对一个求ε最优解算法的修正

邓乃扬[①] 王兆智
（基础科学部）

摘 要 考虑在实际中常用到的非线性最小二乘问题，指出了某文算法收敛性证明中的错误。对该算法加以修正后，证明了有限步终止性。

关键词 非线性最小二乘问题；简单界约束；ε最优解
中图分类号 O241.2

Correction on Algorithm of ε-optimal Solution

Deng Naiyang Wang Zhaozhi
(Department of Basic Sciences)

Abstract An error about the convergence proved in some paper is pointed out. After correcting the algorithm, the convergence is proved.

Key words nonlinear least square problem; simple bounds; ε-optimal solution

考虑在实际应用中经常用到的非线性最小二乘问题[1,2]

$$(\mathrm{P})\qquad \min\ f(x)=\|F(x)\|^2/2=\sum_{i=1}^{m}F_i^2(x)/2$$
$$\text{s.t.}\quad l\leqslant x\leqslant u \qquad x\in \mathrm{R}^n$$

这里 $F(x)=(F_1(x),\cdots,F_m(x))$，$F_i(x)\in C^2$；$l,u$ 是 n 维向量。

定义点 x 的 $\varepsilon(\varepsilon>0)$ 起作用集为

$$A^{\varepsilon}=A_l^{\varepsilon}\cup A_u^{\varepsilon}，其中\ A_l^{\varepsilon}=\{i\,|\,x_i<l_i+\varepsilon\}，A_u^{\varepsilon}=\{i\,|\,x_i>u_i-\varepsilon\}$$

下面条件称为ε最优性条件：

$$\begin{cases} |g_i|<\tilde{\varepsilon} & i\notin A^{\varepsilon} \quad (1)\\ g_i>-\tilde{\varepsilon} & i\in A_l^{\varepsilon} \quad (2)\\ g_i<\tilde{\varepsilon} & i\in A_u^{\varepsilon} \quad (3)\end{cases}$$

其中 $g_i=\partial f(x)/\partial x_i$，$\tilde{\varepsilon}>0$。相应地，满足ε最优性条件的解 x 称为ε最优解。

1 算法及其修正

考虑子问题

收稿日期：1996-01-26
① 邓乃扬，北京清华东路17号中国农业大学（东校区）71信箱，100083

$$(\mathrm{P}_1)\quad \min \psi_1(p)=g^{\mathrm{T}}p+p^{\mathrm{T}}Bp/2$$
$$\text{s.t.}\quad \|Dp\|\leqslant \Delta \qquad p\in \mathrm{R}^n$$
$$p_i=0 \qquad i\in A^{\varepsilon}$$

其中 $B=J(x)^{\mathrm{T}}J(x)$，$J(x)=(\nabla F_1\quad \nabla F_2\quad \cdots\quad \nabla F_n)^{\mathrm{T}}$；$D=\mathrm{diag}(1/d_i)$，$d_i=\min\{x_i-l_i, u_i-x_i\}$，$i=1,\cdots,n$；$\Delta\in(0,1)$。

文[1]算法如下。

1)给 $\varepsilon,\tilde{\varepsilon}(>0)$，$x$ 是初始可行内点，$\Delta\in(0,1)$。

2)确定 $A^{\varepsilon}, A_l^{\varepsilon}, A_u^{\varepsilon}$，进行最优性检验：

a.若 x 为 ε 最优解则停，否则继续；

b.若式(2)或(3)不成立，令 $i^*=\mathrm{argmax}\{|g_i|\,|\,i\in A^{\varepsilon}\}$，$A^{\varepsilon}=A^{\varepsilon}-\{i^*\}$，求解置信域子问题$(\mathrm{P}_1)$，使

$$\rho=\frac{f(x+p)-f(x)}{\psi_1(p)}>\mu,\quad \mu\in(0,1) \tag{4}$$

转 3)，否则继续；

c.求解置信域子问题(P_1)，使式(4)成立，转 3)。

3)令 $x=x+p$，转 2)。

文[1]在证明收敛性时，首先依据

$$d_i\geqslant\varepsilon \qquad i\notin A^{\varepsilon} \tag{5}$$

证明了每迭代一步，对应子问题(P_1)均有

$$-\psi_1(p)\geqslant(\tilde{\varepsilon}/2)\min\{\varepsilon\Delta,\tilde{\varepsilon}L\}$$

即下文中引理 1，而式(5)对 2)中的情况 b 不成立，并且 b 中 i^* 不一定违反式(2)或(3)；因此其算法收敛性的证明是不正确的。

笔者将算法步 2)作如下修改：

2′)进行最优性检验：

a.若 x 为 ε 最优则停，否则继续；

b.若式(1)不成立，求置信域子问题(P_1)，使式(4)成立，转 3)，否则继续；

c.令 $i^*=\mathrm{argmax}\{|g_i|\,\|\,i\in A^{\varepsilon}$，且式(2)或(3)不成立$\}$，$s=-\dfrac{g_{i^*}}{|g_{i^*}|}e_i$，$e_i$ 是单位向量，求单变量置信域子问题

$$(\mathrm{P}_2)\quad \min\psi_2(\lambda)=\lambda g^{\mathrm{T}}s+\lambda^2 s^{\mathrm{T}}Bs/2$$
$$\text{s.t.}\quad 0<\lambda<\tilde{d}_i\Delta$$

其中 $\tilde{d}_i=\max\{x_i-l_i, u_i-x_i\}$，使

$$\rho=\frac{f(x+\lambda s)-f(x)}{\psi_2(\lambda)}>\mu,\quad \mu\in(0,1) \tag{6}$$

转 4)。

4)令 $x=x+\lambda s$，转 2)。

2 收敛性

引理 1[1] x 是(P)的可行内点，p 是(P_1)的解。设存在 $i\notin A^{\varepsilon}$，使 $|g_i|\geqslant\tilde{\varepsilon}$，则存在 $L_1>0$(L_1

与 x,p 无关)，使

$$-\psi_1(p)\geqslant(\tilde{\varepsilon}/2)\min\{\varepsilon\Delta,\tilde{\varepsilon}L_1\}$$

引理 2　x 是(P)的可行内点，λ 是(P_2)的解。$\|s\|=|g_i|\geqslant\tilde{\varepsilon}$，则存在 $L_2>0$(L_2 与 x,λ 无关)，使

$$-\psi_2(\lambda)\geqslant(\tilde{\varepsilon}/2)\min\{\varepsilon'\Delta,\tilde{\varepsilon}L_2\}$$

其中 $\varepsilon'=u_i-l_i-\varepsilon$。

证明　因为

$$\psi_2(t)=-t|g_i|+\frac{1}{2}t^2\|J_i\|^2,\quad J_i=J(x)e_i=(\partial F_1/\partial x_i,\cdots,\partial F_n/\partial x_i)^{\mathrm{T}}$$

所以(P_2)的解为

$$\lambda=\begin{cases}|g_i|/\|J_i\|^2 & |g_i|\leqslant\|J_i\|^2\Delta\tilde{d}_i\\ \Delta\tilde{d}_i & |g_i|>\|J_i\|^2\Delta\tilde{d}_i\end{cases}$$

所以

$$\psi_2(\lambda)=-\frac{1}{2}\frac{|g_i|}{\|J_i\|^2}$$

或

$$\psi_2(\lambda)=-\Delta\tilde{d}_i|g_i|+\frac{1}{2}\Delta^2\tilde{d}_i^2\|J_i\|^2\leqslant-\Delta\tilde{d}_i|g_i|+\frac{1}{2}\Delta\tilde{d}_i|g_i|=-\frac{1}{2}\Delta\tilde{d}_i|g_i|$$

从而

$$\begin{aligned}-\psi_2(\lambda)\geqslant&\min\left\{\frac{1}{2}\Delta\tilde{d}_i|g_i|,\frac{1}{2}\frac{|g_i|}{\|J_i\|^2}\right\}=\\&\frac{1}{2}|g_i|\min\left\{\Delta\max(x_i-l_i,u_i-x_i),\frac{|g_i|}{\|J_i\|^2}\right\}\geqslant\\&\frac{1}{2}|g_i|\min\left\{\Delta(u_i-l_i-\varepsilon),\frac{|g_i|}{\|J_i\|^2}\right\}\geqslant\frac{1}{2}\tilde{\varepsilon}\min\{\Delta\varepsilon',\tilde{\varepsilon}L_2\}\end{aligned}$$

其中 L_2 是 $\frac{1}{\|J_i\|^2}$ 的下界。

定理 3　对给定 $\varepsilon>0$，用上述算法在有限步内可得到 ε 最优解。

证明　用反证法证。

设$\{x\}$是由上述算法产生的无穷点列。首先证明置信域半径序列$\{\Delta\}$有大于零的下界，若不然，则存在子列$\{\Delta'\}$收敛于 0，且$\{\Delta\}/\{\Delta'\}$有大于零的下界。

考虑序列$\{x'\}$。若 x' 是由(P_1)产生的，则由引理 1 知，当 $\Delta'\to0$ 时

$$-\psi_1(p)\geqslant(\tilde{\varepsilon}/2)\min\{\varepsilon\Delta',\tilde{\varepsilon}L_1\}$$

又由连续性及 x,p 的有界性，存在 $K>0$，使

$$f(x+p)-f(x)-\psi_1(p)\leqslant K\|p\|^2/2$$

又$\|Dp\|\leqslant\Delta'$，$\|D^{-1}\|$有界，设$\|D^{-1}\|\leqslant\delta$，则

$$f(x+p)-f(x)-\psi_1(p)\leqslant K\delta^2\Delta'^2/2$$

从而

$$|\rho-1|=\left|\frac{f(x+p)-f(x)-\psi_1(p)}{\psi_1(p)}\right|\leqslant\frac{K\delta^2\Delta'}{\varepsilon\tilde{\varepsilon}}\tag{7}$$

若 x' 是由(P_2)产生的，由引理 2，同理得

$$|\rho-1|=\left|\frac{f(x+\lambda s)-f(x)-\psi_2(\lambda)}{\psi_2(\lambda)}\right|\leqslant\frac{K\delta^2\Delta'}{\varepsilon'\tilde{\varepsilon}} \tag{8}$$

当 $\Delta'\to 0$ 时，式(7)和(8)意味着 $\rho\to 1$。由置信域半径的调整准则，当 $\rho>\mu$ 时，Δ' 不再减小，而 $\{\Delta\}/\{\Delta'\}$ 有大于零的下界，这与假设 $\{\Delta'\}$ 收敛于 0 矛盾，从而 $\{\Delta\}$ 有大于 0 的下界。由式(4)和(5)有

$$f(x+p)<f(x)-(\mu\tilde{\varepsilon}/2)\min\{\varepsilon\Delta,\tilde{\varepsilon}L_1\}$$

或

$$f(x+\lambda s)<f(x)-(\mu\tilde{\varepsilon}/2)\min\{\varepsilon'\Delta,\tilde{\varepsilon}L_2\}$$

即每步迭代目标函数的下降量大于等于一固定量。这样进行无穷步，$f(x)$ 下降到 $-\infty$，与连续函数 $f(x)$ 在闭区间 $l\leqslant x\leqslant u$ 上有界矛盾。所以有限步内即可得 ε 最优解。

参 考 文 献

1 Sagara N, Fukushima M. A hybrid method for the nonlinear least squares problem with simple bounds. J Comp and Appl Math, 1991, 36: 149～157

2 Conn A R, Gould N I M, Toint Ph L. Global convergence of a class of trust region algorithms for optimization with simple bounds. SIAM J Number Anal, 1988, 25: 433～460