系统科学与数学
J. Sys. Sci. & Math. Scis.
11(3) (1991), 233—244

一个具有 n 步二次收敛性的直接法算法

徐以汛
(山东轻工业学院数学教研室，济南250100)

一、引言

在非线性最优化的直接法算法中，Powell 算法[2]具有一定的代表性，但我们已知 Powell 算法对正定二次函数一般不具有二次终结性，为此围绕着改善 Powell 算法的算法特性，出现了一系列 Powell 算法的改进型[3—6]. 其中俞文魮教授提出了一个新的方法——PY 算法[1]，其基本思想基于，对正定二次的目标函数算法迭代过程中每一轮迭代的搜索方向组的共轭性程度，比起它上一轮迭代的搜索方向组的共轭性程度是不减的，因此算法具有对正定二次函数的 n 步二次终结性. 但是对非二次正定函数，例如一致凸函数，算法是否仍具有二次收敛性，这是一个有待解决的问题. 本文的工作是对 PY 方法进行修改，使算法不仅保持有对目标函数为正定二次函数的 n 步二次终结性，而且对一致凸的函数具有 n 步二次收敛性.

二、算法搜索方向组的修改过程

设在算法的第 i 步，起始点为 x，搜索方向组为 $d_1,\cdots,d_{n-i},d_{n-i+1},\cdots,d_n$. 我们用记号 $y=S(x;d)$ 表示目标函数以 x 为起点，沿方向 d 进行线搜索而得到点 y. 并记 $S(x;d_1,d_2)=S(S(x;d_1);\ d_2)$. 定义 $y_0=x$, $y_j=S(y_{j-1};d_j)$, $j=1,2,\cdots,n-i$, $y=y_{n-i}$, $a=y-x$, $z=S(y;d_{n-i+1},\cdots,d_n)$，那么

$$y_j=y_{j-1}+\lambda_j d_j,\ j=1,2,\cdots,n-i, \tag{2.1}$$

$$a=y-x=\sum_{j=1}^{n-i}\lambda_j d_j;\quad z-x=\sum_{j=1}^{n}\lambda_j d_j. \tag{2.2}$$

设 $a\neq 0$，并定义

$$l_j=\begin{cases}\lambda_j, & \lambda_j\geqslant\|a\|,\\ \|a\|, & \lambda_j<\|a\|,\end{cases}\quad j=1,2,\cdots,n-i, \tag{2.3}$$

$$\beta_j=l_j[2(f(y_j-l_jd_j)-f(y_j))]^{-\frac{1}{2}},\quad j=1,2,\cdots,n-i, \tag{2.4}$$

$$v_j=\beta_j d_j,\quad j=1,2,\cdots,n-i, \tag{2.5}$$

得到向量组

$$\{v_1,\cdots,v_{n-i},d_{n-i+1},\cdots,d_n\}=\{V,D'\}. \tag{2.6}$$

1989 年 4 月 28 日收到.

易知当 $a \doteq 0$，且函数 f 一致凸时，$\|v_j\| > 0 \ \forall\ j$. 作 n 阶正交矩阵

$$A_{jk} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos\theta_{jk} & & & \sin\theta_{jk} & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & \sin\theta_{jk} & & & \cos\theta_{jk} & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ j \\ \\ \\ \\ k, \\ \\ \\ \\ \end{matrix} \tag{2.7}$$

其中，

$$\cos\theta_{jk} = \frac{\langle (a_j, a_k)^T, (v_{j1} v_{k1})^T \rangle}{\|(a_j a_k)^T\| \cdot \|(v_{j1} v_{k1})^T\|}, \quad 1 \leqslant j < k \leqslant n, \tag{2.8}$$
$$(a_1, \cdots, a_n)^T = a, \ (v_{11}, v_{21}, \cdots, v_{n1})^T = v_1.$$

并令

$$A = \prod_{1 \leqslant j < k \leqslant n} A_{jk}, \tag{2.9}$$

$$R = (V^T V)^{-1} V^T A V, \tag{2.10}$$

那么

$$VR = AV, \tag{2.11}$$

对 R 按 Gram-Schmidt 过程正交化:

$$b_k = r_k - \sum_{j=1}^{k-1} \langle r_j, q_j \rangle q_j,$$
$$q_k = b_k / \|b_k\|, \ k = 1, 2, \cdots, n - i, \tag{2.12}$$

得正交矩阵 $Q = (q_1, \cdots, q_{n-i})$. 过程(2.12)等价于对 R 右乘以某一上三角正线矩阵 Π，即 $R\Pi = Q$. 所以

$$VQ = VR\Pi = AV\Pi = U = (u_1, \cdots, u_{n-i}), \tag{2.13}$$
$$u_1 \| a.$$

将 $u_j (j = 1, 2, \cdots, n - i)$ 单位化，得向量组

$$\bar{D} = (\bar{d}_1, \cdots, \bar{d}_n), \tag{2.14}$$

其中，

$$\bar{d}_j = u_{j+1} / \|u_{j+1}\|, \ j = 1, 2, \cdots, n - i - 1,$$
$$\bar{d}_j = d_{j+1}, \qquad j = n - i, \cdots, n - 1,$$
$$\bar{d}_n = (z - x) / \|z - x\|.$$

并令

$$\bar{x} = S(z; \bar{d}_n). \tag{2.15}$$

我们称由 x 和 $D = \{d_1, \cdots, d_n\}$ 经过 (2.1)—(2.15) 而得到 $\bar{x}$ 和 $\bar{D} = \{\bar{d}_1, \cdots, \bar{d}_n\}$ 的过程为算法的第 i 步操作过程，并用映射 T_i 表示.

定义 2.1. $(\bar{x};\bar{d}_1,\cdots,\bar{d}_n)=T_i(x;d_1,\cdots,d_n)$.

三、算法的共轭性程度

定理 3.1. 设 f 为正定二次函数，$\|d_j\|=1\quad(j=1,2,\cdots,n-i)$，那么

i) $v_j(j=1,2,\cdots,n-i)$ 为关于 $\nabla^2 f$ 的单位向量，

ii) β_j 与 l_j 的取值无关.

证明从略.

假定

(A1) $f\in C^2$，存在常数 $m,M>0$，使得对任意 $x,y\in R^n$，有

$$my^Ty\leqslant y^T\nabla^2 f(x)y\leqslant My^Ty. \tag{3.1}$$

(A2) $f\in C^2$，存在常数 $L>0$，使得对任意 $x\in R^n$，有

$$\|\nabla^2 f(x)-\nabla^2 f(x^*)\|\leqslant L\|x-x^*\|, \tag{3.2}$$

其中 x^* 是 $f(x)$ 的极小值点.

定理 3.2. 若 f 满足假设(A1)，(A2)，$\|d_j\|=1\ (j=1,2,\cdots,n-i)$，那么对 $j=1,2,\cdots,n-i$ 有

$$M^{-2}\leqslant\beta_j\leqslant m^{-2}. \tag{3.3}$$

证明从略.

向量组 $\{d_1,\cdots,d_n\}$ 关于 n 阶正定矩阵 G 的共轭性程度定义为

$$\Delta_G(d_1,\cdots,d_n)=(\det G)^{\frac{1}{2}}\cdot|\det(d_1,\cdots,d_n)|\cdot\prod_{i=1}^{n}(d_i^TGd_i)^{-\frac{1}{2}}. \tag{3.4}$$

定理 3.3. 设目标函数 f 为正定二次函数，$\{d_1,\cdots,d_n\}$ 为一独立向量组，$\{d_{n-i+1},\cdots,d_n\}$ 关于 $\nabla^2 f$ 互相共轭，如果 $y=S(x;d_1,\cdots,d_{n-i})$，并且

$$(\bar{x};\bar{d}_1,\cdots,\bar{d}_n)=T_i(x;d_1,\cdots,d_n),$$

那么

$$\Delta_{\nabla^2 f}(\bar{d}_1,\cdots,\bar{d}_n)\geqslant\Delta_{\nabla^2 f}(d_1,\cdots,d_n). \tag{3.5}$$

参见[1,2,7]本定理易证，证明从略.

四、算法模型及二次终结性

$$\text{(P)}\qquad \min\,[f(x)\,|\,x\in R^n] \tag{4.1}$$

算法 (Ag):

i) 给定点 x^0 和单位向量组 $\{d_1^1,\cdots,d_n^1\}$，令 $x^1:=S(x^0;d_n^1)$，$k:=1$.

ii) 下面的过程从 $i=1$ 进行到 $i=n-1$，称为算法的第 k 轮，

$$(x^{i+1};d_1^{i+1},\cdots,d_n^{i+1})=T_i(x^i;d_1^i,\cdots,d_n^i), \tag{4.2}$$

并称(4.2)为算法第 k 轮的第 i 步. 如果存在某一 i 使得

$$x^i=S(x^i;d_1^i,\cdots,d_{n-i}^i), \tag{4.3}$$

则令第 k 轮终止，$i[k]:=i$；否则第 k 轮在 $(n-1)$ 步后终止，$i[k]:=n$.

iii) 如果 $i[k]<n$，$x^1:=x^{i[k]}$，$\{d_1^1,\cdots,d_n^1\}:=\{d_1^{i[k]},\cdots,d_n^{i[k]}\}$；否则 $i[k]=n$ 时，$x^1:=x^n$，$\{d_1^1,\cdots,d_n^1\}:=\{d_1^{n-1},\cdots,d_n^{n-1}\}$；$k:=k+1$ 转回 ii)，进行下一轮迭代.

注 1° 算法本身并没有给出终止准则.

注 2° 算法一般只适于一致凸函数，这是为了保证(2.4)式 $\beta_i>0$.

定理 4.1. 设 f 为正定二次函数，x^* 为极小值点，那么对于算法(Ag)，下述结论成立：

i) 在第一轮终止时，已得到了 x^*；

ii) $\Delta(d_1^i,\cdots,d_n^i)\leqslant\Delta(d_1^{i+1},\cdots,d_n^{i+1})$； (4.4)

iii) 当 $i[1]=n$ 时，$\Delta(d_1^{n-1},\cdots,d_n^{n-1})=1$. (4.5)

证. 结论 ii) 显然是定理 3.3 的推论.

现证结论 i). 设算法第一轮在第 i 步时终止，$i\leqslant n$. 由(4.3)，$x^i=S(x^i;d_1^i,\cdots,d_{n-i}^i)$，得

$$(\nabla f(x^i))^T d_j^i=0,\quad j=1,2,\cdots,n-i. \tag{4.6}$$

另一方面，根据 (Ag) ii) 的(4.2)和(2.14)，可知 $\{d_{n-i+2}^{i-1}=d_{n-i+1}^i,\cdots,d_n^{i-1}=d_{n-1}^i,d_n^i\}$ 互相共轭，而 x^i 是在前一步经这些方向的逐次线搜索而得，所以根据扩张子空间定理[9]得

$$(\nabla f(x^i))^T d_j^i=0,\quad j=n-i+1,\cdots,n. \tag{4.7}$$

再根据(4.4)，$\Delta_i\geqslant\Delta_1>0$，综合(4.6)，(4.7)便得结论.

当 $i[1]=n$ 时，根据 (Ag) ii) 轮次终止准则，得 $x^i-x^{i-1}\neq 0$ $(i=1,2,\cdots,n)$. 所以由引理 3.1 可知$\{d_1^{n-1},\cdots,d_n^{n-1}\}$互相共轭，这便证明了结论 iii). 证毕.

五、二次收敛性预备引理

设本节所讨论的函数均满足假设 (A1)，(A2). 不失普遍性，可设函数的极小值点 $x^*=0$，极小值 $f(x^*)=0$. 并记 $g=\nabla f$，$h=\nabla^2 f$，$C=\nabla^2 f(0)$. 定义函数

$$\tilde{f}(x)=\frac{1}{2}x^T Cx, \tag{5.1}$$

称之为对应于 $f(x)$ 的二次逼近函数. 我们将在所有的参数上冠以"～"，表示以 $\tilde{f}$ 为目标时相应的算法过程中的参数.

引理 5.1. i) $|f(x)-\tilde{f}(x)|\leqslant\frac{\Lambda}{6}\|x\|^3$； (5.2)

ii) $\|g(x)-\tilde{g}(x)\|\leqslant\frac{\Lambda}{2}\|x\|^2$； (5.3)

iii) $\|h(x)-\tilde{h}(x)\|\leqslant\Lambda\|x\|$， (5.4)

其中 Λ 为正实数.

证明从略.

引理 5.2 如果 $f(x^1)\leqslant f(x^2)$，那么

$$\|x^1\|\leqslant\left(\frac{M}{m}\right)^{\frac{1}{2}}\|x^2\|. \tag{5.5}$$

证明从略.

引理 5.3. 如果 $x^2 = S(x^1;d)$, $x^2 = x^1 + \lambda d$,那么

$$\lambda = -g(x^1)^T d / d^T h(x^1 + \theta d)d,\ 0 \leqslant |\theta| \leqslant |\lambda|. \tag{5.6}$$

证. $-g(x^1)^T d = g(x^2)^T d - g(x^1)^T d = (g(x^2) - g(x^1))^T d$

$$= \int_0^\lambda d^T h(x^1 + \tau d)d \cdot d\tau = \lambda d^T h(x^1 + \theta d)d,\ 0 < |\theta| < |\lambda|,$$

$$\lambda = -g(x^1)^T d / d^T h(x^1 + \theta d)d.$$

证毕.

引理 5.4. 设 $\bar{x} = S(x; d_1, \cdots, d_n)$,那么

$$\tilde{f}(\bar{x}) \leqslant [1 - \Delta_C^2(d_1, \cdots, d_n)]\tilde{f}(x). \tag{5.7}$$

证明参见[8].

引理 5.5.

$$\|\tilde{x}^i - \tilde{x}^{i-1}\| \geqslant \xi_1 \|\tilde{x}^{i-1}\|, \tag{5.8}$$

其中 ξ_1 仅与 $m, M, \Delta(d_1^1, \cdots, d_n^1)$ 有关.

证. 由(5.6)式 $\|\tilde{x}^i - \tilde{x}^{i-1}\| = |\lambda| = |\tilde{g}(\tilde{x}^{i-1})^T \tilde{d} / d^T C d|$,而 $\tilde{d} = (\tilde{x}^i - \tilde{x}^{i-1})/\|\tilde{x}^i - \tilde{x}^{i-1}\|$,因此

$$\begin{aligned}\|\tilde{x}^i - \tilde{x}^{i-1}\|^2 &= |(\tilde{x}^{i-1T} C \tilde{x}^i - \tilde{x}^{i-1T} C \tilde{x}^{i-1}) / \tilde{d}^T C \tilde{d}| \\ &= |(\tilde{x}^i - \tilde{x}^{i-1})^T C (\tilde{x}^i - \tilde{x}^{i-1}) + \tilde{x}^{i-1T} C \tilde{x}^{i-1} - \tilde{x}^{iT} C \tilde{x}^i| / 2\tilde{d}^T C \tilde{d} \\ &\geqslant \frac{1}{2M}(\tilde{x}^{i-1T} C \tilde{x}^{i-1} - \tilde{x}^{iT} C \tilde{x}^i).\end{aligned}$$

这是根据假设(A2),再由(5.7)和(3.5)便得结论. 证毕.

引理 5.6.

$$\|\tilde{u}_j^i\| \leqslant M_1,\ \forall i, j, \tag{5.9}$$

其中 M_1 仅与 m, M 有关.

证. 由(2.13) $\|\tilde{u}_j^i\|^2 = \|\tilde{V}^i \tilde{q}_j^i\|^2 = \tilde{q}_j^{iT} \tilde{V}^{iT} \tilde{V}^i \tilde{q}_j^i \leqslant \lambda_{\max}$,这里 $\lambda_{\max}$ 是 $\tilde{V}^{iT}\tilde{V}^i$ 的最大特征值,其余为 $\lambda_2, \cdots, \lambda_{n-i}$. 因此

$$\lambda_{\max} + \sum_{j=2}^{n-i} \lambda_j = \sum_{j=1}^{n-i} \tilde{v}_j^{iT} \tilde{v}_j^i = \sum_{j=1}^{n-i} (\beta_j^i)^2 \leqslant (n-i)m^{-1}.$$

证毕.

引理 5.7.

$$(\Delta(D^1))^2 \leqslant \det(\tilde{D}_i^T \tilde{D}_i) \leqslant 1, \tag{5.10}$$

其中 $\tilde{D}_i = (\tilde{d}_1^i, \cdots, \tilde{d}_n^i)$, $D^1 = (d_1^1, \cdots, d_n^1)$.

证. 设 $D_j' = (\tilde{d}_1^i, \cdots, \tilde{d}_j^i)$, $1 \leqslant j < n$,则有

$$\det (D_j', \tilde{d}_{j+1}^i)^T (D_j', \tilde{d}_{j+1}^i) = \det D_j'^T D_j' (1 - \tilde{d}_{j+1}^{iT} D_j' (D_j'^T D_j')^{-1} D_j'^T \tilde{d}_{j+1}^i).$$

因 $D_j'(D_j'^T D_j')^{-1} D_j'^T$ 的特征值为 0 或为 1,所以

$$\det (D_j', \tilde{d}_{j+1}^i)^T (D_j', \tilde{d}_{j+1}^i) \leqslant \det D_j'^T D_j'$$

对 $j = 1, 2, \cdots, n-1$ 成立. 再根据(3.5)便得结论. 证毕.

引理 5.8. 设 $\tilde{a}^i \neq 0$,则

$$\|\tilde{a}^i\| \geqslant \xi_2 \|\tilde{x}^{i-1}\|, \tag{5.11}$$

其中常数 ξ_2 仅与 $m, M, \Delta(D^1)$ 有关.

证. 设(5.1)中的 $C = P^T P$,由(5.6)经整理得

$$P\tilde{y}_1^i = P\tilde{y}_0^i - \frac{(P\tilde{y}_0^i)^T(P\tilde{d}_1^i)}{\tilde{d}_1^{iT}C\tilde{d}_1^i}P\tilde{d}_1^i = \left(E - \frac{P\tilde{d}_1^i}{\|P\tilde{d}_1^i\|}\cdot\frac{(P\tilde{d}_1^i)^T}{\|P\tilde{d}_1^i\|}\right)P\tilde{y}_0^i.$$

因此

$$P\tilde{y}_{n-i}^i = \left(E - \frac{P\tilde{d}_{n-i}^i}{\|P\tilde{d}_{n-i}^i\|}\cdot\frac{(P\tilde{d}_{n-i}^i)^T}{\|P\tilde{d}_{n-i}^i\|}\right)\cdots\left(E - \frac{P\tilde{d}_1^i}{\|P\tilde{d}_1^i\|}\cdot\frac{(P\tilde{d}_1^i)^T}{\|P\tilde{d}_1^i\|}\right)P\tilde{y}_0^i. \quad (5.12)$$

如果设

$$G_K = (P\tilde{d}_1^i/\|P\tilde{d}_1^i\|, \cdots, P\tilde{d}_{n-K}^i/\|P\tilde{d}_{n-K}^i\|),$$

$$B_K = \prod_{j=1}^{K}\left(E - \frac{P\tilde{d}_j^i}{\|P\tilde{d}_j^i\|}\cdot\frac{(P\tilde{d}_j^i)^T}{\|P\tilde{d}_j^i\|}\right), \quad (5.13)$$

$$S_K = \sum_{j=1}^{K}\alpha_j(P\tilde{d}_j^i), \quad K = 1, 2, \cdots, n-i,$$

那么下面不等式成立，即

$$\max_{z\in S_K, \|z\|=1}\{\|B_K z\|^2\} \leqslant (1 - \det C_K^T C_K), \ K = 1, \cdots, n-i. \quad (5.14)$$

因为 $\{\tilde{d}_{n-i+2}^{i-1}, \cdots, \tilde{d}_n^{i-1}, \tilde{z}_n^{i-1} - \tilde{x}^{i-2}\}$ 互相共轭，所以 $g(\tilde{y}_0^i) \perp L(\tilde{d}_{n-i+2}^{i-1}, \cdots, \tilde{d}_n^{i-1}, \tilde{z}^{i-1} - \tilde{x}^{i-2})$，即

$$P\tilde{y}_0^i \in L(P\tilde{d}_1^i, \cdots, P\tilde{d}_{n-i}^i). \quad (5.15)$$

由(5.14)，(5.12)便得

$$\|P\tilde{y}_{n-i}^i\|^2 \leqslant (1 - \det C_{n-i}^T C_{n-i})\|P\tilde{y}_0^i\|^2. \quad (5.16)$$

而根据(5.10)，$\det C_{n-i}^T C_{n-i} \geqslant \left(\frac{m}{M}\right)^{n-i}(\Delta(D^1))^2$，即可得

$$\tilde{f}(\tilde{y}_{n-i}^i) \leqslant \left[1 - \left(\frac{m}{M}\right)^{n-i}(\Delta(D^1))^2\right]\tilde{f}(\tilde{y}_0^i).$$

如引理 5.5 的证明方法，即可证得

$$\|\tilde{a}^i\| = \|\tilde{y}_{n-i}^i - \tilde{y}_0^i\| \geqslant \xi_2\|\tilde{x}^{i-1}\|.$$

证毕.

引理 5.9.

$$\|\tilde{u}_j^i\| \geqslant m_1, \quad (5.17)$$

其中 m_1 仅与 m, M 有关.

证. 由(2.13)，

$$\|\tilde{u}_j^i\|^2 = \|\tilde{V}^i\tilde{q}_j^i\|^2 \geqslant \lambda_{\min},$$

这里 $\lambda_{\min}$ 是 $\tilde{V}^{iT}\tilde{V}^i$ 的最小特征值，其余为 $\lambda_1, \cdots, \lambda_{n-i-1}$. 由(2.5)，(3.3)和(5.10)，得

$$\lambda_{\min}\cdot\lambda_1\cdots\lambda_{n-i-1} = \det(\tilde{V}^{iT}\tilde{V}^i) = \prod_{j=1}^{n-i}\beta_j^2\det(\tilde{D}^{iT}\tilde{D}^i) \geqslant \left(\frac{\Delta(D^1)}{M}\right)^2.$$

再结合条件 $\lambda_j \leqslant (n-i)m^{-4}$ $(j = 1, 2, \cdots, n-i)$，便得结论. 证毕.

引理 5.10.

$$\|S(f, x; d) - S(\tilde{f}, \tilde{x}; \tilde{d})\| \leqslant \delta_1\|\Delta x\| + \delta_2\|\tilde{x}\|\|\Delta d\| + \delta_3\|\Delta g\| + \delta_4\|\tilde{x}\|^2 + \delta_5\|\tilde{x}\|\|x\|, \quad (5.18)$$

其中 $\delta_1, \cdots, \delta_5$ 仅依赖于 m, M, Λ.

证. $S(f, x; d) - S(\tilde{f}, \tilde{x}; \tilde{d}) = (x - \lambda d) - (\tilde{x} - \tilde{\lambda}\tilde{d}) = \Delta x - \Delta\lambda\cdot d - \tilde{\lambda}\Delta d.$

其中

$$\Delta\lambda \triangleq \tilde{\lambda} - \tilde{\lambda} = (\Delta g)^T d/D + (\tilde{g}(\tilde{x}))^T d/D - \tilde{g}(\tilde{x})^T \tilde{d}/D \cdot \tilde{D} \cdot \Delta D, \tag{5.19}$$

$$\Delta D \triangleq D - \tilde{D} = d^T h(x - \vartheta d) d - \tilde{d}^T C d = \Delta d^T h(x - \theta d) d + \tilde{d}^T \Delta h d + \tilde{d}^T C \Delta d, \tag{5.20}$$

$$\Delta h = h(x - \theta d) - C, \quad 0 \leqslant |\theta| < \lambda.$$

由(5.4),(5.5)知

$$\|\Delta h\| \leqslant \Lambda \|x - \theta d\| \leqslant \Lambda \max\{\|x\|, \|x - \lambda d\|\} \leqslant \Lambda \left(\frac{M}{m}\right)^{\frac{1}{2}} \|x\|, \tag{5.21}$$

而由假设(A1)及 $\|d\| = \|\tilde{d}\| = 1$ 可知

$$\|g(x)\| \leqslant M\|x\|, \ \|\tilde{g}(\tilde{x})\| \leqslant M\|\tilde{x}\|, \tag{5.22}$$

$$\|h(x - \theta d)\| \leqslant M, \ \|D\| \geqslant m, \ \|\tilde{D}\| \geqslant m. \tag{5.23}$$

将(5.23),(5.21)代入(5.20)得

$$|\Delta D| \leqslant 2M\|\Delta d\| + \Lambda\left(\frac{M}{m}\right)^{\frac{1}{2}} \|x\|. \tag{5.24}$$

将(5.23),(5.22),(5.24)代入(5.19)得

$$|\Delta\lambda| \leqslant \frac{1}{m}\|\Delta y\| + \frac{M}{m}\|\tilde{x}\| + \frac{M}{m^2}\|\tilde{x}\|\left[M\|\Delta d\| + \Lambda\left(\frac{M}{m}\right)^{\frac{1}{2}}\|x\| + M\|\Delta d\|\right]. \tag{5.25}$$

综合(5.25),(5.23),(5.22)再经整理可得结论．证毕．

引理 5.11.

$$|\beta_j^i - \tilde{\beta}_j^i| \leqslant \frac{\xi_3}{\|\tilde{x}^{i-1}\|}(|\Delta l_j^i| + \|\Delta y_j^i\|), \tag{5.26}$$

其中 $\Delta l_j^i = l_j^i - \tilde{l}_j^i$, $\Delta y_j^i = y_j^i - \tilde{y}_j^i$, ξ_3仅与 M, m, Λ 有关．

证．由(2.6)经整理可得不等式

$$|\beta_j^i - \tilde{\beta}_j^i| \leqslant \frac{|\Delta l_j^i|}{2|l_j^i|} + \frac{\beta_i|[f(y_i - l_i d_i) - f(y_i)]^{\frac{1}{2}} - [\tilde{f}(\tilde{y}_i - \tilde{l}_i\tilde{d}_i) - \tilde{f}(\tilde{y}_i)]^{\frac{1}{2}}|}{2[\tilde{f}(\tilde{y} - \tilde{l}_i\tilde{d}_i) - \tilde{f}(\tilde{y}_i)]^{\frac{1}{2}}}$$

$$\leqslant \frac{|\Delta l_j^i|}{2|l_j^i|} + \frac{\beta_i|f(y_i - l_i d_i) - f(y_i) - \tilde{f}(\tilde{y}_i - \tilde{l}_i\tilde{d}_i) + \tilde{f}(\tilde{y}_i)|}{2(\tilde{f}(\tilde{y} - \tilde{l}_i\tilde{d}_i) - \tilde{f}(\tilde{y}_i))}. \tag{5.27}$$

根据(5.1),(3.1),(2.3),(5.11)知，

$$\tilde{f}(\tilde{y}_i - \tilde{l}_i\tilde{d}_i) - \tilde{f}(\tilde{y}_i) \geqslant \frac{1}{2} m\tilde{l}_i^{i2} \geqslant \frac{1}{2} m\|\tilde{x}^{i-1}\|^2. \tag{5.28}$$

根据(5.2),

$$|f(y_i - l_i d_i) - f(y_i) - \tilde{f}(\tilde{y}_i - \tilde{l}_i\tilde{d}_i) + \tilde{f}(\tilde{y}_i)|$$

$$\leqslant \frac{\Lambda}{6}\|y_i - l_i d_i\|^3 + \frac{\Lambda}{6}\|y_i\|^3 + \|\nabla\tilde{f}(\theta_1)\|\|\Delta(y_i - l_i d_i)\| + \|\nabla\tilde{f}(\theta_2)\|\|\Delta y_i\|, \tag{5.29}$$

其中 $\theta_1 = (y_i - l_i d_i) + \lambda_1\Delta(y_i - l_i d_i)$, $\theta_2 = y_i + \lambda_2\Delta y_i$, $0 < |\lambda_1|, |\lambda_2| < 1$. 而且由(5.1),(3.1)可知，

$$\begin{aligned} \|\nabla\tilde{f}(\theta_1)\| &= \|\theta_1^T C\| \leqslant M\|\theta_1\| \leqslant 2M(\|y_i\| + \|\tilde{y}_i\| + l_i + \tilde{l}_i) \\ \|\nabla\tilde{f}(\theta_2)\| &\leqslant 2M(\|y_i\| + \|\tilde{y}_i\|). \end{aligned} \tag{5.30}$$

将(5.30),(5.29),(5.28)代入(5.27),得

$$|\beta_j^i-\tilde{\beta}_j^i|\leqslant\frac{|\Delta l_j^i|}{2|\bar{l}_j^i|}+\frac{\xi_3'}{\|\tilde{x}^{i-1}\|^2}[\|y_j^i-l_j^id_j^i\|^3$$
$$+(\|y_j^i\|+\|\tilde{y}_j^i\|+l_j^i+\bar{l}_j^i)(\|y_j^i-\tilde{y}_j^i\|+|l_j^i-\bar{l}_j^i|)].$$

再根据(5.5),(2.3),(5.11),经过整理便可得(5.26).

引理 5.12.

$$\|\Delta A^i\|\leqslant\xi_4(\|\Delta d_1^i\|+\|\Delta a^i\|/\|\tilde{x}^{i-1}\|),\tag{5.31}$$

其中 $\Delta A^i=A^i-\tilde{A}^i$, $\Delta d_1^i=d_1^i-\tilde{d}_1^i$, $\|\Delta a^i\|=a^i-\tilde{a}^i$, ξ_4 仅与 ξ_2 有关.

证. 由(2.10)及$\|A_{jk}^i\|\leqslant 1$, $\|\tilde{A}_{jk}^i\|\leqslant 1$ 可知,

$$\|\Delta A^i\|\leqslant\sum_{1\leqslant j<k\leqslant n-i}\|A_{jk}^i-\tilde{A}_{jk}^i\|.\tag{5.32}$$

由于

$$|\cos\theta_{jk}^i-\cos\tilde{\theta}_{jk}^i|\leqslant|\theta_{jk}^i-\tilde{\theta}_{jk}^i|\leqslant|\theta^i-\tilde{\theta}^i|$$
$$\leqslant\|(d_1^i-a^i/\|a^i\|)-(\tilde{d}_1^i-\tilde{a}^i/\|\tilde{a}^i\|)\|\leqslant\|\Delta d_1^i\|+2\|\Delta a^i\|/\|\tilde{a}^i\|\tag{5.33}$$

同理

$$|\sin\theta_{jk}^i-\sin\tilde{\theta}_{jk}^i|\leqslant\|\Delta d_1^i\|+2\|\Delta a^i\|/\|\tilde{a}^i\|.\tag{5.34}$$

因此将(5.34),(5.33)代入(5.32),再结合(5.11)即得(5.31). 证毕.

引理 5.13.

$$\|R^i\|=1.\tag{5.35}$$

证. 由(2.10)知

$$\det R^i=(\det(V^TV))^{-1}\cdot\det(V^TAV).$$

选W为 $L(V)$ 的正交补空间中的单位正交基, 由于A为 $L(V)$ 内的旋转变换, 所以 $AW=W$. 因此

$$\det(V^TAV)=\det(V^-AV\dotplus E)=\det(V,W)^TA(V,W)$$
$$=\det(V^TV\dotplus W^TW)$$

故结论成立. 证毕.

引理 5.14.

$$\|R^i-\tilde{R}^i\|\leqslant\xi_5\left(\|\Delta d^i\|+\sum_{j=1}^{n-i}\|\Delta\beta_j^i\|+\|\Delta A^i\|\right),\tag{536}$$

其中 $\Delta d^i=d^i-\tilde{d}^i$, $\Delta\beta_j^i=\beta_j^i-\tilde{\beta}_j^i$, $\Delta A^i=A^i-\tilde{A}^i$, ξ_5 与 $M,m,\Delta(D^1)$ 有关.

证. 由(2.11)得

$$\|\tilde{V}(R-\tilde{R})\|\leqslant\|R\|\|V-\tilde{V}\|+\|AV-\tilde{A}\tilde{V}\|,\tag{5.37}$$

而根据(2.5),(5.10),有

$$\|\tilde{V}(R-\tilde{R})\|\geqslant\|R-\tilde{R}\|\det^{\frac{1}{2}}(V^TV)\geqslant\|R-\tilde{R}\|\prod_{j=1}^{n-i}\tilde{\beta}_j^i\cdot\Delta(D),\tag{5.38}$$

根据(5.35)$\|R\|=1$,有

$$\|R\|\|V-\tilde{V}\|+\|AV-\tilde{A}\tilde{V}\|\leqslant\prod_{j=1}^{n-i}\tilde{\beta}_i\cdot\sum_{j=1}^{n-i}\|\Delta d_j\|$$

$$+\sum_{j=1}^{n-i}|\Delta\beta_j|+\prod_{j=1}^{n-i}\tilde{\beta}_j\|\Delta A^i\|. \tag{5.39}$$

综合(5.37)—(5.39)便可得结论．证毕．

引理 5.15.

$$\Delta_E(\tilde{R}^i)\geqslant\xi_6>0, \tag{5.40}$$

其中 ξ_6 仅与 m, M 有关．

证．由(3.4),(2.10)以及(5.35),

$$\Delta_E(\tilde{R}^i)=|\det(\tilde{R})|\prod_{j=1}^{n-i}\|\tilde{r}_j\|^{-1}=\prod_{j=1}^{n-i}\|\tilde{r}_j\|^{-1}.$$

设 $\lambda_{\min}$ 为 $\tilde{V}^T\tilde{V}$ 的最小特征值，并根据(2.11),(3.3)知，

$$\lambda_{\min}\|\tilde{r}_j\|^2\leqslant\tilde{r}_j^T\tilde{V}^T\tilde{V}\tilde{r}_j=\tilde{v}_j^T\tilde{A}^T\tilde{A}\tilde{v}_j=\|\tilde{v}_j\|^2\leqslant\tilde{\beta}_j^2\leqslant m^{-4}.$$

同时根据引理 5.9 的证明，可知

$$\lambda_{\min}\geqslant m_1.$$

因此(5.40)成立．证毕．

引理 5.16.

$$1\leqslant\|\tilde{r}_j\|/\|\tilde{b}_j\|\leqslant\xi_7, \tag{5.41}$$

其中 ξ_7 仅与 ξ_6 有关．

证．由(2.12)式知，

$$\|\tilde{r}_j\|^2=(\cos^2(\widehat{\tilde{r}_j,\hat{q}_1})+\cdots+\cos^2(\widehat{\tilde{r}_j,\hat{q}_{j-1}}))\|\tilde{r}_j\|^2+\|\tilde{b}_j\|^2,$$

$$\|\tilde{r}_j\|^2/\|\tilde{b}_j\|^2=\left[1-\sum_{k=1}^{j-1}\cos^2(\widehat{\tilde{r}_j,\hat{q}_k})\right]^{-1}=\cos^{-2}(\widehat{\tilde{r}_j,q_j}).$$

这是因为 $\tilde{r}_j\in L(\hat{q}_1,\cdots,\hat{q}_j)$．又因 $\tilde{r}_j\bar{\in}L(\hat{q}_1,\cdots,\hat{q}_{j-1})$，所以

$$(\widehat{\tilde{r}_j,\hat{q}_j})\leqslant\frac{\pi}{2}-(\widehat{\tilde{r}_j,L(\hat{q}_1,\cdots,\hat{q}_{j-1})})=\frac{\pi}{2}-(\widehat{\tilde{r}_j,L(\tilde{r}_1,\cdots,\tilde{r}_{j-1})}).$$

由(5.40)，$\Delta(\tilde{R}^i)\geqslant\xi_6>0$，故存在 $\theta>0$，使得

$$(\widehat{\tilde{r}_j,L(\tilde{r}_1,\cdots,\tilde{r}_j)})\geqslant\theta>0,$$

$$\cos^2(\widehat{\tilde{r}_j,\hat{q}_j})\geqslant\cos^2\left(\frac{\pi}{2}-\theta\right)=\xi_7^{-1}.$$

证毕．

引理 5.17.

$$\|\Delta D^{i+1}\|\leqslant\xi_8\left[\|\Delta D^i\|+\frac{\|\Delta x^i\|+\|\Delta x^{i-1}\|}{\|\tilde{x}^{i-1}\|}+\sum_{j=1}^{n-i}|\Delta\beta_j^i|+\|\Delta R^i\|\right], \tag{5.42}$$

其中 $\Delta D^i=D^i-\tilde{D}^i$，$\Delta R^i=R^i-\tilde{R}^i$，$\xi_8$ 仅与 m, M, m_1, ξ_1, ξ_7 有关．

证．

$$\|\Delta D^{i+1}\|\leqslant\sum_{j=2}^{n-i}\left\|\frac{u_j^i}{\|u_j^i\|}-\frac{\tilde{u}_j^i}{\|\tilde{u}_j^i\|}\right\|+\left\|\frac{x^i-x^{i-1}}{\|x^i-x^{i-1}\|}-\frac{\tilde{x}^i-\tilde{x}^{i-1}}{\|\tilde{x}^i-\tilde{x}^{i-1}\|}\right\|$$
$$+\|(d_{n-i+1}^i,\cdots,d_n^i)-(\tilde{d}_{n-i+1}^i,\cdots,\tilde{d}_n^i)\|. \tag{5.43}$$

(5.43)右端第三项

$$\|(d^i_{n-i+1},\cdots,d^i_n)-(\tilde{d}^i_{n-i+1},\cdots,\tilde{d}^i_n)\|\leqslant\|\Delta D^i\|, \tag{5.44}$$

(5.43)右端第二项

$$\left\|\frac{x^i-x^{i-1}}{\|x^i-x^{i-1}\|}-\frac{\tilde{x}^i-\tilde{x}^{i-1}}{\|\tilde{x}^i-\tilde{x}^{i-1}\|}\right\|\leqslant\frac{\|x^i-x^{i-1}\|+\|\tilde{x}^i-\tilde{x}^{i-1}\|}{\|\tilde{x}^i-\tilde{x}^{i-1}\|}$$

$$+\frac{\|(x^i-x^{i-1})-(\tilde{x}^i-\tilde{x}^{i-1})\|}{\|\tilde{x}^i-\tilde{x}^{i-1}\|}\leqslant\frac{2}{\xi_1\|\tilde{x}^{i-1}\|}(\|\Delta x^i\|+\|\Delta x^{i-1}\|). \tag{5.45}$$

考查(5.43)右端第一项，根据(5.17)得

$$\|u^i_j/\|u^i_j\|-\tilde{u}^i_j/\|\tilde{u}^i_j\|\|\leqslant\frac{2}{m_1}\|u^i_j-\tilde{u}^i_j\|$$

$$\leqslant\frac{2}{m_1}[\|V^i\|\|q_i-\tilde{q}_i\|+\|q_i\|\|V^i-\tilde{V}^i\|]$$

$$\leqslant\frac{2}{m_1}\left[\prod_{k=1}^{n-i}\beta^i_k\cdot\frac{2\|b^i_j-\tilde{b}^i_j\|}{\|\tilde{b}^i_j\|}+\sum_{k=1}^{n-i}(|\Delta\beta^i_k|+\beta^i_k\|\Delta d^i_k\|)\right], \tag{5.46}$$

其中，

$$\|b_j-\tilde{b}_j\|/\|\tilde{b}_j\|\leqslant\|r_j-\tilde{r}_j\|/\|\tilde{b}_j\|+\frac{1}{\|\tilde{b}_j\|}\sum_{k=1}^{j-1}\left\|\frac{(r_j,b_k)}{(b_k,b_k)}b_k-\frac{(\tilde{r}_j,\tilde{b}_k)}{(\tilde{b}_k,\tilde{b}_k)}\tilde{b}_k\right\|$$

$$\leqslant\|r_j-\tilde{r}_j\|/\|\tilde{b}_j\|+\sum_{k=1}^{j-1}\left[\frac{\|\tilde{r}_j\|}{\|\tilde{b}_j\|}\cdot\frac{\|b_k-\tilde{b}_k\|}{\|\tilde{b}_k\|}\right.$$

$$\left.+\frac{\|b_k\|}{\|\tilde{b}_j\|}\cdot\frac{|\|\tilde{b}_k\|^2(r_j,b_k)-\|b_k\|^2(\tilde{r}_j,\tilde{b}_k)|}{\|b_k\|^2\cdot\|\tilde{b}_k\|^2}\right]$$

$$\leqslant\frac{\|r_j-\tilde{r}_j\|}{\|\tilde{b}_j\|}+\sum_{k=1}^{j-1}\frac{\|\tilde{r}_j\|}{\|\tilde{b}_j\|}\frac{\|b_k-\tilde{b}_k\|}{\|\tilde{b}_k\|}+\sum_{k=1}^{j-1}\frac{|\|b_k\|-\|\tilde{b}_k\||\cdot|(r_j,b_k)|}{\|\tilde{b}_j\|\cdot\|b_k\|^2}$$

$$+\sum_{k=1}^{j-1}\left[\frac{|(\tilde{r}_j,\tilde{b}_k-b_k)|\|b_k\|}{\|\tilde{b}_j\|\cdot\|\tilde{b}_k\|^2}+\frac{|(r_j-\tilde{r}_j,b_k)|\|b_k\|}{\|\tilde{b}_j\|\|\tilde{b}_k\|^2}\right]$$

$$\leqslant\xi'_8\sum_{k=1}^{j}\frac{\|\tilde{r}_k-r_k\|}{\|\tilde{r}_k\|}+\xi'_8\sum_{k=1}^{j-1}\frac{\|b_k-\tilde{b}_k\|}{\|\tilde{b}_k\|}. \tag{5.47}$$

(5.47)的成立是根据不等式

$$(n-j)^{-1}\left(\frac{m}{M}\right)^4\leqslant\|r_k\|,\|\tilde{r}_k\|\leqslant m_1\cdot m^{-4} \tag{5.48}$$

和(5.41)，因此 ξ'_8 仅与 m,M,m_1,ξ_7 有关．对 j 用归纳法便可得

$$\|b^i_j-\tilde{b}^i_j\|/\|\tilde{b}^i_j\|\leqslant\xi''_8\|\Delta R^i\|. \tag{5.49}$$

综合(5.49)，(5.46)—(5.43)便可得(5.42)．证毕．

六、算法的二次收敛性

我们称算法具有 n_0 步二次收敛性，如果存在常数 C 及正整数 n_0，使得

$$\|x^{n_0}-x^*\|\leqslant C\|x^0-x^*\|^2, \tag{6.1}$$

其中 $\{x^n\}$ 为算法产生的点列，x^0 为初始点，x^* 为函数的极小值点．

定理 6.1. 设 f 满足假设 (A1), (A2), $\{x^k\}$ 是算法(Ag)以 f 为目标函数得到的点列,那么 $\{x^k\}$ n 步二次收敛于 f 的极小值点 $x^*=0$,即

$$\|x^n\| \leqslant C\|x^0\|^2,$$

其中 x^0 为起始点, C 仅与 $m, M, \Lambda, \Delta(D^1)$ 有关.

证. 设 $\tilde{f}$ 如(5.1)为 f 的二次逼近函数, $\tilde{x}^0=x^0$, $\tilde{D}^1=D^1$,并设

$$\begin{aligned} &x^0 = y_0^0,\\ &y_j^i = S(y_{j-1}^i; d_j^i), \quad j=1,2,\cdots,n, \; i=0,1,2,\cdots, \qquad (6.2)\\ &x^{i+1} = y_0^{i+1} = S(y_n^i; d_n^{i+1}), \quad i=0,1,2,\cdots. \end{aligned}$$

根据定理 4.1,算法对 $\tilde{f}$ 至多一轮就收敛到 $\tilde{f}$ 的极小值点,所以我们只须对 $i\leqslant n$ 证明

$$\|\Delta y_j^i\| \leqslant \zeta\|x^0\|^2 \qquad (6.3)$$

$$\|\Delta D^{i+1}\| \leqslant \eta\|x^0\|^2/\|\tilde{x}^i\| \qquad (i,j=0,1,\cdots,n) \qquad (6.4)$$

成立. 将数对(i,j)按字典序排列,易知(6.3),(6.4)在$(0,0)$时成立,故设在小于或等于(i,j)时(6.3),(6.4)仍成立. 在大于(i,j)时,分两种情况讨论:

i) $j<n$. 故只须证(6.3),(6.4)在$(i,j+1)$时成立. 由归纳假设(6.4)已经成立,下证(6.3). 由(5.18),

$$\|\Delta y_{j+1}^i\| \leqslant \delta_1\|\Delta y_j^i\| + \delta_2\|\tilde{y}_j^i\|\|\Delta d_j^i\| + \delta_3\|\Delta g_j^i\| + \delta_4\|\tilde{y}_j^i\|^2 + \delta_5\|y_j^i\|\|\tilde{y}_j^i\|. \qquad (6.5)$$

由(5.3)与(5.5),知

$$\|\Delta g_j^i\| \leqslant \|g(y_j^i) - \tilde{g}(y_j^i)\| + \|\tilde{g}(y_j^i) - \tilde{g}(\tilde{y}_j^i)\| \leqslant \frac{\Lambda}{2}\|y_j^i\|^2 + M\|\Delta y_j^i\|, \qquad (6.6)$$

$$\|y_j^i\| \leqslant \left(\frac{M}{m}\right)^{\frac{1}{2}}\|x^0\|, \quad \|\tilde{y}_j^i\| \leqslant \left(\frac{M}{m}\right)^{\frac{1}{2}}\|\tilde{x}^{i-1}\|. \qquad (6.7)$$

将(6.6),(6.7),(6.4)代入(6.5),并根据归纳假设,得证(6.3),其中 ζ 只与 M, m, Λ 有关.

ii) $j=n$. 要证(6.3),(6.4)在$(i+1,0)$时成立. 先证(6.4). 由(5.42),(5.36),(5.31),(5.26)得

$$\|\Delta D^{i+1}\| \leqslant \frac{\zeta_6+\zeta_5+\zeta_4+\zeta_3}{\|\tilde{x}^{i-1}\|}\Big[\|\Delta D^i\| + \|\Delta x^i\| + \|\Delta x^{i-1}\| + \|\Delta a_1^i\| + \sum_{j=1}^{n-i}\|\Delta y_j^i\| + \sum_{j=1}^{n-i}|\Delta l_j^i|\Big]. \qquad (6.8)$$

由归纳假设(6.3)对 i 成立,因此

$$\|\Delta x^i\| + \|\Delta x^{i-1}\| + \|\Delta a^i\| + \sum_{j=1}^{n-i}\|\Delta y_j^i\| \leqslant 4n\zeta\|x^0\|^2. \qquad (6.9)$$

下面我们分四种情况对 $|\Delta l_j^i| = |l_j^i - \tilde{l}_j^i|$ 进行讨论:

i) $|\Delta l_j^i| = |\|a^i\| - \|\tilde{a}^i\|| \leqslant \|\Delta x^i\| + \|\Delta y_{n-i}^i\|$;

ii) $|\Delta l_j^i| = |\lambda_j^i - \tilde{\lambda}_j^i| = |\|y_j^i - y_{j-1}^i\| - \|\tilde{y}_j^i - \tilde{y}_{j-1}^i\|| \leqslant \|\Delta y_j^i\| + \|\Delta y_{j-1}^i\|$;

iii) $|\Delta l_j^i| = |\lambda_j^i - \|\tilde{a}^i\||$;

此时, $\lambda_j^i \geqslant \|a^i\|$, $\|\tilde{a}^i\| \geqslant \tilde{\lambda}_j^i$. 如果 $\lambda_j^i \geqslant \|\tilde{a}^i\|$,那么

$$|\Delta l_j^i| = \lambda_j^i - \|\tilde{a}^i\| \leqslant \lambda_j^i - \tilde{\lambda}_j^i,$$

归结为 ii). 如果$\|\tilde{a}^i\| \geqslant \lambda_j^i$,那么

$$|\Delta l_j^i| = \|\tilde{a}^i\| - \lambda_j^i \leqslant \|\tilde{a}^i\| - \|a^i\|$$

归结为 i)。

iv) $|\Delta l_j^i| = |\bar{\lambda}_j^i - \|a^i\||$。类似 iii)，仅归为 i)或 ii)。综合这四种情况与(6.9)，再根据归纳假设，即得证(6.4)，其中 η 只与 L, m, M, Λ 与 $\Delta(D^1)$ 有关。

又因为 $\|\Delta y_0^{i+1}\| = \|S(y_n^i; d_n^{i+1}) - S(\tilde{y}_n^i; \tilde{d}_n^{i+1})\|$，根据(5.18)及归纳假设，即得(6.3)。至此我们归纳证明了(6.3)，(6.4)对(i, j)成立。

由于算法对 $\tilde{f}$ 至多 n 步终结，即存在 $0 \leqslant n_0 \leqslant n$，使得 $\tilde{x}^{n_0} = 0$，所以由(6.3)得到

$$\|x^{n_0}\| = \|\Delta x^{n_0}\| \leqslant C\|x^0\|^2.$$

证毕。

定理 6.2. 设 f 满足假设(A1)，(A2)，对 f 应用算法(Ag)得点列$\{x^i\}$，那么$\{x^i\}$ n 步二次收敛于 f 的极小值点 x^*，即

$$\|x^n - x^*\| \leqslant C\|x^0 - x^*\|,$$

其中 x^0 为起始点，C 为仅与 $L, m, M, \Lambda, \Delta(D^1)$ 有关的常数。

证。由定理 6.1 即得。证毕。

致谢：本文是在俞文魮教授的指导下完成，在此谨表谢意。

参 考 文 献

[1] Yu Wenci, A New Improvement of Powell Method, 1985, to appear.

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

[3] Zangwill, W. I., Minimizing a function without calculating derivatives, *Computer J.*, 10(1967), 293—296.

[4] Brent, P. R., Algorithms for Minimization Without Derivatives, Prentice-Hall, Englewood Cliffs, 1973.

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

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

[7] Fletcher, R., Practical Mathods of Optimization, V. I, Unconstrained Optimization, John Wiley & Sons, 1980.

[8] Meang, R. K., A matrix inequality, *SIAM, J. Numer Anal.*, 6(1967), 104—107.

[9] Luenburge, D. G., Introduce of Linear and Nonlinear Programming, Addison-Wesley Publishing Company, 1973.

A DIRECT ALGORITHM WITH *n*-STEP QUADRATIC CONVERGENCE

Xu Yi-Fan

(Shandong Institute of Light Industry, Jinan 250100)

ABSTRACT

A modifiel PY method is proposed. Its main character is that the rotation transformation and Gram-Schmidt orthogonalizing process are used to rectify the set of search directions. Hence, the algorithm remains not only the quadratic termination, but also the *n*-step quardratic convergence for the objective function which is uniform convex.