

A NOTE ON THE GRADIENT PROJECTION METHOD WITH EXACT STEPSIZE RULE ^{*1)}

Naihua Xiu

(Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, China
Email: nhxiu@center.njtu.edu.cn)

Changyu Wang

(Operations Research Center, Qufu Teachers' University, Qufu 273165, China
Email: wcy0537@eyou.com)

Lingchen Kong

(Department of Applied Mathematics, Beijing Jiaotong University, Beijing 100044, China
Email: konglchen@126.com)

Abstract

In this paper, we give some convergence results on the gradient projection method with exact stepsize rule for solving the minimization problem with convex constraints. Especially, we show that if the objective function is convex and its gradient is Lipschitz continuous, then the whole sequence of iterations produced by this method with bounded exact stepsizes converges to a solution of the concerned problem.

Mathematics subject classification: 90C30, 65K05.

Key words: Gradient projection method, Exact stepsize rule, Full convergence.

1. Introduction

We consider the convexly constrained minimization problem

$$\min\{f(x) \mid x \in \Omega\}, \quad (1)$$

where $\Omega \subseteq R^n$ is a nonempty closed convex set, and the function $f(x)$ is continuously differentiable on Ω . We say that $x^* \in \Omega$ is a stationary point of problem (1) if it satisfies condition

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0, \quad \forall x \in \Omega, \quad (2)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of R^n . Let Ω^* denote the set of all stationary points of problem (1). Then if f is convex or pseudo-convex, Ω^* becomes the solution set of this problem.

The gradient projection method was first proposed by Goldstein [5] and Levitin and Polyak [9] for solving convexly constrained minimization problems. This method is regarded as an extension of the steepest descent or Cauchy algorithm for solving unconstrained optimization problems. It now has many variants in different setting, and supplies a prototype for various more advanced projection methods. Its iterative scheme is to update $x^k \in \Omega$ according to the formula

$$x^{k+1} = P[x^k - \alpha_k \nabla f(x^k)], \quad (3)$$

* Received February 22, 2005; final revised March 23, 2006; accepted June 29, 2006.

¹⁾ The research was in part supported by the National Natural Science Foundation of China (70471002, 10571106), and NCET040098.

where $P[\cdot]$ denotes the projection from R^n onto Ω , i.e.,

$$P[y] = \arg \min\{\|x - y\| \mid x \in \Omega\}, \quad y \in R^n,$$

∇f is the gradient of f , and $\alpha_k > 0$ is a judiciously chosen stepsize. If α_k is taken as a global minimizer (this assumes, of course, that such a minimum exists) of the subproblem

$$\min\{f(x^k(\alpha)) \mid \alpha \geq 0\}, \quad (4)$$

where $x(\alpha) := P[x - \alpha \nabla f(x)]$, then it is called the exact stepsize. If α_k is taken as the largest element in the set $\{\gamma l^0, \gamma l^1, \gamma l^2, \dots\}$ ($\gamma > 0, l \in (0, 1)$) satisfying

$$f(x^k(\alpha_k)) \leq f(x^k) + \mu \langle \nabla f(x^k), x^k(\alpha_k) - x^k \rangle, \quad \mu \in (0, 1), \quad (5)$$

then it is called the inexact Armijo stepsize.

The exact stepsize rule was first used by McCormick and Tapia [10] and further studied by Phelps [11, 12], Hager and Park [7], et al. Frequently, this rule is not used since it requires to evaluate the projections for all choices of $\alpha \geq 0$ while the inexact Armijo stepsize requires to only evaluate the finite projections at each iteration. However, as Hager and Park pointed out for some difficult optimization problems, where the structure of constraints is relatively simple, the exact stepsize rule is useful since it provides a mechanism for making a large step to escape from one valley of the cost function and move to another (possibly distant) valley with a smaller minimum cost. In [7], they gave an example in which the NP-hard graph partitioning problem can be formulated as a continuous quadratic programming problem whose constraints have a simple structure, and a practical procedure to evaluate the exact stepsize was given for the reformulated problem. This shows that it is quite necessary to study convergence of the gradient projection method with *exact stepsize rule*.

This paper is mainly concerned with the above issue. In Section 3 we characterize some sufficient conditions under which this method with bounded exact searches possesses some encouraging convergent properties. Especially, we obtain the result that if f is convex and ∇f is Lipschitz continuous on Ω , then the full sequence produced by the gradient projection method with bounded exact stepsizes is convergent to a solution of problem (1).

2. Some Lemmas

The analysis of the gradient projection method defined by (3) and (4) requires the following lemmas which are consequences of some basic properties of the projection operator.

Lemma 2.1. *Let P be the projection into Ω . Then for $x \in \Omega$,*

- (i) $\langle x(\alpha) - x + \alpha \nabla f(x), y - x(\alpha) \rangle \geq 0$, for all $y \in \Omega$ and $\alpha > 0$;
- (ii) $\langle \nabla f(x), x - x(\alpha) \rangle \geq \frac{1}{\alpha} \|x(\alpha) - x\|^2$, for all $\alpha > 0$.

Lemma 2.2. *Under the conditions of the above lemma, it holds that*

- (i) $\|x(\alpha) - x\|$ is nondecreasing on $\alpha > 0$ (see [14]);
- (ii) $\frac{1}{\alpha} \|x(\alpha) - x\|$ is nonincreasing on $\alpha > 0$ (see [4] or [3]);
- (iii) $\langle \nabla f(x), x - x(\alpha) \rangle$ is nondecreasing on $\alpha > 0$;
- (iv) $\frac{1}{\alpha} \langle \nabla f(x), x - x(\alpha) \rangle$ is nonincreasing on $\alpha > 0$ (see [15]).

For a given closed convex set Ω , we denote the tangent cone $T(x)$ at $x \in \Omega$ to be the closure of the cone of all feasible directions at x . Since $T(x)$ is a nonempty closed convex set in R^n ,

the projection, denoted by $\nabla_{\Omega}f(x)$, of $-\nabla f(x)$ onto $T(x)$ is uniquely defined, i.e.,

$$\nabla_{\Omega}f(x) = \arg \min\{\|v + \nabla f(x)\| \mid v \in T(x)\}.$$

Calamai and Moré [3] called it the projected gradient, and showed some of its characterizations.

Lemma 2.3. *Let ∇f be continuous on the closed convex set Ω . Then*

- (i) *the point $x^* \in \Omega^*$ if and only if $\nabla_{\Omega}f(x^*) = 0$;*
- (ii) *$\min\{\langle v, \nabla f(x) \rangle \mid v \in T(x), \|v\| \leq 1\} = -\|\nabla_{\Omega}f(x)\|$;*
- (iii) *the mapping $\|\nabla_{\Omega}f(x)\|$ is lower semicontinuous on Ω .*

Define $N = \{0, 1, 2, \dots\}$. Based on Lemma 2.1, we derive the following basic lemma.

Lemma 2.4. *Let $\{x^k\}$ be generated by formula (3). Then we have for any $x \in \Omega$,*

$$\frac{\langle \nabla f(x^k), x^k - x \rangle}{\|x^k - x\|} \leq \frac{\langle \nabla f(x^{k-1}), x^{k-1} - x^k \rangle}{\|x^{k-1} - x^k\|} + \|\nabla f(x^k) - \nabla f(x^{k-1})\|, \quad \forall k \in N.$$

Proof. For any $x \in \Omega$, by Lemma 2.1 (i) we have

$$\langle \nabla f(x^{k-1}), x^k - x \rangle \leq \frac{\langle x^{k-1} - x^k, x^k - x \rangle}{\alpha_{k-1}}, \quad \forall k \in N.$$

From the above inequality and Lemma 2.1 (ii) we obtain for any $k \in N$,

$$\begin{aligned} \frac{\langle \nabla f(x^k), x^k - x \rangle}{\|x^k - x\|} &= \frac{\langle \nabla f(x^{k-1}), x^k - x \rangle}{\|x^k - x\|} + \frac{\langle \nabla f(x^k) - \nabla f(x^{k-1}), x^k - x \rangle}{\|x^k - x\|} \\ &\leq \frac{\langle x^{k-1} - x^k, x^k - x \rangle}{\alpha_{k-1} \|x^k - x\|} + \|\nabla f(x^k) - \nabla f(x^{k-1})\| \\ &\leq \frac{1}{\alpha_{k-1}} \|x^{k-1} - x^k\| + \|\nabla f(x^k) - \nabla f(x^{k-1})\| \\ &\leq \frac{\langle \nabla f(x^{k-1}), x^{k-1} - x^k \rangle}{\|x^{k-1} - x^k\|} + \|\nabla f(x^k) - \nabla f(x^{k-1})\|. \end{aligned}$$

The proof is complete.

Before stating and proving the following key lemma for our convergence analysis, we need to recall the notion of absolute continuity which can be founded in any text of mathematical analysis.

Definition 2.1. *We say that ∇f is absolutely continuous on Ω , if for given $\varepsilon > 0$, there is a number $\delta > 0$ such that for arbitrarily finite set $\{x^1, x^2, \dots, x^t\} \in \Omega$ satisfying $\sum_{i,j \in \{1, \dots, t\}} \|x^i - x^j\| < \delta$,*

$$\sum_{i,j \in \{1, \dots, t\}} \|\nabla f(x^i) - \nabla f(x^j)\| < \varepsilon.$$

Clearly, Lipschitz continuity of ∇f implies absolute continuity, and the latter implies uniform continuity. That is, absolute continuity is between Lipschitz continuity and uniform continuity.

Lemma 2.5. *Assume that f is bounded below and ∇f is absolutely continuous on Ω . Let $\{x^k\}$ be generated by (3) such that*

- (i) *$\{f(x^k)\}$ is monotonically decreasing;*

$$(ii) \lim_{k \rightarrow \infty} \|x^k - x^{k+1}\| = 0;$$

$$(iii) \liminf_{k \rightarrow \infty} \frac{\langle \nabla f(x^k), x^k - x^{k+1} \rangle}{\|x^k - x^{k+1}\|} = 0.$$

Then, $\lim_{k \rightarrow \infty} \nabla_{\Omega} f(x^k) = 0$.

Proof. From the uniform continuity of ∇f and the assumption (ii), we obtain

$$f(x^k) - f(x^{k+1}) = \langle \nabla f(x^k), x^k - x^{k+1} \rangle + o(\|x^k - x^{k+1}\|) \quad (6)$$

for all sufficiently large $k \in N$. For arbitrarily given $\epsilon > 0$, let

$$N(\epsilon) := \{k \in N \mid \frac{\langle \nabla f(x^k), x^k - x^{k+1} \rangle}{\|x^k - x^{k+1}\|} < \epsilon\}, \quad \bar{N}(\epsilon) := N \setminus N(\epsilon).$$

From the assumption (iii), we know that $N(\epsilon)$ is an infinite subset in N .

Suppose that $\bar{N}(\epsilon)$ is also an infinite subset. We will show that this is a contradiction. In view of the assumption (ii) and (6), we have for any sufficiently large $k' \in \bar{N}(\epsilon)$,

$$f(x^k) - f(x^{k+1}) \geq \frac{\epsilon}{2} \|x^k - x^{k+1}\|, \quad k \in \bar{N}(\epsilon), \quad k \geq k'. \quad (7)$$

This, together with the boundedness of f and the assumption (i), implies that

$$\begin{aligned} \frac{\epsilon}{2} \sum_{k \in \bar{N}(\epsilon), k \geq k'} \|x^k - x^{k+1}\| &\leq \sum_{k \in \bar{N}(\epsilon), k \geq k'} \{f(x^k) - f(x^{k+1})\} \\ &\leq \sum_{k \geq k'} \{f(x^k) - f(x^{k+1})\} < +\infty. \end{aligned}$$

Hence,

$$\lim_{k' \rightarrow \infty} \sum_{k \in \bar{N}(\epsilon), k \geq k'} \|x^k - x^{k+1}\| = 0. \quad (8)$$

Let $k'' = \max\{k \in N \mid k < k', k \in N(\epsilon/2)\}$. Then $k \in \bar{N}(\epsilon/2)$ for any $k \in [k'' + 1, k' - 1]$. By using Lemma 2.4 repeatedly, we have

$$\begin{aligned} \epsilon &\leq \frac{\langle \nabla f(x^{k'}), x^{k'} - x^{k'+1} \rangle}{\|x^{k'} - x^{k'+1}\|} \\ &\leq \frac{\langle \nabla f(x^{k'-1}), x^{k'-1} - x^{k'} \rangle}{\|x^{k'-1} - x^{k'}\|} + \|\nabla f(x^{k'}) - \nabla f(x^{k'-1})\| \\ &\leq \frac{\langle \nabla f(x^{k''}), x^{k''} - x^{k''+1} \rangle}{\|x^{k''} - x^{k''+1}\|} + \\ &\quad + \|\nabla f(x^{k''+1}) - \nabla f(x^{k''})\| + \sum_{k \in [k''+1, k'-1]} \|\nabla f(x^{k+1}) - \nabla f(x^k)\| \\ &< \frac{\epsilon}{2} + \|\nabla f(x^{k''+1}) - \nabla f(x^{k''})\| + \sum_{k \in \bar{N}(\epsilon/2), k \geq k''+1} \|\nabla f(x^{k+1}) - \nabla f(x^k)\|. \end{aligned}$$

That is,

$$\frac{\epsilon}{2} < \|\nabla f(x^{k''+1}) - \nabla f(x^{k''})\| + \sum_{k \in \bar{N}(\epsilon/2), k \geq k''+1} \|\nabla f(x^{k+1}) - \nabla f(x^k)\|. \quad (9)$$

From the assumption (ii), the absolute continuity of ∇f , and the claim (whose proof is similar to the one of (8))

$$\lim_{k'' \rightarrow \infty} \sum_{k \in \bar{N}(\epsilon/2), k \geq k''+1} \|x^k - x^{k+1}\| = 0,$$

we obtain

$$\lim_{k'' \rightarrow \infty} \|\nabla f(x^{k''+1}) - \nabla f(x^{k''})\| = 0, \tag{10}$$

and

$$\lim_{k'' \rightarrow \infty} \sum_{k \in \bar{N}(\epsilon/2), k \geq k''+1} \|\nabla f(x^{k+1}) - \nabla f(x^k)\| = 0. \tag{11}$$

Letting $k' \rightarrow \infty$, then by the assumption (iii), we have $k'' \rightarrow \infty$. Thus, taking the limit in (9) as $k' \rightarrow \infty$ and using (10) and (11), we derive $\frac{\epsilon}{2} \leq 0$, a contradiction.

Therefore, $\bar{N}(\epsilon)$ is a finite subset. Since $\epsilon > 0$ is arbitrary, this shows that

$$\lim_{k \rightarrow \infty} \frac{\langle \nabla f(x^k), x^k - x^{k+1} \rangle}{\|x^k - x^{k+1}\|} = 0. \tag{12}$$

Now, we prove the conclusion of this lemma. By the definition of $T(x^k)$, Lemma 2.4 yields

$$-\langle \nabla f(x^k), v \rangle \leq \frac{\langle \nabla f(x^{k-1}), x^{k-1} - x^k \rangle}{\|x^{k-1} - x^k\|} + \|\nabla f(x^k) - \nabla f(x^{k-1})\|, \quad \forall k \in N, \tag{13}$$

for any $v \in T(x^k)$ with $\|v\| = 1$. Hence, from (13) and Lemma 2.3 (ii) we know that

$$\|\nabla_{\Omega} f(x^k)\| \leq \frac{\langle \nabla f(x^{k-1}), x^{k-1} - x^k \rangle}{\|x^{k-1} - x^k\|} + \|\nabla f(x^k) - \nabla f(x^{k-1})\|, \quad \forall k \in N. \tag{14}$$

Taking the limit in (14) as $k \rightarrow \infty$, and by using (12), the assumption (ii) and the uniform continuity of ∇f , we obtain the desired result. The proof is complete.

3. Main Results

In this section, we proceed to the global convergence analysis of the gradient projection method (3) with bounded exact stepsize rule, i.e., α_k is taken as a constrained minimizer of the subproblem

$$\min\{f(x^k(\alpha)) \mid \alpha \in [0, c]\}, \tag{15}$$

where $c > 0$ is a constant. Clearly, this rule guarantees the existence of α_k in any case. The first result is that the distance of two neighbor iterates tends to zero.

Theorem 3.1. *Assume that f is bounded below and ∇f is uniformly continuous on Ω . If $\{x^k\}$ is generated by the method (3) and (15), then $\lim_{k \rightarrow \infty} \|x^k - x^{k+1}\| = 0$.*

Proof. By the bounded exact stepsize rule (15), we have for any $\alpha \in [0, c]$,

$$f(x^k) - f(x^{k+1}) \geq f(x^k) - f(x^k(\alpha)), \quad \forall k \in N. \tag{16}$$

Define $\tilde{x}^{k+1} := x^k(\tilde{\alpha}_k)$ for every x^k , where $\tilde{\alpha}_k$ is the largest element in $\{c, cl, cl^2, \dots\}$ ($l \in (0, 1)$) such that

$$f(x^k) - f(\tilde{x}^{k+1}) \geq \mu \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle. \tag{17}$$

Then by modelling on the proof of Theorem 2.3 in [3], we readily obtain

$$\lim_{k \rightarrow \infty} \frac{1}{\tilde{\alpha}_k} \|x^k - \tilde{x}^{k+1}\| = 0. \tag{18}$$

In fact, assume that there is an infinite subset $N_0 \subseteq N$ such that

$$\frac{1}{\tilde{\alpha}_k} \|x^k - \tilde{x}^{k+1}\| \geq \epsilon > 0, \quad \forall k \in N_0.$$

Then from (16), (17) and Lemma 2.1 (ii), we imply that for $k \in N_0$,

$$\begin{aligned} f(x^k) - f(x^{k+1}) &\geq f(x^k) - f(\tilde{x}^{k+1}) \\ &\geq \mu \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle \geq \frac{\mu}{\tilde{\alpha}_k} \|x^k - \tilde{x}^{k+1}\|^2 \\ &\geq \mu \epsilon \max\{\epsilon \tilde{\alpha}_k, \|x^k - \tilde{x}^{k+1}\|\}. \end{aligned}$$

By convergence of $\{f(x^k)\}$, one has $\lim_{k \in N_0, k \rightarrow \infty} \tilde{\alpha}_k = 0$, and $\lim_{k \in N_0, k \rightarrow \infty} \|x^k - \tilde{x}^{k+1}\| = 0$. Thus, from the way of choosing $\tilde{\alpha}_k$, we obtain that for all sufficiently large $k \in N_0$,

$$f(x^k) - f(x^k(\tilde{\alpha}_k l^{-1})) < \mu \langle \nabla f(x^k), x^k - x^k(\tilde{\alpha}_k l^{-1}) \rangle.$$

This, together with the uniform continuity of ∇f , Lemma 2.2 (ii) and (iii), shows that for all sufficiently large $k \in N_0$,

$$(1 - \mu) < \frac{o(\|x^k - x^k(\tilde{\alpha}_k l^{-1})\|)}{\langle \nabla f(x^k), x^k - x^k(\tilde{\alpha}_k l^{-1}) \rangle} \leq \frac{o(l^{-1}\|x^k - \tilde{x}^{k+1}\|)}{\langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle} \leq \frac{o(l^{-1}\|x^k - \tilde{x}^{k+1}\|)}{\epsilon \|x^k - \tilde{x}^{k+1}\|}.$$

Taking the limits on the above inequality as $k \in N_0, k \rightarrow \infty$, leads to $1 - \mu \leq 0$, a contradiction. Hence (18) is true.

If $\alpha_k \geq \tilde{\alpha}_k$, from Lemma 2.2 (ii) and (15) we have

$$\|x^k - x^{k+1}\| \leq \frac{c}{\tilde{\alpha}_k} \|x^k - \tilde{x}^{k+1}\|. \quad (19)$$

If $\alpha_k < \tilde{\alpha}_k$, from (16), (17), Lemma 2.2 (iii) and Lemma 2.1 (ii) we obtain

$$\begin{aligned} f(x^k) - f(x^{k+1}) &\geq f(x^k) - f(\tilde{x}^{k+1}) \geq \mu \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle \\ &\geq \mu \langle \nabla f(x^k), x^k - x^{k+1} \rangle \geq \frac{\mu}{c} \|x^k - x^{k+1}\|^2. \end{aligned} \quad (20)$$

Combining (19) and (20) yields

$$\|x^k - x^{k+1}\| \leq \max \left\{ \frac{c}{\tilde{\alpha}_k} \|x^k - \tilde{x}^{k+1}\|, \sqrt{\frac{c}{\mu} (f(x^k) - f(x^{k+1}))} \right\}.$$

This, together with the monotonicity of $\{f(x^k)\}$ and the boundedness of f and (18), leads to the desired result. The proof is complete.

If the uniform continuity of ∇f is replaced by the absolute continuity, then we have a further result.

Theorem 3.2. *Assume that f is bounded below and ∇f is absolutely continuous on Ω . If $\{x^k\}$ is generated by the method (3) and (15), then*

$$\lim_{k \rightarrow \infty} \nabla_{\Omega} f(x^k) = 0 \quad \text{or} \quad \lim_{k \rightarrow \infty} x^k = x^* \in \Omega^*.$$

Proof. If there is a small number $\epsilon > 0$ such that

$$\langle \nabla f(x^k), x^k - x^{k+1} \rangle \geq \epsilon \|x^k - x^{k+1}\|, \quad \forall k \in N. \quad (21)$$

Then from Theorem 3.1 and the absolute continuity of ∇f , we know that for all sufficiently large $k \in N$,

$$\begin{aligned} f(x^k) - f(x^{k+1}) &= \langle \nabla f(x^k), x^k - x^{k+1} \rangle + o(\|x^k - x^{k+1}\|) \\ &\geq \frac{\epsilon}{2} \|x^k - x^{k+1}\|. \end{aligned}$$

This implies that for all sufficiently large k' ,

$$\sum_{k \geq k'} \|x^k - x^{k+1}\| \leq \frac{2}{\epsilon} \sum_{k \geq k'} \{f(x^k) - f(x^{k+1})\} < +\infty,$$

i.e., $\lim_{k' \rightarrow \infty} \sum_{k \geq k'} \|x^k - x^{k+1}\| = 0$. This shows that $\{x^k\}$ is a convergent sequence.

Let $\lim_{k \rightarrow \infty} x^k = x^*$. Thus, from (17), Lemma 2.1 (i), (16), $\{f(x^k)\} \downarrow f(x^*)$ and (18) we derive for every $z \in \Omega$,

$$\begin{aligned} \langle \nabla f(x^*), x^* - z \rangle &= \lim_{k \rightarrow \infty} \{ \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle + \langle \nabla f(x^k), \tilde{x}^{k+1} - z \rangle \} \\ &\leq \lim_{k \rightarrow \infty} \left\{ \frac{f(x^k) - f(\tilde{x}^{k+1})}{\mu} + \frac{\langle \tilde{x}^{k+1} - x^k, z - \tilde{x}^{k+1} \rangle}{\tilde{\alpha}_k} \right\} \\ &\leq \lim_{k \rightarrow \infty} \left\{ \frac{f(x^k) - f(x^{k+1})}{\mu} + \frac{\langle \tilde{x}^{k+1} - x^k, z - x^k \rangle}{\tilde{\alpha}_k} - \frac{\|\tilde{x}^{k+1} - x^k\|^2}{\tilde{\alpha}_k} \right\} \\ &\leq \lim_{k \rightarrow \infty} \left\{ \frac{f(x^k) - f(x^{k+1})}{\mu} + \frac{\|\tilde{x}^{k+1} - x^k\| \|z - x^k\|}{\tilde{\alpha}_k} \right\} \\ &= 0. \end{aligned}$$

This shows that $x^* \in \Omega^*$.

If (21) does not hold, then we have

$$\liminf_{k \rightarrow \infty} \frac{\langle \nabla f(x^k), x^k - x^{k+1} \rangle}{\|x^k - x^{k+1}\|} = 0.$$

By applying Lemma 2.5 and using Theorem 3.1 and the above equality, we deduce that $\lim_{k \rightarrow \infty} \nabla_{\Omega} f(x^k) = 0$. The proof is complete.

Moreover, if the above assumptions on f are strengthened so that f is convex and ∇f is Lipschitz continuous, then the gradient projection method with bounded exact stepsize rule generates a convergent sequence whose limit point is a solution of problem (1). This is the main result of the paper.

Theorem 3.3. *Assume that f is convex and ∇f is Lipschitz continuous on Ω , and that Ω^* is nonempty. If $\{x^k\}$ is generated by the method (3) and (15), then $\lim_{k \rightarrow \infty} x^k = x^{\infty} \in \Omega^*$.*

Proof. We first prove that there exists a constant $c_1 > 0$ satisfying

$$\langle \nabla f(x^k), x^k - x^{k+1} \rangle \leq c_1 \{f(x^k) - f(x^{k+1})\}, \quad \forall k \in N. \quad (22)$$

In fact, if $\alpha_k \leq \tilde{\alpha}_k$ then by Lemma 2.2 (iii),

$$\langle \nabla f(x^k), x^k - x^{k+1} \rangle \leq \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle, \quad (23)$$

where $\tilde{x}^{k+1} = x^k(\tilde{\alpha}_k)$ and $\tilde{\alpha}_k$ is produced by the Armijo stepsize rule (17). Let $L > 0$ be Lipschitz constant of ∇f . From (17) and the mean value theorem we derive

$$\tilde{\alpha}_k \geq \min\{c, 2l(1 - \mu)/L\} := c_2 > 0.$$

If $\alpha_k > \tilde{\alpha}_k$, then by the above inequality and $c \geq \alpha_k$, from Lemma 2.2 (iv) we know that

$$\langle \nabla f(x^k), x^k - x^{k+1} \rangle \leq \frac{c}{c_2} \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle. \tag{24}$$

Thus, from (23), (24), (17) and (16) we obtain

$$\begin{aligned} \langle \nabla f(x^k), x^k - x^{k+1} \rangle &\leq \max\{1, \frac{c}{c_2}\} \langle \nabla f(x^k), x^k - \tilde{x}^{k+1} \rangle \\ &\leq \frac{c}{c_2\mu} \{f(x^k) - f(\tilde{x}^{k+1})\} \\ &\leq \frac{c}{c_2\mu} \{f(x^k) - f(x^{k+1})\}. \end{aligned}$$

Setting $c_1 = \frac{c}{c_2\mu}$ yields (22).

We next prove that there exists a constant $c_3 > 0$ such that for any $x^* \in \Omega^*$,

$$\|x^{k+1} - x^*\|^2 + c_3 f(x^{k+1}) \leq \|x^k - x^*\|^2 + c_3 f(x^k), \quad \forall k \in N, \tag{25}$$

where Ω^* is the solution set of problem (1) since f is convex. In fact, for any $x^* \in \Omega^*$ we have

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &= \|x^k - x^*\|^2 + 2\langle x^{k+1} - x^k, x^k - x^* \rangle + \|x^{k+1} - x^k\|^2 \\ &\leq \|x^k - x^*\|^2 + 2\langle x^{k+1} - x^k, x^{k+1} - x^* \rangle \\ &\leq \|x^k - x^*\|^2 + 2\alpha_k \langle \nabla f(x^k), x^* - x^{k+1} \rangle \\ &= \|x^k - x^*\|^2 + 2\alpha_k \langle \nabla f(x^k), x^k - x^{k+1} \rangle + 2\alpha_k \langle \nabla f(x^k), x^* - x^k \rangle \\ &\leq \|x^k - x^*\|^2 + 2cc_1 \{f(x^k) - f(x^{k+1})\}, \end{aligned}$$

where the second inequality uses Lemma 2.1 (i), and the last inequality comes from (15), (22) and the convexity of f . Letting $c_3 = 2cc_1$ yields (25).

Finally, we prove the conclusion of the theorem. Since $\{x^k\}$ is bounded by (25) and by the monotonicity of $\{f(x^k)\}$, there exists a point $x^\infty \in \Omega$ satisfying

$$\lim_{k \in N_0 \subseteq N, k \rightarrow \infty} x^k = x^\infty \in \Omega^*. \tag{26}$$

Replacing x^* in (25) by x^∞ , we know that $\{\|x^k - x^\infty\|^2 + c_3 f(x^k)\}$ is monotonically decreasing, and hence it converges to $c_3 f(x^\infty)$ by (26) and $\{f(x^k)\} \downarrow f(x^\infty)$. Thus,

$$\lim_{k \rightarrow \infty} \|x^k - x^\infty\|^2 = \lim_{k \rightarrow \infty} \{\|x^k - x^\infty\|^2 + c_3 f(x^k)\} - \lim_{k \rightarrow \infty} c_3 f(x^k) = 0.$$

The proof is complete.

In Theorem 3.3, the Lipschitz continuity of ∇f can not be weakened to continuity of ∇f , although full convergence of the gradient projection method with inexact Armijo stepsize needs only the continuity of ∇f (for the special case where $\Omega = R^n$, see [2, 6, 8]; for a general closed convex set Ω , see [13]). In fact, Gonzaga [6] gave a genuine example for the special case where $\Omega = R^n$, in which if f is continuously differentiable, convex and strictly convex at all non-optimal points, the steepest descent method with exact stepsize rule for solving this unconstrained problem generates four distinct accumulation points with $\|x^k - x^{k+1}\|$ bounded away from zero.

Theorem 3.4. *Assume that f is convex and ∇f is Lipschitz continuous on Ω , and that Ω^* is nonempty. If $\{x^k\}$ is generated by the method (3) and (15) and if the sequence $\{\alpha_k\}$ of stepsizes satisfies*

$$c_0 \leq \alpha_k (\leq c), \quad \forall k \in N, \tag{27}$$

where $c_0 > 0$ is a constant, then $\lim_{k \rightarrow \infty} x^k = x^\infty \in \Omega^*$ and $\lim_{k \rightarrow \infty} \nabla_{\Omega} f(x^k) = 0$.

Proof. Observing the proof of Lemma 2.4, we have actually obtained that for any $x \in \Omega$,

$$\frac{\langle \nabla f(x^k), x^k - x \rangle}{\|x^k - x\|} \leq \frac{1}{\alpha_{k-1}} \|x^{k-1} - x^k\| + \|\nabla f(x^k) - \nabla f(x^{k-1})\|, \quad \forall k \in N,$$

from which, similar to the proof of (14), we imply that

$$\|\nabla_{\Omega} f(x^k)\| \leq \frac{1}{\alpha_{k-1}} \|x^{k-1} - x^k\| + \|\nabla f(x^k) - \nabla f(x^{k-1})\|, \quad \forall k \in N.$$

Thus, the desired result follows immediately from Theorems 3.1 and 3.3, and the given assumptions.

4. Concluding Remarks

In this paper, we have presented global convergence properties of the gradient projection method with bounded exact step-size rule (15), where the constant c plays a key role. To verify the convergence behavior of the method and to find a way of choosing c , we have done numerical experiments for solving the constrained optimization problems [1], and some numerical results are shown in a technical report. In future studies, we plan to do further work on both theory and numerical testing for the exact step-size rule.

Acknowledgments. The authors would like to thank the two referees for their careful reading of the paper and for their helpful suggestions.

References

- [1] Bello, L. and Raydan, M., Preconditioned spectral projected gradient method on convex sets, *J. Comput. Math.*, **23** (2005), 225-232.
- [2] Burachik, R., Drummond, L., Iusem, A. and Svaiter, B., Full convergence of the steepest descent method with inexact searches, *Optimization*, **32** (1995), 137-146.
- [3] Calamai, P.H. and Moré, J.J., Projected gradient methods for linearly constrained problems, *Mathematical Programming*, **39** (1987), 93-116.
- [4] Gafni, E.M. and Bertsekas, D.P., Two-metric projection methods for constrained optimization, *SIAM J. Control Optim.*, **22** (1984), 936-964.
- [5] Goldstein, A.A., Convex programming in Hilbert space, *Bulletin of the American Mathematical Society*, **70** (1964), 709-710.
- [6] Gonzaga, C.C., Two facts on the convergence of the Cauchy algorithm, *J. Optim. Theory Appl.*, **107** (2000), 591-600.
- [7] Hager, W.W. and Park, S., The gradient projection method with exact line search, *J. Global Optimization*, **30** (2004), 103-118.
- [8] Kiwiel, K. and Murty, K.G., Convergence of the steepest descent method for minimizing quasi-convex functions, *J. Optim. Theory Appl.*, **89** (1996), 221-226.
- [9] Levitin, E.S. and Polyak, B.T., Constrained minimization problems, *USSR Computational Mathematics and Mathematical Physics*, **6** (1966), 1-50.
- [10] McCormick, G.P. and Tapia, R.A., The gradient projection method under mild differentiability conditions, *SIAM J. Control Optim.*, **10** (1972), 93-98.
- [11] Phelps, R.R., Metric projections and the gradient projection method in Banach space, *SIAM J. Control Optim.*, **23** (1985), 973-977.

- [12] Phelps, R.R., The gradient projection method using Curry's steplength, *SIAM J. Control Optim.*, **24** (1986), 692-699.
- [13] Wang, C.Y. and Xiu, N.H., Convergence of the gradient projection method for generalized convex minimization, *Computational Optim. Appl.*, **16** (2000), 111-120.
- [14] Toint, P.H.L., Global convergence of a class of trust region methods for nonconvex minimization in Hilbert space, *IMA J. Numer. Anal.*, **8** (1988), 231-252.
- [15] Xiu, N.H., Wang, C.Y. and Zhang, J.Z., Convergence properties of projection and contraction methods for variational inequality problems, *Applied Math. Opt.*, **43** (2001), 147-168.