第15卷 增刊 1995年11月 数学研究与评论 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Vol. 15 Supp Nov. 1995

一个改进的SQP型算法*

简 金 宝

(广西大学数学系，南宁 530004)

摘 要 本文建立非线性等式和不等式约束规划问题的一个序列二次规划(SQP)型算法．算法的每次迭代只需解一个确实可解的二次规划，然后对其解进行简单的显式校正，便可产生关于罚函数是下降的搜索方向，克服 Maratos 效应．在适当的假设条件下，还论证了算法的全局收敛性和超线性收敛性．

关键词 一般非线性规划，序列二次规划，精确罚函数，收敛速度．

分类号 AMS(1991) 90C30/CCL O221.2

§1 引 言

本文考虑一般约束的非线性规划问题

$$
(P)\quad \begin{cases} \text{Min} & f(x), \\ \text{s.t.} & g_i(x) \leqslant 0, i \in L_1, \quad (1) \\ & g_j(x) = 0,\ j \in L_2, \quad (2) \end{cases}
$$

其中 $f, g_j (j \in L = L_1 \cup L_2)$ 均为 R^n 中二阶连续可微函数．

由于序列二次规划方法能产生超线性收敛速度，因此它被高度重视和广泛应用．但在解问题(P)的一些 SQP 型方法中，为克服 Maratos 效应获得步长 1，往往在每次迭代中要解多个二次规则[1,2]．因而方法计算量大，降低实用性．本文吸取[1,3,4]的思想，建立一个改进的求解(P)的 SQP 型算法．该算法不但具有全局收敛性和超线性收敛性，而且每次迭代只需解一个二次规划和进行一次显式校正，该二次规划总是可解的．本文还给出罚函数中罚参数的调整方法，且该调整只须进行有限次．

§2 改进的SQP型算法

对于参数 r，正定阵 $B \in E^{n \times n}$，方向 $d \in E^n$．记

$$\varphi(x) = \sum_{j \in L_1} g_j(x)_+ + \sum_{j \in L_2} |g_j(x)|, \tag{3}$$

$$\psi_r(x, d, s, B) = \nabla f(x)^T d + \frac{1}{2} d^T B d - r s \varphi(x). \tag{4}$$

精确罚函数为

* 1993年1月22日收到．广西壮族自治区教委资助．

$$F_r(x) = f(x) + r\varphi(x), \tag{5}$$

其中 $g_j(x)_+ = \max\{0, g_j(x)\}$, $0 \leqslant s \leqslant 1$. Lagrangian 函数为

$$L(x, u) = f(x) + \sum_{j\in L} u_j g_j(x), \nabla^2 L(x,u) = \nabla^2 f(x) + \sum_{j\in L} u_j \nabla^2 g_j(x).$$

对于给定点 x，为寻找搜索方向，人们常常考虑如下最简单的二次规划

$$\mathrm{QP}(x,B) \quad \begin{cases} \mathrm{Min} \quad d^T\nabla f(x) + \dfrac{1}{2}d^TBd, \\ \mathrm{s.t.} \quad g_j(x) + \nabla g_j(x)^T d \leqslant 0, j \in L_1, & (6) \\ \qquad\quad g_i(x) + \nabla g_i(x)^T d = 0, \ i \in L_2. & (7) \end{cases}$$

然而 x 的不可行性常常导致(6)，(7)不相容，使得 QP(x,B)不可求解，算法难以进行. 为此 Powell[3]，Tone[4]和 Biggs[5]分别将 QP(x,B)修正为一个确实可解的二次规划. 我们用如下二次规划代替 QP(x,B).

$$\mathrm{MQP}(x,B) \quad \begin{cases} \mathrm{Min} \quad \nabla f(x)^T d + \dfrac{1}{2}d^TBd - Ms, \\ \mathrm{s.t.} \quad sg_j(x) + \nabla g_j(x)^T d \leqslant 0, j \in L_1, & (8) \\ \qquad\quad sg_i(x) + \nabla g_i(x)^T d = 0, \ i \in L_2, & (9) \\ \qquad\quad 0 \leqslant s \leqslant 1, & (10) \end{cases}$$

其中 M 为正的常数，d，s 为变量.

易见(8)－(10)是相容的，事实上 $d=0$，$s=0$ 为其可行解. 从而由 B 正定，$0 \leqslant s \leqslant 1$ 易知 MQP(x,B)有最优解.

为说明 QP(x,B)与 MQP(x,B)之间的关系，我们有

引理 1* 设(6)和(7)相容，M 充分大，则 QP(x,B)与 MQP(x,B)是等价的. 即如 d 为 QP(x,B)的最优解，则 d 及 $s=1$ 为 MQP(x,B)的最优解；反之如(d,s)为 MQP(x,B)的最优解，则 $s=1$，且 d 为 QP(x,B)的最优解.

证明 1° 设 d^0 为(6)和(7)的一个解，d 为 QP(x,B)的最优解. 对于 MQP(x,B)的任意可行解$(\tilde{d}, \tilde{s})$.

当 $\tilde{s}=1$ 时，$\tilde{d}$ 为 QP(x,B)的可行解，从而

$$\nabla f(x)^T d + \frac{1}{2}d^TBd - Mx \leqslant \nabla f(x)^T\tilde{d} + \frac{1}{2}\tilde{d}^TB\tilde{d} - \tilde{s}M.$$

当 $\tilde{s}<1$ 时，由 M 充分大，故而

$$M(1-\tilde{s}) \geqslant \nabla f(x)^T d + \frac{1}{2}d^TBd - \nabla f(x)^T\tilde{d} - \frac{1}{2}^T B\tilde{d}.$$

由上述可知 d 及 $s=1$ 为 MQP(x,B)的最优解.

2° 设(d,s)为 MQP(x,B)的最优解. 如 $s<1$，则由于$(d^0, 1)$为 MQP(x,B)之可行解，有

$$\nabla f(x)^T d + \frac{1}{2}d^TBd - Ms \leqslant \nabla f(x)^T d^0 + \frac{1}{2}(d^0)^TBd^0 - M,$$

$$M(1-s) \leqslant \nabla f(x)^T(d^0 - d) + \frac{1}{2}(d^0)^TBd^0 - \frac{1}{2}d^TBd.$$

* 该引理证明是不够严格的，但并不影响下文！

这与 M 充分大，$0<1-s$ 矛盾. 故 $s=1$. 于是依定义可知 d 不但为 QP(x,B)的可行解，且是最优解.

在全文的余下部分，不需“M 充分大”.

引理 2 设(d,s)为 MQP(x,B)的 $K-T$ 点，$u_j(j\in L)$为其关于(d,s)的 $K-T$ 乘子，B 正定. 则 i) $d=0$ 时，s 只能是 1 或者 0. 当$(d,s)=(0,1)$时，x 为原问题(P)的 $K-T$ 点； ii) 设罚参数 $r\geqslant\max\{|u_j|,j\in L\}$，若 $d\neq 0$，则 $\psi_r(x,d,s,B)<0$，且 $\psi_r(x,d,s,B)\geqslant D_dF_r(x)+\frac{1}{2}d^TBd>D_dF_r(x)$. 从而 d 为 $F_r(x)$在 x 处的下降方向. 其中 $D_dF_r(x)$为 F_r 在 x 处沿着方向 d 的方向导数.

证明 由定义易知 MQP(x,B)的可行点(d,s)为其 $K-T$ 点当且仅当存在乘子 $u_j(j\in L)$及 μ_1,μ_2，使得

$$\begin{cases} \nabla f(x)+Bd+\sum\limits_{j\in L}u_j\nabla g_j(x)=0, & (11)\\ -M+\sum\limits_{j\in L}u_jg_j(x)-\mu_1+\mu_2=0, & (12)\\ u_j(sg_j(x)+\nabla g_j(x)^Td)=0,\ u_j\geqslant 0,j\in L_1, & (13)\\ \mu_1 s=0,\mu_2(s-1)=0,\ \mu_1\geqslant 0,\mu_2\geqslant 0. & (14)\end{cases}$$

i) 设 $d=0$，反设 $0<s<1$. 由(8)，(9)和(13)，(14)知：$g_j(x)=0,\forall\ j\in L_2;u_jg_j(x)=0,\forall\ j\in L_1;\mu_1=\mu_2=0$. 从而由(12)有 $M=0$. 这与 $M>0$ 矛盾，故 $s=1$ 或 0. 当$(d,s)=(0,1)$时，x 为(P)的可行点. 由(11)，(13)易见 x 为原问题(P)的 $K-T$ 点.

ii)

$$\psi_r(x,d,s,B)\overset{\triangle}{=}\psi=\nabla f(x)^Td+\frac{1}{2}d^TBd-rs\varphi(x),$$

$$\psi=-\frac{1}{2}d^TBd-\sum_{j\in L}u_j\nabla g_j(x)^Td-rs\sum_{j\in L_1}g_j(x)_+-rs\sum_{j\in L_2}|g_j(x)|. \tag{15}$$

由(13)，(9)有

$$\sum_{j\in L_1}u_j\nabla g_j(x)^Td=-s\sum_{j\in L_1}u_jg_j(x),\quad \sum_{j\in L_2}u_j\nabla g_j(x)^Td=-s\sum_{j\in L_2}u_jg_j(x),$$

$$\psi=-\frac{1}{2}d^TBd+\sum_{j\in L_1}su_jg_j(x)-rs\sum_{j\in L_1}g_j(x)_++\sum_{j\in L_2}su_jg_j(x)-rs\sum_{j\in L_2}|g_j(x)|$$

$$\leqslant-\frac{1}{2}d^TBd+\sum_{\substack{g_j>0\\ j\in L_1}}(su_j-rs)g_j(x)+\sum_{j\in L_2}(|u_j|s-rs)|g_j(x)|.$$

由于 $r\geqslant\max\{|u_j|,j\in L\}$，$0\leqslant s\leqslant 1$，$B$ 正定，故 $\psi_r(x,d,s,B)\leqslant-\frac{1}{2}d^TBd<0$.

参考[3]，有

$$D_dF_r(x)=\nabla f(x)^T+r\sum_{j\in L_2,g_j>0}^{+}\nabla g_j(x)^Td-r\sum_{j\in L_2,g_j<0}^{-}\nabla g_j(x)^Td+r\sum_{j\in L_2,g_j=0}^{0}|\nabla g_j(x)^Td|$$

$$+r\sum_{j\in L_1,g_j>0}^{+}\nabla g_j(x)^Td-r\sum_{j\in L_1,g_j=0}\max\{0,\nabla g_j(x)^Td\}.$$

利用(9)，(10)，(11)有

$$D_d F_r(x) = - d^T B d^T - \sum_{j\in L} u_j \nabla g_j(x)^T d - r \sum_{j\in L_2}^{+} s g_j(x) + r \sum_{j\in L_2}^{-} s g_j(x) + r \sum_{j\in L_1}^{+} \nabla g_j(x)^T d,$$

$$D_d F_r(x) = - d^T B d - \sum_{j\in L} u_j \nabla g_j(x)^T d - rs \sum_{j\in L_2} |g_j(x)| + r \sum_{j\in L_1}^{+} \nabla g_j(x)^T d. \tag{16}$$

由(15),(16)得

$$\psi_r(x,d,s,B) - D_d F_r(x) = \frac{1}{2} d^T B d - r \sum_{j\in L_1}^{+} (s g_j(x) + \nabla g_j(x)^T d) \geqslant \frac{1}{2} d^T B d.$$

从而 $0 > \psi_r(x,d,s,B) \geqslant D_d F_r(x) + \frac{1}{2} d^T B d > D_d F_r(x)$,$d$ 为 F_r 在 x 处的下降方向.

在给出算法前对(P)作如下假设①:

(H) 当 $(d,s)=(0,0)$ 为 MQP(x,B) 的 $K-T$ 点时,原问题(P)无可行解或无最优解.

算法具体步骤:

步 0 选定一个正的常数 M②,给定 $x^1 \in E^n$,正定阵 $B_1 \in E^{n\times n}$,$\alpha \in (0,1)$,$\beta \in (0,1)$,$\varepsilon_0 > 0$,$r_0 > 0$. 令 $k=1$.

步 1 确定搜索方向

i) 求二次规划 MQP(x^k, B_k) 的 $K-T$ 点 (d_k, s_k).

记 $I_k = \{ j \in L \mid g_j(x^k) s_k + \nabla g_j(x^k)^T d_k = 0\}$,$u^k = (u_j^k, j\in L)$ 为 $K-T$ 乘子,$N_k = (\nabla g_j(x^k), j \in I_k)$.

如 $(d_k, s_k) = (0,1)$,则 x^k 为原问题(P)的 $K-T$ 点;如 $(d_k, s_k) = (0,0)$,则(P)无解,停. 如 $d_k \neq 0$ 进入 ii).

ii) 如 $\det(N_k^T N_k) = 0$,令 $\bar{d}_k = 0$. 否则令

$$\bar{d}_k = - Q_k \tilde{g}^k, \tag{17}$$

其中 $Q_k = N_k (N_k^T N_k)^{-1}$,$\tilde{g}^k = (g_j(x^k + d_k), j \in I_k)$.

步 2 调整参数 r_k

计算 $\tilde{r}_k = \max\{ |u_j^k|, j\in L\} + r_0$,如 $\tilde{r}_k \leqslant r_{k-1}$,令 $r_k = r_{k-1}$;否则令 $r_k = \max\{\tilde{r}_k, r_{k-1} + \varepsilon_0\}$.

步 3 线搜索:计算 $\psi \overset{\triangle}{=} \psi_{r_k}(x^k, d_k, s_k, B_k)$,求 $\{1, \beta, \beta^2, \cdots\}$ 中满足(18)式的最大值,并记其为 t_k.

$$F_{r_k}(x^k + t d_k + t^2 \bar{d}_k) \leqslant F_{r_k}(x^k) + \alpha t \psi_k. \tag{18}$$

步 4 令 $x^{k+1} = x^k + t_k d_k + t_k^2 \bar{d}_k$,$k := k+1$,利用某种修正公式,如[6,3],修正 B_k 得 B_{k+1},返回步 1.

由引理 2 之 ii)易知线搜索(18)可有限步终止. ①和②见 § 3 之后.

§ 3 算法的收敛速度

由上述讨论知如算法有限步终止,则或者终止于(P)的 $K-T$ 点,或者(P)无解. 为便于讨论,我们不妨假设算法产生无穷点列 $\{x^k\}$,且(P)有可行解. 为获得全局收敛性和超线性收敛速度,假设

(a) 矩阵 $\{B_k\}$ 为一致正定,且 $\{B_k\}$,$\{x^k, d_k, \bar{d}_k\}$ 有界.

定理 1 在假设(a)下,若 $\{|u_j^k|, j\in L\}$ 有界,则

1)　当 k 充分大时，罚参数 r_k 固定不变.

2)　$\{x^k\}$的任意极限点 x^* 为(P)的 $K-T$ 点.

证明　1) 若不然，存在无穷子列 $\mathscr{K}$，使得 $r_k=\max\{\bar{r}_k, r_{k-1}+\varepsilon_0\}, k\in\mathscr{K}$. 故而易见 $r_k\to+\infty$. 又 $r_{k-1}<\bar{r}_k\leqslant\sup\limits_k\{\bar{r}_k\}<+\infty$，因为$\{|u_j^k|, j\in L\}$有界. 由$\{r_k\}$的单调性知：

$$\lim_{k\to\infty} r_{k-1}\leqslant\sup_k\{\bar{r}_k\}<+\infty.$$

这与 $r_k\to+\infty$矛盾，故 1)成立.

至于 2)的证明可参照[7]中定理 1 或[1]，利用 1)的结论完成，为省篇幅，从略.

为讨论超线性收敛性，进一步假设

(b)　$x^k\to x^*, d_k\to 0, k\to\infty$. x^* 为(P)的 $K-T$ 点.

(c)　$s_k\to 1, k\to\infty, 1-s_k=O(\|d_k\|)$.

(d)　$I^*=\{j\in L\,|\,g_j(x^*)=0\}$；向量组$\{\nabla g_j(x^*), j\in I^*\}$线性无关，且在 x^* 处严格互补条件和二阶充分性条件成立. 即 x^* 之 $K-T$ 乘子 $u_j^*>0, \forall\ j\in I^*\cap L_1$. $\nabla^2 L(x^*, u^*)$在子空间$\{d\,|\,P_*d=0\}$上正定. 其中 $P_*=I-N_*(N_*^TN_*)^{-1}N_*^T, N_*=(\nabla g_j(x^*), j\in I^*)$.

引理 3　在(a)－(d)假设下，当 k 充分大时有

1)　$I_k\equiv I^*=\{j\in L_1\,|\,u_j^k>0\}\cup L_2$，从而 $\det(N_k^TN_k)\neq 0$，$\bar{d}_k$ 均由(17)式确定.

2)　$\exists\ \beta_1>0$ 使得：$\varphi_k=\sum\limits_{j\in L_2}|g_j(x^k)|+\sum\limits_{j\in I_1^*}g_j(x^k)_+$，$\sum\limits_{j\in I^*}u_j^kg_j(x^k)-r_k\varphi_k\leqslant-\beta_1\sum\limits_{j\in I^*}|g_j(x^k)|\leqslant-\beta_1|I^*|\cdot\|g^k\|$，其中 $I_1^*=I^*\cap L_1, g^k=(g_j(x^k), j\in I^*)$，$|I^*|$表示 I^*元素个数.

证明　1) 首先在(b)，(c)假设下易知 $I_k\subset I^*$. 又若有无穷子列 $\mathscr{K}$ 及 j_0 使得 $j_0\in I^*\backslash I_k$，$\forall\ k\in\mathscr{K}$. 从而 $j_0\in L_1\cap I^*=I_1^*$. 不妨设 $I_k\equiv I_0, \forall\ k\in\mathscr{K}$. 由(11)，(13)有

$$\nabla f(x^k)+B_kd_k+\sum_{j\in I_0}u_j^k\nabla g_j(x^k)=0. \tag{19}$$

由于 $I_0=I_k\subset I^*$，利用假设条件及(19)有

$$\nabla f(x^*)+\sum_{j\in I_0}u_j^*\nabla g_j(x^*)=0,\quad u_j^k\to u_j^*,\ j\in I_0. \tag{20}$$

由 $j_0\overline{\in} I_0$ 及 $K-T$ 乘子唯一性知 $u_{j_0}^*=0$，从而 $g_{j_0}(x^*)<0$，这与 $j_0\in I^*$矛盾. 故 $I_k\equiv I^*$，从而易证 $I_k=I^*=\{j\in L_1\,|\,u_j^k>0\}\cup L_2$. 而 $\det(N_k^TN_k)\to\det(N_*^TN_*)\neq 0$，于是 $\det(N_k^TN_k)\neq 0$，$\bar{d}_k$ 均由(17)式定义.

2)　首先由(20)及定理 1 知：当 k 充分大时 r_k 固定不变，在下面讨论中设 $r_k\equiv r$.

由于当 $j\in L_1\backslash I_1^*$ 时，$g_j(x^k)\to g_j(x^*)<0$. 从而当 k 充分大时 $g_j(x^k)_+=0$. 于是

$$\varphi_k=\sum_{j\in L_2}|g_j(x^k)|+\sum_{j\in I_1^*}g_j(x^k)_+,$$

$$\sum_{j\in I^*}u_j^kg_j(x^k)-r\varphi_k=\sum_{j\in L_2}[u_j^kg_j(x^k)-r|g_j(x^k)|]+\sum_{j\in I_1^*}u_j^kg_j(x^k)-r\sum_{j\in I_1^*}g_j(x^k)_+.$$

令 $I_{11}^*=\{j\in I_1^*\,|\,g_j(x^k)\leqslant 0\}, I_{12}^*=\{j\in I_1^*\,|\,g_j(x^k)>0\}$，则

$$\sum_{j\in I^*}u_j^k g_j(x^k)-r\varphi_k\leqslant\sum_{j\in L_2}(|u_j^k|-r)|g_j(x^k)|+\sum_{j\in I_{12}^*}(|u_j^k|-r)|g_j(x^k)|-\sum_{j\in I_{11}^*}u_j^k|g_j(x^k)|.$$

由于 $r\geqslant|u_j^k|+r_0,u_j^k\to u_j^*>0,j\in I_1^*$. 于是存在 $\beta_1>0,\beta_2=\beta_1|I^*|>0$ 使得

$$\sum_{j\in I^*}u_j^k g_j(x^k)-r\varphi_k\leqslant\beta_1\sum_{j\in I^*}|g_j(x^k)|\leqslant-\beta_2\|g^k\|.$$

引理 4 $\|\bar d_k\|=O(\|d_k\|^2),g_j(x^k+d_k+\bar d_k)=o(\|d_k\|^2),j\in I^*$.

证明 $\forall\ j\in I_k=I^*,s_kg_j(x^k+d_k)=s_kg_j(x^k)+s_k\nabla g_j(x^k)^Td_k+O(\|d_k\|^2)=(s_k-1)\nabla g_j(x^k)^Td_k+O(\|d_k\|^2)$. 由假设(c) 有 $|s_kg_j(x^k+d_k)|=O(\|d_k\|^2)$,因为 $s_k\to 1$,故 $|g_j(x^k+d_k)|=O(\|d_k\|^2)$.

于是由(17)式知 $\|\bar d_k\|=O(\|d_k\|^2)$,

$$\begin{aligned}g_j(x^k+d_k+\bar d_k)&=g_j(x^k+d_k)+\nabla g_j(x^k+d_k)^T\bar d_k+O(\|\bar d_k\|^2)\\&=g_j(x^k+d_k)+\nabla g_j(x^k)^T\bar d_k+o(\|d_k\|^2)=o(\|d_k\|^2).\end{aligned}$$

设 $P_k=I-N_k(N_k^TN_k)^{-1}N_k^T$,进一步假设

(e) $\|P_k(B_k-\nabla^2L(x^*,u^*))d_k\|=o(\|d_k\|)$.

引理 5 在(a)—(e)假设下,当 k 充分大时,$\bar d_k$ 均由(17)式确定,且 $t_k\equiv1$.

证明 记 $A_k=F_r(x^k+d_k+\bar d_k)-F_r(x^k)-\alpha\psi_k$.

为证 $t_k\equiv1$,即证 $A_k\leqslant0$. 由引理 3,4 及(3)—(5)知 $\bar d_k$ 均由(17)式确定,且

$$\begin{aligned}A_k&=f(x^k+d_k+\bar d_k)-f(x^k)-r\varphi_k-\alpha[\nabla f(x^k)^Td_k+\frac12d_k^TB_kd_k-rs_k\varphi_k]+o(\|d_k\|^2)\\&=(1-\alpha)\nabla f(x^k)^Td_k+\nabla f(x^k)^T\bar d_k+\frac12d_k^T\nabla^2f(x^k)d_k-\frac12\alpha d_k^TB_kd_k\\&\quad-(1-\alpha s_k)r\varphi_k+o(\|d_k\|^2).\end{aligned}$$

利用(11)或(19)及引理 4,有

$$\nabla f(x^k)^T\bar d_k=-\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^T\bar d_k-\bar d_k^TB_kd_k=-\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^T\bar d_k+o(\|d_k\|^2),$$

$$\nabla f(x^k)^Td_k=-\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^Td_k-d_k^TB_kd_k,$$

$$\begin{aligned}A_k&=(\frac12\alpha-1)d_k^TB_kd_k+\frac12d_k^T\nabla^2f(x^k)d_k-\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^T(d_k+\bar d_k)\\&\quad-(1-\alpha s_k)r\varphi_k+\alpha\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^Td_k+o(\|d_k\|^2).\end{aligned}$$

利用引理 4 将 $g_j(x^k+d_k+\bar d_k)$在 x^k 处展开可得

$$-\sum_{j\in I^*}u_j^k\nabla g_j(x^k)^T(d_k+\bar d_k)=\frac12\sum_{j\in I^*}u_j^kd_k^T\nabla^2g_j(x^k)d_k+\sum_{j\in I^*}u_j^kg_j(x^k)+o(\|d_k\|^2)$$

又 $\nabla g_j(x^k)^Td_k=-s_kg_j(x^k),\forall\ j\in I^*$,于是

$$\begin{aligned}A_k&=\frac12(\alpha-1)d_k^TB_kd_k+\frac12d_k^T[\nabla^2L(x^k,u^k)-B_k]d_k-(1-\alpha s_k)r\varphi_k\\&\quad+(1-\alpha s_k)\sum_{j\in I^*}u_j^kg_j(x^k)+o\|d_k\|^2).\end{aligned}$$

又令 $\widetilde{d}_k=d_k-P_kd_k=N_k(N_k^TN_k)^{-1}N_k^Td_k=N_k(N_k^TN_k)^{-1}(-s_kg_j(x^k),j\in I^*)$. 于是

$$\|\widetilde{d}_k\| = O(\|g^k\|).$$

由假设(a)知:∃ $\beta_3>0$ 使得

$$d_k^TB_kd_k \geqslant \beta_3\|d_k\|^2 \text{③} \tag{21}$$

$$A_k=\frac{1}{2}(a-1)d_k^TB_kd_k+\frac{1}{2}d_k^TP_k[\nabla^2L(x^k,u^k)-B_k]d_k-(1-as_k)r\varphi_k$$
$$+(1-as_k)\sum_{j\in I^*}u_j^kg_j(x^k)+\|g^k\|O(\|d_k\|)+o(\|d_k\|^2).$$

由假设(e)及(21)有

$$A_k\leqslant-\frac{1}{2}\beta_3(1-a)\|d_k\|^2+(1-as_k)[\sum_{j\in I^*}u_j^kg_j(x^k)-r\varphi_k]$$
$$+\|g^k\|O(\|d_k\|)+o(\|d_k\|^2).$$

由引理 3 之 2)有(其中 $\beta_2=\beta_1|I^*|$)

$$A_k\leqslant-\frac{1}{2}\beta_3(1-a)\|d_k\|^2-[(1-as_k)\beta_2-O(\|d_k\|]\|g^k\|+o(\|d_k\|^2).$$

由于 $d_k\to0,1-as_k\to1-a>0$,于是由上式知当 k 充分大时 $A_k\leqslant0$,从而 $t_k\equiv1$.

定理 2 在(a)-(e)假设下,$\{x^k\}$超线性收敛于 x^*. 即算法是超线性收敛的.

证明 由[7]或[1],易知在(a)-(e)假设下有

$$\frac{\|x^k+d_k-x^*\|}{\|x^k-x^*\|}\to0,\quad k\to\infty, \tag{22}$$

$$\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|}=\frac{\|x^k+d_k+\bar{d}_k-x^*\|}{\|x^k-x^*\|}$$
$$\leqslant\frac{\|x^k+d^k-x^*\|}{\|x^k-x^*\|}+\frac{\|d_k\|}{\|x^k-x^*\|}\cdot\frac{\|\bar{d}_k\|}{\|d_k\|}. \tag{23}$$

由(22)及引理 4 有$\{\frac{\|d_k\|}{\|x^k-x^*\|}\}$有界,$\frac{\|\bar{d}_k\|}{\|d_k\|}\to0$. 于是由(22),(23)可知

$$\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|}\to0,\quad k\to\infty.$$

即$\{x^k\}$超线性收敛于 x^*.

注① 第二节中假设(H)有时与问题(P)是互相抵触的,如[4]中例子所云. 但当 $g_j(j\in L)$都是线性函数时,假设(H)是成立的. 虽然[5]或[4]中修正的二次规划关于假设(H)成立的可能性也许比本文 MQP(x,B)大些(也不完全成立!),但[4]中方法对罚参数 r 限制较严,很难获得超线性收敛性(就本文方法而言).

注② 为有利于假设(b),(c)中 $d_k\to0,s_k\to1$ 的实现,结合引理 1,步 0 中 M 最好在迭代中不断增加.

注③ $\{B_k\}$一致正定的假设,即(21)式可减弱为:

$$\exists\ \beta_3>0,0\leqslant\beta_4\leqslant2\beta_2\ 使得\ d_k^TB_kd_k\geqslant\beta_3\|d_k\|^2-\beta_4\|g^k\|.$$

作者在完成此文时得到赖炎连研究员和薛声家教授的帮助,在此表示衷心的感谢.

参考文献

[1] Masao Fukushima, *A Successive Quadratic Programming Algorithm with Global and Superlinear Convergence Properties*, Mathematical Programming, 35(1986), 253—264.

[2] Eliane R. Panier and Andre L. Tits, *A superlinearly convergent feasible method for the solution of inequality constrained optimization problems*, SIAM J. Control and Optimization, Vol. 25, No. 4 (1987), 934—950.

[3] M. J. D. Powell, *A fast algorithm for nonlinearly constrained optimization caculations*, In: G. A. Waston, ed., Numerical Analysis Dundee 1977(Springer-Verlag, Berlin, 1979), 144—157.

[4] Karou Tone, *Revisision of constraint approximations in the successive QP method for nonlinear programming problems*, Mathematical Programming, 26(1983), 144—152.

[5] M. C. Bartholomew-Biggs, *Recursive quadratic programming methods for nonlinear constraints*, in: M. J. D. Powell, ed., Nonlinear Optimization, 1981(Academic Press, New York, 1982), 213—221.

[6] M. J. D. Powell, *The convergence of variable metric methods for nonlinearly constrained optimization caculations*, in: O. L. Magasariam, R. R. Meyer and S. M. Robinson, ed., Nonlinear Programming 3(Academic Press, New York, 1978), 27—63.

[7] M. J. D. Powell, *Variable metric methods for constrained optimization*, in: A. Bachem, M. Grötschel and B. Korte, ed., Mathematical Programming: The state of the art, Bonn 1982 (Springer-Verlag, Berlin, 1983), 288—311.

[8] S. P. Han, *A globally convergent method for nonlinear programming*, Journal of Optimization Theory and Applications, 22(1977), 297—309.

A Modified Algorithm of SQP Type

Jian Jinbao

(Dept. of Math., Guangxi University, Nanning 530004)

Abstract

This paper presents an algorithm of successive quadratic programming (SQP) type for programming problems with nonlinear equality and inequality constraints. Under some suitable conditions, we prove the global and superlinear convergence properties.

Keywords general nonlinear programming, successive quadratic programming, exact penalty functions, rate of convergence.