Projected Extragradient Method for Finding Saddle Points of General Convex Programming*

Defeng Sun†

Abstract

A kind of practical extragradient method for finding saddle points of general convex programming by combining Korpelevich's extragradient concept and the method of inexact line searches is proposed. It is proved that the algorithm has global convergence property when the goal function and the constraint functions are continuously differentiable. Moreover, a sufficient and necessary condition on the existence of saddle point is obtained.

Keywords: convex programming; variational inequality problem; saddle points; extragradient; projection; inexact searches.

1 Introduction

Consider the following general convex programming problem:

$$\min\{f(x)\,|\,x \in X,\ g_i(x) \le 0,\ 1 \le i \le m\}, \tag{CP}$$

where $x \in \mathbb{R}^n$, $X \subset \mathbb{R}^n$ is a nonempty closed convex set, and the objective function $f : \mathbb{R}^n \to \mathbb{R}$ as well as the constraint functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, are all convex and continuously differentiable.

For convenience, let

$$G(x) = (g_1(x), \dots, g_m(x))^T, \quad D = \left\{(x, y) | x \in X, y \in \mathbb{R}^m_+\right\}.$$

Define the Lagrangian function of problem (CP) over the domain D as:

$$L(x, y) = f(x) + \langle G(x), y\rangle.$$

If there exists $(\bar{x}, \bar{y}) \in D$ such that

$$L(\bar{x}, y) \le L(\bar{x}, \bar{y}) \le L(x, \bar{y}), \quad \forall (x, y) \in D.$$

Then $(\bar{x}, \bar{y}) \in D$ is called a saddle point of $L(x, y)$ over D, and the set of all such saddle points is denoted by D^*. According to the Kuhn–Tucker theory [4] of convex programming, if $(x^*, y^*) \in D$ is a saddle point of $L(x, y)$, then x^* is an optimal solution to problem (CP). If the saddle point set D^* is nonempty (which may or may not be the case), then solving problem (CP) is equivalent to finding a saddle point of $L(x, y)$ over D.

*This is an English translation of a paper originally published in *Journal of Qufu Normal University*, 19(4):10–17, 1993. Received on September 12, 1991.

†Department of Mathematics, Nanjing University, Nanjing, Jiangsu, China.

Definition 1.1. *Define the operator*

$$T(x,y) = (\partial_x L(x,y),\ -\partial_y L(x,y)),$$

where $\partial_x L(x,y)$ and $\partial_y L(x,y)$ denote the partial derivatives of the Lagrangian $L(x,y)$ with respect to x and y, respectively. Under the assumption that $f(x)$ and $g_i(x)$ for $1 \le i \le m$ are continuously differentiable, the mapping $T(\cdot)$ is continuous on D. By the Kuhn–Tucker theorem [4], if $T(\cdot)$ is continuous, then $u^* = (x^*, y^*) \in D$ is a saddle point of L(x, y) over D if and only if $\langle T(u^*),\ u - u^* \rangle \ge 0, \quad \forall u = (x,y) \in D$, which is known as the variational inequality problem (VIP). In this way, finding a saddle point of $L(x,y)$ over D is transformed into solving a variational inequality problem (VIP).

Since $f(x)$ and $g_i(x)$ are convex and continuously differentiable for all $i = 1, \ldots, m$, it follows that

$$\langle T(u_2) - T(u_1),\ u_2 - u_1 \rangle \ge 0, \quad \forall u_1, u_2 \in D. \tag{1.1}$$

In particular, we have

$$\langle T(u),\ u - u^* \rangle \ge 0, \quad \forall u \in D, \tag{1.2}$$

where u^* is a saddle point of $L(x,y)$ over D.

Recently, He and He [6] proposed a class of algorithms for computing saddle points in convex programming problems in a formal framework. However, so far these methods have only been applied to quadratic programming or linear programming cases. Previously, the Soviet mathematician Korpelevich [8] proposed the extragradient method using an extrapolation technique to find saddle points. Assuming that $T(\cdot)$ is Lipschitz continuous on D and that $D^* \neq \emptyset$, Korpelevich proved the global convergence of the extragradient method. However, the Lipschitz continuity condition is too strong and often hard to satisfy in practice. Even if $T(\cdot)$ is Lipschitz continuous on D, finding a saddle point u^* is not always easy. For more discussion on this, see Polak [9]. Furthermore, no verifiable necessary and sufficient condition concerning the existence of saddle point is established in [8].

The main goal of this paper is to remove the Lipschitz continuity requirement on $T(\cdot)$ imposed in [8] by incorporating inexact techniques to propose a practical iterative method for computing saddle points, along with a verifiable sufficient and necessary condition for their existence during the algorithmic process.

2 Preliminaries and Several Lemmas

Let $Q \subset \mathbb{R}^{n+m}$ be an arbitrary nonempty closed and convex set. The projection of z onto Q, denoted by $P_Q(z)$, is defined as:

$$P_Q(z) = \arg\min_{\nu \in Q} \{\|\nu - z\|\}, \tag{2.1}$$

where $z \in \mathbb{R}^{n+m}$ and $\|\cdot\|$ denotes the L_2-norm. When Q has a relatively simple structure, for instance $Q = \mathbb{R}_+^{n+m}$, computing $P_Q(z)$ is straightforward. Since Q is nonempty closed and convex, $P_Q(z)$ is well-defined and single-valued.

We define the variational inequality problem VIP(T,Q) as follows: find $z^* \in Q$ such that

$$\langle T(z^*),\ z - z^* \rangle \ge 0, \quad \forall z \in Q. \tag{2.2}$$

Accordingly, the solution set of VIP(T,Q) is denoted by:

$$Q^* = \{z^* \mid (2.2) \text{ holds}\}. \tag{2.3}$$

Lemma 2.1. *Let P_Q be the projection operator onto a nonempty, closed and convex set $Q \subset \mathbb{R}^{m+n}$. Then the following properties hold:*

(i) $\|P_Q(v) - z\|^2 \leq \|v - z\|^2 - \|P_Q(v) - v\|^2, \quad \forall z \in Q,\ v \in \mathbb{R}^{m+n}$;

(ii) $\|P_Q(u) - P_Q(v)\| \leq \|u - v\|, \quad \forall u, v \in \mathbb{R}^{m+n}$.

Lemma 2.1 is easy to prove; see Zarantonello (1971, [10]) or Dunn (1981, [2]). Define

$$u_Q(\alpha) := P_Q(u - \alpha T(u)), \tag{2.4}$$

and for simplicity we shall use $u(\alpha)$ in place of $u_Q(\alpha)$ when no confusion arises. The definition of Q can be inferred from context.

The following result was initially proven by Gafni and Bertsekas (1984, [5]), with a simpler proof subsequently developed by Calamai and Moré (1987, [1]).

Lemma 2.2 ([5],[1])**.** *Let P_Q be the projection operator onto a nonempty, closed and convex set $Q \subset \mathbb{R}^{n+m}$. Given $u \in \mathbb{R}^{n+m}$ and $d \in \mathbb{R}^{n+m}$, define the function ψ by*

$$\psi(\alpha) = \frac{\|P_Q(u + \alpha d) - u\|}{\alpha}, \quad \alpha > 0. \tag{2.5}$$

Then $\psi(\alpha)$ is non-increasing in α; that is, if $0 < \alpha_1 \leq \alpha_2$, then $\psi(\alpha_1) \geq \psi(\alpha_2)$.

Lemma 2.3 ([3])**.** *Let P_Q be the projection operator onto a closed convex set $Q \subset \mathbb{R}^{n+m}$. Then $z^* \in Q$ is a solution of the variational inequality problem* VIP(T, Q) *if and only if*

$$z^* = P_Q\left(z^* - \alpha T(z^*)\right), \quad \textit{for some or any } \alpha > 0. \tag{2.6}$$

Lemma 2.4. *Suppose that $T(\cdot)$ is continuous on Q, and P_Q is the projection onto Q. Given a constant $\eta \in (0, 1)$, if $u \in Q \setminus Q^*$, then there exists a positive constant $\delta > 0$ such that*

$$\eta \|u(\alpha) - u\|^2 \geq \alpha^2 \|T(u(\alpha)) - T(u)\|^2, \quad \forall \alpha \in (0, \delta]. \tag{2.7}$$

Proof. By Lemma 2.3, since $u \notin Q^*$, it follows that $\|u(1) - u\| > 0$. Then by Lemma 2.2, we have

$$\frac{\|u(\alpha) - u\|}{\alpha} \geq \|u(1) - u\| > 0, \quad \forall \alpha \in (0, 1]. \tag{2.8}$$

Since $T(\cdot)$ is continuous on Q, and by Lemma 2.1 (ii), we have

$$\|u(\alpha) - u\| = \|P_Q(u - \alpha T(u)) - P_Q(u)\| \leq \alpha \|T(u)\|.$$

Hence $u(\alpha) \to u$ as $\alpha \to 0$, and therefore,

$$\|T(u(\alpha)) - T(u)\| \to 0, \quad \text{as } \alpha \to 0. \tag{2.9}$$

Combining (2.8) and (2.9), we conclude that there exists a constant $\delta > 0$ such that

$$\eta \cdot \frac{\|u(\alpha) - u\|^2}{\alpha^2} \geq \|T(u(\alpha)) - T(u)\|^2, \quad \forall \alpha \in (0, \delta],$$

which is equivalent to (2.7). □

Lemma 2.5 ([7])**.** *Let $Q \subset \mathbb{R}^{n+m}$ be a nonempty compact and convex set. Let P_Q denote the projection operator onto Q, and suppose that $T(\cdot)$ is continuous on Q. Then the solution set of the variational inequality problem* VIP(T, Q)*, denoted by Q^*, is nonempty, i.e.,*

$$Q^* \neq \emptyset.$$

3 Algorithm

We now introduce the following algorithm, Projected Extragradient Method (PEGM).

Algorithm 1 Projected Extragradient Method (PEGM)

Input: Constants $s \in (0, +\infty)$, $\beta \in (0, 1)$, $\eta \in (0, 1)$.

1: Choose arbitrary initial point $u^0 \in D$.

2: **for** $k = 0, 1, 2, \ldots$, while $u^k \notin D^*$ **do**

3: (i) Choose the smallest nonnegative integer m_k such that

$$\eta \|u(\alpha_k) - u^k\|^2 \geq \alpha_k^2 \|T(u(\alpha_k)) - T(u^k)\|^2, \tag{3.1}$$

where $\alpha_k = \beta^{m_k} s$, $u(\alpha_k)$ is defined by equation (2.4) and here $Q = D$.

4: (ii) Update

$$\bar{u}^k = P_D(u^k - \alpha_k T(u^k)) \text{ and } u^{k+1} = P_D(u^k - \alpha_k T(\bar{u}^k)). \tag{3.2}$$

5: **end for**

Remarks on the Algorithm

(i) In practical computations, the following stopping criterion can be used:

$$\frac{\|u^k(s) - u^k\|^2}{s^2} \leq \varepsilon^2,$$

where $\varepsilon > 0$ is a prescribed accuracy.

(ii) By Lemma 2.4, the condition in (3.1) of the algorithm is always satisfiable. In particular, as long as $u^k \notin D^*$, the integer m_k must be finite.

(iii) If $T(\cdot)$ is Lipschitz continuous on D with Lipschitz constant L, then it is easy to verify that

$$\alpha_k \geq \min\left\{\frac{\beta\sqrt{\eta}}{L},\ s\right\}.$$

(iv) In practice, the constant s may be replaced by a sequence $\{s_k\}$. As long as $\{s_k\}$ is bounded and bounded below by a positive number, the convergence properties of the algorithm are preserved.

4 Convergence Properties of the Algorithm

Theorem 4.1. *Suppose that $T(\cdot)$ is continuous on D, and the saddle point set $D^* \neq \emptyset$. Then, for any sequence $\{u^k\}$ generated by (3.1) and (3.2), there exists a subsequence $\{u^k\}$ such that $u^k \to \bar{u} \in D^*$ as $k \to \infty$.*

Proof. Let arbitrary $u^* \in D^*$. From inequality (i) in Lemma 2.1, we have:

$$\begin{aligned}
\|u^{k+1} - u^*\|^2 &\leq \|u^k - \alpha_k T(\bar{u}^k) - u^*\|^2 - \|u^k - \alpha_k T(\bar{u}^k) - u^{k+1}\|^2 \\
&= \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2 + 2\alpha_k \langle T(\bar{u}^k),\ u^* - u^{k+1}\rangle \\
&= \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2 + 2\alpha_k \langle T(\bar{u}^k),\ u^* - \bar{u}^k\rangle + 2\alpha_k \langle T(\bar{u}^k),\ \bar{u}^k - u^{k+1}\rangle \\
&\leq \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2 + 2\alpha_k \langle T(\bar{u}^k),\ \bar{u}^k - u^{k+1}\rangle.
\end{aligned} \tag{4.1}$$

The last inequality uses the result of inequality (1.2). On the other hand, we have

$$\begin{aligned}\|u^{k+1}-u^*\|^2 &\le \|u^k-u^*\|^2-\|u^k-\bar{u}^k\|^2-\|\bar{u}^k-u^{k+1}\|^2\\ &\quad -2\langle u^k-\bar{u}^k,\bar{u}^k-\ u^{k+1}\rangle+2\alpha_k\langle T(\bar{u}^k),\bar{u}^k-\ u^{k+1}\rangle\\ &=\|u^k-u^*\|^2-\|u^k-\bar{u}^k\|^2-\|\bar{u}^k-u^{k+1}\|^2\\ &\quad +2\langle u^k-\alpha_kT(\bar{u}^k)-\bar{u}^k,\ u^{k+1}-\bar{u}^k\rangle.\end{aligned} \tag{4.2}$$

Estimate the last term in (4.2). From Lemma 2.1 (i), we have

$$\langle u^k-\alpha_kT(u^k)-\bar{u}^k,\ u^{k+1}-\bar{u}^k\rangle\le 0.$$

In fact, in Lemma 2.1, let $v=u^k-\alpha_kT(u^k)$, $z=u^{k+1}$, and recall that $\bar{u}^k=P_D(v)$. Substituting into the inequality yields:

$$\begin{aligned}&2\langle u^k-\alpha_kT(\bar{u}^k)-\bar{u}^k,\ u^{k+1}-\bar{u}^k\rangle\\ &=2\langle u^k-\alpha_kT(u^k)-\bar{u}^k,\ u^{k+1}-\bar{u}^k\rangle+2\alpha_k\langle T(u^k)-T(\bar{u}^k),u^{k+1}-\bar{u}^k\rangle\\ &\le 2\alpha_k\langle T(u^k)-T(\bar{u}^k),\ u^{k+1}-\bar{u}^k\rangle\\ &\le \alpha_k^2\|T(u^k)-T(\bar{u}^k)\|^2+\|u^{k+1}-\bar{u}^k\|^2.\end{aligned} \tag{4.3}$$

The last inequality in (4.3) comes from using the Cauchy–Schwarz inequality, and combining (3.1), (4.2), and (4.3), we obtain:

$$\begin{aligned}\|u^{k+1}-u^*\|^2 &\le \|u^k-u^*\|^2-\|u^k-\bar{u}^k\|^2-\|\bar{u}^k-u^{k+1}\|^2\\ &\quad +\eta\|u^k-\bar{u}^k\|^2+\|u^{k+1}-\bar{u}^k\|^2\\ &=\|u^k-u^*\|^2-(1-\eta)\|u^k-\bar{u}^k\|^2.\end{aligned} \tag{4.4}$$

From (4.4), it follows that the sequence $\{\|u^k-u^*\|\}$ is non-increasing, and hence

$$\|u^k-\bar{u}^k\|\to 0 \quad \text{as } k\to\infty. \tag{4.5}$$

Since $\{u^k\}$ is bounded, there exists a subsequence $\{u^{k_i}\}$ and a point $\bar{u}\in D$ such that

$$u^{k_i}\to\bar{u}, \quad \text{as } k_i\to\infty. \tag{4.6}$$

Since $\{\alpha_{k_i}\}$ is also bounded, we may assume (without loss of generality, taking a subsequence if necessary) that

$$\alpha_{k_i}\to\alpha_0 \quad \text{as } k_i\to\infty.$$

To complete the proof, we now proceed by contradiction. Assume that $\bar{u}\notin D^*$. Then by Lemma 2.3 (i), there exists a positive constant $\delta_1>0$ such that

$$\frac{\|P_D\left(\bar{u}-sT(\bar{u})\right)-\bar{u}\|^2}{s^2}\ge\delta_1. \tag{4.7}$$

Since $u^{k_i}\to\bar{u}$, it follows that there exists an integer k_1 such that for all $k_i\ge k_1$, we have

$$\left\|\frac{P_D\left(u^{k_i}-sT(u^{k_i})\right)-u^{k_i}}{s^2}\right\|^2\ge\frac{\delta_1}{2}, \tag{4.8}$$

Using Lemma 2.2, we obtain

$$\frac{1}{\alpha_{k_i}^2}\left\|\bar{u}^{k_i}-u^{k_i}\right\|^2 \geq \frac{1}{s^2}\left\|P_D\left(u^{k_i}-sT(u^{k_i})\right)-u^{k_i}\right\|^2. \tag{4.9}$$

On the other hand, due to the continuity of $T(\cdot)$, there exists $\delta \in (0,s)$ such that

$$\|T(\bar{u}(\alpha))-T(\bar{u})\|^2 \leq \frac{\delta_1}{4}\eta, \quad \forall \alpha \in [0,\delta]. \tag{4.10}$$

Since $u^{k_i} \to \bar{u}$ and $T(\cdot)$ is continuous on D, there exists an integer k_2 such that for all $k_i \geq k_2$,

$$\|T(u^{k_i}(\alpha))-T(u^{k_i})\|^2 \leq \frac{\delta_1}{2}\eta, \quad \forall \alpha \in [0,\delta]. \tag{4.11}$$

Otherwise, there would exist a subsequence $\{j_i\} \subset \{k_i\}$ and corresponding $\{\alpha_{j_i}\} \subset [0,\delta]$ such that

$$\|T(u^{j_i}(\alpha_{j_i}))-T(u^{j_i})\|^2 > \frac{\delta_1}{2}\eta. \tag{4.12}$$

Without loss of generality, we may assume there exists $\bar{\alpha}$ such that $\alpha_{j_i} \to \bar{\alpha} \in [0,\delta]$. Taking the limit in (4.12) on both sides gives

$$\|T(\bar{u}(\bar{\alpha}))-T(\bar{u})\|^2 \geq \frac{\delta_1}{2}\eta,$$

which contradicts (4.10). Therefore, (4.11) holds. Taking $k=\max\{k_1,k_2\}$, from (4.8), (4.9), and (4.11), when $k_i \geq k$ we have

$$\frac{\|u^{k_i}(\alpha)-u^{k_i}\|^2}{\alpha^2} \geq \frac{1}{\eta}\|T(u^{k_i}(\alpha))-T(u^{k_i})\|^2, \quad \forall \alpha \in [0,\delta]. \tag{4.13}$$

Hence for all $k_i \geq k$, since $\alpha_{k_i} \geq \beta\delta$, it follows that $\alpha_0 > 0$. Because $||u^{k_i}-\bar{u}^{k_i}|| \to 0$ when $k_i \to \infty$, we obtain

$$\bar{u} = \lim_{i\to\infty} \bar{u}^{k_i}.$$

Since $\bar{u}^{k_i} = P_D(u^{k_i}-\alpha_{k_i}T(u^{k_i}))$, taking limits on both sides gives

$$\bar{u} = P_D(\bar{u}-\alpha_0 T(\bar{u})). \tag{4.14}$$

It follows from Lemma 2.3 that $\bar{u} \in D^*$, which contradicts the assumption $\bar{u} \notin D^*$. Hence we have shown $\bar{u} \in D^*$.

Since $\{\|u^k-\bar{u}\|\}$ is a monotone decreasing sequence and converges to $\bar{u}$, the full sequence $\{u^k\}$ also converges to $\bar{u}$. This completes the proof of Theorem 4.1. □

Remark 4.1. *When D is replaced by a general nonempty closed convex set, the conclusion of Theorem 4.1 remains valid.*

Theorem 4.1 proves the global convergence of the PEGM algorithm under the assumption $D^* \neq \emptyset$. The following result identifies a necessary and sufficient condition for the existence of a solution point, based on the algorithm's iteration behavior.

Theorem 4.2. *Suppose that $T(\cdot)$ is monotone and continuous on D, and let P_D denote the projection onto D. Then $D^* \neq \emptyset$ if and only if for some or any sequence $\{u^k\}$ generated by* (3.1) *and* (3.2)*, the sequence $\{u^k\}$ is bounded.*

Proof. **(Necessity).** By Theorem 4.1, if $D^* \neq \emptyset$, then the sequence $\{u^k\}$ generated by (3.1) and (3.2) converges. Hence, it is bounded.

(Sufficiency). Assume that for some sequence $\{u^k\}$ generated by (3.1) and (3.2), there exists a constant $r_1 \in (0, +\infty)$ such that

$$\|u^k\| \leq r_1, \quad \forall k.$$

By the definition of $\{\bar{u}^k\}$ and the continuity of $T(\cdot)$, there exist constants $r_2, r_3 \in (0, +\infty)$ such that

$$\|\bar{u}^k\| \leq r_2, \quad \forall k; \quad s\|T(u^k)\| \leq r_3, \quad s\|T(\bar{u}^k)\| \leq r_3.$$

Let $r = \max\{r_1, r_2, r_3\}$. Then, by Lemma 2.1 (ii), we have

$$\begin{aligned} &\|P_D(u^k - \alpha T(u^k))\| \leq 2r, \quad \forall \alpha \in [0, s]. \\ &\|P_D(u^k - \alpha T(\bar{u}^k))\| \leq 2r, \quad \forall \alpha \in [0, s]. \end{aligned} \tag{4.15}$$

Let ν be an arbitrary point in D, define

$$Y = \left\{ z \in \mathbb{R}^{n+m} \mid \|z\|_\infty \leq 2r + \|\nu\|_\infty \right\} \cap D,$$

where $\|\cdot\|_\infty$ is the $L_\infty - norm$. By the definition of Y and (4.15), it follows that

$$P_D(u^k - \alpha T(u^k)) \in Y, \quad P_D(u^k - \alpha T(\bar{u}^k)) \in Y, \quad \forall \alpha \in [0, s]. \tag{4.16}$$

From the definition of projection and the relationship (4.16), we have:

$$\begin{aligned} &P_Y(u^k - \alpha T(u^k)) = P_D(u^k - \alpha T(u^k)), \\ &P_Y(u^k - \alpha T(\bar{u}^k)) = P_D(u^k - \alpha T(\bar{u}^k)), \quad \forall \alpha \in [0, s]. \end{aligned} \tag{4.17}$$

By the PEGM algorithm and (4.17), we have:

$$\bar{u}^k = P_Y(u^k - \alpha T(u^k)), \quad u^{k+1} = P_Y(u^k - \alpha T(\bar{u}^k)). \tag{4.18}$$

That is, the sequence $\{u^k\}$ can be regarded as generated by PEGM for the variational inequality problem VIP(T, Y). Since Y is a nonempty compact convex set, Lemma 2.5 implies that the solution set Y^* of VIP(T, Y) is not empty. By Theorem 4.1 and Remark 4.1, there exists $z^* \in Y^*$ such that

$$u^k \to z^*, \quad \text{as } k \to \infty.$$

Since $\nu \in Y$, it follows that $\langle T(z^*), \nu - z^* \rangle \geq 0$. As z^* is a limit point of the sequence $\{u^k\}$, it is uniquely determined. Because ν is an arbitrary point in D, we have

$$\langle T(z^*), z - z^* \rangle \geq 0, \quad \forall z \in D. \tag{4.19}$$

This shows that z^* is a solution of the variational inequality problem VIP(T, D), hence $D^* \neq \emptyset$. □

Corollary 4.1. *Assume the conditions of Theorem 4.2 hold. If the point sequence* $\{u^k\}$ *generated by* (3.1) *and* (3.2) *is bounded, then there exists* $u^* \in D^*$ *such that* $u^k \to u^*$ *as* $k \to +\infty$.

By combining Theorems 4.1 and 4.2, we obtain the necessary and sufficient conditions for the nonemptiness of the saddle point set (that is, the set of saddle points of the Lagrange function $L(x, y)$ over D) of the original problem.

Corollary 4.2. *Suppose that $T(\cdot)$ is continuous on D. Then the saddle point set D^* of the convex programming problem (CP) is empty if and only if some or any sequence $\{u^k\}$ generated by* (3.1) *and* (3.2) *is unbounded.*

The above conclusion shows that when the saddle point set is nonempty, our algorithm can approximate a saddle point, yielding an approximate solution to problem (CP). Conversely, when the saddle point set is empty, the algorithm provides a natural stopping criterion. In practice, we may predefine a large threshold E, and terminate the algorithm when $\|u^k\| \geq E$.

References

[1] P. H. Calamai and J. J. More. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39:93–116, 1987.

[2] J. C. Dunn. Global and asymptotic convergence rate estimates for a class of projected gradient processes. *SIAM Journal on Control and Optimization*, 19:368–400, 1981.

[3] B. C. Eaves. On the basic theorem of complementarity. *Mathematical Programming*, 1:68–75, 1971.

[4] R. Fletcher. *Practical Method of Optimization, Constrained Optimization.* John Wiley and Sons, 1981.

[5] E. M. Gafni and D. P. Bertsekas. Two-metric projection methods for constrained optimization. *SIAM Journal on Control and Optimization*, 22:936–964, 1984.

[6] B. He and X. He. A class of saddle point methods for solving convex programming. *Journal of Computational Mathematics*, 1:51–57, 1991.

[7] S. Karamardian. Generalized complementarity problems. *Journal of Optimization Theory and Applications*, 8:161–167, 1971.

[8] G. M. Korpelevich. The extragradient method for finding saddle points and other problems. *MATEKON*, Summer:35–49, 1977.

[9] E. Polak. An historical survey of computational methods in optimal control. *SIAM Review*, 15:553–584, 1973.

[10] E. H. Zarantonello. Projections on convex sets in hilbert space and spectral theory. In E. H. Zarantonello, editor, *Contributions to Nonlinear Analysis.* Academic Press, New York, 1971.