系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
22(3) (2002, 7), 355–364

求解 Minimax 优化问题的 SQP 方法*

薛 毅

(北京工业大学应用数理学院, 北京 100022)

摘要 本文提出一类求解 minimax 优化问题的有效算法, 该算法属于序列二次规划方法. 它具有全局收敛性和超线性收敛速率. 数值例子表明, 该算法是非常有效的, 这与算法具有良好的理论结果是分不开的.

关键词 Minimax 优化问题, SQP 方法.

MR(2000) 主题分类号 90C30, 90C47

1 引 言

考虑 minimax 优化问题

$$\min_{x} F(x) = \max_{1\le i\le m} f_i(x), \quad x \in R^n, \tag{1.1}$$

其中 $f_i(x), i=1,2,\cdots,m$ 是定义在 R^n 上的实值函数. 它等价于非线性优化问题

$$\min \quad z, \tag{1.2}$$
$$\text{s.t.} \quad f_i(x) - z \le 0, \quad i=1,2,\cdots,m. \tag{1.3}$$

可用求解约束问题的算法求解问题 (1.2)–(1.3), 从而得到 minimax 问题的解. 但这样会丢掉一些 minimax 问题的固有性质. 从另一角度来讲, SQP 方法是求解约束问题最有效的方法之一, 因此, 将 SQP 方法用于 minimax 问题, 会得到更为有效的方法.

本文提出一类求解 minimax 问题的有效方法, 算法基本思想是: 每步求解一个二次规划子问题, 其子问题的非零解是目标函数 $F(x)$ 的下降方向, 从而达到问题的最优解. 本文证明了所提出的算法具有全局收敛性和超线性收敛速率.

为了便于讨论, 我们先回顾一下 minimax 问题的基本性质. 定义

$$I(x) = \{i \mid f_i(x) = F(x)\}$$

为有效函数指标集, 简称有效集. 由问题 (1.2)–(1.3) 的一阶必要条件得到, 若 x^* 是问题 (1.1)

* 国家自然科学基金 (19971008) 资助课题.
谨以本文纪念许国志院士.

的解，则存在 $\lambda^*=(\lambda_1^*,\lambda_2^*,\cdots,\lambda_m^*)^{\mathrm{T}}$, 使得

$$\sum_{i=1}^{m}\lambda_i^*\nabla f_i(x^*)=0, \tag{1.4}$$

$$\sum_{i=1}^{m}\lambda_i^*=1, \tag{1.5}$$

$$\lambda_i^*\geq 0,\qquad i=1,2,\cdots,m, \tag{1.6}$$

$$\lambda_i\left(f_i(x^*)-F(x^*)\right)=0,\qquad i=1,2,\cdots,m. \tag{1.7}$$

因此称满足条件 (1.4)-(1.7) 的点为问题 (1.1) 的稳定点.

设 $f_i(x)$ 一阶连续可微，若存在 $0\neq d\in R^n$, 满足

$$\nabla f_i(\overline{x})^{\mathrm{T}}d<0,\quad i\in I(\overline{x}),$$

则称 d 是目标函数 $F(x)$ 在点 $\overline{x}$ 处的下降方向.

2 算法与定理

算法 2.1

(1) 选取初始点 $x^{(1)}\in R^n$ 和初始正定对称矩阵 $B_1(=I)$. 取 $\sigma\in(0,\frac{1}{2})$, $\delta\in(0,1]$, $\varepsilon\geq 0$, 置 $k=1$.

(2) 求解二次规划问题

$$\min\quad \frac{1}{2}d^{\mathrm{T}}B_kd+\frac{\delta}{2}t^2+t, \tag{2.1}$$

$$\text{s.t.}\quad \nabla f_i(x^{(k)})^{\mathrm{T}}d-t\leq F(x^{(k)})-f_i(x^{(k)}),\quad i=1,2,\cdots,m, \tag{2.2}$$

其最优解为 $(\overline{d}^{(k)},t_k)$, 相应的 Lagrange 乘子为 $\overline{\lambda}^{(k)}$. 置

$$d^{(k)}=\frac{\overline{d}^{(k)}}{(1+\delta t_k)},\qquad \lambda^{(k)}=\frac{\overline{\lambda}^{(k)}}{(1+\delta t_k)}.$$

(3) 若 $\|d^{(k)}\|\leq\varepsilon$, 则停止计算；否则取 $\beta\in(0,1)$, 令 $j=0$.

(4) 如果

$$F(x^{(k)}+\beta^jd^{(k)})\leq F(x^{(k)})+\sigma\beta^jt_k,$$

则置 $\alpha_k=\beta^j$, $x^{(k+1)}=x^{(k)}+\alpha_kd^{(k)}$, 转 (5); 否则置 $j=j+1$, 转 (4).

(5) 修正 B_k, 令

$$s^{(k)}=x^{(k+1)}-x^{(k)},$$

$$y^{(k)}=\sum_{i=1}^{m}\lambda_i^{(k)}\left[\nabla f_i(x^{(k+1)})-\nabla f_i(x^{(k)})\right],$$

$$\overline{y}^{(k)}=\theta_ky^{(k)}+(1-\theta_k)B_ks^{(k)},$$

其中

$$\theta_k = \begin{cases} 1, & \text{如果}\ (y^{(k)})^{\mathrm{T}} s^{(k)} \geq 0.2 (s^{(k)})^{\mathrm{T}} B_k s^{(k)}, \\ \dfrac{0.8 (s^{(k)})^{\mathrm{T}} B_k s^{(k)}}{(s^{(k)})^{\mathrm{T}} B_k s^{(k)} - (y^{(k)})^{\mathrm{T}} s^{(k)}}, & \text{其它}. \end{cases}$$

置

$$B_{k+1} = B_k - \frac{B_k s^{(k)} (s^{(k)})^{\mathrm{T}} B_k}{(s^{(k)})^{\mathrm{T}} B_k s^{(k)}} + \frac{\overline{y}^{(k)} (\overline{y}^{(k)})^{\mathrm{T}}}{(\overline{y}^{(k)})^{\mathrm{T}} s^{(k)}}.$$

(6) 置 $k = k+1$, 转 (2).

引理 2.2　设 $(\overline{d}^{(k)}, t_k)$ 是由算法得到的. 若 $\overline{d}^{(k)} = 0$, 则 $t_k = 0$.

证　由算法可知, $(\overline{d}^{(k)}, t_k)$ 是二次规划 (2.1)–(2.2) 的解, 因此存在 $\overline{\lambda}^{(k)}$, 满足

$$B_k \overline{d}^{(k)} + \sum_{i=1}^{m} \overline{\lambda}_i^{(k)} \nabla f_i(x^{(k)}) = 0, \tag{2.3}$$

$$\delta t_k + 1 - \sum_{i=1}^{m} \overline{\lambda}_i^{(k)} = 0, \tag{2.4}$$

$$\nabla f_i(x^{(k)})^{\mathrm{T}} \overline{d}^{(k)} - t_k \leq F(x^{(k)}) - f_i(x^{(k)}), \qquad i = 1, 2, \cdots, m, \tag{2.5}$$

$$\overline{\lambda}_i^{(k)} \geq 0, \qquad i = 1, 2, \cdots, m, \tag{2.6}$$

$$\overline{\lambda}_i^{(k)} (\nabla f_i(x^{(k)})^{\mathrm{T}} \overline{d}^{(k)} - t_k - F(x^{(k)}) + f_i(x^{(k)})) = 0, \qquad i = 1, 2, \cdots, m. \tag{2.7}$$

若 $\overline{d}^{(k)} = 0$, 由 (2.5), 得到

$$F(x^{(k)}) - f_i(x^{(k)}) + t_k \geq 0, \qquad i = 1, 2, \cdots, m,$$

对于 $i \in I(x^{(k)})$, 有 $F(x^{(k)}) = f_i(x^{(k)})$, 因此, $t_k \geq 0$.

另一方面, $(d, t) = (0, 0)$ 是问题 (2.1)–(2.2) 的可行解, 因此有

$$\frac{1}{2} (\overline{d}^{(k)})^{\mathrm{T}} B_k \overline{d}^{(k)} + \frac{\delta}{2} t_k^2 + t_k \leq 0, \tag{2.8}$$

因此, $t_k \leq 0$. 所以得到 $t_k = 0$.

注意到, $d^{(k)} = 0$ 与 $\overline{d}^{(k)} = 0$ 等价, 由 (2.3)–(2.7) 式和 (1.4)–(1.7) 式, 得到如下定理.

定理 2.3　若算法产生的 $d^{(k)} = 0$, 则 $x^{(k)}$ 是问题 (1.1) 的稳定点.

这就是算法终止条件为

$$\left\| d^{(k)} \right\| \leq \varepsilon$$

的理由.

定理 2.4　若算法产生的 $d^{(k)} \neq 0$, 则 $t_k < 0$ 且 $d^{(k)}$ 是 $F(x)$ 在点 $x^{(k)}$ 的下降方向.

证　$d^{(k)} \neq 0$ 等价于 $\overline{d}^{(k)} \neq 0$. 由 (2.8) 式, 并注意到 B_k 是正定对称矩阵, 因此得到

$$t_k \leq -\frac{1}{2} (\overline{d}^{(k)})^{\mathrm{T}} B_k \overline{d}^{(k)} - \frac{\delta}{2} t_k^2 \leq -\frac{1}{2} (\overline{d}^{(k)})^{\mathrm{T}} B_k \overline{d}^{(k)} < 0. \tag{2.9}$$

再由 (2.5), 得到

$$\nabla f_i(x^{(k)})^{\mathrm{T}}\overline{d}^{(k)} \le t_k + F(x^{(k)}) - f_i(x^{(k)}) = t_k < 0, \quad i \in I(x^{(k)}).$$

注意, $(1+\delta t_k) > 0$, 因此, $\nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)} < 0$.

从定理 2.4 可知算法第 (4) 步的合理性. 从上面的推导可以看出, 算法是可行的. 在算法中, 令

$$d^{(k)} = \frac{\overline{d}^{(k)}}{(1+\delta t_k)}, \quad \lambda^{(k)} = \frac{\overline{\lambda}^{(k)}}{(1+\delta t_k)}.$$

目的是保证在每步迭代中均有

$$B_k d^{(k)} = -\sum_{i=1}^{m} \lambda_i^{(k)} \nabla f_i(x^{(k)}), \qquad \sum_{i=1}^{m} \lambda^{(k)} = 1. \tag{2.10}$$

3 收敛性分析

在本小节我们对算法进行收敛性分析. 首先讨论算法的全局收敛性, 在建立全局收敛性的定理之前, 先作几点假设:

假设 I

1) $\{x^{(k)}\}$, $\{d^{(k)}\}$ 和 $\{B_k\}$ 有界;

2) 向量 $\begin{bmatrix} \nabla f_i(x) \\ -1 \end{bmatrix}$, $i \in I(x)$ 线性无关;

3) 对每个矩阵 B_k, 满足 $d^{\mathrm{T}} B_k d \ge \eta \|d\|^2$, 这里 η 是大于 0 的常数.

引理 3.1 设 $f_i(x)$ $(i = 1, 2, \cdots, m)$ 连续可微, 并满足假设 I. 且 $d^{(k)}$ 和 α_k 是由算法 2.1 得到的搜索方向及步长. 如果 $d^{(k)} \ne 0$, 则存在 $\overline{\epsilon} > 0$, 使得 $\alpha_k \ge \beta\overline{\epsilon}$.

证 取 $\sigma \in (0, \frac{1}{2})$, 由 $f_i(x)$ 的连续可微性, 有

$$f_i(x^{(k)} + \alpha d^{(k)}) - f_i(x^{(k)}) - \sigma\alpha \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)} = (1-\sigma)\alpha \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)} + o(\|\alpha\|),$$

存在 $\epsilon_1 > 0$, 当 $\alpha < \epsilon_1$ 时, 使得

$$\frac{o(\|\alpha\|)}{\|\alpha\|} < (1-\sigma) \left| \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)} \right|.$$

即

$$f_i(x^{(k)} + \alpha d^{(k)}) < f_i(x^{(k)}) + \sigma\alpha \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)}. \tag{3.1}$$

所以,

$$\max_{i \in I(x^{(k)})} f_i(x^{(k)} + \alpha d^{(k)}) \le \max_{i \in I(x^{(k)})} f_i(x^{(k)}) + \max_{i \in I(x^{(k)})} \sigma \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)}. \tag{3.2}$$

由 f_i 的连续性, 存在 $\epsilon_2 > 0$ 和 $i_0 \in I(x^{(k)})$, 使当 $\alpha < \epsilon_2$ 时, 有 $F(x^{(k)} + \alpha d^{(k)}) = f_{i_0}(x^{(k)} + \alpha d^{(k)})$. 令 $\overline{\epsilon} = \min\{\epsilon_1, \epsilon_2\}$, 由 (3.1) 式和 (3.2) 式, 当 $\alpha < \overline{\epsilon}$ 时, 有

$$\begin{aligned} F(x^{(k)} + \alpha d^{(k)}) &\le F(x^{(k)}) + \sigma\alpha \max_{i \in I(x^{(k)})} \nabla f_i(x^{(k)})^{\mathrm{T}} d^{(k)} \\ &\le F(x^{(k)}) + \sigma\alpha t_k. \end{aligned} \tag{3.3}$$

设 α_k 是由算法 2.1 第 (4) 步得到的步长，若 $\alpha_k \geq \bar{\epsilon}$, 则有 $\alpha_k \geq \beta\bar{\epsilon}$; 否则由上述证明过程可知，当 $\alpha_k < \bar{\epsilon}$ 时，(3.3) 式成立，且 $\beta^{i-1} \geq \bar{\epsilon}$. 因此，$\alpha_k \geq \beta\bar{\epsilon}$.

引理 3.2　设 $f_i(x)$ $(i = 1, 2, \cdots, m)$ 连续可微，并满足假设 I. 且 $d^{(k)}$ 和 α_k 是由算法 2.1 得到的搜索方向及步长. 若取 $\varepsilon = 0$, 且算法 2.1 不有限步终止，则

$$\lim_{k\to\infty} d^{(k)} = 0.$$

证　由算法 2.1 、引理 3.1 、定理 2.4 和 (2.9) 式，有

$$F(x^{(k)}) - F(x^{(k+1)}) \geq -\sigma\alpha_k t_k \geq \frac{1}{2}\sigma\alpha_k (d^{(k)})^{\mathrm{T}} B d^{(k)} \geq \frac{1}{2}\sigma\beta\bar{\epsilon}\eta\|d^{(k)}\|^2,$$

注意序列 $\{F(x^{(k)})\}$ 有界，则级数 $\sum(F(x^{(k+1)}) - F(x^{(k)}))$ 收敛，因此，

$$\lim_{k\to\infty} d^{(k)} = 0.$$

由引理 3.2, 可直接得到全局收敛定理.

定理 3.3　设 $f_i(x)$ $(i = 1, 2, \cdots, m)$ 连续可微，并满足假设条件 I,$\{x^{(k)}\}$ 是由算法 2.1 产生的点列，则 $\{x^{(k)}\}$ 的任意聚点均是问题 (1.1) 的稳定点.

下面证明算法的超线性收敛性. 为证明超线性收敛性的有关性质，我们的第二条假设为：

假设 II

1) $f_i(x), i = 1, 2, \cdots, m$ 是二次连续可微;

2) 算法 2.1 产生步长 $\alpha_k = 1$, 且点列 $\{x^{(k)}\}$ 收敛的问题的最优解是 x^*;

3) 矩阵 $\sum\limits_{i=1}^{m} \lambda_i^* \nabla^2 f_i(x)$ 非奇异，这里 λ^* 是 x^* 处的 Lagrange 乘子向量，即满足 $\sum\limits_{i=1}^{m} \lambda_i^* = 1$, $\lambda_i^* \geq 0$, $i = 1, 2, \cdots, m$ 和 $\sum\limits_{i=1}^{m} \lambda_i^* \nabla f_i(x) = 0$.

4) $\nabla^2 f_i(x)$ 满足 Lipschitz 条件，即存在 $\gamma_i > 0$, 使得

$$\|\nabla^2 f_i(\bar{x}) - \nabla^2 f_i(x)\| \leq \gamma_i \|\bar{x} - x\|.$$

定理 3.4　若假设 II 的 1), 2), 3) 成立，则

$$\lim_{k\to\infty} \frac{\|x^{(k)} + d^{(k)} - x^*\|}{\|x^{(k)} - x^*\|} = 0 \tag{3.4}$$

的充分必要条件是

$$\lim_{k\to\infty} \frac{\left\|\sum\limits_{i=1}^{m} \lambda_i^* \nabla f_i(x^{(k)} + d^{(k)})\right\|}{\|d^{(k)}\|} = 0. \tag{3.5}$$

证　因为 $\sum\limits_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0$, $\sum\limits_{i=1}^{m} \lambda_i^* \nabla^2 f_i(x^*)$ 非奇异，和 $\nabla^2 f_i(x)$ 的连续性，有

$$\begin{aligned}\left\|\sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^{(k)} + d^{(k)})\right\| &= \left\|\sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 f_i(\xi_k)(x^{(k)} + d^{(k)} - x^*)\right\| \\ &= O(\|x^{(k)} + d^{(k)} - x^*\|). \end{aligned} \tag{3.6}$$

若 (3.5) 式成立，由 (3.6) 式可知，

$$\lim_{k\to\infty}\frac{\|x^{(k)}+d^{(k)}-x^*\|}{\|d^{(k)}\|}=0. \tag{3.7}$$

又因为

$$\frac{\|x^{(k)}+d^{(k)}-x^*\|}{\|d^{(k)}\|}\geq\frac{\|x^{(k)}+d^{(k)}-x^*\|}{\|x^{(k)}+d^{(k)}-x^*\|+\|x^{(k)}-x^*\|}=\frac{r_k}{1+r_k},$$

这里 $r_k=\frac{\|x^{(k)}+d^{(k)}-x^*\|}{\|x^{(k)}-x^*\|}$.

由 (3.7) 式，得到 $\lim\limits_{k\to\infty}r_k=0$, 即 (3.4) 式成立.

反之，若 (3.4) 式成立，则有

$$\lim_{k\to\infty}\frac{\|d^{(k)}\|}{\|x^{(k)}-x^*\|}=1. \tag{3.8}$$

由 (3.6) 式和 (3.8) 式，可知 (3.5) 式成立.

定理 3.5 若假设 II 成立，则

$$\lim_{k\to\infty}\frac{\|x^{(k)}+d^{(k)}-x^*\|}{\|x^{(k)}-x^*\|}=0$$

的充分必要条件是

$$\lim_{k\to\infty}\frac{\left\|\left[B_k-\sum\limits_{i=1}^m\lambda_i^*\nabla^2f_i(x^*)\right]d^{(k)}+\sum\limits_{i=1}^m\left(\lambda_i^{(k)}-\lambda_i^*\right)\nabla f_i(x^{(k)})\right\|}{\|d^{(k)}\|}=0. \tag{3.9}$$

证 由 (2.10) 式，可知

$$\begin{aligned}&\left[B_k-\sum_{i=1}^m\lambda_i^*\nabla^2f_i(x^*)\right]d^{(k)}+\sum_{i=1}^m(\lambda_i^{(k)}-\lambda_i^*)\nabla f_i(x^{(k)})\\
=&B_kd^{(k)}-\sum_{i=1}^m\lambda_i^*\nabla^2f_i(x^*)d^{(k)}+\sum_{i=1}^m\lambda_i^{(k)}\nabla f_i(x^{(k)})-\sum_{i=1}^m\lambda_i^*\nabla f_i(x^{(k)})\\
=&-\sum_{i=1}^m\lambda_i^*\nabla f_i(x^{(k)})-\sum_{i=1}^m\lambda_i^*\nabla^2f_i(x^*)d^{(k)}\\
=&\sum_{i=1}^m\lambda_i^*[\nabla f_i(x^{(k)}+d^{(k)})-\nabla f_i(x^{(k)})-\nabla^2f_i(x^*)d^{(k)}]-\sum_{i=1}^m\lambda_i^*\nabla f_i(x^{(k)}+d^{(k)}).\end{aligned}$$

由于 $\nabla^2f_i(x)$ 满足 Lipschitz 条件，因此有

$$\begin{aligned}&\frac{\left\|\sum\limits_{i=1}^m\lambda_i^*\left[\nabla f_i(x^{(k)}+d^{(k)})-\nabla f_i(x^{(k)})-\nabla^2f_i(x^*)d^{(k)}\right]\right\|}{\|d^{(k)}\|}\\
\leq&\frac{1}{2}\sum_{i=1}^m\lambda_i^*\gamma_i\left(\|x^{(k)}-x^*\|+\|x^{(k+1)}-x^*\|\right)\\
=&\gamma\left(\|x^{(k)}-x^*\|+\|x^{(k+1)}-x^*\|\right),\end{aligned}$$

其中 $\gamma=\frac{1}{2}\sum_{i=1}^{m}\lambda_i^*\gamma_i$. 当 $k\to\infty$ 时，$x^{(k)}\to x^*$. 因此，(3.9) 式成立的充分必要条件是

$$\lim_{k\to\infty}\frac{\left\|\sum_{i=1}^{m}\lambda_i^*\nabla f_i(x^{(k)}+d^{(k)})\right\|}{\|d^{(k)}\|}=0.$$

由定理 3.4, 命题得证.

若当 $\lambda^{(k)}=\lambda^*$ 时，则超线性收敛步的充分必要条件可改为

$$\lim_{k\to\infty}\frac{\left\|[B_k-\sum_{i=1}^{m}\lambda_i^*\nabla^2 f_i(x^*)]d^{(k)}\right\|}{\|d^{(k)}\|}=0.$$

当 $m=1$ 时，即只有一个目标函数，由算法可知，$\lambda^{(k)}=1=\lambda^*$. 因此定理 3.5 的结果与一般拟 Newton 法的超线性收敛步的结果相同，这说明，本文提出的方法是原一般拟 Newton 法的推广.

4　数值结果

我们从文献 [1] 和 [2] 中选择了 10 个问题，下面列出这 10 个问题. 问题的基本形式为

$$\min_x F(x)=\max_{1\le i\le m} f_i(x),\qquad x\in R^n.$$

问题 4.1　$n=2, m=3$ 和

$$f_1(x)=x_1^2+x_2^4,\quad f_2(x)=(2-x_1)^2+(2-x_2)^2,\quad f_3(x)=2\exp(-x_1+x_2),$$

取初始点为 $x^{(1)}=(1,-0.1)^{\mathrm{T}}$.

问题 4.2　$n=2, m=3$ 和

$$f_1(x)=x_1^4+x_2^2,\quad f_2(x)=(2-x_1)^2+(2-x_2)^2,\quad f_3(x)=2\exp(-x_1+x_2),$$

取初始点为 $x^{(1)}=(1,-0.1)^{\mathrm{T}}$.

问题 4.3 (Rosen-Suzuki 问题)　$n=4, m=4$ 和

$$\begin{aligned}
f_1(x)=&x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4,\\
f_2(x)=&f_1(x)+10(x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4-8),\\
f_3(x)=&f_1(x)+10(x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4-10),\\
f_4(x)=&f_1(x)+10(x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4-5),
\end{aligned}$$

取初始点为 $x^{(1)}=(0,0,0,0)^{\mathrm{T}}$.

问题 4.4　$n=2, m=3$ 和

$$f_1(x)=x_1^2+x_2^2+x_1x_2,\quad f_2(x)=\sin(x_1),\quad f_3(x)=\cos(x_2),$$

取初始点为 $x^{(1)}=(3,1)^{\mathrm{T}}$.

问题 4.5 $n=3, m=6$ 和

$$\begin{aligned} &f_1(x)=x_1^2+x_2^2+x_3^2-1, && f_2(x)=x_1^2+x_2^2+(x_3-2)^2,\\ &f_3(x)=x_1+x_2+x_3-1, && f_4(x)=x_1+x_2-x_3+1,\\ &f_5(x)=2x_1^3+6x_2^2+2(5x_3-x_1+1)^2, && f_6(x)=x_1^2-9x_3, \end{aligned}$$

取初始点为 $x^{(1)}=(1,1,1)^{\mathrm{T}}$.

问题 4.6 $n=3, m=30$ 和

$$\begin{aligned} &f_i(x)=y_i-x_1-\tfrac{u_i}{x_2v_i+x_3w_i}, && i=1,2,\cdots,15,\\ &f_i(x)=-f_{i-15}(x), && i=16,17,\cdots,30, \end{aligned}$$

其中

$$\begin{aligned} &u_i=i,\quad v_i=16-i,\quad w_i=\min\{u_i,v_i\},\\ &y=(.14,.18,.22,.25,.29,.32,.35,.39,.37,.58,.73,.96,1.34,2.1,4.39), \end{aligned}$$

取初始点为 $x^{(1)}=(1,1,1)^{\mathrm{T}}$.

问题 4.7 $n=3, m=21$ 和

$$\begin{aligned} &f_i(x)=\frac{x_1+x_2t_i}{1+x_3t_i+x_4t_i^2+x_5+t_i^3}-\exp(t_i), && i=1,2,\cdots,11,\\ &f_i(x)=-f_{22-i}(x), && i=12,13,\cdots,21, \end{aligned}$$

其中 $t_i=\frac{i-1}{10}-1$. 取初始点为 $x^{(1)}=(1,1,1)^{\mathrm{T}}$.

问题 4.8(Wong 1 问题) $n=7, m=5$ 和

$$\begin{aligned} f_1(x)=&(x_1-10)^2+5(x_2-12)^2+x_3^4+3(x_4-11)^2+10x_5^6+7x_6^2+x_7^4\\ &-4x_6x_7-10x_6-8x_7,\\ f_2(x)=&f_1(x)+10(2x_1^2+3x_2^4+x_3+4x_4^2+5x_5-127),\\ f_3(x)=&f_1(x)+10(7x_1+3x_2+10x_3^2+x_4-x_5-282),\\ f_4(x)=&f_1(x)+10(23x_1+x_2^2+6x_6^2-8x_7-196),\\ f_5(x)=&f_1(x)+10(4x_1^2+x_2^2-3x_1x_2+2x_3^2+5x_6-11x_7), \end{aligned}$$

取初始点为 $x^{(1)}=(1,2,0,4,0,1,1)^{\mathrm{T}}$.

问题 4.9(Wong 2 问题) $n=10, m=9$ 和

$$\begin{aligned} f_1(x)=&x_1^2+x_2^2+x_1x_2-14x_1-16x_2+(x_3-10)^2+4(x_4-5)^2+(x_5-3)^2\\ &+2(x_6-1)^2+5x_7^2+7(x_8-11)^2+2(x_9-10)^2+(x_{10}-7)^2+45,\\ f_2(x)=&f_1(x)+10(3(x_1-2)^2+4(x_2-3)^2+2x_3^2-7x_4-120),\\ f_3(x)=&f_1(x)+10(5x_1^2+8x_2+(x_3-6)^2-2x_4-40),\\ f_4(x)=&f_1(x)+10(0.5(x_1-8)^2+2(x_2-4)^2+3x_5^2-x_6-30),\\ f_5(x)=&f_1(x)+10(x_1^2+2(x_2-2)^2-2x_1x_2+14x_5-6x_6),\\ f_6(x)=&f_1(x)+10(4x_1+5x_2-3x_7+9x_8-105),\\ f_7(x)=&f_1(x)+10(10x_1-8x_2-17x_7+2x_8),\\ f_8(x)=&f_1(x)+10(-3x_1+6x_2+12(x_9-8)^2-7x_{10}),\\ f_9(x)=&f_1(x)+10(-8x_1+2x_2+5x_9-2x_{10}-12), \end{aligned}$$

取初始点为 $x^{(1)} = (2,3,5,5,1,2,7,3,6,10)^{\mathrm{T}}$.

问题 4.10(Wong 3 问题)　$n=20$, $m=18$ 和

$$\begin{aligned}
f_1(x) =& x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3-10)^2 + 4(x_4-5)^2 \\
&+ (x_5-3)^2 + 2(x_6-1)^2 + 5x_7^2 + 7(x_8-11)^2 + 2(x_9-10)^2 \\
&+ (x_{10}-7)^2 + (x_{11}-9)^2 + 10(x_{12}-1)^2 + 5(x_{13}-7)^2 + 4(x_{14}-14)^2 \\
&+ 27(x_{15}-1)^2 + x_{16}^4 + (x_{17}-2)^2 + 13(x_{18}-2)^2 + (x_{19}-3)^2 + x_{20}^2 + 95, \\
f_2(x) =& f_1(x) + 10(3(x_1-2)^2 + 4(x_2-3)^2 + 2x_3^2 - 7x_4 - 120), \\
f_3(x) =& f_1(x) + 10(5x_1^2 + 8x_2 + (x_3-6)^2 - 2x_4 - 40), \\
f_4(x) =& f_1(x) + 10(0.5(x_1-8)^2 + 2(x_2-4)^2 + 3x_5^2 - x_6 - 30), \\
f_5(x) =& f_1(x) + 10(x_1^2 + 2(x_2-2)^2 - 2x_1x_2 + 14x_5 - 6x_6), \\
f_6(x) =& f_1(x) + 10(4x_1 + 5x_2 - 3x_7 + 9x_8 - 105), \\
f_7(x) =& f_1(x) + 10(10x_1 - 8x_2 - 17x_7 + 2x_8), \\
f_8(x) =& f_1(x) + 10(-3x_1 + 6x_2 + 12(x_9-8)^2 - 7x_{10}), \\
f_9(x) =& f_1(x) + 10(-8x_1 + 2x_2 + 5x_9 - 2x_{10} - 12), \\
f_{10}(x) =& f_1(x) + 10(x_1 + x_2 + 4x_{11} - 21x_{12}), \\
f_{11}(x) =& f_1(x) + 10(x_1^2 + 15x_{11} - 8x_{12} - 28), \\
f_{12}(x) =& f_1(x) + 10(4x_1 + 9x_2 + 5x_{13}^2 - 9x_{14} - 87), \\
f_{13}(x) =& f_1(x) + 10(3x_1 + 4x_2 + 3(x_{13}-6)^2 - 14x_{14} - 10), \\
f_{14}(x) =& f_1(x) + 10(14x_1^2 + 35x_{15} - 79x_{16} - 92), \\
f_{15}(x) =& f_1(x) + 10(15x_2^2 + 11x_{15} - 61x_{16} - 54), \\
f_{16}(x) =& f_1(x) + 10(5x_1^2 + 2x_2 + 9x_{17}^4 - x_{18} - 68), \\
f_{17}(x) =& f_1(x) + 10(x_1^2 - x_9 + 19x_{19} - 20x_{20} + 19), \\
f_{18}(x) =& f_1(x) + 10(7x_1^2 + 5x_2^2 + x_{12}^2 - 30x_{20}),
\end{aligned}$$

取初始点为 $x^{(1)} = (2,3,5,5,1,2,7,3,6,10,2,2,6,15,1,2,1,2,1,3)^{\mathrm{T}}$.

我们是用 Matlab 语言进行计算，在算法中，取 $\beta=\frac{1}{2}$, $\sigma=0.1$, $\delta_0=0.1$ 和 $B_1=I$. 终止准则为

$$\|d^{(k)}\| < 10^{-5}$$

或

$$\|x^{(k+1)} - x^{(k)}\| < 10^{-8}$$

其计算结果见下表.

问题	NI	NF	NG	$F(x)$	$\|d\|$	$I(x)$
1	6	11	6	1.95222	2.9694e–6	1,2
2	5	9	5	2	4.7033e–9	1,2,3
3	10	23	10	–44	3.7726e–5	1,2,4
4	8	15	8	0.616433	1.1560e–6	1,3
5	10	19	10	3.59972	4.3608e–7	2,5
6	8	15	8	0.0508169	2.5070e–7	8,15,24
7	14	27	14	2.75464e–6	1.2423e–6	2,8,11,12,17,21
8	14	40	14	680.63	1.0945e–5	1,2,5
9	17	34	17	24.3062	2.7316e–5	1,2,3,5,6,7,9
10	24	53	24	132.614	7.9283e–6	1,2,3,5,6,7,9,11,12,15,16,17,18

表中符号的意义如下：

NI	— 迭代次数.	$F(x)$	— 最优目标函数值.
NF	— 目标函数的总计算次数.	$\|d\|$	— 二次规划子问题的解 d 的模.
NG	— 梯度函数的总计算次数.	$I(x)$	— 在 x 处的有效函数指标集.

参 考 文 献

[1] Vardi A. New minimax algorithm. *Journal of Optimization Theory and Application*, 1992, **75**(3): 613–634.

[2] Luksan L. A compact variable metric algorithm for nonlinear minimax approximation. *Computing*, **36**: 355–373.

THE SEQUENTIAL QUADRATIC PROGRAMMING METHOD FOR SOLVING MINIMAX PROBLEM

Xue Yi

(*College of Applied Science, Beijing Polytechnic University, Beijing* 100022)

Abstract In this paper, an efficient algorithm for solving minimax optimization problem is proposed. It belongs to the sequential quadratic programming method. The algorithm is globally convergent and superlinearly convergent. Some numerical experiments suggest that the practical efficiency of the methods is related to these theoretical results.

Key words Minimax optimization problem, SQP method.