第 26卷 第 4期 电子科技大学学报 Vol. 26 No. 4
1997年 8月 Journal of UEST of China Aug. 1997

一般约束优化问题的摄动梯度投影法

陈华富*
(电子科技大学应用数学系 成都 610054)

【摘要】 利用梯度投影法与罚函数技巧,将带等式和不等式约束优化问题化成一个无约束问题,提出了求解不等式、等式约束优化问题的摄动梯度投影算法。考虑到计算的误差因素,在搜索方向上进行摄动,得到一个方向不精确的梯度投影法。参数 W_k 取不同的数还可以得到一类梯度投影法。从而保证了在实际应用中更容易实现,在较弱的条件下,证明了该算法的全局收敛性。

关 键 词 不等式和等式约束; 摄动梯度投影; 罚函数; 全局收敛性
中图分类号 O212.2

从 1960年 Rosen[1]提出梯度投影法以来,已有许多有效的算法,但由于计算的不精确性,人们更注重对不精确的方向和步长进行研究,对于步长的不精确性研究已取得了较好的结果,但方向的不精确性研究更为困难和重要。文献[2]提出了对只含不等式约束的摄动梯度投影法,进行了研究。文献[3]在文献[2]的基础上削弱了条件,扩大了它的应用范围,但也只能解决不等式约束问题。最近几年对于等式约束的研究取得了重大成就[4,5],从而为一般的约束问题的研究提供了基础。

本文提出了带不等式、等式约束的摄动梯度投影算法,解决了一般约束问题的摄动梯度投影算法,同时证明了全局收敛性。

1 记号,假设

本文考虑 (NP): $\min\limits_{x\in R} f(x)$

其中 $R=\{x\in E^n \mid h_i(x)\leqslant 0\ i\in L_1, h_i(x)=0\ i\in L_2\}$, $L_1=\{1,2,\cdots,m\}$, $L_2=\{m+1,m+2,\cdots,m+p\}$, $J_0(x)=\{h_j(x)=0\ j\in L_1\}$, $J_W(x)=\{h_j(x)\geqslant -W\ j\in L_1, |h_j(x)|<W\ j\in L_2\}$。设 $J\subset\{1,2,\cdots,m+p\}$,若 $\mathrm{rank}\{\nabla h_i(x)\ i\in J\}=|J|$, $x\in R$,记 $g(x)=-\nabla f(x)$, $N_J(x)=\{N_i(X)\ i\in J\}$, $N_i(x)=\dfrac{\nabla h_i(x)}{\|\nabla h_i(x)\|}\ i\in J$, $A_J=\{A_i(x)\ i\in J\}$, $(U_i(x)\ i\in J)^T=(A_J(x)^T A_J(x))^{-1}A_J(x)g(x)$, $A_i(x)=N_i(x)+W g(x)$, $A_J(x)$为 $N_J(x)$之摄动,其中 $W\geqslant 0$,满足 $\mathrm{rank}A_J(x)=|J|$,相应令

$$\hat{P}_i(x)=I-N_i(x)(N_i(x)^T N_i(x))^{-1}N_i(x)^T$$

$$P_i(x)=I-A_i(x)(A_i(x)^T A_i(x))^{-1}A_i(x)^T$$

$$B_i(x)=(A_i(x)^T A_i(x))^{-1}A_i(x),\ \hat{U}_i(x)=(N_i(x)^T N_i(x))^{-1}N_i(x)$$

$$U_i(x)=(A_i(x)^T A_i(x))^{-1}A_i(x)^T g(x)=-B_i(x)\nabla f(x)$$

对于实数 a,记 $a_+=\max\{0,a\}$。本算法在 x 点迭代方向

$$d_J(x)=-P_J(x)\nabla f(x)-B_J(x)V_J(x) \tag{1}$$

⒇ 1996年 10月 5日收稿
* 男 28岁 硕士 助教

定义罚函数

$$G_C(x) = f(x) + C\Big(\sum_{j\in L_1} h_j(x)_+ + \sum_{j\in L_2} |h_j(x)|\Big) \tag{2}$$

$$V_j(x) = \begin{cases} U_j(x) & U_j(x)\geqslant 0 \quad j\in J\cap L_1 \\ -h_j(x) & U_j(x) < 0 \quad j\in J\cap L_1 \\ h_j(x) & \quad j\in J\cap L_2 \end{cases} \tag{3}$$

本文假设 $\begin{cases} (H_1) \quad f(x), h_i(x)\in C^1 \qquad i\in L_1\cup L_2 \\ (H_2) \quad \nabla h_i(x) \quad i\in J_0\cup L_1 \qquad \text{线性无关。} \end{cases}$

2 算 法

记 $\bar{C}_j(x) = \max\{U_j(x)_+, j\in J\cup L_2\}$,任意取 $x_1\in R$, $k := 1$,常数 $C_0 > 0$, $X > 0$

1) 计算 $J_k = J_{W_k}(x_k)$;

2) 若 $|N_{J_k}(x_k)^T N_{J_k}(x_k)| \leqslant W_k$ 或 $W_k^T U_J(x) \leqslant 1$,则令 $W_k := \frac{W_k}{2}$,转 1),否则转 3);

3) 计算 $U_{J_k}(x_k)$, $P_{J_k}(x_k)$;

4) 若 $P_{J_k}\nabla f(x_k) = 0$,则 x_k 是 $K-T$ 点,停,否则转 5);

5) 若 $\bar{C}_k > C_{k-1}$,令 $C_k = \max\{\bar{C}_k, C_{k-1} + X\}$,否则,令 $C_k := C_{k-1}$;

6) 求 λ_k,λ_k 满足下面的条件

$$G_{C_k}(x_k + \lambda_k d^k) \leqslant G_{C_k}(x_k) \qquad G_{C_k}(x_k + \lambda_k d^k) \leqslant \min G_{C_k}(x_k + \lambda d^k) + X_k$$

其中 $0\leqslant \lambda \leqslant 1$, $X_k \geqslant 0$, $X_k \to 0$, $k\to\infty$。

7) $x_{k+1} = x_k + \lambda_k d^k$, $W_{k+1} := W_k$, $k := k+1$,转 1)。

3 性 质

引理 1 若 $\mathrm{rank} N_J(x) = |J|$, $W^T \hat{U}_J(x) \neq -1$,则必有 $\mathrm{rank}\, A_J(x) = |J|$。

证明 若存在向量 $T \neq 0$,使得 $A_J(x) T_J = 0$,则有 $\sum_{i\in J} T_i N_i(x) = (-\sum_{i\in J} T_i W) g(x)$,又因 $\mathrm{rank} N_J(x) = |J|$,故 $\sum_{i\in J} T_i W \neq 0$,得 $g(x) = \sum_{i\in J} \frac{T_i}{-T_J^T W} N_i(x)$,即知 $\hat{P}_J(x) g(x) = 0$ 又有 $g(x) = \sum_{i\in J} \hat{U}_i(x) N_i(x)$,所以 $\hat{U}_i(x) = \frac{T_i}{-T_J^T W}$ $i\in J$,即 $\hat{U}_J(x) = -\frac{1}{T_J^T W} T_J$,因此 $W^T \hat{U}_J(x) = -1$ 矛盾。

引理 2 若 $J\supseteq J_0(x)$, $W^T U_i(x) \leqslant 1$,当 $P_i(x) g(x) = 0$ 且 $U_i(x) \geqslant 0$, $U_i(x) h_i(x) = 0$, $i\in J\cup L_1$ 时,则 x 是 (NP) 的 $K-T$ 点。

证明 由 $P_J(x) g(x) = 0$ 知 $g(x) = \sum_{i\in J} U_i(x) A_i(x)$,把 $A_i(x) = N_i(x) + W g(x)$ 代入得

$$(1 - W^T U_J(x)) g(x) = \sum_{i\in J} U_i N_i(x)$$

若 $W^T U_J(x) = 0$,由 $\mathrm{rank} N_J(x) = |J|$,知 $U_J(x) = 0$,因此得 $g(x) = 0$

若 $W^T \hat{U}_J(x) < 1$,则 $g(x) = \sum_{i\in J} \frac{U_i(x)}{1 - W^T U_J(x)} N_i(x)$ 必有 $\hat{P}_J(x) g(x) = 0$,又有 $g(x) = \sum_{i\in J} \hat{U}_i(x) N_i(x)$,

推出 $\hat{U}_i(x) = \frac{U_i(x)}{1 - W^T U_J(x)}$, $i\in J$,由引理条件得 $\hat{U}_J(x) \geqslant 0$ 且 $\hat{U}_i(x) h_i(x) = 0$ $i\in J\cup L_1$,即 x

或满足 $g(x)=0$,或满足 $g(x)=\sum_{i\in J}\hat{U}_i(x)N_i(x)$,且 $U_i(x)\geq 0,\hat{U}_i(x)\geq 0,\hat{U}_i(x)h_i(x)=0,i\in J\cup L_1$,因此,$x$ 是 (NP) 的 $K-T$ 点。

引理 3　$G_C(x)$ 在点 x 处方向 d 的方向导数

$$DG_C(x;d)=\nabla f(x)^T d+C\sum_{j\in L_1^+\cup J}\nabla h_j(x)^T d+C\sum_{j\in L_1^0\cup J}(\nabla h_j(x)^T d)_+ + C\sum_{j\in L_1^-\cup J}\nabla h_j(x)^T d+C\sum_{j\in L_2^+\cup J}\nabla h_j(x)^T d+C\sum_{j\in L_2^0\cup J}\nabla h_j(x)^T d-\sum_{j\in L_2^-\cup J}\nabla h_j(x)^T d$$

其中 $L_i^+=\{j\mid j\in L_i\quad h_j(x)>0\},L_i^0=\{j\mid j\in L_i\quad h_j(x)=0\},L_i^-=\{j\mid j\in L_i\quad h_j(x)<0\}\quad i=1,2$。

定理 1　若 $W_k U(x_k)\leq 1,DG_{C_k}(x_k;d^k)=0$,则 x_k 是 $x-T$ 点。

证明　由引理 3有

$$DG_{C_k}(x_k;d^k)=-\|P_k\nabla f(x_k)\|^2+\sum\nolimits_1+\sum\nolimits_2+\sum\nolimits_3+\sum\nolimits_4+\sum\nolimits_5 \qquad (4)$$

其中 $\sum_1=\sum_{\substack{j\in L_1\cup J\\ U_j(x)\leq 0}}-(U_j(x))^2$,$\sum_2=\sum_{\substack{j\in L_1\cup J\\ U_j(x)>0}}U_j(x)h_j(x)$,$\sum_3=\sum_{j\in L_2^+\cup J}(U_j^k-C_k)h_j(x_k)$,$\sum_4=0$,$\sum_5=\sum_{j\in L_2^-\cup J}(U_j^k+C_k)h_j(x_k)$。

由 C_k 和 $\bar{C}_k$ 定义可知:

$\forall j\in L_2^+\cup J$ 有 $C_k\geq U_j^k+C_0,U_j^k-C_k\leq -C_0<0$;$j\in L_2^-\cup J$ 有 $C_k\geq -U_j^k+C_0$,即 $U_j^k+C_k\geq C_0>0$。

由此可知 $\sum_i\leq 0(i=1,2,\cdots,5)$,则 $DG_{C_k}(x_k;d^k)\leq 0$。若 $DG_C(x;d)=0$,由式(4)则 $\sum_i=0$ $i=1,2,\cdots,5$,$\|P_k\nabla f(x_k)\|^2=0$,即得 $P_k\nabla f(x_k)=0$,且 $\sum_{\substack{j\in L_1\cup J\\ U_j(x)>0}}U_j(x)h_j(x)=0,U_j(x)\geq 0\quad j\in L_1\cup J$。又由 $W_k U(x_k)\leq 1$ 及引理知 x 是 (NP) 的 $K-T$ 点。

4　算法的全局收敛性

引理 4　在有聚点情况下,存在正整数 K_0,使得当 $k\geq K_0$ 时,$C_k=C_{k_0}\triangleq C$。

证明　若命题不成立,由算法中 C_k 的定义知存在无穷序列 K',使 $C_k=\max\{\bar{C}_k,C_{k-1}+X\}$,$\bar{C}_k>C_{k-1}$ 对任意 $k\in K'$ 成立,由于 $X>0$ 有 $C_k\geq C_{k-1}+X,\forall k\in K'$,因此 $C_k\to+\infty,\bar{C}_k\to+\infty,k\in K'$。由于 x_k 包含在一紧集中,故存在 $K\subset K'$ 使得 $\{x_k\}_K$ 收敛,设 x^* 为极限点,令 $U^*=\lim_{k\in K,k\to\infty}B(X_k)g(x_k)$,显然,当 k 沿 K 趋于无限时,$U_k\to U^*$,从而有当 $k\in K$,k 充分大时,$\max\{|U_i^k|\quad i\in J\cup L_2\}+\bar{C}_0$ 将不会趋于无穷,这与 $\bar{C}_k\to+\infty,k\in K\subset K'$ 矛盾。

定理 2　算法或有限步终止于 (NP) 的 $K-T$ 点,或得到一无限点列,若该点包含在一紧集中,则其任一聚点为 (NP) 的 $K-T$ 点。

当算法有限步终止于 x_k 时,由算法及定理 1知 x_k 为 (NP) 的 $K-T$ 点。

当算法产生一无限点列,且 x^* 为一聚点,将证明 x^* 为 (NP) 的 $K-T$ 点。

不妨设 $\{x_k\}_K\to x^*$ 具有如下性质:

1) $\forall k\in K,J_0(x_k)$ 不变化,记为 J^*;

2) $k\geq k_0$ 即 $\forall k\in K$,k 充分大有 $C_k\triangleq C_0$,下面完成定理证明。

若 x^* 不为 (NP) 的 $K-T$ 点,令 $d^*=\lim_{k\in K,k\to\infty}d_k$,显然 $J^*\subset J_0$,由定理 2知 $DG_C(x^*;d^*)<0$,由 $G_C(x)$ 的连续性知存在 λ^*,$0<\lambda^*\leq W$ 有

$$G_c(x^* + \lambda^* d^*) - G_c(x^*) < \frac{1}{2}\lambda^* DG_c(x^*; d^*)$$

由 $d_k \to d^*$，当 $k \in K$，$k \to \infty$，故存在 $k_1 \in K$，当 $k \geqslant k_1$，$k \in K$ 时，有

$$G_c(x_k + \lambda^* d^*) - G_c(x_k) < \frac{1}{4}\lambda^* DG_c(x^*; d^*)$$

由 $\min\limits_{0<\lambda<W} G_c(x_k + \lambda d^k) \leqslant G_c(x_k + \lambda^* d^*)$，$G_c(x_k + \lambda^k d) - G_c(x_k) \leqslant G_c(x_k + \lambda^* d^k) + X_k - G_c(x_k)$，$X_k \to 0$，存在 k_2，$k_2 > k_1$，使得 $k \geqslant k_2$，$k \in K$ 有

$$G_c(x_k + \lambda^k d^k) - G_c(x_k) \leqslant \frac{\lambda^*}{5} DG_c(x^*; d^*) \tag{5}$$

由 $0 < \lambda^k < W$ 知，序列 $\{x_k + \lambda^k d^k, k \in K\}$ 存在收敛子序列 $\{x_k + \lambda^k d^k, k \in K\}$，其极限点记为 $\bar{x}$，且不妨设 $k \geqslant k_2$，$\forall k \in K$ 在式(5)中，设 $k \to \infty$，则

$$G_c(\bar{x}) - G_c(x^*) \leqslant \frac{\lambda^*}{5} DG_c(x^*; d^*) < 0 \tag{6}$$

但 $G_c(x_k)$ 是单调下降序列，故有 $G_c(\bar{x}) = G_c(x^*)$，与式(6)矛盾。因此有 $DG_c(x^*; d^*) = 0$

再由 (H_1) 及 $W(x_k)$ 的选法知 $\{U_J(x_k)\}_K \to U_J(x^*)$， $\{d^k\}_K \to d^*$，$W(x^*) U_J(x^*) \leqslant 1$，所以 x^* 为 (NP) 的 $K-T$ 点。 证毕

参 考 文 献

1 Rosen J B. The gradient profection method for nonlinear programming, part 1, linear constraint. J SIAM Appl, 1960, 8: 181~217

2 堵 枉. 非线性约束条件下的梯度投影方法, 应用数学学报, 1985, 8(1): 7~16

3 施保昌. 一族非线性约束条件下的摄动梯度投影法, 应用数学学报, 1989, 12(2): 190~195

4 赖炎连, 高自友, 贺国平. 非线性最优化的广义梯度投影法. 中国科学(A辑), 1992, 9: 916~924

5 何光宗, 陈华富. 初始点任意优化问题的广义梯度投影法. 电子科技大学学报, 1996, 25(3): 330~334

A Perturbed Gradient Projection Method for General Constrained Optimization Problems

Chen Huafu

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

Abstract In this paper, the gradient projection and penalty function are used to make optimization problems for inequality and equality constraints into optimization problems without constraints. An algorithm of perturbed gradient projection for inequality and equality constrained problems is given. Considering the error of calculating and perturbing in search direction, a gradient projection method for inexact search direction is produced, which can get a sort of gradient projection algorithm when parameter W_k is differently chosed. The algorithm is globally convergent under very weak conditions.

Key words inequality and equality constraints; perturbed gradient projection; penalty function; globally convergent

编辑 徐培红