第26卷 第4期
2011年8月
郑州轻工业学院学报(自然科学版)
JOURNAL OF ZHENGZHOU UNIVERSITY OF LIGHT INDUSTRY(Natural Science)
Vol.26 No.4
Aug. 2011

文章编号:1004-1478(2011)04-0117-04

等式约束下一个带线搜索的信赖域算法

李少娟, 景书杰

(河南理工大学 数学与信息科学学院, 河南 焦作 454003)

摘要:提出了一个等式约束下凸二次规划问题的带强 Wolfe 线搜索的信赖域算法.该算法利用增广 Lagrange 函数将约束问题转化为无约束问题,在传统信赖域算法的基础上结合线搜索技术,当试探步失败时不用重解信赖域子问题,减少了计算量.在适当的条件下,证明了此算法的全局收敛性.

关键词:等式约束;强 Wolfe 线搜索;信赖域算法;增广 Lagrange 函数;全局收敛性

中图分类号:O221.2　　**文献标志码**:A

A trust region algorithm with line search under equality constraints

LI Shao-juan, JING Shu-jie

(*School of Math. and Infor. Sci., He'nan Polytechnic Univ., Jiaozuo 454003, China*)

Abstract: A new trust region algorithm with strong Wolfe line search for convex quadratic programming under equality constraints is proposed. This method first uses augmented Lagrange function to transform this restraint question into the non-constraint question, and on the basis of traditional trust region method it takes line search instead of resolving the subproblem when the trail step is not successful. This may allow a considerable computational saving. Global convergence is proved under certain conditions.

Key words: equality constraint; strong Wolfe line search; trust region algorithm; augmented Lagrange function; global convergence

0 引言

考虑等式约束优化问题

$$\min f(x)$$

$$\text{s.t.} \quad c_i(x)=0 \quad i=1,\cdots,m$$

其中,$f(x):R^n\to R$ 及 $c_i(x)$ 都是二次连续可微函数.本文主要讨论利用信赖域算法求解等式约束下凸二次规划问题

$$\min f(x)=c^Tx+\frac{1}{2}x^THx \qquad ①$$

$$\text{s.t.} \quad Ax=b \qquad ②$$

其中,H 为 n 阶正定对称阵,$A\in R^{m\times n}$,$b\in R^m$,$c\in R^n$,$x\in R^n$.信赖域算法最先由 M. J. D. Powell[1] 提出,是指在一个可以“信赖”的区域内将目标函数用一个二次函数替换.该算法不如线搜索成熟,应用也没有线搜索广泛,但由于它具有较强的收敛性和可靠性,对它的研究越来越受到学界的重视.

J. Nocedal 等[2]提出了将信赖域算法和线搜索方法相结合来构造新的计算方法的思想.姚升保等[3]将 Armijo 线搜索应用到了非单调信赖域算法上.李红等[4]将线搜索应用到了利用 R-函数调节信赖域半径的自适应信赖域算法上.本文拟将等式约束问题转化为无约束问题,再借鉴文献[2]的思想,并在文献[3-4]的基础上,将强 Wolfe 线搜索用

收稿日期:2011-04-19

作者简介:李少娟(1984—),女,河南省濮阳市人,河南理工大学硕士研究生,主要研究方向为最优化理论及其应用.

于信赖域算法.当试探步失败时并不重解信赖域子问题,而是采用线搜索以得到下一个迭代点,以减少计算量,并在一定条件下证明此算法的全局收敛性.

1 问题的转化

定义增广 Lagrange 函数

$$L(x,\lambda,\xi) = f(x) - \sum_{i=1}^{m}\lambda_i c_i(x) + \frac{\xi}{2}\sum_{i=1}^{m} c_i^{2}(x)$$

利用增广 Lagrange 函数将约束问题转化为无约束问题,问题①②的 Lagrange 函数为

$$L(x,\lambda,\xi) = f(x) - \lambda^{T}(Ax-b) + \frac{\xi}{2}(Ax-b)^{T}(Ax-b)$$

式中 $\lambda = (\lambda^1,\lambda^2,\cdots,\lambda^m)^T \in R^m$, $\xi > 0$. 于是问题①②可以转化为以下无约束最小值问题

$$\min F(x) = L(x,\lambda,\xi) = f(x) - \lambda^{T}(Ax-b) + \frac{\xi}{2}(Ax-b)^{T}(Ax-b) \quad ③$$

信赖域算法是求解问题③的一类重要的数值计算方法.其迭代格式为

$$x_{k+1} = x_k + \alpha_k d_k$$
$$\lambda_{k+1} = \lambda_k - \xi(Ax_k - b)$$

其中,λ_k 为乘子向量,d_k 为搜索方向,通过求解信赖域子问题

$$\min_{d\in R^n}\varphi_k(d) = g_k^{T}d + \frac{1}{2}d^{T}\boldsymbol{B}_k d \quad ④$$

$$\text{s.t.} \quad \| d \| \leqslant \Delta_k \quad ⑤$$

可知,x_k 是问题③的当前迭代点,记 $F_k = F(x_k)$, $g_k = \nabla F(x_k)$, α_k 为步长,由线搜索确定 $\boldsymbol{B}_k \in R^{n\times n}$ 是近似于 Hessian 阵 $\nabla^2 F(x_k)$ 的对称矩阵,$\Delta_k > 0$ 是信赖域半径,$\|\cdot\|$ 表示向量或矩阵的 2-范数.

2 算法

本文算法中用到的线搜索为强 Wolfe 线搜索[5]

$$F(x_k + \alpha_k d_k) \leqslant F(x_k) + \sigma_1 \alpha_k g_k^{T} d_k \quad ⑥$$

$$|g_{k+1}^{T} d_k| \leqslant \sigma_2 |g_k^{T} d_k| \quad ⑦$$

其中,$0 < \sigma_1 < \sigma_2 < 1$.

求解无约束问题的线搜索信赖域算法具体步骤如下:

1)给定 $x_1 \in R^n$,对称阵 $\boldsymbol{B}_1 \in R^{n\times n}$, $\varepsilon > 0$, $\Delta_1 > 0$, $\tau > 0$, $0 < \mu < 1$, $0 < \sigma_1 < \sigma_2 < 1$, $0 < \beta_1 < \beta_2 < 1 < \beta_3$,乘子向量初始估计 λ_1,参数 $\xi > 0$, $k:=1$.

2)计算 g_k,如果 $\| g_k \| \leqslant \varepsilon$,则停止.

3)求解子问题④⑤得近似解 d_k,且满足

$$\varphi_k(0) - \varphi_k(d_k) \geqslant \tau \| g_k \| \min\left\{\Delta_k, \frac{\| g_k \|}{\| \boldsymbol{B}_k \|}\right\} \quad ⑧$$

$$g_k^{T} d_k \leqslant -\tau \| g_k \| \min\left\{\Delta_k, \frac{\| g_k \|}{\| \boldsymbol{B}_k \|}\right\}$$

其中 $\varphi_k(0) - \varphi_k(d_k) = \mathrm{Pred}_k = -g_k^{T} d_k - \frac{1}{2} d_k^{T} \boldsymbol{B}_k d_k$.

4)计算 $\rho_k = \frac{Ared_k}{\mathrm{Pred}_k}$,其中 $Ared_k = F_k - F(x_k + d_k)$.

5)若 $\rho_k \geqslant \mu$,令 $x_{k+1} = x_k + d_k$, $\Delta_{k+1} \in [\| d_k \|, \beta_3 \| d_k \|]$;否则进行线搜索求 α_k,使其满足式⑥⑦,令

$$x_{k+1} = x_k + \alpha_k d_k, \Delta_{k+1} \in [\beta_1 \| d_k \|, \beta_2 \| d_k \|]$$

6)令 $\lambda_{k+1} = \lambda_k - \xi(Ax_k - b)$,用 BFGS 公式修正 B_{k+1}, $k:=k+1$,转 2).

3 收敛性分析

为证明算法的全局收敛性,作以下假设:

H_1 $\{x_k\}$ 是由算法产生的点列,数列 $\{F_k\}$ 有下界;

H_2 $\nabla F(x)$ 在 R^n 上一致连续.

引理 1 若假设 H_2 成立,且存在常数 $\delta > 0$,使对 $\forall k$ 有 $\| g_k \| \geqslant \delta$,则存在常数 $m > 0$,使得

$$\| d_k \| \geqslant \frac{m}{Z_k} \quad k = 1,2,\cdots \quad ⑨$$

其中,$Z_k := \max_{1\leqslant i\leqslant k} \| B_i \| + 1$.

证明 由假设 $\nabla F(x)$ 在 R^n 上一致连续可知,对于 $\forall \varepsilon > 0$,存在 $\mu > 0$,当 $\| d \| \leqslant \mu$ 时,有

$$d_k^{T}(\nabla F(x+d) - \nabla F(x)) < \varepsilon \| d_k \| \quad \forall x \in R^n$$

取 $\varepsilon = \frac{\tau\delta(1-\mu)}{2}$,其中 $0 < \mu < 1$, $\tau > 0$,则

$$d_k^{T}(\nabla F(x+d) - \nabla F(x)) < \frac{\tau\delta(1-\mu)}{2} \| d_k \| \quad \forall x \in R^n \quad ⑩$$

令

$$m = \{\beta_1\delta, \Delta_1 Z_1, \beta_1\eta Z_1, \tau\delta\beta_1(1-\mu)\} \quad ⑪$$

当 $\| d_k \| < \Delta_k$ 时,由 d_k 的最优性条件知 $g_k = -\boldsymbol{B}_k d_k$,于是有

$$\| g_k \| = \| -\boldsymbol{B}_k d_k \| = \| \boldsymbol{B}_k d_k \| \leqslant \| \boldsymbol{B}_k \| \cdot \| d_k \|$$

由⑪式及 $0 < \beta_1 < 1$ 有 $m \leqslant \beta_1\delta < \delta$,由 Z_k 的定义

显然有 $Z_k > \|B_k\|$,从而有

$$\|d_k\| \geqslant \frac{\|B_k d_k\|}{\|B_k\|} = \frac{\|g_k\|}{\|B_k\|} \geqslant \frac{\delta}{\|B_k\|} > \frac{\delta}{Z_k} > \frac{m}{Z_k}$$

当 $\|d_k\| = \Delta_k$ 时,用数学归纳法证明不等式⑨成立.

当 $\|d_k\| = \Delta_1$ 时,由⑪式有 $m \leqslant \Delta_1 Z_1$,从而 $\|d_1\| = \Delta_1 \geqslant \frac{m}{Z_1}$,故当 $k=1$ 时,不等式⑨成立. 现在假设⑨对 k 成立,即当 $\|d_k\| = \Delta_k$ 时,有 $\|d_k\| \geqslant \frac{m}{Z_k}$. 下证不等式⑨对 $k+1$ 也成立. 当 $\|d_{k+1}\| = \Delta_{k+1}$ 时,由序列 $\{Z_k\}$ 单调上升得 $\frac{m}{Z_k} \geqslant \frac{m}{Z_{k+1}}$,所以只需证明下式成立即可.

$$\Delta_{k+1} \geqslant \frac{m}{Z_k} \qquad ⑫$$

此时,若 $\Delta_{k+1} \geqslant \|d_k\|$,则由归纳假设 $\|d_k\| \geqslant \frac{m}{Z_k}$ 易知⑫式成立,因此只需考虑算法中 $\Delta_{k+1} \geqslant \|d_k\|$ 的情形,即 $\Delta_{k+1} \in [\beta_1\|d_k\|, \beta_2\|d_k\|]$ 的情形. 如果 $\|d_k\| \geqslant \eta$,则有

$$\Delta_{k+1} \geqslant \beta_1\|d_k\| \geqslant \beta_1\eta \geqslant \frac{\beta_1\eta Z_1}{Z_k} \geqslant \frac{m}{Z_k}$$

所以下面只需证明当 $\Delta_{k+1} < \|d_k\| < \eta$ 时⑫式成立即可. 此时,由算法易知 $\rho_k < \mu$,即有

$$F(x_k + d_k) - F(x_k) > -\mu \mathrm{Pred}_k =$$

$$-\mu(\varphi_k(0) - \varphi_k(d_k)) = \mu\left(g_k^T d_k + \frac{1}{2} d_k^T B_k d_k\right) \qquad ⑬$$

由⑩式知

$$F(x_k + d_k) - F(x_k) = \int_0^1 d_k^T \nabla F(x_k + \theta d_k)\mathrm{d}\theta =$$

$$\int_0^1 d_k^T(\nabla F(x_k + \theta d_k) - \nabla F(x_k))\mathrm{d}\theta +$$

$$d_k^T \nabla F(x_k) \leqslant d_k^T g_k + \frac{\tau\delta(1-\mu)}{2}\|d_k\| \qquad ⑭$$

由⑬⑭式可得

$$(1-\mu)\left(d_k^T g_k + \frac{\tau\delta}{2}\|d_k\|\right) \geqslant \frac{\mu}{2} d_k^T B_k d_k \qquad ⑮$$

而由⑧式有

$$-d_k^T g_k - \frac{1}{2} d_k^T B_k d_k \geqslant \tau\delta \min\left\{\|d_k\|, \frac{\delta}{\|B_k\|}\right\} \qquad ⑯$$

将⑯式两边乘以 $(1-\mu)$ 加到⑮式,得

$$-d_k^T B_k d_k >$$

$$\tau\delta(1-\mu)\min\left\{\|d_k\|, \frac{2\delta}{\|B_k\|} - \|d_k\|\right\}$$

注意到 $-d_k^T B_k d_k \leqslant \|d_k\|^2 \cdot \|B_k\|$,于是有

$$\|d_k\|^2 \cdot \|B_k\| >$$

$$\tau\delta(1-\mu)\min\left\{\|d_k\|, \frac{2\delta}{\|B_k\|} - \|d_k\|\right\}$$

考虑当 $\|d_k\| \leqslant \frac{\delta}{\|B_k\|}$ 时

$$\min\left\{\|d_k\|, \frac{2\delta}{\|B_k\|} - \|d_k\|\right\} = \|d_k\|$$

代入⑰式有

$$\|d_k\| \cdot \|B_k\| > \tau\delta(1-\mu)$$

当 $\|d_k\| > \frac{\delta}{\|B_k\|}$ 时,直接可得 $\|d_k\| \cdot \|B_k\| > \delta$,根据 m 的取法,有

$$\|d_k\| \cdot \|B_k\| > \min\{\tau\delta(1-\mu), \delta\} =$$

$$\frac{1}{\beta_1}\min\{\beta_1\delta, \tau\delta\beta_1(1-\mu)\} \geqslant \frac{m}{\beta_1}$$

所以有 $\|d_k\| \cdot \beta_1 \geqslant \frac{m}{\|B_k\|}$,从而可得

$$\Delta_{k+1} \geqslant \beta_1\|d_k\| \geqslant \frac{m}{\|B_k\|} > \frac{m}{Z_k}$$

故当 $\|d_{k+1}\| = \Delta_{k+1}$ 时,⑫式成立. 因此,对任意 k,当 $\|d_k\| = \Delta_k$ 时,不等式⑨成立. 从而引理 1 成立.

引理 2[6]　设 $\{\Delta_k\}$ 和 $\{Z_k\}$ 是 2 个正数列,若存在正常数 $m>0$, $0<\beta_2<1$, $\beta_3>1$ 以及 $\{1,2,3,\cdots\}$ 的 1 个子集 I,使得

$$\Delta_{k+1} \leqslant \beta_3\Delta_k, \forall k \in I \qquad \Delta_{k+1} \leqslant \beta_2\Delta_k, \forall k \notin I$$

$$\Delta_k \geqslant \frac{m}{Z_k}, \forall k \qquad Z_{k+1} \geqslant Z_k, \forall k \qquad \sum_{k\in I}\frac{1}{Z_k} < +\infty$$

则必有

$$\sum_{k=1}^{+\infty}\frac{1}{Z_k} < +\infty$$

定理 1　若假设 H_1, H_2 均成立,且 Z_k 满足

$$\sum_{k=1}^{+\infty}\frac{1}{Z_k} = +\infty \qquad ⑱$$

则由算法产生的点列 $\{x_k\}$ 必满足

$$\liminf_{k\to+\infty}\|g_k\| = 0$$

证明　假设定理不成立,那么存在常数 $\delta>0$,使 $\|g_k\| \geqslant \delta$, $\forall k$. 由引理 1 可知⑨式成立. 定义 $I = \{k \mid \rho_k \geqslant \mu\}$,由假设 H_1 中 $\{F_k\}$ 有下界及算法易知 F_k 是单调不增的,所以有 $\lim\limits_{k\to\infty}(F_k - F_{k+1}) = 0$,从而有 $\sum\limits_{k=1}^{+\infty}(F_k - F_{k+1}) < +\infty$. 又由引理 1 和⑧式及 $\|d_k\| \leqslant \Delta_k$ 有

$$\sum_{k=1}^{+\infty}(F_k - F_{k+1}) \geqslant \sum_{k\in I}(F_k - F_{k+1}) \geqslant$$

$$\sum_{k\in I}\mu \mathrm{Pred}_k \geqslant \sum_{k\in I}\mu\tau\delta\min\left\{\|d_k\|,\frac{\delta}{Z_k}\right\} \geqslant$$

$$\sum_{k\in I}\mu\tau\delta\frac{1}{Z_k}\min\{\|d_k\|\cdot Z_k,\delta\} \geqslant$$

$$\sum_{k\in I}\mu\tau\delta\frac{1}{Z_k}\min\{m,\delta\}$$

因此 $\sum_{k\in I}\frac{1}{Z_k} < +\infty$,应用引理2有

$$\sum_{k=1}^{+\infty}\frac{1}{Z_k} < +\infty$$

这与⑱式矛盾,故定理成立.

参考文献:

[1] Powell M J D. A New Algorithm for Unconstrained Optimization[C]//Nonlinear Programming, New York: Academic Press, 1970:31-66.

[2] Nocedal J, Yuan Y X. Combining trust region and line search techniques[C]//Advances in Nonlinear Programming, Berlin: Kluwer Academic Publishers, 1998:153.

[3] 姚升保,施保昌,彭叶辉.一类带线搜索的非单调信赖域算法[J].数学杂志,2003,23(3):290.

[4] 李红,焦宝聪.一类带线搜索的自适应信赖域算法[J].运筹学学报,2008,12(2):97.

[5] 王宜举,修乃华.非线性规划理论与算法[M].2版.西安:陕西科学技术出版社,2008:25-26.

[6] 袁亚湘,孙文瑜.最优化理论与方法[M].北京:科学出版社,1997:563-565.

(上接第78页)

上进行优化的,试验证明该算法是有效和成功的.

参考文献:

[1] Zhang Defu, Kang Yan, Deng Ansheng. A new heuristic recursive algorithm for the strip rectangular packing problem [J]. Computer&Operations Research, 2006, 33(8):2209.

[2] Hopper E, Turton B. A genetic algorithm for a 2D industrial packing problem[J]. Comp and Ind Eng, 1999, 37(1):375.

[3] 贾志欣,殷国富,罗阳.二维不规则零件排样问题的遗传算法求解[J].计算机辅助设计与图形学学报,2002,14(5):467.

[4] 王小平,曹立明.遗传算法理论与应用[M].西安:西安交通大学出版社,2001:4-9.

[5] 李明,宋成芳,周泽魁.二维不规则零件排样问题的粒子群算法求解[J].江南大学学报:自然科学版,2005,4(3):266.

[6] 刘心雄,许昌永,王弘.不规则图形样件的自动排料设计[J].华中科技大学学报:自然科学版,2002,30(2):19.