第10卷 第2期
华侨大学学报自然科学版
JOURNAL OF HUAQIAO UNIVERSITY
(NATURAL SCIENCE)
Vol.10 No.2
1989年4月
Apr. 1989

函数极小化的最优单纯形方法

蔡火萤

(应用数学系)

摘 要

本文提出一个函数极小化的最优单纯形方法，它是用一维最优点 X^* 来代替当前单纯形中函数最高值顶点 X^h。

一、引 言

1962年，Spendley，Hext 和 Himsworth 首先在文[1]中，提出用单纯形方法求解无约束极小化问题，他们所使用的单纯形，只限于正规单纯形，且在应用过程中没有延伸步骤和压缩步骤，只是采用一系列的反射步骤来搜索目标函数的极小点。1965年，Nelder 和 Mead 在文[2]中对单纯形方法加以改善，他们把正规单纯形扩大到任意单纯形。而且在使用中增加延伸步骤和压缩步骤，使方法变得灵活多样，提高了计算效率，成为现代数值方法中常用而有效的方法之一[3]。本文对目前使用的单纯形方法，提出进一步的改善，它的主要思想是用一维最优点 X^* 来代替当前单纯形中的函数最高值顶点 X^h。

二、方法的建立

考虑 R^n 上的实值函数 $f(X)$。为求 $f(X)$ 在 R^n 中的极小点，首先在 R^n 中适当取出 $(n+1)$ 个点 X_1，X_2，…，X_{n+1}。使它们构成 R^n 中的一个单纯形，然后求出目标函数 $f(X)$ 在这个单纯形各点处的值：$f(X_1)$，$f(X_2)$，…，$f(X_{n+1})$，并从中选出目标函数的最高值顶点 X^h 和最低值顶点 X^l，即

$$f(X^h)=\max\{f(X_1),\ f(X_2),\ \cdots,\ f(X_{n+1})\},$$
$$f(X^l)=\min\{f(X_1),\ f(X_2),\ \cdots,\ f(X_{n+1})\}.$$

再求单纯形除最高值顶点 X^h 以外的各顶点的平均形心点 $\bar{X}$，即 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n+1}X_i$，$X_i\neq X^h$。然

本文1987年12月22日收到。

后在直线 $X^h+\alpha(\bar{X}-X^h)$，$\alpha\in R$ 上求目标函数 $f(X)$ 的一维最优点 X^*，即 $X^*=X^h+\alpha^*(\bar{X}-X^h)$。而

$$f\{X^h+\alpha^*(\bar{X}-X^h)\}=\min_{\alpha\in R}f\{X^h+\alpha(\bar{X}-X^h)\},$$

若 $f(X^l)\geqslant f(X^*)$，则用 X^* 代替 X^h，并得到一个新的单纯形. 否则，把原单纯形的各顶点与目标函数最低值顶点 X^l 之间的距离各缩小一半，即命

$$\hat{X}_i=X_i+(1/2)(X^l-X_i),\qquad i=1,2,\cdots,n+1,$$

再用 $\hat{X}_i$ 代替 X_i，$i=1,2,\cdots,n+1$，综上所述，得到下述的计算步骤：

（1） 给出允许精度 $\varepsilon>0$，给出初始单纯形的各顶点 X_1，X_2，…，X_{n+1}。

（2） 计算目标函数 $f(X)$ 在单纯形各顶点处的值：$f(X_1)$，$f(X_2)$，…，$f(X_{n+1})$。

（3） 选出单纯形中目标函数最高值顶点 X^h 和最低点 X^l，即

$$f(X^h)=\max\{f(X_1),f(X_2),\cdots,f(X_{n+1})\};\ f(X^l)=\min\{f(X_1),f(X_2),\cdots,f(X_{n+1})\}.$$

（4） 求出单纯形中目标函数最高值点 X^h 以外的各顶点的平均形心点 $\bar{X}$，即

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n+1}X_i,\qquad X_i\neq X^h.$$

（5） 检查精度. 若 $\left\{\frac{1}{n+1}\sum_{i=1}^{n+1}[f(X_i)-f(\bar{X})]^2\right\}^{\frac{1}{2}}\leqslant\varepsilon$，则停机，并以 X^l 作为目标函数 $f(X)$ 在 R^n 中满足精度要求的近似极小点，否则转去执行（6）。

（6） 作一维搜索，求出目标函数的一维极小点 X^*，即 $X^*=X^h+\alpha^*(\bar{X}-X^h)$。而 $f(X^h+\alpha^*(\bar{X}-X^h))=\min_{\alpha\in R}f(X^h+\alpha(\bar{X}-X^h))$。若 $f(X^l)\geqslant f(X^*)$，则用 X^* 代替 X^h，得到下一个单纯形，其顶点仍记 X_1，X_2，…，X_{n+1}，并且转去执行（2）。否则令 $\hat{X}_i=X_i+(1/2)(X^L-X_i)$，$i=1,2,\cdots,n+1$。并把 $\hat{X}_i$ 代替 X_i，$i=1,2,\cdots,n+1$。得到一个新的单纯形，仍把它的顶点记作 X_1，X_2，…，X_{n+1}，并转去执行（2）。

三、在 n 维空间中一维极小点的计算方法

在 n 维空间中，计算目标函数的一维极小点可采用下面两步：第一步，建立"两头高中间低"的三个基本点. 首先作一个辅助函数，令

$$\Phi(\alpha)=f(X^h+\alpha(\bar{X}-X^h)),\qquad \alpha\in R,$$

然后适当选取一个初始点 α_0 和一个初始步长 $h_0>0$，并计算 $\Phi(\alpha_0)$ 和 $\Phi(\alpha_0+h_0)$。若 $\Phi(\alpha_0)\geqslant(\alpha_0+h_0)$，则记 $\alpha_1=\alpha_0+h_0$，$h_1=2h_0$，计算 $\Phi(\alpha_1)$、$\Phi(\alpha_1+h_1)$、若 $\Phi(\alpha_1)\geqslant\Phi(\alpha_1+h_1)$，则记 $\alpha_2=\alpha_1+h_1$，$h_2=2h_1$。这样一直做下去，直到求出一个 K，使得 $\Phi(\alpha_k)<\Phi(\alpha_k+h_k)$ 时为止。若 $\Phi(\alpha_0)<\Phi(\alpha_0+h_0)$，则记 $\alpha_1=\alpha_0$，$h_1=-(1/2)h^0$。计算 $\Phi(\alpha_1)$、$\Phi(\alpha_1+h_1)$。若 $\Phi(\alpha_1)\geqslant\Phi(\alpha_1+h_1)$，则记 $\alpha_2=\alpha_1+h_1$，$h_2=2h_1$。这样一直进行下去，直到某一个 K，使得成立关系式 $\Phi(\alpha_k)<(\alpha_k+h_k)$ 时为止。不论上面的两种情况哪一种成立，其最后三点都具有两头高中间低的性质。第二步，在上面所得到的三个基本点处，建立一个二次插值多项式函数，并且用它的准确极小点来近似目标函数的极小点。假设三个基本点是 $\bar{\alpha}_1$、$\bar{\alpha}_2$、$\bar{\alpha}_3$、二次插值多

项式是 $P(\alpha)=a\alpha^2+b\alpha+c$，其中 a、b、c 是待定系数，为确定这些系数，可设

$$P(\bar{\alpha}_i)=\Phi(\bar{\alpha}_i),\qquad i=1,\ 2,\ 3$$

由于目标函数 $\Phi(\alpha)$ 在三个基本点处具有两头高中间低的性质，因此这个二次插值多项式也必然具有二头高中间低的性质，由此推知 $a>0$，且二次插值多项式的准确极小点为 $\alpha^*=-b/2a$。若 $\|\Phi(\alpha^*)-P(\alpha^*)\|\leqslant\varepsilon$，则用 α^* 作为目标函数 $\Phi(\alpha)$ 满足精度要求的极小点的近似值。否则在四个点 $\bar{\alpha}_1$、$\bar{\alpha}_2$、$\bar{\alpha}_3$、α^* 中挑选出三个基本点，其方法是：若 $\Phi(\alpha^*)\geqslant\Phi(\bar{\alpha}_2)$，则用 $\bar{\alpha}_2$ 作为新的 $\bar{\alpha}_2$，而 $\bar{\alpha}_2$ 的左、右邻接点分别为新的 $\bar{\alpha}_1$ 和 $\bar{\alpha}_3$；若 $\Phi(\alpha^*)<\Phi(\bar{\alpha}_2)$，则用 α^* 作为新的 $\bar{\alpha}_2$，而它的左、右邻近点分别为新的 $\bar{\alpha}_1$ 和 $\bar{\alpha}_3$。不管哪种情况成立，我们都得到新的三个基本点 $\bar{\alpha}_1$、$\bar{\alpha}_2$、$\bar{\alpha}_3$。重复这些步骤可求出目标函数满足精度要求的极小点的近似值。若 $\alpha^*=\bar{\alpha}_2$，且未满足精度要求，可对 α^* 作微小摄动。

四、效 果 分 析

为了说明最优单纯形方法的理论意义和实用价值，对以下三种情况将最优单纯形方法和单纯形方法进行比较。

（1） 设目标函数为

$$f(X)=\begin{cases}2 & X\leqslant 0\\(X-1)^2 & 0<X<2,\\1 & X\geqslant 2\end{cases}$$

如图 1 所示，自变量区域 R 中的单纯形为 $\overline{AB}$，其中 $A=0$，$B=2$，因为 $f(A)>f(B)$，由单纯形方法知，要求下一个单纯形，必须求 A 点关于 B 点的反射点 C，从而得到下一个单纯形

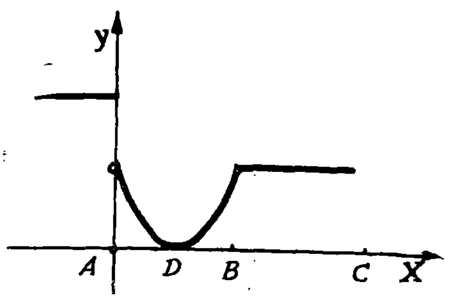

图 1

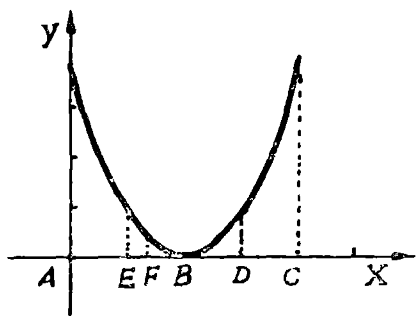

图 2

$\overline{BC}$。这样目标函数的真正最优点 $D=1$ 被“跳过去”了，说明在这种情况下，采用单纯形方法找不到目标函数的最优点，而采用最优单纯形方法就不会发生这种情况。

（2） 设目标函数为

$$f(X)=(X-2)^2,\qquad -\infty<X<+\infty,$$

如图 2 所示，自变量区域 R 中的单纯形为 $\overline{AB}$，其中 $A=0$，$B=2$，因为 $f(A)>f(B)$，由单

纯形方法知，要求下一个单纯形，必须求 A 点关于 B 点的反射点 C。这时 $f(A)\geqslant f(C)$，$f(C)>f(B)$，因此必须作压缩运算，求出压缩点 D。新的单纯形为 $\overline{BD}$，由于 $f(D)>f(B)$，要求下一个单纯形，先求 D 点关于 B 点的反射点 E。由于 $f(D)\geqslant f(E)$，且 $f(E)>f(B)$，必须作压缩运算，求出压缩点 F，得到下一个单纯形 $\overline{BF}$。反复进行这种运算，得到一个衰减的单纯形序列。这样虽然可以求出目标函数的近似最优点，但计算量较大，而采用最优单纯形方法，就不会发生这种情况。

（3） 设目标函数为

$$f(X)=\begin{cases}(X+2)^2, & -\infty<X\leqslant 0\\ 1/2\quad, & 0<X<+\infty\end{cases}$$

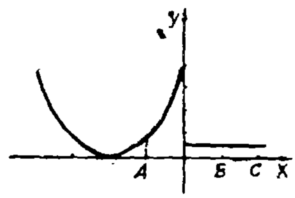

图 3

如图3所示，自变量区域 R 中的单纯形为 $\overline{AB}$，因为 $f(A)>f(B)$，$A=-1$，$B=1$。由单纯形方法知，要求下一个单纯形，必须求 A 点关于 B 点的反射点 C，从而得到下一个单纯形 $\overline{BC}$，而目标函数的真正最优点是在 A 的左方，采用单纯形方法永远达不到目标函数的最优点，而采用最优单纯形方法就不会发生这种情况。以上三种情况，不仅在理论上而且在实用上都说明最优单纯形方法比单纯形方法好。

最后，我们再用一个数值例子来对比两种算法的收敛速度。今以目标函数

$$f(X)=X_1^2+X_2^2-12X_1-8X_2+62$$

为例，用10组不同初始数据(初始点和步长)，分别用两种不同算法，并将计算结果列于表1、2(本程序在VAX-11/750机运行通过)。

从表1和表2对比可以看出，最优单纯形方法比单纯形方法的收敛速度块一倍。

表　1　　最优单纯形方法

初始点	步长	迭代次数	最优点	最优值
(12.000000, 10.000000)	8.000000	14	(6.000483,4.000767)	10.000000
(15.000000, 15.000000)	10.000000	12	(6.001857,4.002287)	10.0000008
(−5.000000, 4.000000)	5.000000	20	(6.018037,3.997044)	10.000328
(−5.000000, 6.000000)	8.000000	13	(6.000603,4.000083)	10.000000
(20.000000, 14.000000)	10.000000	18	(6.000060,3.999912)	9.999996
(−2.000000, −4.000000)	8.000000	13	(5.999424,3.999985)	10.000000
(16.000000, −4.000000)	8.000000	11	(5.999446,3.998163)	10.000000
(1.500000, 2.900000)	4.000000	12	(5.997469,3.999163)	10.000004
(−8.000000, 6.000000)	8.000000	17	(5.999419,3.999913)	9.999996
(−12.000000,−15.600000)	11.000000	17	(5.992192,3.997092)	10.000065

表 2　单纯形方法

初始点	步长	迭代次数	最优点	最优值
(12.000000, 10.000000)	8.000000	28	(5.998097,3.997369)	10.000008
(15.000000, 15.000000)	10.000000	27	(5.997071,3.997914)	10.000011
(−5.000000, 4.000000)	5.000000	27	(5.997920,4.003235)	10.000015
(−5.000000, 6.000000)	8.000000	24	(6.005905,3.997972)	10.000038
(20.000000, 14.000000)	10.000000	29	(6.001657,4.000678)	10.000004
(−2.000000, −4.000000)	8.000000	25	(6.004334,3.997890)	10.000019
(16.000000, −4.000000)	8.000000	26	(6.002945,3.998688)	10.000008
(1.500000, 2.900000)	4.000000	23	(6.001921,3.997730)	10.000008
(−8.000000, 6.000000)	8.000000	26	(6.000759,4.004056)	10.000023
(−12.000000, −15.600000)	11.000000	30	(6.002944,3.998425)	10.000008

参考文献

[1] Spendley, W., Hext, G.R. and Himsworih, F.R., Seqauential Application of Simplex Designs in Optimisation and Evolutionary Operation, *Technometrics*, 4 (1962), 441—461.

[2] Nelder, J.A. and Mead, R., A Simplex Method for Function Minimization, *Computer J.*, 7 (1965), 308—313.

[3] Avriel, M., *Nonlinear Programming: Analysis and Methods*, Prentice-Hall Internaticnal INC., London, (1976).

An Optimal Simplex Method for Function Minimization

Cai Huoying

Abstract

For function minimization, this paper poses an optimal simplex method which is the further improvement of the work of Nelder and Mead. The author's main idea lies in the replacement of X^h, the vertex of highest functional value in the existing simplex method with X^*, the one-dimensional optimal point. Here,

$$X^* = \overline{X} + \alpha^*(\overline{X} - X^h)$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n+1} X_i, \quad X_i \neq X^h$$

If $f(X^L) \geqslant f(X^*)$, then X^h is replaced by X^* and a new optimal simplex is resulted; otherwise, the distance between the vertices of original simplex and the vertex X^l of lowest functional value will be reduced to a half.