimization without a penalty function or filter, *SIAM J. Optimization*, **16** (2005), 146-169.
- [7] C. Lemaréchal, F. Oustry, Nonsmooth algorithms to solve semidefinite programs, *Society for Industrial and Applied Mathematics*, Philadelphia (2000), 57-77.
- [8] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey(1970).

