第4卷 第2期
1984年4月

数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION

Vol.4 No.2
April 1984

解非线性约束拟凸规划的一个梯度投影法*

薛声家

（广西大学数学系）

目前国内外所流行的梯度投影法（包括Rosen的原有算法和一些修正算法）还存在以下几个问题：一、要增加Polak程序以保证算法的收敛性。二、在计算投影梯度时，每步一般要作两次投影。三、对于非线性约束问题，负梯度投影方向是不可行的，因此必须在此方向的基础上构造出能保证算法收敛的新可行下降方向。而目前为构造出这个新方向所作的计算都比较复杂。

1981年[5]提出了一个处理线性约束条件的梯度投影法，基本上解决了线性约束条件下的上述第一、二两个问题。

本文吸取[5]的思想，提出了一个处理非线性约束拟凸规划的算法。这算法免去了Polak程序。也不需要在每一步作两次梯度投影；同时给出一个构造新可行下降方向的简单方法。加上由于把目标函数转化为线性，故在每步中能免去计算量大的线搜索。所以，这是解非线性约束拟凸规划一个计算量小、简单易行的算法。

中国科学院应用数学研究所桂湘云和赖炎连两位同志对本文提出了很多宝贵意见，谨在此表示衷心的感谢。

§1 方法的叙述

由于带有非线性目标函数的规划问题总可用增加一个变量和一个约束条件的方法转化为目标函数是线性的情形，故本文只考虑如下的规划问题：

$$(P)\quad\begin{cases}\min\ c^Tx\\ g_j(x)\leqslant 0\quad (j=1,\cdots,m)\\ x\in E^n\end{cases}$$

对于问题（P），我们作如下假设：

H_1。$g_j(x)$ （$j=1,\cdots,m$）连续可微，拟凸，从而约束集合R为闭凸集。

H_2。存在x^0使得$g_j(x^0)<0$ （$j=1,\cdots,m$），并且集合$R(x^0)=\{x|x\in R,\ c^Tx\leqslant c^Tx^0\}$有界，故问题（$P$）的最优解存在。

H_3。$\forall x\in R(x^0)$，记$J_0(x)=\{j|g_j(x)=0,\ j=1,\cdots,m\}$，

设$\{\nabla g_j(x)\}_{j\in J_0(x)}$是线性无关的。

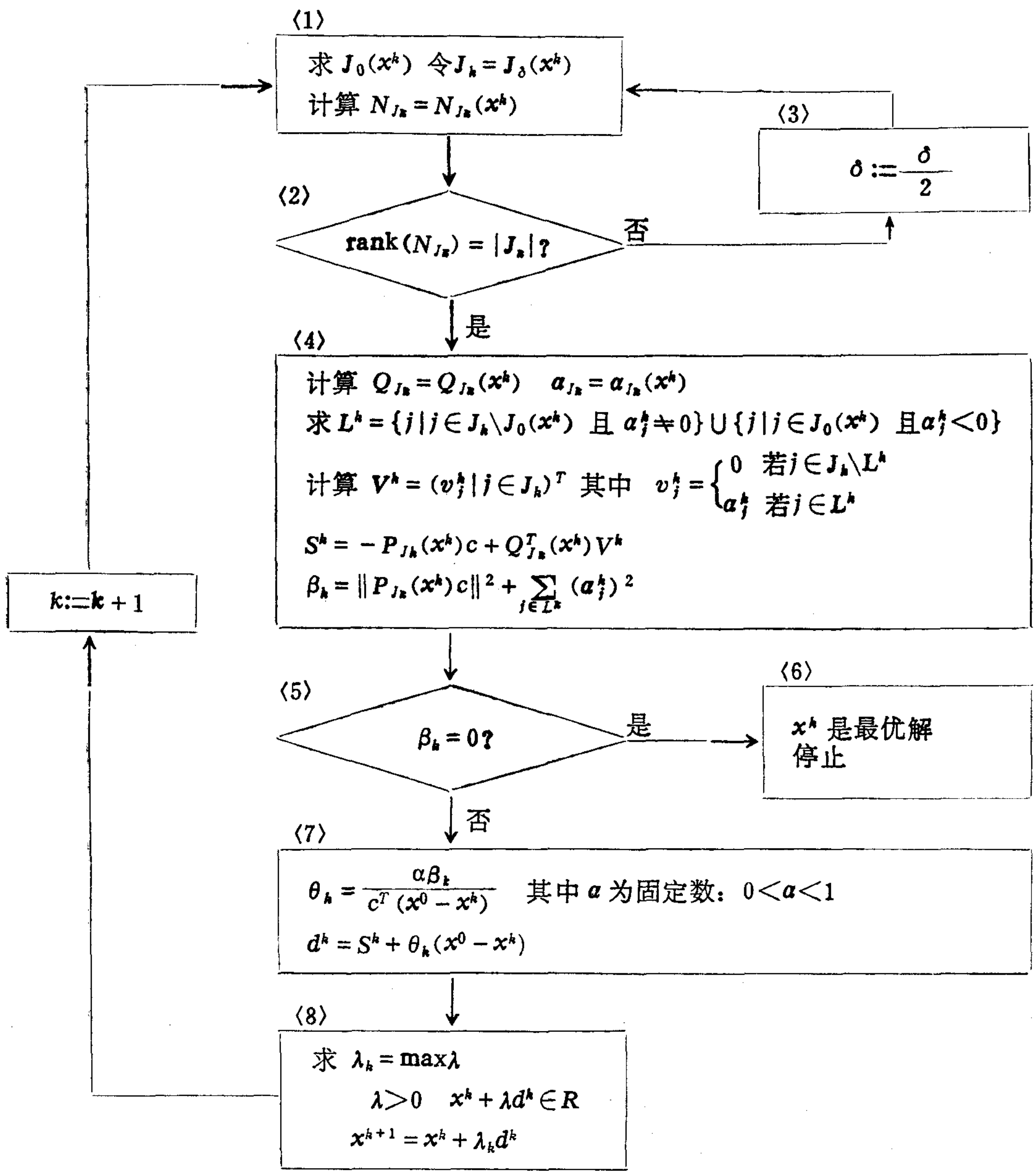
*1981年7月18日收到。

引理 1 在上述假设条件下，$\exists\delta_0>0$，使$\forall x\in R(x^0)$ 当$0\leqslant\delta\leqslant\delta_0$时，$\{\nabla g_j(x)\}_{j\in J_\delta(x)}$线性无关，其中 $J_{\delta(x)}=\{j\,|\,g_j(x)\geqslant-\delta,\ j=1,\cdots,m\}$。

证 可参看［4］§3 定理 1 的证明。

设 J 为指标集 $\{1,\cdots,m\}$ 的某子集，记 $n\times|J|$矩阵$N_J=N_J(x)=(\nabla g_j(x)\,|\,j\in J)$，其中$|J|$为 J 的元素个数。N_J的广义逆为$Q_J(x)=(N_J^TN_J)^{-1}N_J^T$。投影矩阵$P_J(x)=I-N_JQ_J$，其中 I 为$n\times n$单位阵。$a_J(x)=(a_j\,|\,j\in J)^T=Q_J(x)(-c)$ 为$|J|\times 1$ 矩阵。

算法的叙述



步骤 0，求 x^0 满足 $g_j(x^0)<0(j=1,\cdots,m)$

求 $\lambda_0=\max\limits_{\lambda>0}\lambda$

$x^0+\lambda(-c)\in R$

令 $x^1=x^0+\lambda_0(-c)$，取充分小的正数 δ，令 $k:=1$，进入框图计算。

下述各引理及定理说明了算法框图的功能。

引理 2　迭代由〈2〉进入〈3〉的次数是有限的。

证　由引理 1 即得。

引理 3　给定 $x\in R(x^0)$，若存在某指标集 J：$J_0(x)\subseteq J\subseteq\{1,\cdots,m\}$，使得按步骤〈4〉所求得之 $L=\phi$ 且 $P_J(x)c=0$，则 x 是问题（P）的最优解。

证　我们有 $-c=-P_J(x)c+\sum\limits_{j\in J}a_j\nabla g_j(x)$ 由假设可得 $-c=\sum\limits_{j\in J_0(x)}a_j\nabla g_j(x)$，其中 $a_j\geqslant0$，即 x 为问题(P)的 K——T 点，由于目标函数是线性而约束函数是拟凸的，故 x 为问题(P)的最优解。

定理1　x^k是问题（P）的最优解当且仅当 $\beta_k=0$。

证　充分性：由引理 3 即得。

必要性：我们证明若 $\beta_k\neq0$（此时 $\beta_k>0$），则d^k是目标函数在点 x^k 的可行下降方向，因而x^k不是最优解。

为此，只需证 $c^Td^k<0$ 和 $\forall j\in J_0(x^k)$ 有 $\nabla g_j(x^{k,T})d^k<0$。我们有 $c^Td^k=-c^TP_{J_k}(x^k)c+c^TQ^T_{J_k}(x^k)V^k+\theta_kc^T(x^0-x^k)=-\|P_{J_k}(x^k)c\|^2+(Q_{J_k}(x^k)c)^TV^k+a\beta_k$

$$=-\|P_{J_k}(x^k)c\|^2-\sum_{j\in L^k}(a_j^k)^2+a\beta_k=(a-1)\beta_k<0$$

任取 $j\in J_0(x^k)$，注意到 $\nabla g_j(x^k)^T(P_{J_k}(x^k)c)=0$ 和$\nabla g_j(x^k)^TQ^T_{J_k}(x^k)V^k=v_j^k$则有：

$$\nabla g_j(x^k)^Td^k=v_j^k+\frac{a\beta_k\nabla g_j(x^k)^T(x^0-x^k)}{c^T(x^0-x^k)}$$

若 $j\in L^k,v_j^k=a_j^k<0$（注意 $j\in J_0(x^k)$）；若 $j\bar{\in}L^k$，$v_j^k=0$；而 $\nabla g_j(x^k)^T(x^0-x^k)<0$，$c^T(x^0-x^k)>0$，$a\beta_k>0$，所以有$\nabla g_j(x^k)^Td^k<0$。证毕。

§2　算法的收敛性

引理4　若算法产生无穷点列 $\{x^k\}$，设$\{x^{\bar{k}}\}$为$\{x^k\}$的子列：$\{x^{\bar{k}}\}\to x^\infty$，则当$\bar{k}$充分大时 $J_{\bar{k}}\supseteq J_0(x^\infty)$。

证　由引理 1 和 2 可见当$\bar{k}\geqslant K_0$(K_0为某自然数）δ 的值将保持不变，记为δ_c。因此当 $\bar{k}\geqslant K_0,J_{\bar{k}}=J_{\delta_c}(x^{\bar{k}})=\{j|g_j(x^{\bar{k}})\geqslant-\delta_c,j=1,\cdots,m\}$。

另一方面，$\forall j\in J_0(x^\infty)$，存在自然数 K_j 使当 $\bar{k}\geqslant K_j$ 时，$|g_j(x^{\bar{k}})-g_j(x^\infty)|=|g_j(x^{\bar{k}})|<\frac{\delta_0}{2}$（$\delta_0$ 为引理 1 所提到的，显然$\frac{\delta_0}{2}<\delta_c$），从而$g_j(x^{\bar{k}})>-\frac{\delta_0}{2}>-\delta_c$。所以当 $\bar{k}\geqslant\max\{K_0,K_j|j\in J_0(x^\infty)\}$，$\forall j\in J_0(x^\infty)$ 均有 $g_j(x^{\bar{k}})>-\delta_c$，亦即 $j\in J_{\bar{k}}$。证毕。

定理2　（收敛性定理）若算法产生无穷点列$\{x^k\}$，则$\{x^k\}$的任一极限点都是问题（P）的最优解。

我们用反证法证明定理2，设$\{x^{\bar{k}}\}$为$\{x^k\}$的某子列：$\{x^{\bar{k}}\}\to x^\infty$，而$x^\infty$不是问题(P)的最优解。

由于$\{J_{\bar{k}}\}$，$\{J_0(x^{\bar{k}})\}$，$\{L^{\bar{k}}\}$都是有限集的子集所产生的无限序列，故必存在$\{\bar{k}\}$的子列$\{\bar{\bar{k}}\}$使得$J_{\bar{\bar{k}}}=J$，$J_0(x^{\bar{\bar{k}}})=J_0$，$L^{\bar{\bar{k}}}=L$，其中$J$，$J_0$，$L$都为固定指标集。当然有$J\supseteq J_0$，$J\supseteq L$且由引理4 $J\supseteq J_0(x^\infty)$。

记$\alpha_J^\infty=(\alpha_j^\infty \mid j\in J)^T=-Q_J(x^\infty)c$；$V_J^\infty=(v_j^\infty \mid j\in J)^T$，其中$v_j^\infty=\begin{cases}0 & 若 j\in J\setminus L\\ \alpha_j^\infty & 若 j\in L\end{cases}$

记$S^\infty=-P_J(x^\infty)c+Q_J^T(x^\infty)V_J^\infty$；$\beta_\infty=\|P_J(x^\infty)c\|^2+\sum_{j\in L}(\alpha_j^\infty)^2$，

$$\theta_\infty=\frac{\alpha\beta_\infty}{c^T(x^0-x^\infty)};\quad d^\infty=S^\infty+\theta_\infty(x^0-x^\infty)$$

易见当$\bar{\bar{k}}\to\infty$时有：

$$\alpha_J^{\bar{\bar{k}}}\to\alpha_J^\infty \qquad V^{\bar{\bar{k}}}\to V_J^\infty \qquad s^{\bar{\bar{k}}}\to s^\infty$$
$$\beta_{\bar{\bar{k}}}\to\beta_\infty \qquad \theta_{\bar{\bar{k}}}\to\theta_\infty \qquad d^{\bar{\bar{k}}}\to d^\infty$$

我们先证明下面几个引理：

引理5 $L\supseteq L^\infty$。

其中 $L=L^{\bar{\bar{k}}}=\{j\mid j\in J\setminus J_0 且\alpha_j^{\bar{\bar{k}}}\neq0\}\cup\{j\mid j\in J_0 且\alpha_j^{\bar{\bar{k}}}>0\}$

$L^\infty=\{j\mid j\in J\setminus J_0(x^\infty) 且 \alpha_j^\infty\neq0\}\cup\{j\mid j\in J_0(x^\infty) 且 \alpha_j^\infty<0\}$

证 设$j\in L^\infty$，则有两种情形发生：

1° $j\in J\setminus J_0(x^\infty)$ 且 $\alpha_j^\infty\neq0$

由$g_j(x^{\bar{\bar{k}}})\to g_j(x^\infty)\neq0$，$\alpha_j^{\bar{\bar{k}}}\to\alpha_j^\infty\neq0$ 得知当$\bar{\bar{k}}$充分大时$g_j(x^{\bar{\bar{k}}})\neq0$，$\alpha_j^{\bar{\bar{k}}}\neq0$，即$j\in J\setminus J_0(x^{\bar{\bar{k}}})$ 且$\alpha_j^{\bar{\bar{k}}}\neq0$因而$j\in L^{\bar{\bar{k}}}=L$。

2° $j\in J_0(x^\infty)$ 且$\alpha_j^\infty<0$。

这时当$\bar{\bar{k}}$充分大有$\alpha_j^{\bar{\bar{k}}}<0$，故不管$j$是否属于$J_0(x^{\bar{\bar{k}}})=J_0$，总有$j\in L^{\bar{\bar{k}}}=L$。

引理6 $\beta_\infty>0$，$\theta_\infty>0$

证 $\beta_\infty=\|P_J(x^\infty)c\|^2+\sum_{j\in L}(\alpha_j^\infty)^2\geq\|P_J(x^\infty)c\|^2+\sum_{j\in L^\infty}(\alpha_j^\infty)^2$，由于$x^\infty$非最优解，或$P_J(x^\infty)c\neq0$，或$L^\infty\neq\varnothing$，故上式右边必大于0。$\theta_\infty>0$是显然的。

引理7 $c^Td^\infty<0$

证 与定理1必要性第一部分的证明类似我们有$c^Td^\infty=(\alpha-1)\beta_\infty<0$

引理8 存在$\varepsilon>0$和$N>0$，使得当$\bar{\bar{k}}\geq N$时有$x^{\bar{\bar{k}}}+\lambda d^{\bar{\bar{k}}}\in R$ $\forall\lambda\in[0,\varepsilon]$

证 若$j\bar{\in}J_0(x^\infty)$ 则$g_j(x^\infty)<0$，这时由g_j的连续性和$d^{\bar{\bar{k}}}$的有界性易见$\exists\varepsilon_1>0$和$N_1>0$使得当$\bar{\bar{k}}\geq N_1$，$\forall\lambda\in[0,\varepsilon_1]$，$\forall j\bar{\in}J_0(x^\infty)$ 均有$g_j(x^{\bar{\bar{k}}}+\lambda d^{\bar{\bar{k}}})<0$。

下面讨论 $j\in J_0(x^\infty)\subseteq J$的情形。

易见 $\theta_\infty \nabla g_j(x^\infty)^T(x^0-x^\infty)<0$，故存在 $\gamma>0$ 使 $\forall j\in J_0(x^\infty)$ 有 $\theta_\infty \nabla g_j(x^\infty)^T(x^0-x^\infty)\leqslant -4\gamma<0$。所以当 $\bar{\bar{k}}$ 充分大，$\forall j\in J_0(x^\infty)$ 有 $\theta_{\bar{k}}\nabla g_j(x^{\bar{k}})^T(x^0-x^{\bar{k}})\leqslant -2\gamma<0$。此外

$-\nabla g_j(x^{\bar{k}})^T P_J(x^{\bar{k}})c=0$，

$$\nabla g_j(x^{\bar{k}})^T Q_J^T(x^{\bar{k}})V^{\bar{k}}=v_j^{\bar{k}}=\begin{cases}0 & \text{若 } j\overline{\in} L\\ a_j^{\bar{k}}<0 & \text{若 } j\in L.\end{cases}$$

所以 $\nabla g_j(x^{\bar{k}})^T(-P_J(x^{\bar{k}})c)+\nabla g_j(x^{\bar{k}})^T Q_J^T(x^{\bar{k}})V^{\bar{k}}+\theta_{\bar{k}}\nabla g_j(x^{\bar{k}})^T(x^0-x^{\bar{k}})$

$$\leqslant\theta_{\bar{k}}\nabla g_j(x^{\bar{k}})^T(x^0-x^{\bar{k}})\leqslant -2\gamma<0.$$

因此，存在 $\varepsilon_2>0$ 和 $N_2>0$，使得 $\forall\bar{\bar{k}}\geqslant N_2$，$\forall\lambda\in[0,\varepsilon_2]$ $\forall j\in J_0(x^\infty)$ 均有

$$\nabla g_j(x^{\bar{k}}+\lambda d^{\bar{k}})^T(-P_J(x^{\bar{k}})c)+\nabla g_j(x^{\bar{k}}+\lambda d^{\bar{k}})^T Q_J^T(x^{\bar{k}})V^{\bar{k}}+$$
$$\theta_{\bar{k}}\nabla g_j(x^{\bar{k}}+\lambda d^{\bar{k}})^T(x^0-x^{\bar{k}})\leqslant -\gamma<0$$

即
$$\nabla g_j(x^{\bar{k}}+\lambda d^{\bar{k}})^T d^{\bar{k}}\leqslant -\gamma<0$$

另一方面，使用中值定理，我们有

$$g_j(x^{\bar{k}}+\lambda d^{\bar{k}})=g_j(x^{\bar{k}})+\lambda\nabla g_j(x^{\bar{k}}+\lambda a_{j\bar{k}}d^{\bar{k}})^T d^{\bar{k}}$$
$$\leqslant\lambda\nabla g_j(x^{\bar{k}}+\lambda a_{j\bar{k}}d^{\bar{k}})^T d^{\bar{k}}$$

其中 $a_{j\bar{k}}\in[0,1]$，因此有 $g_j(x^{\bar{k}}+\lambda d^{\bar{k}})\leqslant -\lambda\gamma\leqslant 0$

令 $\varepsilon=\min\{\varepsilon_1,\varepsilon_2\}$，$N=\max\{N_1,N_2\}$ 则当 $\bar{\bar{k}}\geqslant N$，$\forall\lambda\in[0,\varepsilon]$，$\forall j=1,\cdots,m$ 都有 $g_j(x^{\bar{k}}+\lambda d^{\bar{k}})\leqslant 0$，即 $x^{\bar{k}}+\lambda d^{\bar{k}}\in R$。证毕。

引理 7 和引理 8 的结论与参考文献[10]中的引理10．2．6矛盾，收敛性定理因此得证

参　考　文　献

[1] 越民义、韩继业，一个新的既约梯度法及其收敛性，中国科学，1979年第4期，345—356.

[2] 赖炎连，非线性约束凸规划的一个解法及其收敛性，应用数学学报，1980年第4期，322—331.

[3] 章祥荪，改进的Rosen—Polak方法，应用数学学报，1979年第3期，257—267.

[4] 章祥荪，关于非线性约束条件下的Polak算法的一些讨论，应用数学学报1981年第1期1—13.

[5] 堵丁柱、孙捷，一个新的梯度投影方法，计算数学，1983年第4期，378—386.

[6] Rosen, J. B. The Gradient Projection Method for Nonlinear Programming. Part I, Linear Constraints, SIAMJ. *Applied Mathematics* 8 (1960) 181—217.

[7] ——. Part II, Nonlinear Constraints, SIAMJ. *Applied Mathematics* 9 (1961) 514—532.

[8] Polak, E., On the Convergence of Optimization Algorithms. Rev. Francaise, *Informat, Recherche Operationnelle* 3, 16, (1969), 17—34.

[9] Polak, E., Computational Methods in Optimization, Academic Press, New York, 1971.

[10] Bazaraa, M. S. and shetty, C. M., Nonlinear Programming, Theory and Algorithms, *John Wiley and Sons, Inc.*, 1979.

A Gradient Projection Method for Quasiconvex Programming with Nonlinear Constraints

Xue Sheng Jia

Abstract

This paper proposes a gradient projection algorithm to hanlde quasiconvex programming with nonlinear constraints. The algorithm not only has avoided Polak's perturbation procedure, but it needs only one gradient projection at each iteration instead of projecting two times as in most of the projection methods; and a simple method for constructing a new improved feasible direction is given. It is proved that the algorithm either terminates at an optimal solution after finitely many steps or it generates a sequence of feasible points whose every limit point is an optimal solution of the original problem.