A NEW PROXIMAL CHEBYSHEV CENTER CUTTING PLANE ALGORITHM FOR NONSMOOTH OPTIMIZATION AND ITS CONVERGENCE

JIE SHEN∗
School of Mathematics
Liaoning Normal University
Dalian, 116029, China

JIAN LV AND FANG-FANG GUO
School of Finance
Zhejiang University of Finance and Economics
Hangzhou, 310018, China
and
School of Mathematical Sciences
Dalian University of Technology
Dalian, 116024, China

YA-LI GAO AND RUI ZHAO
School of Mathematics
Liaoning Normal University
Dalian, 116029, China

(Communicated by Liqun Qi)

Abstract. Motivated by the proximal-like bundle method [K. C. Kiwiel, Journal of Optimization Theory and Applications, 104(3) (2000), 589-603], we establish a new proximal Chebyshev center cutting plane algorithm for a type of nonsmooth optimization problems. At each step of the algorithm, a new optimality measure is investigated instead of the classical optimality measure. The convergence analysis shows that an \(\varepsilon\)-optimal solution can be obtained within \(O(1/\varepsilon^3)\) iterations. The numerical result is presented to show the validity of the conclusion and it shows that the method is competitive to the classical proximal-like bundle method.

2010 Mathematics Subject Classification. Primary: 90C25; Secondary: 65K05.
Key words and phrases. Nonsmooth optimization, proximal bundle method, subgradient, localization set, Chebyshev center.

The first author is supported by the National Natural Science Foundation of China under Project No. 11301246, No. 11671183, No. 11601061 and the Natural Science Foundation Plan Project of Liaoning Province No.20170540573, the Foundation of Educational Committee of Liaoning Province No.LF201783607 and the Fundamental Research Funds for the Central Universities of China No.DUT16LK07.
∗ Corresponding author: Jie Shen.
1. Introduction. Nonsmooth optimization problems (NSO) arise from many fields of applications, for example, in economics [18], mechanics [16], engineering [15] and optimal control [2]. Consider the unconstrained convex minimization problem

$$\min_{x} f(x)$$

s.t. $x \in \mathbb{R}^n$, (1)

where $f : \mathbb{R}^n \to \mathbb{R}$ is a nonsmooth closed proper convex function. We denote the optimal value of (1) by $f^*$ and the optimal solution set by $X^*$. The nonlinear conjugate gradient method is one of the effective algorithms for solving (1), some ideal results in recent years demonstrate its satisfactory performance under special conditions, and the search direction not only satisfies the sufficient descent condition but also belongs to a trust region, see [28, 29, 27, 7]. Proximal-like bundle method is another promising and efficient algorithm for nonsmooth optimization problems, its convergence rate can be very rapid when compared with the conjugate gradient method, and when it comes to the search direction, unlike nonlinear conjugate gradient method, it is less strict for accepting a candidate as a useful direction since it only concerns with the descent of the objective function. Proximal-like bundle methods [13, 17, 24, 20, 23, 21] approximate the objective function by a regularized cutting plane model which is the sum of a piecewise linear function and a quadratic function, and it has already been generalized to situations using closed convex functions with certain properties in place of a quadratic function. These methods can also be used to solve variational inequality problems, see [22, 30, 31, 25]. Based on identical ideas and techniques in [3, 1, 8, 9], the authors in [17] extend Elzinga-Moore cutting plane algorithm by enforcing the next trial point to be not far away from the previous ones, which removes the compactness assumption. Instead of lower approximations used in proximal bundle methods, the approach in [17] is based on the object regularizing translated functions of the objective function, and it can be viewed as a double regularization approach.

In this paper, motivated by the work [13] we present a proximal Chebychev center cutting plane algorithm (pc$^3$pa for short) and analyze its convergence from a new point of view which is quite different from the traditional ones for proximal-like bundle methods. Under the assumption that for each $z^i \in \mathbb{R}^n$, the function value $f(z^i)$ and one arbitrary subgradient $g^i \in \partial f(z^i)$ can be computed through an oracle, we focus on the estimation of the negative optimal value $w_k$ of subproblem of searching for the next trial point, and we find that $w_k$ decreases significantly after a null step and it may serve as a new optimality measure of current iterative point $x_k$. The following question is also answered: after how many iterations at most, an approximate solution with certain finite precision can be obtained and how the approximation accuracy depends on the iteration numbers. We refer the readers to [9, 11, 10, 12] for other discussions of similar efficiency estimations for subgradient projection methods, analytic center cutting plane methods and so on.

The paper is organized as follows: in Section 2, we propose a new pc$^3$pa algorithm and apply it to solving (1) by adjusting its update for proximity control parameters and eliminating the approximate stopping criterion. The convergence analysis for the proposed algorithm is presented in Section 3. Section 4 reports some numerical performance of our pc$^3$pa algorithm for solving some nondifferentiable problems. In Section 5 we make some conclusions and comparisons.

We denote the usual inner product and norm in $\mathbb{R}^n$ by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. The subdifferential of a convex function $f$ at $x$ is defined by $\partial f(x) = \{p \in \mathbb{R}^n :$
2. Proximal Chebychev center cutting plane algorithm. In this part, by eliminating approximate stopping criterion we present a new pc³pa algorithm with the update for proximity control parameters. The pc³pa algorithm proposed in our paper generates a sequence of iterative points \( \{x^k\} \) called Chebychev centers, and some trial points \( z^i \) are generated at the same time, we can evaluate the subgradient \( g^i \in \partial f(z^i) \) and the function value \( f(z^i) \) through an oracle as usual. Given current Chebychev center \( x^k \), the following outer approximation to the epigraph which is below \( f(x^k) \), i.e., the set

\[
\tilde{X}_{x^k,k} = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : r \leq f(x^k), f(z^i) + \langle g^i, x - z^i \rangle \leq r, i \in I^k \}
\]

is defined to be the localization set, where \( I^k = \{1, 2, \cdots, k\} \). Obviously, we have \( X^* \times \{x^*\} \subset \tilde{X}_{x^k,k} \). Therefore, the basic issue for solving (1) is how to choose the next iterative point \( x^{k+1} \) so as to shrink the localization set \( \tilde{X}_{x^k,k} \). By decreasing the upper bounds \( f(x^k) \) we find that the radii of the largest ball inside \( \tilde{X}_{x^k,k} \) shrink to zero and the Chebychev centers \( \{x^k\} \) of the largest ball inside \( \tilde{X}_{x^k,k} \) converge to the minimizer of \( f \) if any. The next Chebychev center can be determined by solving the problem

\[
\min_{x \in \mathbb{R}^n} \psi_{x^k}(x),
\]

where

\[
\psi_{x^k}(x) = \max \{ \frac{\langle g, x^k - z \rangle + f(z) - f(x^k)}{1 + \sqrt{1 + \|g\|^2}} \mid z \in \mathbb{R}^n, g \in \partial f(z) \}. \tag{4}
\]

The optimal value of (3) gives the negative value of the radius of the largest ball inside \( \tilde{X}_{x^k,k} \). The optimal solution of (3) is the next Chebychev center. Unfortunately, the minimization of \( \psi_{x^k} \) has no reason to be easy since computing the value of \( \psi_{x^k} \) at any point is already a difficult issue. However, with the trial points \( z^i, i \in I^k \), we can build the following simpler function

\[
\tilde{\psi}_{x^k,k}(x) = \max_{i \in I^k} \left\{ \frac{f(z^i) + \langle g^i, x - z^i \rangle - f(x^k)}{1 + \sqrt{1 + \|g^i\|^2}} \right\} \tag{5}
\]

to approximate \( \psi_{x^k} \). Therefore, computing the candidate Chebychev center of the localization set \( \tilde{X}_{x^k,k} \) amounts to solving

\[
\min_{x \in \mathbb{R}^n} \tilde{\psi}_{x^k,k}(x). \tag{6}
\]

The model function \( \tilde{\psi}_{x^k,k} \) approximates function \( \psi_{x^k} \) in the neighbourhood of current iterative point and this approximation is unlikely to be reliable when it is far away from current iterative point, it is reasonable to enforce the search for the next trial point not too far away from previous ones. By employing the idea of Moreau-Yosida regularization, the next candidate \( z^{k+1} \) is found by solving the following strongly convex quadratic programming associated with the proximal control parameter \( \mu^k \)

\[
\min_{x \in \mathbb{R}^n} \{ \tilde{\psi}_{x^k,k}(x) + \frac{\mu^k}{2} \|x - x^k\|^2 \}. \tag{7}
\]

Note that here we employ the main idea of proximal-like bundle methods for (3) which proceeds by minimizing the model function \( \tilde{\psi}_{x^k,k} \) and intends to use the
resulting solutions to improve the model function $\tilde{\psi}_{z,k}$ again. A quadratic regularization term is needed to avoid the solution to oscillate, it can make the approach more efficient. Obviously, problem (7) is equivalent to

$$\min \quad v + \frac{\mu_k}{2} \| x - x^k \|^2$$
$$\text{s.t.} \quad \frac{f(z^i) + \langle g^i, x - z^i \rangle - f(x^k)}{1 + \sqrt{1 + \| g^i \|^2}} \leq v, \; i \in I^k,$$  \hspace{1cm} (8)

where $v$ represents the negative value of the radius of the largest ball inside $\tilde{X}_{x,k}$. The optimization model in (7) can be found in many other science fields, such as [4, 6, 5, 14, 19, 26].

Let us introduce some useful notations which will be used in the sequel. For each $i \in I^k$, let $\alpha(x^k, z^i) = f(x^k) - [f(z^i) + \langle g^i, x^k - z^i \rangle](\geq 0)$ to be the linearization error between $z^i$ and $x^k$.

Let $\gamma_{g^i} = (1 + \sqrt{1 + \| g^i \|^2})^{-1}$ for any $g^i \in \mathbb{R}^n$. Let $g^i_\alpha = \gamma_{g^i} g^i$ and $\alpha_{i,k} = \gamma_{g^i} \alpha(x^k, z^i)$ be scaled subgradient and scaled linearization error. Problem (8) can be expressed with the notations above as follows

$$\min \quad v + \frac{\mu_k}{2} \| x - x^k \|^2$$
$$\text{s.t.} \quad \langle g^i_\alpha, x - x^k \rangle - \alpha_{i,k} \leq v, \; i \in I^k.$$ \hspace{1cm} (9)

If we define $\tilde{f}_k(x) = \max_{i \in I^k} \{ \langle g^i_\alpha, x - x^k \rangle - \alpha_{i,k} \}$, (9) is equivalent to

$$\min \quad \varphi_k(x) := \tilde{f}_k(x) + \frac{\mu_k}{2} \| x - x^k \|^2$$
$$\text{s.t.} \quad x \in \mathbb{R}^n.$$ \hspace{1cm} (10)

Define the linearization of the translated function $f_{x,k}(x) = f(x) - f(x^k)$ to be $f_i(x) = \langle g^i_\alpha, x - x^k \rangle - \alpha_{i,k}$, it is easy to know that

$$f_i(x) \leq \frac{1}{2} \| f(x) - f(x^k) \|, \; \forall x \in \mathbb{R}^n,$$ \hspace{1cm} (11)

since $\gamma_{g^i} \leq \frac{1}{2}$ and $g^i \in \partial f(z^i)$, we have

$$\bar{f}_k(x) = \max_{i \in I^k} \{ f_i(x) \} \leq \frac{1}{2} \| f(x) - f(x^k) \|, \; \forall x \in \mathbb{R}^n.$$ \hspace{1cm} (12)

The dual problem of (9) can be easily obtained

$$\min \left\{ \frac{1}{2\mu_k} \| \sum_{i=1}^k \lambda_i g^i_\alpha \|^2 + \sum_{i=1}^k \lambda_i \alpha_{i,k} : \lambda \in \Lambda_k \right\},$$ \hspace{1cm} (13)

where $\Lambda_k = \{ \lambda \in \mathbb{R}^k : \sum_{i=1}^k \lambda_i = 1, \lambda_i \geq 0 \}$ is the unit simplex of $\mathbb{R}^k$. Let $\lambda_i^k$ denote the optimal solution of (13) and define the aggregate scaled subgradient and aggregate scaled linearization error respectively by

$$g^k_a = \sum_{i=1}^k \lambda_i^k g^i_\alpha$$
$$\alpha^k_a = \sum_{i=1}^k \lambda_i^k \alpha_{i,k}.$$ \hspace{1cm} (14)

Based on optimality condition of (10): $0 \in \partial \varphi_k(z^{k+1})$, the optimal solution of (9) is given by

$$z^{k+1} = x^k - \frac{g^k_a}{\mu_k} \; \text{and} \; v^k = \bar{f}_k(z^{k+1}) = \tilde{\psi}_{z^{k+1}, k}(z^{k+1}) = - \frac{\| g^k_a \|^2}{\mu_k} - \alpha^k_a.$$ \hspace{1cm} (15)

Problem (9) appears to be the same type as the subproblem arising in proximal-like bundle methods, but here $g^i_\alpha$ and $\alpha_{i,k}$ are used in place of the “ordinary” subgradient
\(g^i\) and linearization \(\alpha(x^k, z^i)\), and \(v\) does not represent \(\dot{f}_k\)-values either since \(\psi_{x^k, k}\) is not a model for \(f\). Define

\[
\gamma_a^k = \left(\sum_{i=1}^{k} \lambda_i^k \gamma_{g_i}^k\right)^{-1},
\]

(16)
since \(\lambda_i^k \in \Lambda_k; \gamma_{g_i} \leq \frac{1}{2}\) for each \(i \in I_k\), we have \(0 < \sum_{i=1}^{k} \lambda_i^k \gamma_{g_i}^k \leq \frac{1}{2}\) and \(2 \leq \gamma_a^k < \infty\).

We have all the necessary ingredients to state our implementable algorithm.

**Proximal Chebyshev Center Cutting Plane Algorithm for (1) (pc\(^3\)pa):**

**Step 0:** Select the parameter \(0 < \kappa < 1\) and the proximity control parameter bounds \(0 < \mu_{\text{min}} \leq \mu_{\text{max}} < \infty\). Choose \(\mu^1 \in [\mu_{\text{min}}, \mu_{\text{max}}]\) and an initial point \(z^1 \in \mathbb{R}^n\). Compute \(f(z^1), g^1 \in \partial f(z^1)\). Set \(x^1 = z^1\) and \(l^1 = \{1\}, k = k(l) = 1, l = 1\). \([k(l+1) - 1\) denotes the iteration number of the \(l\)th descent step.]

**Step 1:** If \(g^k = 0\), terminate.

**Step 2:** Solve (13) to obtain \(\lambda_i^k, i \in I_k\). Compute \(\gamma^k_a, \alpha^k_a\) and \(\gamma^k_a\) by (14) and (16). Set

\[
z^{k+1} = x^k - g^k_a, \quad \sigma^k = \frac{\|\gamma^k_a\|^2}{\mu^k} + \alpha^k_a.
\]

(17)

**Step 3:** Compute \(f(z^{k+1})\) and \(g^{k+1} \in \partial f(z^{k+1})\). If \(f(z^{k+1}) \leq f(x^k) - 2\kappa\sigma^k\), set \(x^{k+1} = z^{k+1}, k(l+1) = k + 1\) and increase \(l\) by 1. Otherwise set \(x^{k+1} = x^k\).

**Step 4:** If \(x^{k+1} \neq x^k\), set \(\mu^{k+1} = \mu^k\), otherwise, choose \(\mu^{k+1} \in [\mu^k, \mu_{\text{max}}]\).

**Step 5:** Choose \(I^{k+1} \supset \{I^k \cup \{k+1\}\}\), where \(I^k = \{i \in I_k : \lambda_i^k \neq 0\}\). Increase \(k\) by 1 and go to Step 1.

**Remark 1.** By imitating the analysis in [17] we have the following convergence result: If there are infinitely many Chebyshev centers \(\{x^k\}_{k \in \mathbb{N}}\), then \(f(x^k) \rightarrow f^*\) as \(k \rightarrow \infty\). Furthermore, if \(X^*\) is nonempty, then the sequence \(\{\sigma^k\}_{k \in \mathbb{N}}\) (the radius of the largest ball inside \(X_{x^k, k}\)) tends to 0 and the sequence \(\{x^k\}_{k \in \mathbb{N}}\) converges to an optimal solution of problem (1) as \(k \rightarrow \infty\). For the case when the algorithm stops at some point \(x^{k_0}\), \((x^{k_0}\) denotes the last Chebyshev center generated by pc\(^3\)pa algorithm), it is shown that the sequence \(\{x^k\}_{k \in \mathbb{N}}\) tends to 0 as \(k \rightarrow \infty\) and the optimality of \(x^{k_0}\) is obtained.

**Remark 2.** We have more freedom in the way of choosing proximity control parameter \(\mu^k\). Since it controls the strength of the quadratic term in (9), its choice is a difficult task. Here we employ the technique in [18] to update \(\mu^k\), and other update techniques have been proposed in the literatures, for example, see [9].

3. **Convergence analysis.** The presented work in this section follows a line of investigation initiated in [13]. We expand and generalize the central idea in [13] to nonsmooth optimization problems based on the so-called localization sets and Chebyshev centers. Some techniques have to be adjusted to the new situations. We start this section by introducing several technical results from [13]. Define \(w_k\) to be the negative optimal value of subproblem (10)

\[
w_k = -\varphi_{x^k}(z^{k+1}) = -\min_{x \in \mathbb{R}^n} \varphi_{x^k}(x).
\]

(18)

Define the aggregate linearization of the translated function \(f_{x^k}(x) = f(x) - f(x^k)\) to be
The following conclusions hold:

\[ \tilde{f}_k(\cdot) = \tilde{f}_k(z^{k+1}) + \langle g^k_a, \cdot - z^{k+1} \rangle. \]  

(19)

From (12) and (19), it is easy to know that

\[ \frac{1}{2}(f(x) - f(x^k)) \geq \tilde{f}_k(x) = \tilde{f}_k(z^{k+1}) + \langle g^k_a, x - z^{k+1} \rangle, \quad \forall x \in \mathbb{R}^n, \]

where

\[ g^k_a = -\mu^k(z^{k+1} - x^k). \]

(21)

And hence we have

\[ \frac{1}{2}(f(x) - f(x^k)) \geq \tilde{f}_k(x) = \tilde{f}_k(z^{k+1}) + \langle g^k_a, x - z^{k+1} \rangle = \tilde{f}_k(x^k) + \langle g^k_a, x - x^k \rangle = \langle g^k_a, x - x^k \rangle - \varepsilon_k, \]

where

\[ \varepsilon_k = -\tilde{f}_k(x^k) \geq 0. \]

(23)

The following conclusion characterizes the relationships between \( v_k \) and \( w_k \).

**Lemma 3.1.** The following conclusions hold:

\[ -v_k = \mu^k ||z^{k+1} - x^k||^2 + \varepsilon_k = \frac{||g^k_a||^2}{\mu^k} + \varepsilon_k \geq 0. \]

(24)

\[ w_k = \frac{1}{2\mu^k} ||g^k_a||^2 + \varepsilon_k \geq 0, \quad v_k \leq w_k \leq \frac{v_k}{2}. \]

(25)

\[ f(x^k) - f(x) \leq 2 \sqrt{2\mu^k w_k} ||x - x^k|| + 2w_k, \quad \forall x \in \mathbb{R}^n. \]

(26)

**Proof.** By (15), (19), (21) and (23), we obtain

\[ -v_k = -\tilde{f}_k(z^{k+1}) = \mu^k ||z^{k+1} - x^k||^2 - \tilde{f}_k(x^k) = \frac{||g^k_a||^2}{\mu^k} + \varepsilon_k = \mu^k ||z^{k+1} - x^k||^2 + \varepsilon_k. \]

According to the definition of \( w_k \) and (24), it follows that

\[ w_k = -\varphi_k(z^{k+1}) = -\tilde{f}_k(z^{k+1}) - \frac{\mu^k}{2} ||z^{k+1} - x^k||^2 = -v_k - \frac{\mu^k}{2} ||z^{k+1} - x^k||^2 = \frac{||g^k_a||^2}{2\mu^k} + \varepsilon_k \geq 0, \]

\[ \frac{-v_k}{2} = \frac{||g^k_a||^2}{2\mu^k} + \frac{1}{2} \varepsilon_k \leq w_k \leq -v_k. \]

From (22) and Cauchy-Schwartz inequality

\[ f(x^k) - f(x) \leq -2\langle g^k_a, x - x^k \rangle + 2\varepsilon_k \leq 2||g^k_a|| \cdot ||x - x^k|| + 2\varepsilon_k, \]

and in light of \( 0 \leq \varepsilon_k \leq w_k, ||g^k_a||^2 = 2\mu^k (w_k - \varepsilon_k) \), we have

\[ f(x^k) - f(x) \leq 2 \sqrt{2\mu^k w_k} ||x - x^k|| + 2w_k. \]

According to the boundedness of sequence \( \{g^k_a\} \) and (27), we can derive a global optimality estimation which involves \( w_k \).

**Lemma 3.2.** The following conclusions hold: \( w_k \to 0 \) as \( k \to \infty \), and \( G < 0 \), \( D < \infty \),

\[ f(x^k) - f^* \leq 4 \max_k \sqrt{2\mu_{\max} w_k D}, w_k =: \delta_k \to 0 \quad (k \to \infty), \]

where

\[ G := \sup_k ||g^k_a||, \quad D := \sup_k \{d_{\Delta^*}(x^k)\}, \quad d_{\Delta^*}(x^k) := \min_{x \in \Delta^*} ||x^k - x||. \]  

(27)
The desired result (27) is obtained.

The result (27) points out that our convergence analysis can boil down to estimating how fast \( w_k \) decreases. Lemma 3.3 below discusses how to bound the decrease \( w_{k-1} - w_k \) via Lagrangian relaxation after a null step.

**Lemma 3.3.** If \( x^{k+1} = x^k \), then

\[
0 \leq w_{k+1} \leq w_k \left\{ 1 - \frac{1}{2} (1 - 2\kappa \gamma_{g^k}) \min \left\{ \frac{1}{\| \gamma_{g^{k+1}} g^{k+1} - g^k \|^2} \right\} \right\} < w_k. \tag{29}
\]

**Proof.** According to Step 4 and 5 of the proximal Chebychev center cutting plane algorithm, it follows from \( \tilde{f}_{k+1} = \max \{ f_k, f_{k+1} \} \) and \( \mu^{k+1} \geq \mu^k \) that \( \varphi_{k+1}(\cdot) \geq \max \{ f_k(\cdot), f_{k+1}(\cdot) \} + \frac{\mu_k}{2} \| \cdot - x^k \|^2 \). Therefore,

\[
\varphi_{k+1}(\cdot) \geq L(\beta, \cdot) := (1 - \beta) \tilde{f}_k(\cdot) + \beta f_{k+1}(\cdot) + \frac{\mu_k}{2} \| \cdot - x^k \|^2, \quad \forall \beta \in [0, 1], \tag{30}
\]

and

\[
\min \varphi_{k+1}(\cdot) \geq q(\beta) := L(\beta, x(\beta)), \quad x(\beta) = \arg \min_{\beta \in [0, 1]} L(\beta, \cdot).
\]

Using the relations (19) and (11), we obtain

\[
L(\beta, \cdot) = \tilde{f}_k(\cdot) - \beta \tilde{f}_k(\cdot) + \beta f_{k+1}(\cdot) + \frac{\mu_k}{2} \| \cdot - x^k \|^2
= \tilde{f}_k(z^{k+1}) + \beta [\gamma_{g^{k+1}} f(z^{k+1}) - \tilde{f}_k(z^{k+1})]
+ (g_a^k + \beta (\gamma_{g^{k+1}} g^{k+1} - g^k) \cdot - z^{k+1}) - \beta \gamma_{g^{k+1}} f(x^k) + \frac{\mu_k}{2} \| \cdot - x^k \|^2. \tag{31}
\]

Hence (30) and \( z^{k+1} = x^k - g_a^k / \mu_k \) yield

\[
x(\beta) = x^k - \frac{g_a^k + \beta (\gamma_{g^{k+1}} g^{k+1} - g^k)}{\mu_k} = z^{k+1} - \beta \gamma_{g^{k+1}} g^{k+1} - g_a^k, \tag{32}
\]

\[
q'(\beta) = \gamma_{g^{k+1}} f(z^{k+1}) - \tilde{f}_k(z^{k+1}) + \langle \gamma_{g^{k+1}} g^{k+1} - g^k, x(\beta) - z^{k+1} \rangle
= \gamma_{g^{k+1}} f(z^{k+1}) - \tilde{f}_k(z^{k+1}) - \beta \| \gamma_{g^{k+1}} g^{k+1} - g_a^k \|^2 / \mu_k - \gamma_{g^{k+1}} f(x^k). \tag{33}
\]

By (33) and Taylor formula, \( q(\beta) = q(0) + q'(0)\beta + q''(0)\beta^2 / 2 \) and

\[
q(0) = L(0, x(0)) = \tilde{f}_k(z^{k+1}) + \frac{\mu_k}{2} \| z^{k+1} - x^k \|^2 = v_k + \| g_a^k \|^2 / 2 \mu_k = -w_k \leq 0, \tag{34}
\]

\[
q'(0) = \gamma_{g^{k+1}} f(z^{k+1}) - \tilde{f}_k(z^{k+1}) - \gamma_{g^{k+1}} f(x^k)
= \gamma_{g^{k+1}} f(z^{k+1}) - v_k - \gamma_{g^{k+1}} f(x^k)
> \gamma_{g^{k+1}} 2\kappa v_k - v_k = (2\kappa \gamma_{g^{k+1}} - 1) v_k. \tag{35}
\]

\[
q''(0) = -\| \gamma_{g^{k+1}} g^{k+1} - g_a^k \|^2 / \mu_k. \tag{36}
\]
Therefore,
\[ \hat{\beta} := \arg \max_{\beta \in [0,1]} q(\beta) = \min \{ \hat{\beta}, 1 \}, \]
(37)
where
\[ \hat{\beta} := \arg \max_{\beta \in \mathbb{R}} q(\beta) = \begin{cases} -q'(0)/q''(0), & \text{if } q''(0) < 0, \\ +\infty, & \text{if } q''(0) = 0. \end{cases} \]

Now \( q(1) - q(0) = q'(0) + \frac{q''(0)}{2} > \frac{q'(0)}{2}, \) if \( \hat{\beta} > 1; q(\hat{\beta}) - q(0) = -\frac{q'(0)^2}{2q''(0)} \leq \frac{q'(0)}{2}, \) if \( \hat{\beta} \leq 1, \) so
\[ q(\hat{\beta}) - q(0) \geq \min \left\{ \frac{q'(0)}{2}, -\frac{q'(0)^2}{2q''(0)} \right\}. \]
(38)

Using relations (34)-(37) and \(-w_{k+1} = \min \varphi_{k+1} \geq q(\hat{\beta}),\)
\[ w_{k+1} \leq -q(\hat{\beta}) = -q(\hat{\beta}) + q(0) + w_k \\
= w_k - (q(\hat{\beta}) - q(0)) \\
\leq w_k - [\min \{ \frac{q'(0)}{2}, -\frac{q'(0)^2}{2q''(0)} \}] \\
= w_k - \min \left\{ \frac{(2\kappa \gamma_{k+1} - 1)w_k}{2}, \frac{q'(0)}{2q''(0)} \right\} \\
= w_k - \min \left\{ \frac{1 - 2\kappa \gamma_{k+1} |v_k|^2}{2}, \frac{(1 - 2\kappa \gamma_{k+1})^2 |v_k|^2}{2 |\gamma_{k+1}^2 + 1 - g_k|} \right\}. \]
The desired conclusion follows from \(|v_k| \geq w_k. \]

The following Lemma 3.4 still considers the null step cases.

**Lemma 3.4.** The following results hold:

(a) \( w_k \leq w_{k(l)} \leq \frac{||g_k^l||^2}{2\mu_{k(l)}} \leq \frac{G^2}{2\mu_{\min}}, \) for \( k \geq k(l); \)

(b) If \( x^{k+1} = x^k, \) then \( w_{k+1} \leq w_k (1 - \frac{c_l}{c_l}), \) where \( c_l = \frac{8G^2}{(1 - 2\kappa \gamma_{k+1})^2 \mu_{\min}} > 16w_{k(l)}; \)

(c) If \( x^{k+1} = x^k, \) then \( w_k \leq \frac{c}{k(k+1)16}, \) where \( c := \frac{8G^2}{(1 - \kappa)^2 \mu_{\min}} \geq \sup \{ c_l \}. \)

**Proof.** (a) If \( k > k(l), \) then \( w_k \leq w_{k(l)} \) by Lemma 3.3. If \( k = k(l), \) according to (12), we have
\[ \varphi_k(\cdot) \geq \psi_k(\cdot) := f_k(\cdot) + \frac{\mu_k || - x^k ||^2}{2}, \]
where \( f_k(\cdot) = \langle g_k, - x^k \rangle - \alpha_{k,k}. \) Hence \(-w_k = \min \varphi_k \geq \psi_k = -\frac{||g_k||^2}{2\mu_k}. \) The desired conclusion (a) follows from \( ||g_k^l|| \leq G \) and \( \mu_{k(l)} = \mu_k \geq \mu_{\min}. \)

(b) By Lemma 3.3 and the definition of \( c_l, \) we have \( \frac{w_k}{c_l} \leq \frac{w_{k(l)}}{c_l} \leq \frac{G^2/(2\mu_{k(l)})}{c_l} = \frac{(1 - 2\kappa \gamma_{k+1}^2)^2}{16} < \frac{1}{16}, \) whereas \( ||g_k - g_{k+1}^l||^2 \leq (||g_k|| + ||g_{k+1}||)^2 \leq 4G^2 \) since \( g_k \in \text{conv} \{ g_i^k \}_{i=1}^k \) and \( ||g_i^k|| \leq G. \) The facts \( \mu_k \geq \mu_{k(l)} \) and \( w_k \geq 0 \) yield that
\[ \frac{w_k}{||\gamma_{k+1}^2 g_{k+1}^k - g_k||^2} = \frac{(1 - 2\kappa \gamma_{k+1}^2)w_k \mu_k}{||g_k^k + g_{k+1}^k||^2} \geq \frac{2G^2}{(1 - 2\kappa \gamma_{k+1}^2)w_k \mu_{k(l)}} \]
\[ \leq \frac{2G^2}{(1 - 2\kappa \gamma_{k+1}^2)\mu_l} < 1, \]
therefore, if $x^{k+1} = x^k$, by (29)

$$w_{k+1} \leq w_k \left\{ 1 - \frac{1}{\xi} \left( 1 - 2\kappa \gamma_{g_k+1} \right) \min \left\{ 1, \frac{(1-2\kappa \gamma_{g_k+1}) w_k \mu^k}{\| \gamma_{g_k+1} g_k + g_k \|} \right\} \right\}$$

$$= w_k \left\{ 1 - \frac{1}{\xi} \left( 1 - 2\kappa \gamma_{g_k+1} \right) \left( \frac{(1-2\kappa \gamma_{g_k+1}) w_k \mu^k}{\| \gamma_{g_k+1} g_k + g_k \|} \right) \right\}$$

$$\leq w_k \left( 1 - \frac{m}{\epsilon c} \right).$$

(c) The conclusion can be obtained by imitating the proof of Lemma 3.4 in [13].

Now we are ready to state and prove our principle result.

**Theorem 3.5.** For $\varepsilon > 0$, let

$$k \geq \begin{cases} \frac{215 G^2 \mu_{\max}^2 D^4}{\kappa (1-2\kappa)^2 \mu_{\max}^2}, & \text{if } \varepsilon \leq 8\mu_{\max} D^2, \\ \frac{215 G^2 \mu_{\max}^2 D^4}{\kappa (1-2\kappa)^2 \mu_{\max}^2}, & \text{otherwise}, \end{cases}$$

(39)

then $f(x_k) - f^* < \varepsilon$, which means that for any acceptance tolerance $\varepsilon > 0$, the pc$^3$pa algorithm finds an $\varepsilon$-solution $x_k$ such that $f(x_k) - f^* < \varepsilon$ after at most $k = O(1/\varepsilon^3)$ iterations.

**Proof.** Suppose for contradiction that for current iteration $k = \tilde{k}, \tilde{l} = l$, $f(x_k) - f^* \geq \varepsilon$. Let $\tilde{k}(l) = k(l + 1) - 1$ be the iteration index of the $l$th descent step for $l = 1: \tilde{l}$. Define $K_l := \{ k(l) : \tilde{k}(l) \}$ for $l = 1: \tilde{l}$. Since $f(x_k)$ is nonincreasing, $f(x_k) \geq f(x(\tilde{k}(l))) \geq f^* + \delta_k$ for $k = 1: \tilde{k}(l)$. Hence, by $\delta_k \geq \varepsilon$ we have the equivalent expression of (27)

$$w_k \geq \min \left\{ \frac{\varepsilon^2}{32D^2 \mu_{\max}}, \frac{\varepsilon}{4} \right\} =: \varepsilon_w, \text{ for } k = 1: \tilde{k}(\tilde{l}).$$

(40)

Now, divide these indices $l$ into $\tilde{m}$ groups

$$L_m := \{ l \leq \tilde{l} : 2^{m-1} \varepsilon_w \leq w_{k(l)} \leq 2^m \varepsilon_w \}, (m = 1: \tilde{m}),$$

(41)

to make sure $L_m \neq \emptyset$. Suppose $L_m \neq \emptyset$, by $w_{k(l)} \leq 2^m \varepsilon_w$ and (27)

$$f(x(\tilde{k}(l))) - f^* \leq 4 \max \left\{ \frac{2\mu_{\max} 2^m \varepsilon_w D}{2^{\kappa m-1} \varepsilon_w} \right\}, \forall l \in L_m.$$  

(42)

Note that $f(x(\tilde{k}(l)))$ is decreasing and for each $l \in L_m$, $f(x(\tilde{k}(l))) - f(x(k(l)+1)) \geq -\kappa \delta_{k(l)}$, where $-\delta_{k(l)} \geq w_{k(l)} \geq 2^{m-1} \varepsilon_w$, so the $m$th group reduces $f$ at least $|L_m| \kappa 2^{m-1} \varepsilon_w$. This reduction can’t be greater than the initial reduction $f(x(\tilde{k}(l))) - f^*$, where $l = \min_{l \in L_m} l$, i.e., $|L_m| \kappa 2^{m-1} \varepsilon_w \leq f(x(\tilde{k}(l))) - f^*$. This inequality and (42) yield

$$|L_m| \leq \frac{4 \max \left\{ \frac{\sqrt{2\mu_{\max} 2^m \varepsilon_w} D}{2^{\kappa m-1} \varepsilon_w} \right\}}{\kappa 2^{m-1} \varepsilon_w} = 8 \max \left\{ \frac{\sqrt{2\mu_{\max} D}}{\sqrt{2^{\kappa m-1} \varepsilon_w}}, 1 \right\} / \kappa.$$  

(43)

By combining the definition of $K_l$, Lemma 3.4 (c) ($k = \tilde{k}(l)$) with (41), we have

$$|K_l| = \tilde{k}(l) - k(l) + 1 < \frac{c}{w_{k(l)}} \leq \frac{c}{2^{m-1} \varepsilon_w}, \quad \forall l \in L_m.$$  

(44)

For $K_m := \cup_{l \in L_m} K_l$, we get

$$|K_m| \leq |L_m| \max \left\{ |K_l| : l \in L_m \right\}$$

$$< 8 \max \left\{ \frac{\sqrt{2\mu_{\max} D}}{\sqrt{2^{\kappa m-1} \varepsilon_w}}, 1 \right\} \cdot \frac{c}{2^{m-1} \varepsilon_w}$$

$$= \left( \frac{16c}{\kappa} \right) \max \left\{ \frac{\sqrt{2\mu_{\max} D}}{\sqrt{2^{\kappa m-1} \varepsilon_w}}, 1 \right\} / 2^m.$$  

(45)
Since $\tilde{k} \leq \tilde{k}(\tilde{l}) \in \bigcup_{m=1}^{\tilde{m}} K_m$ and $\sum_{m=1}^{\tilde{m}} 2^{-m} < 1$, it follows from (45) that

$$\tilde{k} \leq \tilde{k}(\tilde{l}) \leq \bigcup_{m=1}^{\tilde{m}} |K_m| < (16c/\kappa) \max\left\{ \frac{\sqrt{2\mu_{\max}D}}{\varepsilon w^{3/2}}, 1/\varepsilon \right\}. \quad (46)$$

By (40) and $\varepsilon_w = (\varepsilon/4) \min\{\varepsilon/8, D_{\max}\}$, after a simple calculation the maximum of (46) equals to $28\mu_{\max}D^4/\varepsilon^3$ if $\varepsilon \leq 8\mu_{\max}D^2$, and equals to $4/\varepsilon$ otherwise. The right-hand side of (45) coincides with the expression of (39). It’s obvious that if $\tilde{k} = k$ satisfies (39), $f(x^k) - f^* < \varepsilon$, the proof is completed.

4. Numerical test. In this section, we report numerical results on the computational behaviour of the proposed $pc^3pa$ algorithm and illustrate the presented convergence results. All numerical experiments were implemented by using MATLAB R2012a and on a PC with 1.80GHz CPU. The quadratic programming solver is QuadProg.M, which is available in the Optimization Toolbox.

We first introduce a subclass of polynomial functions $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \ldots, n$,

$$h_i(x) = \sum_{j=1}^{n} x_j + (ix^2_i - 2x_i), \quad (47)$$

Then we define several test functions

$$f_1(x) := \sum_{i=1}^{n} \max\{h_i(x), 0\}, \quad (48)$$

$$f_2(x) := \sum_{i=1}^{n} \max\{h_i(x), 0\} + \frac{1}{2} |x|^2, \quad (49)$$

$$f_3(x) := \sum_{i=1}^{n} \max\{h_i(x), 0\} + \frac{1}{2} |x|. \quad (50)$$

It has been shown that all the test functions are convex. It can be obtained that $0 = \min f_k(x)$, and $\{0\} \subseteq \arg \min f_k(x)$ for $k = 1, 2, 3$. In our experiments we chose the values for all the parameters and the initial point as follows:

- the initial point $x_0 = (1, 1, \ldots, 1)^T$,
- the accuracy tolerance $\varepsilon_{g^k} = 10^{-6}$,
- the Armijo-like parameter $\kappa = 0.05$,
- the proximity control parameter bounds $\mu_{\min} = 1, \mu_{\max} = 50$.

The numerical results are listed in the following Tables 1–3 in which $n$ denotes the dimension of the problem, and and $\|g^k\|_{\text{final}}$ − the final value of $\|g^k\|$, $f_{\text{initial}}$ − the initial objective value, $f_{\text{final}}$ − the final objective value, $\text{Time}$ − the CPU time(sec.), $\text{NI}$ − the number of iterations.

We have tested three examples with different dimensions. Tables 1–3 show that by using $pc^3pa$ algorithm the final objective value is much smaller than $10^{-4}$ in most test problems, whose theoretical results are zero. Our limited computational experiments suggest the good performance and viability of our proposed method for a large class of problems.
Table 1. Test results obtained by $pc^3pa$ algorithm for $\min_{x \in \mathbb{R}^n} f_1(x)$.

| $n$ | $x^*$ | $f_{\text{final}}$ | $\|g^k\|_{\text{final}}$ | Ni    | Time   |
|-----|-------|---------------------|--------------------------|-------|--------|
| 6   | (-0.0000, 0.0000, -0.0000, 0.0000, 0.0000, -0.0000) | 0.0000 | 6.62e-07 | 13    | 1.1073 |
| 7   | 1.0e-05(0.12, -0.12, 0.42, -0.02, -0.02, 0.06) | 2.01e-05 | 3.02e-06 | 16    | 1.596  |
| 8   | (0.0000, 0.0000, 0.0000, -0.0001, -0.0005, 0.0001, 0.0001) | 4.31e-7 | 3.40e-07 | 22    | 1.9153 |
| 9   | 1.0e-04(0.00, 0.00, 0.00, 0.00, 0.01, 0.02, 0.01) | 1.30e-05 | 4.10e-07 | 36    | 2.5103 |
| 10  | 1.0e-04(0.06, -0.07, -0.09, -0.16, 0.22, 0.25, 0.03) | 3.26e-04 | 6.25e-07 | 36    | 1.8299 |

Table 2. Test results obtained by $pc^3pa$ algorithm for $\min_{x \in \mathbb{R}^n} f_2(x)$.

| $n$ | $x^*$ | $f_{\text{final}}$ | $\|g^k\|_{\text{final}}$ | Ni    | Time   |
|-----|-------|---------------------|--------------------------|-------|--------|
| 6   | (0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000) | 0.0000 | 1.06e-07 | 13    | 0.5323 |
| 7   | (-0.0000, 0.0000, 0.0000, -0.0000, -0.0000, 0.0000, 0.0000) | 0.0000 | 1.03e-07 | 10    | 1.1167 |
| 8   | (-0.0000, -0.0000, -0.0000, -0.0000, -0.0000, -0.0000, 0.0000) | 0.0000 | 2.08e-08 | 27    | 2.1463 |
| 9   | 1.0e-05(0.01, 0.01, 0.01, 0.01, 0.01, -0.02, -0.02) | 1.37e-8 | 2.91e-07 | 36    | 2.6105 |
| 10  | 1.0e-06(-0.01, 0.02, -0.02, 0.02, -0.02, 0.01, -0.02) | 4.32e-07 | 3.41e-07 | 39    | 3.3611 |

Table 3. Test results obtained by $pc^3pa$ algorithm for $\min_{x \in \mathbb{R}^n} f_3(x)$.

| $n$ | $x^*$ | $f_{\text{final}}$ | $\|g^k\|_{\text{final}}$ | Ni    | Time   |
|-----|-------|---------------------|--------------------------|-------|--------|
| 6   | (0.0000, -0.0000, -0.0000, 0.0000, 0.0000, 0.0000) | 0.0000 | 3.64e-07 | 16    | 1.0614 |
| 7   | (0.0000, -0.0000, -0.0000, 0.0000, 0.0000, 0.0000, 0.0000) | 0.0000 | 3.02e-07 | 19    | 1.9637 |
| 8   | (0.0000, -0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000) | 0.0000 | 4.01e-07 | 26    | 1.9437 |
| 9   | 1.0e-06(0.03, 0.03, 0.03, 0.01, -0.01, 0.00, 0.00) | 2.07e-7 | 2.14e-07 | 33    | 2.8025 |
| 10  | 1.0e-06(0.02, -0.02, 0.04, -0.02, 0.04, 0.02, -0.02) | 4.30e-08 | 7.19e-08 | 41    | 3.3061 |

5. Conclusions. The $pc^3pa$ algorithm in this paper is based on the so-called localization set $\tilde{X}_{x^k,k}$ and its Chebychev center which is the center of the largest ball.
inside it. This kind of algorithms can be viewed as a serious alternative to proximal-like bundle methods, therefore, its convergence analysis is especially important. We present a new optimality measure $\omega_k$, the negative optimal value of subproblem of searching for the next trial point, which can be computed easily with the process of iterations. Without additional boundedness assumptions, we conclude that for any $\varepsilon > 0$, after at most $O(1/\varepsilon^3)$ iterations, we obtain an $\varepsilon$- approximate solution with the help of $\omega_k$. Compared with the convergence result in [13] in which it only tells us that the sequence $\{x^k\}_{k \in \mathbb{N}}$ converges to some optimal solution, our result says exactly after how many iterations at most, what kind of approximate solution can be obtained. It is more convenient for users who would like to acquire an approximate solution with some kind of acceptance tolerance.

Acknowledgments. The authors are grateful to referees for their constructive comments which significantly helped improve the quality of the paper.

REFERENCES

[1] J. Baptiste, H. Urruty and C. Lemaréchal, Convex Analysis and Minimization Algorithms, Springer, Berlin, 1993.
[2] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, Nonsmooth Analysis and Control Theory, Springer, New York, 1998.
[3] R. Correa and C. Lemaréchal, Convergence of some algorithms for convex minimization, Math. Program., 62 (1993), 261–275.
[4] Z. Fu, K. Ren, J. Shu, X. Sun and F. Huang, Enabling personalized search over encrypted out-sourced data with efficiency improvement, IEEE Transactions on Parallel and Distributed Systems, (2015).
[5] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano and S. Li, Incremental support vector learning for ordinal regression, IEEE Transactions on Neural Networks and Learning Systems, 26 (2015), 1403–1416.
[6] B. Gu and V. S. Sheng, A robust regularization path algorithm for $\nu$-support vector classification, IEEE Transactions on Neural Networks and Learning Systems, 28 (2017), 1241–1248.
[7] J. Gu, X. Xiao and L. Zhang, A subgradient-based convex approximations method for DC programming and its applications, Journal of Industrial Management Optimization, 12 (2016), 1349–1366.
[8] K. C. Kiwiel, Methods of Descent for Nondifferentiable Optimization, Lectures Notes in Mathematics, Springer, Berlin, 1985.
[9] K. C. Kiwiel, Proximity control in bundle methods for convex nondifferentiable minimization, Math. Program., 46 (1990), 105–122.
[10] K. C. Kiwiel, Efficiency of the analytic center cutting plane method for convex minimization, SIAM J. Optim., 7 (1997), 336–346.
[11] K. C. Kiwiel, The efficiency of subgradient projection methods for convex optimization. Part 1: General level methods, SIAM Journal on Control and Optimization, 34 (1996), 660–676.
[12] K. C. Kiwiel, T. Larsson and P. O. Lindberg, The efficiency of ball step subgradient level methods for convex optimization, Mathematics of Operations Research, 24 (1999), 237–254.
[13] K. C. Kiwiel, Efficiency of proximal bundle methods, Journal of Optimization Theory and Applications, 104 (2000), 589–603.
[14] J. Li, X. Li, B. Yang and X. Sun, Segmentation-based image copy-move forgery detection scheme, IEEE Transactions on Information Forensics and Security, 10 (2015), 507–518.
[15] E. S. Mistakidis and G. E. Stavroulakis, Nonconvex Optimization in Mechanics. Smooth and Nonsmooth Algorithms, Heuristics and Engineering Applications, F.E.M. Kluwer Academic Publisher, Dordrecht, 1998.
[16] J. J. Moreau, P. D. Panagiotopoulos and G. Strang (Eds.), Topics in Nonsmooth Mechanics, Birkhäuser Verlag, Basel, 1988.
[17] A. Ouorou, A proximal cutting plane method using Chebychev center for nonsmooth convex optimization, Math. Program. Ser. A, 119 (2009), 239–271.
[18] 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.
[19] Z. Pan, Y. Zhang and S. Kwong, Efficient motion and disparity estimation optimization for low complexity multiview video coding, IEEE Transactions on Broadcasting, 61 (2015), 166–176.

[20] H. Schramm and J. Zowe, A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results, SIAM J. Optim., 2 (1992), 121–152.

[21] J. Shen, D. Li and L. Pang, A cutting plane and level stabilization bundle method with inexact data for minimizing nonsmooth nonconvex functions, Abstract and Applied Analysis, 2014 (2014), Article ID 192893, 6pp.

[22] J. Shen and L. Pang, An approximate bundle-type auxiliary problem method for generalized variational inequality, Mathematical and Computer Modeling, 48 (2008), 769–775.

[23] J. Shen, X. Liu, F. Guo and S. Wang, An approximate redistributed proximal bundle method with inexact data for minimizing nonsmooth nonconvex functions, Mathematical Problems in Engineering, 2015 (2015), Article ID 215310, 9pp.

[24] J. Shen, Z. Xia and L. Pang, A proximal bundle method with inexact data for convex nondifferentiable minimization, Nonlinear Analysis A: Theory, Methods and Applications, 66 (2007), 2016–2027.

[25] K. Wang, L. Xu and D. Han, A new parallel splitting descent method for structured variational inequalities, Journal of Industrial Management Optimization, 10 (2014), 461–476.

[26] Z. Xia, X. Wang, X. Sun and Q. Wang, A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data, IEEE Transactions on Parallel and Distributed Systems, 27 (2016), 340–352.

[27] G. Yuan, Z. Wei and G. Li, A modified Polak-Ribi`ere-Polyak conjugate gradient algorithm for nonsmooth convex programs, Journal of Computational and Applied Mathematics, 255 (2014), 86–96.

[28] G. Yuan, Z. Meng and Y. Li, A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations, Journal of Optimization Theory and Applications, 168 (2016), 129–152.

[29] G. Yuan and M. Zhang, A three-terms Polak-Ribi`ere-Polyak conjugate gradient algorithm for large-scale nonlinear equations, Journal of Computational and Applied Mathematics, 286 (2015), 186–195.

[30] J. Zhang, Y. Li and L. Zhang, On the coderivative of the solution mapping to a second-order cone constrained parametric variational inequality, Journal of Global Optimization, 61 (2015), 379–396.

[31] J. Zhang, S. Lin and L. Zhang, A log-exponential regularization method for a mathematical program with general vertical complementarity constraints, Journal of Industrial Management Optimization, 9 (2013), 561–577.

Received April 2016; revised August 2017.

E-mail address: tt010725@163.com
E-mail address: lvjian328@163.com
E-mail address: guoff@dlut.edu.cn
E-mail address: 704756727@qq.com
E-mail address: 1005514336@qq.com