第4卷 第2期
1984年4月

数学研究与评论
JOURNAL OF MATHEMATICAL
RESEARCH AND EXPOSITION

Vol.4 No.2
April 1984

Rate of Convergence of an Improved Reduced Gradient Method*

Zhou Huishan (周惠山)

(Qufu Teachers' College)

Abstract

An improved reduced gradient method was proposed in [4] to solve the nonlinear programming (P) with linear constraints:

$$(P)\quad \min_{X\in R} f(x) \qquad R=\{x\in E^n \mid AX=b,\ x\geqslant 0\} \quad b\in E^m.$$

In this paper we introduce parameters ρ_k which is the skill used in [5] to the algorithm of [4] to obtain a reduced gradient method which is linearly convergent under the conditions of R being non-degenerate, f being second-order continuously differentiable and strong convex.

I Hypotheses and Notations

We shall study the following nonlinear programming with linear constraints:

$$(P)\quad \min_{X\in R} f(x) \qquad R=\{x \mid Ax=b,\ x\geqslant 0,\quad x\in E^n\}$$

where A is a $m\times n$ matrix ($m\leqslant n$), $b\in E^m$, E^n and E^m are n-dimension and m-dimension Euclidean space respectively. Suppose that the rank of A is equal to m. We assume that

($H1$) $R\neq\phi$, every extreme point of the polyhedron R is non-degenerate.

($H2$) the function f: $E^n\to E^1$ is real-valued first-order continuously differentiable in E^n.

R* denotes the set of optimal solutions of (P). A_L^J is the submatrix of A consisting of elements a_{ij}, $(i,j)\in L\times J$, where $J\subseteq\{1,\cdots,n\}$, $L\subseteq\{1,\cdots,m\}$. If $L=\{1,\cdots,m\}$, A_L^J is denoted by A^J for brief. $I\subseteq\{1,\cdots,n\}$ is called a basis if both the number of elements in I and the rank of A^I are equal to m. $\bar{I}=\{1,\cdots,n\}\backslash I$. $T(I)=(A^I)^{-1}A$, $T^{\bar{I}}(I)=(A^I)^{-1}A^{\bar{I}}$. $t(I)=(A^I)^{-1}b$. $T_i^{\bar{I}}(I)$ denotes the i row vector in $T^{\bar{I}}(I)$, T_i^j denotes the (i,j) element in $T^{\bar{I}}(I)$.

*Received June 17, 1981.

$\nabla_I f(x)=\left(\frac{\partial f(x)}{\partial x_i},\ j\in I\right)$ and $\nabla_{\bar I} f(x)=\left(\frac{\partial f(x)}{\partial x_i},\ j\in \bar I\right)$ denote column vectors. x is a column vector, the transpose x' is a row vector.

If $x=(x_I^t, x_{\bar I}^t)^t\in R$, then for any basis I, we have $x_I=t(I)-T^{\bar I}(I)x_{\bar I}$,

We define
$$\bar f(x_{\bar I})=f(t(I)-T^{\bar I}(I)x_{\bar I},\ x_{\bar I})$$
and $\nabla\bar f(x_{\bar I})$ is called the "reduced gradient", which is equal to
$$\nabla\bar f(x_{\bar I})=\nabla_{\bar I} f(x)-T^{\bar I}(I)'\nabla_I f(x).$$

II Algorithm

We perform pivotal operations given in [3] and have the lemma 1.

Lemma 1 If $(H1)$ is satisfied, then for any feasible point $x\in R$, any basis I, any positive number $\varepsilon<1$ and any index set $D\subset\bar I$, the pivotal process must terminate after at most m times of pivotal operations. Furthermore, provided that I_p, ε_p, and D_p denote the final I, ε and D respectively, we have $\varepsilon_p<1$, $D_p\subset\bar I_p$ and $\min\{x_i\,|\,i\in I_p\}>\frac{\varepsilon_p}{2}$,

Now we shall give an iterative algorithm of the reduced gradient method to solve the problem (P).

Algorithm Starting from an arbitrary feasible point $x^1\in R$, an arbitrary basis I_0, a positive number $\varepsilon_0<1$ and an index set $D_0=\phi$, let $k=1$.

(1) Perform the pivotal operations for $(x^k,\ I_{k-1},\ \varepsilon_{k-1},\ D_{k-1})$, set $I_k=I_p$, $\varepsilon_k=\varepsilon_p$, $D_k=D_p$, we get $(x^k,\ I_k,\ \varepsilon_k,\ D_k)$ and go on to (2).

(2) Compute $T(I_k)$ and $\nabla\bar f(x^k_{\bar I_k})$, and go on to (3).

(3) given $\rho_k>0$, and go on to (4).

(4) Define vector $\tilde x^k_{\bar I_k}$:

$$\tilde x_j^k=\begin{cases}0 & \text{if } x_j^k\leqslant\rho_k\frac{\partial\bar f(x^k_{\bar I_k})}{\partial x_i},\ j\in\bar I_k.\\[2mm] x_j^k-\rho_k\frac{\partial\bar f(x^k_{\bar I_k})}{\partial x_i} & \text{if } x_j^k>\rho_k\frac{\partial\bar f(x^k_{\bar I_k})}{\partial x_i},\ j\in\bar I_k.\end{cases}$$

If $\tilde x^k_{\bar I_k}=x^k_{\bar I_k}$, stop; otherwise go on to (5).

(5) Let
$$x^{k\prime}_{I_k}=t(I_k)-T^{\bar I_k}(I_k)\ \tilde x^k_{\bar I_k}$$
$$x^{k+1}=x^k+\lambda_k(\tilde x^k-x^k)$$
where λ_k is the maximum of the sequence $1,\frac{1}{2},\cdots,\frac{1}{2^n},\cdots$ which satisfy
$$x^{k+1}\in R,\qquad f(x^k)-f(x^{k+1})\geqslant-\frac{\lambda_k}{2}\nabla\bar f(x^k_{\bar I_k})'(\tilde x^k_{\bar I_k}-x^k_{\bar I_k})$$
Set $k=k+1$, then go back to (1).

We obtain the following lemma from the proof of the proposition 4 in [3]:

Lemma 2 There exists a positive integer k_0 such that $\varepsilon_k=\varepsilon_{k_0}$ for all $k\geqslant k_0$.

As presented above, assuming that $(H1)$ and $(H2)$ are satisfied in the following lemmas, we obtain lemmas 3-5 similar to lemmas 2-3 in [4].

Lemma 3 Let $\varphi(x_{\bar I_k})=\|x_{\bar I_k}-x^k_{\bar I_k}+\rho_k\nabla\bar f(x^k_{\bar I_k})\|^2$, then

(1) $\tilde x^k_{\bar I_k}$ is the solution of $\min\{\varphi(x_{\bar I_k})\,|\,x_{\bar I_k}\geqslant 0\}$;

(2) For any $x_{\bar I_k}\geqslant 0$ we have $(\tilde x^k_{\bar I_k}-x^k_{\bar I_k})^t(\tilde x^k_{\bar I_k}-x_{\bar I_k})\leqslant-\rho_k\nabla\bar f\,(x^k_{\bar I_k})^t(\tilde x^k_{\bar I_k}-x_{\bar I_k})$.

Lemma 4 If $\tilde x^k_{\bar I_k}=x^k_{\bar I_k}$, x^k is a K-T point of (p).

Lemma 5 For any $x_{\bar I_k}\geqslant 0$, the inequality

$$f(x^{k+1})-f(x^k)\leqslant\frac{1}{4}\rho_k^{-1}\{\|x^k_{\bar I_k}-x_{\bar I_k}\|^2-\|x^{k+1}_{\bar I_k}-x_{\bar I_k}\|^2\}+\frac{1}{2}\lambda_k\nabla\bar f(x^k_{\bar I_k})^t\,(x_{\bar I}-x^k_{\bar I_k})$$

holds.

Now It is obvious that there exists the step size λ_k satisfying (5) of the algorithm. Set $x_{\bar I_k}=x^k_{\bar I_k}$ in Lemma 5, we have

Lemma 6 $\frac{1}{4}\rho_k^{-1}\|x^{k+1}_{\bar I_k}-x^k_{\bar I_k}\|^2\leqslant f(x^k)-f(x^{k+1})$. Hence $f(x^k)$ is always not increasing as $k\to\infty$.

Let $R'_k=\{x_{\bar I_k}\geqslant 0\,|\,\bar f(x_{\bar I_k})<\bar f(x^k_{\bar I_k})\}$. If f$(x)$ is pseudo-convex, then for $x_{\bar I_k}\in R'_k$ we have

$$\nabla\bar f(x^k_{\bar I_k})^t\,(x_{\bar I_k}-x^k_{\bar I_k})=\nabla f(x^k)^t(x-x^k)<0$$

Hence we deduce the following

Lemma 7 If f(x) is pseudo-convex, then

$$f(x^{k+1})-f(x^k)<\frac{1}{4}\rho_k^{-1}\{\|x^k_{\bar I_k}-x_{\bar I_k}\|^2-\|x^{k+1}_{\bar I_k}-x_{\bar I_k}\|^2\}$$

holds for any x: $x_{\bar I_k}\in R'_k$.

Lemma 8 Given any a: $Aa=0$, for any k, $a=(a^t_{I_k},\ a^t_{\bar I_k})^t$ there exist two constants $\mu_1^k\geqslant\mu_2^k>0$ which depends on I_k and is irrelevant to a such that

$$\mu_2\|a_{\bar I_k}\|^2\leqslant\|a\|^2\leqslant\mu_1^k\|a_{\bar I_k}\|^2$$

Proof Since $\|a\|^2=\|a_{\bar I_k}\|^2+\|a_{I_k}\|^2=a^t_{\bar I_k}\mathrm{E}a_{\bar I_k}+a_{\bar I_k}(T^{\bar I_k}(I_k))^tT^{\bar I}\ (\mathrm{I}_k)a_{\bar I_k}=a^t_{\bar I_k}(E+B_k)a_{\bar I_k}$ where E is a unit matrix $(\mathrm{n}-\mathrm{m})\times(\mathrm{n}-\mathrm{m})$, E and $B_k=T^{\bar I_k}(I_k)^tT^{\bar I_k}(I_k)$ are both positive definite matrix. Let μ_1^k and μ_2^k be the maximum and minimum eigenvalues of matrix $E+B_k$ respectively, the result is followed.

Theorem 1 Assume that $(H1)$ and $(H2)$ are satisfied, and that $\{\rho_k\}$ is a bounded sequence. Let x^1 be an arbitrary feasible solution of (P). Then either the algorithm leads to a $K.-T.$ point in a finite number of steps, or every cluster point of $\{x^k\}$ generated by it is a $K.-T.$ point.

The proof is similar to that of Theorem 1 in [4].

The total number of pivotal operations is finite if there exist an integer $k_0 \geqslant 0$, a basis I_* such that $I_k = I_*$ for all $k \geqslant k_0$. From Lemma [3] and [4] we obtain Theorem 2.

Theorem 2 Assume that $(H1)$ and $(H2)$ are satisfied, $f(x)$ is pseudo-convex and there exist $B > b > 0$ such that $B \geqslant \rho_k \geqslant b$, then either the algorithm leads to a $K.-T.$ point in a finite number of steps or the algorithm generates a sequence $\{x^k\}$ satisfying the following property:

(1) If $R^* \neq \phi$, the total number of pivotal operations is finite, then there exist a positive integer k_0, a basis I_* and $\beta > 0$ such that $\|x^k_{\bar{I}_*} - x_{\bar{I}_*}\|^2 + \beta f(x^k)$ is monotone decreasing for $k \geqslant k_0$ where $x \in R^*$;

(2) The necessary and sufficient condition for $R^* \neq \phi$ and the total number of pivotal operations being finite is that the sequence $\{x^k\}$ is convergent.

Proof If there exists a k such that $\tilde{x}^k = x^k$, then x^k is a K.-T. point of (P), and an optimal solution of (P) too.

Now we suppose that $\tilde{x}^k \neq x^k (k = 1, 2, \cdots)$. From the condition of (1). there exist a positive integer k_0 and a basis I_* such that

$$R' = \{x_{\bar{I}_*} | x \in R^*\} \subset R'_k$$

holds for all $k \geqslant k_0$. From Lemma 7, when $x \in R^*$

$$f(x^{k+1}) - f(x^k) < \frac{1}{4}\rho_k^{-1}\{\|x^k_{\bar{I}_k} - x_{\bar{I}_k}\|^2 - \|x^{k+1}_{\bar{I}_*} - x_{\bar{I}_*}\|^2\}$$

holds. Take $\beta = 4B$, then $4\rho_k \leqslant \beta$. From the above inequality we have

$$\|x^{k+1}_{\bar{I}_*} - x_{\bar{I}_*}\|^2 + \beta f(x^{k+1}) < \|x^k_{\bar{I}_*} - x_{\bar{I}_*}\|^2 + \beta f(x^k).$$

Secondarily, we shall prove (2). The sufficiency can be proved as follows. Suppose that $\lim\limits_{k \to \infty} x^k = x^*$, then we know that the total number of pivotal operations is finite from Theorem 7 in [3], and x^* is a $K.-T.$ point from Theorem 1. Since $f(x)$ is pseudo-convex, then x^* is an optimal solution .i. e. $R^* \neq \phi$. The proof of the necessity is similar to that of Theorem 3 in [4].

Ⅲ Rate of convergence

In order to estimate the rate of convergence we must assume further that

$(H3)$ $f(x)$ is second-order continuously differentiable.

Let $x^k \in R$, I_k is a basis, set

$$\lambda'_k = \min_{i \in I_k} x^k_i \left(\max_{i \in I_k} \|T^{\bar{I}_k}_i(I_k)\| \right)^{-1} \tag{3.1}$$

$$\Omega_k = \{x | x_{\bar{I}_k} \geqslant 0,\ x \in R,\ \|x_{\bar{I}_k} - x^k_{\bar{I}_k}\| \leqslant \lambda'_k\} \tag{3.2}$$

$$B_k = (T^{\bar{I}}\ (I_k))^t T^{\bar{I}_k}(I_k)$$

To estimate a minimal positive integer S_k satisfying

$$S_k \geqslant n\|E+B_k\| \max_{1\leqslant i\leqslant j\leqslant n} \max_{x\in\Omega_k} \left|\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right| \tag{3.3}$$

where E is a $(n-m)\times(n-m)$ unit matrix, take

$$\rho_k = \min\{\lambda_k' \|\nabla \bar{f}(x^k_{\bar{I}_k})\|^{-1},\ S_k^{-1}\} \tag{3.4}$$

We assume further that

(H4) f(x) is a convex function.

(H5) f(x) is a strong convex fuuction i. e. there exists $\delta>0$ such that

$$\delta\|y\|^2 \leqslant y'\nabla^2 f(x) y$$

holds for all $y\in E^n$, $x\in R$, where $\nabla^2 f(x)$ denotes the matrix of second-order differentiative of f at x.

$\nabla^2 f(x)$ is a symmetric nonnegative matrix if $(H4)$ is satisfied. Hence the maximal absolute value among its elements surely appears on its diagonal, thus S_k can be estimated by a simpler formula:

$$S_k \geqslant n\|E+B_k\| \max_{1\leqslant i\leqslant n} \max_{x\in\Omega_k} \left|\frac{\partial^2 f(x)}{\partial x_i^2}\right| \tag{3.3'}$$

If f(x) is convex, estimating S_k by (3.3) is equivalent to estimating S_k by (3.3′). But the quantity of computing by (3.3′) is much smaller. Therefore S_k is estimated by (3.3′) if f(x) is convex; otherwise S_k is estimated by (3.3).

In lemma 9—14 assume that $(H1)$ and $(H3)$ are satisfied and ρ_k is defined by (3.1)—(3.4).

Lemma 9 $\tilde{x}^k\in R$ for $k=1,2,\cdots$

Proof From (3.1) and (3.3)

$$\begin{aligned}\tilde{x}_i^k &= x_i^k - T_i^{\bar{I}_k}(I_k)\,(\tilde{x}^k_{\bar{I}_k} - x^k_{\bar{I}_k}) \geqslant x_i^k - \|T_i^{\bar{I}_k}(I_k)\|\,\rho_k\|\nabla\bar{f}(x^k_{\bar{I}_k})\| \\ &\geqslant x_i^k - \min_{i\in I_k} x_i^k \geqslant 0\end{aligned} \tag{3.5}$$

hold for $i\in I_k$.

Lemma 10 $f(\tilde{x}^k) - f(x^k) \leqslant \nabla\bar{f}(x^k_{\bar{I}_k})'(\tilde{x}^k_{\bar{I}_k} - x^k_{\bar{I}_k}) + \frac{1}{2}\rho_k^{-1}\|\tilde{x}^k_{\bar{I}_k} - x^k_{\bar{I}_k}\|^2$.

Proof Applying Taylor's theorem we have

$$\begin{aligned}f(\tilde{x}^k) - f(x^k) &= \nabla f(x^k)'(\tilde{x}^k - x^k) + \frac{1}{2}(\tilde{x}^k - x^k)'\,\nabla^2 f(Z^k(\theta_k))\,(\tilde{x}^k - x^k) \\ &\leqslant \nabla\bar{f}(x^k_{\bar{I}_k})'(\tilde{x}^k_{\bar{I}_k} - x^k_{\bar{I}_k}) + \frac{1}{2}n\left\{\max_{1\leqslant i\leqslant j\leqslant n}\max_{0\leqslant\lambda\leqslant 1}\left|\frac{\partial^2 f(Z^k(\lambda))}{\partial x_i\partial x_j}\right|\right\}\|E+B_k\|\ \|\tilde{x}^k_{\bar{I}_k} - x^k_{\bar{I}_k}\|^2,\end{aligned} \tag{3.6}$$

where
$$z^k(\theta_k) = x^k + \theta_k(\tilde{x}^k - x^k),\quad 0\leqslant\theta_k\leqslant 1;$$
$$z^k(\lambda) = x^k + \lambda(\tilde{x}^k - x^k),\quad 0\leqslant\lambda\leqslant 1.$$

From (3.2) we have $\tilde{x}^k\in\Omega_k$, $Z^k(\theta_k)\in\Omega_k$, $Z^k(\lambda)\in\Omega_k$, and from (3.3) we have

$$n\left\{\max_{1\leq i<j\leq n}\max_{0\leq\lambda\leq 1}\left|\frac{\partial^2 f(Z^k(\lambda))}{\partial x_i\partial x_j}\right|\right\}\|E+B_k\|\leq\rho_k^{-1},$$

the result is followed.

Lemma 11 $x^{k+1}=\tilde{x}^k\ (k=1,2,\cdots)$.

Proof From Lemma 10, Lemma 3(1) and Lemma 9, we have $\lambda_k=1$ for all k.

Lemma 12 (1) $\rho_k\leq 1$; (2) If $\{x^k\}$ is bounded, then $\inf_k \rho_k>0$

Proof (1) is obvious.

(2) In view of $\{x^k\}$ being bounded and the number of different basis being finite, $\{\lambda_k'\}$ must be bounded; And from (3.3), the continuity of $\nabla f(x)$ and $\nabla^2 f(x)$ there exists $S>0$ such that

$$S_k\leq S,\qquad \|\nabla\bar{f}(x^k_{\bar{I}_k})\|\leq S$$

holds for any k. According to (3.1), (3.4) and Lemma 2, we have

$$\rho_k\geq S^{-1}\min\{(\inf_k\min_{i\in I_k}x_i^k)(\max_{I_k}\max_{i\in I_k}\|T_i^{\bar{I}_k}(I_k)\|)^{-1},1\}$$

hence
$$\inf_k\rho_k>0$$

Lemma 13 If $(H4)$ is satisfied, then

$$f(x^{k+1})-f(x)\leq\frac{1}{2}\rho_k^{-1}\{\|x^k_{\bar{I}_k}-x_{\bar{I}_k}\|^2-\|x^{k+1}_{\bar{I}_k}-x_{\bar{I}_k}\|^2\}$$

holds for any $x\in R$.

Proof Because of the convexity of $f(x)$

$$f(x^k)-f(x)\leq\nabla f(x^k)^t(x^k-x)=\nabla\bar{f}(x^k_{\bar{I}_k})^t(x^k_{\bar{I}_k}-x_{\bar{I}_k})$$

holds for any x. Because $x^{k+1}=\tilde{x}^k$, we can add this inequality to the inequality in Lemma 10:

$$f(x^{k+1})-f(x)\leq\nabla\bar{f}(x^k_{\bar{I}_k})^t(x^{k+1}_{\bar{I}_k}-x_{\bar{I}_k})+\frac{1}{2}\rho_k^{-1}\|x^{k+1}_{\bar{I}_k}-x^k_{\bar{I}_k}\|^2$$
$$=\frac{1}{2}\rho_k^{-1}\{\|x^k_{\bar{I}_k}-x_{\bar{I}_k}\|^2-\|x^{k+1}_{\bar{I}_k}-x_{\bar{I}_k}\|^2\}.$$

Theorem 3 Assume that $(H1)$, $(H3)$ and (H4) are satisfied; If the parameters ρ_k are defined by (3.1)—(3.4), then $\lambda_k=1$ for all k, and either the algorithm leads to a optimal solution in a finite number of steps, or the algorithm generates an infinite sequence $\{x^k\}$ satisfying the following properties:

(1) If $R^*\neq\phi$ and the total number of pivotal operations is finite, then there exist a positive integer k_0, a basis I_* such that the sequence $\{\|x^k_{\bar{I}_*}-x_{\bar{I}_*}\|\}$ is monotone decreasing for $k\geq k_0$; $(x\in R^*)$

(2) The necessary and sufficient condition for $R^*\neq\phi$ and the total number of pivotal operations being finite is that the sequence $\{x^k\}$ is convergent.

Proof From Lemma 11, $\lambda_k=1$. If $x_{k+1}\neq x_k\,(k=1,2,\cdots)$, then (1) is proved by applying Lemma 13. Hence $\{x^k\}$ is bounded, and we can apply Lemma 12 to show that ρ_k are satisfied with conditions of Theorem 2. Thus the other part of this theorem can be obtained by Theorem 2;

Lemma 14 Assume that $\lim_{k\to\infty} x^k = x^*$; If $f(x)$ is convex in a neighborhood of x^*

$$N(x^*) = \{x \mid \|x - x^*\| \leqslant \varepsilon\} \quad (\varepsilon > 0),$$

then there exist a positive integer k_0, a basis I_* such that

$$f(x^{k+1}) - f(x) \leqslant \frac{1}{2}\rho_k^{-1}\{\|x^k_{\bar{I}_*} - x_{\bar{I}_*}\|^2 - \|x^{k+1}_{\bar{I}_*} - x_{\bar{I}_*}\|^2\}$$

for all $k \geqslant k_0$ and all $x \in N(x^*) \cap R$;

Proof There exists k_1 such that $x^k \in N(x^*)$ for all $k \geqslant k_1$, and by virtue of convexity of $f(x)$ in $N(x^*)$, the inequality in Lemma 10 and applying the argument analogous to those in Lemma 13 we obtain

$$f(x^{k+1}) - f(x) \leqslant \frac{1}{2}\rho_k^{-1}\{\|x^k_{\bar{I}_k} - x_{\bar{I}_k}\|^2 - \|x^{k+1}_{\bar{I}_k} - x_{\bar{I}_k}\|^2\}.$$

According to the convergence of $\{x^k\}$ and the total number of pivotal operations being finite, there exists $k_0 \geqslant k_1$ such that $I_k = I_*$ for all $k \geqslant k_0$; This proves the lemma;

Theorem 4 Assume that $(H1)$ and $(H3)$ are satisfied, the parameters ρ_k are defined by (3.1)—(3.4), $\lim_{k\to\infty} x^k = x^*$ and $\nabla^2 f(x^*)$ is positive definite. Then there exist $0 < \alpha < 1$, a basis I_*, a positive integer k_0 such that

$$\|x^{k+1}_{\bar{I}_*} - x^*_{\bar{I}_*}\| \leqslant \alpha \|x^k_{\bar{I}_*} - x^*_{\bar{I}_*}\|$$

holds for all $k \geqslant k_0$.

Proof From $(H3)$ and $\nabla^2 f(x^*)$ being positive definite, there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$y^t \nabla^2 f(x) y \geqslant \delta \|y\|^2 \tag{3.7}$$

holds for each $x \in N(x^*) = \{x \mid \|x - x^*\| < \varepsilon\}$ and $y \in E^n$. Hence from Lemma 14, there exist positive integer k_0, a basis I_* such that

$$x^k \in N(x^*) \quad \text{for} \quad k \geqslant k_0$$

and

$$f(x^{k+1}) - f(x) \leqslant \frac{1}{2}\rho_k^{-1}\{\|x^k_{\bar{I}_*} - x_{\bar{I}_*}\|^2 - \|x^{k+1}_{\bar{I}_*} - x_{\bar{I}_*}\|^2\} \tag{3.8}$$

holds for $x \in N(x^*)$ and $k \geqslant k_0$; By applying $\nabla f(x^*)^t (x - x^*) \geqslant 0$ $(x \in R)$, $x_\lambda = x^* + \lambda(x^{k+1} - x^*) \in N(x^*)$ $(k \geqslant k_0)$, Taylor's theorem, (3.7) and Lemma 8 we have

$$\begin{aligned} f(x^{k+1}) - f(x^*) &= \nabla f(x^*)^t (x^{k+1} - x^*) + \frac{1}{2}(x^{k+1} - x^*)^t \nabla^2 f(x_\lambda)(x^{k+1} - x^*) \\ &\geqslant \frac{1}{2}\delta \|x^{k+1} - x^*\|^2 \geqslant \frac{1}{2}\delta \mu_2^k \|x^{k+1}_{\bar{I}_*} - x^*_{\bar{I}_*}\|^2 \\ &\geqslant \frac{1}{2}\delta \mu_2 \|x^{k+1}_{\bar{I}_*} - x^*_{\bar{I}_*}\|^2 \end{aligned} \tag{3.9}$$

where $\mu_2=\inf_k\{\mu_2^k\}>0$. Compare (3.8) and (3.9) with the fact that $\rho_k\geqslant b>0$,

$$\|x_{I_*}^{k+1}-x_{I_*}^*\|\leqslant a\|x_{I_*}^k-x_{I_*}^*\|$$

hold for $k\geqslant k_0$, where $0<a=(1+\delta\mu_2 b)^{-\frac{1}{2}}<1$

Theorem 5 Assume that $(H1)$, $(H3)$ and $(H5)$ are satisfied, the parameters ρ_k are defined by (3.1)—(3.4). Let x^1 be an arbitrary feasible solution, then either the algorithm leads to an optimal solution in a finite number of steps, or the algorithm generates an infinite sequence $\{x^k\}$ converging to the unique optimal solution of (P), and there exist a basis I_* $0<a<1$, a positive integer k_0 such that

$$\|x_{I_*}^{k+1}-x_{I_*}^*\|\leqslant a\|x_{I_*}^k-x_{I_*}^*\|$$

hold for all $k\geqslant k_0$.

Proof From $(H5)$ R^* contains a unique point x^*, and

$$E=\{x\,|\,x\in R,\ f(x)\leqslant f(x^1)\}.$$

is bounded. If the sequence $\{x^k\}$ is infinite, $\{x^k\}$ itself must converge to x^* from Theorem 1. The rate of convergence is obtained from Theorem 4.

Acknowledgement The author is grateful to professor Wang Changyu for his instructions.

References

[1] Wolfe, P., Methods of Nonlinear Programming, Recent Advances in Mathematical Programming (Eds. Graves-Wolfe), McGraw-Hill (1963).

[2] Wolfe, P., On the Convergence of Gradient Methods Under Constraints, IBM Journal of Res. and Dev., 16 (1972), pp. 407—411.

[3] Yue Minyi and Han Jiye, A New Reduced Gradient Method, SCIENTIA SINICA, Vol. XXII, No. 10, pp. 1099-1113.

[4] Wang Changyu, On the Convergence of An Improved Reduced Gradient Method, KEXUE TONGBAO, 17(1982), pp. 1030-1033.

[5] Wang Changyu, A New Pivotal Operation Method and the Simplification of Levitin-Polyak Gradient Projection Method And Its Convergent Property, Acta Mathematicae Applicatae Sinica, Vol. 14, No. 1(1981), pp.37-52.