求非线性规划问题的一种非可行域方法

姜爱萍,濮定国
(同济大学 数学系,上海 200092)

摘要: 提出一种带滤子的QP-free非可行域方法,用来解满足不等式约束的非线性规划问题.此方法通过乘子函数和4-1线性互补函数构造一个等价于原约束问题的一阶KKT条件的非光滑方程组,并在此基础上给出解这个方城组的迭代算法.这个方法的每一步迭代都可以看作是对求KKT条件解的牛顿或拟牛顿迭代的扰动,在线性搜索时我们用到滤子方法.这个方法是可实行的且具有全局性,并且在适当的条件下我们还可以得到此方法的超线性收敛性.

关键词: 滤子;NCP函数;收敛性
中图分类号:O221.2 **文献标识码**:A **文章编号**:1007-855X(2007)05-0118-04

A Kind of Infeasible Method for Nonlinear Programming Problems

JIANG Ai-ping, PU Ding-guo
(Deparment of Mathematics, Tongji University, Shanghai 200092, China)

Abstract: A filter QP-free infeasible method is proposed for minimizing a function subject to smooth inequality constraints. This iterative method is based on the solution of nonsmooth equations which are obtained by the multiplier and some NCP functions for the KKT first-order optimality conditions. Locally, each iteration of this method can be viewed as a perturbation of a Newton or quasi-Newton iteration on both the primal and dual variables for the solution of the KKT optimality conditions. The filter is also used in linear search. This method is implementable and globally convergent, which proves that the method has superlinear convergence rate under sortie mild conditions.

Key words: filter; NCP function; convergence

0 引言

我们考虑如下的约束非线性规化问题(NLP):

$$\min f(x),\ x \in R^n,$$
$$s.t.\quad G(x) \leqslant 0 \tag{1}$$

其中:$f: R^n \to R$ 和 $G(x) = (g_1(x), g_2(x), \cdots, g_m(x))^{\mathrm{T}}: R^n \to R^m$ 是Lipschitz连续可微函数.

我们定义约束非线性优化问题(NLP)的可行域为

$$D = \{x \in R^n \mid G(x) \leqslant 0\}$$

约束非线性优化问题(NLP)的Lagrangian乘子函数为

$$L(x,\mu) = f(x) + \mu^T G(x)$$

其中 $\mu = (\mu_1, \mu_2, \cdots, \mu_m)^{\mathrm{T}} \in R^m$ 是乘子向量,为方便起见,用 $(x,\mu)^{\mathrm{T}}$ 表示列向量 $(x^{\mathrm{T}}, \mu^{\mathrm{T}})^{\mathrm{T}}$.

Karush-Kuhn-Tucker($K-K-T$)点 $(\bar{x}, \bar{\mu}) \in \mathrm{R}^n \times \mathrm{R}^m$ 是满足 NLP 问题的一阶最优必要条件的点

$$\nabla_x L(\bar{x}, \bar{\mu}) = 0,\ G(\bar{x}) \leqslant 0,\ \bar{\mu} \geq 0,\ \mu_i g_i(\bar{x}) = 0 \tag{2}$$

收稿日期:2007-03-27. **基金项目**:国家自然科学基金资助项目(项目编号:10571137,10371089).
第一作者简介:姜爱萍(1979-),女,博士研究生. 主要研究方向:最优化论理论与方法. **E-mail**:ap724@126.com

已经提出解约束非线性优化问题的 $QP-free$ 方法大多是可行域方法(见文献[1~3]). 本文提出了一种新的可实现的 $QP-free$ 非可行域方法. 与文献[2]中提出的方法相比,我们的方法不要求严格可行性. 其中我们用 $QP-free$ 方法来找目标函数 f(x)的下降方向,用滤子使目标函数或者违反约束函数下降. 我们还证明了在一定条件下,该方法具有超线性收敛性.

本文结构如下:在第 1 节给出一些预备知识;第 2 节给出这个带滤子的 $QP-free$ 非可行域方法的具体算法,并在第 3 节指出该方法是可行的,并证明了该方法在较弱的假设条件下具有全局收敛性和超线性收敛性;第 4 节给出了本文的结论.

1 预备知识

4-1 线性 NCP 函数如下定义:

$$\psi(a,b)=\begin{cases}k^2a & \text{如果 } b\geq k|a|,\\ 2kb-b^2/a & \text{如果 } a\geq |b|/k \text{ 且 } a>0,\\ 2ka+2kb+b^2/a & \text{如果 } a\leq -|b|/k \text{ 且 } a<0,\\ k^2a+4kb & \text{如果 } b\leq k|a|.\end{cases}\tag{3}$$

在解决非线性规划问题时,我们既要使目标函数下降还要满足可行域条件,为此我们引进滤子的概念[4]. 我们令

$$p(G(x),\mu)=\|\phi_1(x,\mu)\|^2$$

其中 $\phi_1(x,\mu)=(\phi_1(x,\mu),\cdots,\phi_m(x,\mu))^{\mathrm{T}},\phi_i(x,\mu_i)=\psi(-g_i(x),\mu_i)$

定义 1.1 $(p(G(x_k),\mu_k),f(x_k))$ 优于 $(p(G(x_l),\mu_l),f(x_l))$ 当且仅当 $(p(G(x_k),\mu_k)\leq(p(G(x_l),\mu_l)$ 且 $f(x_k)\leq f(x_l)$.

定义 1.2 滤子是由一对对的 $(p(G(x_k),\mu_k),f(x_k))$ 组成的集合,使得其中没有任何一对优于其他对. 如果滤子中的任何一对都不比 $(p(G(x_k),\mu_k),f(x_k))$ 优,则 $(p(G(x_k),\mu_k),f(x_k))$ 可被加到滤子中去.

我们用 F_k 定义迭代步指标 $l(l<k)$ 的集合使得 $(p(G(x_l),\mu_l),f(x_l))$ 是当前步滤子中的元素.

2 算 法

如果 $(g(x^k),\mu^k)\neq(0,0)$,令

$$\xi_j^k=\xi_j(x^k,\mu^k)=\begin{cases}-k^2 & \text{如果 } \mu_i\geq k|g_i(x)|,\\ -\mu_i^2/g_i(x)^2 & \text{如果 } if -g_i(x)\geq|\mu_i|/k \text{ 且 } -g_i(x)>0,\\ -2k\mu_i+\mu_i^2/g_i(x)^2 & \text{如果 } g_i(x)\geq|\mu_i|/k \text{ 且 } -g_i(x)<0,\\ -k^2 & \text{如果 } \mu_i\leq k|g_i(x)|,\end{cases}$$

$$\eta_j^k=\eta_j(x^k,\mu^k)=\begin{cases}0 & \text{如果 } \mu_i\geq k|g_i(x)|,\\ 2k+2\mu_i/g_i(x) & \text{如果 } if -g_i(x)\geq|\mu_i|/k \text{ 且 } -g_i(x)>0,\\ -2kg_i(x)+2k-2\mu_i/g_i(x) & \text{如果 } g_i(x)\geq|\mu_i|/k \text{ 且 } -g_i(x)<0,\\ 4k & \text{如果 } \mu_i\leq k|g_i(x)|,\end{cases}$$

令

$$V^k=\begin{pmatrix}V_{11}^k & V_{12}^k\\ V_{21}^k & V_{22}^k\end{pmatrix}=\begin{pmatrix}H^k & \nabla G^k\\ \mathrm{diag}(\xi^k)(\nabla G^k)^{\mathrm{T}} & \mathrm{diag}(\eta^k-c^k)\end{pmatrix}.\tag{4}$$

和

$$\hat{V}^k=\begin{pmatrix}V_{11}^k & V_{12}^k\\ V_{21}^k & V_{22}^k\end{pmatrix}=\begin{pmatrix}H^k & \nabla G^k\\ \mathrm{diag}(\xi^k)(\nabla G^k)^{\mathrm{T}} & \mathrm{diag}(\eta^k-\hat{c}^k)\end{pmatrix}.\tag{5}$$

其中 H^k 是可以用 BFGS 方法来校正的正定矩阵. $\mathrm{diag}(\xi^k)$,$\mathrm{diag}(\hat{c}^k)$ 和 $\mathrm{diag}(\eta^k-c^k)$ 表示第 j 个元素分别是

$\xi_j^k,\hat{c}_j^k$ 和 $\eta_j^k-c_j^k$ 的对角矩阵.

$$\hat{c}_j^k = m_1\min\{1,\ \|\phi^k\|^v\}.$$

$$c_j^k = \begin{cases}0 & n_j^k \neq 0,\\ \hat{c}_j^k & \eta_j^k = 0.\end{cases}$$

其中 $m_1>0,v>1$ 是给定的参数.

算法 2.1

步骤0 初始化,给一个初始点 $x^0\in R^n,\tau\in(0,1),\bar{\mu}>0,1>\theta_1>\theta>0$,对称正定阵 H^0. 令 $k:=0$.

步骤1 计算搜索方向. 解下面线性方程组中的 (d,λ),把得到的解记为 d^{k0} 和 λ^{k0}

$$\hat{V}^k\begin{pmatrix}d\\ \lambda\end{pmatrix} = \begin{pmatrix}-\nabla f^k\\ 0\end{pmatrix} \tag{6}$$

解下面线性方程组中的 (d,λ),把得到的解记为 $d^{k1}\lambda^{k1}$

$$V^k\begin{pmatrix}d\\ \lambda\end{pmatrix} = \begin{pmatrix}-\nabla L^k\\ -\phi_1^k\end{pmatrix} \tag{7}$$

如果 $\eta_i=0$,那么就令 $\lambda_i^{k0}=\bar{\lambda}_i^{k0}$,否则令 $\lambda_i^{k0}=(\eta_i-\hat{c}_i)\bar{\lambda}_i^{k0}/\eta_i$.

步骤2 用滤子进行线搜索.

1) 如果

$$\|\phi_1(x^k+d^{k1},\mu^k+\lambda^{k1})\|^2 \leqslant \theta_1\|\phi_1^k\|^2, \tag{8}$$

那么令 $x^{k+1}=x^k+d^{k1}$ 且 $\bar{\mu}^k=\mu^k+\lambda^{k1}$ 转步 3.

2) 否则找到能被滤子接受的 x^{k+1}. 令 $x^{k+1}=x^k+\alpha^k d^k$ 且 $\mu^{k+1}=\mu^k+\alpha^k\lambda^k$,如果 $d^{k0}=0$,则令 $b^k=0,\rho_k=1$,否则就记 $b^k=(1-\rho^k)$,

$$\rho^k = \begin{cases}1 & ,\text{如果}(d^{k1})^T\ \nabla f^k \leqslant \theta(d^{k0})^T\ \nabla f^k,\\ (1-\theta)\dfrac{(d^{k0})^T\ \nabla f^k}{(d^{k0}-d^{k1})^T\ \nabla f^k} & ,\text{其他}\end{cases}$$

并且令

$$\begin{pmatrix}d^k\\ \lambda^k\end{pmatrix} = b^k t^k\begin{pmatrix}d^{k0}\\ \lambda^{k0}\end{pmatrix} + \rho^k(t^k)^2\begin{pmatrix}d^{k1}\\ \lambda^{k1}\end{pmatrix},$$

其中 $t^k=\tau^j$,j 是使得

或者
$$\|\phi_1(x^{k+1},\mu^{k+1})\|^2 \leqslant (1-\alpha_k)\theta\|\phi_1^k\|^2 \tag{9}$$

$$f(x^{k+1})-f(x^l) <^k\theta\|\phi_1^k\| \tag{10}$$

对所有的 $(f(x^l),\|\phi_1^l\|^2)\in F$ 成立的最小非负整数.

步骤3 更新.

如果 x^{k+1} 是 KKT 点,则停止;否则,如果 $\mu_i^{k+1}\leqslant\bar{\mu}$ 那么 $\mu_i^{k+1}=\mu_i^{k+1}$;若不然则令 $\mu_i^{k+1}=\bar{\mu}$,更新 H^{k+1},$F^{k+1}=F^k\cup(F(x(k+1)),\mu^{k+1})$,且删除在 $Fk+1$ 所有被 $(F(x(k+1)),\mu^{k+1})$ 优于的 $(f(x^l),\|\phi_1^l\|^2)$,$k=k+1$ 且转步骤1.

3 算法的实现及其收敛性

我们假设 A1 ~ A3 成立.

A1 水平集 $\{x|f(x)\leqslant f(x^0)\}$ 是有界的,且对任意大的 k,有 $\|\mu^k+\lambda^{k0}+\lambda^{k1}\|<\bar{\mu}$ 成立.

A2 f 和 g_i 是二次 Lipschitz 连续可微的,且对任意的 $y,z\in R^{n+m}$,满足

$$\|\nabla L(y)-\nabla L(z)\| \leqslant m_1\|y-z\|,\quad \|\phi(y)-\phi(z)\| \leqslant m_1\|y-z\|,$$

其中 $m_1>0$ 是 Lipschitz 常数.

A3 H^k 是正定矩阵且存在正数 m_2 和 m_3 使得对所有的 $d\in R^n$ 和 k,有

$$m_2 \| d \|^2 \leq d^T H^k d \leq m_3 \| d \|^2$$

成立.

引理 3.1　如果 $\phi^k \neq 0$,那么 V^k 和 $\hat{V}^k$ 是非奇异的.

引理 3.2　存在 $m_4 > 0$ 使得对任意的 $- < t \leq 1$,有

$$\| \phi_1(x^k + td^{k0}, \mu^k + t\lambda^{k0}) \|^2 - \| \phi_1 \|^2 \leq m_4 t^2.$$

引理 3.3　若 $\phi_1^k \neq 0$, $b^k = 0$,则对任意给定的 $\varepsilon > 0$,存在 $\bar{t} > 0$,使得对任意的 $0 < t \leq \bar{t}$,有

$$\| \phi_1^k \|^2 - \| \phi_1(x^k + d^{kt}, \mu^k + \lambda^{kt})^2 \geq (2 - \varepsilon) t^2 \| \phi_1^k \|^2 .$$

若 $\phi_1^k = 0$ 或者 $\phi_1^k \neq 0$ 且 $b^k \neq 0$,则存在 $m_5 > 0$,使得对任意的 $0 < t \leq 1$,有

$$\| \phi_1^k \|^2 - \| \phi_1(x^k + d^{kt}, \mu^k + \lambda^{kt}) \|^2 \geq - m_5 t$$

引理 3.4　考虑序列 $\{ \| \phi_1(x^k) \|^2 \}$ 和 $\{f^k\}$ 使得 $\{ \| \phi_1(x^k) \|^2 \geq 0 \}$ 和 $\{f^k\}$ 单调下降且下有界. 对所有的 k,令常数 η 和 γ 满足

$$\| \phi_1(x^k) \|^2 \leq (1 - \alpha) \eta \| \phi_1(x^l) \|^2 \tag{11}$$

或者

$$f(x) - f(x^l) < \gamma \phi_1(x) \tag{12}$$

其中 $\alpha > 0$ 是步长, η 和 $\gamma \in (0,1)$ 是给定的正的常数,那么 $\phi_1(x^k) \to 0$.

定理 3.1　如果 (x^*, μ_*) 是 $\{(x^k, \mu^k)\}$ 的聚点,那么 x^* 是 NLP 问题的 KKT 点.

以下假设:

A4　$\{ \nabla g_i(x^*) \}$ 是线性无关的,其中 $i \in I = \{ i | g_i(x^*) = 0 \}$, x^* 是 $\{x^k\}$ 的一个聚点,也是 NLP 问题的 KKT 点;

A5　数列 $\{H^k\}$ 满足

$$\frac{\| (H^k - \nabla_x^2 L(x^k, \mu^k)) d^{k1} \|}{\| d^{k1} \|} \to 0$$

A6　在每个 KKT 点 (x^*, μ^*),严格互补条件成立;

A7　f 和 g_i 都是三次 Lipschitz 连续可微的.

假设 A6 说明了 ϕ 在每个 KKT 点 (x^*, μ^*) 处是连续可微的.

定量 3.2　假设 A1 – A7 成立. 设 $\{(x^k, \lambda^k)\}$ 是由算法 3.1 产生的序列, (x^*, λ^*) 是它的一个聚点,则 (x^*, λ^*) 是 NLP 问题的 KKT 点, (x^k, λ^k) 超线性收敛于 (x^*, λ^*).

4 结 论

本文提出一种求解带不等式约束的非线性规划问题的 QP – free 非可行域方法. 此方法是有效并且可行的,并且在适当的条件下我们还可以得到此方法的超线性收敛性. 所以这是对求解带不等式约束的非线性规划问题的一种切实有效的新方法.

参考文献:

[1] E R Panier, A L Tits, J N Herskovits. A QP – free, globally, locally superlinear convergent method for the inequality constrained optimization problems[J]. SIAM J. Control Optim, 1988, 36: 788 – 811.

[2] H Qi, L Qi. A new QP – free, globally convergent, locally superlinearly convergent algorithm for inequalityconstrainedoptimization[J]. SIAM J. Optim, 2000(11): 113 – 132.

[3] D Pu, Y Zhou, Z Zhang. A QP Free Feasible Method[J]. Journal of Computational Mathematics, 2004, 22(5): 651 – 660.

[4] Roger Fletcher, Seven Leyffer. Nonlinear Programming Without a Penalty ruction[J]. Mathematical Programming, 2002, 91(2): 239 – 269, 125 – 139.

[5] Schittkowski K. More test examples for nonlinear programming codes[M]. Springer – Verlag, 1988.