数 学 年 刊
3 (5) 1982

关于二次规划局部最优解的充要条件

刘嘉荃　宋天泰　堵丁柱
(中国科学技术大学研究生院)

一、问 题 叙 述

考虑二次规划

$$\begin{cases}\min f(x)=q^Tx+\dfrac{1}{2}x^Tcx \\ a_i^Tx\leqslant b_i,\ i\in J=\{1,\ 2,\ \cdots,\ N\},\end{cases}\tag{1.1}$$

熟知 Kuhn-Tucker 条件是非线性规划局部最优解的必要条件. 具体到二次规划(1.1), 设 x_0 是局部最优解, 则必存在一组非负数 λ_i, $i\in J$, 使得

$$\begin{cases}q+cx_0+\sum\limits_{i\in J}\lambda_ia_i=0, \\ \lambda_i(a_i^Tx_0-b_i)=0,\ i\in J.\end{cases}\tag{1.2}$$

如果按 $a_i^Tx_0-b_i=0$ 是否成立, 可将标号集 J 分为二组: J_1 和 J_2

$$\begin{aligned}&J_1=\{i\in J\,|\,a_i^Tx_0-b_i=0\},\\ &J_2=J\setminus J_1,\end{aligned}\tag{1.3}$$

与 J_1 中标号相应的约束条件称为积极约束. 条件(1.2)又可改写成: 存在一组非负数 $\{\lambda_i|i\in J_1\}$, 使得

$$q+cx_0+\sum_{i\in J_1}\lambda_ia_i=0.\tag{1.4}$$

至于说局部最优解成立的充要条件, 对于一般非凸规划甚为困难. 对于二次规划(1.1)的局部最优解的充要条件, 文献[1, 2]曾作过讨论. 如果令

$$J^*=\{i\in J_1\,|\,\lambda_i>0\},\tag{1.5}$$

[1]认为满足 Kuhn-Tucker 条件的 x_0 成为局部最优解的充要条件是二次型 $\frac{1}{2}x^Tcx$ 在子空间

$$R=\{x\in R^n\,|\,a_i^Tx=0,\ i\in J^*\}$$

上非负. 从后面将要证明的定理可以看出这是一个充分条件, 但同时也不难举出反例说明它并不是必要条件. [2]纠正了这一错误, 提出了正确的充要条件, 但其证明是错误的. 本文给出二次规划(1.1)局部最优解的充要条件的一个严格证明, 并提出对于一般目标函数的平行结果

本文1980年7月7日收到.

二、凸分析中的有关结果

我们的证明基于凸分析的一些结果，即多面锥的 Minkowski 有限基定理和空间的锥投影分解定理[3]. 为自封起见，我们给出这些结果的证明.

定理 2.1 (Minkowski) 多面锥

$$S=\{x\in R^n \mid a_i^T x\leqslant 0,\ i=1,\ 2,\ \cdots,\ m\}$$

具有有限基，即有一组矢量 $y_1,\ y_2,\ \cdots,\ y_k\in S$, S 中每一矢量 x 可用 $y_1,\ y_2,\ \cdots,\ y_k$ 的非负系数线性组合表出:

$$x=\sum_{i=1}^{k}\alpha_i y_i,\quad \alpha_i\geqslant 0.$$

证 对 m 用归纳法证明

当 $m=1$ 时，$S_1=\{x\mid a_1^T x\leqslant 0\}$ 是半空间，取空间 R^n 的一组正交基 $a_1,\ e_2,\ \cdots,\ e_n$, 那末显然有

$$S_1=C(-a_1,\ e_2,\ -e_2,\ \cdots,\ e_n,\ -e_n), \tag{2.1}$$

这里记号 $C(z_1,\ z_2,\ \cdots,\ z_l)$ 表示由诸矢量 $z_1,\ z_2,\ \cdots,\ z_l$ 的非负系数线性组合生成的凸锥.

设对 m 命题已成立，即有一组 $y_1,\ y_2,\ \cdots,\ y_k$, 使得

$$S_m=\{x\mid a_i^T x\leqslant 0,\ i=1,\ 2,\ \cdots,\ m\}=C(y_1,\ y_2,\ \cdots,\ y_k). \tag{2.2}$$

考虑

$$\begin{aligned}S_{m+1}&=\{x\mid a_i^T x\leqslant 0,\ i=1,\ 2,\ \cdots,\ m,\ m+1\}=S_m\cap\{x\mid a_{m+1}^T x\leqslant 0\}\\&=C(y_1,\ \cdots,\ y_k)\cap\{x\mid a_{m+1}^T x\leqslant 0\},\end{aligned} \tag{2.3}$$

只要作一个正系数的比例伸缩后，可以认为诸 y_i 分成三组:

1) $e_i,\ i=1,\ 2,\ \cdots,\ r,\ a_{m+1}^T e_i=-1,$

2) $f_j,\ j=1,\ 2,\ \cdots,\ t,\ a_{m+1}^T f_j=1,$

3) $g_l,\ l=1,\ 2,\ \cdots,\ s,\ a_{m+1}^T g_l=0.$ (2.4)

定义

$$d_{ij}=e_i+f_j,\ i=1,\ 2,\ \cdots,\ r,\ j=1,\ 2,\ \cdots,\ t, \tag{2.5}$$

那末 d_{ij} 和 e_i, g_l 一起都属于 S_{m+1}. 今证所有这些矢量就生成 S_{m+1}. 任取 $x\in S_{m+1}\subset S_m$, 依归纳假设，存在非负数 $\alpha_i,\ i=1,\ 2,\ \cdots,\ r,\ \beta_j,\ j=1,\ 2,\ \cdots,\ t,\ \gamma_l,\ l=1,\ 2,\ \cdots,\ s$, 使得

$$x=\sum_i\alpha_i e_i+\sum_j\beta_j f_j+\sum_l\gamma_l g_l, \tag{2.6}$$

并且满足

$$a_{m+1}^T x=-\sum_i\alpha_i+\sum_j\beta_j\leqslant 0. \tag{2.7}$$

如果 $\sum\limits_j\beta_j=0$, 则一切 $\beta_j=0$, 这时(2.6)已将 x 表成 e_i 和 g_l 的非负系数线性组合. 于是可设 $0<\sum\limits_j\beta_j$. 这时注意(2.7)，命题在 $m+1$ 时成立，可由下列分解式直接看出

$$\begin{aligned}x&=\sum_i\alpha_i e_i+\sum_j\beta_j f_j+\sum_l\gamma_l g_l\\&=\sum_i\alpha_i\left(\frac{\sum\limits_i\alpha_i-\sum\limits_j\beta_j}{\sum\limits_i\alpha_i}\right)e_i+\sum_{i,j}\left(\frac{\beta_j\alpha_i}{\sum\limits_i\alpha_i}\right)(e_i+f_j)+\sum_l\gamma_l g_l.\end{aligned} \tag{2.8}$$

定理 2.2　设 S 是多面锥，那么时空间任一矢量 x，存在到 S 上的投影分解

$$x=y+z,$$

这里 $y\in S$，y 和 z 正交，且

$$\|x-y\|=\min_{w\in S}\|x-w\|.$$

x 的这种分解是唯一的．因而我们可定义两个算子

$$y=Px,\quad z=Qx.$$

如 S 是子空间，定理 2.2 就是平常的正交投影分解．对任一凸闭集(多面锥是凸闭集!)，这种分解仍然成立，而且证明方法完全一样，故从略．只是对一般凸闭集，投影算子 P 不再象到子空间的投影算子那样是线性算子了．

三、二次规划局部最优解的充要条件

对线性约束来说，无需约束品质方面的假定，非线性规划问题的局部最优解必满足 Kuhn-Tucker 条件，即必为 K-T 点．(见[4])因此，我们只须考虑 K-T 点．利用第二节叙述的凸分析的结果，可以证明

定理 3.1　设 x_0 是二次规划的 K-T 点，即有非负数 λ_i，使(1.4)成立．那末 x_0 成为局部最优解的充要条件是二次型 h^Tch 在多面锥

$$S=\{h\,|\,a_j^Th=0,\ j\in J^*;\ a_j^Th\leqslant 0,\ j\in J_1\backslash J^*\}\tag{3.1}$$

上非负，这里标号集 J_1 和 J^* 如(1.3)，(1.5)所定义．

证　必要性

$$f(x_0+\varepsilon h)-f(x_0)=(q^T+x_0^Tc)\,\varepsilon h+\frac{\varepsilon^2}{2}h^Tch=-\varepsilon\sum_{j\in J^*}\lambda_j a_j^Th+\frac{\varepsilon^2}{2}h^Tch.\tag{3.2}$$

设 $h\in S$，则 $a_j^Th=0$，$\forall j\in J^*$，故

$$f(x_0+\varepsilon h)-f(x_0)=\frac{\varepsilon^2}{2}h^Tch.$$

注意 S 是(1.1)在 x_0 处的可行方向锥．当 $\varepsilon>0$ 充分小时，依 x_0 的局部最优性质，知道上式右端非负，从而 c 在 S 上非负．

充分性　依定理 2.2，存在到锥 S(3.1)的投影算子 P 和另一算子 Q，使对每一 $h\in R^n$，可分解为

$$h=Ph+Qh,$$

其中 $Ph\in S$，Ph 和 Qh 正交，且

$$\|Qh\|=\min_{w\in S}\|h-w\|.\tag{3.3}$$

首先证明，存在正数 M，对一切 $h\in\mathscr{D}$

$$\mathscr{D}=\{h\,|\,a_i^Th\leqslant 0,\ i\in J_1\}\tag{3.4}$$

都有

$$\|Qh\|\leqslant -M\sum_{j\in J^*}\lambda_j a_j^Th.\tag{3.5}$$

依 Minkowski 定理 2.1，$\mathscr{D}$ 有一组基矢量，我们把这些基矢量分成二类：一类基矢量 $e_i\in S$，$i=1,\cdots,r$，对 e_i 而言，$a_j^Te_i(j\in J^*)$ 全部为零；另一类基矢量 $f_k\in\mathscr{D}\backslash S$，$k=1,\cdots,t$，

对 f_k 而言，$a_j^T f_k (j\in J^*)$ 不全为零，从而 $\sum_{j\in J^*}\lambda_j a_j^T f_k$ 是一个负数，所以存在常数 $M>0$，使

$$\|Qf_k\|\leqslant -M\sum_{j\in J^*}\lambda_j a_j^T f_k,\quad \forall k=1,\ \cdots,\ t, \tag{3.6}$$

$\mathscr{D}$ 中任一矢量 h 可以表示为

$$\begin{aligned} h&=\sum_i \alpha_i e_i+\sum_k \beta_k f_k \quad (\alpha_i,\ \beta_k\geqslant 0)\\ &=\Big(\sum_i \alpha_i e_i+\sum_k \beta_k P f_k\Big)+\sum_k \beta_k Q f_k. \end{aligned}$$

由于 $\sum_i \alpha_i e_i+\sum_k \beta_k P f_k\in S$，依(3.3)，(3.6)，有

$$\begin{aligned} \|Qh\|&=\min_{w\in S}\|h-w\|\leqslant \Big\|h-\Big(\sum_i \alpha_i e_i+\sum_k \beta_k P f_k\Big)\Big\|=\Big\|\sum_k \beta_k Q f_k\Big\|\\ &\leqslant \sum_k \beta_k\|Qf_k\|\leqslant -M\sum_k \beta_k\sum_{j\in J^*}\lambda_j a_j^T f_k\\ &=-M\sum_{j\in J^*}\lambda_j a_j^T\Big(\sum_i \alpha_i e_i+\sum_k \beta_k f_k\Big)=-M\sum_{j\in J^*}\lambda_j a_j^T h. \end{aligned}$$

现在我们设二次型 h^Tch 在 S 上非负．对 $h\in\mathscr{D}$，依(3.5)有

$$\begin{aligned} h^Tch&=(Ph)^Tc(Ph)+(h+Ph)^Tc(h-Ph)\geqslant (h+Ph)^Tc(h-Ph)\\ &=(h+Ph)^Tc(Qh)\geqslant -\|h+Ph\|\cdot\|c\|\cdot\|Qh\|\geqslant M_1\|h\|\sum_{j\in J^*}\lambda_j a_j^T h, \end{aligned} \tag{3.7}$$

这里 $M_1=2M\|c\|$，在(3.7)推导中利用了 $\|Ph\|\leqslant\|h\|$，这从 h 的正交分解 $h=Ph+Qh$ 看出．

由于当 $i\in J\setminus J_1$ 时，$a_i^T x_0<b_i$，于是存在充分小的正数 ε，当 $\|h\|\leqslant\varepsilon$ 时，$x=x_0+h$ 满足二次规划(1.1)约束条件的充要条件是 $h\in\mathscr{D}$．这时依(3.2)，(3.7)，有

$$\begin{aligned} f(x_0+h)-f(x_0)&=-\sum_{j\in J^*}\lambda_j a_j^T h+\frac{1}{2}h^Tch\geqslant -\sum_{j\in J^*}\lambda_j a_j^T h+\frac{1}{2}M_1\|h\|\sum_{j\in J^*}\lambda_j a_j^T h\\ &=-\Big(1-\frac{1}{2}M_1\|h\|\Big)\sum_{j\in J^*}\lambda_j a_j^T h\geqslant 0, \end{aligned}$$

只要再让 $\|h\|\leqslant\dfrac{2}{M_1}$．这就证明了 x_0 是二次规划(1.1)的局部最优解．

系 3.2　考虑带有等式约束的一般二次规划

$$\begin{cases} \min f(x)=q^Tx+\dfrac{1}{2}x^Tcx\\ a_j^Tx\leqslant b_j,\ j\in J=(1,\ 2,\ \cdots,\ N),\\ w_i^Tx=v_i\ \ i\in I=(1,\ 2,\ \cdots,\ M). \end{cases} \tag{3.8}$$

设 x_0 是 K-T 点，即有实数 μ_i，$i\in I$ 及非负数 λ_j，$j\in J$，使

$$\begin{cases} q+cx_0+\sum_{i\in I}\mu_i w_i+\sum_{j\in J}\lambda_j a_j=0\\ \sum_{j\in J}\lambda_j(a_j^Tx_0-b_j)=0. \end{cases}$$

和不带等式约束情况一样．

定义
$$J_1=\{j\in J\,|\,a_j^Tx_0-b_j=0\},$$
$$J^*=\{j\in J_1\,|\,\lambda_j>0\},$$

那末 x_0 是二次规划(3.8)局部最优解的充要条件是二次型 h^Tch 在锥

$$S=\left\{h\in R^n\,\middle|\,\begin{array}{l} a_j^Th=0,\ j\in J^*;\ a_j^Th\leqslant 0,\ j\in J_1\setminus J^*;\\ w_i^Th=0,\ i\in I \end{array}\right\}$$

上非负.

由于我们的证明主要只依赖于目标函数的线性部份，对一般目标函数可以得出平行的结果，但我们只叙述成简化的形式.

系 3.3 考虑规划

$$\begin{cases}\min f(x)=-\sum\limits_{j\in J^*}a_j^T x+g(x),\\ a_j^T x\leqslant 0,\ j\in J_1,\ (J^*\subset J_1).\end{cases}\tag{3.9}$$

设函数 $g(x)$ 可微分，$\nabla g(0)=0$，那末 $x=0$ 是(3.9)局部最优解的充要条件是 $x=0$ 是下列规划的局部最优解

$$\begin{cases}\min g(x)\\ a_j^T x=0,\ j\in J^*;\ a_j^T x\leqslant 0,\ j\in J_1\backslash J^*.\end{cases}$$

利用估计式 3.5，系 3.3 和定理 3.1 一样证明.

本文是在中国科学院应用数学所韩继业老师帮助下完成的，谨致谢意.

参 考 文 献

[1] Ritter, K., Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, **4**(1965), 149—153.
[2] Majthay, A., Mathematical Programming, **1**(1971) 359—365.
[3] Stoer, J., Witzgall, C., Convexity and Optimization in Finite Dimensions I, Springes Verley 1970.
[4] Bazaraa, M. S., Shetty, C. M., Nonlinear Programming Theory and Algorithms., Jahn Wiley & Sons, 1979.

ON THE NECESSARY AND SUFFICIENT CONDITI ON OF THE LOCAL OPTIMAI SOLUTION OF QUADRATIC PROGRAMMING

LIU JIAQUAN SONG TIANTAI DU DINGZHU
(Graduate School, University of Science and Technology of China)

ABSTRACT

In this paper a necessary and sufficient aptimality condition of quadratic programming is giver. The main result is Theorem 3.1. It is shown that a point x_0 is local optimal solution of quadratic programming (1, 1) if and only if x_0 is a Kuhn-Tucker point and $h^T ch$ is non-negative for every point in $\{h \mid a_j^T h=0,\ j\in J^*,\ a_j^T h\leqslant 0,\ j\in J_1\backslash J^*\}$.