文章编号: 1000- 0887(2009)05- 0613- 10

非线性互补约束均衡问题的一个 SQP 算法*

朱志斌[1], 简金宝[2], 张 聪[1]

(1. 桂林电子科技大学 数学与计算科学学院, 广西 桂林 541004;
2. 广西大学 数学与信息科学学院, 南宁 530004)

(张石生推荐)

摘要: 提出了一个求解非线性互补约束均衡问题(MPCC)的逐步逼近光滑 SQP 算法. 通过一系列光滑优化来逼近MPCC. 引入 l_1 精确罚函数, 线搜索保证算法具有全局收敛性. 进而, 在严格互补及二阶充分条件下,算法是超线性收敛的. 此外,当算法有限步终止, 当前迭代点即为MPEC 的一个精确稳定点.

关 键 词: 均衡问题; 序列二次规划算法; 逐步逼近; 全局收敛; 超线性收敛速率
中图分类号: O221.2 文献标识码: A
DOI: 10.3879/j.issn.1000- 0887.2009.05.012

引 言

本文考虑如下带非线性互补约束的均衡问题(MPEC):

$$\begin{cases} \min \quad f(\boldsymbol{x}, \boldsymbol{y}), \\ \text{s.t.} \quad g(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0,\ 0 \leqslant F(\boldsymbol{x}, \boldsymbol{y}) \perp \boldsymbol{y} \geqslant 0, \end{cases} \tag{1}$$

其中$f: R^{n+m} \to \mathbf{R}$, $g: R^{n+m} \to R^l$, $F: R^{n+m} \to R^m$ 二阶连续可微, $\boldsymbol{a} \perp \boldsymbol{b}$ 表示向量$\boldsymbol{a}$ 与向量$\boldsymbol{b}$ 正交. 该类问题大量出现在经济、工程设计等各领域中.

众所周知, 由于互补约束条件的存在, MPEC 问题(1) 的求解是一个十分有意义而又比较困难的课题. 因为文献[1] 指出: 对于该类优化问题, 即使较弱的MFCQ 条件, 在任意可行点处都不满足, 从而非线性规划一些经典的算法一般不能直接应用到均衡问题上来.

利用与 Fischer- Burmeister 函数 $\phi(a, b) = a + b - \sqrt{(a^2 + b^2)}$ 性质类似的光滑函数族 $\phi(\cdot, \cdot, \mu)$, 文献[2] 提出了两个全局收敛的 SQP 算法. 然而, 该文并未对算法的超线性收敛速率作出分析. 此外, 在理论上, 由于参数 μ 的存在, 显光滑 SQP 算法在有限步终止时得不到问题(1) 的精确稳定点.

本文对文献[2- 4] 的方法作进一步研究, 利用 Fischer- Burmeister 函数和逐次逼近思想[4], 给出了一个 MPEC 问题(1) 的等价形式. 进一步, 用SQP 算法求解此等价形式. 若非退化条件成立, 算法全局收敛. 不同于文献[2- 3], 我们的算法在理论上保证了当算法有限步终止时, 当

* 收稿日期: 2008- 07-19; 修订日期: 2009-02-27
基金项目: 国家自然科学基金资助项目(10501009; 10771040); 广西壮族自治区自然科学基金资助项目(0728206; 0640001); 中国博士后基金资助项目(20070410228)
作者简介: 朱志斌(1974—), 男, 湖南双峰人, 教授, 博士(联系人. E- mail: zhuzbma@hotmail.com).

前迭代点即为 MPEC 问题(1)的一个精确解.

1 预备知识

记问题(1)的可行集为

$$\mathscr{F}=\{z=(x,y): g(z)\geqslant 0, 0\leqslant F(z)\perp y\geqslant 0\},$$

且简记

$$X=\{z: g(z)\geqslant 0\}, L_1=\{1,\cdots,l\}, L_2=\{1,\cdots,m\},$$
$$I(z)=\{j\in L_1: g(z)=0\}.$$

对 $\forall z=(x,y)$, 将指标集 $L_2=\{1,2,\cdots,m\}$ 分解为如下 3 个互不相交的子集:

$$\alpha(z)=\{1\leqslant i\leqslant m: F_i(z)<y_i\},\quad \beta(z)=\{1\leqslant i\leqslant m: F_i(z)=y_i\},$$
$$\gamma(z)=\{1\leqslant i\leqslant m: F_i(z)>y_i\},$$

且在 z 点处定义指标集 $\mathscr{A}(z)$ 如下:

$$\mathscr{A}(z)=\{(\mathscr{J},\mathscr{K}): \mathscr{J}\supseteq\alpha(z), \mathscr{K}\supseteq\gamma(z), \mathscr{J}\cup\mathscr{K}=\{1,2,\cdots,m\}, \mathscr{J}\cap\mathscr{K}=\varnothing\}.$$

定义 1 称点 $z^*=(x^*,y^*)\in\mathscr{F}$ 是 MPEC 问题(1)的一个分段稳定点或者原始-对偶稳定点, 如果 $\forall(\mathscr{J},\mathscr{K})\in\mathscr{A}(z^*)$, 存在 KKT 乘子 $\xi^*\in R^l$, $\eta^*\in R^m$, $\pi^*\in R^m$, 使得

$$\begin{cases}\nabla f(x^*,y^*)-\nabla g(x^*,y^*)\xi^*-\nabla F(x^*,y^*)\eta^*-\begin{pmatrix}\mathbf{0}_{n\times m}\\ \boldsymbol{I}_{m\times m}\end{pmatrix}\pi^*=0,\\ F_j(x^*,y^*)=0,\ 0\leqslant y_j^*\perp\pi_j^*\geqslant 0,\qquad \forall j\in\mathscr{J},\\ \qquad 0\leqslant F_j(x^*,y^*)\perp\eta_j^*\geqslant 0,\ y_j^*=0,\qquad \forall j\in\mathscr{K},\\ 0\leqslant g(x^*,y^*)\perp\xi^*\geqslant 0.\end{cases}\tag{2}$$

定义 2 称问题 MPEC(1)在点 $(x,y)\in R^{n+m}$ 满足下层非退化条件, 如果

$$y_j\neq F_j(x,y),\qquad \forall j=1,2,\cdots,m.$$

引入 Fischer-Burmeister 函数

$$\phi(a,b)=a+b-\sqrt{a^2+b^2}.\tag{3}$$

类似于文献[5], 利用函数 $\phi(a,b)$, 我们考虑如下与问题(1)等价的光滑优化问题:

$$\begin{cases}\min\ f(x,y)\\ \text{s.t.}\ \ g(x,y)\geqslant 0,\ F(x,y)-w=0,\ \Phi_\varepsilon(y,w)=\mathbf{0},\end{cases}\tag{4}$$

其中

$$\Phi(y,w)=\Phi_\varepsilon(y,w)+\Phi_\varepsilon(y,w),\tag{5}$$

$$\Phi(y,w)=\begin{pmatrix}\phi(y_1,w_1)\\ \vdots\\ \phi(y_m,w_m)\end{pmatrix},\quad \Phi_\varepsilon(y,w)=\begin{pmatrix}\phi_\varepsilon(y_1,w_1)\\ \vdots\\ \phi_\varepsilon(y_m,w_m)\end{pmatrix},$$

$$\Phi_\varepsilon(y,w)=\begin{pmatrix}\phi_\varepsilon(y_1,w_1)\\ \vdots\\ \phi_\varepsilon(y_m,w_m)\end{pmatrix},$$

且对于 $\varepsilon>0$ 及指标集 $E(z,\varepsilon)\stackrel{\triangle}{=}\{j\in L_2\mid\sqrt{y_j^2+w_j^2}<\varepsilon\}$,

$$\phi_\varepsilon(y_j,w_j)=\begin{cases}\phi(y_j,w_j), & j\in L_2\setminus E(z,\varepsilon),\\ \dfrac{2\varepsilon-y_j}{2\varepsilon}y_j+\dfrac{2\varepsilon-w_j}{2\varepsilon}w_j-\dfrac{\varepsilon}{2}, & j\in E(z,\varepsilon).\end{cases}\tag{6}$$

$$\phi_\varepsilon(y_j, w_j) = \begin{cases} 0, & j \in L_2 \setminus E(z, \varepsilon), \\ \dfrac{(\sqrt{y_j^2 + w_j^2} - \varepsilon)^2}{2\varepsilon}, & j \in E(z, \varepsilon). \end{cases} \tag{7}$$

显然, 函数 $\Phi_\varepsilon(\boldsymbol{y}, \boldsymbol{w})$ 是光滑的且如下结论成立.

命题 1 对 $(\boldsymbol{y}, \boldsymbol{w}) \in R^{2m}$, 有

$$\nabla_a \phi_\varepsilon^2(y_j, w_j) + \nabla_b \phi_\varepsilon^2(y_j, w_j) \geqslant 3 - 2\sqrt{2} > 0, \quad j = 1, \cdots, m. \tag{8}$$

显然, 如果 $\boldsymbol{z}^* = (\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{w}^*)$ 是问题(4) 的 KKT 点, 则存在 KKT 乘子向量 $(\lambda_g, \lambda_F, \lambda_\Phi) \in R^{l+2m}$ 使得

$$\begin{cases} \begin{pmatrix} \nabla f(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \boldsymbol{0} \end{pmatrix} - \begin{pmatrix} \nabla g(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \boldsymbol{0} \end{pmatrix} \lambda_g + \begin{pmatrix} \nabla F(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ -\boldsymbol{I} \end{pmatrix} \lambda_F + \\ \qquad \begin{pmatrix} \boldsymbol{0} \\ \nabla \Phi_\varepsilon(\boldsymbol{y}^*, \boldsymbol{w}^*) \end{pmatrix} \lambda_\Phi = \boldsymbol{0}, \\ 0 \leqslant g(\boldsymbol{x}^*, \boldsymbol{y}^*) \perp \lambda_g \geqslant 0, \ F(\boldsymbol{x}^*, \boldsymbol{y}^*) - \boldsymbol{w}^* = 0, \ \Phi_\varepsilon(\boldsymbol{y}^*, \boldsymbol{w}^*) = \boldsymbol{0}. \end{cases} \tag{9}$$

下面的结论表明了 MPEC 问题(1) 与问题(4) 的之间解的联系, 为算法提供了理论依据.

命题 2 假设总是 MPEC(1) 在点 $(\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathscr{F}$ 满足下层非退化条件且 $\Phi(\boldsymbol{y}^*, \boldsymbol{w}^*) = \Phi_\varepsilon(\boldsymbol{y}^*, \boldsymbol{w}^*)$, 则如下两论断是等价的:

1) $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ 是问题 MPEC(1) 的一个分段稳定点;

2) $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{w}^*)$ 是问题(4) 的一个 KKT 点, 其中 $\boldsymbol{w}^* = F(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

2 算 法

根据命题 2, 我们首先考虑求解问题(4). 对任意给定 $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w})$ 及 $\boldsymbol{d} = (\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{y}, \mathrm{d}\boldsymbol{w})$, 我们引入如下二次规划(QP) 来逼近问题(4):

$$\begin{cases} \min & \nabla f(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} + \dfrac{1}{2} \boldsymbol{d}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{d} \\ \text{s.t.} & g(\boldsymbol{x}, \boldsymbol{y}) + \nabla g(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} \geqslant 0, \\ & F(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{w} + \nabla F(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} - \mathrm{d}\boldsymbol{w} = 0, \\ & \Phi_\varepsilon(\boldsymbol{y}, \boldsymbol{w}) + \boldsymbol{A}\mathrm{d}\boldsymbol{y} + \boldsymbol{B}\mathrm{d}\boldsymbol{w} = \boldsymbol{0}, \end{cases} \tag{10}$$

其中 $(\boldsymbol{A}, \boldsymbol{B}) = \nabla \Phi_\varepsilon(\boldsymbol{y}, \boldsymbol{w})$, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为适当的对角矩阵.

有关 QP(10) 的可行性的讨论是 MPEC 问题的一个焦点. 其中, 文献[3] 和[6] 都分别作出了相应论述.

对 $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}) \in R^{n+2m}$, $(\rho, \varepsilon) \in (0, \infty) \times [0, \infty)$, 定义关于问题(4) 的 l_1 精确罚函数 $\theta: R^{n+2m} \times (0, \infty) \times [0, \infty) \to \mathbf{R}$ 如下:

$$\theta(z, \rho, \varepsilon) = f(\boldsymbol{x}, \boldsymbol{y}) + \rho\left[\sum_{j=1}^{m} |F_j(\boldsymbol{x}, \boldsymbol{y}) - w_j| + \sum_{j=1}^{m} |\phi_\varepsilon(y_j, w_j)|\right]. \tag{11}$$

显然, 函数 $\theta(\cdot, \rho, \varepsilon)$ 有方向导数. 为方便, 引入如下指标集:

$$J_+ \equiv J_+(z) \equiv \left\{j : F_j(\boldsymbol{x}, \boldsymbol{y}) > w_j\right\}, \quad J_0 \equiv J_0(z) \equiv \left\{j : F_j(\boldsymbol{x}, \boldsymbol{y}) = w_j\right\},$$

$$J_- \equiv J_-(z) \equiv \left\{j : F_j(\boldsymbol{x}, \boldsymbol{y}) < w_j\right\}, \quad L_+ \equiv L_+(\boldsymbol{y}, \boldsymbol{w}, \varepsilon) \equiv \left\{j : \phi_\varepsilon(y_j, w_j) > 0\right\},$$

$L_0 \equiv L_0(\boldsymbol{y}, \boldsymbol{w}, \varepsilon) \equiv \left\{ j: \phi_\varepsilon(y_j, w_j) = 0 \right\}$, $L_- \equiv L_-(\boldsymbol{y}, \boldsymbol{w}, \varepsilon) \equiv \left\{ j: \phi_\varepsilon(y_j, w_j) < 0 \right\}$,

记 $\theta'(\boldsymbol{z}, \rho, \varepsilon; \boldsymbol{d})$ 为函数 $\theta(\cdot, \rho, \varepsilon)$ 在点 $\boldsymbol{z}$ 处沿方向 $\boldsymbol{d} = (\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{y}, \mathrm{d}\boldsymbol{w})$ 的方向导数, 则有结论:

$$
\begin{aligned}
\theta'(\boldsymbol{z}, \rho, \varepsilon; \boldsymbol{d}) = {} & \nabla f(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} + \rho \sum_{j \in J_+} \begin{pmatrix} \nabla F_j(\boldsymbol{x}, \boldsymbol{y}) \\ -1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \\ \mathrm{d}w_j \end{pmatrix} + \\
& \rho \sum_{j \in J_0} \left| \begin{pmatrix} \nabla F_j(\boldsymbol{x}, \boldsymbol{y}) \\ -1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \\ \mathrm{d}w_j \end{pmatrix} \right| - \rho \sum_{j \in J_-} \begin{pmatrix} \nabla F_j(\boldsymbol{x}, \boldsymbol{y}) \\ -1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \\ \mathrm{d}w_j \end{pmatrix} + \\
& \rho \sum_{j \in L_+} \nabla \phi_\varepsilon(y_j, w_j)^{\mathrm{T}} \begin{pmatrix} \mathrm{d}y_j \\ \mathrm{d}w_j \end{pmatrix} + \rho \sum_{j \in L_0} \left| \nabla \phi_\varepsilon(y_j, w_j)^{\mathrm{T}} \begin{pmatrix} \mathrm{d}y_j \\ \mathrm{d}w_j \end{pmatrix} \right| - \\
& \rho \sum_{j \in L_-} \nabla \phi_\varepsilon(y_j, w_j)^{\mathrm{T}} \begin{pmatrix} \mathrm{d}y_j \\ \mathrm{d}w_j \end{pmatrix}.
\end{aligned} \tag{12}
$$

命题 3 假设 $\boldsymbol{d} = (\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{y}, \mathrm{d}\boldsymbol{w})$ 是 QP 子问题(10) 的一个解且 $(\lambda_g, \lambda_F, \lambda_\Phi)$ 为其对应的乘子向量. 若 $g(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0$ 且 $\rho \geqslant \max\left\{ \left|(\lambda_F)_i\right|, \left|(\lambda_\Phi)_j\right|, 1 \leqslant i, j \leqslant m \right\}$, 则有

$$\theta'(\boldsymbol{z}, \rho, \varepsilon; \boldsymbol{d}) \leqslant -\boldsymbol{d}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{d}.$$

证明 注意到式(10), 我们有

$$
\begin{cases}
\nabla \phi_\varepsilon(y_j, w_j)^{\mathrm{T}} \begin{pmatrix} \mathrm{d}y_j \\ \mathrm{d}w_j \end{pmatrix} = -\phi_\varepsilon(y_j, w_j), & j = 1, \cdots, m, \\
\begin{pmatrix} \nabla F_j(\boldsymbol{x}, \boldsymbol{y}) \\ -1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \\ \mathrm{d}w_j \end{pmatrix} = -F_j(\boldsymbol{x}, \boldsymbol{y}) + w_j, & j = 1, \cdots, m.
\end{cases} \tag{13}
$$

故由式(12), 可知

$$
\begin{aligned}
& \theta'(\boldsymbol{z}, \rho, \varepsilon; \boldsymbol{d}) = \\
& \quad \nabla f(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} - \rho \sum_{j=1}^{m} \left| F_j(\boldsymbol{x}, \boldsymbol{y}) - w_j \right| - \rho \sum_{j=1}^{m} \left| \phi_\varepsilon(y_j, w_j) \right|.
\end{aligned} \tag{14}
$$

又由式(10) 的 KKT 条件可得

$$
\begin{aligned}
& \nabla f(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} + \boldsymbol{d}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{d} - \lambda_g^{\mathrm{T}} \nabla g(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \end{pmatrix} + \\
& \quad \lambda_F^{\mathrm{T}} \begin{pmatrix} \nabla F(\boldsymbol{x}, \boldsymbol{y}) \\ -\boldsymbol{I} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{y} \\ \mathrm{d}\boldsymbol{w} \end{pmatrix} + \lambda_\Phi^{\mathrm{T}} \nabla \Phi_\varepsilon(\boldsymbol{y}, \boldsymbol{w})^{\mathrm{T}} \begin{pmatrix} \mathrm{d}\boldsymbol{y} \\ \mathrm{d}\boldsymbol{w} \end{pmatrix} = \boldsymbol{0}.
\end{aligned}
$$

故而

$$
\begin{aligned}
\theta'(\boldsymbol{z}, \rho, \varepsilon; \boldsymbol{d}) \leqslant {} & -\boldsymbol{d}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{d} + \sum_{j \in J_+} ((\lambda_F)_j - \rho)(F_j(\boldsymbol{x}, \boldsymbol{y}) - w_j) + \\
& \sum_{j \in J_-} ((\lambda_F)_j + \rho)(F_j(\boldsymbol{x}, \boldsymbol{y}) - w_j) + \sum_{j \in L_+} ((\lambda_\Phi)_j - \rho) \phi_\varepsilon(y_j, w_j) + \\
& \sum_{j \in L_-} ((\lambda_\Phi)_j + \rho) \phi_\varepsilon(y_j, w_j) \leqslant \\
& -\boldsymbol{d}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{d}.
\end{aligned}
$$

证毕.

显然, 在命题 3 中, 条件 $g(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0$ 是很重要的. 所以, 在算法迭代中, 我们希望迭代方

向 $\boldsymbol{d}=(\mathrm{d}\boldsymbol{x},\mathrm{d}\boldsymbol{y},\mathrm{d}\boldsymbol{w})$ 满足如下条件:

$$\nabla g_j(\boldsymbol{x},\boldsymbol{y})^{\mathrm{T}}\begin{pmatrix}\mathrm{d}\boldsymbol{x}\\ \mathrm{d}\boldsymbol{y}\end{pmatrix}>0,\quad \forall j\in I(\boldsymbol{x},\boldsymbol{y})=\left\{j:g_j(\boldsymbol{x},\boldsymbol{y})=0\right\}.$$

为此,我们修正QP(10)的方向 $\mathrm{d}z=(\mathrm{d}\boldsymbol{x},\mathrm{d}\boldsymbol{y},\mathrm{d}\boldsymbol{w})$. 假设问题(4)满足MFCQ约束规则,则存在方向 $\boldsymbol{d}=(\overline{\mathrm{d}\boldsymbol{x}},\overline{\mathrm{d}\boldsymbol{y}},\overline{\mathrm{d}\boldsymbol{w}})$,使得

$$\begin{cases}\nabla g_j(\boldsymbol{x},\boldsymbol{y})^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}\\ \overline{\mathrm{d}\boldsymbol{y}}\end{pmatrix}>0,\ j\in I(\boldsymbol{x},\boldsymbol{y})=\left\{j:g_j(\boldsymbol{x},\boldsymbol{y})=0\right\},\\ \begin{pmatrix}\nabla F(\boldsymbol{x},\boldsymbol{y})\\ -\boldsymbol{I}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}\\ \overline{\mathrm{d}\boldsymbol{y}}\\ \overline{\mathrm{d}\boldsymbol{w}}\end{pmatrix}=\boldsymbol{0},\quad \nabla\Phi_\varepsilon(\boldsymbol{y},\boldsymbol{w})^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{y}}\\ \overline{\mathrm{d}\boldsymbol{w}}\end{pmatrix}=\boldsymbol{0}.\end{cases}\tag{15}$$

类似于文献[7],我们考虑方向 $\boldsymbol{d}=(\mathrm{d}\boldsymbol{x},\mathrm{d}\boldsymbol{y},\mathrm{d}\boldsymbol{w})$ 与 $\boldsymbol{d}=(\overline{\mathrm{d}\boldsymbol{x}},\overline{\mathrm{d}\boldsymbol{y}},\overline{\mathrm{d}\boldsymbol{w}})$ 的凸组合:

$$\boldsymbol{q}=(\boldsymbol{qx},\boldsymbol{qy},\boldsymbol{qw})=(1-\tau)\boldsymbol{d}+\tau\bar{\boldsymbol{d}},\tag{16}$$

其中,$\tau\in(0,1]$ 且 τ满足

$$(1-\tau)\theta'(z,\rho,\varepsilon;\boldsymbol{d})+\tau\nabla f(\boldsymbol{x},\boldsymbol{y})^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}\\ \overline{\mathrm{d}\boldsymbol{y}}\end{pmatrix}\leqslant\sigma\theta'(z,\rho,\varepsilon;\boldsymbol{d})\tag{17}$$

和 $\sigma\in(0,1)$. 如果矩阵 $\boldsymbol{W}$正定,则 $\theta'(z,\rho,\varepsilon;\boldsymbol{d})\leqslant-\boldsymbol{d}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{d}/2<0$,故而存在常数 $\tau\in(0,1]$,使得式(17)成立.

命题4 假设矩阵 $\boldsymbol{W}$ 正定,且方向 $\boldsymbol{q}$ 由式(16)所确定,则有

$$\begin{cases}\theta'(z,\rho,\varepsilon;\boldsymbol{q})\leqslant\sigma\theta'(z,\rho,\varepsilon;\boldsymbol{d})<0,\\ \nabla g_j(\boldsymbol{x},\boldsymbol{y})^{\mathrm{T}}\begin{pmatrix}\boldsymbol{qx}\\ \boldsymbol{qy}\end{pmatrix}>0,\qquad j\in I(\boldsymbol{x},\boldsymbol{y}).\end{cases}\tag{18}$$

下面描述求解MPEC问题(1)的光滑SQP算法.

光滑SQP算法(1)

步骤0 初始化

令 $\rho_{-1},\varepsilon_0,\delta\in(0,\infty)$, $\alpha\in(0,1/2)$, $\upsilon,\beta,\sigma\in(0,1)$, $\theta\in(2,3)$. 选取向量 $z^0=(\boldsymbol{x}^0,\boldsymbol{y}^0,\boldsymbol{w}^0)\in R^{n+2m}$ 使得 $g(\boldsymbol{x}^0,\boldsymbol{y}^0)\geqslant 0$ 及对称正定阵 $\boldsymbol{W}_0\in R^{(n+2m)\times(n+2m)}$. 置 $k=0$.

步骤1 计算QP子问题(10)的方向 $\boldsymbol{d}^k$

在迭代点 $z^k=(\boldsymbol{x}^k,\boldsymbol{y}^k,\boldsymbol{w}^k)$ 及 $\varepsilon=\varepsilon_k$,求解QP子问题(10)得到其KKT点 $\boldsymbol{d}^k=(\mathrm{d}\boldsymbol{x}^k,\mathrm{d}\boldsymbol{y}^k,\mathrm{d}\boldsymbol{w}^k)$ 及相应的KKT乘子 $\lambda^k=(\lambda_g^k,\lambda_F^k,\lambda_\Phi^k)$. 若 $\boldsymbol{d}^k=\boldsymbol{0}$,转入步骤7.

步骤2 罚参数更新

$$令\quad \rho_k=\begin{cases}\rho_{k-1}, & \rho_{k-1}\geqslant\|\lambda^k\|_\infty+\delta,\\ \max\left\{\|\lambda^k\|_\infty+\delta,\ \rho_{k-1}+2\delta\right\}, & 否则,\end{cases}\tag{19}$$

其中 $\|\lambda^k\|_\infty=\max\limits_{1\leqslant j\leqslant m}\left\{|(\lambda_g^k)_i|,|(\lambda_F^k)_j|,|(\lambda_\Phi^k)_t|:1\leqslant i\leqslant l,1\leqslant j,t\leqslant m\right\}$.

步骤3 计算高阶校正方向 $\boldsymbol{d}^k$

$$令\quad I_g^k=\left\{j:g_j(\boldsymbol{x}^k,\boldsymbol{y}^k)+\nabla g_j(\boldsymbol{x}^k,\boldsymbol{y}^k)^{\mathrm{T}}\begin{pmatrix}\mathrm{d}\boldsymbol{x}^k\\ \mathrm{d}\boldsymbol{y}^k\end{pmatrix}=0\right\}.$$

求解如下线性系统:

$$V_{I_g^k}\begin{pmatrix}\mathrm{d}\boldsymbol{x}\\ \mathrm{d}\boldsymbol{y}\\ \mathrm{d}\boldsymbol{w}\end{pmatrix} = -\begin{pmatrix} g_{I_g^k}(\boldsymbol{x}^k + \mathrm{d}\boldsymbol{x}^k, \boldsymbol{y}^k + \mathrm{d}\boldsymbol{y}^k) + \|\boldsymbol{d}^k\|^{\theta}\boldsymbol{e} \\ F(\boldsymbol{x}^k + \mathrm{d}\boldsymbol{x}^k, \boldsymbol{y}^k + \mathrm{d}\boldsymbol{y}^k) - (\boldsymbol{w}^k + \mathrm{d}\boldsymbol{w}^k) + \|\boldsymbol{d}^k\|^{\theta}\boldsymbol{e} \\ \Phi_{\varepsilon_k}(\boldsymbol{y}^k + \mathrm{d}\boldsymbol{y}^k, \boldsymbol{w}^k + \mathrm{d}\boldsymbol{w}^k) + \|\boldsymbol{d}^k\|^{\theta}\boldsymbol{e}\end{pmatrix}, \tag{20}$$

其中, $\boldsymbol{e} = (1, \cdots, 1)^{\mathrm{T}} \in R^m$, $\boldsymbol{e} = (1, \cdots, 1)^{\mathrm{T}} \in R^{|I_g^k|}$, $g_{I_g^k}(\cdot, \cdot) = (g_j(\cdot, \cdot), j \in I_g^k)$ 且

$$V_{I_g^k} = \begin{pmatrix} (\nabla_x g)_{I_g^k}(\boldsymbol{x}^k, \boldsymbol{y}^k)^{\mathrm{T}} & (\nabla_y g)_{I_g^k}(\boldsymbol{x}^k, \boldsymbol{y}^k)^{\mathrm{T}} & \boldsymbol{0} \\ \nabla_x F(\boldsymbol{x}^k, \boldsymbol{y}^k)^{\mathrm{T}} & \nabla_y F(\boldsymbol{x}^k, \boldsymbol{y}^k)^{T} & -\boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{A}_k & \boldsymbol{B}_k \end{pmatrix},$$

而 $(\boldsymbol{A}_k, \boldsymbol{B}_k) = \nabla\Phi_{\varepsilon_k}(\boldsymbol{y}^k, \boldsymbol{w}^k)$, $\boldsymbol{A}_k$ 和 $\boldsymbol{B}_k$ 是适当的对角阵. 记其解为 $\boldsymbol{d}^k = (\mathrm{d}\boldsymbol{x}^k, \mathrm{d}\boldsymbol{y}^k, \mathrm{d}\boldsymbol{w}^k)$. 若式(20) 无解或者 $\|\boldsymbol{d}^k\| > \|\boldsymbol{d}^k\|$, 则令 $\boldsymbol{d}^k = \boldsymbol{0}$.

步骤 4　试探步搜索

若下列不等式组成立:

$$\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{d}^k) \leqslant \min\left\{-\xi\|\boldsymbol{d}^k\|^{\delta}, -\xi\|\boldsymbol{d}^k + \boldsymbol{d}^k\|^{\delta}\right\}, \tag{21}$$

$$\theta(\boldsymbol{z}^k + \boldsymbol{d}^k + \boldsymbol{d}^k, \rho_k, \varepsilon_k) \leqslant \theta(\boldsymbol{z}^k, \rho_k, \varepsilon_k) + \alpha\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{d}^k), \tag{22}$$

$$g_j(\boldsymbol{x}^k + \mathrm{d}\boldsymbol{x}^k + \mathrm{d}\boldsymbol{x}^k, \boldsymbol{y}^k + \mathrm{d}\boldsymbol{y}^k + \mathrm{d}\boldsymbol{y}^k) \geqslant 0, \quad j \in I. \tag{23}$$

令 $t_k = 1$, $\boldsymbol{q}^k = \boldsymbol{d}^k + \boldsymbol{d}^k$, 转步骤7.

步骤 5　计算辅助方向 $\boldsymbol{q}^k$

步骤 5.1　由式(15), MFCQ 在 $\boldsymbol{z}^k$ 处成立, 求解 $\overline{\boldsymbol{d}}^k = (\overline{\mathrm{d}\boldsymbol{x}}^k, \overline{\mathrm{d}\boldsymbol{y}}^k, \overline{\mathrm{d}\boldsymbol{w}}^k)$ 使

$$\nabla g_j(\boldsymbol{x}^k, \boldsymbol{y}^k)^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}^k\\ \overline{\mathrm{d}\boldsymbol{y}}^k\end{pmatrix} > \boldsymbol{0}, \quad j \in I(\boldsymbol{x}^k, \boldsymbol{y}^k),$$

$$\begin{pmatrix}\nabla F(\boldsymbol{x}^k, \boldsymbol{y}^k)\\ -\boldsymbol{I}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}^k\\ \overline{\mathrm{d}\boldsymbol{y}}^k\\ \overline{\mathrm{d}\boldsymbol{w}}^k\end{pmatrix} = \boldsymbol{0}, \quad \nabla\Phi_{\varepsilon_k}(\boldsymbol{y}^k, \boldsymbol{w}^k)^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{y}}^k\\ \overline{\mathrm{d}\boldsymbol{w}}^k\end{pmatrix} = \boldsymbol{0}.$$

步骤 5.2　计算 $\zeta_k = -\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{d}^k)$. 作方向 $\boldsymbol{d}^k$ 和 $\overline{\boldsymbol{d}}^k$ 的凸组合:

$$\begin{cases} \boldsymbol{q}^k = \zeta_k[(1-\tau_k)\boldsymbol{d}^k + \tau_k\overline{\boldsymbol{d}}^k], \\ \tau_k = \max\left\{\tau \in (0, 1] : (1-\tau)\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{d}^k) + \right. \\ \qquad \left. \tau\nabla f(\boldsymbol{x}^k, \boldsymbol{y}^k)^{\mathrm{T}}\begin{pmatrix}\overline{\mathrm{d}\boldsymbol{x}}^k\\ \overline{\mathrm{d}\boldsymbol{y}}^k\end{pmatrix} \leqslant \sigma\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{d}^k)\right\}. \end{cases} \tag{24}$$

步骤 6　线搜索

计算序列 $\{1, \beta, \beta^2, \cdots\}$ 中满足下列不等式组之最大的常数 t, 记其为 t_k:

$$\theta(\boldsymbol{z}^k + t\boldsymbol{q}^k, \rho_k, \varepsilon_k) \leqslant \theta(\boldsymbol{z}^k, \rho_k, \varepsilon_k) + \upsilon t\theta'(\boldsymbol{z}^k, \rho_k, \varepsilon_k; \boldsymbol{q}^k), \tag{25}$$

$$g_j(\boldsymbol{x}^k + t\boldsymbol{qx}^k, \boldsymbol{y}^k + t\boldsymbol{qy}^k) \geqslant 0, \quad j \in I. \tag{26}$$

步骤 7　终止及更新准则

若 $\boldsymbol{d}^k = \boldsymbol{0}$, $E(\boldsymbol{z}^k, \varepsilon_k) = \varnothing$, 停止. 若 $\boldsymbol{d}^k = \boldsymbol{0}$, $E(\boldsymbol{z}^k, \varepsilon_k) \neq \varnothing$, 令

$$\boldsymbol{z}^{k+1} = \boldsymbol{z}^k, \quad \boldsymbol{W}_{k+1} = \boldsymbol{W}_k, \quad \varepsilon_{k+1} = \varepsilon_k/2,$$

转入步骤 8. 否则, 更新 W_k 得到对称正定阵 W_{k+1} 且令

$$z^{k+1}=z^k+t_k q^k,\quad r_{k+1}=\min\left\{\sqrt{(y_j^{k+1})^2+(w_j^{k+1})^2}\mid j=1,\cdots,m\right\},$$

$$\varepsilon_{k+1}=\begin{cases}\varepsilon_k/2, & r_{k+1}<\varepsilon_k,\\ \varepsilon_k, & \text{否则}.\end{cases}$$

步骤 8 循环

令 $k=k+1$, 返回步骤 1.

注 在上述算法中, 由步骤 1 到步骤 8 有两个种循环, 即 I: 1—2—3—4—7—8 和 II: 1—2—5—6—7—8. 在循环I 中, 如果试探步成功, 则当前迭代结束, 即循环II 不参与此次迭代. 否则, 算法执行循环II 以保证算法的良定性. 命题 4 保证了算法执行步骤6 时, 一定可以搜索到适当的步长 t.

3 SQP 算法的全局收敛性

本节讨论算法的全局收敛性. 首先, 我们阐述上面算法是良定的.

下面结论表明, 当算法有限步终止于迭代点 z^k 时, 当前点即为问题(1) 的一个精确稳定点.

类似于文献[8] 之引理 3.3, 下面结论成立.

引理 1 假设 (d^k, λ^k) 是 QP 子问题(10) 的 KKT 点对. 如果 $d^k=\mathbf{0}$, 则或者 z^k 是问题(1) 的一个精确原始对偶稳定点, 或者 $d^{k+1}\neq\mathbf{0}$.

下面结论表明, 线搜索步骤6 是良定的.

引理 2 线搜索是良定的, 即步骤6 可在有限次内确定一个 j_k, 使得步长 $t_k=\beta^{j_k}$.

证明 由算法结构及命题 4 可知结论显然成立.

引理 2 表明算法是良定的且由引理 1 可知, 如果算法有限步终止于迭代点 z^k, 则当前点即为问题(1) 的一个原始对偶稳定点. 下面假设算法产生无限点列 $\{z^k\}$.

首先, 对序列 $\{z^k\}$ 作如下假设:

H1 序列 $\{z^k\}$ 有界, 且对于任意可行点 z, 向量组 $\{\nabla g_j(x,y), j\in I(x,y)\}$ 线性无关.

H2 存在常数 $a, b>0$, 使得

$$a\|z\|^2\leqslant z^{\mathrm T}W_k z\leqslant b\|z\|^2,\qquad \forall z\in R^{n+2m}\ \text{及}\ k=1,2,\cdots.$$

H3 序列 $\{z^k\}$ 的极限点 $z^*=(x^*, y^*, w^*)$ 满足下层非退化条件.

H4 在序列 $\{z^k\}$ 的极限点 $z^*=(x^*, y^*, w^*)$ 处, 子矩阵 $(\nabla_y F(x,y))_{\alpha^*\alpha^*}$ 是非奇异的, 其中 $\alpha^*=\{j: w_j^*=F_j(x^*, y^*)=0\}$ (参考文献[3]).

下面关于序列 $\{z^k, k\in K\}$ 的一些性质, 具体可参考文献[3].

引理 3 假设 $\{z^k, k\in K\}\to z^*$, $k\to\infty$, 则下面结论成立:

(a) 存在常数 $p>0$, 使得

$$\left\|\begin{pmatrix}\nabla_y F(x^k, y^k)^{\mathrm T} & -I\\ A_k & B_k\end{pmatrix}\right\|\leqslant p,\qquad \forall k\in K;$$

(b) 序列 $\{d^k: k\in K\}$, $\{\lambda^k: k\in K\}$, $\{d^k: k\in K\}$ 和 $\{q^k: k\in K\}$ 都有界;

(c) 存在正整数 k_0, 使得对 $\forall k\geqslant k_0$, 有 $\rho_k\equiv\rho_{k_0}=\rho$;

(d) 存在正整数 k_1, 使得对 $\forall k\geqslant k_1$, 有 $\varepsilon_k\equiv\varepsilon_{k_1}=\varepsilon$.

在下文中, 我们总假设 $\rho_k\equiv\rho$, $\varepsilon_k\equiv\varepsilon$. 当 k 充分大时, 由引理3 可知 $\Phi(y^k, w^k)\equiv\Phi_\varepsilon(y^k, w^k)$ 且或 $d^k=\mathbf{0}$(z^k 是问题(1) 的原始对偶稳定点), 或 $d^k\neq\mathbf{0}$.

根据假设H1、H2及引理3, 我们不妨假设存在无限子集 K, 使得

$$x^k \to x^*,\quad W_k \to W_*,\quad d^k \to d^*,\quad d^k \to d^*,\quad q^k \to q^*,\quad \lambda^k \to \lambda^*,\qquad k \in K. \tag{27}$$

下面给出算法的收敛性定理.

定理1　若上述所有假设条件成立, 则算法或有限步终止于问题(1)的原始-对偶稳定点 z^k, 或产生一无限点列 $\{z^k\}$ 且其任意聚点 z^* 都是问题(1)的原始-对偶稳定点.

证明　根据上述分析, 第一部分显然成立. 假设算法产生无限点列 $\{z^k\}$ 且式(27)成立, 由式(21)、(22)、(25)及命题4, 显然有

$$\theta(z^k, \rho, \varepsilon) \to \theta(z^*, \rho, \varepsilon),\qquad k \to \infty. \tag{28}$$

通过以上分析, 只需证明 $d^k \to 0$, $k \in K$.

如果存在无限子集 K_1, 使得点 $z^{k+1} = z^k + t_k q^k$, $k \in K_1$ 由步骤4和步骤7产生, 则

$$0 = \lim_{k\in K_1}(\theta(z^{k+1}, \rho, \varepsilon) - \theta(z^k, \rho, \varepsilon)) \leqslant \lim_{k\in K_1} \alpha\theta'(z^k, \rho, \varepsilon; d^k) \leqslant -\alpha(d^k)^{\mathrm T} W_k d^k \leqslant 0.$$

因而, 有 $d^k \to 0$, $k \in K$, $k \to \infty$.

不失一般性, 我们不妨假设点 $z^{k+1} = z^k + t_k q^k$, $\forall k \in K$ 由步骤6和步骤7产生. 由命题4可知 $d^k \to 0$, $k \in K$. 结论得证.

4 收敛速率

下面继续讨论算法的超线性收敛速率. 鉴于问题(1)与问题(4)的等价性, 我们只需讨论问题(4)即可.

首先, 当 k 充分大时, 由 $E(z^k, \varepsilon) = \varnothing$ 可知函数 $\Phi_\varepsilon(y^k, w^k)$ 二次连续可微. 下面继续讨论算法的超线性收敛性, 为此, 另作如下假设:

H5　$W_k \to W_*$, $k \to \infty$.

H6　上层严格互补条件成立, 即 $(\lambda_g^*)_j > 0, j \in I(x^*, y^*)$.

H7　问题(1)在其原始-对偶稳定点 z^* 及相应乘子 λ^* 满足二阶充分条件, 等价地, 问题(4)在点 z^* 及相应乘子向量(ξ^*, η^*, π^*) 满足二阶充分条件.

在上述条件下, 由文献[9]之命题4.1可得如下结论成立.

引理4　$\{z^k\}$ 整列收敛到 z^*.

引理5　对于充分大的 k, 有 $I_g^k \equiv I(x^*, y^*)$ 且

$$\lim_{k\to\infty} d^k = 0,\quad \lim_{k\to\infty}\lambda_g^k = \lambda_g^*,\quad \lim_{k\to\infty}\lambda_F^k = \lambda_F^*,\quad \lim_{k\to\infty}\lambda_\Phi^k = \lambda_\Phi^*.$$

证明　注意到定理1的证明过程及 $z^k \to z^*$, $k \to \infty$, 易知 $d^k \to 0$, $k \to \infty$. 故对充分大的 k, 有 $I_g^k \subseteq I(x^*, y^*)$. 另一方面, 如果 $I(x^*, y^*) \subseteq I_g^k$ 不成立, 则存在一固定指标 j_0 及子列 K, 使得 $j_0 \in I(x^*, y^*) \setminus I_g^k$, $k \in K$. 因 I_g^k 在 K 上是有界的, 不失一般性, 对于 $k \in K$ 充分大, 我们不妨假设 $I_g^k \equiv J \subseteq I(x^*, y^*)$. 记

$$(\lambda_g^k)_J = (\lambda_j^k, j \in J),$$
$$(\nabla_x g)_J^k = (\nabla_x g_j(x^k, y^k), j \in J),$$
$$(\nabla_y g)_J^k = (\nabla_y g_j(x^k, y^k), j \in J).$$

因 $(\lambda_g^k)_j = 0, j \in I \setminus J$, 由式(10)可知

$$\begin{pmatrix} \nabla f(\boldsymbol{x}^k, \boldsymbol{y}^k) \\ \boldsymbol{0} \end{pmatrix} + \boldsymbol{W}_k \boldsymbol{d}^k + \begin{pmatrix} (\nabla_x g)_J^k & \nabla_x F(\boldsymbol{x}^k, \boldsymbol{y}^k) & \boldsymbol{0} \\ (\nabla_y g)_J^k & \nabla_y F(\boldsymbol{x}^k, \boldsymbol{y}^k) & \boldsymbol{A}_k \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{B}_k \end{pmatrix} \begin{pmatrix} (\lambda_g^k)_J \\ \lambda_F^k \\ \lambda_\Phi^k \end{pmatrix} = \boldsymbol{0}, \quad k \in K. \tag{29}$$

记

$$\boldsymbol{R}_k = \begin{pmatrix} (\nabla_x g)_J^k & \nabla_x F(\boldsymbol{x}^k, \boldsymbol{y}^k) & \boldsymbol{0} \\ (\nabla_y g)_J^k & \nabla_y F(\boldsymbol{x}^k, \boldsymbol{y}^k) & \boldsymbol{A}_k \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{B}_k \end{pmatrix}, \quad \boldsymbol{R} = \begin{pmatrix} (\nabla_x g)_J^* & \nabla_x F(\boldsymbol{x}^*, \boldsymbol{y}^*) & \boldsymbol{0} \\ (\nabla_y g)_J^* & \nabla_y F(\boldsymbol{x}^*, \boldsymbol{y}^*) & \boldsymbol{A}_* \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{B}_* \end{pmatrix}.$$

注意到 $J \subseteq I(\boldsymbol{x}^*, \boldsymbol{y}^*)$ 及 $\nabla_y F$ 是一个 $P-$ 矩阵, 由假设 H1 可知 $(\boldsymbol{R}_k^{\mathrm{T}} \boldsymbol{R}_k)^{-1}$ 和 $(\boldsymbol{R}^{\mathrm{T}} \boldsymbol{R})^{-1}$ 存在, 且

$$(\boldsymbol{R}_k^{\mathrm{T}} \boldsymbol{R}_k)^{-1} \to (\boldsymbol{R}^{\mathrm{T}} \boldsymbol{R})^{-1}, \qquad k \in K.$$

故由式(29) 可得

$$\begin{pmatrix} (\lambda_g^k)_J \\ \lambda_F^k \\ \lambda_\Phi^k \end{pmatrix} \to -(\boldsymbol{R}^{\mathrm{T}} \boldsymbol{R})^{-1} \boldsymbol{R}^{\mathrm{T}} \begin{pmatrix} \nabla f(\boldsymbol{x}^*, \boldsymbol{y}^*) \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} (\bar{\lambda}_g^*)_J \\ \bar{\lambda}_F^* \\ \bar{\lambda}_\Phi^* \end{pmatrix}, \qquad k \in K.$$

而 $(\lambda_g^k)_j = 0 = (\bar{\lambda}_g^*)_j, j \in I \setminus J, k \in K$. 故由 $\boldsymbol{z}^*$ 是问题(4) 的 KKT 点及乘子的唯一性可知

$$\left(\lambda_g^k, \lambda_F^k, \lambda_\Phi^k\right) \to \left(\bar{\lambda}_g^*, \bar{\lambda}_F^*, \bar{\lambda}_\Phi^*\right) = \left(\lambda_g^*, \lambda_F^*, \lambda_\Phi^*\right), \qquad k \in K, k \to \infty. \tag{30}$$

因此, $g_{j_0}(\boldsymbol{x}^*, \boldsymbol{y}^*) = 0, (\lambda_g^*)_{j_0} = 0$. 这与假设 H6 矛盾, 故而 $J_k \equiv I(\boldsymbol{x}^*, \boldsymbol{y}^*)$.

进而, 由 $I_g^k \equiv I(\boldsymbol{x}^*, \boldsymbol{y}^*)$, 类似于式(30) 的证明, 可知引理之余下结论都成立.

引理 6 假设问题(4) 满足 LICQ, 则对充分大的 k, 步骤 3 所产生的方向 $\tilde{\boldsymbol{d}}^k$ 满足

$$\| \tilde{\boldsymbol{d}}^k \| = O\left(\| \boldsymbol{d}^k \|^2 \right).$$

证明 根据步骤 3 中方向 $\tilde{\boldsymbol{d}}^k$ 的定义, 由引理 5 可知结论成立.

记 $L\left(\boldsymbol{z}, \lambda_g, \lambda_F, \lambda_\Phi\right)$ 为问题(4) 的 Lagrange 函数, 则有

$$L(\boldsymbol{z}^k, \lambda_g^k, \lambda_F^k, \lambda_\Phi^k) = f(\boldsymbol{x}^k, \boldsymbol{y}^k) + (\lambda_g^k)^{\mathrm{T}} g(\boldsymbol{x}^k, \boldsymbol{y}^k) + (\lambda_F^k)^{\mathrm{T}} (F(\boldsymbol{x}^k, \boldsymbol{y}^k) - \boldsymbol{w}^k) + (\lambda_\Phi^k)^{\mathrm{T}} \Phi_\varepsilon(\boldsymbol{y}^k, \boldsymbol{w}^k).$$

为得到超线性收敛速率, 做进一步假设:

H8 矩阵序列 $\left\{ \boldsymbol{W}_k \right\}$ 满足

$$\| \boldsymbol{P}_k (\boldsymbol{W}_k - \nabla_{zz}^2 L\left(\boldsymbol{z}^*, \lambda_g^*, \lambda_F^*, \lambda_\Phi^* \right)) \boldsymbol{d}^k \| = o(\| \boldsymbol{d}^k \|),$$

其中 $\boldsymbol{P}_k = \boldsymbol{I}_{n+2m} - \boldsymbol{R}_k (\boldsymbol{R}_k^{\mathrm{T}} \boldsymbol{R}_k)^{-1} \boldsymbol{R}_k^{\mathrm{T}}$, $\boldsymbol{R}_k$ 的表达式见引理 5.

类似于文献[8] 之引理 4.4 之证明过程, 我们有如下结论:

引理 7 当 k 充分大时, 试探步搜索一定成立, 即 $\boldsymbol{z}^{k+1} = \boldsymbol{z}^k + \boldsymbol{d}^k + \tilde{\boldsymbol{d}}^k$.

此外, 注意到引理 7 及文献[10] 之定理 5.2, 我们有如下收敛定理:

定理 2 若上述所有条件成立, 则算法超线性收敛, 即算法所产生的点列 $\left\{ \boldsymbol{z}^k \right\}$ 满足

$$\| \boldsymbol{z}^{k+1} - \boldsymbol{z}^* \| = o\left(\| \boldsymbol{z}^k - \boldsymbol{z}^* \| \right).$$

[参 考 文 献]

[1] Outrate J V, Kocvare M, Zowe J. Nonsmooth Approach to Optimization Problems With Equilibrium Consrtaints[M]. The Netherlands: Kluwer Academic Publishers, 1998.

[2] Jiang H, Ralph D. Smooth SQP method for mathematical programs with nonlinear complementarity constraints[J]. SIAM J Optimization, 2000, **10**(3): 779- 808.

[3] Fukushima M, Luo Z Q, Pang J S. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints[J]. Comp Opti Appl, 1998, **10**(1): 5- 34.

[4] Ma C F, Liang G P. A new successive approximation damped Newton method for nonlinear complementarity problems[J]. Journal of Mathematical Research and Exposition, 2003, **23**(1): 1- 6.

[5] 朱志斌, 罗志军, 曾吉文. 互补约束均衡问题一个新的磨光技术[J]. 应用数学和力学, 2007, **28**(10): 1253- 1260.

[6] Fukushima M, Pang J S. Some feasibility issues in mathematical programs with equilibrium constraints[J]. SIAM J Optimization, 1998, **8**(3): 673- 681.

[7] Panier E R, Tits A L. On combining feasibility, descent and superlinear convergence in inequality constrained optimization[J]. Mathematical Programming, 1993, **59**(1): 261- 276.

[8] Zhu Z B, Zhang K C. A superlinearly convergent SQP algorithm for mathematical programs with linear complementarity constraints[J]. Applied Mathematics and Computation, 2006, **172**(1): 222- 244.

[9] Panier E R, Tits A L. A superlinearly convergent feasible method for the solution of inequality constrained optimization problems[J]. SIAM J Control Optim, 1987, **25**(3): 934- 950.

[10] Facchinei F, Lucidi S. Quadraticly and superlinearly convergent for the solution of inequality constrained optimization problem[J]. J Optim Theory Appl, 1995, **85**(2): 265- 289.

An SQP Algorithm for Mathematical Programs With Nonlinear Complementarity Constraints

ZHU Zhi- bin[1], JIAN Jin- bao[2], ZHANG Cong[1]

(1. College of Computational Science and Mathematics, Guilin University of Electronic Technology, Guilin, Guangxi 541004, P. R. China;

2. College of Mathematics and Information Science, Guangxi University, Nanning 530004, P. R. China)

Abstract: A successive approximation and smooth SQP method for mathematical programs with nonlinear complementarity constraints (MPCC) is described. A class of smooth programs to approximate the MPCC was introduced. Using an l_1 penalty function, the line search assures the global convergence, while superlinear convergence rate is shown under strictly complementary conditions and the second order sufficient condition. Moreover, it was proved that the current iterated point is an exact stationary point of the MPEC when the algorithm terminates finitely.

Key words: MPEC; SQP algorithm; successive approximation; global convergence; superlinear convergence rate