DOI：10.13718/j.cnki.xdzk.2018.11.011

修正共轭投影梯度滤子法①

王祥玲[1]，　左双勇[1]，　朱志斌[2]

1. 宜春幼儿师范高等专科学校 初等教育学院，江西 宜春 330814；
2. 桂林电子科技大学 数学与计算科学学院，广西 桂林 541004

摘要：利用共轭投影梯度技术，结合滤子算法的思想，通过修正搜索方向，建立了一个新的共轭投影梯度滤子算法. 该算法不需要求解二次规划子问题，而且能有效避免常规滤子算法中的恢复算法. 在适当的条件下，证明了算法的全局收敛性.

关　键　词：共轭投影梯度；滤子；非线性规划；全局收敛性

中图分类号：O221.2　　**文献标志码**：A　　**文章编号**：1673－9868(2018)11－0074－07

本文考虑如下非线性规划问题：

$$\min_{x \in \mathbb{R}^n} f(\boldsymbol{x})$$
$$\text{s.t.}\ \boldsymbol{c}_j(\boldsymbol{x}) \leqslant 0,\ j \in I = \{1, 2, \cdots, m\} \tag{1}$$

其中：$\boldsymbol{x} \in \mathbb{R}^n$，$f: \mathbb{R}^n \longrightarrow \mathbb{R}$，$\boldsymbol{c}_j(j \in I): \mathbb{R}^n \longrightarrow \mathbb{R}$ 是二次连续可微的. 可行集 $X = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{c}_j(\boldsymbol{x}) \leqslant 0, j \in I\}$，有效集 $I(\boldsymbol{x}) = \{j \in I \mid \boldsymbol{c}_j(\boldsymbol{x}) = 0\}$.

非线性规划问题广泛应用于工程和社会生活的各个领域，并在其中起着举足轻重的作用. 非线性规划问题的求解主要是最优性和可行性. 为了使这两个规则都能满足，算法需要在迭代过程中每一步都保持这两种规则的平衡. 目前关于非线性规划问题的求解方法多种多样[1-12]. 近年来，滤子法引起很多学者的密切关注. 滤子法最早由 Fletecher 和 Leyffer 提出[1]，它能有效避免选择罚函数的困难且具有良好的数值结果，从而近年来被广泛应用于各种优化问题的求解中[2-6]. 共轭投影梯度算法是解决非线性规划问题的另外一种有效算法[7]. 文献[8]将 SQP 方法和广义投影技巧结合，提出一个具有显式主搜索方向的共轭投影梯度算法，有效地简化了算法结构及计算工作量. 为此，在文献[8]的基础上，本文提出一个修正的共轭投影梯度算法. 该方法保证了每个试探点都不会远离可行域，同时在合适的条件下证明了算法的收敛性.

1　相关理论知识

1.1　相关符号

对于问题(1)的近似解 $\boldsymbol{x}^k \in X$，$\sigma_k > 0$，正定矩阵 $\boldsymbol{B}_k = \boldsymbol{B}(\boldsymbol{x}^k)$，集合 $L_k \subseteq I$，为了便于讨论给出以下记号：

$$\boldsymbol{g}_j^k = \boldsymbol{g}_j(\boldsymbol{x}^k) = \nabla \boldsymbol{c}_j(\boldsymbol{x}^k),\ \boldsymbol{H}_j^k = \boldsymbol{H}_j(\boldsymbol{x}^k) = \nabla^2 \boldsymbol{c}_j(\boldsymbol{x}^k)$$
$$\boldsymbol{F}(\boldsymbol{x}^k) = (\boldsymbol{c}_j(\boldsymbol{x}^k),\ j \in L_k),\ \boldsymbol{A}_k = \boldsymbol{A}(\boldsymbol{x}^k) = (\boldsymbol{g}_j(\boldsymbol{x}^k),\ j \in L_k) \qquad j = 1, \cdots, m \tag{2}$$

本文算法引用了文献[7]提出的共轭投影梯度技术，并采用文献[7]中定义的以下各量：

---

① 收稿日期：2017－05－24
基金项目：国家自然科学基金项目(11361018)；广西自然科学基金资助项目(2014GXNSFFA118001)；宜春市社科研究“十三五”规划项目(YCSK2018－106，YCSK2018－115).
作者简介：王祥玲(1986－)，女，讲师，主要从事最优化理论研究.

$$Q_k = Q(x^k) = (A_k^{\mathrm{T}} B_k^{-1} A_k)^{-1} A_k^{\mathrm{T}} B_k^{-1},\ P_k = P(x^k) = B_k^{-1}(E_n - A_k Q_k) \tag{3}$$

$$\pi^k = \pi(x^k) = -Q_k \nabla f(x^k),\ d_0^k = d_0(x^k) = -P_k \nabla f(x^k) + Q_k^{\mathrm{T}} V^k \tag{4}$$

$$V^k = V(x^k) = (V_j^k,\ j \in L_k),\ V_j^k = \begin{cases} -f_j(x^k) & \pi_j^k > 0 \\ \pi_j^k & \pi_j^k \leqslant 0 \end{cases} \tag{5}$$

$$d_1^k = -Q_k^{\mathrm{T}}(\| d_0^k \|^{\tau} e + F(x^k + d_0^k)),\ e = (1, \cdots, 1),\ d^k = d_0^k + d_1^k \tag{6}$$

$$\rho_k = -\nabla f(x^k)^{\mathrm{T}} d_0^k,\ d_2^k = \frac{-\rho_k}{1 + 2\,|e^{\mathrm{T}} \pi^k|} Q_k^{\mathrm{T}} e,\ q^k = \rho_k (d_0^k + d_2^k) \tag{7}$$

定义 1[7] 如果函数 $\mu(x)$: $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ 满足以下条件:

1) $\mu(x)$ 连续.

2) 当 x^* 是问题(1)的 K-T 点时, $\mu(x^*)$ 为相应的 K-T 乘子.

则称 $\mu(x)$ 为乘子函数.

1.2 滤子技术

滤子技术能很好地平衡目标函数和约束条件, 对于每次迭代, 试探点被接受当且仅当约束违反度函数值或目标函数值相对于当前的滤子集中的点有充分的下降.

本文定义约束违反度函数为

$$h(x) = \| c(x)^+ \|_\infty \tag{8}$$

其中

$$c(x)^+ = \max\{0,\ c_j(x):\ j \in I\}$$

易知 $h(x)=0$ 的充要条件是保证迭代点 x 是可行点, 所以试探点应该使得约束违反度函数值降低或目标函数值减小. 为了保证这两项至少有一项满足, 参照文献[1]引入滤子的相关定义.

定义 2 称点 x^1 支配点 x^2 当且仅当

$$h(x^1) \leqslant h(x^2) \text{ 且 } f(x^1) \leqslant f(x^2) \tag{9}$$

定义 3 滤子为具有形式为 (h_i, f_i) 的点集 Z, 使得任意点不能被其他点支配, 即 $h(x^i) \leqslant h(x^j)$ 或 $f(x^i) \leqslant f(x^j)$ 对所有的 $i \neq j$ 成立.

在实际计算中, 为了获得全局收敛性, 需要添加一些条件, 如文献[1]中定义了一个"包络"来阻止点任意接近滤子. 从而有以下滤子准则:

定义 4 一个点被滤子接受当且仅当对任意的 $(h_j, f_j) \in Z$ 有

$$h(x) \leqslant \beta h_j \text{ 或 } f(x) \leqslant f_j - \gamma h_j \tag{10}$$

其中: $h_j = h(x^j)$; $f_j = f(x^j)$; $0 < \gamma < \beta < 1$ 且 $\gamma \to 0$, $\beta \to 1$.

在算法的进程中需要将 (h_j, f_j) 对添加到滤子中, 如果一个点 x^k 被滤子集 Z 接受, 则令 $x^{k+1} = x^k + t_k d^k$ 且滤子按如下的规则更新

$$Z_{k+1} = Z_k \cup \{(h_{k+1}, f_{k+1})\} \backslash D_{k+1} \tag{11}$$

其中

$$D_{k+1} = \{(h_j, f_j) \mid h_j \geqslant h_k,\ f_j - \gamma h_j \geqslant f_k - \gamma h_k,\ \forall (h_j, f_j) \in Z_k\}$$

事实上, 这里"点 x 被滤子集 Z 接受"严格上说是将 (h_j, f_j) 对添加到滤子集中. 如果点 x^k 被滤子集接受或者已在滤子集中, 那么满足

$$h(x) \leqslant (1-\gamma) h_k \text{ 且 } f(x) \leqslant f_k - \gamma h_k \tag{12}$$

的点 x 也是被滤子接受的.

2 算法模型

算法 1

步 0(初始化) 给定 $x^0 \in \mathbb{R}^n$, $\mu(x^0) \in \mathbb{R}^m$, B_0 是对称正定矩阵, 选择 $\sigma_0, \xi, \varepsilon, v \in (0, 1)$, $\tau \in (2, 3)$, $\delta > 2$ 且 $\alpha_1, \alpha_2, \eta \in (0, 1)$, $t_0 = 1$, 初始化滤子集 Z_0, 并令 $k = 0$.

步 1（有效集 L_k）

(a) 令 $i=0$，$\sigma_{k,i}=\sigma_0$；

(b) 若 $\det(\boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{A}_k)>\sigma_{k,i}$，令 $L_k=L_{k,i}$，$i_k=i$，转步 2，否则转步 1(c)，其中：

$$L_{k,i}=\{j\in I\mid -\sigma_{k,i}\mid\boldsymbol{\mu}_j(\boldsymbol{x}^k)\mid\leqslant f_j(\boldsymbol{x}^k)\leqslant 0\},\ \boldsymbol{A}_{k,i}=(\boldsymbol{g}_j(\boldsymbol{x}^k),\ j\in L_{k,i})$$

(c) 令 $i=i+1$，$\sigma_{k,i}=0.5\sigma_{k,i-1}$，返回步 1(b).

步 2（搜索方向）　通过式(2)－(5)获得 $\boldsymbol{d}_0^k$. 若 $\boldsymbol{d}_0^k=0$，则获得一个 K－T 点，停止. 否则，通过式(6)计算 $\boldsymbol{d}^k$.

若

$$\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{d}_0^k\leqslant\min\{-\xi\|\boldsymbol{d}_0^k\|^{\delta},-\xi\|\boldsymbol{d}^k\|^{\delta}\}\tag{13}$$

则转步 3. 否则，转步 4.

步 3（滤子准则）　若 $\boldsymbol{x}^k+t_k\boldsymbol{d}^k$ 不被滤子接受，则令 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k$，$t_{k+1}=0.5t_k$，$k=k+1$，转步 4. 若 $\boldsymbol{x}^k+t_k\boldsymbol{d}^k$ 被滤子接受，则令 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+\boldsymbol{d}^k$，$f_{k+1}=f(\boldsymbol{x}^k+\boldsymbol{d}^k)$，$h_{k+1}=h(\boldsymbol{x}^k+\boldsymbol{d}^k)$，并去除那些被 $\boldsymbol{x}^{k+1}$ 支配的点，更新滤子集 Z_{k+1}，转步 5.

步 4（线搜索）　通过式(7)获得 $\boldsymbol{q}^k$，取 $\left\{1,\frac{1}{2},\frac{1}{4},\cdots\right\}$ 的最大值为 λ_k，使其满足

$$f(\boldsymbol{x}^k+\lambda_k\boldsymbol{q}^k)\leqslant f(\boldsymbol{x}^k)+\upsilon\lambda_k\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{q}^k\tag{14}$$

$$\boldsymbol{c}_j(\boldsymbol{x}^k+\lambda_k\boldsymbol{q}^k)\leqslant 0,\ j\in I\tag{15}$$

令 $\boldsymbol{d}^k=\boldsymbol{q}^k$，$t_k=\lambda_k$.

步 5（迭代更新）　令 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{d}^k$，更新正定对称矩阵 $\boldsymbol{B}_{k+1}$，$k=k+1$. 返回步 1.

3　算法的可行性和收敛性分析

算法 A 需要以下理论的支持.

假设 A

(A1) 迭代序列 $\{\boldsymbol{x}^k\}\subseteq X$ 非空且位于 $\mathbb{R}^n$ 中的有界闭凸子集内；

(A2) 对任意的 $\boldsymbol{x}\in\mathbb{R}^n$，向量组 $\{\nabla\boldsymbol{c}_j(\boldsymbol{x}),\ j\in I(\boldsymbol{x})\}$ 是线性无关的；

(A3) $\forall k$，$\boldsymbol{B}_k\in\mathbb{R}^n$ 是一个 n 阶对称正定矩阵且 $\{\boldsymbol{B}_k\}$ 一致有界；

(A4) 目标函数 $f(\boldsymbol{x})$ 和约束违反度函数 $h(\boldsymbol{x})$ 都是二次连续可微的.

(A5) 存在两个常数 $0<a\leqslant b$，使得 $a\|\boldsymbol{d}\|^2\leqslant\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{d}\leqslant b\|\boldsymbol{d}\|^2$，$\forall k$，$\forall\boldsymbol{d}\in\mathbb{R}^n$.

由于集合 $L_k\subseteq I$ 的有限性，不妨设存在无穷子列 K，使得

$$\boldsymbol{x}^k\to\boldsymbol{x}^*,\ \boldsymbol{B}_k\to\boldsymbol{B}_*,\ L_k\equiv L,\ k\in K$$

其中 L 为一个固定指标集. 记 $\boldsymbol{A}_*$，$\boldsymbol{Q}_*$，$\boldsymbol{P}_*$，$\boldsymbol{\pi}^*$，$\boldsymbol{V}^*$，$\boldsymbol{d}_0^*$，$\boldsymbol{d}^*$，$\boldsymbol{q}^*$ 分别为式(3)－(7)中各式在 $\boldsymbol{x}^*$ 处以 $\boldsymbol{B}_*$ 代替 $\boldsymbol{B}(\boldsymbol{x}^*)$ 的相应各量，则 $\boldsymbol{A}_k\to\boldsymbol{A}_*$，$\boldsymbol{Q}_k\to\boldsymbol{Q}_*$，$\boldsymbol{P}_k\to\boldsymbol{P}_*$，$\boldsymbol{d}_0^k\to\boldsymbol{d}_0^*$，$\boldsymbol{q}^k\to\boldsymbol{q}^*$，$k\in K$，$k\to\infty$.

引理 1　对任意的迭代指标 k，算法 1 步 1(b) 和步 1(c) 之间的循环有限步终止.

证　利用反证法.

假设算法在步 1(b) 和步 1(c) 之间无限循环，则由算法结构知，

$$\det(\boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{A}_k)\leqslant\frac{1}{2^i}\sigma_0\qquad i=1,2,\cdots$$

由 $L_{k,i}$ 的定义有

$$L_{k,i}\supseteq L_{k,i+1}$$

又集合 I 是个有限集，从而对充分大的 i 有

$$L_{k,i}=L_{k,i+1}$$

并定义为 L_k^*. 令 $i\to\infty$，则

$$\det(\boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{A}_k)=0 \qquad J_k^*=I(\boldsymbol{x}^k)$$

这与假设(A2)矛盾.

引理2 若假设A满足且迭代点 $\boldsymbol{x}^k\in X$,则有下列结论成立

1) $\boldsymbol{x}^k$ 是问题(1)的K-T点当且仅当 $\boldsymbol{d}_0^k=0$.

2) 若 $\boldsymbol{x}^k$ 不是问题(1)的K-T点,则 $\boldsymbol{d}^k$ 仍是问题(1)在点 $\boldsymbol{x}^k$ 处的一个可行下降方向,即有

$$\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{d}_0^k<0,\ \nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{q}^k\leqslant-\frac{1}{2}\boldsymbol{\rho}_k^2<0,\ (\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{d}_0^k\leqslant0,\ (\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{q}^k<0,\ j\in I(\boldsymbol{x}^k)$$

证 参考文献[7]中定理1的证明.

引理3[7] 对于算法1产生的序列$\{\boldsymbol{x}^k\}$,设 $\boldsymbol{x}^*$ 是它的任意一个聚点.若 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{d}^k$ 均由步4和步5产生,且 $\boldsymbol{x}^*$ 不是问题(1)的K-T点,则有

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{q}^*<0,\ \boldsymbol{g}_j(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{q}^*<0,\ j\in I(\boldsymbol{x}^*)$$

$$\lambda_k\geqslant\lambda_*=\inf\{\lambda_k\}>0,\ k\in K$$

引理4 算法是可执行的.

证 只需证明在假设A成立的前提下,当k充分大时,算法不再执行步4.由算法知,只需证明由步4产生的迭代点也一定能被滤子接受.假设算法不能有限步终止,则有 $t_k\to0$,$k\to\infty$,且新的迭代点 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{q}^k$ 不能被滤子接受.下面分两种情况讨论.

1) 若 $h_k=0$.由 $h(\boldsymbol{x})$ 的定义可得,

$$h_{k+1}=\max\{0,\boldsymbol{c}_j(\boldsymbol{x}^{k+1})\}=\max\{0,\boldsymbol{c}_j(\boldsymbol{x}^k)+t_k(\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{q}^k+O\|t_k\boldsymbol{q}^k\|^2\}$$

由引理1知,

$$(\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{q}^k<0$$

结合 $t_k\to0$,$k\to\infty$,则存在一个常数β,使得

$$h_{k+1}\leqslant\max\{0,\beta\boldsymbol{c}_j(\boldsymbol{x}^k)\}\leqslant\beta h_k \tag{16}$$

再由式(14)和引理2知

$$f(\boldsymbol{x}^{k+1})\leqslant f(\boldsymbol{x}^k)+v\lambda_k\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{q}^k\leqslant f(\boldsymbol{x}^k)-v\lambda_k\frac{1}{2}\boldsymbol{\rho}_k^2\leqslant f(\boldsymbol{x}^k)-\boldsymbol{c}\boldsymbol{\rho}_k^2,\ 0<\boldsymbol{c}<1$$

故

$$f(\boldsymbol{x}^{k+1})\leqslant f(\boldsymbol{x}^k) \tag{17}$$

由式(16)-(17)可知新的迭代点 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{q}^k$ 能被滤子和 $\boldsymbol{x}^k$ 接受.矛盾.

2) 若 $h_k\neq0$.类似于情况(1),可得 $h_{k+1}\leqslant\beta h_k$.由于 $\boldsymbol{x}^k$ 被滤子接受,所以有

$$h_k\leqslant\beta h_j\ \text{或}\ f_k\leqslant f_j-\gamma h_j,\ \forall(h_j,f_j)\in Z$$

由假设迭代点 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{q}^k$ 不能被滤子接受,则有

$$h_j\geqslant h_k \tag{18}$$

和

$$f_j-\gamma h_j\geqslant f_k-\gamma h_k \tag{19}$$

对于点 $\boldsymbol{x}^k$,若 $h_k\leqslant\beta h_i$ 成立,则由引理1知,

$$(\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{q}^k<0$$

结合 $t_k\to0$,$k\to\infty$,则存在一个常数β,使得

$$h_{k+1}=\max\{0,\boldsymbol{c}_j(\boldsymbol{x}^{k+1})\}\leqslant\max\{0,\boldsymbol{c}_j(\boldsymbol{x}^k)\}\leqslant h_k\leqslant\beta h_j$$

这与式(18)矛盾.

对于点 $\boldsymbol{x}^k$,若 $f_k\leqslant f_j-\beta h_j$ 成立,则由引理1知,

$$(\boldsymbol{g}_j^k)^{\mathrm{T}}\boldsymbol{q}^k<0$$

结合 $t_k\to0$,$k\to\infty$,则存在一个常数β,使得

$$f(\boldsymbol{x}^{k+1})=f(\boldsymbol{x}^k)+t_k\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{q}^k+O(\|t_k\boldsymbol{q}^k\|^2)\leqslant f(\boldsymbol{x}^k)\leqslant f_j-\beta h_j$$

这与式(19) 矛盾.

综上分析，当 k 充分大时，算法不再执行步 4，即算法是有效的.

引理 5[9]　考虑$\{\boldsymbol{x}^k\}$是进入滤子集的无穷序列，与其对应的 $h(\boldsymbol{x}^k)>0$ 且$\{f(\boldsymbol{x}^k)\}$有下界，则

$$\lim_{k\to\infty} h(\boldsymbol{x}^k)=0$$

引理 6[9]　1）如果有无限多个点被添加到滤子集中，则

$$\lim_{k\to\infty} h(\boldsymbol{x}^k)=0$$

2）如果有有限多个点被添加到滤子集中，则存在一个 k_1，使得当 $k>k_1$ 时，有

$$\lim_{k\to\infty} h(\boldsymbol{x}^k)=0$$

引理 7　按式(6) 计算的主方向 $\boldsymbol{d}^k$ 满足 $\|\boldsymbol{d}^k\|\sim\|\boldsymbol{d}_0^k\|$，$\|\boldsymbol{d}_1^k\|=O(\|\boldsymbol{d}_0^k\|^2)$.

证　由算法 1 步 4 中 λ_k 的取法易知，当 k 充分大时，有 $\lambda_k\equiv 1$，此时 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+\boldsymbol{d}^k$. 当 $j\in I_*$ 时有

$$\begin{aligned}\boldsymbol{c}_j(\boldsymbol{x}^{k+1})=\boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}^k)=\boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}_0^k+\boldsymbol{d}_1^k)=&\\ \boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}_0^k)+(\boldsymbol{g}_j^k)^{\mathrm T}\boldsymbol{d}_1^k+O(\|\boldsymbol{d}_1^k\|^2)=\boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}_0^k)+(\boldsymbol{g}_j^k)^{\mathrm T}\boldsymbol{d}_1^k+O(\|\boldsymbol{d}_0^k\|^3)&\end{aligned}$$

而

$$\boldsymbol{A}_k^{\mathrm T}\boldsymbol{d}_1^k=-\|\boldsymbol{d}_0^k\|^\tau\boldsymbol{e}-\boldsymbol{F}(\boldsymbol{x}^k+\boldsymbol{d}_0^k)，故\ \boldsymbol{g}_j(\boldsymbol{x}^k)^{\mathrm T}\boldsymbol{d}_1^k=-\|\boldsymbol{d}_0^k\|^\tau-\boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}_0^k)$$

所以，

$$\boldsymbol{c}_j(\boldsymbol{x}^{k+1})=\boldsymbol{c}_j(\boldsymbol{x}^k+\boldsymbol{d}^k)=-\|\boldsymbol{d}_0^k\|^\tau+O(\|\boldsymbol{d}_0^k\|^3)，\tau\in(2,3)$$

从而当 k 充分大时有，

$$\|\boldsymbol{d}_1^k\|=O(\|\boldsymbol{d}_0^k\|^2)，\ \|\boldsymbol{d}^k\|\sim\|\boldsymbol{d}_0^k\|$$

定理 1　在假设 $\boldsymbol{A}$ 成立的前提下，算法 1 或有限次迭代终止于问题(1) 的一个 K－T 点 $\boldsymbol{x}^k$，或产生一个无穷迭代序列$\{\boldsymbol{x}^k\}$，使得它的任意一个聚点 $\boldsymbol{x}^*$ 是问题(1) 的 K－T 点.

证　若算法 1 终止于点 $\boldsymbol{x}^k$，则 $\boldsymbol{d}_0^k=0$. 故由引理 2 可知迭代点 $\boldsymbol{x}^k$ 是问题(1) 的 K－T 点. 即证明了定理 1 的前半部分. 下面假设 $\boldsymbol{x}^*$ 为无穷迭代序列$\{\boldsymbol{x}^k\}$的任意一个聚点，不妨设当 $k\to\infty$ 时有 $\boldsymbol{x}^k\to\boldsymbol{x}^*$，$k\in K$，其中 K 为无穷指标集. 下面分情况讨论.

1）若存在一个无穷序列 $K_1\subseteq K$，使得迭代点 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{d}^k$ 均由算法 1 步 2 和步 3 产生. 结合引理 5、引理 6 知 $\lim\limits_{k\to\infty}h(\boldsymbol{x}^k)=0$. 若 $\lim\limits_{k\in K_1,k\to\infty}\|\boldsymbol{d}^k\|=0$，则易知 $\boldsymbol{x}^*$ 是一个 K－T 点. 由 $\lim\limits_{k\to\infty}h(\boldsymbol{x}^k)=0$ 和假设 A 中的(A5) 可设存在一个 k_0，使得当 $k>k_0$，$k\in K_1$，有 $h(\boldsymbol{x}^k)\leqslant\dfrac{a\varepsilon^2}{2M}\leqslant\dfrac{a\|\boldsymbol{d}^k\|^2}{2M}\leqslant\dfrac{(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k}{2M}$，结合 K－T 条件，对所有的 $k>k_0$，$k\in K_1$ 有

$$\begin{aligned}\nabla f(\boldsymbol{x}^k)^{\mathrm T}\boldsymbol{d}^k=-(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k-(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{A}_k\omega_k=-(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k-\omega_k^{\mathrm T}\boldsymbol{c}(\boldsymbol{x}^k)=&\\ -(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k+\|\omega_k\|_\infty h(\boldsymbol{x}^k)\leqslant Mh(\boldsymbol{x}^k)-(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k\leqslant&\\ -0.5(\boldsymbol{d}^k)^{\mathrm T}\boldsymbol{B}_k\boldsymbol{d}^k\leqslant-0.5a\|\boldsymbol{d}^k\|^2&\end{aligned}$$

其中 ω_k 为拉格朗日乘子. 从而，

$$\begin{aligned}0=\lim_{k\in K_1,k\to\infty}(f(\boldsymbol{x}^{k+1})-f(\boldsymbol{x}^k))=&\\ \lim_{k\in K_1,k\to\infty}(t_k\nabla f(\boldsymbol{x}^k)^{\mathrm T}\boldsymbol{d}^k+O(\|t_k\boldsymbol{d}^k\|^2)\leqslant&\\ \lim_{k\in K_1,k\to\infty}t_k\nabla f(\boldsymbol{x}^k)^{\mathrm T}\boldsymbol{d}^k\leqslant 0&\end{aligned}$$

因此

$$\lim_{k\in K_1,k\to\infty}\|\boldsymbol{d}^k\|=0$$

再结合引理 7 可得

$$\lim_{k\in K_1,k\to\infty}\|\boldsymbol{d}_0^k\|=0$$

而

$$\lim_{k\in K_1,k\to\infty}\boldsymbol{d}_0^k=\boldsymbol{d}_0^*$$

因此，根据引理2的证明，可证 $\boldsymbol{x}^*$ 是问题(1)的一个K－T点.

2) 若存在一个无穷序列 $K_1\subseteq K$，使得若迭代点 $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k+t_k\boldsymbol{d}^k$ 均由算法1步4和步5产生，反设 $\boldsymbol{x}^*$ 不是问题(1)的一个K－T点，则由引理3知

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{q}^*<0,\ \boldsymbol{g}_j(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{q}^*<0,\ j\in I(\boldsymbol{x}^*),\ \lambda_k\geqslant\lambda_*=\inf\{\lambda_k\}>0,\ k\in K$$

再由式(14)知

$$0=\lim_{k\in K_1,k\to\infty}(f(\boldsymbol{x}^{k+1})-f(\boldsymbol{x}^k))\leqslant\lim_{k\in K_1,k\to\infty}v\lambda_k\nabla f(\boldsymbol{x}^k)^{\mathrm{T}}\boldsymbol{q}^k\leqslant0.5v\lambda_*\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{q}^*<0$$

这是一个矛盾. 因此 $\boldsymbol{x}^*$ 是问题(1)的一个K－T点.

4 数值实验

在这一部分从数值计算的角度进一步说明本文给出的算法的有效性，算法的编程使用MATLAB软件. 算法中涉及的参数取值如下：

$$\sigma_0=0.01,\ \xi=0.01,\ \varepsilon=0.05,\ v=0.1,\ \tau=2.25,\ \delta=2.5,\ \gamma=0.05,\ \beta=0.95,\ \boldsymbol{B}_0=I\in\mathbb{R}^{n\times n}$$

$\boldsymbol{B}_k$ 的更新选自BFGS公式修正[10]：

$$\boldsymbol{B}_{k+1}=BFGS(\boldsymbol{B}_k,\ \boldsymbol{s}^k,\ \boldsymbol{y}^k)$$

其中：

$$\boldsymbol{s}^k=\boldsymbol{x}^{k+1}-\boldsymbol{x}^k,\ \hat{\boldsymbol{y}}^k=\nabla f(\boldsymbol{x}^{k+1})-\nabla f(\boldsymbol{x}^k)+\sum_{j=1}^m\boldsymbol{\mu}_j^k(\boldsymbol{g}_j(\boldsymbol{x}^{k+1})-\boldsymbol{g}_j(\boldsymbol{x}^k)),\ \boldsymbol{y}^k=\theta\hat{\boldsymbol{y}}^k+(1-\theta)\boldsymbol{B}_k\boldsymbol{s}^k$$

$$\theta=\begin{cases}1 & \text{如果}(\hat{\boldsymbol{y}}^k)^{\mathrm{T}}\boldsymbol{s}^k\geqslant0.2(\boldsymbol{s}^k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{s}^k\\ \dfrac{0.8(\boldsymbol{s}^k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{s}^k}{(\boldsymbol{s}^k)^{\mathrm{T}}\boldsymbol{B}_k\boldsymbol{s}^k-(\hat{\boldsymbol{y}}^k)^{\mathrm{T}}\boldsymbol{s}^k} & \text{否则}\end{cases}$$

而乘子函数 $\boldsymbol{\mu}(\boldsymbol{x})=-(\boldsymbol{N}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{N}(\boldsymbol{x})+\boldsymbol{D}(\boldsymbol{x}))^{-1}\boldsymbol{N}(\boldsymbol{x})^{\mathrm{T}}\nabla f(\boldsymbol{x})$，其中 $\boldsymbol{N}(\boldsymbol{x})=(\boldsymbol{g}_j(\boldsymbol{x}),\ j\in I)$，$\boldsymbol{D}(\boldsymbol{x})=\mathrm{diag}(\boldsymbol{c}_j^2(\boldsymbol{x}),\ j\in I)$. 终止准则为 $\|\boldsymbol{d}_0^k\|<10^{-6}$.

从文献[11]中选取了7个测试问题，对本文的算法进行测试，测试结果见表1.

表1 本文算法与文献[11]的数值结果

问题编号	n	m	$\boldsymbol{X}^0$	NF_1	NG_1	NIT_1	NF_2	NG_2	NIT_2
hs012	2	1	$(0,0)^{\mathrm{T}}$	10	8	7	13	60	12
hs033	3	6	$(0,0,3)^{\mathrm{T}}$	4	4	3	142	159	48
hs043	4	3	$(0,0,0,0)^{\mathrm{T}}$	19	12	11	19	12	11
hs076	4	3	$(0.5,0.5,0.5,0.5)^{\mathrm{T}}$	7	7	6	17	48	13
hs100	7	4	$(1,2,0,4,0,1,1)^{\mathrm{T}}$	37	19	18	65	113	23
hs110	10	20	$(9,9,9,9,9,9,9,9,9,9)^{\mathrm{T}}$	10	8	9	8	16	8
hs113	10	8	$(2,3,5,5,1,2,7,3,6,10)^{\mathrm{T}}$	25	15	14	62	204	33

表1问题编号即测试问题在文献[11]中的编号，n 表示问题中变量的维数，m 表示问题中约束的维数，$\boldsymbol{X}^0$ 表示问题的初始点，NF_1，NF_2 分别表示本文算法和文献[11]的目标函数和约束函数的计算次数，NG_1，NG_2 分别表示本文算法和文献[11]的梯度的计算次数，NIT_1，NIT_2 分别表示本文算法和文献[11]的算法的迭代次数.

从上述分析可看出，本文给出的修正共轭投影梯度滤子算法是有效可行的．将共轭投影技术引入到滤子算法使得该算法不需要求解二次规划子问题，而且能有效避免常规滤子算法中的恢复算法，简化了计算．

参考文献：

［1］ FLETCHER R，LEYFFER S．Nonlinear Programming Without a Penalty Function［J］．Mathematica Programming，2002，91(2)：239－269．

［2］ NIE P Y．Composite-Step like Filter Methods for Equality Constriant Problems［J］．Journal of Computational Mathematics，2003，21(5)：613－624．

［3］ 苏　珂，刘　英．求解非线性规划的修正滤子信赖域方法［J］．数学学报(中文版)，2009，52(6)：1157－1164．

［4］ 王祥玲，朱志斌，杨　萌．一种基于步长的 SQP 滤子法［J］．应用数学，2010，23(3)：670－674．

［5］ WANG X L，ZHU Z B，ZUO S Y，et al．An SQP-Filter Method for Inequality Constrained Optimization and Its Global Convergence［J］．Applied Mathematics and Computation，2011，217(24)：10224－10230．

［6］ HUANG Q Q，ZHU Z B，WANG X L．A Predictor-Corrector Algorithm Combined Conjugate Gradient with Homotopy Interior Point for General Nonlinear Programming［J］．Applied Mathematics and Computation，2013，219(9)：4379－4386．

［7］ 朱志斌，张可村．一个新的共轭投影梯度算法及其超线性收敛性［J］．应用数学学报，2004，27(1)：149－161．

［8］ 王祥玲，朱志斌，周志轩．共轭投影梯度滤子算法及其全局收敛性［J］．桂林电子科技大学学报，2012，32(6)：496－498．

［9］ FLETCHER R，LEYFFER S，TOINT P L．On the Global Convergence of a Filter-SQP Algorithm［J］．SIAM Journal on Optimization，2002，13(1)：44－59．

［10］ 袁亚湘，孙文瑜．最优化理论与方法［M］．北京：科学出版社，1997：232－238．

［11］ HOCK W，SCHITTKOWSKI K．Test Examples for Nonlinear Programming Codes［J］．Journal of Optimization Theory and Applications，1980，30(1)：127－129．

［12］ 柳　馨．两个修正的 DL 共轭梯度法［J］．重庆工商大学学报(自然科学版)，2017，34(5)：13－18．

A Modified Conjugate Projection Gradient Filter Method

WANG Xiang-ling[1]，　ZUO Shuang-yong[1]，　ZHU Zhi-bin[2]

1. Primary Education College，Yichun Early Childhood Teachers College，Yichun Jiangxi 330814，China；

2. Department of Mathematics and Computational Science，Guilin University of Electronic Technology，Guilin Guangxi 541004，China

Abstract：A new conjugate projection gradient filter algorithm is established by modifying the search direction. In this algorithm，conjugate projection gradient technology and filter method are combined. By the introduction of the filter，this algorithm does not need to solve a QP sub-problem. With the idea of the conjugate projection gradient，this method is effective to avoid the restoration algorithm in general filter algorithms. Under some conditions，its global convergence is obtained.

Key words：conjugate projection gradient；filter；nonlinear programming；global convergence

责任编辑　张　构