Chin. Ann. of Math.
8B (4) 1987

A MODIFICATION OF POWELL-ZANGWILL'S METHOD AND ITS RATE OF CONVERGENCE

HE LIMIN (何利民)*

Abstract

In this paper, a modified version of Powell–Zangwill's method for function minimization without calculating derivatives is proposed. The new method possesses following properties: quadratic termination, global convergence for strictly convex function and Q–linear convergence rate for uniformly convex function. Furthermore, the main part of this paper is to show that the rate of convergence of the new method is quadratic for every $n(2n+1)$ line searches if the objective function is a uniformly convex and suitably smooth function on R^n.

§ 1. Introduction

There is a class of direct methods using one-dimensional line serach for solving nonlinear unconstrained optimization problems, and its typical model is the basic Powell's method (the first method in [1]), others are its modifications somehow[1–5]. In [2] W. I. Zangwill proposed a modified version of the basic Powell's method, which is called Powell–Zangwill's method, and proved that it possesses the properties of quadratic termination and global convergence for strictly convex function with bounded level set. In [5] P. L. Toint and F. M. Callier proposed an algorithm model, a generalization of the basic Powell's method and Powell–Zangwill's method and showed that it has Q–linear rate of convergence for uniformly convex function provided that the search directions are uniformly nonsingular. They also showed that under some hypotheses for the points generated by the algorithm model, it has n-iteration Q–superlinear rate of convergence[6,7]. However, they could not verify the tenableness of these hypotheses for the basic Powell's method or Powell–Zangwill's method.

The purpose of this paper is to propose a modified version of Powell–Zangwill's method and to estimate its rate of convergence. In section 2, we will describe the new method——PZM method and give some simple properties of this method, such as quadratic termination and global convergence, etc. Then, in sections 3, 4, we will make great effort to estimate the convergence rate of PZM method for uniformly

Manuscript received January 17, 1985. Revised March 17, 1986.

* Department of Statistics and Operations Research, Fudan University, Shanghai, China.

convex function whose Hessian matrix has Lipschitz condition at minimum point. Using the thought of comparing the uniformly convex function with its quadratic approximation[8], we will estimate the difference of two points found by several line searches with respect to the two functions. By these estimations and quadratic termination property of the method, we will show that the rate of convergence of PZM method is Q-linear and n-iteration quadratic.

§ 2. PZM Method

Let us consider the following minimization problem

$$\min_{x\in R^n} f(x).$$

The procedure of PZM method is as follows:

Step 1. Given an initial point x^0, n fixed search directions $e_1, \cdots, e_n$, which are linearly independent, and n nonzero initial variable search directions $p_1^1, \cdots, p_n^1$, set $k=1$.

Step 2. Set $t_0^k = x^{k-1}$, choose $t_1^k \in S(f, t_0^k, p_n^k)$.

Step 3. For $i=1, \cdots, n$, choose $t_{i+1}^k \in S(f, t_i^k, e_i)$.

Step 4. For $i=1, \cdots, n$, choose $t_{n+i+1}^k \in S(f, t_{n+i}^k, p_i^k)$.

Step 5. Set

$$p_i^{k+1} = p_{i+1}^k,\ i=1, \cdots, n-1, \tag{2.1}$$

$$p_n^{k+1} = t_{2n+1}^k - t_1^k, \tag{2.2}$$

$$x^k = t_{2n+1}^k, \tag{2.3}$$

$k=k+1$, and go to Step 2.

The notation $S(f, x, d)$ above is defined by

$$S(f, x, d) = \{x+\alpha' d \mid f(x+\alpha' d) = \min_{\alpha\in R} f(x+\alpha d),\ \alpha' \in R\}. \tag{2.4}$$

The sequence of $\{x^k\}$, $\{t_i^k\}$ ($i=0, \cdots, 2n+1$) and $\{p_1^k, \cdots, p_n^k\}$ are said to be the sequence of iterative points, auxiliary points, and variable search directions, respectively. The process from one iterative point to the next is called one iteration.

The PZM method adds line searches along n linearly independent fixed search directions, and dose not like Powell-Zangwill's method which adds line searches only along one or several alternate coordinate directions. In a way similar to the proof in [2], we can prove that PZM method has the properties of quadratic termination and global convergence for strictly convex function with bounded level set.

For later discussion, following lemmas about PZM method are given.

Lemma 1. (i) *Suppose that $k\geqslant 1$ and $1\leqslant i\leqslant n$. Then*

$$p_i^k=\begin{cases}p_{k+i-1}^1, \ if\ k+i\leqslant n+1,\\ t_{2n+1}^{k+i-n-1}-t_1^{k+i-n-1}, \ otherwise.\end{cases} \tag{2.5}$$

(ii) *Suppose that the objective function f is a continuously defferentiable and strictly convex function. If $p_{i_0}^{k_0}=0$, then $x^{k_0+i_0-n-1}$ is the minimum point of f.*

Proof (i) By (2.1) and (2.2), (2.5) can be derived immediately.

(ii) From Lemma 1. (i), we know that $k_0+i_0>n+1$ and $t_{2n+1}^{k_0+i_0-n-1}=t_1^{k_0+i_0-n-1}$. Since PZM method is a descent method and f is a strictly convex function, it can be shown that $t_i^{k_0+i_0-n-1}=x^{k_0+i_0-n-1}$ for $1\leqslant i\leqslant 2n+1$. Taking notice of the linear independence of the fixed search directions, we are sure that $x^{k_0+i_0-n-1}$ is the minimum point of f.

Lemma 2. *Let Q be a nonsingular $n\times n$ matrix and f a strictly convex function. If we make two parallel applications of PZM method, the first to the function f with initial point x^0, fixed search directions $e_1, \cdots, e_n$ and nonzero initial variable search directions $p_1^1, \cdots, p_n^1$, the second to the function $\bar{f}$, which satisfies $\bar{f}(x)=f(Qx)$, with initial point $Q^{-1}x^0$, fixed search directions $Q^{-1}e_1, \cdots, Q^{-1}e_n$ and nonzero initial variable search directions $Q^{-1}p_1^1, \cdots, Q^{-1}p_n^1$, then the following equations are satisfied for all positive integer k:*

$$\bar{x}^{k-1}=Q^{-1}x^{k-1}, \tag{2.6}$$

$$\bar{p}_i^k=Q^{-1}p_i^k,\ i=1, \cdots, n, \tag{2.7}$$

where the bars distinguish the second application of PZM method.

Proof It is easy to verify these equations by induction. The detail is omitted.

§ 3. Some Estimations

Definition 1. *Let Q be an $n\times n$ positive definit matrix and $d_1, \cdots, d_n$ be nonzero vectors of R^n. We define*

$$\Delta_Q(d_1, \cdots, d_n)=(\det Q)^{\frac{1}{2}}|\det(d_1, \cdots, d_n)|/\prod_{i=1}^{n}(d_i^TQd_i)^{\frac{1}{2}} \tag{3.1}$$

to be the conjugacy of $d_1 \cdots, d_n$ with respect to Q.

Definition 2. *Let f be a function on R^n and $x, d_1, \cdots, d_l\in R^n$, where l is an integer. $S(f, x, d_1, \cdots, d_l)$ is defined to be a set of points found by l successive one-dimensional line searches along the directions $d_1, \cdots, d_l$ with the stating point x with respect to the function f. For convenience, we usually do not distinguish the set $S(f, x, d_1, \cdots, d_l)$ from its point if $S(f, x, d_1, \cdots, d_l)$ has only one point.*

Lemma 3. *Let f be a quadratic function on R^n with positive definite Hessian matrix A and $d_1, \cdots, d_n$ be n nonzero vectors of R^n. If $z_2=S(f, z_1, d_1, \cdots, d_n)$, then*

$$f(z_2)-f(x^*)\leqslant[1-\Delta_A^2(d_1, \cdots, d_n)][f(z_1)-f(x^*)], \tag{3.2}$$

where x^* is the minimum point of f[9].

In this paper, the vector norm and matrix norm are always Euclidean norm and l_2 matrix norm respectively.

Assumption 1. *f is a twice continuously differentiable function on R^n and there exist positive numbers m, M such that for $x, y\in R^n$,*

$$my^Ty\leqslant y^T\nabla^2 f(x)y\leqslant My^Ty. \tag{3.3}$$

Assumption 2. *f is a twice continuously differentiable function on R^n and there exists a Lipschitz number L such that for $x\in R^n$,*

$$\|\nabla^2 f(x)-\nabla^2 f(x^*)\|\leqslant L\|x-x^*\|, \tag{3.4}$$

where x^* is the minimum point of f.

Lemma 4. *Suppose that f satisfies Assumption 1. If $f(z)\leqslant f(y)$, then*

$$\|z-x^*\|\leqslant\sqrt{M/m}\,\|y-x^*\|, \tag{3.5}$$

where x^* is the minimum point of f.

Proof By the second order Talor expansion formula, we have

$$f(x)-f(x^*)=\frac{1}{2}(x-x^*)^T\int_0^1\nabla^2 f(x^*+t(x-x^*))\,dt\,(x-x^*) \tag{3.6}$$

for all $x\in R^n$. In virtue of (3.6), (3.3) and $f(z)\leqslant f(y)$, (3.5) can be obtained.

Lemma 5. *Suppose that f satisfies Assumption 1. If d is a nonzero vector of R^n and $z_2=S(f, z_1, d)$, then*

$$z_2=z_1-(\nabla^T f(z_1)d/d^T\nabla^2 f(z_3)d)d, \tag{3.7}$$

where z_3 lies between z_1 and z_2, and $f(z_3)\leqslant f(z_1)$.

Proof Let λ be real number such that $z_2=z_1+\lambda d$. By mean value theorem,

$$\nabla^T f(z_2)d-\nabla^T f(z_1)d=d^T\nabla^2 f(z_3)(z_2-z_1)=\lambda d^T\nabla^2 f(z_3)d, \tag{3.8}$$

where z_3 lies between z_1 and z_2, Noting that f is a convex function and $f(z_2)\geqslant f(z_1)$, we know $f(z_3)\leqslant f(z_1)$. By (3.8) and $\nabla^T f(z_2)d=0$, (3.7) can be derived.

Lemma 6. *Suppose that f satisfies Assumptions 1, 2 and its Hessian matrix at minimum point x^* is the identity matrix I. Let $\tilde{f}$ be the quadratic approximation of f at x^*, i.e.*

$$\tilde{f}(x)=\frac{1}{2}\|x-x^*\|^2+f(x^*). \tag{3.9}$$

If d, $\tilde{d}$ are unit vectors and $z_2=S(f, z_1, d)$, $\tilde{z}_2=S(\tilde{f}, \tilde{z}_1, \tilde{d})$, then

$$\|z_2-\tilde{z}_2\|\leqslant 2\|z_1-\tilde{z}_1\|+2\|\tilde{z}_1-x^*\|\,\|d-\tilde{d}\|+c_1\|z_1-x^*\|^2, \tag{3.10}$$

where constant c_1 depends only on m, M, L.

Proof Let λ, $\tilde{\lambda}$ be real numbers such that $z_2=z_1+\lambda d$, $\tilde{z}_2=\tilde{z}_1+\tilde{\lambda}\tilde{d}$. By Lemma 5, we have

$$\lambda=-\nabla^T f(z_1)d/d^T\nabla^2 f(z_3)d, \tag{3.11}$$

$$\tilde{\lambda}=-(\tilde{z}_1-x^*)^T\tilde{d}, \tag{3.12}$$

where z_3 satisfies $f(z_3)\leqslant f(z_1)$. By triangle inequality and Schwarz inequality, we have

$$\begin{aligned}\|z_2-\tilde{z}_2\| &\leqslant \|z_1-\tilde{z}_1\| + \|\lambda d-\tilde{\lambda}\tilde{d}\| \\ &\leqslant \|z_1-\tilde{z}_1\| + \|\lambda d+(z_1-x^*)^T dd\| + \|-(z_1-x^*)^T dd+(\tilde{z}_1-x^*)^T dd\| \\ &\quad + \|-(\tilde{z}_1-x^*)^T dd+(\tilde{z}_1-x^*)^T \tilde{d}d\| + \|-(\tilde{z}_1-x^*)^T \tilde{d}d+(\tilde{z}_1-x^*)^T \tilde{d}\tilde{d}\| \\ &\leqslant 2\|z_1-\tilde{z}_1\| + 2\|\tilde{z}_1-x^*\|\|d-\tilde{d}\| + |\lambda+(z_1-x^*)^T d|. \end{aligned} \tag{3.13}$$

Taking notice of (3.11) and $\nabla f(x^*)=0$, and using mean value theorem, we also have

$$\begin{aligned}\lambda+(z_1-x^*)^T d &= \frac{1}{d^T\nabla^2 f(z_3)d}\{-\nabla^T f(z_1)d+\nabla^T f(x^*)d+(z_1-x^*)^T d \\ &\quad +(d^T\nabla^2 f(z_3)d-1)(z_1-x^*)^T d\} \\ &= \frac{1}{d^T\nabla^2 f(z_3)d}\{-(z_1-x^*)^T(\nabla^2 f(z_4)-I)d \\ &\quad +(z_1-x^*)^T dd^T(\nabla^2 f(z_3)-I)d\},\end{aligned} \tag{3.14}$$

where z_4 lies between z_1 and x^*. By Assumptions 1, 2 and Lemma 4

$$\begin{aligned}|\lambda+(z_1-x^*)^T d| &\leqslant \frac{1}{m} L\|z_1-x^*\|[\|z_4-x^*\|+\|z_3-x^*\|] \\ &\leqslant \frac{2L}{m}\sqrt{M/m}\,\|z_1-x^*\|^2.\end{aligned} \tag{3.15}$$

Combining (3.13) and (3.15), and setting $c_1=2L\sqrt{M/m}/m$, we obtain (3.10).

Lemma 7. *Suppose that f and $\tilde{f}$ satisfy the conditions of Lemma 6, and suppose that*

(i) $f(z_1)\leqslant f(z_0)$, $z_1\neq x^*$, $\tilde{z}_1\neq x^*$;

(ii) $z_2=S(f, z_1, e_1, \cdots, e_n)$, $\tilde{z}_2=S(\tilde{f}, \tilde{z}_1, e_1, \cdots, e_n)$;

(iii) $f(z_4)\leqslant f(z_3)\leqslant f(z_2)$, $\tilde{f}(\tilde{z}_4)\leqslant \tilde{f}(\tilde{z}_3)\leqslant \tilde{f}(\tilde{z}_2)$;

(iv) $d=z_3-z_1$, $\tilde{d}=\tilde{z}_3-\tilde{z}_1$;

(v) $z_5=S(f, z_4, d)$, $\tilde{z}_5=S(\tilde{f}, \tilde{z}_4, \tilde{d})$;

where $e_1, \cdots, e_n$ are linearly independent. Then, we have

(i) *d, $\tilde{d}$ are nonzero vectors, and*

$$\|d/\|d\|-\tilde{d}/\|\tilde{d}\|\|\leqslant c_2(\|z_1-\tilde{z}_1\|+\|z_3-\tilde{z}_3\|)/\|\tilde{z}_1-x^*\|. \tag{3.16}$$

(ii) $\|z_5-\tilde{z}_5\|\leqslant 2\|z_4-\tilde{z}_4\|+2c_2(\|z_1-\tilde{z}_1\|+\|z_3-\tilde{z}_3\|)+c_3\|z_0-x^*\|$, (3.17)

where constant c_2 depends only on the conjugacy of $e_1, \cdots, e_n$ with respect to I, and $c_3=(M/m)c_1$.

Proof (i) By hypotheses (i), (ii) and (iii), we know

$$f(z_3)\leqslant f(z_2)<f(z_1), \tag{3.18}$$

$$\tilde{f}(\tilde{z}_3)\leqslant \tilde{f}(\tilde{z}_2)<\tilde{f}(\tilde{z}_1), \tag{3.19}$$

hence d, $\tilde{d}$ are nonzero. By triangle inequality, we have

$$\begin{aligned}\|d/\|d\|-\tilde{d}/\|\tilde{d}\|\| &\leqslant \|d/\|d\|-d/\|\tilde{d}\|\|+\|d/\|\tilde{d}\|-\tilde{d}/\|\tilde{d}\|\| \\ &\leqslant 2\|d-\tilde{d}\|/\|\tilde{d}\|\leqslant 2(\|z_1-\tilde{z}_1\|+\|z_3-\tilde{z}_3\|)/\|\tilde{d}\|.\end{aligned} \tag{3.20}$$

On the other hand, by Lemma 3

$$2\|\tilde{z}_3-x^*\|^2\leqslant 2\|\tilde{z}_2-x^*\|^2=\tilde{f}(\tilde{z}_2)-\tilde{f}(x^*)$$
$$\leqslant[1-\Delta_I^2(e_1, \cdots, e_n)][\tilde{f}(\tilde{z}_1)-\tilde{f}(x^*)]$$
$$=2[1-\Delta_I^2(e_1, \cdots, e_n)]\|\tilde{z}_1-x^*\|^2. \tag{3.21}$$

Hence, if set $c_2'=1-[1-\Delta_I^2(e_1, \cdots, e_n)]^{1/2}$, we have

$$\|\tilde{d}\|\geqslant\|\tilde{z}_1-x^*\|-\|\tilde{z}_3-x^*\|\geqslant c_2'\|\tilde{z}_1-x^*\|. \tag{3.22}$$

Combining (3.20) and (3.22), and setting $c_2=2/c_2'$, we get (3.16).

(ii) Taking notice of $\tilde{f}(\tilde{z}_4)\leqslant\tilde{f}(\tilde{z}_1)$ and $f(z_4)\leqslant f(z_0)$, we have

$$\|\tilde{z}_4-x^*\|\leqslant\|\tilde{z}_1-x^*\|, \tag{3.23}$$

$$\|z_4-x^*\|\leqslant\sqrt{M/m}\,\|z_0-x^*\|c \tag{3.24}$$

By Lemma 6 and (3.16), we can get (3.17) immediately.

Lemma 8. *Suppose that f and $\tilde{f}$ satisfy the conditions of Lemma 6. Let l be a positive integer and $d_1, \cdots, d_l$ be nonzero vectors. If*

(i) $f(z_1)\leqslant f(z_0)$,

(ii) $z_2=S(f, z_1, d_1, \cdots, d_l)$, $\tilde{z}_2=S(\tilde{f}, \tilde{z}_1, d_1, \cdots, d_l)$,

then

$$\|z_2-\tilde{z}_2\|\leqslant 2^l(\|z_1-\tilde{z}_1\|+c_3\|z_0-x^*\|^2). \tag{3.25}$$

Proof Let $y_1=z_1$, $\tilde{y}_1=\tilde{z}_1$, $y_{i+1}=S(f, y_i, d_i)$ and $\tilde{y}_{i+1}=S(\tilde{f}, \tilde{y}_i, d_i)$ for $1\leqslant i\leqslant l$. By Lemma 6 and taking notice of $f(y_i)\leqslant f(z_0)$, we have

$$\|y_{i+1}-\tilde{y}_{i+1}\|\leqslant 2\|y_i-\tilde{y}_i\|+c_1\|y_i-x^*\|^2\leqslant 2\|y_i-\tilde{y}_i\|+c_3\|z_0-x^*\|^2$$
$$\leqslant 2^i\|y_1-\tilde{y}_1\|+(2^{i-1}+2^{i-2}+\cdots+1)c_3\|z_0-x^*\|^2, \tag{3.26}$$

therefore, (3.25) holds.

§ 4. Convergence Rate

Theorem 1. *Suppose that f satisfies Assumptions 1, 2 and A is its Hessian matrix at minimum point x^*. If the sequence of iterative points $\{x^k\}$ generated by PZM method for function f is not finite terminate, then the sequence $\{x^k\}$ converges to x^* with Q-linear convergence rate and*

$$\overline{\lim_{k\to\infty}}\left[\frac{(x^{k+1}-x^*)^TA(x^{k+1}-x^*)}{(x^k-x^*)^TA(x^k-x^*)}\right]^{\frac{1}{2}}\leqslant[1-\Delta_A^2(e_1, \cdots, e_n)]^{\frac{1}{2}} \tag{4.1}$$

holds, where $e_1, \cdots, e_n$ are the fixed search directions of the method.

Proof If we let Q be the square root of A^{-1}, then I is the Hessian matrix at minimum point of function $\bar{f}$, which satisfies $\bar{f}(x)=f(Qx)$. So by Lemma 2 we can assume that $A=I$ without loss of generality. Let $\tilde{f}$ be the quadratic approximation of f at point x^*, and let

$$\tilde{x}^{k+1}=S(\tilde{f}, x^k, p_n^{k+1}, e_1, \cdots, e_n, p_1^{k+1}, \cdots, p_n^{k+1}). \tag{4.2}$$

Using Lemma 8 with $z_0=x^k$, we have

$$\|x^{k+1}-\tilde{x}^{k+1}\|\leqslant 2^{2n+1}c_3\|x^k-x^*\|^2. \tag{4.3}$$

On the other hand, using Lemma 3 and taking notice of the descent property of one-dimensional line search, we have

$$\frac{1}{2}\|\tilde{x}^{k+1}-x^*\|^2=\tilde{f}(\tilde{x}^{k+1})-\tilde{f}(x^*)\leqslant[1-\Delta_I^2(e_1,\cdots,e_n)][\tilde{f}(x^k)-\tilde{f}(x^*)]$$
$$=\frac{1}{2}[1-\Delta_I^2(e_1,\cdots,e_n)]\|x^k-x^*\|^2. \tag{4.4}$$

In virtue of (4.3), (4.4), we obtain

$$\|x^{k+1}-x^*\|\leqslant\|\tilde{x}^{k+1}-x^*\|+\|x^{k+1}-\tilde{x}^{k+1}\|$$
$$\leqslant[1-\Delta_I^2(e_1,\cdots,e_n)]^{\frac{1}{2}}\|x^k-x^*\|+2^{2n+1}c_3\|x^k-x^*\|^2. \tag{4.5}$$

Since function satisfying Assumption 1 is a strictly convex function with bounded level set, the sequence $\{x^k\}$ converges to x^*. Hence, by (4.5), we get (4.1).

Theorem 2. *Suppose that f satisfies Assumptions 1, 2 and A is its Hessian matrix at minimum point x^*. If the sequence of iterative points $\{x^k\}$ generated by PZM method for function f is not finite terminate, then the sequence $\{x^k\}$ is n-iterations quadratic convergent to x^*, that is*

$$\|x^{k+n}-x^*\|/\|x^k-x^*\|^2\leqslant c_4, \tag{4.6}$$

where constant c_4 depends only on m, M, L, n and $\Delta_A(e_1,\cdots,e_n)$, $e_1,\cdots,e_n$ are the fixed search directions.

Proof In a way similar to the proof of (4.1), we can assume that $A=I$. Moreover, because any iterative point x^k can be regarded as an initial point, we may prove (4.6) only for $k=0$.

Let $\tilde{f}$ be the quadratic approximation of f at x^*. With the same initial point and search directions as the application of PZM method to f, we make an application of PZM method to $\tilde{f}$ and denote by $\{\tilde{x}_i^k\}$, $\{\tilde{t}_i^k\}$ $(i=0,\cdots,2n+1)$ and $\{\tilde{p}_1^k,\cdots,\tilde{p}_n^k\}$ the sequence of iterative points, auxiliary points and variable search directions respectively. Since $\{\tilde{x}^k\}$ is quadratic terminate in at most n iterations, we can assume that $\tilde{t}_{i_0}^{j_0}$ is the first auxiliary point which equals to x^*, of course, $1\leqslant j_0\leqslant n$, $1\leqslant i_0\leqslant 2n+1$.

(i) We first prove it by induction that

$$s_j=\max_{0\leqslant i\leqslant 2n+1}\|t_i^j-\tilde{t}_i^j\|\leqslant(c_5)^jc_3\|x^0-x^*\|^2,\ j=1,\cdots,j_0, \tag{4.7}$$

where $c_5=2^{2n+1}(2+4c_2)$, c_2, c_3 are constants in Lemma 7.

For $j=1$, let

$$d_i=\begin{cases}p_n^1, & i=1,\\ e_{i-1}, & 2\leqslant i\leqslant n+1,\\ p_{i-n-1}^1, & n+2\leqslant i\leqslant 2n+1.\end{cases} \tag{4.8}$$

According to the process of PZM method

$$t_i^1=S(f,x^0,d_1,\cdots,d_i), \tag{4.9}$$
$$\tilde{t}_i^1=S(\tilde{f},x^0,d_1,\cdots,d_i). \tag{4.10}$$

Using Lemma 8 with $z_0=x^0$, we have

$$\|t_i^1-\tilde{t}_i^1\|\leqslant 2^i c_3\|x^0-x^*\|^2,\ i=1,\ \cdots,\ 2n+1. \tag{4.11}$$

Therefore (4.7) holds for $j=1$.

Now assume that (4.7) holds for $1\leqslant j\leqslant r$. If $r=j_0$, the induction proof is complete. If $r<j_0$, by Lemma 1,

$$p_i^{r+1}=p_{i+r}^1=\tilde{p}_{i+r}^1=\tilde{p}_i^{r+1}\quad(1\leqslant i\leqslant n-r), \tag{4.12}$$

$$\tilde{p}_i^{r+1}=t_{2n+1}^{r+i-n}-t_1^{r+i-n}\quad(n-r+1\leqslant i\leqslant n), \tag{4.13}$$

$$\tilde{p}_i^{r+1}=\tilde{t}_{2n+1}^{r+i-n}-\tilde{t}_1^{r+i-n}\quad(n-r+1\leqslant i\leqslant n). \tag{4.14}$$

We estimate $\|t_i^{r+1}-\tilde{t}_i^{r+1}\|$ by two parts.

(a) Considering the first and last r line searches in $(r+1)$-th iteration, we have

$$t_1^{r+1}=S(f,\ t_0^{r+1},\ p_n^{r+1}), \tag{4.15}$$

$$\tilde{t}_1^{r+1}=S(\tilde{f},\ \tilde{t}_0^{r+1},\ \tilde{p}_n^{r+1}), \tag{4.16}$$

and

$$t_{n+i+1}^{r+1}=S(\tilde{f},\ t_{n+i}^{r+1},\ p_i^{r+1})\quad(n-r+1\leqslant i\leqslant n), \tag{4.17}$$

$$\tilde{t}_{n+i+i}^{r+1}=S(\tilde{f},\ \tilde{t}_{n+i}^{r+1},\ \tilde{p}_i^{r+1})\quad(n-r+1\leqslant i\leqslant n). \tag{4.18}$$

In order to use Lemma 7, we set $z_0=x^0$, $z_1=t_1^r$, $\tilde{z}_1=\tilde{t}_1^r$, $z_3=t_{2n+1}^r$, $\tilde{z}_3=\tilde{t}_{2n+1}^r$, $z_4=t_0^{r+1}$, $\tilde{z}_4=\tilde{t}_0^{r+1}$ for (4.15), (4.16), and set $z_0=x^0$, $z_1=t_1^{r+i-n}$, $\tilde{z}_1=\tilde{t}_1^{r+i-n}$, $z_3=t_{2n+1}^{r+i-n}$, $\tilde{z}_3=\tilde{t}_{2n+1}^{r+i-n}$, $z_4=t_{n+i}^{r+1}$, $\tilde{z}_4=t_{n+i}^{r+1}$ for (4.17), (4.18). Since $\{x^k\}$ is not finite terminate and $r<j_0$, we know that $z_1\neq x^*$ and $\tilde{z}_1\neq x^*$. By the structure of the method and (4.13) (4.14), we can easily verify that the conditions of Lemma 7 are satisfied. Hence, by (3.17) of Lemma 7 and (4.7) for $j\leqslant r$, we have

$$\begin{aligned}\|t_1^{r+1}-\tilde{t}_1^{r+1}\|&\leqslant 2\|t_0^{r+1}-\tilde{t}_0^{r+1}\|+2c_2(\|t_1^r-\tilde{t}_1^r\|+\|t_{2n+1}^r-\tilde{t}_{2n+1}^r\|)+c_3\|x^0-x^*\|^2\\&\leqslant 2\|t_0^{r+1}-\tilde{t}_0^{r+1}\|+[4c_2(c_5)^r+1]c_3\|x^0-x^*\|^2,\end{aligned} \tag{4.19}$$

and

$$\begin{aligned}\|t_{n+i+1}^{r+1}-\tilde{t}_{n+i+i}^{r+1}\|&\leqslant 2\|t_{n+i}^{r+1}-\tilde{t}_{n+i}^{r+1}\|+2c_2(\|t_1^{r+i-n}-\tilde{t}_1^{r+i-n}\|\\&\quad+\|t_{2n+1}^{r+i-n}-\tilde{t}_{2n+1}^{r+i-n}\|)+c_3\|x^0-x^*\|^2\\&\leqslant 2\|t_{n+i}^{r+1}-\tilde{t}_{n+i}^{r+1}\|+[4c_2(c_5)^{r+i-n}+1]c_3\|x^0-x^*\|^2\\&\qquad(n-r+1\leqslant i\leqslant n).\end{aligned} \tag{4.20}$$

(b) Considering the other line searches of $(r+1)$-th iteration, if set

$$d_i=\begin{cases}e_i, & 1\leqslant i\leqslant n,\\ p_{i-n}^{r+1}, & n+1\leqslant i\leqslant 2n-r,\end{cases} \tag{4.21}$$

we have

$$t_{i+1}^{r+1}=S(f,\ t_i^{r+1},\ d_i)\qquad(1\leqslant i\leqslant 2n-r), \tag{4.22}$$

$$\tilde{t}_{i+1}^{r+1}=S(\tilde{f},\tilde{t}_i^{r+1},\ d_i)\qquad(1\leqslant i\leqslant 2n-r). \tag{4.23}$$

Using Lemma 8 with $z_0=x^0$, we get

$$\|t_{i+1}^{r+1}-\tilde{t}_{i+1}^{r+1}\|\leqslant 2\|t_i^{r+1}-\tilde{t}_i^{r+1}\|+2c_3\|x^0-x^*\|^2 \tag{4.24}$$

for $1\leqslant i\leqslant 2n-r$.

Combining (4.19), (4.20), (4.24) we have

$$\begin{aligned}\|t_{i+1}^{r+1}-\tilde{t}_{i+1}^{r+1}\| &\leqslant 2\|t_i^{r+1}-\tilde{t}_i^{r+1}\|+(4c_2+1)(c_5)^r c_3\|x^0-x^*\|^2 \\ &\leqslant 2^{i+1}\|t_0^{r+1}-\tilde{t}_0^{r+1}\|+(2^i+\cdots+1)(4c_2+1)c_5^r c_3\|x^0-x^*\|^2 \\ &\leqslant 2^{i+1}[\|t_0^{r+1}-\tilde{t}_0^{r+1}\|+(4c_2+1)c_5^r c_3\|x^0-x^*\|^2]. \end{aligned} \tag{4.25}$$

Noting $t_0^{r+1}=t_{2n+1}^r$, $\tilde{t}_0^{r+1}=\tilde{t}_{2n+1}^r$ and inequality (4.7) for $j=r$, we can prove that (4.7) holds for $j=r+1$.

So far the induction proof is complete.

(ii) Since $j_0\leqslant n$, we have

$$f(x^n)=f(t_{2n+1}^n)\leqslant f(t_{i_0}^{j_0}), \tag{4.26}$$

By Lemma 4

$$\|x^n-x^*\|\leqslant\sqrt{M/m}\,\|t_{i_0}^{j_0}-x^*\|. \tag{4.27}$$

Taking notice of that $\tilde{t}_{i_0}^{j_0}=x^*$ and inequality (4.7), we have

$$\|x^n-x^*\|\leqslant\sqrt{M/m}\,\|t_{i_0}^{j_0}-\tilde{t}_{i_0}^{j_0}\|\leqslant\sqrt{M/m}\,(c_5)^{j_0}c_3\|x^0-x^*\|^2. \tag{4.28}$$

Hence, if set $c_4=M/m(c_5)^n\, c_3$, we get (4.6).

Remark. Suppose that f satisfys the inequalities (3.1) and (3.4) for all points in an open neighbourhood of point x^*, where x^* is not assumed to be the minimum point of f. If the sequences of auxiliary points $\{t_i^k\}$ ($i=0, 1, \cdots, 2n+1$) generated by PZM method for function f are convergent to x^* and not finite terminate, then x^* is a local minimum point of f and inequalities (4.1) of Theorem 1 and (4.6) of Theorem 2 also hold.

I would like to thank Professor W. Yu for his guidance from the bottom of my heart.

References

[1] Powell, M. J. D., An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comp. J.*, **7** (1964), 155—162.

[2] Zangwill, W. I., Minimizating a function without calculating derivatives, *Comp. J.*, **10** (1967), 293—296.

[3] Brent, P. R., Algorithm for minimization without derivatives, Prentice-Hall, *Englewood Cliffs, N. J.*, (1973).

[4] Sargent, R. W. H., Minimization without constraints, *Optimization and Design*, M. Avriel, etc. (Eds), Prentice-Hall, Englewood Cliffs, N. J., (1973).

[5] Toint, P. L. and Callier, F. M., On the uniform nonsingularity of matrices of search directions and the rate of convergence in minimization algorithms, *J. O. T. A.*, **23** (1977), 511—529.

[6] Toint, P. L. and Callier, F. M., On the accelerating property of an algorithm for function minimization without calculating derivatives, *J. O. T. A.*, **23** (1977), 531—547.

[7] Toint, P. L. and Callier, F. M., On the accelerating property of an algorithm for function minimization without calculating derivatives (Errata Corrige), *J. O. T. A.*, **26** (1978), 465—467.

[8] Cohen, A. I., Rate of convergence of several conjugate gradient algorithms, *SIAM. J. Numer. Anal.*, **9** (1972), 248—259.

[9] Deng, N. Y. and Zu, M. F., Research on the theory of Powell's method, *Kexue Tongbao*, **10** (1979), 433—437 (in chinese).