1991年1月　　重 庆 大 学 学 报　　Vol. 14, №. 1
第14卷第1期　　JOURNAL OF CHONGQING UNIVERSITY　　Jan. 1991

无约束优化的信赖域算法*

TRUST REGION METHOD FOR UNCONSTRAINED OPTIMIZATION

唐 健*　　段虞荣
Tang Jian　　Duan Yurong
（重庆建筑工程学院）　　（重庆大学）

摘 要 提出了一种无约束优化问题的信赖域算法。根据原优化问题的二次近似模型，运用拟牛顿方向与最速下降方向之凸组合作为搜索方向，采用了新的策略。进行了收敛性分析，得到整体收敛及局部二次收敛性结果，并给出了算法的执行过程及算例。

关键词 信赖域方法；二次规划；无约束优化。

中国图书资料分类法分类号： O221.2

ABSTRACT A new trust region algorithm is proposed for solving unconstrained optimization problems. According to the quadratic approximate model of the original optimization problem, the trust region algorithm uses directions, a convex combination of the quasi-newton direction and the steepest descent direction. This algorithm with new strategy is analyzed and the global and local quadratic convergent theorems are proved. At last, the implementation and computational results of the algorithm are demonstrated.

KEY WORDS trust region; quadratic programming; unconstrained optimization.

0 引 言

近年来人们对用信赖域方法求解无约束优化问题越来越重视，无约束优化的信赖域方法已得到发展和应用，见文[2-6]。如何获得理论上和实际计算中均很成熟的信赖域算法正是人们研究的焦点所在，下面我们将给出一个新的算法，并研究其收敛性。

1 算 法

考虑无约束优化问题：

$$\mathrm{MIN}\ f(x),\ x\in R^n \tag{1}$$

其中 $f: R^n\to R, f\in C^2$。 记 $g_K=\nabla f(x_K)$， $f_K=f(x_K)$， B_K 为 $f(x)$ 的 Hesse 阵或近似阵，

$\psi_K(w)=g_K^T w+\frac{1}{2}w^T B_K w$。下面给出问题(1)的信赖域算法。

1） 给定初值及常数 x_0， B_0^{-1}， $u_1\leqslant u_0\leqslant u_2$， $0<r_3<r_2<r<1<r_1$， $\theta\in(0,1)$，$\sigma\in(r,1)$， ε；

---

* 收文日期 1989－04－10

*原系我校硕士研究生

2） $k:=0$，　$L(k):=0$，计算 f_k, g_k；

3） 如 $\|g_k\|_2 \leq \varepsilon$，停止执行；

4） 如 $\|B_k^{-1}g_k\|_2 \leq v_k$，令 $d_k=-B_k^{-1}g_k$，取 $p_k=\tau_k d_k$，其中 $\tau_k=\sigma^j (j=0,1,2,\cdots)$ 为使下式成立的最大正数：

$$f(x_k+\tau_k d_k) \leq \max_{0\leq j\leq L(k)} f_{k-j}+r\tau_k g_k^T d_k$$

否则，令 $d_k=-[\theta B_k^{-1}g_k+(1-\theta)g_k]$，　取 $p_k=\tau_k d_k$，　$\tau_k=\sigma^j (j=(0,1,2,\cdots)$ 为使下式成立的最大正数：

$$\begin{cases} f(x_k+\tau_k d_k) \leq \max\limits_{0\leq j\leq L(k)} f_{k-j}+r\tau_k g_k^T d_k \\ \|\tau_k d_k\|_2 \leq v_k \end{cases}$$

5） $x_{k+1}:=x_k+p_k$，计算 f_{k+1}，　g_{k+1}，　$L(k+1)=\mathrm{Min}(M, L(k)+1)$；

6） 修正 B_k^{-1} 得 B_{k+1}^{-1}，　$g_k=\dfrac{f_{k+1}-f_k}{\psi_k(p_k)}$，　$v_{k+1}=\begin{cases}\mathrm{Min}(u_2, r_1 v_k)，若\ q_k \geq r_3 \\ \mathrm{Max}(u_1, r_1 v_k)，否则\end{cases}$

7） $k:=k+1$，　转第三步。

算法中 $u_1, u_2, r_1, r_2, r_3, r, \theta, \sigma, \varepsilon$　均为事先选定的常数，M 为正整数，一般应有 $M\leq 10$，ε 为误差参数。称算法第四步中步长收索策略

$$f(x_k+\tau_k d_k) \leq \max_{0\leq j\leq L(k)} f_{k-j}+r\tau_k g_k^T d_k$$

为非单调 Armijo 型规则，它有利于计算稳定性。

2　收敛性分析

由修正公式可保证 B_k^{-1} 对称正定。

2.1定理1：　设 $\{\tau_k\}$ 为非单调 Armijo 规则每步终止时的 τ_k 值，则有正数 ε，k 存在，$\forall\ k>k_0$，$\tau_k \geq \varepsilon$。

证：设结论不真。则存在子列 $\{\tau_{\bar{k}}\}$ 趋于零，这里 $\bar{k}\in\{k\}$。由算法，存在 $\bar{k}_0$，$\forall\ \bar{k}\geq\bar{k}_0$，有

$$f(x_{\bar{k}}+\frac{\tau_{\bar{k}}}{\sigma}d_{\bar{k}}) > \max_{0\leq j\leq L(\bar{k})} f_{\bar{k}-j}+\frac{r}{\sigma}\tau_{\bar{k}} g_{\bar{k}}^T d_{\bar{k}} \geq f_{\bar{k}}+\frac{r}{\sigma}\tau_{\bar{k}} g_{\bar{k}}^T d_{\bar{k}}$$

由中值定理，

$$g(x_{\bar{k}}+\theta_{\bar{k}}\frac{\tau_{\bar{k}}}{\sigma}d_{\bar{k}})^T d_{\bar{k}} > r g_{\bar{k}}^T d_{\bar{k}}$$

这里 $\theta_{\bar{k}}\in(0,1)$，$\forall\ \bar{k}$。由算法易知

$$g_k^T d_k<0, \forall\ k$$

$$\lim_{\bar{k}\to\infty}\tau_{\bar{k}}=0$$

又 $f\in C^1$，故 $\bar{k}$ 充分大时，$g(x_{\bar{k}}+\frac{1}{\sigma}\theta_{\bar{k}}\tau_{\bar{k}}d_{\bar{k}})^T d_{\bar{k}}<0$

$$\lim_{\bar{k}\to\infty} g(x_{\bar{k}}+\frac{1}{\sigma}\theta_{\bar{k}}\tau_{\bar{k}}d_{\bar{k}})^T d_{\bar{k}}=\lim_{\bar{k}\to\infty} g_{\bar{k}}^T d_{\bar{k}}$$

故由(2)式得

$$0<g(x_{\bar{k}}+\frac{1}{\sigma}\theta_{\bar{k}}\tau_{\bar{k}}d_{\bar{k}})^T d_{\bar{k}}/g_{\bar{k}}^T d_{\bar{k}}<r<1$$

这与上式矛盾。

2.2　定理2：　设　$f: D\subset R^n\to R$，在紧集 $D_0\subset D$ 上连续可微，在 D_0 内驻点有限。$\{x_k\}\subset D_0$，$\lim\limits_{k\to\infty}(x_{k+1}-x_k)=0$，$\lim\limits_{k\to\infty} f'(x_k)=0$。则 $\lim\limits_{k\to\infty} x_k=x^*$，且 $f'(x^*)=0$

证：此结论容易得到，略。

2.3 定理3：设$\{x_k\}$由算法导出，$\nabla f(x_k)\neq 0,\forall k=1,2,\cdots$且(i)水平集$Z_0=\{x|f(x)\leqslant f(x_0)\}$有界；(ii) $\forall k,\beta\leqslant\|B_k^{-1}\|\leqslant\lambda,\beta,\lambda$为常数。则(a)$\{x_k\}\subset Z_0$；(b)如$\bar{x}$为$\{x_k\}$的任一聚点，则$\nabla f(\bar{x})=0$；(c)$\{x_k\}$的聚点均非$f(x)$的局部极大值点；(d)如在$Z_0$内驻点有限，则驻点唯一，即$\lim\limits_{k\to\infty}x_k=x^*$，且$x^*$为问题(1)的驻点。

证：由算法及定理1

$$g_k^Tp_k=\begin{cases}-\tau_kg_k^TB_k^{-1}g_k\leqslant-\varepsilon\beta\|g_k\|_2^2,\\ \text{或}-\tau_kg_k^T[\theta B_k^{-1}+(1-\theta)I_n]g_k\leqslant-\varepsilon[\theta\beta+1-\theta]\|g_k\|_2^2,\end{cases}$$

$$\|p_k\|_2=\begin{cases}\|-\tau_kB_k^{-1}g_k\|_2\leqslant\lambda\|g_k\|_2,\\ \text{或}\|-\tau_k[\theta B_k^{-1}+(1-\theta)I_n]g_k\|_2\leqslant(\theta\lambda+1-\theta)\|g_k\|_2.\end{cases}$$

令 $c_1=\mathrm{Min}\{\varepsilon\beta,\varepsilon(\theta\beta+1-\theta)\}$，　$c_2=\mathrm{Max}\{\lambda,\theta\lambda+1-\theta\}$，有

$$g_k^Tp_k\leqslant-c_1\|g_k\|_2^2,\quad \|p_k\|_2\leqslant c_2\|g_k\|_2,k=1,2,\cdots \tag{3}$$

结论(a)可由算法直接推出。

现证明结论(b)。分二步证明$\lim\limits_{k\to\infty}\tau_k\|d_k\|_2=0$

设$\{l_{(k)}\}$为$\{k\}$的子集，使得$f(x_{l(k)})=\mathop{\mathrm{Max}}\limits_{0\leqslant j\leqslant l(k)}f_{k-j}$

显然　$k-L(k)\leqslant l(k)\leqslant k$，　$L(k+1)\leqslant L(k)+1$，

由(3)式

$$f[x_{l(k+1)}]=\mathop{\mathrm{Max}}\limits_{0\leqslant j\leqslant L(k+1)}f_{k+1-j}\leqslant\mathop{\mathrm{Max}}\limits_{0\leqslant j\leqslant l(k)+1}f_{k+1-j}$$
$$=\mathrm{Max}(f[x_{l(k)}],f(x_{k+1}))\leqslant f[x_{l(k)}]$$

因此$\{f_{l(k)}\}$为单调下降序列，又Z_0为紧集，$\{x_0\}\subset Z_0$，故$\{f_{l(k)}\}$有下界，即$f[x_{l(k)}]$的极限存在且唯一，设为f^*。对充分大的正数k，$\forall k>M$，由(3)式

$$f[x_{l(k)}]=f[x_{l(k)-1}+\tau_{l(k)-1}d_{l(k)-1}]$$
$$\leqslant\mathop{\mathrm{Max}}\limits_{0\leqslant j\leqslant l(l(k)-1)}f_{l(k)-1-j}+r\tau_{l(k)-1}g_{l(k)-1}^Td_{l(k)-1}\tag{5}$$

又$\tau_k>0$，　$g_k^Td_k<0$，　由(5)式

$$\lim_{k\to\infty}\tau_{l(k)-1}g_{l(k)-1}^Td_{l(k)-1}\geqslant 0$$

由(3)式

$$\tau_kg_k^Td_k\leqslant-c_1\|g_k\|_2^2\leqslant-\frac{c_1}{c_2^2}\|p_k\|_2^2$$

$$\lim_{k\to\infty}\tau_{l(k)-1}\|d_{l(k)-1}\|=0\tag{6}$$

即

$$\lim_{k\to\infty} f[x_{l(k)}] = \lim_{k\to\infty} f[x_{l(k)-1} + \tau_{l(k)-1} d_{l(k)-1}] = \bar{f} = \lim_{k\to\infty} f[x_{l(k)-1}] \tag{7}$$

要证：

$$\forall\ m \leqslant l(k),\quad \begin{cases} \lim\limits_{k\to\infty} \tau_{l(k)-m} \|d_{l(k)-m}\|_2 = 0 \\ \lim\limits_{k\to\infty} f[x_{l(k)-m}] = \bar{f} \end{cases} \tag{8}$$

$m=1$ 时，由 (6)、(7)式，知(8)式成立。

设 $m=j$ 时，(8)式成立、则当 $m=j+1$时，由(5)式

$$f[x_{l(k)-j}] \leqslant f[x_{l(l(k)-j-1)}] + r\tau_{l(k)-j-1} g^T_{l(k)-j-1} d_{l(k)-j-1}$$

对固定的 j，由假设：

$$\lim_{k\to\infty} f[x_{l(k)-j}] = \bar{f} = \lim_{k\to\infty} f[x_{l(l(k)-j-1)}]$$

$$\lim_{k\to\infty} r\tau_{l(k)-j-1} g^T_{l(k)-j-1} d_{l(k)-j-1} \geqslant 0$$

同样可推出

$$\lim_{k\to\infty} \tau_{l(k)-j-1} \|d_{l(k)-j-1}\|_2 = 0$$

$$\lim_{k\to\infty} f[x_{l(k)-j-1}] = \lim_{k\to\infty} f[x_{l(k)-j}] = \bar{f}$$

故对任何正整数 m，($m \leqslant l(k)$)，(8)式成立。

由 $k-L(k) \leqslant l(k) \leqslant k$， $k \leqslant l(k+L(k)) \leqslant k+L(k)$

注意到，当 $k \geqslant M$ 时， $L(k) \equiv M$， $l(k+L(k)) = k+M$， 又

$$x_k = x_{l(k+L(k))} - \sum_{j=1}^{l(k+L(k))-k} \tau_{l(k+L(k))-j} d_{l(k+L(k))-j}$$

由(8)式

$$\lim_{k\to\infty} \|x_k - x_{l(k+L(k))}\|_2 = 0 \tag{9}$$

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} f[x_{l(k+M)}] = \bar{f} \tag{10}$$

由(3)式及定理1：

$$f(x_{k+1}) \leqslant f[x_{l(k)}] + r g_k^T d_k \leqslant f[x_{l(k)}] - r c_1 \|g_k\|_2^2$$

由(10)式推出：

$$\lim_{k\to\infty} \|g_k\|_2 = 0, \quad \lim_{k\to\infty} \|p_k\|_2 = \lim_{k\to\infty} \tau_k \|d_k\|_2 = 0$$

因而如 $\bar{x}$ 为$\{x_k\}$的任一聚点，则$\nabla f(\bar{x}) = 0$。

现证明结论(c)。设 $\bar{x}$ 为 f 的局部极大值点，由(9)式推出，存在$\{x_{l(k)}\}$的子列$\{x_{l(k)}\}_x$ 收敛于 $\bar{x}$. 因 $f[x_{l(k)}]$单调下降，$\lim\limits_{k\to\infty} f[x_{l(k)}] = f(\bar{x})$，且 $f[x_{l(k)}] \geqslant f(\bar{x})$。由(5)式，∃ $k' > k$，使得

$$f[x_{l(k')}] < f[x_{l(k)}]$$

对充分大的 k 有 $\bar{x}$ 的邻域内之 $x_{l(k)}$，使得 $f[x_{l(k)}] > f(\bar{x})$， 这与假设矛盾。

结论(d)可由(3)、(4)式推出

$$\lim_{k\to\infty}(x_k - x_{k+1}) = 0, \qquad \lim_{k\to\infty} g(x_k) = 0 .$$

又$\{x_k\}\subset Z_0$(紧集)，由定理2，结论成立。

2.4 引理1： 设 $f: D\subset R^n \to R$，在紧集 $D_0\subset D$ 上可微，$H: D_0\to L(R^n)$是连续映射，并且 $H(x)$正定，$\forall\ x\in D_0$。令 $P(x) = -H(x)^{-1}f'(x), x\in D_0$，则存在常数 $c\in(0,1]$，使得

$$[f'(x)]^T P(x) \leqslant -c\|f'(x)\|\|p(x)\|, \forall\ x\in D_0。$$

证明参见文[6]中定理14.4.1的证明。

2.5 引理2： 设 $g: R^n\to R^n$， $\varepsilon>0$， x^*是 $g\in C^1$在 $s(x^*,\varepsilon)$球邻域上的零点，且 $\nabla g(x^*)\neq 0$。则

(a)存在 $\delta>0$，使得，如 $x_0\in s(x^*,\delta)$，则由 $x_{k+1} = x_k - [\nabla g(x_k)]^{-1}g(x_k)$导出的序列$\{x_k\}$有定义，$x_k\in S(x^*,\delta)$，$\forall\ k$，并收敛于 x^*；

(b)如$\forall\ k, x_k\neq x^*$，则 $\lim\limits_{k\to\infty}\|x_{k+1}-x^*\|/\|x_k-x^*\| = 0$，且如$\forall\ r>0$，存在 $\delta_r>0$，使得如 $x_k\in s(x^*,\delta_r)$，则$\|x_{k+1}-x^*\|\leqslant r\|x_k - x^*\|$；

(c)如存在常数 L, M，使得$\|\nabla g(x)-\nabla g(y)\|\leqslant L\|x-y\|$， $\|\nabla g(x)^{-1}\|\leqslant M$，$\forall\ x,y\in s(x^*,\varepsilon)$，那么$\|x_{k+1}-x^*\|\leqslant\dfrac{LM}{2}\|x_k-x^*\|^2$。

证明类似于文[7]中第三节相应定理。

2.6 定理4： 设存在 $c_1\in(0,1]$，每步迭代都有 $-\psi_k(p_k)\geqslant -\mathrm{Min}\{g_k^T d+\frac{1}{2}d^T B_k d;\ \ \|d\|\leqslant v_k, d\in[d_k]\}$。其中 d_k 满足：$d_k^T g_k\leqslant c_1\|d_k\|\|g_k\|$，$[d_k]$表示 d_k 的线性组合之集。则存在 $c_2,\delta>0$，使得 $\forall\ B_k, v_k, g_k$，有$-\psi_k(p_k)\geqslant c_2\|g_k\|\mathrm{Min}\{v_k, \|g_k\|/\|B_k\|\}$。

证： 设 $\alpha^* d_k$ 为 $\mathrm{Min}\{g_k^T d+\frac{1}{2}d^T B_k d;\ \ \|d\|\leqslant v_k, d\in[d_k]\}$之解，令 $h(\alpha) = \alpha g_k^T d_k + \frac{1}{2}\alpha^2 d_k^T B_k d_k$。显然 α^* 极小化 $h(\alpha)$，并满足约束 $\alpha^*\|d_k\|\leqslant v_k$，且

$$\alpha^* = \begin{cases} -g_k^T d_k/d_k^T B_k d_k, \text{如果} & -g_k^T d_k\leqslant\dfrac{v_k}{\|d_k\|}d_k^T B_k d_k; \\ v_k/\|d_k\|, & \text{否则} \end{cases}$$

当 α^*为第一种情况时：

$$-h(\alpha^*) = \frac{(g_k^T d_k)^2}{2d_k^T B_k d_k}\geqslant\frac{c_1^2\|g_k\|^2}{2\|B_k\|}$$

当 α^*为第二种情况时：

$$-h(\alpha^*) = -\frac{v_k}{\|d_k\|}g_k^T d_k - \frac{1}{2}\frac{v_k^2}{\|d_k\|^2}d_k^T B_k d_k\geqslant\frac{c_1}{2}v_k\|g_k\|$$

取 $c_2=\mathrm{Min}\{\frac{c_1^2}{2},\frac{c_1}{2}\}$，则结论成立。

2.7　定理5：　设定理4结论成立，$r<\frac{1}{2}$，取 $B_k^{-1}=[\nabla^2 f(x_k)]^{-1}$，$\forall$ k，又设 $\|\nabla^2 f(x)-\nabla^2 f(y)\|\leqslant L\|x-y\|$，$x^*$ 为 $\{x_k\}$ 之一聚点，且 $\nabla^2 f(x^*)$ 正定。则 x_k 二次收敛于 x^*。

证：　由泰勒展式

$$\|(f_k-f_{k+1})-(-g_k^T p_k-\frac{1}{2}p_k^T B_k p_k)\|=|\frac{1}{2}p_k^T B_k p_k-\int_0^1 p_k^T\nabla^2 f(x_k+\lambda p_k)p_k(1-\lambda)d\lambda|$$
$$\leqslant\|p_k\|^2\int_0^1\|B_k-\nabla^2 f(x_k+\lambda p_k)\|(1-\lambda)d\lambda$$

令 $q_k=\frac{f_{k+1}-f_k}{\psi_k(p_k)}$，则

$$|q_k-1|\leqslant\frac{\|p_k\|^2}{|\psi_k(p_k)|}\int_0^1\|B_k-\nabla^2 f(x_k+\lambda p_k)\|(1-\lambda)d\lambda \tag{11}$$

由假设，存在 $\{x_k\}$ 的子列 $\{x_{k_j}\}$ 收敛于 x^*。设 $\Omega_1=\{x\,|\,\|x-x^*\|<\delta_1\}$，使得 $\nabla^2 f(x)$ 在 Ω_1 内正定；又设 $0<\delta_2<\frac{\delta_1}{4}$，使得在 $\Omega_2=\{x\,|\,\|x-x^*\|<\delta_2\}$ 内有 $\|\nabla^2 f^{-1}(x)g(x)\|\leqslant\frac{\delta_1}{2}$。由定理3知此假定合理。

设 $x_{l(k_{j_0})}\in\Omega_2$ 且满足：$f[x_{l(k_{j_0})}]<\inf\{f(x)\,|\,x\in\Omega_1-\Omega_2\}$。这样的 $x_{l(k_{j_0})}$ 存在，因为 $f[x_{l(k)}]$ 单调下降，$g(x^*)=0$，$\nabla^2 f(x^*)$ 正定。显然 x^* 为局部严格极小点。

当 $k\geqslant k_{j_0}$ 而 $x_k\in\Omega_2$ 时，要证 $x_{k+1}\in\Omega_2$。如若不然，因　$f_{k+1}<f_{l(k)}\leqslant f_{l(k_{j_0})}$，故　$x_{k+1}\notin\Omega_1$

$$v_k\geqslant\|x_{k+1}-x_k\|\geqslant\|x_{k+1}-x^*\|-\|x_k-x^*\|$$
$$>\frac{\delta_1}{2}\geqslant\|B_k^{-1}g_k\|$$

由算法，$p_k=-\tau_k B_k^{-1}g_k$，但 $\|p_k\|<\frac{\delta_1}{2}$，即 $x_{k+1}\in\Omega_1$，矛盾。故 $\forall$ $k\geqslant k_{j_0}$，　$x_k\in\Omega_2$。由于 $\{f_{l(k)}\}$ 单调下降且 x^* 为 f 在 Ω_2 内唯一的极小点，故 x_k 收敛于 x^*。

在算法中，有 $\|B_k^{-1}g_k\|\leqslant v_k$，或者 $\|p_k\|\leqslant v_k\leqslant\|B_k^{-1}g_k\|$ 总有　　$\|p_k\|\leqslant\|B_k^{-1}\|\|g_k\|$

由假设得

$$-\psi_k(p_k)\geqslant c_2\|g_k\|\mathrm{Min}\{v_k,\|g_k\|/\|B_k\|\}$$
$$\geqslant c_2\frac{\|p_k\|}{\|B_k^{-1}\|}\mathrm{Min}\{\|p_k\|,\|p_k\|/\|B_k\|\|B_k^{-1}\|\}$$
$$\geqslant c_2\frac{\|p_k\|^2}{\|B_k\|\|B_k^{-1}\|}$$

因 $f\in C^2$，$\nabla^2 f(x^*)$ 正定，对充分大之 k 有

$$-\psi_k(p_k)\geqslant\frac{c^2}{2}\frac{\|p_k\|^2}{\|\nabla^2 f(x^*)\|\|[\nabla^2 f(x^*)]^{-1}\|}$$

由(11)式　$|q_k - 1| \leqslant \dfrac{\|[\nabla^2 f(x^*)]^{-1}\| \|\nabla^2 f(x^*)\|}{c_2} \|p_k\|$

$$\lim_{k\to\infty} q_k = 1$$

由算法知，存在 k_0，$\forall\ k \geqslant k_0$，$v_{k0} < v_k$，又 $\lim\limits_{k\to\infty} g(x_k) = 0$，所以 $\|B_k^{-1} g_k\| \leqslant v_{k_0} < v_k$。故 k 充分大时，$p_k$ 恒为 $-\tau_k B_k^{-1} g_k$ 的形式，τ_k 由非单调 Armijo 规则决定。

因 $\lim\limits_{k\to\infty} q_k = 1$，$r < \dfrac{1}{2}$，存在 k_0，$\forall\ k \geqslant k_0$，$\|q_k\| \geqslant 2r$，取 $p_k = -B_k^{-1} g_k$，因 $\|B_k^{-1} g_k\| \leqslant v_{k_0} < v_k$，当 k 充分大时，

$$|q_k| = \frac{f_{k+1} - f_k}{\frac{1}{2} g_k^T B_k g_k} = \frac{f_{k+1} - f_k}{-\frac{1}{2} g_k^T p_k} \geqslant 2r$$

$$f(x_k + p_k) \leqslant f_k + r g_k^T p_k \leqslant f[x_{l(k)}] + r g_k^T p_k$$

故对充分大之 k，$\tau_k \equiv 1$，并使算法中第四步成立。即每次搜索向量 p_k 均为 Newton 方向，由引理2，$\{x_k\}$ 二次收敛于 x^*。

3　数值计算

在算法的执行中，为了节省校正 B_k^{-1} 的工作量，采用 Powell 在文中[8]提出的办法。另外，$\psi_k(p_k)$ 的计算直接采用公式：$\psi_k(p_k) = \dfrac{1}{2} g_k^T p_k$。

我们在文献和专著上选了四个经典试验问题，编制程序进行计算。

3.1　问题1：(L. R. Foulds)

$$\text{Min } f(x) = x_1^2 - 2x_1x_2 + x_2^2 - 4x_1,\quad x^* = (4,2), f^* = -8$$

3.2　问题2：(M. J. D. Powell)

$$\text{Min } f(x) = x_1^4 + 2x_1^2x_2 + x_2^4 - 21x_1^2 - 13x_2^2 - 14x_1 - 22x_2 + 170,$$
$$x^* = (3.58443, -1.84813)$$

3.3　问题3：(L. R. Foulds)

$$\text{Min } f(x) = 5x_1^2 + x_2^2 + x_3^2 - 4x_1x_2 - 2x_1 - 6x_3,\quad x^* = (1,2,3)$$

3.4　问题4：(H. Y. Huang and J. P. Chambliss)

$$Min\ f(x) = (x_1 + 10x_3)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^2 + 10(x_1 - x_4)^2 + (x_4 - x_5)^2,$$
$$x^* = (0,0,0,0,0), f^* = 0。$$

参数选取及计算结果见表1，表2。数值结果表明了算法的整体及局部快速收敛性质。要作进一步比较还需作实例计算。

表 1　参数的选取

r	r_1	r_2	r_3	v_c	u_1	u_2	M	θ	δ	L
0.3	1.2	0.2	0.1	1	2	0.5	2	0.7	0.5	0
0.2	1.1	0.3	0.2	0.8	1.5	0.9	4	0.6	0.4	1
0.1					0.55	5×10^{-1}				

表 2 计算结果

№	x_0	ε	k	x(k)	f_k	$\|\triangle x_k\|_\infty$
1	(1,4) (2,1) (6,4)	10^{-2} 10^{-3} 10^{-4}	7 12 10	(3.993203, 1.996031) (4.003139, 2.003031) (3.995599, 1.999149)	−7.999977 −7.999993 −7.999987	6.8×10^{-3} 3.1×10^{-3} 4.4×10^{-3}
2	(2,3) (4,0) (6,−2) (10,−1)	10^{-4} 10^{-4} 10^{-4} 10^{-5}	10 19 12 21	(3.584428, −1.848128) (3.584429, −1.848130) (3.584428, −1.848127) (3.584428, −1.848126)	1.55×10^{-11} 1.31×10^{-10} 1.39×10^{-11} 9.00×10^{-12}	2.00×10^{-5} 2.00×10^{-5} 3.00×10^{-6} 4.00×10^{-6}
3	(0.833, 1.55, 2.33) (2, 3.55, 5.33) (−1, 0, 7)	10^{-5} 10^{-4} 10^{-4}	10 15 20	(0.9999985, 1.9999923, 2.999991) (0.9998796, 1.999660, 2.997650) (0.999982, 1.999972, 2.999958)	−9.999903 −9.999943 −9.999987	8.3×10^{-6} 2.3×10^{-3} 4.2×10^{-5}
4	(1,1,1,1,1) (0.1,0.1,0.1,0.1,0.1) (−10,−13,−4,−7,−8) (10,10,10,−10,10)	10^{-4} — — —	30 10 60 70	(1.07×10^{-3}, 2.18×10^{-3}, -2.74×10^{-4}, 1.13×10^{-3}, 5.94×10^{-3}) (1.38×10^{-3}, 3.76×10^{-3}, 4.89×10^{-4}, 1.50×10^{-3}, 7.89×10^{-3}) (3.07×10^{-4}, -1.81×10^{-4}, 1.40×10^{-4}, 1.46×10^{-4}, 9.97×10^{-4}) (5.68×10^{-3}, 1.00×10^{-2}, 2.46×10^{-3}, 5.59×10^{-3}, 2.00×10^{-2})	4.32×10^{-8} 9.56×10^{-5} 1.60×10^{-4} 9.8×10^{-2}	5.94×10^{-3} 7.89×10^{-3} 9.97×10^{-4} 2.00×10^{-2}

如果注意到对参数的选取，在程序中加以自动调节，会加速收敛过程，使我们得到更精确的解。这些均有必要进一步研究。

参 考 文 献

1 AL—Baali M. Descent property and global convergence of the Fletcher—Reeves method with inexact line search. IMA. J. Numer. Analysis. 1985, 41(5): 18—23

2 Bultedu JP, Vial JP. A restricted trust region algorithm for unconstrained optimization. JOTA. 1985, 47(4): 413—435

3 More JJ. Recent developments in algorithms and software for trust region method. Math Prog. —the state of art. Spring V+erlag. Bonn. 1983

4 Schnable R B, Byrd R H. A family of trust region based algorithms for unconstrained minimization with strong global convergent property. SIAM J. Numer. Anal. 1985, 22(1): 23—40

5 张建中. 约束极值问题的 SQP 方法. 高校应用数学学报, 1987, 2(2): 21—25

6 Ortega JM, Rheinboldt WC. Iterative solution of nonlinear equations in several variables. Academic Press. 1970

7 Bertsekas DP. Constrained optimization and Lagrange multiplier methods. Academic Press. New York and London. 1982

8 Powell MJD. Updating conjugate directions by the BFGS formula. Math. Prog. 1987, 38(1): 29—46