非线性 l_1 问题的一种解法

马圣容[1],尤兴华[2]

(1. 南京晓庄学院数学与信息技术学院,,江苏 南京 211171)
(2. 南京工程学院基础部,江苏 南京 211167)

[**摘要**]　本文对非线性 l_1 问题 $\min\limits_{x\in\mathbf{R}^n} F(\boldsymbol{x})=\sum\limits_{i=1}^{m}|f_i(\boldsymbol{x})|$,从理论上研究了 $F(\boldsymbol{x})$ 的下降方向、最优解与某种盒式约束最小二乘问题的最优解之间的关系,进而构造了一个非线性 l_1 问题的下降算法,并证明了该算法的收敛性. 数值例子说明所给的非线性 l_1 问题的下降算法是有效的.

[**关键词**]　不可微,l_1 问题,内点算法,盒式约束最小二乘问题

[**中图分类号**]O221.2　[**文献标志码**]A　[**文章编号**]1001-4616(2013)02-0031-08

An Algorithm for Nonlinear l_1 Problem

Ma Shengrong[1],You Xinghua[2]

(1. School of Mathematics and Information Technology,Nanjing Xiaozhuang University,Nanjing 211171,China)
(2. Department of Basic Courses,Nanjing Institute of Technology,Nanjing 211167,China)

Abstract:This paper studied the nonlinear l_1 problem:$\min\limits_{x\in\mathbf{R}^n} F(\boldsymbol{x})=\sum\limits_{i=1}^{m}|f_i(\boldsymbol{x})|$. We first discuss the descent direction of the objective function $F(\boldsymbol{x})$ in theory,further more,we study the relation between the optimal solution of nonlinear l_1 problem and the optimal solution of some kind of quadratic programming problem with box constrains. Hence,we construct a descent algorithm for nonlinear l_1 problem and prove the convergence of the algorithm. An example shows that the new descent algorithm for nonlinear l_1 problem is effective.

Key words:nondifferentiable,l_1 problem,interior point algorithm,quadratic programming problem with box constrains

1　非线性 l_1 问题

非线性 l_1 问题一般表述为:

$$\min_{x\in\mathbf{R}^n} F(\boldsymbol{x})=\sum_{i=1}^{m}|f_i(\boldsymbol{x})|,$$

其中 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 是连续可微函数,且至少有一个是非线性的[1-5]. 而非线性 l_1 问题在经济数学、工程数学、目标规划等领域有着极其广泛的应用.

目前,解决非线性 l_1 问题有好几种算法[6,7],在本文中我们将构造一种新的方法. 首先,我们通过求解盒式约束最小二乘问题[8,9]

$$Q(\boldsymbol{x},\delta)\begin{cases}\min\left\|\sum\limits_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})v_i-\sum\limits_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})\right\|_2\\ s.t.\ -1\leqslant v_i\leqslant 1,\quad i\in E(\boldsymbol{x},\delta),\end{cases}$$

并利用其最优解 $\bar{v}_i,i\in E(\boldsymbol{x},\delta)$ 构造 $F(\boldsymbol{x})$ 的下降方向

$$\boldsymbol{d}=\sum_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})\bar{v}_i-\sum_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x}),$$

收稿日期:2012-11-28.

基金项目:国家青年科学基金(11101216).

通讯联系人:马圣容,讲师,研究方向:计算数学理论及应用. E-mail:srong2008@sina.com

然后利用该方向给出非线性 l_1 问题一个下降算法:算法1,最后,我们通过一个数值实例来验算算法的有效性.

2 非线性 l_1 问题的一个算法

2.1 非线性 l_1 问题的最优性条件

以下的结果属于文献[10].

引理1 假设函数 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$,在 $\boldsymbol{x}_0$ 的某邻域 $\boldsymbol{S}_\delta(\boldsymbol{x}_0)$ 内是连续可微的,$\boldsymbol{S}_\delta(\boldsymbol{x}_0)=\{\boldsymbol{x}\mid \|\boldsymbol{x}-\boldsymbol{x}_0\|<\delta\}$,$\delta>0$,则 $\forall \boldsymbol{d}\in\mathbf{R}^n$, $\|\boldsymbol{d}\|=1$,都 $\exists\varepsilon=\varepsilon(\boldsymbol{d})>0$,使得对 $\forall\alpha\in[0,\varepsilon)$,有 $F(\boldsymbol{x}_0+\alpha\boldsymbol{d})=F(\boldsymbol{x}_0)+\alpha\Delta F(\boldsymbol{x}_0,\boldsymbol{d})+o(\boldsymbol{d};\alpha)$,其中,

$$\Delta F(\boldsymbol{x}_0,\boldsymbol{d})=\sum_{i\notin A}\sigma_i\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}_0)+\sum_{i\in A}|\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}_0)|,$$

$$A=A(\boldsymbol{x}_0)=\{i\mid f_i(\boldsymbol{x}_0)=0,i\in 1,2,\cdots,m\},$$

$$\sigma_i=\mathrm{sign}(f_i(\boldsymbol{x}_0)),i\notin A,$$

$\nabla f_i(\boldsymbol{x})$ 是 $f_i(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 的梯度向量,

$$\lim_{\alpha\to0^+}\frac{o(\boldsymbol{d};\alpha)}{\alpha}=0.$$

引理2 设函数 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 在 $\boldsymbol{x}_0$ 的某邻域 $\boldsymbol{S}_\delta(\boldsymbol{x}_0)$ 内是二阶连续可微的,则 $\forall\boldsymbol{d}\in\mathbf{R}^n$, $\|\boldsymbol{d}\|=1$,都 $\exists\varepsilon=\varepsilon(\boldsymbol{d})>0$,使得对 $\forall\alpha\in[0,\varepsilon)$,有

$$F(\boldsymbol{x}_0+\alpha\boldsymbol{d})=F(\boldsymbol{x}_0)+\alpha\Delta F(\boldsymbol{x}_0;\boldsymbol{d})+\frac{1}{2}\alpha^2\Delta^2F(\boldsymbol{x}_0,\boldsymbol{d})+o(\boldsymbol{d};\alpha^2),$$

其中,

$$\lim_{\alpha\to0^+}\frac{o(\boldsymbol{d};\alpha^2)}{\alpha^2}=0,$$

$$\Delta^2F(\boldsymbol{x}_0,\boldsymbol{d})=\sum_{i\notin A}\sigma_i\boldsymbol{d}^{\mathrm T}\nabla^2f_i(\boldsymbol{x}_0)\boldsymbol{d}+\sum_{i\in A\setminus B}s_i\boldsymbol{d}^{\mathrm T}\nabla^2f_i(\boldsymbol{x}_0)\boldsymbol{d}+\sum_{i\in B}|\boldsymbol{d}^{\mathrm T}\nabla^2f_i(\boldsymbol{x}_0)\boldsymbol{d}|,$$

$$B=B(\boldsymbol{x}_0)=\{i\mid\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}_0)=0,i\in A\},s_i=\mathrm{sign}(\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}_0)),i\in A\setminus B,$$

$\nabla^2f_i(\boldsymbol{x}_0)$ 是 $f_i(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 的 Hessian 矩阵,$\Delta F(\boldsymbol{x},\boldsymbol{d})$ 与 A 如前定义.

引理3(最优性必要条件) 设函数 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 连续可微,则 $\boldsymbol{x}^*$ 是 $F(\boldsymbol{x})$ 的一个局部最小值点的必要条件是:$\Delta F(\boldsymbol{x}^*,\boldsymbol{d})\geqslant0$, $\forall\boldsymbol{d}\in\mathbf{R}^n$.

引理4 以下两命题等价:

(1) $\forall\boldsymbol{d}\in\mathbf{R}^n$,$\Delta F(\boldsymbol{x}^*,\boldsymbol{d})\geqslant0$,即 $\sum\limits_{i\notin A}\sigma_i\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)+\sum\limits_{i\in A}|\boldsymbol{d}^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)|\geqslant0$.

(2) $\exists v_i$,使得 $\sum\limits_{i\notin A}\sigma_i\nabla f_i(\boldsymbol{x}^*)+\sum\limits_{i\in A}v_i\nabla f_i(\boldsymbol{x}^*)=0$,其中 $-1\leqslant v_i\leqslant1$, $i\in A$.

引理5 设 $F(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 满足最优性的必要条件,则存在 $\boldsymbol{d}^0\in\mathbf{R}^n$, $\|\boldsymbol{d}^0\|=1$,使得

$$\sum_{i\notin A}\sigma_i(\boldsymbol{d}^0)^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)+\sum_{i\in A}|(\boldsymbol{d}^0)^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)|=0$$

的充分必要条件是:

$$\begin{cases}(\boldsymbol{d}^0)^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)=0, & |v_i|\neq1,i\in A,\\ (\boldsymbol{d}^0)^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)\geqslant0, & v_i=1,i\in A,\\ (\boldsymbol{d}^0)^{\mathrm T}\nabla f_i(\boldsymbol{x}^*)\leqslant0, & v_i=-1,i\in A.\end{cases}$$

引理6(二阶充分条件) 设函数 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 为二次连续可微函数,则 $\boldsymbol{x}^*$ 是 $F(\boldsymbol{x})$ 的一个严格局部极小点的一个充分条件是:

(1) $\exists v_i$,使得 $\sum\limits_{i\notin A}\sigma_i\nabla f_i(\boldsymbol{x}^*)+\sum\limits_{i\in A}v_i\nabla f_i(\boldsymbol{x}^*)=0$,其中 $-1\leqslant v_i\leqslant1$, $i\in A$.

(2)任意满足下列条件的非零向量 $\boldsymbol{d}$

$$\begin{cases}(\boldsymbol{d})^{\mathrm{T}}\nabla f_i(\boldsymbol{x}^*)=0, & |v_i|\neq 1, i\in A,\\ (\boldsymbol{d})^{\mathrm{T}}\nabla f_i(\boldsymbol{x}^*)\geqslant 0, & v_i=1, i\in A,\\ (\boldsymbol{d})^{\mathrm{T}}\nabla f_i(\boldsymbol{x}^*)\leqslant 0, & v_i=-1, i\in A\end{cases}$$

都有 $\boldsymbol{d}^{\mathrm{T}}[\sum\limits_{i\notin A}\sigma_i\nabla^2 f_i(\boldsymbol{x}^*)+\sum\limits_{i\in A}v_i\nabla^2 f_i(\boldsymbol{x}^*)]\boldsymbol{d}>0$.

利用以上结果,我们在 2.2 中将给出非线性 l_1 问题的一个下降算法.

2.2 非线性 l_1 问题的算法

记 $E(\boldsymbol{x},\delta)=\{i\,|\,|f_i(\boldsymbol{x})|\leqslant\delta, i=1,2,\cdots,m\}$,其中 $\delta\geqslant 0$,假设盒式约束最小二乘问题

$$Q(\boldsymbol{x},\delta)\begin{cases}\min\left\|\sum\limits_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})v_i-\sum\limits_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})\right\|_2\\ s.\,t.\ -1\leqslant v_i\leqslant 1, \quad i\in E(\boldsymbol{x},\delta)\end{cases}$$

的最优解为 $v(\boldsymbol{x},\delta)$,则有如下结论:

定理 1 设 $f_i(\boldsymbol{x}), i=1,2,\cdots,m$ 连续可微,$v(\boldsymbol{x},\delta)$ 为 $Q(\boldsymbol{x},\delta)$ 的最优解,$\boldsymbol{d}(\boldsymbol{x},\delta)=\sum\limits_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},\delta)-\sum\limits_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})$,则

(1) $\sum\limits_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum\limits_{i\in E(\boldsymbol{x},\delta)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)|=-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2$.

(2) 若 $\boldsymbol{d}(\boldsymbol{x},\delta)\neq 0$,则 $\boldsymbol{d}(\boldsymbol{x},\delta)$ 为 $F(\boldsymbol{x})=\sum\limits_{i=1}^{m}|f_i(\boldsymbol{x})|$ 在 $\boldsymbol{x}$ 处的一个下降方向.

(3) 若 $\delta_2\geqslant\delta_1\geqslant 0$,则 $\|\boldsymbol{d}(\boldsymbol{x},\delta_1)\|_2\geqslant\|\boldsymbol{d}(\boldsymbol{x},\delta_2)\|_2$,即

$$\sum_{i\notin E(\boldsymbol{x},\delta_1)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_1)+\sum_{i\in E(\boldsymbol{x},\delta_1)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_1)|\leqslant\sum_{i\notin E(\boldsymbol{x},\delta_2)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_2)+\sum_{i\in E(\boldsymbol{x},\delta_2)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_2)|.$$

证明 (1)由 $\boldsymbol{d}(\boldsymbol{x},\delta)$ 的定义知,

$$\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2=\sum_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)-\sum_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta),$$

所以,

$$\sum_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},\delta)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)|=-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2+\sum_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},\delta)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)|\geqslant-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2.$$

记 $\boldsymbol{S}(\boldsymbol{x},\delta)=\{\sum\limits_{i\in E(\boldsymbol{x},\delta)}v_i\nabla f_i(\boldsymbol{x})\,|\,-1\leqslant v_i\leqslant 1, i\in E(\boldsymbol{x},\delta)\}$,由 $\boldsymbol{S}(\boldsymbol{x},\delta)$ 为一个有界闭凸集及凸集分离定理知,$\forall\boldsymbol{y}\in\boldsymbol{S}(\boldsymbol{x},\delta)$,有 $(\boldsymbol{y}-\sum\limits_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)\geqslant 0$,

因为

$$\sum_{i\in E(\boldsymbol{x},\delta)}[-\mathrm{sign}(\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x}))]\nabla f_i(\boldsymbol{x})\in\boldsymbol{S}(\boldsymbol{x},\delta),$$

所以

$$\sum_{i\in E(\boldsymbol{x},\delta)}-|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|-\sum_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)\geqslant 0,$$

即

$$\sum_{i\in E(\boldsymbol{x},\delta)}|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|\leqslant-\sum_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta).$$

又

$$\sum_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},\delta)}|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|\leqslant\sum_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)-\sum_{i\in E(\boldsymbol{x},\delta)}v_i(\boldsymbol{x},\delta)\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)=-(\sum_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},\delta)-\sum_{i\notin E(\boldsymbol{x},\delta)}\mathrm{sign}(f_i(\boldsymbol{x}))\nabla f_i(\boldsymbol{x}))^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)=-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2.$$

综上所述,可得

$$\sum_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},\delta)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)|=-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2.$$

(2)若 $\boldsymbol{d}(\boldsymbol{x},\delta)\neq 0$,

因为

$$E(\boldsymbol{x},0)\subseteq E(\boldsymbol{x},\delta),$$

所以

$$\Delta F(\boldsymbol{x},\boldsymbol{d}(\boldsymbol{x},\delta))=\sum_{i\notin E(\boldsymbol{x},0)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},0)}|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|=$$

$$\sum_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+\sum_{i\in E(\boldsymbol{x},\delta)\backslash E(\boldsymbol{x},0)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+$$

$$\sum_{i\in E(\boldsymbol{x},0)}|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|\leqslant\sum_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta)+$$

$$\sum_{i\in E(\boldsymbol{x},\delta)}|\boldsymbol{d}(\boldsymbol{x},\delta)^{\mathrm{T}}\nabla f_i(\boldsymbol{x})|=-\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2^2<0,$$

因而,由引理 1 可知,$\boldsymbol{d}(\boldsymbol{x},\delta)$ 为 $F(\boldsymbol{x})=\sum_{i=1}^{m}|f_i(\boldsymbol{x})|$ 在 x 处的一个下降方向.

(3)若 $\delta_2\geqslant\delta_1\geqslant 0$,则 $E(\boldsymbol{x},\delta_1)\subseteq E(\boldsymbol{x},\delta_2)$,

$$\boldsymbol{d}(\boldsymbol{x},\delta_1)=-\sum_{i\notin E(\boldsymbol{x},\delta_1)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})+\sum_{i\in E(\boldsymbol{x},\delta_1)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},\delta_1)=-\sum_{i\notin E(\boldsymbol{x},\delta_2)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})+$$

$$\sum_{i\in E(\boldsymbol{x},\delta_2)\backslash E(\boldsymbol{x},\delta_1)}-\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})+\sum_{i\in E(\boldsymbol{x},\delta_1)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},\delta_1).$$

令

$$v_i=-\operatorname{sign}(f_i(\boldsymbol{x})),\ i\in E(\boldsymbol{x},\delta_2)\backslash E(\boldsymbol{x},\delta_1),\ v_i=v_i(\boldsymbol{x},\delta_1),\ i\in E(\boldsymbol{x},\delta_1),$$

则

$$-1\leqslant v_i\leqslant 1,\ i\in E(\boldsymbol{x},\delta_2),$$

于是

$$\boldsymbol{d}(\boldsymbol{x},\delta_1)=-\sum_{i\notin E(\boldsymbol{x},\delta_2)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})+\sum_{i\in E(\boldsymbol{x},\delta_2)}\nabla f_i(\boldsymbol{x})v_i,$$

再由

$$\|\boldsymbol{d}(\boldsymbol{x},\delta_2)\|_2=\min\{\|\sum_{i\in E(\boldsymbol{x},\delta_2)}\nabla f_i(\boldsymbol{x})v_i-\sum_{i\notin E(\boldsymbol{x},\delta_2)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})\|_2,-1\leqslant v_i\leqslant 1,i\in E(\boldsymbol{x},\delta_2)\},$$

可知

$$\|\boldsymbol{d}(\boldsymbol{x},\delta_1)\|_2\geqslant\|\boldsymbol{d}(\boldsymbol{x},\delta_2)\|_2.$$

由(1)可知

$$\sum_{i\notin E(\boldsymbol{x},\delta_1)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_1)+\sum_{i\in E(\boldsymbol{x},\delta_1)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_1)|\leqslant$$

$$\sum_{i\notin E(\boldsymbol{x},\delta_2)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_2)+\sum_{i\in E(\boldsymbol{x},\delta_2)}|\nabla f_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}(\boldsymbol{x},\delta_2)|.$$

定理 2　设 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$,连续可微,令

$$\bar{\delta}=\begin{cases}\min\{|f_i(\boldsymbol{x})|\,|\,|f_i(\boldsymbol{x})|>0,i=1,2,\cdots,m\}, & \text{存在}|f_i(\boldsymbol{x})|>0,\\ 1, & \text{不存在}|f_i(\boldsymbol{x})|>0,\end{cases}$$

若 $0\leqslant\delta\leqslant\bar{\delta}$,则 $Q(\boldsymbol{x},\delta)$ 的最优目标函数值

$$\|\boldsymbol{d}(\boldsymbol{x},\delta)\|_2=\|\sum_{i\in E(\boldsymbol{x},\delta)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},\delta)-\sum_{i\notin E(\boldsymbol{x},\delta)}\operatorname{sign}(f_i(\boldsymbol{x}))\,\nabla f_i(\boldsymbol{x})\|_2=0$$

的充分必要条件为:

$$\Delta F(\boldsymbol{x},\boldsymbol{d})\geqslant 0,\ \forall\,\boldsymbol{d}\in\mathbf{R}^n.$$

证明　显然 $E(\boldsymbol{x},\delta)=E(\boldsymbol{x},0)$,$Q(\boldsymbol{x},\delta)$ 即 $Q(\boldsymbol{x},0)$,故只需对 $Q(\boldsymbol{x},0)$ 证明.

显然　$\|\boldsymbol{d}(\boldsymbol{x},0)\|_2=0\Leftrightarrow\sum_{i\in E(\boldsymbol{x},0)}\nabla f_i(\boldsymbol{x})v_i(\boldsymbol{x},0)=\sum_{i\notin E(\boldsymbol{x},0)}\operatorname{sign}f_i(\boldsymbol{x})\,\nabla f_i(\boldsymbol{x})$

$$\Leftrightarrow\Delta F(\boldsymbol{x},\boldsymbol{d})\geqslant 0,\ \forall\,\boldsymbol{d}\in\mathbf{R}^n.$$

由定理1、2,可以构造如下算法.

算法1

(1)初始点 $\boldsymbol{x}^0\in\mathbf{R}^n$,取适当 $\delta^0>0$,置 $k=0$.

(2)若 $F(\boldsymbol{x}^k)=\sum_{i=1}^{m}|f_i(\boldsymbol{x}^k)|=0$,则 $\boldsymbol{x}^k$ 为最优解,停算;否则确立 $E(\boldsymbol{x}^k,\delta^k)=\{i\,|\,|f_i(\boldsymbol{x}^k)|\leqslant\delta^k,i=1,2,\cdots,m\}$,并求解 $Q(\boldsymbol{x}^k,\delta^k)\begin{cases}\min\ \left\|\sum_{i\in E(\boldsymbol{x}^k,\delta^k)}\nabla f_i(\boldsymbol{x}^k)v_i-\sum_{i\notin E(\boldsymbol{x}^k,\delta^k)}\operatorname{sign}(f_i(\boldsymbol{x}^k))\nabla f_i(\boldsymbol{x}^k)\right\|_2\\ s.t.\ -1\leqslant v_i\leqslant1,\quad i\in E(\boldsymbol{x}^k,\delta^k)\end{cases}$

得最优解 $v_i(\boldsymbol{x}^k,\delta^k),i\in E(\boldsymbol{x}^k,\delta^k)$.

(3)置 $\boldsymbol{d}^k=\boldsymbol{d}(\boldsymbol{x}^k,\delta^k)=\sum_{i\in E(\boldsymbol{x}^k,\delta^k)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta^k)-\sum_{i\notin E(\boldsymbol{x}^k,\delta^k)}\operatorname{sign}(f_i(\boldsymbol{x}^k))\nabla f_i(\boldsymbol{x}^k)$.

(4)若 $\|\boldsymbol{d}^k\|=0$,转(5),否则转(6).

(5)置 $\bar{\delta}^k=\min\{|f_i(\boldsymbol{x}^k)|\,|\,|f_i(\boldsymbol{x}^k)|>0,i=1,2,\cdots,m\}$,若 $\delta^k<\bar{\delta}^k$,则停算;否则置 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k,\delta^{k+1}=\frac{1}{2}\bar{\delta}^k,k=k+1$,转(2).

(6)求 $\lambda_k>0$,使得 $F(\boldsymbol{x}^k+\lambda_k\boldsymbol{d}^k)=\min\{F(\boldsymbol{x}^k+\lambda\boldsymbol{d}^k)\,|\,\lambda\geqslant0\}$,并且置 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+\lambda_k\boldsymbol{d}^k,\delta^{k+1}=\begin{cases}\delta^k, & \|\boldsymbol{d}^k\|_2^2\geqslant\delta^k,\\ \frac{1}{2}\delta^k, & \|\boldsymbol{d}^k\|_2^2<\delta^k,\end{cases}$ $k=k+1$,转(2).

2.3 算法1的收敛性

引理7 设 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 连续,$\{\boldsymbol{x}^k\}\subset\mathbf{R}^n$ 且 $\boldsymbol{x}^k\to\boldsymbol{x}^\infty$,令

$$\bar{\delta}=\begin{cases}\frac{1}{2}\min\{|f_i(\boldsymbol{x}^\infty)|\,|\,|f_i(\boldsymbol{x}^\infty)|>0,i=1,2,\cdots,m\}, & \text{存在}|f_i(\boldsymbol{x}^\infty)|>0,\\ 1, & \text{不存在}|f_i(\boldsymbol{x}^\infty)|>0,\end{cases}$$

则

(1) $\forall\delta\in[0,\bar{\delta}]$,有 $E(\boldsymbol{x}^\infty,0)=E(\boldsymbol{x}^\infty,\delta)$ 及对于充分大的 k 成立 $E(\boldsymbol{x}^k,\delta)\subset E(\boldsymbol{x}^\infty,0)$.

(2) $\forall\delta>0$,有 $E(\boldsymbol{x}^\infty,0)\subset E(\boldsymbol{x}^k,\delta)$ 对充分大的 k 成立.

证明 (1) $\forall\delta\in[0,\bar{\delta}]$,由 $\bar{\delta}$ 的定义知

$$E(\boldsymbol{x}^\infty,0)=E(\boldsymbol{x}^\infty,\delta).$$

对 $i\notin E(\boldsymbol{x}^\infty,0)=E(\boldsymbol{x}^\infty,\bar{\delta})$,知

$$|f_i(\boldsymbol{x}^\infty)|>\bar{\delta},$$

所以当 k 充分大时,有

$$|f_i(\boldsymbol{x}^k)|>\bar{\delta},$$

即 $i\notin E(\boldsymbol{x}^k,\bar{\delta})$,也即 $E(\boldsymbol{x}^k,\delta)\subset E(\boldsymbol{x}^\infty,0)$.

(2)设 $i\in E(\boldsymbol{x}^\infty,0)$,由 $|f_i(\boldsymbol{x}^\infty)|=0$ 知,k 充分大时,有

$$|f_i(\boldsymbol{x}^k)|\leqslant\delta,$$

即 $i\in E(\boldsymbol{x}^k,\delta)$,也即 $E(\boldsymbol{x}^\infty,0)\subset E(\boldsymbol{x}^k,\delta)$.

引理8 设 $f_i(\boldsymbol{x}),i=1,2,\cdots,m$ 连续可微,$\{\boldsymbol{x}^k\}\subset\mathbf{R}^n$ 且 $\boldsymbol{x}^k\to\boldsymbol{x}^\infty$,$\bar{\delta}$ 的取法与引理7相同,若 $Q(\boldsymbol{x}^\infty,0)$ 的最优目标函数值

$$\|\boldsymbol{d}(\boldsymbol{x}^\infty,0)\|_2=\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2>0,$$

则当 k 充分大以后,对 $\forall\delta\in[0,\bar{\delta}]$,均有

$$\|\boldsymbol{d}(\boldsymbol{x}^k,\delta)\|_2=\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)-\sum_{i\notin E(\boldsymbol{x}^k,\delta)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)\right\|_2\geqslant\frac{1}{2}\|\boldsymbol{d}(\boldsymbol{x}^\infty,0)\|_2=$$

$$\frac{1}{2}\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2.$$

证明　由引理7知,k充分大时,有$E(\boldsymbol{x}^k,\delta)=E(\boldsymbol{x}^\infty,0)$, $\forall\delta\in[0,\bar{\delta}]$.

$$\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2(\text{由}\ v(\boldsymbol{x}^\infty,0)\text{最优得})\leqslant$$
$$\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^k,\delta)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2=\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^k,\delta)-\right.$$
$$\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)+\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)+$$
$$\left.\sum_{i\in E(\boldsymbol{x}^k,\delta)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)-\sum_{i\notin E(\boldsymbol{x}^k,\delta)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)\right\|_2\leqslant\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)-\right.$$
$$\left.\sum_{i\notin E(\boldsymbol{x}^k,\delta)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)\right\|_2+\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}[\nabla f_i(\boldsymbol{x}^\infty)-\nabla f_i(\boldsymbol{x}^k)]v_i(\boldsymbol{x}^k,\delta)\right\|_2+$$
$$\left\|\sum_{i\notin E(\boldsymbol{x}^\infty,\delta)}[\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)-\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)]\right\|_2.$$

对于任意的$\varepsilon_1>0$,$\varepsilon_2>0$,当k充分大时,有

$$\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}[\nabla f_i(\boldsymbol{x}^\infty)-\nabla f_i(\boldsymbol{x}^k)]v_i(\boldsymbol{x}^k,\delta)\right\|_2\leqslant\varepsilon_1\ \left\|\sum_{i\notin E(\boldsymbol{x}^\infty,0)}[\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)-\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)]\right\|_2=$$
$$\left\|\sum_{i\notin E(\boldsymbol{x}^\infty,0)}[\operatorname{sign}f_i(\boldsymbol{x}^\infty)[\nabla f_i(\boldsymbol{x}^k)-\nabla f_i(\boldsymbol{x}^\infty)]]\right\|_2\leqslant\varepsilon_2.$$

故当k充分大时,

$$\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2\leqslant$$
$$\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)-\sum_{i\notin E(\boldsymbol{x}^k,\delta)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)\right\|_2+\varepsilon_1+\varepsilon_2.$$

只要取

$$\varepsilon_1+\varepsilon_2\leqslant\frac{1}{2}\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2,$$

即可得

$$\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta)-\sum_{i\notin E(\boldsymbol{x}^k,\delta)}\operatorname{sign}f_i(\boldsymbol{x}^k)\nabla f_i(\boldsymbol{x}^k)\right\|_2\geqslant$$
$$\frac{1}{2}\left\|\sum_{i\in E(\boldsymbol{x}^\infty,0)}\nabla f_i(\boldsymbol{x}^\infty)v_i(\boldsymbol{x}^\infty,0)-\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^\infty)\nabla f_i(\boldsymbol{x}^\infty)\right\|_2.$$

引理9　设$f_i(\boldsymbol{x})$,$i=1,2,\cdots,m$,连续可微,算法1产生的$\{\boldsymbol{x}^k\}$有界,则算法1中有$\lim\limits_{k\to\infty}\delta^k=0$.

证明　$\{\delta^k\}$单调递减有下界,设$\delta^\infty=\lim\limits_{k\to\infty}\delta^k$,若$\delta^\infty>0$,则$\delta^{k+1}=\frac{1}{2}\delta^k$或$\delta^{k+1}=\frac{1}{2}\bar{\delta}^k\leqslant\frac{1}{2}\delta^k$不会出现无数次,即

$$\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta^k)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta^k)-\sum_{i\notin E(\boldsymbol{x}^k,\delta^k)}\operatorname{sign}(f_i(\boldsymbol{x}^k))\nabla f_i(\boldsymbol{x}^k)\right\|_2^2<\delta^k$$

不会出现无数次,因而存在$\bar{k}>0$,当$k\geqslant\bar{k}$时有

$$\left\|\sum_{i\in E(\boldsymbol{x}^k,\delta^k)}\nabla f_i(\boldsymbol{x}^k)v_i(\boldsymbol{x}^k,\delta^k)-\sum_{i\notin E(\boldsymbol{x}^k,\delta^k)}\operatorname{sign}(f_i(\boldsymbol{x}^k))\nabla f_i(\boldsymbol{x}^k)\right\|_2^2\geqslant\delta^k\geqslant\delta^\infty.$$

由于$-1\leqslant v_i(\boldsymbol{x}^k,\delta^k)\leqslant1$, $\nabla f_i(\boldsymbol{x})$连续,$\boldsymbol{x}^k$有界,故$\{\boldsymbol{d}(\boldsymbol{x}^k,\delta^k)\}$有界,所以存在$\{k_i\}\subset\{k\}$,使得$\lim\limits_{k_i\to\infty}\boldsymbol{x}^{k_i}=\boldsymbol{x}^\infty$, $\lim\limits_{k_i\to\infty}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})=\boldsymbol{d}^\infty$,再由引理7的(2)以及$\delta^{k_i}\geqslant\delta^\infty>0$知$k_i$充分大以后成立$E(\boldsymbol{x}^\infty,0)\subset E(\boldsymbol{x}^{k_i},\delta^{k_i})\subset E(\boldsymbol{x}^{k_i},\delta^{k_i})$,

所以

$$\sum_{i\notin E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^{k_i})\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})+\sum_{i\in E(\boldsymbol{x}^\infty,0)}|\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})|=$$
$$\sum_{i\notin E(\boldsymbol{x}^{k_i},\delta^{k_i})}\operatorname{sign}f_i(\boldsymbol{x}^{k_i})\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})+\sum_{i\in E(\boldsymbol{x}^{k_i},\delta^{k_i})\backslash E(\boldsymbol{x}^\infty,0)}\operatorname{sign}f_i(\boldsymbol{x}^{k_i})\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})+$$
$$\sum_{i\in E(\boldsymbol{x}^\infty,0)}|\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})|\leqslant\sum_{i\notin E(\boldsymbol{x}^{k_i},\delta^{k_i})}\operatorname{sign}f_i(\boldsymbol{x}^{k_i})\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})+$$
$$\sum_{i\in E(\boldsymbol{x}^{k_i},\delta^{k_i})}|\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})|=-\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})\|_2^2\leqslant-\delta^{k_i}\leqslant-\delta^\infty,$$

$$\beta=\Delta F(\boldsymbol{x}^{\infty},\boldsymbol{d}^{\infty})=\lim_{k_i\to\infty}\Big[\sum_{i\notin E(\boldsymbol{x}^{\infty},0)}\operatorname{sign} f_i(\boldsymbol{x}^{k_i})\,\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})+\sum_{i\in E(\boldsymbol{x}^{\infty},0)}|\nabla f_i(\boldsymbol{x}^{k_i})^{\mathrm T}\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})|\leqslant-\delta^{\infty}.$$

由 $F(\boldsymbol{x}^{\infty}+\alpha\boldsymbol{d}^{\infty})=F(\boldsymbol{x}^{\infty})+\alpha\beta+o(\boldsymbol{d}^{\infty},\alpha)$, $\forall\alpha\in[0,\varepsilon)$,知存在 $\bar\alpha<\varepsilon$,当 $\alpha\in[0,\bar\alpha]$ 时,$F(\boldsymbol{x}^{\infty}+\alpha\boldsymbol{d}^{\infty})\leqslant F(\boldsymbol{x}^{\infty})+\frac{5}{6}\alpha\beta.$

对固定的 $\alpha\in(0,\bar\alpha]$,当 k_i 充分大时有

$$|F(\boldsymbol{x}^{k_i})-F(\boldsymbol{x}^{\infty})|\leqslant-\frac{1}{6}\alpha\beta,\quad |F(\boldsymbol{x}^{k_i}+\alpha\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i}))-F(\boldsymbol{x}^{\infty}+\alpha\boldsymbol{d}^{\infty})|\leqslant-\frac{1}{6}\alpha\beta,$$

因此

$$F(\boldsymbol{x}^{k_i}+\alpha\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i}))\leqslant F(\boldsymbol{x}^{\infty}+\alpha\boldsymbol{d}^{\infty})-\frac{1}{6}\alpha\beta\leqslant F(\boldsymbol{x}^{\infty})+\frac{4}{6}\alpha\beta\leqslant F(\boldsymbol{x}^{k_i})+\frac{3}{6}\alpha\beta<F(\boldsymbol{x}^{k_i})+\frac{1}{6}\alpha\beta\leqslant F(\boldsymbol{x}^{\infty}).$$

另外,由 $k_i\geqslant\bar k$ 时,$\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})\|_2^2\geqslant\delta^{\infty}>0$ 可知,$\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})$ 为 $F(\boldsymbol{x})$ 在 $\boldsymbol{x}^{k_i}$ 处下降方向,所以 $F(\boldsymbol{x}^{k_i+1})\leqslant F(\boldsymbol{x}^{k_i}+\alpha\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i}))<F(\boldsymbol{x}^{\infty})$,但是 $\{F(\boldsymbol{x}^k)\}$ 是单调下降的,所以应该有 $F(\boldsymbol{x}^{k_i+1})\geqslant F(\boldsymbol{x}^{\infty})$,矛盾. 所以必定有 $\delta^{\infty}=0$. 即

$$\lim_{k\to\infty}\delta^k=0.$$

定理 3 假设函数 $f_i(\boldsymbol{x})$, $i=1,2,\cdots,m$,是连续可微的,并且水平集 $\boldsymbol{L}(\boldsymbol{x}^0)=\{\boldsymbol{x}\,|\,F(\boldsymbol{x})\leqslant F(\boldsymbol{x}^0),\boldsymbol{x}\in\mathbf{R}^n\}$ 有界,则

(1) $\lim\limits_{k\to\infty}\inf\|\boldsymbol{d}(\boldsymbol{x}^k,\delta^k)\|_2=0$,

(2) 必存在 $\{\boldsymbol{x}^k\}$ 的极限点 $\bar{\boldsymbol{x}}$ 使 $\bar{\boldsymbol{x}}$ 处成立:$\Delta F(\bar{\boldsymbol{x}},\boldsymbol{d})\geqslant0$, $\forall\boldsymbol{d}\in\mathbf{R}^n$.

证明

(1) 由引理 9 可知 $\delta^k\to0$,故存在 $\{k_i\}\subset\{k\}$,使 $\delta^{k_i+1}=\frac{1}{2}\delta^{k_i}$ 或 $\delta^{k_i+1}=\frac{1}{2}\bar\delta^{k_i}\leqslant\frac{1}{2}\delta^{k_i}$,故 $\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})\|_2^2\leqslant\delta^{k_i}$,由 $\delta^{k_i}\to0$ 知 $\lim\limits_{k\to\infty}\inf\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})\|_2=0$.

(2) 对(1)中所给的 $\{\boldsymbol{x}^{k_i}\}$,由于 $\{\boldsymbol{x}^{k_i}\}$ 有界,故不妨假设 $\lim\limits_{k_i\to\infty}\boldsymbol{x}^{k_i}=\bar{\boldsymbol{x}}$,如果 $\Delta F(\bar x,\boldsymbol{d})\geqslant0$, $\forall\boldsymbol{d}\in\mathbf{R}^n$ 不成立,则由定理 2 知 $\|\boldsymbol{d}(\bar{\boldsymbol{x}},0)\|_2>0$,由引理 8 知存在 $\bar\delta>0$,当 k_i 充分大以后,对 $\forall\delta\in[0,\bar\delta]$,有

$$\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta)\|_2\geqslant\frac{1}{2}\|\boldsymbol{d}(\bar{\boldsymbol{x}},0)\|_2>0.$$

故当 k_i 充分大以后,由 $\forall\delta^{k_i}\in[0,\bar\delta]$,(1)及上式知

$$0<\frac{1}{2}\|\boldsymbol{d}(\bar{\boldsymbol{x}},0)\|_2^2\leqslant\|\boldsymbol{d}(\boldsymbol{x}^{k_i},\delta^{k_i})\|_2^2\leqslant\delta^{k_i},$$

此与 $\delta^{k_i}\to0$ 矛盾.

3 数值实例

$$\min_{\boldsymbol{x}\in\mathbf{R}^2}F(\boldsymbol{x})=\sum_{i=1}^{3}|f_i(\boldsymbol{x})|,$$

其中

$$f_1(\boldsymbol{x})=\boldsymbol{x}_1^2+\boldsymbol{x}_2^2-\boldsymbol{x}_1\boldsymbol{x}_2-1,$$
$$f_2(\boldsymbol{x})=\cos\boldsymbol{x}_1,$$
$$f_3(\boldsymbol{x})=-\sin\boldsymbol{x}_2.$$

解 选取初始点 $\boldsymbol{x}^{(0)}=(1,1)^{\mathrm T}$, $\delta^0=0.1$. 迭代结果见表 1.

迭代到第 11 步时,由算法 1 的(5)可知当 $\boldsymbol{x}^{11}=(1.000\,0\quad 0.000\,0)$ 时,问题有最小值为 $f(\boldsymbol{x}^{11})=0.540\,3$.

表 1 迭代结果

Table 1 The iteration result

k	δ^k	λ_k	$\boldsymbol{d}^{(k)\mathrm T}$	$\boldsymbol{x}^{(k)\mathrm T}$	$F(\boldsymbol{x}^{(k)})$
0	0.1	0.413 7	(0.690 9 −0.690 9)	(1 1)	1.381 8
1	0.1	0.178 0	(−0.897 7 −0.898 2)	(1.285 8 0.714 2)	1.181 2
2	0.1	0.594 2	(−0.008 6 −0.840 9)	(1.126 0 0.554 3)	1.005 6
3	0.1	0.072 3	(−1.286 7 0.013 1)	(1.120 9 0.054 6)	0.687 6
4	0.1		(0 0)	(1.027 8 0.055 5)	0.574 6
5	0.05	0.110 8	(−0.229 6 −0.500 8)	(1.027 8 0.055 5)	0.574 6
6	0.05		(0 0)	(1.002 4 0.000 0)	0.543 1
7	0.025		(0 0)	(1.002 4 0.000 0)	0.543 1
8	0.012 5		(0 0)	(1.002 4 0.000 0)	0.543 1
9	0.006 3		(0 0)	(1.002 4 0.000 0)	0.543 1
10	0.003 1	0.002 1	(−1.162 0 0.002 4)	(1.002 4 0.000 0)	0.543 1
11	0.003 1			(1.000 0 0.000 0)	0.540 3

［参考文献］

［1］ 袁亚湘,孙文瑜. 最优化理论与方法［M］. 北京:科学出版社,1997:373-451.
［2］ 邓乃扬. 无约束最优化计算方法［M］. 北京:科学出版社,1982:206-280.
［3］ 袁亚湘. 非线性规划数值方法［M］. 上海:上海科技出版社,1993:200-223.
［4］ 陈宝林. 最优化理论与算法［M］. 北京:清华大学出版社,2000:334-383.
［5］ 席少霖. 非线性最优化方法［M］. 北京:高等教育出版社,1992:187-246.
［6］ Avriel M. Nonlinear Programming［M］. New Jersey:Prentice-Hall,Inc,1976:150-186.
［7］ 颜世建,葛福生. 非线性 l_1 问题的一个算法［J］. 南京师大学报:自然科学版,1999,22(2):1-7.
［8］ 张艺. 框式约束凸二次规划问题的内点算法［J］. 高等学校计算数学学报,2002(2):163-168.
［9］ 马圣容. 框式约束凸二次规划问题的内点算法［J］. 南京晓庄学院学报:自然科学版,2011,27(3):19-22.
［10］ Chris Charalambous. On conditions for optimality of the nonlinear l_1 problem［J］. Math Programming,1997,17:123-135.

［责任编辑:丁　蓉］

(上接第30页)

［参考文献］

［1］ Asmussen S. Ruin Probabilities［M］. Singapore:World Scientific,2000.
［2］ Cai J,Dickson D C M. On the expected discounted penalty function at ruin of a surplus process with interest［J］. Insurance:Mathematics and Economics,2002,30(3):389-404.
［3］ Palsen J,Gjessing H K. Ruin theory with stochastic economic environment［J］. Advances in Applied Probability,1997,29(4):965-985.
［4］ Cai J,Yang H L. Ruin in the perturbed compound poisson risk process under interest force［J］. Advances in Applied Probability,2005,37(3):819-835.
［5］ Sundt B,Teugels J L. Ruin estimates under interest force［J］. Insurance:Mathematics and Economics,1995,16(1):7-22.
［6］ 方世祖,罗建华. 双复合 Poisson 风险模型［J］. 纯粹数学与应用数学,2006,22(2):271-278.
［7］ Paul E,Claudia K,Thomas M. Modelling Extremal Events for Insurance and Finance［M］. New York:Springer-Verlag,1997.
［8］ 魏广华,高启兵. 常利力下双复合泊松风险模型破产概率的上界［J］. 南京师大学报:自然科学版,2009,32(1):30-34.
［9］ 魏广华,高启兵,王晓谦. 常利力下带干扰的双复合 Poisson 风险过程的生存概率［J］. 应用概率统计,2012,28(1):31-42.

［责任编辑:丁　蓉］