(20) 125-130

非线性规划的一种全局收敛算法*

A Globally Convergent Algorithm for Nonlinear Programming

O221.2

赵云彬 段虞荣
Zhao Yunbin Duan Yurong
(重庆大学系统工程及应用数学系，重庆，630044)

A

摘　要　定义了一种偏离 Kuhn-Tucker 三元点的度量函数的基础上，对一般连续可微非线性规划提出了一个新的全局收敛算法。利用这个算法在获得问题最优解的同时，还得到了与最优解相应的 Lagrange 乘子。把这种算法应用于二次规划，得到了二次规划的一种新的迭代法。最后给出了一个计算实例。

关键词　偏离 Kuhn-Tucker 三元点的度量函数；全局收敛；可行域；垂直射影

中国图书资料分类法分类号　O221.2

3 非线性规划

ABSTRACT　A new globally convergent algorithm was presented for a continuous differentiable nonlinear programming by defining a measure function deviating from the Kuhn-Tucker point. With this algorithm which can be used to get the optimum solution of the problem, the optimal lagrangian multiplier corresponding the optimal solution of the problem was also obtained. A new iterative algorithm for quadratic programming is obtained when applying the general algorithm to quadratic programming. Finally, a numerical example was given.

KEYWORDS　measure function deviating from the Kuhn-Tucker triad point; global convergence; feasible region; vertical projection

0 引　言

考虑如下非线性规划问题

$$(NLP) \quad \begin{cases} \min \quad f(x) \quad x \in R^n \\ s.t. \quad g_i(x) \geqslant 0, \quad i = 1, \cdots, m \\ \qquad\quad h_j(x) = 0, \quad j = 1, \cdots, l \end{cases} \tag{1}$$

其中 $f(x), g_i(x), i = 1, \cdots, m$； $h_j(x), j = 1, \cdots, l$ 都是连续可微函数，约束规格(constraint qualification)假设为：起作用约束的梯度向量线性无关。

记 $x = (x_1, \cdots, x_n)^T \in R^n, \omega = (\omega_1, \cdots, \omega_m)^T \in R^m, v = (v_1, \cdots, v_l)^T \in R^l, u = (x^T, \omega^T, v^T)^T \in R^{n+m+l}$. 为方便起见，把 u 简记成 $u = (x, \omega, v)$.

---

* 收文日期　1991-06-21

问题(1)的广义 Lagrange 函数为

$$L(u) = f(x) - \omega^T g(x) - v^T h(x) \tag{2}$$

这里 $g(x) = (g_1(x), \cdots, g_m(x))^T, h(x) = (h_1(x), \cdots, h_l(x))$.

点 $u \in R^{n+m+l}$ 是问题(1)的 Kuhn-Tucker 三元点。(Kuhn-Tucker triad point)的充要条件是：

$$\begin{cases} \nabla_x L(u) = \nabla f(x) - \sum\limits_{i=1}^{m} \omega_i g_i(x) - \sum\limits_{j=1}^{l} v_j \nabla h_j(x) = 0 \\ g_i(x) \geqslant 0, \quad i = 1, \cdots, m \\ h_j(x) = 0, \quad j = 1, \cdots, l \\ \omega_i g_i(x) = 0, \quad i = 1, \cdots, m \\ \omega_i \geqslant 0, \quad i = 1, \cdots, m \end{cases} \tag{3}$$

定义如下函数

$$Q(u) = Q(x, \omega, v) = \|\nabla_x L(u)\|^2 + \sum_{i=1}^{m} [\omega_i^2 g_i^2(x) + \omega_{i+}^2] \tag{4}$$

其中 $\omega_{i+} = \max\{0, \omega_i\}$. 显然 $Q(u)$ 是个连续函数.记 $S = \{x \in R^n \mid g_i(x) \geqslant 0, i = 1, \cdots, m; h_j(x) = 0, j = 1, \cdots, l\}$ 为原问题(1) 的可行域。假设 S 是闭集。下面的结论是明显的。

引理 1　若 $u = (x, \omega, v) \in R^{n+m+l}, x \in S$ 且满足

$$Q(u) = 0 \tag{5}$$

则 $u = (x, \omega, v)$ 是问题(1) 的 Kuhn-Tucker 点,反之若 u 是问题(1) 的 Kuhn-Tucker 三元点,则必有 $x \in S$ 且(5) 式成立。

以上结论正是我们以下算法的实质,求解问题(1) 转化成寻找 $Q(u)$ 在可行域 S 中的零点。我们把(4) 式定义的函数 $Q(u)$ 称为偏离 Kuhn-Tucker 三元点的度量函数。因为若 $x \in S$, u 不是 Kuhn-Tucker 三元点,则 $Q(u) > 0$,若 $x \in S$,u 是 Kuhn-Tucker 三元点,则 $Q(u) = 0$. 因此 $Q(u)$ 反映了 u 点偏离 Kuhn-Tucker 三元点的程度。

1　算法及其全局收敛性

定义 1　设问题(1) 的可行域是闭集.对任意 $\bar{x} \in R^n$,子问题

$$\begin{cases} \min \quad \dfrac{1}{2}(x - \bar{x})^T (x - \bar{x}) \\ \text{s.t.} \quad x \in S \end{cases} \tag{6}$$

的最优解为 $\bar{x}^*$,则称 $\bar{x}^*$ 为 $\bar{x}$ 在可行域 S 中的垂直射影。

定义 2　若存在连续函数 $\beta(u): R^{n+m+l} \to R^{n+m+l}$ 使得对任意 $u' \in R^{n+m+l}$ 有

$$(u' - u)^T \beta(u) \leqslant Q(u') - Q(u) \tag{7}$$

则称 $\beta(u)$ 为 $Q(u)$ 在 u 点的广义下降方向。

特别地取 $u' = u^* \in K^*$(Kuhn-Tucker 三元点全体),则 $Q(u^*) = 0$,对 $u \in R^{n+m+l}$,若 $\beta(u)$ 存在,则(7) 式可以简写成

$$(u^* - u)^T \beta(u) \leqslant -Q(u) \tag{8}$$

若 $u \notin K^*$,且 $x \in S$ 时 $Q(u) > 0$,由(8) 式知 $\beta(u) \neq 0$,这保证了后面算法中的 d^k 计算有意义(参见(10) 式)。

如果 $Q(u)$ 是连续可微的凸函数时，必然存在 $\beta(u)$ 使(7)式成立，事实上 $\beta(u)=\nabla Q(u)$ 满足要求，即

$$(u'-u)\nabla Q(u)\leqslant Q(u')-Q(u) \tag{9}$$

下面在假设问题(1)的函数 $Q(u)$ 存在广义下降方向条件下，设计了问题(1)的一个比较一般的迭代算法。

算法 1

1) 给定 $\varepsilon>0$，任意初始点 $u^0\in R^{n+m+l}$，令 $k:=0$.

2) 假设已得到 u^k，如果 $Q(u^k)<\varepsilon$，停止，否则按下列步骤计算 $\bar{u}^{k+1}=(\bar{x}^{k+1},\bar{\omega}^{k+1},\bar{v}^{k+1})$

$$d^k=Q(u^k)\ /\ \|\beta(u^k)\|^2 \tag{10}$$

$$\bar{u}^{k+1}=u^k-d^k\beta(u^k) \tag{11}$$

3)求解如下极值子问题

$$\begin{cases}\min & \dfrac{1}{2}(x-\bar{x}^{k+1})^T(x-\bar{x}^{k+1})\\ s.t. & g_i(x)\geqslant 0,i=1,\cdots,m\\ & h_j(x)=0,j=1,\cdots,l\end{cases} \tag{12}$$

得到 $\bar{x}^{k+1}$ 在可行域 S 中的垂直射影 $\hat{x}^{k+1}$.

4) 令 $u^{k+1}=(\hat{x}^{k+1},\bar{\omega}^{k+1},\bar{v}^{k+1})$，$k:=k+1$，返回 2).

算法3)的作用是明显的，将可行域外的当前迭代点 $\bar{u}^{k+1}$ 投影到可行域中。如果点 $\bar{x}^{k+1}$ 在可行域中，则其投影 $\hat{x}^{k+1}=\bar{x}^{k+1}$. 这样算法产生的点列 $\{u^k\}$ 包含在可行域中。我们对可行域作一个要求，假设它满足如下性质。

(AS)　若 $\bar{x}^*$ 是 $\bar{x}$ 在可行域中的投影，则

$$\|\bar{x}^*-x\|\leqslant\|\bar{x}-x\|,\forall\ x\in S \tag{13}$$

如果 S 是闭凸集，则(13)式自然成立。

引理 2　假设问题(1)的可行域是闭集且满足性质(13). 设 $u^*\in K^*$(问题(1)的 Kuhn-Tucker 三元点全体)，$u^0\in R^{n+m+l}$ 是任一迭代初始点，$\{u^k\}$ 是算法 1 产生的序列，且 $u^k\notin K^*$，则

1)
$$\|\bar{u}^{k+1}-u^*\|^2\leqslant\|u^k-u^*\|^2-d^kQ(u^k) \tag{14}$$

2)
$$\|u^{k+1}-u^*\|^2\leqslant\|\bar{u}^{k+1}-u^*\|^2 \tag{15}$$

证　1) 由(8)，(10)，(11)三式得

$$\begin{aligned}\|\bar{u}^{k+1}-u^*\|^2&=\|u^k-u^*-d^k\beta(u^k)\|^2\\&=\|u^k-u^*\|^2-2d^k(u^k-u^*)^T\beta(u^k)+d^k\cdot\|\beta(u^k)\|^2\\&\leqslant\|u^k-u^*\|^2-2d^kQ(u^k)+(d^k)^2\cdot\|\beta(u^k)\|^2\\&=\|u^k-u^*\|^2-2d^kQ(u^k)+(d^k)^2\cdot\frac{Q(u^k)}{d^k}\\&=\|u^k-u^*\|^2-d^kQ(u^k)\end{aligned}$$

2) $\bar{u}^{k+1}=(\bar{x}^{k+1},\bar{\omega}^{k+1},\bar{v}^{k+1})$，$u^{k+1}=(\hat{x}^{k+1},\bar{\omega}^{k+1},\bar{v}^{k+1})$，$u^*=(x^*,\omega^*,v^*)$，$\hat{x}^{k+1}$ 是 $\bar{x}^{k+1}$ 在可行域 S 中的垂直射影，满足性质(13)，从而

$$\|u^{k+1}-u^*\|^2=\|\hat{x}^{k+1}-x^*\|+\|\bar{\omega}^{k+1}-\omega^*\|^2+\|\bar{v}^{k+1}-v^*\|^2$$

$$\leqslant \|\bar{x}^{k+1} - x^*\|^2 + \|\bar{\omega}^{k+1} - \omega^*\|^2 + \|\bar{v}^{k+1} - v^*\|^2$$
$$= \|\bar{u}^{k+1} - u^*\|^2$$

推论1 在引理2条件下，有

$$\|u^{k+1} - u^*\|^2 \leqslant \|u^k - u^*\|^2 - d^k \cdot Q(u^k) \tag{16}$$

定理1 问题(1)的可行域是闭集且满足性质(13)，算法1产生的序列$\{u^k\}$，若在某一步有$Q(u^k)=0$，则u^k是问题(1)的Kuhn-Tucker三元点，若产生的无穷序列$\{u^k\} \cap K^* = \Phi$，则$\{u^k\}$存在收敛子序列，且任意这样的收敛子序列其极限点在K^*中，从而也得到问题的Kuhn-Tucher三元点。

证 若$u^k = (x^k, \omega^k, v^k)$是由算法1产生的点，则$x^k$是可行域中的点。如果$Q(u^k)=0$，由引理1知$u^k$已是问题(1)的Kuhn-Tucker点。若$\{u^k\} \cap K^* = \Phi$，当然有$Q(u^k) > 0$，由(16)式知

$$0 \leqslant \|u^{k+1} - u^*\| \leqslant \|u^k - u^*\| \tag{17}$$

序列$\{\|u^k - u^*\|\}$为严格单调下降非负序列，从而存在$C \geqslant 0$，使

$$\lim_{k\to\infty} \|u^k - u^*\| = C \tag{18}$$

从而对任意k，有

$$C \leqslant \|u^{k+1} - u^*\| < \|u^k - u^*\| \tag{19}$$

记$T(u) = d(u) \cdot Q(u)$，其中$d(u) = Q(u) / \|\beta(u)\|^2$，由(16)式知

$$0 \leqslant T(u^k) \leqslant \|u^k - u^*\|^2 - \|u^{k+1} - u^*\|^2 \tag{20}$$

由$\lim\limits_{k\to\infty}(\|u^k - u^*\|^2 - \|u^{k+1} - u^*\|^2) = C^2 - C^2 = 0$，再注意到(19)式可知

$$\lim_{k\to\infty} T(u^k) = 0 \tag{21}$$

从(20)式可知$\{u^k\}$包含在有界子集$\{u \in R^{n+m+l} | C \leqslant \|u - u^*\| \leqslant \|u^0 - u^*\|\}$之中，必然有收敛子序列。设$\{u^{k_j}\}$为$\{u^k\}$的任一收敛子序列，记$\lim\limits_{k\to\infty} u^{k_j} = \bar{u}^*$，由(21)式得

$$0 = \lim_{k\to\infty} T(u^{k_j}) = \lim_{k\to\infty} d(u^{k_j}) \cdot Q(u^{k_j}) = \lim_{k\to\infty} \frac{Q^2(u^{k_j})}{\|\beta(u^{k_j})\|^2} \tag{22}$$

由于$\{u^{k_j}\}$包含在紧子集中以及$\beta(u)$的连续性(定义2)知$\{\|\beta(u^{k_j})\|^2\}$是有界的。从而由(22)式得

$$\lim_{k\to\infty} Q(u^{k_j}) = 0$$

再由$Q(u)$的连续性得

$$\lim_{k\to\infty} = Q(u^{k_j})Q(\bar{u}^*) = 0$$

从而$\bar{u}^* \in K^*$(引理1)。 证毕

由于算法对u^0不作任何限制，从R^{n+m+l}中任一点开始迭代，算法都收敛到问题的$K-T$三元点，从而算法是全局收敛的。如果问题(1)的广义下降方向$\beta(u)$存在而且子问题(5)易解的情况下，算法具有迭代简单的特点，因而不失为一种较好的方法。特别是当约束是线性约束时，子问题(5)是个线性规划问题，算法1求解这类问题将是很方便的。

但是我们也注意到算法依赖于$\beta(u)$的存在性，这无疑对算法的直接应用起了限制作用，另外算法步骤3)要求解一个子规划问题，在一般情况下，其工作量不亚于对原问题的求解。不过上述算法仍然具有积极的意义。下面我们用算法1来求解二次规划问题。得到了二次规划的一种新的迭代解法。

2　二次规划的迭代解法

考虑二次规划问题

$$\begin{cases}\min & \frac{1}{2}x^THx + b^Tx + c \\ s.t. & Ax = P\end{cases} \tag{23}$$

其中 A 是 $m\times n$ 矩阵，$m\leqslant n$，$\mathrm{rank}(A) = m$，$x\in R^n$，$P\in R^m$，H 是 $n\times n$ 对称矩阵。

上述问题的 Lagrange 函数为

$$L(u) = \frac{1}{2}x^THx + b^Tx + C - v^T(Ax - P) \tag{24}$$

因此

$$\nabla xL(u) = Hx + b - A^Tv \tag{25}$$

$$Q(u) = \|\nabla xL(u)\|^2 = \|Hx - A^Tv + b\|^2 \tag{26}$$

显然 $Q(u)$ 关于 $u = (x,v)$ 是凸函数，因而 $\beta(u)$ 存在，即

$$\beta(u) = \nabla Q(u) = 2\begin{pmatrix}H \\ -A\end{pmatrix}\left[(H,\ -A^T)\begin{pmatrix}x \\ v\end{pmatrix} + b\right] \tag{27}$$

子规划(5)可写成

$$\begin{cases}\min\ \frac{1}{2}(x-\bar{x})^T(x-\bar{x}) \\ s.t.\quad Ax = P\end{cases} \tag{28}$$

其 Kuhn-Tucker 条件为

$$x - \bar{x} - A^TZ = 0 \tag{29}$$

$$Ax = P \tag{30}$$

这里 z 表示 Lagrange 乘子向量。用 A 左乘(29)两边得

$$A(x-\bar{x}) = AA^Tz$$

再利用(30)式，从上述等式中解出 z 得

$$z = (AA^T)^{-1}(P - A\bar{x}) \tag{31}$$

将(31)式代入(29)式得

$$\begin{aligned}x &= \bar{x} + A^T(AA^T)^{-1}(P - A\bar{x}) \\ &= (I - A^T(AA^T)^{-1}A)\bar{x} + A^T(AA^T)^{-1}P\end{aligned} \tag{32}$$

(32)式就是 $\bar{x}$ 在可行域 $\{x\mid x\in R^n, Ax = P\}$ 中的垂直射影。这样，从算法1我们得到了如下二次规划的一种新的迭代解法。

1) 给定 $\varepsilon > 0$，$\forall\ u^0\in R^{n+m}$，令 $k := 0$

2) 设已得到 $u^k = (x^k, v^k)$，令 $r^k = (H,\ -A^T)\begin{pmatrix}x^k \\ v^k\end{pmatrix} + b$

如果 $\|r^k\|^2 < \varepsilon$，停止，否则按下述步骤计算 $u^{k+1} = (x^{k+1}, v^{k+1})$

$$d^k = \|r^k\|^2 / 4\left\|\begin{pmatrix} H \\ -A \end{pmatrix} r^k\right\|^2$$

$$\bar{x}^{k+1} = x^k - 2d^k H r^k$$

$$\bar{v}^{k+1} = v^k + 2d^k A r^k$$

$$x^{k+1} = (I - A^T(AA^T)^{-1}A)\bar{x}^{k+1} + A^T(AA^T)^{-1}P$$

$$v^{k+1} = \bar{v}^{k+1}$$

3) 令 $k := k+1$，返回 2)。

3 数值例子

求解如下二次规划问题[1]

$$\min f(x) \triangleq x_1^2 + 2x_2^2 + x_3^2 + x_3 - 2x_1x_2$$

$$s.t. \quad x_1 + x_2 + x_3 = 4$$

$$2x_1 - x_2 + x_3 = 2$$

这里

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

这个问题的最优解是 $x^* = \left(\frac{21}{11}, \frac{43}{22}, \frac{3}{22}\right)$.

采用本文中二次规划迭代解法，结果如下

EPS＝0. 0000000001

THE ITERATIVE POINT OF BEGINNING

U_0＝(－8. 9000000 －5. 2000000 7. 3200000 －9. 5000000 0. 5000000)

THE OPTIMAL SOLUTION：U. ＝(1. 9090930 1. 9545470 0. 1363602 2. 6363580 －1. 3636340)

Q(U.)＝0. 00000000009

其中 EPS ＝ ε 是迭代精度，$Q(u^k) <$ EPS 时终止程序。U_0 是迭代初始点，U_* 是计算最优解，U_* 的前 3 个分量是问题的计算最优解，与 x^* 是完全相符的。U_* 后 2 个分量是最优 Lagrange 乘子值，数值计算结果表明算法是完全可靠的。

参考文献

1 陈宝林. 最优化理论与算法. 北京：清华大学出版社，1989

2 Pardalos P. , Rosen J. Constrained Global Optimization：Algorithms and Applications. Lecture Notes in Computer Science No. 268 (Springer, Berlin, 1988)

3 何炳生. 一类求解凸规划的鞍点法. 计算数学，1991，(2)

4 赵凤治. 几个最优化数值方法新成果. 中国电机工程学报，1990，10增刊：35～40