第 30 卷　第 1 期　　　　电 子 科 技 大 学 学 报　　　　Vol.30　No.1
2001 年 2 月　　　　Journal of UEST of China　　　　Feb. 2001

非线性规划一般约束条件的 SQP 方法

滕召波*　　张世永　　陈华富　　何光中
(四川大学应用数学系　成都　610065)　(电子科技大学应用数学系　成都　610054)　(四川大学应用数学系　成都　610065)

【摘要】 提出了一种新的处理等式和不等式约束条件优化问题的 SQP 方法，计算过程中每一步迭代只需解一个二次规划．在一定条件下，证明了算法的全局和二步超线性收敛性，其优点是具有较小的计算量，避免了 Maratos 现象的发生．

关　键　词　非线性规划；　序列二次规划；　马洛托斯现象；　全局和超线性收敛性
中图分类号　O212.2

由于序列二次规划(SQP)算法对于非线性不等式约束下的最优化问题是非常有效的，并且具有良好的超线性收敛性质，因而此类算法在非线性规划中已占有非常重要的地位。为了处理等式约束条件，通常要使用罚函数，但使用非精确罚函数常常产生 Maratos 现象。解决此问题的方法有文献[1]的方法和文献[2]的 Watchdog 技术。本文提出了一个新的 SQP 方法，该方法不需要计算目标函数和约束条件的二阶偏导，克服了文献[3]中计算量太大的不足。在一定的条件下，证明了该算法具有全局收敛性和二步超线性收敛性。

1　主算法及性质

本文讨论如下非线性规划问题

$$\text{(P)}\min\ f(x)$$

$$\text{s.t}\begin{cases} h_i(x)\leqslant 0 & i\in K=\{1,2,\cdots,m'\} \\ h_i(x)=0 & i\in J=\{m'+1,m'+2,\cdots,m\} \end{cases}$$

假设：1) 可行域非空，$f(x),h_i(x),i\in K\cup J$ 为三阶连续可微函数；2) 对 $\forall x\in R$, 记 $I'(x)=\{i\in K \mid h_i(x)=0\},L=I'(x)\cup J$ 。假设向量组 $\{\nabla h_i(x) \mid i\in L(x)\}$ 线性无关。对于任何 $I\subset L,x\in R,N_I=\{\nabla h_i(x) \mid i\in I\}$，$G_c(x)=f(x)+C\sum\limits_{i\in J}|h_i(x)|,C>0,C>\sup\{\max\limits_{1\leqslant i\leqslant m}|\tilde{m}_i|\}$，$\tilde{m}_i$ 是的最优 Lagrange 乘子，$DG_c(x,d)$ 表示 x 处 $G_c(x)$ 在方向 d 上的方向导数。

算法步骤如下：

步骤 1　选定参数，$\boldsymbol{a}\in\left(0,\frac{1}{2}\right),\boldsymbol{h}>2,\boldsymbol{t}\in(2,3),\boldsymbol{b}\in(0,1),x^1\in R,\boldsymbol{H}_1$ 为对称正定阵，$k:=1$，$\boldsymbol{s}\in(0,1)$；

步骤 2　计算搜索方向：

1) 求 d_k^0，解

$$\min\frac{1}{2}d^TH_kd+\langle\nabla f(x^k),d\rangle$$

$$\text{s.t}\begin{cases} h_i(x^k)+\nabla h_i(x^k)^Td\leqslant 0 & i\in K \\ h_i(x^k)+\nabla h_i(x^k)^Td=0 & i\in J \end{cases}\qquad(1)$$

若式(1)无 $K\text{–}T$ 点，或 $\|H_kd_k^0\|>\|d_k^0\|^{\frac{1}{2}}$，则转4)；若 d_k^0=0，则 x^k 为(P)的 $K\text{–}T$ 点，停，否则转2)。

1999年9月1日收稿
* 男　28岁　大学　讲师

2) 令$I_k=\{i\in I_0^k\cup J\left|h_i(x^k)+\nabla h_i(x^k)^T d=0\right.\}, I_0^k=\{i\in K\left|h_i(x^k)+\nabla h_i(x^k)^T d=0\right.\}$，再令$N_k=N_{I_k}(x^k)$，如果$\det(N_k^T N_k)\neq 0$，则计算$Q_k=(N_k^T N_k)^{-1}N_k^T$，计算方向

$$d^k=\begin{cases} d_k^0 & \det(N_k^T N_k)=0 \\ d_k^0-Q_k^T(\left\|d_k^0\right\|^t e^k+h^{\tilde{k}}) & \det(N_k^T N_k)\neq 0 \end{cases} \tag{2}$$

式中 $e^k=(1,1,\cdots,1)^T$，$h^k=h_i(x^k+d_k^0)$，$i\in I_k$。

3) 若$q_k=\nabla f(x^k)^T d_k^0\leqslant\min\{-\left\|d_k^0\right\|^h,-\left\|d^k\right\|^h\}$，$f(x^k+d^k)\leqslant f(x^k)+\alpha q_k$，$h_i(x^k+d^k)\leqslant 0$，$i\in K$；$-sDG_c(x^k,d^k)\leqslant G_c(x^k)-G_c(x^k+d^k)$，则令$l_k=1$，转步骤4，否则转到4)。

4) 计算一个满足一阶下降条件的方向d^k

$$\begin{cases} \nabla f(x^k)^T d_k^0<0, \nabla h_i(x^k)^T d^k<0, i\in I(x^k) \\ -s\,lDG_c(x^k,d^k)\leqslant G_c(x^k)-G_c(x^k+ld^k) \end{cases} \tag{3}$$

存在$m>0,a_0>1$，使

$$\nabla f(x^k)^T d^k\leqslant -c\left\|d^k\right\|^{a_0} \tag{4}$$

步骤3 线搜索：求$\{1,b,b^2,\cdots\}$中，最大值l^k满足

$$f(x^k+ld^k)<f(x^k)+\alpha lq_k, h_i(x^k+ld^k)\leqslant 0 \qquad i\in K \tag{5}$$

$$-s\,lDG_c(x^k,d^k)\leqslant G_c(x^k)-G_c(x^k+ld^k) \tag{6}$$

步骤4 用某种修正公式修正H_k得正定阵H_{k+1} 令$x^{k+1}=x^k+l_k d^k, k:=k+1$，转回步骤2。

命题1[5,6] 对$\forall x\in R^n, d\in R^n$，有

$$DG_c(x,d)=\nabla f(x)^T d+C\sum_{\substack{i\in J\\ h_i(x)>0}}\nabla h_i(x)^T d+C\sum_{\substack{i\in J\\ h_i(x)=0}}\left|\nabla h_i(x)^T d\right|-C\sum_{\substack{i\in J\\ h_i(x)<0}}\nabla h_i(x)^T d$$

命题2 步骤3满足线性搜索的l_k一定存在，即存在非负整数j，使$l_k=b^j$。

2 算法的收敛性证明

定理1 若$d_k^0=0$为式(1)的$K\text{–}T$点，则x^k为原问题(P)的$K\text{–}T$点。

命题3 设$x^k\to x^*, k\in N_1, (\left|N_1\right|<\infty)$对于每一个$k\in N$。$d_k^0, d^k$，由步骤2中1)和2)确定，$x^{k+1}=x^k+d^k$；由步骤2中3)确定，则$d_k^0\to 0, k\to\infty, k\in N_1$。

定理2 对于满足R_1,R_2的一阶下降可行方向，若算法有限步终止于x^k，则x^k为原问题(P)的$K\text{–}T$点，否则其任一极限点均为(P)的$K\text{-}T$点。

证明 由定理1，只需证明定理2后面一部分即可。设$x^k\to x^*, k\in N_1$，由$\{f(x^k)\}_{k=1}^{\infty}$单调下降及$f(x^k)\to f(x^*), k\in N_1$($f$连续可微)，知$f(x^k)\to f(x^*), k\to\infty$。分为以下两种情形讨论：

1) 存在无穷子列$N_2\subset N_1$，使$d^k(\forall k\in N_2)$由步骤2中4)确定。

假设x^k不为原问题(P)的$K\text{–}T$点，有

$$\begin{cases} h_i(x^k+ld^k)=h_i(x^k)+l\nabla h_i(x^k)^T d^k+o(l) & i\in I(x^*) \\ f(x^k+ld^k)=f(x^k)+l\nabla f(x^k)^T d^k+o(l) \end{cases} \tag{7}$$

由式(3)有

$$\begin{cases} h_i(x^k+ld^k)\leqslant h_i(x^k)-ld+o(l) & i\in I(x^*) \\ f(x^k+ld^k)\leqslant f(x^k)-ld+o(l) \end{cases} \tag{8}$$

又因$\{\|d^k\|\}_{k\in N_2}$有界，$h_i(x^k)\to h_i(x^*)<0, i\in K/I(x^*)$。

由式(8)和步骤3知，存在 $\bar{l} > 0$，使 $l^k \geqslant l, \forall k \in N_2$。故 $f(x^{k+1}) \leqslant f(x^k) + l_k q_k \leqslant f(x^k) - \bar{l} d, \forall k \in N_2$，令 $k \to \infty, k \in N_2$，有 $l\,d \leqslant 0$，矛盾。

2) 设 $\forall k \in N_1, d_k^0, d^k$ 由步骤2中1)和2)确定；$x^{k+1} = x^k + d^k$ 由步骤2中3)确定。由命题3可知 $d_k^0 \to 0, k \to \infty, k \in N_1$。由步骤2中1)知 $\|H_k d_k^0\| < \|d_k^0\|^{\frac{1}{2}}$，故 $H_k d_k^0 \to 0, k \to \infty, k \in N_1$。

假设 $I_k \equiv I, \forall k \in N_1$，由 d_k^0 为 $K\text{–}T$ 点，有

$$\begin{cases} \nabla f(x^k) + H_k d_k^0 + \sum\limits_{i \in K} m_i^k \nabla h_i(x^k) + \sum\limits_{i \in J} n_i^k \nabla h_i(x^k) = 0 \\ m_i^k h_i(x^k) = 0 \qquad m_i^k \leqslant 0,\ i \in K \end{cases} \tag{9}$$

写成矩阵的形式为

$$\begin{cases} \nabla f(x^k) + H_k d_k^0 + N_k m_0^k = 0 \\ m_i^k \geqslant 0 \qquad i \in K \end{cases} \tag{10}$$

因 $d_k^0 \to 0, x^k \to x^*, k \to \infty, k \in N_1$，知 $I \subset I(x^*)$。

令 $N^* = (\nabla h_i(x^*), i \in I)$，当 $k \in N_1$ 充分大时，$\det(N_k^T N_k) \neq 0$，并且有 $(N_k^T N_k)^{-1} \to (N^{*T} N^*)^{-1}$，$N_k \to N^*, k \in N_1$，由式(9)有 $m_0^k = -(N_k^T N_k) N_k^T (\nabla f(x^k) + H_k d_k^0) - (N^{*T} N^*)^{-1} N^{*T} \nabla f(x^*) \triangleq m_0^*$，其中 $(m_0^k)^T = (m_1^k, \cdots, m_{m'}^k, n_{m'+1}^k, \cdots, n_m^k)$。在式(9)中，令 $k \to \infty, k \in N_1$，有

$$\begin{cases} \nabla f(x^*) + \sum\limits_{i \in K} m_i^* \nabla h_i(x^*) + \sum\limits_{i \in J} n_i^* \nabla h_i(x^*) = 0 \\ m_i^* h_i(x^*) = 0 \qquad m_i^* \geqslant 0,\ i \in K \end{cases}$$

即 x^* 为(P)的 $K\text{–}T$ 点。

命题 4　若一阶下降方向满足 R_2，且算法产生一无穷点列 $\{x^k\}$，并且有聚点，那么 $\{\| x^{k+1} - x^k \|\} \to 0, k \to \infty$。

3　算法的超线性收敛性

假设点列 $\{x^k\}$ 具有满足二阶充分条件和严格互补条件的极限点。

命题 5　算法迭代能产生点列 $\{x^k\}$，并有聚点 x^*，则整个点列收敛到 x^*。

记 $N^* = (\nabla h_i(x^*), i \in I)$，$P^* = E - N^*(N^{*T} N^*)^{-1} N^{*T}$；充分靠近 x^* 时，定义如下：$N_k = (\nabla h_i(x^k), i \in I)$，$P_k = E - N_k (N_k^T N_k)^{-1} N_k^T$，$E$ 为单位矩阵。文中再假设：矩阵序列 $\{H_k\}$ 有极限 H^*，且 $P^* H^* P^* = P^* \nabla_{xx}^2 L(x^*, m^*, n^*) P^*$，其中 $\nabla_{xx}^2 L(x^*, m^*, n^*) = \nabla^2 f(x^*) + \sum\limits_{i=1}^{m'} m^* \nabla^2 h_i(x^*) + \sum\limits_{i=m'+1}^{m} n^* \nabla^2 h_i(x^*)$ 表示 Lagrangian 函数 $L(x, m, n)$ 的 Hessian 阵，则存在 $r > 0$，当 k 充分大时，$d^T P_k H_k P_k d \geqslant r \| P_k d \|^2, \forall d \in E^m$。

命题 6　当 k 充分大时：

1) 二次规划 $K\text{–}T$ 点 d_k^0，$d_k^0 \to 0, k \to \infty$；

2) $I_k = \{i \mid h_i(x^k) + \nabla h_i(x^k)^T d_k^0 = 0\} = I(x^*) \triangleq I_0$，$m_i^k \to m_i^*, i \in I(x^*) = I_0$，$n_i^k \to n_i^*, k \to \infty$。$m_i^k, n_i^k (i \in I_k)$ 为 d_k^0 的 $K\text{–}T$ 乘子；

3) $\det(N_k^T N_k) \neq 0$，步骤2中2)定义的方向 $d_k^0 \to 0, k \to \infty$，且 $\|d^k\| \sim \|d_k^0\|$。

定理 3　当 k 充分大时，方向 d^k 均由式(1)确定，且 $x^{k+1} = x^k + d^k$ 由步骤2中3)确定。

定理 4 本文算法产生的点列$\{x^k\}$二步超线性收敛于问题(P)的严格局部最优解x^*，即

$$\frac{\| x^{k+1} - x^* \|}{\| x^{x-1} - x^* \|} = \frac{\| x^{k-1} + d^{k-1} + d^k - x^* \|}{\| x^{k-1} - x^* \|} \to 0$$

参 考 文 献

1 Mayne D Q, Polak E. A superlinearly convergent algorithm for constrained optimization problem. Mathematical Programming Study,1982,16: 45~61

2 Chamberlain R M, Powell M J D, Lemarechal C, *et al*. The Watchdog technique for forcing convergence in algorithms for constrained optimization. Mathematical Programming Study,1982,16: 1~17

3 Fukushima Masao. A successive quadratic programming algorithm with global and superlinear convergence properties.Math Prog, 1986, 35: 253~264

4 Panier E R,Tits A L. A superlinear convergence feasible method for the solution of in-equality constrained optimization problems, SIAM J Control and Optimization, 1987, 25(4): 934~950

5 陈华富. 一般约束极大极小值的梯度投影算法. 电子科技大学学报, 2000, 29(6): 662~665

6 陈华富, 田益祥. 一般约束极大极小问题的广义梯度投影算法. 电子科技大学学报, 2000, 29(3): 319~322

A Successive Quadratic Programming Algorithm with Global and Superlinear Convergence Properties

Teng Zhaobo Zhang Shiyong

(Dept. of Applied Math., Sichuan University Chengdu 610065)

Chen Huafu

(Dept. of Applied Math., UEST of China Chengdu 610054)

He Guangzhong

(Dept. of Applied Math., Sichuan University Chengdu 610065)

Abstract This paper present a new variant of successive quadratic programming methods with inequality and equality constraint. In calculation process, this methods need only to solve a quadratic programming problem at each iteration. In certain conditions, the global and fast local convergence properties of this algorithm is proved, whose step size of unity is acceptable and whose Maratos effect is obviated.

Key words nonlinear programming; successive quadratic programming; Maratos effect; global and superlinear convergence