运筹学与控制论

DOI:10.11721/cqnuj20140201

一类等式约束非线性优化问题的序列二次规划新方法*

夏红卫，文传军
(常州工学院 理学院，江苏 常州 213022)

摘要:本文提出一类新的序列二次规划方法来求解等式约束的非线性优化问题,方法不使用罚函数,避开了罚因子的选取对数值结果的影响,也不采用滤子技巧,去除了滤子方法中的恢复过程。在两个温和条件的假设下,步长的选取不需要目标函数和约束违反度的充分下降,扩大了算法的适用范围,证明了算法的全局收敛性。使用 Matlab 软件,编写了算法的程序,进行了数值试验,并与著名的优化软件 LANCELOT 比较,结果表明算法强健有效。

关键词:等式约束;序列二次规划;全局收敛

中图分类号:O224.2　　**文献标志码**:A　　**文章编号**:1672-6693(2014)02-0001-04

考虑如下一般的等式约束的非线性优化问题

$$\begin{aligned}&\min\ f(x)\\&\text{s. t. }\ h_j(x)=0,j\in\varepsilon\end{aligned}\tag{1}$$

其中 $x\in\mathbf{R}^n$,$\varepsilon=\{1,2,\cdots,m\}$,$f$,$h_j(j\in\varepsilon)$是定义在 $\mathbf{R}^n$ 上的二次连续可微函数。

序列二次规划方法(SQP)是求解问题(1)的一种典型方法。假定 x^k 是当前迭代点,SQP 方法求解如下的二次规划子问题

$$\begin{aligned}&\min\ (g^k)^{\mathrm{T}}d+\frac{1}{2}d^{\mathrm{T}}B_kd\\&\text{s. t. }\ h_j^k+(a_j^k)^{\mathrm{T}}d=0,j\in\varepsilon\end{aligned}\tag{2}$$

其中 $g^k=\nabla f(x^k)$,$h_j^k=h_j(x^k)$,$a_j^k=\nabla h_j(x^k)$,$B_k\in\mathbf{R}^{m\times n}$是 x^k 处拉格朗日函数的 Hessian 阵或其近似。取 d^k 是子问题(2)的解,新的迭代由线搜索 $x^{k+1}=x^k+\alpha_kd^k$ 产生,其中步长 $\alpha_k\in(0,1]$,它的选取使得带罚因子的罚函数满足充分下降条件。如果 α_j^k,$j\in\varepsilon$ 线性无关,B_k 正定,SQP 方法全局收敛到 KKT 点,尤其 B_k 选取适当,算法的收敛速度很快[1-3]。大量的数值试验表明,罚函数方法,包括 SQP 方法和内点法的有效性取决于罚因子的选取,尽管许多 SQP 方法和内点法使用自适应的罚因子的修正方法,但由于罚因子的初值选取的任意性,常常引起数值计算上的困难。为了避开罚因子的选取对实际计算的影响,Fletcher 和 Leyffer 引进滤子技巧,这是一种非线性优化的全局方法,数值结果表明带滤子技巧的 SQP 信赖域方法效果很好,Fletcher、Leyffer 和 Toint[4]还证明了滤子 SQP 信赖域方法的全局收敛性,从那以后,不断产生全局收敛的滤子方法,其中也有许多算法是局部超线性收敛的[5]。最近,Gould 和 Toint[6]提出求解问题(1)的另一种方法,该方法不使用罚函数和滤子,也具有全局收敛性。而 Ulbrich 提出自由罚函数的非单调信赖域方法。数值试验表明这两个方法效果不错;Yamashita 和 Yabe 还提出一种信赖域 SQP 方法来求解非线性的非负约束的优化问题,该方法也不使用罚函数和滤子,而通过信赖域技巧来产生新的迭代,求解几个不同的子问题来处理目标函数和约束条件的非线性性质[7-8]。受这些方法的启发,本文提出一类不带罚函数和滤子的序列二次规划新方法来求解等式约束优化问题(1)。算法通过线搜索产生新的迭代。在每次迭代,二次规划的线性化约束条件放松到满足两个温和条件,步长的选取不需要目标函数和约束违反度都充分下降。本文提出的方法有两个优点:首先不需要迭代序列的有界性,扩大了算法的搜索范围,更有可能找到最优解。同时,由于算法不使用滤子技巧,避开了滤子方法中的恢复过程,进一步简化计算,提高了算法的计算效率。在温和条件的假设下证明了算法在有限步迭代后一定能找到问题(1)的

* **收稿日期**:2013-04-09　　**修回日期**:2013-06-19　　**网络出版时间**:2014-03-10 19:23
资助项目:常州工学院校级基金(No. YN1312)
作者简介:夏红卫,男,讲师,研究方向为非线性优化,E-mail:xiahw@czu.cn
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20140310.1923.001.html

KKT点,或是问题(1)的FJ点,或是最小化约束违反度2范数的稳定点。通过控制线性化约束的精确度,使用二阶修正技巧,去掉了线性无关的约束要求。初步的数值试验表明,对于中小规模的优化问题,算法是强健而有效的。

1 算法

众所周知,即使初始问题(1)的解 x^* 是固定的。也就是约束条件线性无关,二阶充分条件成立,对于远离 x^* 的迭代点 x^k,向量 $a_j^k, j\in\varepsilon$ 仍然可能线性相关,在内点算法中,矩阵$[a_j^k, j\in\varepsilon]$可能会逐渐退化,线性相关也可能导致子问题(2)的解在 x^k 处是不可行的,而逐渐退化将导致子问题解的范数逐渐变大。这些缺点最终导致众多非线性优化方法的失败,因此,一直以来许多学者致力于放松子问题的约束条件的研究,并已经取得一定的成效[9-10]。在当前迭代点 x_k,Liu和Yuan提出求解如下二次规划的子问题

$$\begin{aligned}&\min\ (g^k)^{\mathrm T}d+\frac{1}{2}d^{\mathrm T}B_kd\\&\text{s. t. }\ A_k^{\mathrm T}d=A_k^{\mathrm T}d_p^k\end{aligned}\tag{3}$$

其中 $d_p^k\in\mathbf R^n$ 是 $\|h^k+A_k^{\mathrm T}d\|$ 的近似最小化[11],$A(x)=[\nabla h_1(x),\nabla h_2(x),\cdots,\nabla h_m(x)]$。且满足下列2个温和条件:1) $\|d_p^k\|\leqslant\kappa_1\|A_kh^k\|$,其中 $\kappa_1>0$ 是一个常数;2) 假如 $\|h^k\|\neq0$, $\|h^k\|-\|h^k+A_k^{\mathrm T}d_p^k\|\geqslant\kappa_2\|A_kh^k\|^2/\|h^k\|$,其中 $\kappa_2\in(0,1)$是一个常数。

当 B_k 在 $A_k^{\mathrm T}$ 的零空间上正定时,问题(3)有唯一解,设其解是 d_k。由 $h^k+A_k^{\mathrm T}d^k=h^k+A_k^{\mathrm T}d_p^k$ 得 $\|h^k+A_k^{\mathrm T}d^k\|\leqslant(1-\kappa_2\|A_kh^k\|^2/\|h^k\|)\|h^k\|$。假定 $\|h^k\|\neq0$,取 $d_p^k=-\alpha_k^cA_kh^k$,$\alpha_k^c=\arg\min\limits_{\alpha\in(0,1]}\|h^k-\alpha A_k^{\mathrm T}A_kh^k\|$。则 $\|h^k\|-\|h^k+A_k^{\mathrm T}d_p^k\|\geqslant\frac{1}{2}\min\{1,\theta_k\}\|A_kh^k\|^2/\|h^2\|$,其中 $\theta_k=\|A_kh^k\|^2/\|A_k^{\mathrm T}A_kh^k\|^2$。

构造拉格朗日函数 $L(x,\lambda)=f(x)+\lambda^{\mathrm T}h(x)$,则问题(1)的KKT条件为 $\nabla_xL(x^*,\lambda^*)=g^*+A_*\lambda^*=0,h^*=0$,$x^*\in\mathbf R^n$ 是一个KKT点,$\lambda^*\in\mathbf R^m$ 是相应的拉格朗日乘子。

对每一个 $x\in\mathbf R^n$,定义约束违反度的价值函数为

$$v(x)=\|h(x)\|\tag{4}$$

另外定义 $\varphi(x,d)=\|h(x)+A^{\mathrm T}(x)d\|-\|h(x)\|,d\in\mathbf R^n$。

由此提出如下的算法:步0,初值 $x_0\in\mathbf R^n$,$\sigma\in(0,1/2)$,常数 $\xi_1>0,\xi_2>0$,$\forall\varepsilon>0$,计算 g^0,h^0,v^0,A_0,B_0,取 $v_{\max}^0=0$,$r_0=0.9$,$k=0$;步1,如果 $\max(\|\nabla_xL(x^k,\lambda^k)\|,\|h^k\|)<\varepsilon$,且 $\|A_kh^k\|<\varepsilon\min\{\|h^k\|,1\}$。返回 x^k 作为解,否则转步2;步2,计算满足条件1)、2)的 $\|h^k+A_k^{\mathrm T}d\|$ 的近似最小解 d_p^k。求解QP子问题(3)得解为 d^k;步3,选取步长 $\alpha_k\in(0,1]$,满足

$$f(x^k+\alpha_kd^k)-f(x^k)\leqslant\min\{\sigma\alpha_k(g^k)^{\mathrm T}d^k,-\xi_1v(x^k+\alpha_kd^k)\tag{5}$$

$$v(x^k+\alpha_kd^k)\leqslant\max\{(r_k+1)/2,0.95\}v_{\max}^k,v_{\max}^k\neq0\tag{6}$$

或者

$$v(x^k+\alpha_kd^k)-v(x^k)\leqslant\min\{\sigma\alpha_k\varphi(x^k,d^k),-\xi_2\alpha_k^2\ \|d^k\|^2\}\tag{7}$$

令 $x^{k+1}=x^k+\alpha_kd^k$;步4,如果(7)式在 x^{k+1} 处成立,而在 x^k 处不成立,取 $v_{\max}^{k+1}=v^k$,否则 $v_{\max}^{k+1}=v_{\max}^k$;步5,计算 g^{k+1},h^{k+1},v^{k+1},A_{k+1},并且修正 B_k 到 B_{k+1}。如果(7)式成立,计算 $r_{k+1}=v^{k+1}/v^k$,否则 $r_{k+1}=r_k$。$k=k+1$,返步1。

说明 1) 若(5)、(6)式成立的迭代称为f-型迭代,此时目标函数值下降。(7)式成立的迭代称为h-型迭代,此时约束违反度减小。

2) 由算法可知 $r_k<(r_k+1)/2<1$,且 $(r_k+1)/2\geqslant1/2$。

3) $\max\{(r_k+1)/2,0.95\}=\begin{cases}(r_k+1)/2,\text{若 } r_k\geqslant0.9\\0.95,\text{否则}\end{cases}$,这样对(6)式接受的最大步长是0.95。

2 全局收敛性

为了讨论算法的收敛性,提出如下假设:1) 在 $\mathbf R^n$ 上,$f,h_j,j\in\varepsilon$ 二次连续可微;2) $\liminf\limits_{k\to\infty}f^k>-\infty$,水平集 $C=\{x\in\mathbf R^n;v(x)\leqslant v_0\}$ 有界;3) B_k 下有界且在 $A_k^{\mathrm T}$ 的零空间上正定;4) d_p^k 满足条件1)、2)。

引理1[11] 若假设1)~4)成立,则序列$\{d^k\}$,$\{A_k\lambda^k\}$有界。其中$\{d^k,\lambda^k\}$是问题(3)的一对KKT点。

引理2[11] 令 $\psi(\alpha)=c_1+c_2\alpha^2-c_3\alpha$,其中 $c_1\geqslant0,c_2>0,c_3>0$。则当 $c_1+c_2\alpha^2\leqslant c_3\alpha$ 时,有 $c_1\leqslant\frac{c_3^2}{8c_2}$,且 $\alpha\in$

$\left[\frac{2c_1}{c_3},\frac{c_3}{2c_2}\right]$。尤其是如果 $c_1=0$，则 $\psi(\alpha)\leqslant 0,\alpha\in\left[0,\frac{c_3}{2c_2}\right]$。

引理 3[11]　假设条件 1)～4)成立。对所有的 $k\geqslant 0$，有 $\|A_k h^k\|\geqslant\eta\|h^k\|$，其中 $\eta\in(0,1)$是一个常数。则存在一个常数 $\tilde{t}\in(0,1]$，使得 $1\geqslant\alpha_k\geqslant\tilde{t}$。也就是不论(5)、(6)式成立还是(7)式满足，α_k 在大于零处有界。

引理 4　假设引理 3 中的条件成立，且对充分大的 k，(7)式成立，则 $\lim\limits_{k\to\infty} v^k=0$ 且 $\lim\limits_{k\to\infty}\|d^k\|=0$。

证明　由引理 3，如果对 k 次迭代，(7)式成立，则 $v^{k+1}-v^k\leqslant\sigma\alpha_k\varphi(x^k,d^k)\leqslant-\alpha_k\kappa_2\eta^2\|h^k\|\leqslant-\eta'\|h^k\|$，$\eta'\in(0,1)$。由(4)式，$v^{k+1}\leqslant(1-\eta')v^k$。从而 $\lim\limits_{k\to\infty} v^k=0$。

再由(7)式，有 $v^{k+1}-v^k\leqslant-\xi_2\alpha_k^2\|d^k\|^2\leqslant-\xi_2\tilde{t}^2\|d^k\|^2\leqslant 0$。取极限后得 $\lim\limits_{k\to\infty}\|d^k\|=0$。　证毕

引理 5　假设引理 3 中的条件成立，且对充分大的 k，(5)式成立，则 $\lim\limits_{k\to\infty} v^k=0$ 且 $\lim\limits_{k\to\infty}\|d^k\|=0$。

证明　如果对充分大的 k，(5)式成立，则$\{f^k\}$在有限次迭代后单调下降。而 $f^{k+1}-f^k\leqslant-\xi_1 v^{k+1}$，两边取极限得 $\lim\limits_{k\to\infty} v^k=0$。

再证 $\lim\limits_{k\to\infty}\|d^k\|=0$。反证：假设存在一个无限指标集 K，使得 $\lim\limits_{\substack{k\to\infty\\k\in K}}\|d^k\|\neq 0$。因为 $\lim\limits_{k\to\infty} v^k=0$，则 $(g^k)^{\mathrm T}d^k=(d_p^k)^{\mathrm T}A_k\lambda^k-(d^k)^{\mathrm T}B_k d^k\leqslant-\gamma_0\|d^k\|^2$，其中 $\gamma_0>0,k\in K$，这样对充分大的 $k\in K$，有 $f^{k+1}-f^k\leqslant\sigma\alpha_k(g^k)^{\mathrm T}d^k\leqslant-\sigma\tilde{t}\gamma_0\|d^k\|^2$。结合$\{f^k\}$单调递减下有界，上式两边取极限，即得矛盾。于是 $\lim\limits_{k\to\infty}\|d^k\|=0$。　证毕

引理 6　假设引理 3 中的条件成立，则

$$\lim_{k\to\infty}\inf\|\nabla_x L(x^k,\lambda^k)\|=0\tag{8}$$

证明　如果 $\lim\limits_{k\to\infty}\|d^k\|=0$，因为 d^k 满足：$g^k+B_k d^k+A_k\lambda^k=0$，所以有 $\lim\limits_{k\to\infty}\|\nabla_x L(x^k,\lambda^k)\|=\lim\limits_{k\to\infty}\|B_k d^k\|=0$。如果所有的迭代在有限次终止后都是 h-型迭代或是 f-型迭代，则由引理 4、引理 5，可知(8)式成立。　证毕

定理　若假设 1)～4)成立，ξ_1,ξ_2 由引理 3 给定，则对算法中的充分小的 $\varepsilon>0$，算法经有限次迭代后终止在：或者是一个近似 KKT 点，或者是一个近似不可行的稳定点，或者是一个线性无关条件不满足的近似可行点。

证明　若算法不是有限次迭代后终止，分 2 种情况讨论。

1) 对充分大的 k 和常数 $\eta>0$，$\|A_k h^k\|\geqslant\eta\|h^k\|$。则对充分小的 $\varepsilon>0$，$\|\nabla_x L(x^k,\lambda^k)\|\leqslant\|h^k\|\leqslant\varepsilon$。由引理知，算法将终止，且终止点是问题(1)的一个近似 KKT 点。

2) 存在一个无限指标集 K，使得 $\lim\limits_{\substack{k\to\infty\\k\in K}}\frac{\|A_k h^k\|}{\|h^k\|}=0$。此时算法终止于 x^k，且 $\|A_k h^k\|\leqslant\varepsilon\min\{\|h^k\|,1\}$。对充分小的 $\varepsilon>0$，再分 2 种情况讨论。

a) $\|h^k\|$ 充分小，则 $\|h^k\|<\frac{\|A_k h^k\|}{\|h^k\|}\leqslant\varepsilon$，取 $w^k=\frac{h^k}{\|h^k\|}$，则 $\|A_k w^k\|\leqslant\varepsilon$。也就是说，存在一个单位向量 $w^k\in\mathbf{R}^m$ 使得 $\|A_k w^k\|$ 充分小。这样 x^k 是线性无关条件不成立的一个近似可行点。

b) $\|h^k\|$ 在大于零处有界，不妨设 $\|h^k\|>1$，则 $\|A_k h^k\|=O(\varepsilon)$，$x^k$ 就通过一个最小化约束违反度的近似驻点，这样 x^k 为问题(1)一个近似不可行驻点。

3 数值试验

为了检验算法的有效性，使用 Matlab 语言编写了算法的程序，进行了数值试验，测试问题取自 CUTE 测试问题集[12]。算法中 $A_k^{\mathrm T}$ 的零空间矩阵 W_k 直接由 Matlab 的零空间程序得到，子问题(3)的解由 Matlab 的双共轭梯度法获得，B_k 的修正由 BFGS 方法产生，拉格朗日乘子 λ^k 由 Matlab 的 LSQR 方法给定。其余参数如下：$\xi_1=10^{-10}$，$\xi_2=10^{-4}$，$\sigma=10^{-2}$，步长 α_k 由 Armijo 线搜索程序决定。为了比较算法的效率，把著名的优化软件 Lancelot 的运算结果列于后面。“n,m”分别代表问题的变量个数和约束条件个数。“f,h”“g,A”分别表示计算函数值和函数梯度的次数。计算结果列于表 1。

表 1　测试问题的计算结果

Tab. 1　The result of the test problems

测试问题	问题维数		新算法		LANCELOT	
	n	m	f,h	g,A	f,h	g,A
BT1	2	1	11	8	57	47
BT8	5	2	11	11	27	25
BT11	5	3	13	13	23	20
BT12	5	4	9	8	23	19
CBRATU3D	686	250	3	3	5	5
HATFLDG	25	25	25	7	24	20
HS7	2	1	14	11	58	42
HS26	3	1	36	26	33	31
HS39	4	2	57	41	674	639
HS42	4	2	11	9	13	13
HS50	5	3	25	15	19	19

从计算结果可以看出,对中小规模的问题,新算法是强健而有效的。对于大规模的优化计算,有待于进行进一步的试验。此外,算法的局部收敛性质还可以进行深入的分析。

参考文献:

[1] 袁亚湘. 非线性优化计算方法[M]. 北京:科学出版社,2008.
Yuan Y X.Nonlinear optimization method[M].Beijing:Science Press,2008.

[2] Nocedal J,Wright S. Numerical optimization [M]. New York:Springer-Verlag,1999.

[3] Sun W Y,Yuan Y X. Optimization Theory and Methods: Nonlinear Programming[M]. New York: Springer Press,2006.

[4] Fletcher R,Leyffer S,Toint P L. On the global convergence of a filter-SQP algorithm[J]. SIAM J Optim,2012,13(1):44-59.

[5] Ulbrich M,Ulbrich S,Vicente L N. A globally convergent primal-dual interior-point filter method for nonlinear programming[J]. Math Program,2004,100(2),379-410.

[6] Gould N I M,Toint P L. Nonlinear programming without a penalty function or a filter[J]. Math Program,2009,122(1):155-196.

[7] Ulbrich M,Ulbrich S. Non-monotone trust region methods for nonlinear equality constrained optimization without a penalty function[J]. Math Program,2003,95(1):103-135.

[8] Solodov M. Global convergence of an SQP method without boundedness assumptions on any of the iterative sequences [J]. Math. Program,2009,118(1):1-12.

[9] Andreani R,Birgin E G,Martinez J M,et al. Augmented Lagrangian methods under the constant positive linear dependence constraint qualification[J]. Mathematical Programming,2007,111(1):5-32

[10] Chen L F,Goldfarb D. Interior-point-penalty methods for nonlinear programming with strong global convergence properties[J].Mathematical Programming,2006,108(1):1-36.

[11] Liu X W,Yuan Y X. A null-space primal-dual interior-point algorithm for nonlinear optimization with nice convergence properties[J]. Mathematical Programming,2010,125(1):163-193.

[12] Bongartz I,Conn A R,Gould N I M,et al. CUTE:constrained and unconstrained testing environment[J]. ACM Tran Math Software,1995,21(1):123-160.

Operations Research and Cybernetics

A New Sequential Quadratic Programming Method for Nonlinear Equality-Constrained Optimization

XIA Hong-wei, WEN Chuan-jun

(School of Science, Changzhou Institute of Technology, Changzhou Jiangsu 213022, China)

Abstract: This paper presented a new sequence quadratic programming method for solving nonlinear equality constrained optimization problems. The new method did not use the penalty function, avoided the selection of penalty factor for the numerical results, and also did not use filter tips, removing the filter method in the recovery process. In two mild conditions under the assumption, step size did not need the objective function and the constraint violation for the sufficient descent, expanding the scope of application of the algorithm, the global convergence was proved. Using Matlab software, the preparation of the algorithm procedures, numerical experiments are carried out, and compared with the well-known optimization software LANCELOT, the results show that the algorithm is robust and effective.

Key words: equality-constrained; sequential quadratic programming method; global convergence

(责任编辑 黄 颖)