非单调QP-free非可行域方法

濮定国, 金中

(同济大学 应用数学系,上海 200092)

摘要: 提出了带有 Fischer-Burmeister 非线性互补 (NCP) 函数的非单调 QP-free 非可行域算法. 根据优化问题的一阶 KKT 条件, 利用乘子和 NCP 函数, 得到非光滑方程, 给出解这个非光滑方程的迭代算法. 该算法包含原始-对偶变量, 在局部意义下, 可看成关于一阶 KKT 最优条件的扰动牛顿–拟牛顿迭代算法. 在线性搜索时, 此算法采用非单调方法. 给出的算法是可实现的并具有全局收敛性, 且在适当假设下具有超线性收敛性.

关键词: 非单调; QP-free 方法; 收敛性; 非线性互补函数

中图分类号: O 221.2 **文献表识码**: A **文章编号**: 0253-374x(2010)02-

Nonmonotone Line Search Technique for QP-free Infeasible Method

PU Dingguo, JIN Zhong

(Department of Applied Mathematics, Tongji University, Shanghai 200092, China)

Abstract: In this paper, a new QP-free infeasible method with nonmonotone line search techniqueis and the Fischer-Burmeister NCP function is proposed for minimizing a smooth function subject to smooth inequality constraints. This iterative method is based on the solution of nonsmooth equations which are obtained by the multiplier function and the Fischer-Burmeister NCP function for the KKT first-order optimality conditions. We use nonmonotone line search techniqueis on line searches. This method is implementable and globally convergent. We also prove that the method has superlinear convergence rate under some mild conditions.

Key Words: nonmonotone; QP-free method, convergence; nonlinear complementarity function.

考虑非线性约束优化问题 (NLP):

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

其中 $f : \mathbf{R}^n \to \mathbf{R}$ 和 $G(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \cdots, g_m(\mathbf{x}))^T : \mathbf{R}^n \to \mathbf{R}^m$ 是 Lipchitz 连续可微函数. 分别用 $D = \{\mathbf{x} \in \mathbf{R}^n | G(\mathbf{x}) < 0\}$ 和 $\bar{D} = cl(D)$ 来表示 NLP 问题的严格可行集和可行集. NLP 问题的 Lagrange 乘子函数为

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda^T G(\mathbf{x}), \tag{2}$$

其中 $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \in \mathbf{R}^m$ 是乘子向量. 为方便起见, 有时用 $(\mathbf{x}, \lambda)$ 表示列向量 $(\mathbf{x}^T, \lambda^T)^T$.

Karush-Kuhn-Tucker (KKT) 点 $(\bar{\mathbf{x}}, \bar{\lambda}) \in \mathbf{R}^n \times \mathbf{R}^m$ 是指满足 NLP 问题的一阶最优必要条件:

$$\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\lambda}) = 0, \ G(\bar{\mathbf{x}}) \le 0, \ \bar{\lambda} \ge 0, \ \bar{\lambda}_i g_i(\bar{\mathbf{x}}) = 0, \quad 1 \le i \le m, \tag{3}$$

的点. 如果存在 $\bar{\lambda}$ 使得 $(\bar{\mathbf{x}}, \bar{\lambda})$ 满足 (3) 式,也称 $\bar{\mathbf{x}}$ 为 KKT 点. 寻找 NLP 问题的 KKT 点等价于求解 (3) 中的混合非线性互补问题 (NCP). 在 NCP 函数的基础上, 文献 ([1–7]) 提出了解 NLP 问题的具有全局收敛性和在适当条件下具有超线性收敛性的 QP-free 方法.

收稿日期:2009-04-17

基金项目:国家自然科学基金资助项目(10771162)

作者介绍:濮定国(1948-),男,教授,哲学(理学)博士,博士生导师,E-mail:madpu@mail.tongji.edu.cn

另一方面, 当迭代点靠近一个狭长的峡谷区域, 可能导致非常短的步长或者锯齿现象, 降低收敛速度. 为了克服以上缺点, 可采用非单调线搜索技术.

本文提出一种新的带有 Fischer-Burmeister 非线性互补问题 (NCP) 函数的非单调 QP-free 非可行域算法. 根据优化问题的一阶 KKT 条件, 利用乘子和 NCP 函数, 得到非光滑方程, 本文给出解这个非光滑方程的迭代算法. 特别地, 定义了一个价值函数, 并使其在每一步迭代中不断下降. 这个算法是可实现的并具有全局收敛性. 可证明在适当的假设下, 此算法是超线性收敛的.

定义 1 NCP 对 如果 $a \geq 0$, $b \geq 0$, $ab = 0$, 则称数对 $(a, b) \in \mathbf{R}^2$ 是一个 NCP 对.

定义 2 NCP 函数 如果 $\phi(a, b) = 0$ 当且仅当 (a, b) 是一个 NCP 对, 则称 $\phi : \mathbf{R}^2 \to \mathbf{R}$ 是一个 NCP 函数.

最著名的两个 NCP 函数是 min 函数和 Fischer-Burmeister NCP 函数. 其中 Fischer-Burmeister 函数具有非常简单的结构

$$\psi(a, b) = \sqrt{a^2 + b^2} - a - b. \tag{4}$$

函数 ψ 除了在原点之外处处连续可微, 但是它在原点处是强半光滑的. 也就是说, 如果 $a \neq 0$ 或者 $b \neq 0$, 则 ψ 在点 $(a, b) \in \mathbf{R}^2$ 连续可微, 且

$$\nabla\psi(a, b) = \left(\frac{a}{\sqrt{a^2 + b^2}} - 1,\ \frac{b}{\sqrt{a^2 + b^2}} - 1\right); \tag{5}$$

如果 $a = 0$, $b = 0$, 则 ψ 在 $(0, 0)$ 的广义 Jacobi 矩阵为

$$\partial\psi(0, 0) = \{\eta - 1, \theta - 1 | \eta^2 + \theta^2 = 1\}. \tag{6}$$

构造等价于 KKT 点条件的半光滑方程 $\Phi(\mathbf{x}, \Lambda) = 0$. 令 $\Phi(\mathbf{x}, \Lambda) = (\nabla_{x_1} L(\mathbf{x}, \Lambda), \cdots, \nabla_{x_n} L(\mathbf{x}, \Lambda),$ $\phi_1(\mathbf{x}, \Lambda), \cdots, \phi_m(\mathbf{x}, \Lambda))^T = (\nabla_{\mathbf{X}} L(\mathbf{x}, \Lambda), \Phi_1(\mathbf{x}, \Lambda))$, 其中 $\phi_i(\mathbf{x}, \Lambda) = \psi((-g_i(\mathbf{x})), \lambda_i)$, $1 \leq i \leq m$.

定义 3 F-函数 如果对于任一序列 $\{T_i \in [0, +\infty)\}$

$$\lim_{i \to +\infty} \omega(T_i) = 0 \quad 则 \quad \lim_{i \to +\infty} T_i = 0. \tag{7}$$

称函数 $\omega : [0, +\infty) \to [0, \infty)$ 是强迫函数(简称F-函数).

定义 4 线搜索的非单调F-规则 设 M 是一个非负整数. 对每个 k, 令 $m(k)$ 满足

$$m(0) = 0, \quad 0 \leq m(k) \leq \min[m(k-1) + 1, M], \quad k \geq 1. \tag{8}$$

令 $t^k \geq 0$ 有上界且满足

$$f(\mathbf{x}^k + t^k \mathbf{d}^k) \leq \max_{0 \leq j \leq m(k)} [f(\mathbf{x}^{k-j})] - \omega(T_k). \tag{9}$$

其中 ω 是一个强迫函数.

本文安排如下: 第二节提出了一种非可行域 QP-free 算法. 在第三节中证明算法的可行性. 第四节证明算法的全局收敛性, 且在适当假设下具有超线性收敛性. 第五节给出了简要的总结和数值实验.

1 算法

定义指标集 I_0 和 I_1如下:

$$I_1(\mathbf{x}, \mu) = \{i | (g_i(\mathbf{x}), \mu_i) \neq (0, 0)\}; \quad I_0(\mathbf{x}, \mu) = \{i | (g_i(\mathbf{x}), \mu_i) = (0, 0)\}. \tag{10}$$

如果 $j \in I_0$, 令 $(\xi_j(\mathbf{x}, \mu), \eta_j(\mathbf{x}, \mu)) = (1 - \frac{\sqrt{2}}{2}, -1 + \frac{\sqrt{2}}{2})$; 否则, 令 $(\xi_j(\mathbf{x}, \mu), \eta_j(\mathbf{x}, \mu)) = \nabla\psi(-a, b)|_{a = g_j(\mathbf{x}), b = \mu_j}$. 显然有 $\xi_j(\mathbf{x}, \mu) \geq 0$, $\eta_j(\mathbf{x}, \mu)) \leq 0$ 且

$$\xi_j(\mathbf{x}, \mu) = \left(\frac{g_j(\mathbf{x})}{\sqrt{(g_j(\mathbf{x}))^2 + \mu_j^2}} + 1\right), \qquad \eta_j(\mathbf{x}, \mu) = \left(\frac{\mu_j}{\sqrt{(g_j(\mathbf{x}))^2 + \mu_j^2}} - 1\right). \tag{11}$$

令 $\xi_j^k = \xi_j(\mathbf{x}^k, \mu^k)$, $\eta_j^k = \eta_j(\mathbf{x}^k, \mu^k)$,

$$\mathbf{V}^k = \begin{pmatrix} \mathbf{V}_{11}^k & \mathbf{V}_{12}^k \\ \mathbf{V}_{21}^k & \mathbf{V}_{22}^k \end{pmatrix} = \begin{pmatrix} \mathbf{H}^k + \hat{c}_1^k \mathbf{I}_n & \nabla\mathbf{G}^k \\ diag(\xi^k)(\nabla\mathbf{G}^k)^T & diag(\eta^k - c^k) \end{pmatrix}, \tag{12}$$

其中 $\mathbf{I}_n$ 是 n 维单位矩阵, $\hat{c}_1^k = m_1 min\{1, \|\Phi^k\|^\nu\}$, $m_1 > 0$. 当 $\eta_j^k = 0$ 时, $c_j^k = c min\{1, \|\Phi^k\|^\nu\}$, $c > 0$; 当 $\eta_j^k \neq 0$ 时, $c_j^k = 0$. $diag(\xi^k)$ 和 $diag(\eta^k - c^k)$ 分别表示第 j 个对角元素为 $\xi_j(\mathbf{x}^k, \mu^k)$ 和 $\eta_j(\mathbf{x}^k, \mu^k) - c_j^k$ 的对角矩阵.

算法 1.

步骤 0. 初始化. 给出初始点 $\mathbf{x}^0 \in D$, $0 < m_1 < 1$, $m_2 > 1$, $\tau \in (0,1)$, $\nu > 1$, $\bar{\nu} > 1$, $\bar{\mu} > 0$, $1 > \theta > \theta_1 > 0$, $\theta_2 \in (0,1)$, $c > 0$. 令 $\mu^0 = 0$, $P^0 = 0$, $S^0 = S(\mathbf{x}^0, \mu^0) = f(\mathbf{x}^0) + m_2(\|\Phi_1(\mathbf{x}^0, \mu^0)\|^2 + P^0)$. 给出初始正定矩阵 $\mathbf{H}^0$, 令 $k = 0$.

步骤 1. 计算搜索方向. 通过解下面线性方程组(13)获得 $\mathbf{d}^{k0}$ 和 λ^{k0}.

$$\mathbf{V}^k \begin{pmatrix} \mathbf{d} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f^k \\ 0 \end{pmatrix}. \tag{13}$$

通过解下面线性方程组(14)获得 $\mathbf{d}^{k1}$ 和 λ^{k1}:

$$\mathbf{V}^k \begin{pmatrix} \mathbf{d} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla L^k \\ -\Phi_1^k \end{pmatrix}. \tag{14}$$

步骤 2. 线搜索. 如果下面的不等式成立.

$$\begin{aligned} &S^k - [f(\mathbf{x}^k + \mathbf{d}^{k1}) + m_2(\|\Phi_1(\mathbf{x}^k + \mathbf{d}^{k1}, \mu^k + \lambda^{k1})\|^2 + \gamma^k)] \\ &\geq -\min\{1, \|\Phi^k\|^\nu\}\theta_1(\mathbf{d}^{k0})^T \nabla f^k + \theta_2 m_2 \|\Phi_1^k\|^2, \end{aligned} \tag{15}$$

其中$\gamma^k = \min\{P^k, \|\Phi(\mathbf{x}^k + \mathbf{d}^{k1}, \mu^k + \lambda^{k1})\|^{1/2}\}$, $k \geq 1$, 则令 $\left(\mathbf{d}^k, \lambda^k\right) = \left(\mathbf{d}^{k1}, \lambda^{k1}\right)$, 转步骤 3. 否则, 若 $\mathbf{d}^{k0} = 0$ 则令 $b^k = 0$, $\rho^k = 1$; 若 $\mathbf{d}^{k0} \neq 0$ 则定义 $b^k = (1 - \rho^k)$ 且

$$\rho^k = \begin{cases} 1 & \text{如果}(\mathbf{d}^{k1})^T \nabla f^k \leq \theta(\mathbf{d}^{k0})^T \nabla f^k, \\ (1-\theta)\frac{(\mathbf{d}^{k0})^T \nabla f^k}{(\mathbf{d}^{k0} - \mathbf{d}^{k1})^T \nabla f^k} & \text{否则}; \end{cases} \tag{16}$$

F-规则 线搜索. 设 $\eta \in [0,1]$, $M \geq 1$ 是一个正整数, 定义 $m(k) = \min[k+1, M]$, 选择

$$\eta_{kr} \geq \eta, \quad r = 0, 1, 2, \cdots, m(k) - 1 \ \text{ 和 } \ \sum_{r=0}^{m(k)-1} \eta_{kr} = 1$$

令

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

且

$$\begin{aligned} & S(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k) = f(\mathbf{x}^k + \mathbf{d}^k) \\ + \quad & m_2(\|\Phi_1(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k)\|^2 + \min\{P^k, \|\Phi(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k)\|^{1/2}\}, \end{aligned} \tag{17}$$

寻找 $t^k = \tau^j$, j 是满足下式(18)的最小非负整数,

$$S(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k) \leq \max\{S(\mathbf{x}^k, \mu^k), \sum_{j=0}^{m(k)-1} \eta_{kr} S(\mathbf{x}^{k-j}, \mu^{k-j})\} - \omega(T_k) \tag{18}$$

其中 ω 是一个强迫函数且

$$T_k = -t^k \min\{1, \|\Phi^k\|^\nu\}\theta_1 b^k (\mathbf{d}^{k0})^T \nabla f^k + \theta_2 m_2 (t^k)^2 \rho^k \|\Phi_1^k\|^2.$$

步骤 3. 更新. 令 $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{d}^k$. 如果 $\|\mu^k + \lambda^k\| \leq \bar{\mu}$, 则 $\mu^{k+1} = \mu^k + \lambda^k$; 否则 $\mu_i^{k+1} = \bar{\mu}$, $\bar{\mu} = \bar{\nu}\bar{\mu}$. 令 $P^1 = f^0 + m_2\|\Phi_1^0\|^{1/2} - f^1 - m_2\|\Phi_1^1\|^{1/2}$, 对于 $k \geq 1$, $P^{k+1} = \min\{P^k, \|\Phi(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k)\|^{1/2}\}$. $S^{k+1} = f(\mathbf{x}^{k+1}) + m_2(\|\Phi_1^{k+1}\|^2 + P^{k+1})\}$. 如果 $\mathbf{x}^{k+1}$ 是一个 KKT 点, 算法停止; 否则更新 $\mathbf{H}^k$ 并得到一个对称正定矩阵 $\mathbf{H}^{k+1}$. 令 $k = k+1$ 并转步1.

2 算法的可执行性

本节假定:

A1 水平集 $\{\mathbf{x}|f(\mathbf{x}) \le f(\mathbf{x}^0)\}$ 有界, 且对充分大的 k, $\|\mu^k + b^k\lambda^{k0} + \rho^k\lambda^{k1}\| < \bar{\mu}$.

A2 f 和 g_i 是二次 Lipschitz 连续可微的, 且对所有 $\mathbf{y}, \mathbf{z} \in \mathbf{R}^{n+m}$,

$$\|\nabla L(\mathbf{y}) - \nabla L(\mathbf{z})\| \le m_3\|\mathbf{y}-\mathbf{z}\|, \qquad \|\Phi(\mathbf{y}) - \Phi(\mathbf{z})\| \le m_3\|\mathbf{y}-\mathbf{z}\|,$$

其中 $m_3 > 0$ 是 Lipschitz 常数.

A3 $\mathbf{H}^k$ 是正定的且存在一个正数 m_4 使得对于所有的 $\mathbf{d} \in \mathbf{R}^n$ 和所有的 k 有 $0 < \mathbf{d}^T\mathbf{H}^k\mathbf{d} \le m_4\|\mathbf{d}\|^2$.

易证下面引理成立(参见[5]).

引理 1 如果 $\Phi^k \ne 0$, 则 $\mathbf{V}^k$ 是非奇异的.

引理 1 说明 (13) 或 (14) 总有唯一解.

引理 2 对任意 $0 < t \le 1$, 存在一个 $m_5 > 0$ 使得,

$$\|\Phi_1(\mathbf{x}^k + t\mathbf{d}^{k0}, \mu^k + t\lambda^{k0})\|^2 - \|\Phi_1^k\|^2 \le m_5 t^2.$$

证明: 当 $\Phi_1^k = 0$ 时

$$\|\Phi_1(\mathbf{x}^k + t\mathbf{d}^{k0}, \mu^k + t\lambda^{k0})\|^2 = \|\Phi_1(\mathbf{x}^k + t\mathbf{d}^{k0}, \mu^k + t\lambda^{k0}) - \Phi_1^k\|^2 \le t^2 m_3^2\|(\mathbf{d}^{k0}, \lambda^{k0})\|^2,$$

其中 m_3 是 Lipschitz 常数. 引理对 $\Phi_1^k = 0$ 成立.

对于 $\Phi_1^k \ne 0$ 成立. 易证下面式子成立(参见 [4] 引理 3).

$$\begin{aligned}
&\|\Phi_1(\mathbf{x}^k + t\mathbf{d}^{k0}, \mu^k + t\lambda^{k0})\|^2 \\
&\le (\|\Phi_1(\mathbf{x}^k + t\mathbf{d}^{k0}, \mu^k + t\lambda^{k0}) - \Phi_1^k - t(diag(\bar{\xi}^{k0})(\nabla\mathbf{G}^k)^T\mathbf{d}^{k0} + diag(\bar{\eta}^{k0})\lambda^{k0})\| \\
&+\|\Phi_1^k + t(diag(\bar{\xi}^{k0})(\nabla\mathbf{G}^k)^T\mathbf{d}^{k0} + diag(\bar{\eta}^{k0})\lambda^{k0})\|)^2 = \|\Phi_1^k\|^2 + O(t^2).
\end{aligned}$$

引理得证. □

设

$$\begin{pmatrix} \mathbf{d}^{kt} \\ \lambda^{kt} \end{pmatrix} = tb^k \begin{pmatrix} \mathbf{d}^{k0} \\ \lambda^{k0} \end{pmatrix} + t^2\rho^k \begin{pmatrix} \mathbf{d}^{k1} \\ \lambda^{k1} \end{pmatrix}.$$

引理 3 如果 $\Phi_1^k \ne 0$, $b^k = 0$ 则对任意 $\varepsilon > 0$ 存在 一个 $\bar{t} > 0$ 满足, 对任意 $0 < t \le \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 \ge (2-\varepsilon)t^2\|\Phi_1^k\|^2.$$

如果 $\Phi_1^k = 0$ 且 $b^k \ne 0$, 或者如果 $\Phi_1^k \ne 0$ 且 $b^k \ne 0$, 则存在一个 $m_6 > 0$ 满足, 对任意 $0 < t \le 1$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 \ge -m_6 t^2.$$

如果 $\Phi_1^k = 0$ 且 $b^k = 0$, 则存在一个 $m_6 > 0$ 满足, 对任意$0 < t \le 1$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 \ge -m_6 t^4.$$

证明: 如果 $\Phi_1^k \ne 0$ 且 $b^k = 0$ 则 $\rho^k = 1$, $(\mathbf{d}^{kt}, \lambda^{kt}) = t^2(\mathbf{d}^{k1}, \lambda^{k1})$. (14)式说明

$$diag(\xi^k)(\nabla\mathbf{G}^k)^T\mathbf{d}^{k1} + diag(\eta^k - c^k)\lambda^{k1} = -\Phi_1^k. \tag{19}$$

对于所有 i, 易得(参见[5])

$$\phi_i(\mathbf{x}^k + t^2\mathbf{d}^{k1}, \mu^k + t^2\lambda^{k1}) - \phi_i^k - t^2(\bar{\xi}_i^{k1}(\nabla g_i^k)^T\mathbf{d}^{k1} + (\bar{\eta}_i^{k1})\lambda_i^{k1}) = o(t^2). \tag{20}$$

因为当 $c_i^k \ne 0$ 时, $\eta_i^k = 0$, $\phi_i^k = 0$, 所以 $(\Phi_1^k)^T(diag(\bar{\xi}^{k1})(\nabla\mathbf{G}^k)^T, diag(\bar{\eta}^{k1})) = (\Phi_1^k)^T (diag(\xi^k)(\nabla\mathbf{G}^k)^T, diag(\eta^k - c^k))$. 且对任意 $\varepsilon > 0$, 存在一个 $\bar{t} > 0$ 使得, 对于任意 $0 < t \le \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + t^2\mathbf{d}^{k1}, \mu^k + t^2\lambda^{k1})\|^2 \geq (2-\varepsilon)t^2\|\Phi_1^k\|^2.$$

因此,引理的第一部分得证.

如果 $\Phi_1^k = 0$ 且 $b^k \neq 0$, 则

$$\|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 = \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt}) - \Phi_1^k\|^2 \leq (m_3)^2\|(\mathbf{d}^{kt}, \lambda^{kt})\|^2.$$

引理对于 $\Phi_1^k = 0$ 成立.

如果 $\Phi_1^k \neq 0$ 且 $b^k \neq 0$, 则由引理 2 得

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + tb^k\mathbf{d}^{k0}, \mu^k + tb^k\lambda^{k0})\|^2 = O(t^2)$$

从 $(\mathbf{d}^{kt}, \lambda^{kt}) - (tb^k\mathbf{d}^{k0}, tb^k\lambda^{k0}) = t^2\rho^k(\mathbf{d}^{k1}, \lambda^{k1})$ 中得

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 = O(t^2).$$

于是, 引理的第二部分得证.

这个引理的第三个结论可由在引理 2 的证明过程中分别用 $k1$ 和 t^2 代替 $k0$ 和 t 得到. □

易证以下引理 4-5 (见文献 [5]).

引理 4 如果 $\mathbf{d}^{k0} \neq 0$, 则

$$(\mathbf{d}^{k0})^T(\mathbf{H}^k + \hat{c}_1^k\mathbf{I}_n)\mathbf{d}^{k0} \leq -(\mathbf{d}^{k0})^T\nabla f^k$$

且对于所有 $0< t < 1$,

$$(\mathbf{d}^{kt})^T\nabla f^k \leq [(t-t^2)b^k + t^2\theta^k](\mathbf{d}^{k0})^T\nabla f^k.$$

引理 5 $\mathbf{d}^{k0} = 0$ 当且仅当 $\nabla f^k = 0$. 如果 $\mathbf{d}^{k0} = 0$ 则 $\lambda^{k0} = 0$.

从引理 5 得出如果 $\mathbf{d}^{k0} = 0$, $\Phi_1^k = 0$ 则 $\mathbf{x}^k$ 是一个可行点且 $(\mathbf{x}^k, 0)$ 是一个 KKT 点. 不失一般性, 假定这一方法不在任一 k 终止, 即对于所有的 k, $(\mathbf{d}^{k0}, \Phi_1^k) \neq (0,0)$.

引理 6 如果 $\mathbf{d}^{k0} \neq 0$ 或者 $\Phi_1^k \neq 0$ 则存在一个 $\bar{t} > 0$ 使得, 对于任意的 $t = t^k \in (0, \bar{t})$, (18) 成立.

证明: 显然,对于任意的 t, $P^k - \min\{P^k, \|\Phi(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^{1/2}\} \geq 0$. 分三种情况证明这个引理.

情况 1. 如果 $\mathbf{d}^{k0} \neq 0$ 且 $b^k \neq 0$, 由f 和 g_i 的连续可微性以及引理 4, 有

$$f^k - f(\mathbf{x}^k + \mathbf{d}^{kt}) \geq -[(t-t^2)b^k + t^2\theta^k](\mathbf{d}^{k0})^T\nabla f^k + o(t).$$

引理 2 和 3 说明存在一个 m_7, 对于任意的 $0 < t \leq 1$ 满足,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + \mathbf{d}^{kt}, \mu^k + \lambda^{kt})\|^2 \geq -m_7t^2.$$

由引理 4 和假设 A3, 知 $-(\mathbf{d}^{k0})^T\nabla f^k \geq (\mathbf{d}^{k0})^T\mathbf{H}^k\mathbf{d}^{k0} \geq 0$. 引理对于这种情况成立.

情况 2. 如果 $\mathbf{d}^{k0} \neq 0$, $b^k = 0$, 由 b^k 的定义, 有 $\rho^k = 1$, $(\mathbf{d}^{k1})^T\nabla f^k \leq \theta(\mathbf{d}^{k0})^T\nabla f^k$ 且

$$f^k - f(\mathbf{x}^k + t^2\mathbf{d}^{k1}) \geq -t^2\theta(\mathbf{d}^{k0})^T\nabla f^k + o(t^2).$$

如果 $\Phi_1^k = 0$ 则由引理 3 知, 对于任意的 $0 < t \leq 1$,

$$\|\Phi_1(\mathbf{x}^k + t^2\rho^k\mathbf{d}^{k1}, \mu^k + t^2\rho^k\lambda^{k1})\|^2 \leq m_6t^4.$$

如果 $\Phi_1^k \neq 0$ 则由引理 3 知, 给定任意的 $\varepsilon > 0$ 存在一个 $\bar{t} > 0$ 使得, 对于任意的 $0 < t \leq \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + t^2\mathbf{d}^{k1}, \mu^k + t^2\lambda^{k1})\|^2 \geq (2-\varepsilon)t^2\|\Phi_1^k\|^2.$$

所以, 引理对于这种情况成立.

情况 3. 如果 $\mathbf{d}^{k0} = 0$ 则由引理 5 有 $\lambda^{k0} = 0$ 和 $\nabla f^k = 0$. 又 $b^k = 0$ 和 $\rho^k = 1$. 由这个引理的假设, 可知 $\Phi_1^k \neq 0$. 又由引理 3 得对任意的 $\varepsilon > 0$ 存在一个 $\bar{t} > 0$ 使得, 对于任意的 $0 < t \leq \bar{t}$,

$$\|\Phi_1^k\|^2 - \|\Phi_1(\mathbf{x}^k + t^2\mathbf{d}^{k1}, \mu^k + t^2\lambda^{k1})\|^2 \geq (2-\varepsilon)t^2\|\Phi_1^k\|^2.$$

且由 f 的连续可微性, 有

$$f^k - f(\mathbf{x}^k + t^2\rho^k\mathbf{d}^{k1}) = o(t^2). \tag{21}$$

所以, 引理对于这种情况也成立. 引理得证. □

引理 1-6 充分证明了算法 1 是可以实现的.

3 收敛性

本节同样假设 A1-A3 成立. 从 $\mathbf{x}^0 \in D$ 和 $\mu^0 = 0$ 得出 $\Phi_1^0 = 0$ 和 $\Phi^0 = (\nabla f^0, 0)$. 从 S^k 单调下降, 得 $f^k \le S^k \le S^1 = S^0 = f^0$. 而假设 A1 意味着 $\{(\mathbf{x}^k, \mu^k)\}$ 是有界的.

引理 7 如果 $(\mathbf{x}^*, \mu^*)$ 是 $\{(\mathbf{x}^k, \mu^k)\}$ 的一个聚点, $S^k - S^{k+1} > 0$, 则存在无穷指标集 $\mathcal{K}$, 自然数 K, $t > 0$ 和正数 $\delta > 0$, 使得对于任意的 $k(i) > K$ 和 $k(i) \in \mathcal{K}$, 有

$$S^{k(i)} - S^{k(i)}(t) > \delta/2 > 0. \tag{22}$$

其中对于某一 $0 < \bar{t} \le t \le \bar{t}_1 \le 1$,

$$\begin{aligned} & S^{k(i)}(t) = f(\mathbf{x}^k + \mathbf{d}^k) \\ + \quad & m_2(\|\Phi_1(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k)\|^2 + \min\{P^k, \|\Phi(\mathbf{x}^k + \mathbf{d}^k, \mu^k + \lambda^k)\|^{1/2}\}). \end{aligned} \tag{23}$$

证明: 不妨假定 $b^{k(i)}(\mathbf{x}^{k(i)}, \mu^{k(i)}) \to b^*(\mathbf{x}^*, \mu^*)$, $\rho^{k(i)}(\mathbf{d}^{k(i)0}, \lambda^{k(i)0}) \to \rho^*(\mathbf{d}^{*0}, \lambda^{*0})$, $(\mathbf{d}^{k(i)1}, \lambda^{k(i)1}) \to (\mathbf{d}^{*1}, \lambda^{*1})$, $S^{k(i)} \to S^*$, $\mathbf{H}^{k(i)} \to \mathbf{H}^*$, $\hat{c}_1^{k(i)} \to \hat{c}_1^*$ 和 $c^{k(i)} \to c^*$.

令 $\Phi^* = \Phi(\mathbf{x}^*, \mu^*)$. 如果 $\Phi^* = 0$, 则 $(\mathbf{x}^*, \mu^*)$ 是问题 (NLP) 的一个 KKT 点. 如果 $\Phi^* \ne 0$, $\Phi_1^* = 0$, $\mathbf{d}^{*0} = 0$, 则 $\mathbf{x}^*$ 是一个可行点. 在引理 5 中用 $*$ 来代替 $k0$, 从 $c\min\{1, \|\Phi^*\|^\nu\} \ne 0$ 可推出 $\nabla f^* = 0$, 且 $(\mathbf{x}^*, 0)$ 是问题(NLP)的一个KKT点. 如果 $\Phi_1^* \ne 0$ 或 $\mathbf{d}^{k0} \ne 0$, 由引理 2-6 的证明以及 f 和 g_i 的连续可微性, 知存在正常数 $\bar{t}, \bar{t}_1, \bar{\theta}_1, \bar{\theta}_2$, $0 < \bar{t} \le t \le \bar{t}_1 \le 1, \bar{t} < \tau\bar{t}_1, \theta_1 < \bar{\theta}_1 \le 1$ 和 $\theta_2 < \bar{\theta}_2 < 1$, 使得

$$\begin{aligned} & S^* - [f(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}) + m_2(\|\Phi_1(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}, \mu^* + t^2\rho^*\lambda^{*1})\|^2 \\ & + \min\{P^k, \|\Phi(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}, \mu^* + t^2\rho^*\lambda^{*1})\|^{1/2}\})] \\ \ge \quad & -\min\{1, \|\Phi^*\|^\nu\}\bar{\theta}_1 t(\mathbf{d}^{*0})^T\nabla f^* + m_2\bar{\theta}_2 t^2\rho^*\|\Phi_1^*\|^2. \end{aligned} \tag{24}$$

从 (24) 推出存在一个 $\delta > 0$ 满足

$$\begin{aligned} & S^* - [f(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}) \\ & + m_2(\|\Phi_1(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}, \mu^* + t^2\rho^*\lambda^{*1})\|^2 \\ & + \min\{P^k, \|\Phi(\mathbf{x}^* + tb^*\mathbf{d}^{*0} + t^2\rho^*\mathbf{d}^{*1}, \mu^* + t^2\rho^*\lambda^{*1})\|^{1/2}\})] \ge \delta > 0. \end{aligned} \tag{25}$$

f 和 g_i 的连续可微性说明对于 $t \in (0,1)$, $\Phi(\mathbf{x}^{k(i)} + \mathbf{d}^{k(i)t}, \mu^{k(i)} + \lambda^{k(i)t})$ 一致收敛于 $\Phi(\mathbf{x}^* + \mathbf{d}^{*t}, \mu^* + \lambda^{*t})$. 所以, 存在一个 K 使得, 对于所有的 $k(i) > K$ 和 $0 < \bar{t} \le t \le \bar{t}_1 \le 1$,

$$S^{k(i)} - S^{k(i)}(t) > \delta/2 > 0. \tag{26}$$

引理得证. □

类似于文献 [8] 中定理 2 的证明, 有

引理 8 对于所有 k

$$S(\mathbf{x}^0) - S(\mathbf{x}^{k+1}) \ge \sum_{r=0}^{m(k)-1} \omega(T_k)$$

从引理 8 可推出 $\omega(T_k) \to 0$ 和

$$-t^k \min\{1, \|\Phi^k\|^\nu\}\theta_1 b^k(\mathbf{d}^{k0})^T\nabla f^k + \theta_2 m_2(t^k)^2\rho^k\|\Phi_1^k\|^2 \to 0.$$

从以上引理 7-8 得出 $\|\Phi_1^k\|^2 \to 0$, $\Phi^* = 0$ 和 $\mathbf{d}^{k0} \to 0$. 显然以下定理成立.

定理 1 如果 $(\mathbf{x}^*, \mu^*)$ 是 $\{(\mathbf{x}^k, \mu^k)\}$ 的一个聚点, 则 $\mathbf{x}^*$ 是问题 (NLP) 的一个 KKT 点.

在假设 A1-A4 的基础上补充以下4个假设:

A4 $\{\nabla g_i(\mathbf{x}^*)\}$ 是线性独立的, 其中 $i \in I = \{i | g_i(\mathbf{x}^*) = 0\}$, $\mathbf{x}^*$ 是$\{\mathbf{x}^k\}$的一个聚点且是问题 (NLP) 的一个 KKT 点.

A5 序列 $\{\mathbf{H}^k\}$ 满足

$$\frac{\|(\mathbf{H}^k - \nabla^2_{\mathbf{xx}}L(\mathbf{x}^k, \mu^k))\mathbf{d}^{k1}\|}{\|\mathbf{d}^{k1}\|} \to 0.$$

A6 严格互补条件在每一个 KKT 点 $(\mathbf{x}^*, \mu^*)$ 成立.

A7 f 和 g_i 是三次 Lipschitz 连续可微的.

类似引理 1, 易证下面引理.

引理 9 假设 A1-A7 成立. 则 $\{\|(\mathbf{V}^k)^{-1}\|\}$ 是有界的. 此外, 如果 $\mathbf{V}^*$ 是 $\{\mathbf{V}^k\}$ 的一个聚点矩阵, 则 $\mathbf{V}^*$ 是非奇异的.

引理 10 假设 A1-A7 成立且对于所有 k 有 $\Phi^k \neq 0$, 则 $\lim\limits_{k\to\infty} \|\Phi(\mathbf{x}^k+\mathbf{d}^{k1}, \mu^k+\lambda^{k1})\|/\|\Phi^k\| = 0$.

证明: 由反证法假设存在一个 $m_8 > 0$ 和一个 $\{(\mathbf{x}^k, \mu^k)\}$ 的子序列 $\{(\mathbf{x}^{k(i)}, \mu^{k(i)})\}$, $(\mathbf{x}^{k(i)}, \mu^{k(i)}) \to (\mathbf{x}^*, \mu^*)$, 满足

$$\lim_{k\to\infty} \|\Phi(\mathbf{x}^{k(i)} + \mathbf{d}^{k(i)1}, \mu^{k(i)} + \lambda^{k(i)1})\|/\|\Phi^{k(i)}\| = m_8 > 0. \tag{27}$$

令

$$\bar{\mathbf{V}}^{k(i)} = \begin{pmatrix} \bar{\mathbf{V}}_{11}^{k(i)} & \bar{\mathbf{V}}_{12}^{k(i)} \\ \bar{\mathbf{V}}_{21}^{k(i)} & \bar{\mathbf{V}}_{22}^{k(i)} \end{pmatrix} = \begin{pmatrix} \nabla^2_{\mathbf{xx}} L^{k(i)} & \nabla \mathbf{G}^{k(i)} \\ diag(\xi^{k(i)})(\nabla \mathbf{G}^{k(i)})^T & diag(\eta^{k(i)}) \end{pmatrix}.$$

由假设 A6 得, 对于充分大的 k, $\nabla\Phi^{k(i)}$ 存在且 $\bar{\mathbf{V}}^{k(i)} = \nabla\Phi^{k(i)}$. 有

$$\|\bar{\mathbf{V}}^{k(i)}(\mathbf{d}^{k(i)1}, \lambda^{k(i)1}) - \mathbf{V}^{k(i)}(\mathbf{d}^{k(i)1}, \lambda^{k(i)1})\| = o(\|\mathbf{d}^{k(i)1}\|), \tag{28}$$

从 Φ 的可微性推出

$$\lim_{k\to\infty} [\|\Phi(\mathbf{x}^{k(i)} + \mathbf{d}^{k(i)1}, \mu^{k(i)} + \lambda^{k(i)1}) - \Phi^{k(i)} - \bar{\mathbf{V}}^{k(i)}(\mathbf{d}^{k(i)1}, \lambda^{k(i)1})\|]/\|(\mathbf{d}^{k(i)1}, \lambda^{k(i)1})\| = 0. \tag{29}$$

(29) 和 (28) 说明

$$\|\Phi(\mathbf{x}^{k(i)} + \mathbf{d}^{k(i)1}, \mu^{k(i)} + \lambda^{k(i)1})\| = o(\|\Phi^{k(i)}\|). \tag{30}$$

(27) 和 (30) 矛盾, 故引理得证. □

易证下面引理(参见[4])

引理 11 假定 A1-A7 成立, 则对充分大的 k, $(\mathbf{d}^k, \lambda^k) = (\mathbf{d}^{k1}, \lambda^{k1})$.

由引理 8-11 得出以下定理:

定理 2 假定 A1-A7 成立. 设执行 算法 1 后生成一个序列 $\{(\mathbf{x}^k, \lambda^k)\}$ 且 $(\mathbf{x}^*, \lambda^*)$ 是 $\{(\mathbf{x}^k, \lambda^k)\}$ 的一个聚点, 则 $(\mathbf{x}^*, \lambda^*)$ 是 问题 (NLP) 的一个 KKT 点, 且 $(\mathbf{x}^k, \lambda^k)$ 超线性收敛于 $(\mathbf{x}^*, \lambda^*)$.

4 数值实验.

本节对于求解约束优化问题的算法 1 给出一些数值结果, 执行的细节描述如下.

1. 终止准则. 因为 $\Phi(\mathbf{x}, \mu) = 0$ 当且仅当 $(\mathbf{x}, \mu)$ 是一个 KKT 点, 终止准则为 $\|\Phi^k\| \le 10^{-5}$.

2. 更新 $\mathbf{H}^k$. $\mathbf{H}^k$ 由 BFGS 方法更新(参见[6]).

3. 参数设置如下: $\omega(t)$=t. $m_1 = 0.1$, $m_2 = 10$, $\beta = \bar{\nu} = 1.5$, $\nu = 2$, $\tau = 0.5$, $\alpha = 0.2$, $\alpha_1 = \alpha_2 = 0.1$ 以及 $\bar{\mu} = 10000$.

文献 [9] 中问题的号码	方法	
	迭代的次数	计算目标函数值和约束的次数
29	8	13
30	8	10
33	10	23
35	7	12
37	17	46
43	12	25
44	17	39
76	10	42
100	15	37
113	22	58

表 1

表 1 中的数值结果显示这算法是非常有效的 .

参考文献

[1] Li K, Pu D, Tian W. Filter QP-free Method with 3-piecewise linear NCP function[J]. OR Transactions, 2008, 12(2):49.
[2] Li K, Pu D, Tian W. Filter QP-free Method with piecewise linear NCP function[J]. J. Pacific J. of Optimization, 2008, 4(1):43.
[3] Panier E R, Tits A L, Herskovits J N. A QP-free, globally, locally superlinear convergent method for the inequality constrained optimization problems[J]. SIAM J. Control Optimization, 1988, 36:788.
[4] Pu D, Li K, Xue W. Convergence of QP-Free Infeasible Methods for Nonlinear Inequality Constrained Optimization Problems[J]. J. of Tongji University:Natural Science, 2005, 33(4):525.
[5] Pu D, Zhou Y, Zhang Z. A QP Free Feasible Method[J]. J. of Computational Mathematics, 2004, 22:651.
[6] Qi H, Qi L. A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization[J]. SIAM J. Optimization, 2000, 11:113.
[7] Zhou Y, Pu D. A new QP-free feasible method for inequality constrained optimization[J]. OR Transactions, 2007, 11(3):31.
[8] Yu Z, Pu D. A new nonmonotone line search technique for unconstrained optimization[J]. J. of Computational and Applied Mathematics, 2008, 219(1):134.
[9] Schittkowski K. More test examples for nonlinear programming codes[M]. Springer-Verlag, 1988.