第 26 卷　第 6 期　　电 子 科 技 大 学 学 报　　Vol. 26　No. 6
1997 年 12 月　　Journal of UEST of China　　Dec. 1997

# 初始点任意的摄动梯度投影法

陈华富 *
（电子科技大学应用数学系　成都　610054）

【摘要】 利用梯度投影与罚函数相结合的技巧，将带不等式和等式约束的优化问题化成一个无约束问题，提出了初始点可任意的求解不等式、等式约束优化问题的摄动梯度投影算法；参数 $\delta_k$ 取不同的数还可以得到一类梯度投影算法，从而得出了在搜索方向和步长不精确条件下的梯度投影法，保证了在实际应用中更容易实现；在较弱条件下，证明了该算法的全局收敛性。

关 键 词　不等式、等式约束；　摄动梯度投影；　罚函数；　初始点任意

中图分类号　O212.2

从 1960 年 Rosen[1] 提出梯度投影法以来，已取得了许多有效的算法，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。对于步长不精确性的研究已取得较好的结果，但方向的不精确性研究更为重要，文献[2, 3] 提出了对含不等式约束的摄动梯度投影法，开始了这方面的研究；文献[4 ~6] 利用不同的方法提出了初始点可任意的梯度投影算法，对初始的任意性进行了研究。本文提出了初始点任意的带不等式、等式约束的摄动梯度投影算法，解决了初始点的任意性、方向和步长的不精确性。

## 1　记号，假设

考虑

$$(\mathrm{NP}): \min_{x\in R} f(x)$$

其中　$R=\{x\in E^n \mid h_i(x)\leqslant 0,\ i\in L_1,\ h_i(x)=0,\ i\in L_2\}$，$L_1=\{1, 2, \cdots, m\}$，$L_2=\{m+1, m+2, \cdots, m+p\}$，$J_\delta(x)=\{h_i(x)\geqslant -\delta, i\in L_1, |h_i(x)|< \delta, i\in L_2\}$。

设 $J=\{1, 2, \cdots, m+p\}$，若 rank$\{\nabla h_i(x)\quad i\in J\}$，$=|J|$，$x\in R$。记 $g(x)=-\Delta f(x)$，$N_J(x)=\{N_i(x), i\in J\}$，$N_i(x)=\Delta h_i(x)/\|\nabla h_i(x)\|$，$i\in J$，$A_J(x)=\{A_i(x), i\in J\}$，$A_i(x)=N_i(x)+\delta g(x)$，$(u_i(x), i\in J)^T=(A_J(x)^T A_J(x))^{-1} A_J(x) g(x)$。$A_J(x)$ 为 $N_J(x)$ 的摄动，其中 $\delta\geqslant 0$，满足 rank $A_J(x)=|J|$，相应令 $\hat{p_i}(x)=I-N_i(x)(N_i(x)^T N_i(x))^{-1} N_i(x)^T$，$p_i(x)=I-A_i(x)(A_i(x)^T A_i(x))^{-1} A_i(x)^T$，$B_i(x)=(A_i(x)^T A_i(x))^{-1} A_i(x)$，$\hat{u_i}=(N_i(x)^T N_i(x))^{-1} N_i(x)$，$u_i(x)=(A_i(x)^T A_i(x))^{-1} A_i(x)^T g(x)=-B_i(x)\nabla f(x)$，对于实数 $a$，记 $a_+=\max\{0, a\}$，$\rho(x)=\psi(x)+\varphi(x)+\max\{(-u_j(x))_+, |u_j(x) g_j(x)|\quad j\in L_1\}$，其中 $\psi(x)=\max\{h_i(x)_+, j\in L_1\}$，$\varphi(x)=\{|h_i(x)|, j\in L_2\}$。

本算法在 $x$ 点的可行方向设为

---

1997 年 4 月 8 日收稿

* 男　29 岁　硕士　讲师

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

其中 $V_J(x)=(v_j(x),\ j\in J)$。

$$v_j=\begin{cases}(-u_j(x))_+ + h_j(x) & j\in J\cap L_1 \text{ 且 } h(x)>0\\ (-u_j(x))_+ \ u_j(x)h_j(x) & j\in J\cap L_1 \text{ 且 } h(x)\leqslant 0\\ h_j(x) & j\in J\cap L_2\end{cases} \tag{2}$$

99

定义罚函数

$$G_c(x)=f(x)+C\sum_{j\in L_1}h_i(x)_+ + \sum_{j\in L_2}|h_i(x)| \tag{3}$$

假设$(H_1)$: $f(x)$, $h_i(x)\ i\in L_1\cup L_2\in C^1$; $(H_2)$: $\{\nabla h_i(x),\ i\in J_0\cup L_2\}$线性无关。

其中, $J_0=\{i|\ h_i(x)=0,\ i\in L_1|\}$。

## 2 算 法

常数 $C_0>0$, $\delta_0>0$, $\varepsilon>0$ 任取 $x\in E^n$, $\overline{C}_j(x)=\max\{u_j(x)_+,\ j\in J_0\cup L_2\}$。

1）计算 $J_k=J_{\delta_k}(x_k)$;

2）若$|N_{J_k}(x_k)^T N_{J_k}(x_k)_1|\leqslant \delta_k$ 或 $\theta^T U_J(x)>1$, 则令 $\delta_k=\delta_k/2$ 转 1), 否则转 3);

3）计算 $U_{J_k}(x_k)$, $P_{J_k}(x_k)$, $\rho_k(x_k)$;

4）若 $P_k\nabla f(x_k)=0$, $\rho_k(x_k)=0$, 则 $x_k$ 是$K-T$ 点, 否则转 5);

5）若 $\overline{C}_k>C_{K-1}$, 令 $C_k=\max(\overline{C}_k,\ C_{k-1}+\varepsilon)$; 否则令 $C_k=C_{K-1}$;

6）求 $\lambda_k$: 使得满足不等式 $G_c(x_k+\lambda_k d^k)\leqslant G_c(x_k)$和 $G_c(x_k+\lambda_k d^k)\leqslant\min G_c(x_k+\lambda d^k)+\varepsilon_k$。其中, $0\leqslant\lambda\leqslant 1$, $\varepsilon_k\geqslant 0$, 当 $k\to\infty$时; $\varepsilon_k\to 0$;

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

## 3 性 质

**引理 1** 若 rank $N_J(x)=|J|$, $\theta^T U_J(x)\neq -1$, 则必有 rank $A_J(x)=|J|$。

证明 若存在向量 $X_J\neq 0$, 使得 $A_J(x)X_J=0$, 则有$\sum_{i\in J}\alpha_i N_i(x)=(\sum_{i\in J}\alpha_i\theta)g(x)$, 又因为 rank $N_J(x)=|J|$, 故$\sum_{i\in J}a_i\theta\neq 0$, 可得 $g(x)=\sum_{i\in J}\frac{\alpha_i}{-\alpha_J^T\theta}N_i(x)$, 于是有 $P_J(x)g(x)=0$。由于 $g(x)=\sum_{i\in J}\hat{u_i}(x)N_i(x)$, 所以 $\hat{u_i}(x)=\frac{\alpha_i}{-\alpha_J^T\theta}$ $i\in J$, 即 $U_J(x)=\frac{\alpha_J}{-\alpha_J^T\theta}$。因此, $\theta^T U(x)=-1$ 矛盾。

**引理 2** 若 $J\supseteq J_0(x)$, $\theta^T U_J(x)=-1$, $\rho_k(x_k)=0$, $P_J(x)g(x)=0$, 则 $x$ 是(NP)的 $K-T$ 点。

证明 由$\rho(x)=\psi(x)+\varphi(x)+\max\{(-u_j(x))_+,\ |u_j(x)h_j(x)|\ j\in L_1\}$, $\rho(x)=0$, 以及 $\psi(x)$和 $\varphi(x)$定义得 $h_i(x)\leqslant 0$, $i\in L_1$; $h_i(x)=0$, $i\in L_2$, $u_j(x)\geqslant 0$, $u_j(x)h_j(x)=0$, $j\in L_1$, 即 $x$ 是可行点。由 $P_J(x)g(x)=0$, 知 $g(x)=\sum_{i\in J}u_i(x)A_i(x)$, 把 $A_i(x)=N_i(x)+\theta g(x)$ 代入可得$(1-\theta^T)U_J(x)g(x)=\sum_{i\in J}u_iN_i(x)$。

若 $\theta^T U_J(x)=0$, 由 rank $N_J(x)=|J|$, 知 $U_J(x)=0$, 得 $g(x)=0$。

若 $\theta^T U_J(x)<1$, 则 $g(x)=\sum_{i\in J}\frac{u_i(x)}{1-\theta^T U_J(x)}N_i(x)$, 因此必有 $P_Jg(x)=0$。又由于 $g(x)=$

$\sum_{i\in J}\hat{u_i}(x)N_i(x)$, 推出 $\hat{u_i}(x)=\dfrac{u_i(x)}{1-\delta^T U_J(x)}\quad i\in J$。

由引理条件得 $u_j(x)h_j(x)=0$, $u_j(x)\geqslant 0\ j\in L_1$, 所以 $x$ 是(NP)的 $K-T$ 点。

**引理 3**　若 $i\in L_1$, 则 $\|P_{L/(i)}g(x)\|^2=\|P_Lg(x)\|^2+U_i^2\|P_{L/(i)}\nabla h_i(x)\|^2$

**引理 4**　$G_c(x)$在点 $x$ 处沿方向$d$ 的方向导数

$$DG_c(x;d)=\nabla f(x)^Td+C\sum_{j\in L_1^+}\nabla h_i(x)^Td+C\sum_{j\in L_1^0}(\nabla h_i(x)^Td)_++C\sum_{j\in L_1^+}\nabla h_i(x)^Td+$$
$$C\sum_{j\in L_2^+}|\nabla h_i(x)^Td|-C\sum_{j\in L_2^-}\nabla h_i(x)^Td$$

其中

$$L_i^+=L_i^+(x)=\{i\mid i\in L_i,\ h_i(x)>0\}$$
$$L_i^0=L_i^0(x)=\{i\mid i\in L_i,\ h_i(x)=0\}$$
$$L_i^-=L_i^-(x)=\{i\mid i\in L_i,\ h_i(x)<0\}$$

定理 1　若 $\delta^TU(x)\leqslant 1$, $DG_c(x;d)=0$, $\rho(x)=0$, 则 $x$ 是(NP)的 $K-T$ 点。

证明　由式(1)~(3)及引理 4 有

$$A(x)^Td(x)=-A(x)^TP(x)\nabla f(x)-\rho(x)A(x)^TB(x)^TV(x)=-\rho(x)V(x)\tag{4}$$
$$\nabla f(x)^Td=-\|P\nabla f(x)\|^2+\sum_{i\in L_2}\rho u_jv_j\tag{5}$$

于是

$$\begin{aligned}DG_c(x;d)=&-\|P\nabla f(x)\|^2+\sum_{i\in L_2}\rho u_jv_j-\rho C\{\sum_{j\in L_1^+}v_j-\sum_{j\in L_1^0}(-v_j)_+-\sum_{j\in L_1^-}v_j+\sum_{j\in L_2^+}-\sum_{i\in L_1^-}v_j\}=\\&-\|P\nabla f(x)\|^2-\rho\{\sum_{j\in L_1^+}(C-u_j)(h_j(x)+(-u_j)_+)+\\&\sum_{j\in L_1^0}(-u_j)_+(-u_j)+\sum_{j\in L_1^-}(-u_j)((-u_j)_++u_j(x)h_j(x))+\\&\sum_{j\in L_2^+}(C-u_j)h_j(x)+\sum_{j\in L_2^-}(C+u_j)(-h_j(x))\}=\\&-\|P\nabla f(x)\|^2+\rho\{\sum_{j\in L_1^+}(u_j-C)(h_j(x)_+(-u_j)_+)-\sum_{j\in L_1^0\cup L_1}u_{j+}^2+\\&\sum_{j\in L_1^-}(u_j)^2h_j(x)+\sum_{j\in L_2^+}(-C+u_j)h_j(x)+\sum_{j\in L_2^-}(C+u_j)h_j(x)\}\triangleq\\&-\|P\nabla f(x)\|^2+\rho(M_1+M_2+M_3+M_4+M_5)\end{aligned}$$

由 $C$ 和$\tau$ 定义可得

$\forall j\in L_i^+$ 有 $C\geqslant C_j+C_0$, $u_j-C\leqslant -C_0<0$; $j\in L_i^-$ 有 $C\geqslant -u_j+C_0$, 即 $u_j+C\geqslant C_0>0$, 因此 $\sum_i\leqslant 0$, $(i=1,2,3,4,5)$ 则 $DG_c(x;d)\leqslant 0$。由 $\rho\neq 0$, 可以证明 $DG_c(x;d)<0$。

若$DG_c(x;d)=0$, 则$P\nabla f(x)=0$, $\rho=0$。故 $\psi(x)=0$, $\varphi(x)=0$, $\max\{(-u_j(x))+,\ |u_j(x)h_j(x)|\quad j\in L_1\}=0$, $x$是可行点。于是有$u_j(x)\geqslant 0$, $u_j(x)h_j(x)=0$, $P(x)\nabla f(x)=0$, $j\in L_1$。

由引理 1 知 $x$ 是(NP)的 $K-T$ 点。

# 4 算法的收敛性

引理 5 在有聚点情况下，存在正整数 $K_0$，当 $k\geqslant k_0$ 时，$C_k\equiv C_{k0}\triangleq C_0$。

证明 假设命题不成立，由算法 $C_k$ 的调整方法知，存在无穷子序列 $K$，使 $C_k=\max\{\bar{C}_k, C_{k-1}+\varepsilon\}$，$\bar{C}>C_{k-1}$，$\forall k\in K$ 成立，由于 $\varepsilon>0$ 有 $C_k\geqslant C_{k-1}+\varepsilon$，$\forall k\in K'$，因此 $k\to+\infty$，有 $C_k\to\infty$，以及 $\bar{C}\to+\infty$，由于 $x_k$ 包含在一紧集中，故 $K\subset K'$，使得$\{x_k, K\}$收敛，设 $x^*$ 为极限点，且由于 $J_0(x_k)$选取的有限性，不妨设 $\forall k\in K$，$J_0(x_k)$均相同，令 $u^*=\lim\limits_{k\in k\ k\to\infty}B(x_k)g(x_k)$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，$\max\{|u_j^k|, j\in J_0(x_k)\}+\bar{C}$ 将不会趋于无穷，这与 $\bar{C}_k\to+\infty$，$k\in K\subset K'$矛盾。

设$\{x_k\}_K\to x^*$具有下面的性质：

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

2） $k\geqslant k_0$ 时，$\forall k\in K$ 有 $C_k\triangleq C_0$

记 $d^*=Pg(x^*)-\rho^*B(x^*)^Tv(x^*)$

$$\rho^*=\psi(x^*)+\varphi(x^*)+\max\{(-u_j(x^*))_+,\ |u_j(x^*)h_j(x^*)|\ j\in L_1\}$$

引理6 对 $J\subset J_0(x)$，记 $\bar{L}=J\cup L_2$，$\bar{\rho}=\psi(x)+\varphi(x)+\max\{(-u_j(x))_+, |u_j(x)h_j(x)|\ j\in\bar{J}\}$，$\bar{d}=P_Lg(x)-\bar{\rho}\bar{B}(x)V(x)$，$\bar{B}(x)=(A_L(x)^TA_L(x))^{-1}A_L(x)$，若 $DG_c(x, \bar{d})=0$，则 $x$ 为(NP)的 $K-T$ 点

证明 由定理1知$DG_c(x;\bar{d})\leqslant 0$，当$DG_c(x;\bar{d})=0$，则 $\|P_Lg(x)\|=0$，$\bar{\rho}=0$；由引理3知 $\|P_{L_2\cup J_0}g(x)\|=0$，则 $u_i=0$，$i\in J_0/\bar{J}$，因此，$\bar{\rho}=\rho$，故 $DG_c(x;\bar{d})=0$。再由引理2知 $x$ 为(NP)的 $K-T$ 点。下面完成定理证明。

如果 $x^*$不是(NP)的 $K-T$ 点，则令 $d^*=\lim\limits_{k\in K,\ k\to\infty}d_k$，显然有 $J^*\subset J_0(x^*)$。由引理6可以知道 $DG_C(x^*;d^*)<0$，则由 $G_c(x)$的连续性知存在 $\lambda^*$，$0<\lambda^*<\delta$，使得

$$G_c(x^*+\lambda^*d^*)-G_c(x^*)<\lambda^*DG_c(x^*;d^*)/2$$

由 $d_k\to d^*$，当 $k\in K$，$k\to+\infty$时，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
$$\min_{0<\lambda<\delta}G_c(x_k+\lambda d^k)\leqslant G_c(x_k+\lambda^*d^*)$$

$$G_C(x_k+\lambda^kd^k)-G_c(x_k)\leqslant C_c(x_k+\lambda^*d^*)+\varepsilon_k-G_c(x_k)$$

由于 $\varepsilon_k\to 0$，存在 $k_2\in K$，$k_2<k_1$，使得当 $k_1\geqslant k_2$，$k\in K$ 时有

$$G_c(x_k+\lambda^kd^k)-G_c(x_k)\leqslant\frac{\lambda^*}{5}DG_c(x^*;d^*)\tag{6}$$

由 $0<\lambda^k<\delta$知序列$\{x_k+\lambda^kd^k, k\in K\}$，存在一收敛子序列$\{x_k+\lambda^kd^k, k\in K''\}$其极限点记为 $\bar{x}$，设 $k\geqslant k_2$，$\forall k\in K''$，在式(8)中让 $k$ 沿$K''$趋于无穷得

$$G_c(\bar{x})-G_c(x^*)\leqslant\frac{\lambda^*}{5}DG_c(x^*;d^*)<0\tag{7}$$

但 $G_c(x_k)$是单调下降序列，故 $G_c(\bar{x})=G_c(x^*)$与式(7)矛盾，因此 $DG_c(x^*;d^*)=0$ 所以 $x^*$为(NP)的 $K-T$ 点。

## 参 考 文 献

1 Rosen J B. The gradient projection method for nonlinear programming, Part 1. Linear Constraint J SIAM APPL, 1960, 8 :181 ~ 217

2 堵 丁柱. 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. 一族非线约束条件下的摄动梯度投影法. 应用数数学报, 1989, 12(2) :190 ~ 195

4 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

5 赖炎连, 简金宝. 初始点任意的一个非线性优化问题的广义梯度投影法. 系统科学与数学, 1995, 15(4) :374 ~ 380

6 薛声家, 简金宝. 解一般约束最优化问题的一个初始点任意的梯度投影法. 暨南大学学报(自然科学版), 1994, 15(1) :19 ~ 23

7 陈华富, 钮海, 何光宗. 一般约束优化问题的广义梯度投影法. 电子科技大学学报, 1997, 26(增刊) :83 ~ 88

8 陈华富. 一般约束优化问题的摄动梯度投影法. 电子科技大学学报, 1997, 26(4) :445 ~ 448

# A Pertrubed Gradient Projection Method with Arbitrary Initial Point 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 optimination problem without constraints. The algorithm of perturbed gradient projcetion with aribitrary initial point for inequality and equality constrained problem is given, which can get a sort of gradient projection method in accurate search direction and step, when parameter $\alpha$ is different chosed. It is proved to be convenient in application. The algorithm is gobally convergent under very weak conditions.

**Key words**　inequality and equality constraints;　perturbed gradient projection;　arbitrary initial point

编辑　徐培红