文章编号：0253-374X(2011)06-0914-05 DOI：10.3969/j.issn.0253-374x.2011.06.023

一种全局收敛的线搜索滤子 SQP 方法

金 中[1,2]，王玉青[3]

(1.同济大学 数学系，上海 200092；2.上海海事大学 数学系，上海 201306；
3.嘉兴学院 数学与信息工程学院，浙江 嘉兴 314001)

摘要：对于求解不等式约束优化问题，将线搜索和滤子方法相结合提出了一种新的线搜索滤子序列二次规划(filter SQP)方法.该方法克服了传统的SQP方法二次子问题不相容的困难，并利用滤子避免了罚函数的使用.同时在合理条件下证明了此方法具有全局收敛性质.

关键词：线搜索；滤子方法；序列二次规划；全局收敛性

中图分类号：O 221.2 **文献标识码**：A

A Global Convergent Line Search Filter SQP Method

JIN Zhong[1,2]，*WANG Yuqing*[3]

(1. Department of Mathematics，Tongji University，Shanghai 200092，China；2. Department of Mathematics，Sanghai Maritime University，Shanghai 201306，China；3. College of Mathematics and Information Engineering，Jiaxing University，Jiaxing 314001，China)

Abstract：The paper presents a line search filter sequential quadratic programming (SQP) method for inequality constrained optimization. Compared to traditional SQP methods，the advantages are that the quadratic programming (QP) subproblem is always consistent and the penalty function is not required by filter strategy. Under some mild conditions the global convergence can be induced.

Key words：line search；filter method；sequential quadratic programming (SQP)；global convergence

1 引言

考虑如下的非线性不等式约束优化问题(P)：

$$\begin{aligned}&\min\quad f(\boldsymbol{x})\\&\text{s.t.}\quad g_j(\boldsymbol{x})\leqslant 0,\ j\in I=\{1,2,\cdots,m\}\end{aligned}\tag{1}$$

其中 $x\in\mathbf{R}^n$，$f:\mathbf{R}^n\to\mathbf{R}$，$g_j(j\in I):\mathbf{R}^n\to\mathbf{R}$ 是连续可微函数.

序列二次规划(SQP)方法是求解问题(P)的最有效方法之一，由于其好的收敛性质，已得到了广泛深入的研究，关于SQP方法的一个好的综述见文献[1].SQP方法通过求解如下的二次规划得到搜索方向：

$$\begin{aligned}&\min\quad \nabla f(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}+\frac{1}{2}\boldsymbol{d}^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}\\&\text{s.t.}\quad g_j(\boldsymbol{x}_k)+\nabla g_j(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}\leqslant 0,\\&\quad j\in I=\{1,2,\cdots,m\}\end{aligned}\tag{2}$$

其中 $\boldsymbol{B}_k\in\mathbf{R}^{n\times n}$ 是一对称正定矩阵.但是此QP子问题可能是不相容的，即它的可行域可能为空集，这在实际应用中将会产生困难.而且，传统的SQP方法使用了一个罚函数作为价值函数，其中的罚因子与最优解处的信息相关，即常存在一个临界值，若罚因子小于此临界值，则罚函数在问题(P)的解处没有局部最小值.而此临界值事先是不知道的.在计算中，如果罚因子选择得太小，则可能得到问题(P)的不可行点，甚至可能会造成惩罚项的无限增大.另一方面，如果罚因子选择得太大，则会降低目标函数的作用和影响，导致收敛过慢.

对于子问题不相容的困难，可以通过“大M”方法[1-3]，把二次子问题(2)替换成如下问题：

$$\begin{aligned}Q(\boldsymbol{x}_k,\boldsymbol{B}_k,\boldsymbol{b}_k):\quad&\min \nabla f(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}+\frac{1}{2}\boldsymbol{d}^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}+b_k t\\&\text{s.t.}\quad g_j(\boldsymbol{x}_k)+\nabla g_j(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}\leqslant te\\&j\in I,t\geqslant 0.\end{aligned}\tag{3}$$

其中 b_k 是一正参数.显然，此子问题总是相容的而且当 $\boldsymbol{B}_k$ 是半正定矩阵时是凸规划.

而罚因子选取的难题，Fletcher 和 Leyffer[4] 提

收稿日期：2010-03-01
基金项目：国家自然科学基金资助项目(10771162)
第一作者：金 中(1981—)，男，博士生，主要研究方向为数学规划.E-mail:asdjz1234@163.com

出了滤子(filter)的新思想,产生了一种结合信赖域的新算法.这种新技术可以替代传统的价值函数方法,避免了罚因子的选取.近年来,滤子方法及其收敛性质得到了极大的关注,除了信赖域方法,还有很多其他的方法和filter方法结合使用的,比如线性搜索方法、内点法、QP-free方法、解决非光滑优化问题的bundle方法和derivative-free方法[5-9].其中,文献[6-7]将线搜索方法和滤子方法相结合,求解只带有等式约束的优化问题,得到了好的收敛性质.

本文将线搜索方法和滤子方法相结合来求解不等式约束优化问题,得到了一种新的线搜索滤子SQP方法.这种方法可以看成是文献[6-7]中方法的一种推广.而且此方法采用了二次子问题(3),克服了SQP二次子问题不相容的困难.并在合理条件下证明了算法的全局收敛性质.

2 预备知识

首先,定义约束违反度函数为

$$\theta(\boldsymbol{x}_k)=\max\{g_j(\boldsymbol{x}_k),j\in I;0\}$$

众所周知,在线搜索方法中,若当前迭代点为$\boldsymbol{x}_k$,搜索方向$\boldsymbol{d}_k$确定之后,再通过某种方式给出一个步长$\alpha_k\in(0,1]$,那么下一个迭代点为$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{d}_k$.

本文的搜索方向$\boldsymbol{d}_k$由子问题(3)给出.令$(\boldsymbol{d}_k,t_k)\in(\mathbf{R}^n,\mathbf{R}^1)$是子问题(3)的解.存在Lagrange乘子$(\boldsymbol{\lambda}_k,\upsilon_k)\in(\mathbf{R}^n,\mathbf{R}^1)$满足如下的$Q(\boldsymbol{x}_k,\boldsymbol{B}_k,b_k)$的KKT(Karush-Kuhn-Tucker)系统:

$$\begin{cases}g_j(\boldsymbol{x}_k)+\nabla g_j(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k\leqslant t_k e\ (j\in L_k),t_k\geqslant 0,\\ \nabla f(\boldsymbol{x}_k)+\boldsymbol{B}_k\boldsymbol{d}_k+\sum\limits_{j\in I}\nabla g_j(\boldsymbol{x}_k)(\boldsymbol{\lambda}_k)_j=0,\\ b_k-\|\boldsymbol{\lambda}_k\|_1-\upsilon_k=0,\\ (\boldsymbol{\lambda}_k)_j(g_j(\boldsymbol{x}_k)+\nabla g_j(\boldsymbol{x}_k)^{\mathrm{T}}\bar{\boldsymbol{d}}_k-t_k e)=0\\ \quad(j\in I),\upsilon_k t_k=0,\\ \boldsymbol{\lambda}_k\geqslant 0,\ \upsilon_k\geqslant 0.\end{cases}\tag{4}$$

考虑采用回溯技术(backtracking line search procedure)得出步长,即依次检验一个下降的步长序列$\alpha_{k,l}\in(0,1]$直至试验点$\boldsymbol{x}(\alpha_{k,l})=\boldsymbol{x}_k+\alpha_{k,l}\boldsymbol{d}_k$能够满足某种接受准则.在这里,滤子接受条件就起到了接受准则的作用.滤子方法的基本思想是把目标函数和约束违反度函数分开处理,如果某个迭代步使得目标函数下降,或者使得约束违反度函数下降,那么就接受这一步.

滤子接受条件:如果式(5)成立,步长$\alpha_{k,l}$就能够被当前点$\boldsymbol{x}_k$接受.

$$\theta(\boldsymbol{x}_k(\alpha_{k,l}))\leqslant(1-r_\theta)\theta(\boldsymbol{x}_k),\tag{5}$$

或者

$$f(\boldsymbol{x}_k(\alpha_{k,l}))\leqslant f(\boldsymbol{x}_k)-r_f\theta(\boldsymbol{x}_k),\tag{6}$$

其中$r_\theta,r_f\in(0,1)$.

但是仅仅依靠以上的滤子接受条件可能会产生一个仅满足条件(5)的点列,从而导致算法收敛到一个非优的可行点.为了避免这种情况发生,文献[6-7]采用了以下的f-型转换条件.

f-型转换条件:

$$m_k(\alpha_{k,l})<0\quad 且$$

$$[-m_k(\alpha_{k,l})]^{S_f}[\alpha_{k,l}]^{1-S_f}>\delta[\theta(\boldsymbol{x}_k)]^{S_\theta},\tag{7}$$

其中$\delta>0,s_\theta>1,s_f\geqslant 1$是固定常数,$m_k(\alpha_{k,l})=\alpha\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k$.

如果条件(7)成立,那么$\boldsymbol{d}_k$是目标函数的下降方向,此时要求步长$\alpha_{k,l}$满足Armijo形条件

$$f(\boldsymbol{x}_k(\alpha_{k,l}))<f(\boldsymbol{x}_k)+\eta_f m_k(\alpha_{k,l}),\tag{8}$$

其中$\eta_f\in(0,\frac{1}{2})$是一个固定常数.

不难看出,试验步长$\alpha_{k,l},l=1,\cdots,\hat{l}$仅满足f-型转换条件(7)而不满足Armijo形条件(8)时,当步长缩小到一定程度后,条件(7)中的后一个式子将不再成立,因此算法将转回滤子接受条件(5)和(6).

满足f-型转换条件(7)的步长称为f-型步长.相应地,如果在第k步最终被接受的步长为f-型步长,则称第k步为f-型迭代.另外,在算法中一般令初始滤子为空集或对其约束违反度给定一上界.滤子的更新公式为(见文献[6-7,9]):

$$\mathscr{F}_{k+1}=\mathscr{F}_k\cup\{(\theta,f)\in\mathbf{R}^2:\theta\geqslant(1-r_\theta)\theta(\boldsymbol{x}_k)\ 且\ f\geqslant f(\boldsymbol{x}_k)-r_f\theta(\boldsymbol{x}_k)\},\tag{9}$$

文献[6-7]给定了在等式约束下每一步中线搜索步长的下界,在式(6)和目标函数线性估计的基础上,本文类似地给出不等式约束下进行线搜索时步长的下界为

$$\alpha_k^{\min}=r_\alpha\min\left\{r_\theta,\frac{r_f\theta(\boldsymbol{x}_k)}{-\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k}\right\}.$$

其中因子$r_\alpha\in(0,1]$.

3 算法

现在给出求解问题(1)的算法1.

算法1

步骤1　初始化:给定初始点$\boldsymbol{x}_1\in\mathbf{R}^n$,$\mathscr{F}_1=\varnothing$,参数$b_1>0$和一个对称正定矩阵$\boldsymbol{B}_1\in\mathbf{R}^{n\times n}$.选择参

数 $\rho\in(0,1)$，$\delta_1,\delta_2>0$，$r_\theta,r_f\in(0,1)$，$\delta>0$，$s_\theta>1$，$s_f\geqslant1$，$r_\alpha\in(0,1]$. 令 $k=1$.

步骤2　求解QP子问题(3)得解$(\boldsymbol{d}_k,t_k)$以及对应的Lagrange乘子$(\boldsymbol{\lambda}_k,v_k)$. 如果$\boldsymbol{d}_k=0$且$t_k=0$，算法停止.

步骤3　如果$\boldsymbol{d}_k=0$但$t_k\neq0$，则令$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k$转步骤8.

步骤4　如果$\boldsymbol{d}_k\neq0$且$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k\geqslant0$，转可行性恢复阶段步骤10.

步骤5　如果$\boldsymbol{d}_k\neq0$且$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k<0$，进行线搜索：

步骤5.1　线搜索初始化. 令$\alpha_{k,l}=1$，$l=1$.

步骤5.2　计算求得一个新试验点. 如果试验步长$\alpha_{k,l}<\alpha_k^{\min}$，转可行性恢复阶段步骤10. 否则，计算新的试验点$\boldsymbol{x}_k(\alpha_{k,l})=\boldsymbol{x}_k+\alpha_{k,l}\boldsymbol{d}_k$，并检验其能否被滤子接受，如果$\boldsymbol{x}_k(\alpha_{k,l})\in\mathscr{F}_k$，拒绝该试验步长并转步骤5.4.

步骤5.3　检验是否满足接受准则.

步骤5.3.1　如果步长是f-型，接受该试验步长转步骤6.

步骤5.3.2　如果步长不是f-型，而滤子接受条件成立，则也接受该试验步长转步骤6.

步骤5.4　选择一新的试验步长$\alpha_{k,l+1}=\rho\alpha_{k,l}$. 令$l=l+1$并转步骤5.2.

步骤6　接受试验点. 令$\alpha_k=\alpha_{k,l}$，$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k(\alpha_k)$.

步骤7　更新滤子. 如果k不是f-型迭代，利用式(9)增加滤子；否则保持滤子不变，即令$\mathscr{F}_{k+1}=\mathscr{F}_k$.

步骤8　更新参数. 计算a_k

$$a_k=\min\{\|\boldsymbol{d}_k\|^{-1},\|\boldsymbol{\lambda}_k\|_1+\delta_1\}.\tag{10}$$

令

$$b_{k+1}=\begin{cases}b_k, & 若\ b_k\geqslant a_k,\\ b_k+\delta_2, & 否则.\end{cases}\tag{11}$$

步骤9　使用适当的拟牛顿公式更新$\boldsymbol{B}_k$得到$\boldsymbol{B}_{k+1}$. 令$k=k+1$转步骤2.

步骤10　可行性恢复阶段. 通过减小当前的约束违背函数θ求得新迭代点$\boldsymbol{x}_{k+1}$，使得$\boldsymbol{x}_{k+1}$满足充分下降条件(5)并被滤子接受，即$(\theta(\boldsymbol{x}_{k+1}),f(\boldsymbol{x}_{k+1}))\notin\mathscr{F}_k$，接着使用式(9)更新滤子，令$\boldsymbol{B}_{k+1}=B_1$，$k=k+1$，转步骤2.

注1：可行性恢复阶段在滤子方法中普遍采用，它的目的是得到比当前约束违背函数值小的不可行点，见文献[6-7,9].

注2：滤子的更新机制使得对所有的k，有$(\theta(\boldsymbol{x}_k),f(\boldsymbol{x}_k))\notin\mathscr{F}_k$.

4　算法的可执行性

本文中，假设以下条件成立.

假设：

A1　可行集非空，即$\boldsymbol{X}=\{\boldsymbol{x}\in\mathbf{R}^n\mid g_j(\boldsymbol{x})\leqslant0,j\in I\}\neq\varnothing$.

A2　函数$f(\boldsymbol{x})$和$g_j(\boldsymbol{x})$，$j=1,2,\cdots,m$二次连续可微.

A3　对任意$\boldsymbol{x}\in\boldsymbol{X}$，向量$\{\nabla g_j(\boldsymbol{x}),j\in I(\boldsymbol{x})\}$线性无关，其中$I(\boldsymbol{x})=\{j\in I\mid g_j(\boldsymbol{x})=\theta(\boldsymbol{x})\}$.

A4　算法1产生的序列$\{\boldsymbol{x}^k\}$有界，且对任意k和$\boldsymbol{d}\in\mathbf{R}^n$，存在两个常数$b\geqslant a>0$使得矩阵序列$\{\boldsymbol{B}_k\}$满足$a\|\boldsymbol{d}\|^2\leqslant\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}\leqslant b\|\boldsymbol{d}\|^2$.

下面第一个引理表明算法1中的步骤3和步骤8之间不存在死循环.

引理1　设$\{\boldsymbol{x}_k\}$和$\{\boldsymbol{d}_k\}$是由算法1产生的序列. 如果$\boldsymbol{d}_k=0$，$t_k>0$，则存在某个正整数h使得$\boldsymbol{d}_{k+h}\neq0$.

证明　反证法. 假设结论不成立，对任意的$h\geqslant1$，有$\boldsymbol{d}_{k+h}=0$，$t_{k+h}>0$. 那么$\boldsymbol{x}_{x+h}=\boldsymbol{x}_k$，$t_{t+h}=\theta(\boldsymbol{x}_{x+h})=\theta(\boldsymbol{x}_k)=t_k>0$，$v_{k+h}=0$. 由式(4)和式(10)可知

$$b_{k+h}=\|\boldsymbol{\lambda}_{k+h}\|_1<a_{k+h}.$$

再结合式(11)知对任意的$h\geqslant1$，$b_{k+h+1}=b_{k+h}+\delta_2$，则

$$b_{k+h}\to+\infty,当\ h\to\infty\ 时\ \|\boldsymbol{\lambda}_{k+h}\|_1\to+\infty$$

又因对任意的$h\geqslant1$，有$\boldsymbol{x}_{k+h}=\boldsymbol{x}_k$，从而定义$\boldsymbol{x}_{k+h}=\bar{\boldsymbol{x}}$. 因此

$$0<t_k=\lim_{h\to\infty}t_{k+h}=\lim_{h\to\infty}\theta(\boldsymbol{x}_{k+h})=\theta(\bar{\boldsymbol{x}})$$

且

$$\lim_{h\to\infty}(g_i(\boldsymbol{x}_{x+h})+\nabla g_i(\boldsymbol{x}_{k+h})^{\mathrm{T}}\boldsymbol{d}_{k+h})=$$

$$\lim_{h\to\infty}g_i(\boldsymbol{x}_{k+h})=g_i(\bar{\boldsymbol{x}})<\theta(\bar{\boldsymbol{x}})=t_k,\forall i\notin I(\bar{\boldsymbol{x}}).$$

对任意的$h\geqslant1$易得

$$g_i(\boldsymbol{x}_{k+h})+\nabla g_i(\boldsymbol{x}_{k+h})^{\mathrm{T}}\boldsymbol{d}_{k+h}<t_{k+h},\quad\forall i\notin I(\bar{\boldsymbol{x}}).\tag{12}$$

对任意的$h\geqslant1$，又由式(4)，(12)可得$(\boldsymbol{\lambda}_{k+h})_i=0$，$\forall i\notin I(\bar{\boldsymbol{x}})$. 从而有

$$\nabla f_{k+h}+\sum_{j\in I(\bar{\boldsymbol{x}})}\nabla g_j(\boldsymbol{x}_{k+h})(\boldsymbol{\lambda}_{k+h})_j=0.\tag{13}$$

不失一般性，假设$\lim_{h\to+\infty}\boldsymbol{\lambda}_{k+h}/\|\boldsymbol{\lambda}_{k+h}\|_1=\hat{\boldsymbol{\lambda}}$. 对式(13)两端除以$\|\boldsymbol{\lambda}_{k+h}\|_1$并取极限$h\to+\infty$，得

$$\sum_{j\in I(\bar{x})}\hat{\boldsymbol{\lambda}}_j\nabla g_j(\bar{\boldsymbol{x}})=0,\hat{\boldsymbol{\lambda}}_j\geqslant 0,j\in I(\bar{x}).$$

那么至少存在一个 $j\in I(\bar{x})$ 满足 $\hat{\boldsymbol{\lambda}}_j>0$,与假设 A3 矛盾,结论得证.

类似引理 1 的证明可得如下引理 2.它表明参数 b_k 仅在有限次增加.

引理 2　设$\{\boldsymbol{x}_k\}$是由算法 1 产生的无穷序列,则存在整数 k_0使得对所有的 $k\geqslant k_0$,有 $b_k=b_0$.其中,$b_0=\max\{b_k\}$.

5　算法的全局收敛性

为描述简便,用 $\mathscr{A}\subseteq\mathbb{N}$ 表示满足滤子增加的迭代步的指标集,即

$$\mathscr{F}_k\subsetneq\mathscr{F}_{k+1}\Leftrightarrow k\in\mathscr{A}.$$

引理 3　如果 $\theta(\boldsymbol{x}_k)=0$,则对所有的 $\alpha\in(0,1]$,$m_k(\alpha)<0$.而且对任意 k,有 $\Theta_k:=\min\{\theta:(\theta,f)\in\mathscr{F}_k\}>0$.

证明　如果 $\theta(\boldsymbol{x}_k)=0$,则$(0,\cdots,0)^{\mathrm{T}}\in R^{n+1}$位于 QP 子问题(3) 的可行域内,因此

$$m_k(\alpha)/\alpha=\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k\leqslant -\frac{1}{2}(\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k-b_kt_k<0.$$

第 2 个结论证明可见文献[6]中引理 4.

定理 1　设所有假设成立,则

$$\lim_{k\to\infty}\theta(\boldsymbol{x}_k)=0. \tag{14}$$

证明　参见文献[6]中定理 1.

引理 4　如果算法 1 产生一无穷序列$\{\boldsymbol{x}_k\}$,则序列$\{\boldsymbol{d}_k\}$有界,即存在常数 $M_d>0$ 使得

$$\|\boldsymbol{d}_k\|\leqslant M_d. \tag{15}$$

证明　易知$(0,\cdots,0,\theta(\boldsymbol{x}_k))^{\mathrm{T}}\in\mathbf{R}^{n+1}$位于 QP 子问题(3) 的可行域内,则

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k+\frac{1}{2}(\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k+b_kt_k\leqslant b_k\theta(\boldsymbol{x}_k).$$

由定理 1 知存在常数 M_θ使得$\theta(\boldsymbol{x}_k)\leqslant M_\theta$,那么

$$-\|\nabla f(\boldsymbol{x}_k)\|\cdot\|\boldsymbol{d}_k\|+\frac{1}{2}\alpha\|\boldsymbol{d}_k\|^2\leqslant \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k+ \frac{1}{2}(\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}_k+b_kt_k\leqslant b_k\theta(\boldsymbol{x}_k)\leqslant b_0M_\theta,$$

这说明$\{\boldsymbol{d}_k\}$有界,即存在常数 $M_d>0$ 使得

$$\|\boldsymbol{d}_k\|\leqslant M_d.$$

引理 5　如果有一子序列$\{\boldsymbol{x}_{k_i}\}$满足$\|\boldsymbol{d}_{k_i}\|\geqslant\varepsilon$,其中 $\varepsilon>0$ 取值与 i 无关,则存在常数 $\varepsilon_1,\varepsilon_2>0$ 使得

$$\theta(\boldsymbol{x}_{k_i})\leqslant\varepsilon_1\Rightarrow m_{k_i}(\alpha)\leqslant-\varepsilon_2\alpha$$

对所有 i 和$\alpha\in(0,1]$成立.

证明　$(0,\cdots,0,\theta(\boldsymbol{x}_k))^{\mathrm{T}}\in\mathbf{R}^{n+1}$位于 QP 子问题(3)的可行域内,那么

$$\nabla f(\boldsymbol{x}_{k_i})^{\mathrm{T}}\boldsymbol{d}_{k_i}+\frac{1}{2}(\boldsymbol{d}_{k_i})^{\mathrm{T}}\boldsymbol{B}_{k_i}\boldsymbol{d}_{k_i}+b_{k_i}t_{k_i}<b_{k_i}\theta(\boldsymbol{x}_{k_i}).$$

令 $\varepsilon_1:=\dfrac{a\varepsilon^2}{4b_0}$,则 $b_{k_i}\theta(\boldsymbol{x}_{k_i})\leqslant\dfrac{ab_{k_i}\varepsilon^2}{4b_0}\leqslant\dfrac{1}{4}(\boldsymbol{d}_{k_i})^{\mathrm{T}}\boldsymbol{B}_{k_i}\boldsymbol{d}_{k_i}$,因此

$$\begin{aligned}m_{k_i}(\alpha)/\alpha&=\nabla f(\boldsymbol{x}_{k_i})^{\mathrm{T}}\boldsymbol{d}_{k_i}<-\frac{1}{2}(\boldsymbol{d}_{k_i})^{\mathrm{T}}\boldsymbol{B}_{k_i}\boldsymbol{d}_{k_i}-\\&b_{k_i}t_{k_i}+b_{k_i}\theta(\boldsymbol{x}_{k_i})<-\frac{1}{2}(\boldsymbol{d}_{k_i})^{\mathrm{T}}\boldsymbol{B}_{k_i}\boldsymbol{d}_{k_i}\\&+b_{k_i}\theta(\boldsymbol{x}_{k_i})\leqslant-\frac{1}{4}(\boldsymbol{d}_{k_i})^{\mathrm{T}}\boldsymbol{B}_{k_i}\boldsymbol{d}_{k_i}\leqslant\\&-\frac{a\varepsilon^2}{4},\end{aligned}$$

令 $\varepsilon_2:=\dfrac{a\varepsilon^2}{4}$,则 $m_{k_i}(\alpha)\leqslant-\varepsilon_2\alpha$.

引理 6　假设滤子仅在有限步增加,即$|\mathscr{A}|<\infty$,则$\lim\limits_{k\to\infty}\boldsymbol{d}_k=0$.

证明　根据定理 1,引理 5 及引理 3,证明类似于文献[6]中引理 8.

引理 7　存在常数 $C_f>0$ 使得对 $\alpha\in(0,1]$,

$$f(\boldsymbol{x}_k+\alpha\boldsymbol{d}_k)-f(\boldsymbol{x}_k)-m_k(\alpha)\leqslant C_f\alpha^2\|\boldsymbol{d}_k\|^2. \tag{16}$$

证明　结论由二阶泰勒公式即可得证.

引理 8　如果有一子序列$\{\boldsymbol{x}_{k_i}\}$满足 $m_{k_i}(\alpha)\leqslant-\alpha\varepsilon_2$,其中常数 $\varepsilon_2>0$ 取值与 k_i无关,任意的 $\alpha\in(0,1]$,则存在常数 $\bar{\alpha}>0$ 使得

$$f(\boldsymbol{x}_{k_i}+\alpha\boldsymbol{d}_{k_i})-f(\boldsymbol{x}_{k_i})\leqslant\eta_f m_{k_i}(\alpha)$$

对所有的 k_i和$\alpha\leqslant\bar{\alpha}$成立.

证明　设 M_d和C_f是引理 4 和引理 7 中定义的常数.由引理 7 知,对任意的 $\alpha\leqslant\bar{\alpha}$,$\bar{\alpha}:=\dfrac{(1-\eta_f)\varepsilon_2}{C_fM_d^2}$,

$$f(\boldsymbol{x}_{k_i}+\alpha\boldsymbol{d}_{k_i})-f(\boldsymbol{x}_{k_i})-m_{k_i}(\alpha)\leqslant C_f\alpha^2\|\boldsymbol{d}_{k_i}\|^2\leqslant\alpha(1-\eta_f)\varepsilon_2\leqslant-(1-\eta_f)m_{k_i}(\alpha),$$

即

$$f(\boldsymbol{x}_{k_i}+\alpha\boldsymbol{d}_{k_i})-f(\boldsymbol{x}_{k_i})\leqslant\eta_f m_{k_i}(\alpha).$$

易证下面引理 9 和引理 10(见文献[6])

引理 9　如果有一子序列$\{\boldsymbol{x}_{k_i}\}$满足 $m_{k_i}(\alpha)\leqslant-\alpha\varepsilon_2$,其中 $\varepsilon_2>0$ 取值与 k_i无关,任意的 $\alpha\in(0,1]$,则存在常数 $C_1,C_2>0$ 使得对所有的 k_i和$\alpha\leqslant$

$\min\{C_1, C_2\theta(\boldsymbol{x}_{k_i})\}$,

$(\theta(\boldsymbol{x}_{k_i} + \alpha \boldsymbol{d}_{k_i}), f(\boldsymbol{x}_{k_i} + \alpha \boldsymbol{d}_{k_i})) \notin \mathscr{F}_{k_i}$.

引理 10 如果有一子序列$\{\boldsymbol{x}_{k_i}\}$满足$\|\boldsymbol{d}_{k_i}\| \geqslant \varepsilon$,其中常数$\varepsilon > 0$取值与$k_i$无关,则存在$k \in \mathbb{N}$使得滤子在第$k$步不再增加,即$k \notin \mathscr{A}$.

引理 11 如果有一子序列$\{\boldsymbol{x}_{k_i}\}$满足$\boldsymbol{d}_{k_i} \to 0$,则当$k_i$充分大时,有$t_{k_i} = 0$.

证明 由$\boldsymbol{d}_{k_i} \to 0$结合引理2可得当$k$充分大时,

$$b_{k_i} \geqslant a_{k_i} = \min\{\|\boldsymbol{d}_{k_i}\|^{-1}, \|\boldsymbol{\lambda}_{k_i}\|_1 + \delta_1\} = \|\boldsymbol{\lambda}_{k_i}\|_1 + \delta_1. \tag{17}$$

又由式(4)得

$$b_{k_i} = \|\boldsymbol{\lambda}_{k_i}\|_1 + v_{k_i}. \tag{18}$$

因此当k_i充分大时,$v_{k_i} > 0$.此时,$t_{k_i} = 0$.

定理 2 设所有假设成立,则

$$\lim_{k\to\infty} \theta(\boldsymbol{x}_k) = 0, \tag{19}$$

且存在子序列$\mathfrak{K}$使得

$$\lim_{k\to\infty, k\in\mathfrak{K}} \|\boldsymbol{d}_k\| = \lim_{k\to\infty, k\in\mathfrak{K}} t_k = 0. \tag{20}$$

也就是说,所有的极限点都是可行点而且存在序列$\{\boldsymbol{x}_k\}$的一个聚点$\boldsymbol{x}_*$是不等式约束问题(1)的KKT点.

证明 由定理1知式(19)显然成立.下面分两种情况证明式(20):

(i) 滤子仅在有限步增加时,由引理9和引理11结论成立.

(ii) 存在子序列$\{\boldsymbol{x}_{k_i}\}$使得对所有的i,$k_i \in \mathscr{A}$.现假设$\lim_i \sup \boldsymbol{d}_{k_i} > 0$,则存在$\boldsymbol{x}_{k_i}$的子序列$\{\boldsymbol{x}_{k_{i_j}}\}$($\mathfrak{K}$)和常数$\varepsilon > 0$使得$\lim_{j\to\infty} \theta(\boldsymbol{x}_{k_{i_j}}) = 0$且对所有的$k_{i_j}$,$\|\boldsymbol{d}_{k_{i_j}}\| > \varepsilon$.对序列$\{\boldsymbol{x}_{i_j}\}$应用引理10得,存在一迭代步$k_{i_j}$,在该步滤子并不增加,即$k_{i_j} \notin \mathscr{A}$.这与$\boldsymbol{x}_{k_i}$的选取矛盾,因此$\lim_{k\to\infty, k\in\mathfrak{K}} \boldsymbol{d}_k = 0$.由引理11得式(20)成立.

6 数值实验

对算法1给出了一些数值试验.算例来自文献[10],其结果见表1(基于篇幅,只列出其中一部分),表1中的结果说明本文的算法是有效的.

表1 数值结果

Tab.1 Numerical results

文献[10]中算例的号码	迭代的次数	计算目标函数值和约束的次数
43	16	35
76	13	50
86	6	24
100	17	45

参考文献:

[1] Boggs P T, Tolle J W. Sequential quadratic programming[C]// Acta Numerica. Cambridge: Cambridge University Press, 1995: 1-51.

[2] Liu T, Zeng J. An SQP algorithm with cautious updating criteria for nonlinear degenerate problems[J]. Acta Mathematicae Applicatae Sinica, English Series, 2009, 25(1): 33.

[3] Liu T, Li D. Convergence of the BFGS-SQP method for degenerate probloms[J]. Numerical Functional Analysis and Optimisition, 2007, 28(7-8): 927.

[4] Fletcher R, Leyffer S. Nonlinear programming without a penalty function[J]. Math Program, 2002, 91: 239.

[5] Fletcher R, Leyffer S. A bundle filter method for nonsmooth nonlinear optimization[R]. Dundee: University of Dundee. Department of Mathematics, 1999.

[6] Wätcher A, Biogle L T. Line search filter methods for nonlinear programming: motivation and global convergence[J]. SIAM J Optim, 2005, 16: 1.

[7] Wätcher A, Bigole L T. Line search filter methods for nonlinear programming: local convergence[J]. SIAM J Optim, 2005, 16: 32.

[8] Su K, Pu D. Active set sequential quadratic programing filter method[J]. Journal of Tongji University: Natural Science, 2008, 36(5): 690.

[9] Fletcher R, Gould N I M, Leyffer S, et al. Global convergence of a trust region SQP-filter algorithm for general nonlinear programming[J]. SIAM J Optim, 2002, 13: 635.

[10] Schittkowski K. More test examples for nonlinear programming codes[M]. Berlin: Springer-Verlag, 1988.