文章编号:1672 - 058X(2013)04 - 0013 - 04

不等式约束非线性规划的拉格朗日乘子的收敛

覃亚梅
(重庆师范大学 数学学院,重庆 401331)

摘　要:对于等式约束的非线性规划问题,一般的解决方法是在每次迭代中更新拉格朗日乘子且逐渐增大拉格朗日函数的惩罚因子,当罚因子充分大或充分接近局部最优解时,二阶充分条件是满足的;对不等式约束问题也采用了相应的方法.在凸的情况下,对于任意的罚因子或者在每次迭代中不要求精确极小化,就能全局收敛到最优解;证明了拉格朗日乘子是收敛的.

关键词:非线性规划 ;增广拉格朗日函数; 拉格朗日乘子收敛
中图分类号:O182.1　　**文献标志码:**A

1 引　言

下面的思想,是 Hestens and Powell 提出的.

考虑一个等式约束非线性规划问题:

$$\begin{aligned} &\min f(x) \\ &\text{s.t.}\ \ g_i(x) = 0 \quad i = 1,2,\cdots,m \end{aligned} \tag{1}$$

其中 f,g 是 $R^n - R$ 的可微函数,若把的项加到其拉格朗日函数上,问题的最优解是不会改变的. 即

$$k_r(x,y) = f(x) + \sum_{i=1}^{m}[\lambda_i g_i(x) + r g_i^2(x)]$$

其中常数 $r>0$ 是罚因子. 若$\bar{x}$是一个局部最优解则乘子向量$\bar{\lambda}$满足:

$$\nabla f(\bar{x}) + \sum_{i=1}^{m}\overline{\lambda_i} g_i(\bar{x}) = 0$$

当 r 充分大时,则 $k_r(x,\bar{\lambda})$在$\bar{x}$有一个局部最优解,(若$\bar{x}$满足式(1)的局部最优的二阶充分条件,这结论是显然的). 下面通过一个无约束序列问题来解决式(1).

$$\min k_r(x,\lambda^k)$$

这里,如果收敛显得不是足够的快,r 可能被一次次的增大,x^k 满足:

$$0 = \nabla_x k_r(x^k,\lambda^k) = \nabla f(x^k) + \sum_{i=1}^{m}(\lambda_i^k + 2rg_i(x^k))\nabla g_i(x^k)$$

便产生了一个序列$\{\lambda^k\}$,其中

$$\lambda_i^{k+1} = \lambda_i^k + 2rg_i(x^k) \quad i = 1,2,\cdots,m$$

在文献[1] 中表明:不需要 $r\to\infty$,二阶最优性条件满足,且这个算法是线性收敛的. Miele, et al[2-5] 已经尝

收稿日期:2012 - 11 - 04;**修回日期:**2012 - 12 - 01.
作者简介:覃亚梅(1987-),女,重庆大足人,硕士研究生,从事最优化理论与算法方向研究.

试对这种算法的不同的修正. Fletcher[6,7]在 $k_r(x,y)$ 的极小化过程中,乘子向量 λ 不断被调整. Arrow 和 Solow 表明可以通过微分方程来求解鞍点. Rockafellar 已经研究了对不等式约束的类似的拉格朗日函数 k_r,在凸的情形下,已经确立了一些相关的性质,对于非凸,不等式约束的鞍点理论已经被 Arrow,Gould 和 Howe 证明.

不等式约束拉格朗日函数,其中 $g_i(x)=0$ 被 $g_i(x)\leqslant 0$ 代替,其形式表示如下:

$$L_r(x,\lambda)=f(x)+\frac{1}{4r}\sum_{i=1}^{m}[(\theta(\lambda_i+2rg_i(x)))^2-\lambda_i^2]$$

其中,
$$\theta(t)=\max\{t,0\}$$

因此
$$\frac{1}{4r}\sum_{i=1}^{m}[(\theta(\lambda_i+2rg_i(x)))^2-\lambda_i^2]=\begin{cases} rg_i^2(x)+\lambda_i g_i(x) & \text{当 } g_i(x)\geqslant -\dfrac{\lambda_i}{2r} \\ -\dfrac{\lambda_i}{4r} & \text{当 } g_i(x)\leqslant -\dfrac{\lambda_i}{2r}\end{cases}$$

这里乘子 λ_i 没有被限制是非负的.

算法的基本程序是,给定 λ^k(且 $r>0$),通过极小化 $L_r(x,\lambda^k)$ 来确定 x^k.

$$\min L_r(x,\lambda^k) \tag{2}$$

即

$$0=\nabla_x L_r(x^k,\lambda^k)=\nabla f(x^k)+\sum_{i=1}^{m}(\lambda_i^k+2rg_i(x^k))\nabla g_i(x^k)$$

设

$$\lambda_i^{k+1}=\theta(\lambda_i^k+2rg_i(x^k))\geqslant 0 \tag{3}$$

或者等价于

$$\lambda^{k+1}=\lambda^k+2r\nabla_y L_r(x^k,y^k) \tag{4}$$

由文献[8]知,函数 f,g 是凸的和不必要可微,通过式(2)、(3)这种算法如果满足 slater 约束品性,则任意固定的 r,近似极小化全局收敛. 下面由 MORCAU'S 映射定理,来证明乘子序列 $\{\lambda^k\}$ 是收敛的.

2 乘子的收敛性

假设 X 是实向量空间 E(可能是无限维)的一个凸集 $f,g_1\cdots g_m$ 是在 X 的实值凸函数. 将关注下面的问题:

$$\begin{aligned} &\min f(x) \\ &\quad g_i(x)\leqslant 0 \quad i=1,2,\cdots m \\ &\quad x\in X \end{aligned}$$

由于 $g_i(x)$ 是凸的,也隐含着 $L_r(x,\lambda)$ 在 $x\in X$ 中也是凸的,$\lambda\in R^m$ 是凹的(文献 8 定理 1).

这个算法依赖于初试选择 $r>0,\lambda^0\in R^m$ 和一个序列 $\{\varepsilon_k\}$,其中 $\varepsilon_k\geqslant 0$ 且 $\lim\limits_{k\to\infty}\varepsilon_k=0$. 给定 $\lambda^k\in R^m$,确定 $x^k\in X$ 使得

$$L_r(x^k,y^k)\leqslant \inf_{x\in X}L_r(x,y^k)+\varepsilon_k \tag{5}$$

其中 λ^{k+1} 是式(3)或式(4).

设 u 是问题 P 的最优解,即满足约束的最小值. 对于问题 P 的一个 $K-T$ 向量 $\overline{\lambda}\in R_+^m$ 有性质,即

$$\inf_{x\in X}\{f(x)+\overline{\lambda}_1 g_1(x)+\cdots+\overline{\lambda}_m g_m(x)\}=u>-\infty$$

定理 2.1 假设 P 具有至少一阶 $K-T$ 向量

$$\sum_{K=1}^{\infty}\varepsilon_k<+\infty \tag{6}$$

那么 $\{\lambda^K\}$ 收敛于某个 $K-T$ 向量,其中 $\{x^k\}$ 满足

$$\limsup_{k\to\infty} g_i(x^k) \leqslant 0 \quad i = 1,2,\cdots,m \tag{7}$$

$$\lim_{k\to\infty} f(x^k) = u = \inf(p)$$

假定最优值 u 是有极限的,如果 Slater 条件满足,即存在一个点 $x \in X$ 使得 $g_i(x) < 0, i = 1,2,\cdots m$ 则这 $K-T$向量是存在的. 若一个 $K-T$ 向量是存在的,在 X 中的序列不可能既满足式(7)也满足

$$\lim_{k\to\infty} f(x^k) < u$$

证明　由文献 8 知这个函数

$$g_r(\lambda) = \inf_{x\in X} L_r(x,\lambda)$$

是凹的. 在 $K-T$ 向量是存在的假设下,g_r 是处处有限,在 R^m 中连续可微,且这个 $K-T$ 向量来自它的最大值(文献 8,定理 2),更进一步说

$$g_r(\lambda) = \max_{w\in R^m}\{g_0(w) - \frac{1}{4r}|w-\lambda|^2\} \tag{8}$$

其中

$$g_0(\lambda) = \begin{cases} \inf\{f(x) + \lambda_1 g_1(x) + \cdots \lambda_m g_m(x)\} & \text{当 } \lambda \geqslant 0 \\ -\infty & \text{当 } \lambda \leqslant 0 \end{cases} \tag{9}$$

记 $w = M(\lambda)$是使式(8)获得最大值. 其中$\nabla g_r(y)$表示在 $W = M(\lambda)$对 $g_0(w) - \frac{1}{4r}|w-\lambda|^2$ 的 λ 的梯度,即

$$M(\lambda) = \lambda + 2r\nabla g_r(\lambda) \tag{10}$$

由(文献 8 定理 2)条件式(5)产生这个估算

$$|\nabla L_r(x^k,\lambda^k) - \nabla g_r(\lambda^k)|^2 \leqslant \varepsilon_k$$

由式(4),(10)

$$\frac{1}{r^2}|\lambda^{k+1} - M(\lambda^k)|^2 \leqslant \varepsilon_k \tag{11}$$

因此

$$\lim_{k\to\infty}[\lambda^{k+1} - M(\lambda^k)] = 0 \tag{12}$$

由 MORCAU'S 映射,这里 M 是非膨胀映射,即对任意 y,z 有

$$|M(y) - M(z)| \leqslant |y - z| \tag{13}$$

M 中一个固定的点,使得$\nabla g_r(\lambda) = 0$ 则 λ 是问题(P)的一个 $K-T$ 向量. 设$\bar{\lambda}$是任意的一个 $K-T$ 向量. 则

$$|M(\lambda^k) - \bar{\lambda}| = |M(\lambda^k) - M(\bar{\lambda})| \leqslant |\lambda^k - \bar{\lambda}|$$

由式(11),(13)知:

$$|\lambda^{k+1} - \bar{\lambda}| \leqslant |\lambda^{k+1} - M(\lambda^k)| + |M(\lambda^k) - \bar{\lambda}| \leqslant (r^2\varepsilon_k)^{\frac{1}{2}} + |\lambda^k - \bar{\lambda}|$$

由式(6)知

$$|\lambda^l - \bar{\lambda}| \leqslant |\lambda^k - \bar{\lambda}| + \sum_{s=k}^{\infty}(r^2\varepsilon_s)^{\frac{1}{2}} < -\infty, l > k$$

由于$\{\lambda^k\}$是一个有界序列,g_r 在有界集上是连续一致的,则有式(12)得

$$\lim_{k\to\infty}[g_r(\lambda^{k+1}) - g_r(M(\lambda^k))] = 0 \tag{14}$$

由式(8)得 $g_r(\lambda^k) = g_0(M(\lambda^k) - \frac{1}{4r}|M(\lambda^k) - \lambda^k|^2$ 或者有式(10)得

$$g_r(\lambda^k) + r|\nabla g_r(\lambda^k)|^2 = g_0(M(\lambda^k)) \tag{15}$$

由式(8)也得

$$g_r(M(\lambda^k)) \geqslant g_0(M(\lambda^k)) - \frac{1}{4r}|M(\lambda^k) - M(\lambda^k)|^2 = g_0(M(\lambda^k)) \tag{16}$$

由式(15)、(16)得

$$g_r(M(\lambda^k)) \geqslant g_r(\lambda^k) + r\,|\,\nabla g_r(\lambda^k)\,|^2 \tag{17}$$

由式(14)、(17)和 g_r 有上界的. 可得 $\lim\limits_{k\to\infty} |\nabla g_r(\lambda^k)| = 0$. 所以,若 $\bar{\lambda}$ 是有界数列 $\{\lambda^k\}$ 的极限点,那么 $\nabla g_r(\bar{\lambda}) = 0$. 由于 g_r 是凹的,这意味着 $\bar{\lambda}$ 是 g_r 的最大值,且是一个 $K-T$ 向量. 有(30)对任意的 $K-T$ 向量都成立,那么 $\bar{\lambda}$ 是 $\{\lambda^k\}$ 的唯一极限点. 因此 $\{\lambda^k\}$ 是极大化 g_r 的序列,且收敛于问题 P 的某一 $K-T$ 向量. 证明完毕.

参考文献:

[1] POWELL M J D. A Method for Nonlinear Constraints in Minimization Problems[J]. Optimization,1972

[2] MIELEA C, IVER E E. Use of the Augmented Penalty Function in Mathematical Programming Problems [J]. Journal of Optimization Theory and Applications,1971(8):115-130

[3] MIELEA C, LEVY A V. Use of the Augmented Penalty Function in Mathematical Programming Problems [J]. Journal of Optimization Theory and Applications,1971(8):131-153

[4] MIELEA M,CRAGG E E. A Modification of the Method of Multipliers for Mathematical Programming Problems[M]. techniques of Optimization,New York:BALAKRISHNAN,1972

[5] MIELEA M,LEVY A V,COGGINS G M. On the Method of Multipliers for Mathematical Programming of Optimization Theory and Applications[J]. 1972(10):1-33

[6] FLETCHER R. A Class of Methods for Nonlinear Programming with Termination and Convergence Properties[J]. Integer and Nonlinear Programming,1971(30):107-111

[7] FLETCHER R,Hill S A. A class of Methods for Nonlinear Programming2:Computational Experience[J]. Nonlinear Programming, 1971(27):100-105

[8] ROCKAFELLAR R T. A Dual Approach to Solving Nonlinear Programming problems by UFAN Unconstrained Optimization[J]. Mathematical Programming,1976(8):101-110

The Convergence of Lagrange Multiplier for Nonlinear Programming with General Inequality Constraints

QIN Ya-mei

(School of Mathematics,Chongqing Normal University,Chongqing 401331,China)

Abstract:As for nonlinear programming problem with equality constraints,a general method of solution is to gradually increase penalty factor of Lagrange function and to renew Lagrange multipliers in the iteration of each cycle. if penalty factor is sufficiently large or is close to local optimal solution, the second-order sufficient conditions are satisfied. This paper uses the corresponding method for inequality-constrained problems. Global convergence to an optimal solution is established in the convex case for an arbitrary penalty factor or without the requirement of an exact minimization in the iteration of each cycle. Furthermore, the Lagrange multipliers are proved to converge.

Key words:nonlinear programming;the Augmented Lagrange Function ;convergence of Lagrange multipliers

责任编辑:田 静