国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第12卷第1期　1990年3月　Vol.12 No.1

一种无约束最优化方法

——恰当共轭方向法

张干宗　粟塔山

（系统工程与应用数学系）

摘　要　本文提出了一种求解无约束最优化问题的方法，称之为"恰当共轭方向法"。与通常的共轭方向法（如共轭梯度法）相比，这种方法可以节省大量的一维搜索工作量，并且还能提高收敛速率。

关键词　收敛，无约束最优化问题，恰当共轭方向法，共轭梯度法，一维搜索

分类号　O221.224

共轭方向法是求解无约束最优化问题 $\min\limits_{X\in E^n} f(X)$ 的一类重要的迭代算法。通常的共轭方向法，在每一个迭代循环中，构造一组相互共轭的向量 $Z_1^{(k)}$，…，$Z_n^{(k)}$，从第 k 个近似点 X_k 出发，依次沿着上述方向（或近似于它们的方向）对 f 作一维搜索，依次得点 $X_1^{(k)}$，…，$X_n^{(k)}$，然后以 $X_n^{(k)}$ 作为第 $k+1$ 个近似点 X_{k+1}（[1]、[2]）。本文提出的恰当共轭方向法，在每一个迭代循环中，构造一组特定的共轭向量，使得无需对每个方向作一维搜索，而只需沿其中一个方向进行一维搜索。此法可以节省一维搜索的工作量，还能提高收敛速率。

1　算法的基本思想

先考虑二次函数

$$f(X)=\frac{1}{2}X^TAX+b^TX+c \qquad (X\in E^n) \tag{1}$$

式中 A 为对称正定矩阵。设已知 $X_0\in E^n$，记 $Z=X-X_0$，则

$$f(X)=f(X_0+Z)=\frac{1}{2}Z^TAZ+(AX_0+b)^TZ+f(X_0) \tag{2}$$

如果 Z_1，…，Z_n 是一组关于 A 共轭的向量，则从 X_0 出发依次沿 Z_1，…，Z_n 作一维搜索，便能得出 f 的极小点。下面指出，如果取其中 $n-1$ 个方向，比如说 Z_1，…，Z_{n-1} 都与 $\nabla f(X_0)=AX_0+b$ 正交，则从 X_0 出发只需沿 Z_n 作一维搜索，便可得出 f 的极小点。

学术论文　1988年11月10日收稿

定理 1　对于正定二次函数(式(1))，设$Z_1,\cdots,Z_n$是关于A共轭的一组非零向量，且Z_1，…，Z_{n-1}都与$\nabla f(X_0)$正交，则

$$\min_{X\in E^n} f(X)=\min_{t\in R} f(X_0+tZ_n)$$

证明　由Z_1，…，Z_n线性无关，对任意$Z\in E^n$，有

$$Z=\sum_{i=1}^{n} t_i Z_i,\ t_i\in R\ (i=1,\cdots,n)$$

并由式(2)可得

$$f(X)=f(X_0+Z)=\frac{1}{2}\sum_{i=1}^{n} t_i^2(Z_i^T A Z_i)+t_n(AX_0+b)^T Z_n+f(X_0)$$

$$=\frac{1}{2}\sum_{i=1}^{n-1} t_i^2(Z_i^T A Z_i)+f(X_0+t_n Z_n)$$

再注意到

$$\min_{(t_1,\cdots,t_{n-1})\in E^{n-1}}\frac{1}{2}\sum_{i=1}^{n-1} t_i^2(Z_i^T A Z_i)=0$$

且必在$t_1=\cdots=t_{n-1}=0$达到其最小值。所以

$$\min_{X\in E^n} f(X)=\min_{(t_1,\cdots,t_n)\in E^n} f\left(X_0+\sum_{i=1}^{n} t_i Z_i\right)$$
$$=\min_{t_n\in R} f(X_0+t_n Z_n)$$

下面来构造满足定理1条件的共轭向量组。设

$$\left.\begin{aligned}&\nabla f(X_0)=\langle y_1,y_2,\cdots,y_n\rangle\neq 0^*\\ &|y_{i_0}|=\max_{1<i<n}\{|y_i|\}\end{aligned}\right\}\quad(3)$$

取y_{i_0}满足

令

$$\left.\begin{aligned}&u_1=\langle 1,0,\cdots,-y_1/y_{i_0},\quad 0,\cdots,0\rangle\\ &x_2=\langle 0,1,\cdots,-y_2/y_{i_0},\quad 0,\cdots,0\rangle\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &v_{i_0}=\langle 0,0,\cdots,-y_{i_0+1}/y_{i_0},1,\cdots,0\rangle\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &u_{n-1}=\langle 0,0,\cdots,-y_n/y_{i_0},\ 0,\cdots,1\rangle\end{aligned}\right\}\quad(4)$$

上述每一向量的分量中，除两个分量外，其他分量都取零值。然后令

$$\left.\begin{aligned}&Z_1=u_1\\ &Z_i=u_i-\sum_{j=1}^{i-1}\frac{u_i^T A Z_j}{Z_j^T A Z_j}Z_j\quad(i=2,\cdots,n-1)\\ &Z_n=-\nabla f(X_0)+\sum_{j=1}^{n-1}\frac{\nabla f(X_0)AZ_j}{Z_j^T A Z_j}Z_j\end{aligned}\right\}\quad(5)$$

* <…>表示列向量

定理 2 对于正定二次函数(方程(1))，按(3)～(5)式确定的Z_i $(i=1,\cdots,n)$ 是一组关于A共轭的非零向量，且满足

$$Z_i^T \Delta f(X_0)=0 \quad (i=1,\cdots,n-1) \tag{6}$$

证明 由(4)式可知$u_i^T \nabla f(X_0)=0$ $(i=1,\cdots,n-1)$，由(5)式可知Z_i是$u_1,\cdots,u_i$的线性组合，因此式(6)成立。

用数学归纳法容易验证$Z_1,\cdots Z_{n-1}$关于A两两共轭。又有

$$Z_i^T A Z_n = Z_i^T A\left[-\nabla f(X_0)+\sum_{j=1}^{n-1}\frac{\nabla f(X_0)^T A Z_j}{Z_j^T A Z_j}Z_j\right]=0 \quad (i=1,\cdots,n-1)$$

所以$Z_1,\cdots Z_n$是一组关于A共轭的向量。

最后证明$Z_i\neq 0$ $(i=1,\cdots,n)$。对于$1\leqslant i\leqslant n-1$，若有$Z_i=0$，由$Z_i$是$u_1,\cdots,u_i$的线性组合且组合系数不全为零可知，这时$u_1,\cdots,u_{n-1}$必线性相关。但由(4)式可知，这是不可能的。若$Z_n=0$，则有

$$\nabla f(X_0)=\sum_{j=1}^{n-1}\frac{\nabla f(X_0)^T A Z_j}{Z_j^T A Z_j}Z_j$$

上式两端与$\nabla f(X_0)$作内积，由(6)式可得$\nabla f(X_0)^T\nabla f(X_0)=0$，此与前提$\nabla f(X_0)\neq 0$相矛盾。

一般目标函数在其极小点附近可以用正定二次函数来近似。

$$f(X)\approx f(X_k)+\nabla f(X_k)^T(X-X_k)+\frac{1}{2}(X-X_k)^T\nabla^2 f(X_k)(X-X_k)$$

因此可考虑用$\nabla^2 f(X_k)$代替(5)式中的矩阵A．但是，如在迭代过程中直接计算二阶导数矩阵，其计算量很大。由于，当$\|Z\|$充分小时，有

$$\nabla^2 f(X_k)Z\approx\nabla f(X_k+Z)-\nabla f(X_k)$$

于是，对充分大的正数γ，有

$$\nabla^2 f(X_k)Z_j\approx\gamma\|Z_j\|\left[\nabla f\left(X_k+\frac{Z_j}{\gamma\|Z_j\|}\right)-\nabla f(X_k)\right] \tag{7}$$

因此，对于一般目标函数f，可按如下方法确定在X_k处的搜索方向Z_k．设

$$\left.\begin{aligned}&\nabla f(X_k)=\langle y_1^{(k)},y_2^{(k)},\cdots,y_n^{(k)}\rangle\neq 0\\ \text{取}y_{i_0}^{(k)}\text{满足}\quad &|y_{i_0}^{(k)}|=\max_{1\leqslant i\leqslant n}\{|y_i^{(k)}|\}\end{aligned}\right\} \tag{8}$$

令

$$\left.\begin{aligned}u_1^{(k)}&=\langle 1,0,\cdots,-y_1^{(k)}/y_{i_0}^{(k)},\quad 0,\cdots,0\rangle\\ u_2^{(k)}&=\langle 0,1,\cdots,-y_2^{(k)}/y_{i_0}^{(k)},\quad 0,\cdots,0\rangle\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ u_{i_0}^{(k)}&=\langle 0,0,\cdots,-y_{i_0+1}^{(k)}/x_{i_0}^{(k)},1,\cdots,0\rangle\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ u_{n-1}^{(k)}&=\langle 0,0,\cdots,-y_n^{(k)}/y_{i_0}^{(k)},\quad 0,\cdots,1\rangle\end{aligned}\right\} \tag{9}$$

$$\left.\begin{aligned} &Z_1^{(k)}=u_1^{(k)} \\ &Z_i^{(k)}=u_i^{(k)}-\sum_{j=1}^{i-1}a_{ij}^{(k)}Z_j^{(k)} \quad (i=2,\cdots,n-1) \end{aligned}\right\} \tag{10}$$

式中

$$a_{ij}^{(k)}=\frac{(u_i^{(k)})^T\nabla^2 f(X_k)Z_j^{(k)}}{(Z_j^{(k)})^T\nabla^2 f(X_k)Z_j^{(k)}} \quad \begin{pmatrix} i=2,\cdots,n-1; \\ j=1,\cdots,i-1 \end{pmatrix} \tag{11}$$

利用式(7)，并注意到$u_j^{(k)}$和$Z_j^{(k)}$ $(j=1,\cdots,n-1)$ 均与$\nabla f(X_k)$正交，实际计算时按如下公式确定(10)式中的系数：

$$a_{ij}^{(k)}=\frac{(u_i^{(k)})^T\nabla f\left(X_k+\dfrac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)}{(Z_j^{(k)})^T\nabla f\left(X_k+\dfrac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)} \quad \begin{pmatrix} i=2,\cdots,n-1; \\ j=1,\cdots,i-1 \end{pmatrix} \tag{12}$$

最后令

$$Z_k=Z_n^{(k)}=-\nabla f(X_k)+\sum_{j=1}^{n-1}\frac{\nabla f(X_k)^T\nabla^2 f(X_k)Z_j^{(k)}}{(Z_j^{(k)})^T\nabla^2 f(X_k)Z_j^{(k)}}Z_j^{(k)} \tag{13}$$

同理，实际计算时按下述公式确定搜索方向：

$$Z_k=-\nabla f(X_k)+\sum_{j=1}^{n-1}\frac{\nabla f(X_k)^T\left[\nabla f\left(X_k+\dfrac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)-\nabla f(X_k)\right]}{(Z_j^{(k)})^T\nabla f\left(X_k+\dfrac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)}Z_j^{(k)} \tag{14}$$

如果$\nabla^2 f(X_k)$正定，取γ充分大，可以保证

$$(Z_j^{(k)})^T\nabla f\left(X_k+\frac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)>0$$

这由下式可知：

$$\begin{aligned}(Z_j^{(k)})^T\nabla f\left(X_k+\frac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right)&=(Z_j^{(k)})^T\left[\nabla f(X_k)\right.\\ &\left.+\frac{1}{\gamma\|Z_j^{(k)}\|}\nabla^2 f(X_k)Z_j^{(k)}+0\left(\frac{1}{\gamma}\right)\frac{Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}\right]\\ &=\frac{(Z_j^{(k)})^T\nabla^2 f(X_k)Z_j^{(k)}}{\gamma\|Z_j^{(k)}\|}+0\left(\frac{1}{\gamma}\right)\frac{\|Z_j^{(k)}\|}{\gamma}\end{aligned}$$

对一般的连续可微函数，按上述方法得出的向量$Z_1^{(k)}$，$\cdots$，$Z_{n-1}^{(k)}$仍与$\nabla f(X_k)$正交，但$Z_1^{(k)}$，$\cdots$，$Z_{n-1}^{(k)}$，$Z_n^{(k)}$一般不再具有共轭的性质。从X_k出发，沿$Z_k=Z_n^{(k)}$对f作一维搜索，得点X_{k+1}. 若$\nabla f(X_{k+1})\neq 0$，便再从X_{k+1}出发重复上述做法。如此形成一个迭代算法，我们称它为恰当共轭方向法。

在实际应用上述方法时应取γ适当大。因为，从式(7)考虑，γ取得越大越好；但考虑到计算机的精度限制和误差的影响，γ又不能取得太大。为了提高收敛速率，应当针对不同的X_k和Z_j分别选取γ的值。为此在计算(12)和(14)式之前，可插入一个子程序去确定参数γ的值。但这样又会增加计算量。从实算情况来看，把γ的值固定影响不大。根据实算经验，取$\gamma=10$，效果就很好（在IBM—PC机上计算）。

2 算法步骤与收敛性

设$f(X)(X\in E^n)$连续可微。求解问题$\min\limits_{X\in E^n} f(X)$的恰当共轭方向法的步骤如下：

(1) 选定初始点X_0和精度要求ε. 并置$k:=0$.

(2) 计算$\nabla f(X_k)=\langle y_1^{(k)}, y_2^{(k)}, \cdots, y_n^{(k)}\rangle$，并检验$\|\nabla f(X_k)\|\leqslant\varepsilon$是否满足（也可改用其他停止准则）。若满足，停止计算，得近似解X_k；否则，转(3)。

(3) 按$|y_{i_0}^{(k)}|=\max\limits_{1\leqslant i\leqslant n}\{|y_i^{(k)}|\}$确定$i_0$（如有多个分量的绝对值同时达到最大，约定取其中下标最小者）。然后按(9)、(10)、(12)和(14)式算出$Z_i^{(k)}$ $(i=1,\cdots,n)$，并令$Z_k=Z_n^{(k)}$.

(4) 从X_k出发，沿Z_k进行一维搜索，求得λ_k，使

$$f(X_k+\lambda_k Z_k)=\min_{\lambda>0} f(X_k+\lambda Z_k)$$

并令$X_{k+1}=X_k+\lambda_k Z_k$. 然后置$k:=k+1$，返回(2).

定理 3 设$f:E^n\to R$具有一阶连续偏导数，给定$X_0\in E^n$，并设水平集$\{X\,|f(X)\leqslant f(X_0), X\in E^n\}$有界。令$\{X_k\}$是由上述恰当共轭方向法产生的点列，则$\{f(X_k)\}$为严格单调下降数列，且$\lim\limits_{k\to\infty} f(X_k)$存在。并且，$\{X_k\}$的任意聚点$X^*$都满足$\nabla f(X^*)=0$.

证明 由$Z_i^{(k)}$的构造可知

$$\nabla f(X_k)^T Z_i^{(k)}=0 \qquad (i=1,\cdots,n-1)$$

从而可得

$$\nabla f(X_k)^T Z_n^{(k)}=-\|\nabla f(X_k)\|^2$$

故知，当$\nabla f(X_k)\neq 0$时，必有$\nabla f(X_k)^T Z_n^{(k)}<0$. 从而有（用$\delta$表充分小正数）

$$f(X_{k+1})=f(X_k+\lambda_k Z_k)=\min_{\lambda} f(X_k+\lambda Z_k)\leqslant f(X_k+\delta Z_k)$$

$$=f(X_k)+\delta\nabla f(X_k)^T Z_k+0(\delta\|Z_k\|)<f(X_k)$$

所以$\{f(X_k)\}$是严格单调下降数列。由定理的条件可知数列$\{f(X_k)\}$有下界。因此$\lim\limits_{k\to\infty} f(X_k)$存在。记$\lim\limits_{k\to\infty} f(X_k)=f^*$. 因$\{X_k\}$为有界点列，故必有收敛子列。设子列$\{X_k: k\in\mathscr{K}\}$收敛于$X^*$. 下证$\nabla f(X^*)=0$.

假如$\nabla f(X^*)\neq 0$，则可按恰当共轭方向法产生f在X^*处的下降搜索方向Z^*. 即

$$Z^*=-\nabla f(X^*)+\sum_{j=1}^{n-1}\frac{\nabla f(X^*)^T\left[\nabla f\left(X^*+\dfrac{Z_j^*}{\gamma\|Z_j^*\|}\right)-\nabla f(X^*)\right]}{(Z_j^*)^T\nabla f\left(X^*+\dfrac{Z_j^*}{\gamma\|Z_j^*\|}\right)}Z_j^*$$

式中Z_j^* $(j=1,\cdots,n-1)$的计算公式可相应列出。于是，对充分小的正数λ，有

$$f(X^*+\lambda Z^*)<f(X^*)$$

同时，又有

$$f(X_{k+1})=f(X_k+\lambda_k Z_k)\leqslant f(X_k+\lambda Z_k)$$

再注意到$\{X_{k+1}: k\in\mathscr{K}\}$也是有界点列，设其子列$\{X_{k+1}: k\in\mathscr{K}'\}$收敛于$\bar{X}$，这里$\mathscr{K}'\subset\mathscr{K}$. 在上述不等式两端，令$k\in\mathscr{K}'$，$k\to\infty$，取极限。由$\nabla f(X)$连续和$Z_k$的构造可知

$\lim\limits_{k\to\infty} Z_k = Z^*$. 因此可得

$$f(\bar{X}) \leqslant f(X^* + \lambda Z^*) < f(X^*)$$

但是，由 f 的连续性，

$$f(\bar{X}) = \lim_{\substack{k\to\infty \\ k\in\mathscr{K}'}} f(X_{k+1}) = f^* = \lim_{\substack{k\to\infty \\ k\in\mathscr{K}'}} f(X_k) = f(X^*)$$

至此得出矛盾。所以 $\nabla f(X^*) = 0$.

推论　若 $f: E^n \to R$ 是满足定理3条件的凸函数，则由恰当共轭方向法产生的点列 $\{X_k\}$ 的任意聚点都是问题 $\min\limits_{X\in E^n} f(X)$ 的最优解。

由上节的论述可知，恰当共轭方向法具有二次终止性。关于此法的收敛速率，以下述结论可得到初略了解。

定理 4　对于二阶连续可微函数 f，设 $\nabla f(X_k) \neq 0$，且 $\nabla^2 f(X_k)$ 正定，则按(8)、(9)、(10)、(11)和(13)式确定的搜索方向 Z_k 与牛顿方向一致。即有下式成立（α 为某一正数）：

$$-[\nabla^2 f(X_k)]^{-1} \nabla f(X_k) = \alpha Z_k$$

证明　由定理2可知，按(8)～(11)和(13)式确定的 $Z_1^{(k)}, \cdots, Z_{n-1}^{(k)}$，$Z_n^{(k)}$ 是一组关于 $\nabla^2 f(X_k)$ 共轭的非零向量，且 $Z_1^{(k)}, \cdots, Z_{n-1}^{(k)}$ 均与 $\nabla f(X_k)$ 正交。于是向量 $-[\nabla^2 f\cdot(X_k)]^{-1} \nabla f(X_k)$ 可由 $Z_1^{(k)}, \cdots, Z_n^{(k)}$ 线性表出：

$$-[\nabla^2 f(X_k)]^{-1} \nabla f(X_k) = \alpha_1 Z_1^{(k)} + \cdots + \alpha_{n-1} Z_{n-1}^{(k)} + \alpha_n Z_n^{(k)}$$

上式两端同乘 $\nabla^2 f(X_k)$，然后与 $Z_i^{(k)}$ $(i=1,\cdots,n-1)$ 作内积，便可得知 $\alpha_i = 0$ $(i=1, \cdots, n-1)$。因此

$$-[\nabla^2 f(X_k)]^{-1} \nabla f(X_k) = \alpha_n Z_n^{(k)}$$

式中

$$\alpha_n = \frac{-(Z_n^{(k)})^T \nabla f(X_k)}{(Z_n^{(k)})^T \nabla^2 f(X_k) Z_n^{(k)}} = \frac{\|\nabla f(X_k)\|^2}{(Z_n^{(k)})^T \nabla^2 f(X_k) Z_n^{(k)}} > 0$$

由此可知，在极小点附近，由(8)、(9)、(10)、(12)和(14)式所产生的恰当共轭方向法的搜索方向近似于牛顿方向。但恰当共轭方向法的搜索方向并不等同于牛顿方向。它只用到目标函数的梯度，而勿需计算Hesse矩阵。

3　实算举例

本文用恰当共轭方向法在IBM—PC/XT计算机上解算了大量题目。一维搜索采用进退法。实算结果表明，恰当共轭方向法的收敛速率明显优于共轭梯度法。从迭代次数看，也明显优于DFP变尺度法([3]、[4])。下面列举几个计算结果(参数 γ 都取为10)。

例 1　目标函数是香蕉函数

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

求最小值。问题的准确最优解为 $X^* = \langle 1, 1\rangle$，最优值为 $f^* = 0$. 取初始点 $X^0 = \langle -1.2, 1\rangle$. 用恰当共轭方向法迭代16次得

$$x_1 = 1.0000692510,\quad x_2 = 1.0001453045$$

$$f(x_1,x_2)=9.4166899682\times10^{-9}$$

用DFP变尺度法计算，达到同级精度的结果，需迭代32次。

例 2　$\min f(x_1,x_2)=-x_1^2\exp[1-x_1^2-2.25(x_1-x_2)^2]$ ．此问题有两个最优解：$X_1^*=\langle1,1\rangle$，$X_2^*=\langle-1,-1\rangle$．最优值$f^*=-1$．取初始点 $X_1^0=\langle0.1,0.1\rangle$．用恰当共轭方向法迭代 6 次得

$$x_1=0.99927266533,\ x_2=0.99917796979$$

$$f(x_1,x_2)=-0.99999892153$$

用DFP变尺度计算，达到同级精度的结果，需迭代10次。取初始点$X_2^0=\langle0.1,-0.2\rangle$．用恰当共轭方向法迭代 6 次得

$$x_1=-0.99948567980,\ x_2=-0.99984578714$$

$$f(x_1,x_2)=-0.99999917908$$

用DFP变尺度法计算，达到同级精度的结果，需迭代12次。

例 3　$\min f(x_1,x_2,x_3,x_4)=(x_1+10x_2)^2+5(x_3-x_4)^2+(x_2-2x_3)^4+10(x_1-x_4)^4$．此问题的最优解为$X^*=\langle0,0,0,0\rangle$，最优值$f^*=0$．取初始点$X^0=\langle-3,-1,0,1\rangle$．用恰当共轭方向法迭代12次得：

$$x_1=5.6136169394\times10^{-3},\ x_2=-5.5550429290\times10^{-4}$$

$$x_3=3.0748492971\times10^{-3},\ x_4=3.0687649197\times10^{-3}$$

$$f(x_1,x_2,x_3,x_4)=6.0568126517\times10^{-9}$$

用DFP变尺度法计算，达得同级精度的结果，需迭代34次。

例 4　$\min f(x_1,\cdots,x_{10})=\sum_{i=1}^{9}(x_i-x_{i+1}^2)^2+(1-x_1)^2+(1-x_{10})^2$

此问题的最优解为 $X^*=\langle1,1,\cdots,1\rangle$，最优值 $f^*=0$．取初始点 $X^0=\langle1.5,0.5,2,\cdots,2\rangle$．用恰当共轭方向法迭代 9 次得

$$x_1=1.0000009403,\ x_2=1.0000014903,\ x_3=1.0000037625$$

$$x_4=1.0000071435,\ x_5=1.0000016902,\ x_6=1.0000002067$$

$$x^7=0.99999975829, x_8=0.99999972949, x_9=1.0000001349$$

$$x_{10}=1.0000004028,\ f(x_1,\cdots,x_{10})=1.6949465213\times10^{-10}$$

运算时间41秒。用DFP变尺度法计算，达到同级精度的结果，需迭代57次，运算时间87秒。

其他计算实例不一一列举。总之，从实算情况看，恰当共轭方向法的收敛速率和计算稳定性都是良好的。关于此法的收敛速率的确切估计有待进一步研究。

参考文献

[1] Bazaraa M S, Shetty C M 著；王化存，张春柏译．非线性规划——理论与算法．贵州人民出版社，1986：382～407

[2] 邓乃扬，诸梅芳．最优化方法．辽宁教育出版社，1987

[3] 魏权龄等编著．数学规划与优化设计．国防工业出版社，1984：299～307

[4] Himmelblau D M 著；张义燊等译．实用非线性规划．科学出版社，1981

国 防 科 技 大 学 学 报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第12卷第1期　1990年8月　Vol.12 No.1

一类集值映射的高阶微分

吴 孟 达

（系统工程与应用数学系）

摘　要　在文[1]中，H.T.Banks与M.Q.Jacobs讨论了集映$\Omega: R^m \to \mathscr{B}(R^n)$的一阶微分，分别给出了$\Omega$可微的充分条件与必要条件。本文改进了文[1]的工作，证明了[1]中条件(3.2)的限制是不必要的，从而得到了Ω可微的充分必要条件，及较为简洁的Ω微分解析表达式。在文中也讨论了Ω的k阶微分，得到了Ωk阶可微的充分必要条件及k阶微分的解析表达式。

关键词　集值映射，Hausdorff距离，集映微分，集映高阶微分

分类号　O172.1

1　预备知识

R^n上的范数取作$\|x\| = \max\limits_{1<j<n} |x_j|$，$R^n$中非空紧凸集全体按Hausdorff距离$d_H$所成的度量空间记作$\mathscr{B}(R^n)$，$(d_H(A,B) \triangleq \max\{\sup\limits_{x\in A} d(x,B), \sup\limits_{y\in B} d(y,A)\})$，根据Radstrom嵌入

研究简报　1988年6月6日收稿

An Unconstrained Optimization Technique —Proper Conjugate Direction Method

Zhang Ganzong　Su Tashan

(Department of Applied Mathematics and System Engineering)

Abstract

In this paper, a method for solving unconstrained optimization problems, which is called "proper conjugate direction method" , is developed. Compared with general conjugate direction method,such as the conjugate gradient method, this method will save a great deal of one-dimensional search work and raise the rate of convergence.

Key words: convergence, unconstrained optimizatien problem, proper conjugate direction method, conjugate gradient method, one-dimensional search