文章编号: 1000-0887(2008) 04-0483-06

求解一类 MPEC 问题的一种 $\mathcal{UV}$- 分解方法*

单 锋[1], 庞丽萍[2], 朱丽梅[1], 夏尊铨[2]

(1. 沈阳航空工业学院 理学院, 沈阳 110136;
2. 大连理工大学 最优化研究与应用中心, 应用数学系, 辽宁 大连 116024)

(郭兴明推荐)

摘要: 给出了求解具有线性互补约束的 MPEC 问题的一种 $\mathcal{UV}$- 分解方法. 首先将 MPEC 问题化为非线性规划问题, 给出一种相应的罚函数的次微分结构及其 $\mathcal{UV}$- 分解的结果, 根据所得到的结果构造一个具有超线性收敛速度的概念型算法.

关 键 词: 非光滑优化; 非线性规划; 次微分; $\mathcal{UV}$- 分解; $\mathcal{U}$- Lagrange 函数; MPEC 问题

中图分类号: O224; O221. 2 文献标识码: A

引 言

考虑问题

$$(\mathrm{MPEC})\quad \begin{cases} \min f(\boldsymbol{x}) \\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x}) \leqslant \boldsymbol{0}, \\ \quad \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \\ \quad \boldsymbol{G}(\boldsymbol{x}) \geqslant \boldsymbol{0},\ \boldsymbol{H}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \\ \quad \boldsymbol{G}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{H}(\boldsymbol{x}) = 0, \end{cases} \tag{1}$$

其中 $f: R^n \to R^1$, $\boldsymbol{G}: R^n \to R^m$, $\boldsymbol{H}: R^n \to R^m$, $\boldsymbol{g}: R^n \to R^p$, $\boldsymbol{h}: R^n \to R^q$. 更一般地, (MPEC) 问题可以写成

$$\begin{cases} \min f(\boldsymbol{x}, \boldsymbol{y}) \\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}, \\ \quad \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) \geqslant \boldsymbol{0}, \boldsymbol{y} \geqslant \boldsymbol{0},\ \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y})^{\mathrm{T}}\boldsymbol{y} = 0, \end{cases} \tag{2}$$

它可以视为具有下述变分不等式约束的一类优化问题, 或视为以 $\boldsymbol{x}$ 为参数, 关于 $\boldsymbol{y}$ 的均衡约束问题.

$$\begin{cases} \min f(\boldsymbol{x}, \boldsymbol{y}), \\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \leqslant \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}, \\ \quad \boldsymbol{y} \in \left\{ \boldsymbol{y} \mid \langle \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{y}' - \boldsymbol{y} \rangle \geqslant 0,\ \forall \boldsymbol{y}' \in \Omega, \boldsymbol{y} \in \Omega \right\}, \end{cases} \tag{3}$$

其中 $\Omega = R^m_+$, 见文献[1-2]. 换言之, 问题(1)是问题(2)、(3)的一类具体问题(亦称之为广义

---

* 收稿日期: 2007-09-18; 修订日期: 2008-02-28
基金项目: 国家自然科学基金资助项目(10372063; 10771026; 10471015)
作者简介: 单锋(1960—), 男, 辽宁大连人, 教授, 硕士(联系人. Tel: + 86-24-89724422; E-mail: shanfengsh@163.com).

双层规划[1]或具有均衡约束的数学规划问题[2],还可参阅文献[3].

设 $\boldsymbol{x}^*$ 是(MPEC)的一个可行点,定义指标集:

$$I_g = \left\{ i \mid g_i(\boldsymbol{x}^*) = 0 \right\},$$
$$\alpha = \alpha(\boldsymbol{x}^*) = \left\{ i \mid G_i(\boldsymbol{x}^*) = 0,\ H_i(\boldsymbol{x}^*) > 0 \right\},$$
$$\beta = \beta(\boldsymbol{x}^*) = \left\{ i \mid G_i(\boldsymbol{x}^*) = 0,\ H_i(\boldsymbol{x}^*) = 0 \right\},$$
$$\gamma = \gamma(\boldsymbol{x}^*) = \left\{ i \mid G_i(\boldsymbol{x}^*) > 0,\ H_i(\boldsymbol{x}^*) = 0 \right\}.$$

称 $\boldsymbol{x}^*$ 是(MPEC)的弱稳定点(W-稳定点),若存在 $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ 使条件

$$0 = \nabla f(\boldsymbol{x}^*) + \sum_{i \in I_g} \lambda_i^g \nabla g_i(\boldsymbol{x}^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(\boldsymbol{x}^*) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(\boldsymbol{x}^*) + \lambda_i^H \nabla H_i(\boldsymbol{x}^*)],$$

$$\lambda_{I_g}^g \geqslant \boldsymbol{0},\quad \lambda_\gamma^G \geqslant \boldsymbol{0},\quad \lambda_\alpha^H \geqslant \boldsymbol{0}$$

成立,并且W-稳定条件是严格(MPEC)问题的KKT条件,

$$(\text{TMPEC}) \quad \begin{cases} \min f(\boldsymbol{x}) \\ \text{s.t. } \boldsymbol{g}(\boldsymbol{x}) \leqslant \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \\ \qquad G_i(\boldsymbol{x}) \geqslant 0, \qquad i \in \alpha, \\ \qquad H_i(\boldsymbol{x}) = 0, \qquad i \in \gamma, \\ \qquad G_i(\boldsymbol{x}) = 0,\ H_i(\boldsymbol{x}) = 0, \qquad i \in \beta, \end{cases} \tag{4}$$

见文献[1],换言之,(TMPEC)的KKT点是(MPEC)的W-稳定点. 根据Lemarchal, Oustry和Sagastizûbal[4]的 $\mathcal{UV}$-分解理论,如果凸函数的次微分内部为空集但相对内部非空,则可以将空间进行 $\mathcal{UV}$-分解. 例如,若 g 是凸函数 F 在 $\boldsymbol{x}$ 处的次微分的一个相对内点,则 $\partial F(\boldsymbol{x}, \mu) - g$ 可生成一子空间,记为 $\mathcal{V}$,其正交补子空间为 $\mathcal{U} = \mathcal{V}^\perp$. 第1节将给出一种相应的罚函数的次微分结构及其 $\mathcal{UV}$-分解的结果,第2节中将给出一个 $\mathcal{UV}$-分解算法的格式及算法的超线性收敛的性质.

1 一些有关 $\mathcal{UV}$-分解理论的结果

为方便起见,定义

$$\begin{cases} \alpha\beta(\boldsymbol{x}) = \alpha(\boldsymbol{x}) \cup \beta(\boldsymbol{x}),\quad \beta\gamma(\boldsymbol{x}) = \beta(\boldsymbol{x}) \cup \gamma(\boldsymbol{x}), \\ \alpha\beta\gamma(\boldsymbol{x}) = \alpha(\boldsymbol{x}) \cup \beta(\boldsymbol{x}) \cup \gamma(\boldsymbol{x}), \\ \lambda_g^{\mathrm T} \cdot \boldsymbol{g}(\boldsymbol{x}) = \sum\limits_{i \in I_g(\boldsymbol{x})} \lambda_i \cdot g_i(\boldsymbol{x}),\quad \lambda_h^{\mathrm T} \cdot \boldsymbol{h}(\boldsymbol{x}) = \sum\limits_{j=1}^{q} \lambda_j \cdot h_j(\boldsymbol{x}), \\ \lambda_{\alpha\beta}^{\mathrm T} \cdot \boldsymbol{G}(\boldsymbol{x}) = \sum\limits_{k \in \alpha\beta(\boldsymbol{x})} \lambda_k \cdot G_k(\boldsymbol{x}),\quad \lambda_{\gamma\beta}^{\mathrm T} \cdot \boldsymbol{H}(\boldsymbol{x}) = \sum\limits_{l \in \beta\gamma(\boldsymbol{x})} \lambda_l \cdot H_l(\boldsymbol{x}) \end{cases} \tag{5}$$

与

$$\begin{cases} P = \left\{1, \cdots, p\right\},\ Q = \left\{1, \cdots, q\right\},\ \lambda_j^- = \lambda_j^{'-''} = \lambda_j' - \lambda_j'' \in R^1, \\ \lambda_k^- = \lambda_k^{'-''} = \lambda_k' - \lambda_k'' \in R^1,\ \lambda_l^- = \lambda_l^{'-''} = \lambda_l' - \lambda_l'' \in R^1, \\ \lambda_j^+ = \lambda_j^{'+''} = \lambda_j' + \lambda_j'' \in R^1,\ \lambda_k^+ = \lambda_k^{'+''} = \lambda_k' + \lambda_k'' \in R^1, \\ \lambda_l^+ = \lambda_l^{'+''} = \lambda_l' + \lambda_l'' \in R^1,\ \lambda_i',\ \lambda_j',\ \lambda_j'',\ \lambda_k',\ \lambda_k'' \geqslant 0. \end{cases} \tag{6}$$

下面两个基本假设将在后面用到.

假设[4]

(A1) $G_i(\boldsymbol{x})$、$H_i(\boldsymbol{x})$ $(i\in\alpha\beta\gamma(\boldsymbol{x}))$ 及 $h_i(i\in Q)$ 是线性函数,f 与 $g_i(i\in P)$ 是二次连续可微的凸函数;

(A2) $\nabla g_i(\boldsymbol{x}^*)$ $(i\in I_g)$、$\nabla h_i(\boldsymbol{x}^*)$ $(i\in Q)$、$\nabla G_i(\boldsymbol{x}^*)$ $(i\in\alpha\beta(\boldsymbol{x}^*))$ 及 $\nabla H_i(\boldsymbol{x}^*)$ $(i\in\beta\gamma(\boldsymbol{x}^*))$ 是线性无关的.

定义

$$F(\boldsymbol{x},\mu)=f(\boldsymbol{x})+\mu\|\boldsymbol{c}(\boldsymbol{x})\|_\infty=$$
$$f(\boldsymbol{x})+\mu[\lambda_g^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})+\lambda_h^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})+\lambda_{\alpha\beta(\boldsymbol{x})}^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{x})+\lambda_{\beta\gamma(\boldsymbol{x})}^{\mathrm{T}}\boldsymbol{H}(\boldsymbol{x})].$$

定理 1.1 假设(A1)和(A2)成立. 令 $F(\boldsymbol{x},\mu)$ 为(TMPEC)问题的精确罚函数, 且 $g_0(\boldsymbol{x})=0$, $\nabla g_0(\boldsymbol{x})=\boldsymbol{0}$, 其中 μ 是罚参数. 则有

1) 对任意的 $\boldsymbol{x}\in R^n$, 有

$$\partial F(\boldsymbol{x},\mu)=\nabla f(\boldsymbol{x})+\mu[\lambda_g^{\mathrm{T}}\nabla\boldsymbol{g}(\boldsymbol{x})+\lambda_h^{\mathrm{T}}\nabla\boldsymbol{h}(\boldsymbol{x})+$$
$$\lambda_{\alpha\beta}^{\mathrm{T}}\nabla\boldsymbol{G}(\boldsymbol{x})+\lambda_{\beta\gamma}^{\mathrm{T}}\nabla\boldsymbol{H}(\boldsymbol{x})],\tag{7}$$

其中 $$\sum_{i\in I_g(\boldsymbol{x})}\lambda_i+\sum_{i\in Q}\lambda_i^{'+''}+\sum_{k\in\alpha\beta(\boldsymbol{x})}\lambda_k^{'+''}+\sum_{l\in\beta\gamma(\boldsymbol{x})}\lambda_l^{'+''}=1-\lambda_0\leqslant 1;$$

2) 记 $\mathcal{V}$ 是由 $F(\cdot,\mu)$ 在 $\boldsymbol{x}^*\in R^n$ 的次微分生成的子空间. 则有

$$\mathcal{V}=\mathrm{lin}\left\{\begin{matrix}\nabla g_i(\boldsymbol{x}^*),\ \nabla h_j(\boldsymbol{x}^*),\\ \nabla G_k(\boldsymbol{x}^*),\ \nabla H_l(\boldsymbol{x}^*)\end{matrix}\ \middle|\ \begin{matrix}i\in\{0\}\cup P,\ j\in Q,\\ k\in\alpha\beta(\boldsymbol{x}^*),\ l\in\beta\gamma(\boldsymbol{x}^*)\end{matrix}\right\}=$$
$$\mathrm{lin}(\partial F(\boldsymbol{x}^*,\mu)-\nabla f(\boldsymbol{x}^*)),\tag{8}$$
$$\mathcal{U}=\mathcal{V}^{\perp}=\left\{d\ \middle|\ \begin{matrix}d^{\perp}(\nabla g_i(\boldsymbol{x}^*),\ \nabla h_j(\boldsymbol{x}^*),\ \nabla G_k(\boldsymbol{x}^*),\ H_l(\boldsymbol{x}^*))=\boldsymbol{0},\\ i\in\{0\}\cup P,\ j\in Q,\ k\in\alpha\beta(\boldsymbol{x}^*),\ l\in\beta\gamma(\boldsymbol{x}^*)\end{matrix}\right\},$$

其中 $\mathrm{lin}M$ 表示由集合 M 生成的子空间.

证明 令 $F(\boldsymbol{x},\mu)=f(\boldsymbol{x})+\mu\|\boldsymbol{c}(\boldsymbol{x})\|_\infty$, 其中无穷范数的定义为

$$\|\boldsymbol{c}(\boldsymbol{x})\|_\infty=\max\{|c_i(\boldsymbol{x})|\}=$$
$$\max\left\{\begin{matrix}(g_i(\boldsymbol{x}))^+,\ |h_j(\boldsymbol{x})|,\\ |G_k(\boldsymbol{x})|,\ |H_l(\boldsymbol{x})|\end{matrix}\ \middle|\ \begin{matrix}i\in\{0\}\cup P,\ j\in Q,\\ k\in\alpha\beta(\boldsymbol{x}),\ l\in\beta\gamma(\boldsymbol{x})\end{matrix}\right\}.$$

令 $(g_i(\boldsymbol{x}))^+=\max\{g_i(\boldsymbol{x}),0\}$, $g_0(\boldsymbol{x})=0$. 定义 $B(\boldsymbol{x})=\|\boldsymbol{c}(\boldsymbol{x})\|_\infty$. 于是, $F(\boldsymbol{x},\mu)$ 可写为 $F(\boldsymbol{x},\mu)=f(\boldsymbol{x})+\mu B(\boldsymbol{x})$. 由(A1)知, $F(\boldsymbol{x},\mu)$ 和 $B(\boldsymbol{x})$ 是凸的. 令

$$I(\boldsymbol{x})=\{i\mid g_i(\boldsymbol{x})=h_i(\boldsymbol{x})=G_i(\boldsymbol{x})=H_i(\boldsymbol{x})=0,\ i\in P\cup Q\cup\alpha\beta\gamma(\boldsymbol{x})\}.$$

下面, 首先计算 $F(\boldsymbol{x},\mu)$ 在 $\boldsymbol{x}$ 处的次微分. 由于 f 是可微的并且 $\boldsymbol{G}$ 和 $\boldsymbol{H}$ 是线性的, 所以, $\partial F(\boldsymbol{x},\mu)$ 可以表示为

$$\partial F(\boldsymbol{x},\mu)=\nabla f(\boldsymbol{x})+\mu\partial B(\boldsymbol{x})=\nabla f(\boldsymbol{x})+$$
$$\mu\,\mathrm{co}\{\boldsymbol{0},\ \nabla g_i(\boldsymbol{x}),\ \pm\nabla h_j(\boldsymbol{x}),\ \pm\nabla G_k(\boldsymbol{x}),\ \pm\nabla H_l(\boldsymbol{x}),\ i,j,k,l\in I(\boldsymbol{x})\}.$$

B 在 $\boldsymbol{x}$ 处的次微分可以写成

$$\partial B(\boldsymbol{x})=\sum_{i\in\{0\}\cup I_g(\boldsymbol{x})}\lambda_i\nabla g_i(\boldsymbol{x})+\sum_{j=1}^{q}\lambda_j^{'-''}\nabla h_j(\boldsymbol{x})+$$
$$\sum_{k\in\alpha\beta(\boldsymbol{x})}\lambda_k^{'-''}\nabla G_k(\boldsymbol{x})+\sum_{l\in\beta\gamma(\boldsymbol{x})}\lambda_l^{'-''}\nabla H_l(\boldsymbol{x})=$$
$$\sum_{i\in\{0\}\cup I_g(\boldsymbol{x})}\lambda_i\nabla g_i(\boldsymbol{x})+\lambda_h^{'-''\mathrm{T}}\nabla\boldsymbol{h}(\boldsymbol{x})+\lambda_{\alpha\beta}^{'-''\mathrm{T}}\nabla\boldsymbol{G}(\boldsymbol{x})+\lambda_{\beta\gamma}^{'-''\mathrm{T}}\nabla\boldsymbol{H}(\boldsymbol{x}),$$

$$g_0(\boldsymbol{x}) = 0, \nabla g_0(\boldsymbol{x}) = \boldsymbol{0},$$

即 $\partial B(\boldsymbol{x}) = \lambda_g^{\mathrm{T}} \nabla g_i(\boldsymbol{x}) + \bar{\lambda}_h^{'-''\mathrm{T}} \nabla \boldsymbol{h}(\boldsymbol{x}) + \bar{\lambda}_{\alpha\beta}^{'-''\mathrm{T}} \nabla \boldsymbol{G}(\boldsymbol{x}) + \bar{\lambda}_{\beta\gamma}^{'-''\mathrm{T}} \nabla \boldsymbol{H}(\boldsymbol{x})$. 注意 λ_g、$\bar{\lambda}_h^{'-''}$、$\bar{\lambda}_{\alpha\beta}^{'-''}$ 和 $\bar{\lambda}_{\beta\gamma}^{'-''}$ 是集合. 相应地, $F(\cdot, \mu)$ 的次微分可表示为

$$\partial F(\boldsymbol{x}, \mu) = \nabla f(\boldsymbol{x}) + \mu\Big[\sum_{i\in I_g(\boldsymbol{x})} \lambda_i \nabla g_i(\boldsymbol{x}) + \sum_{i=1}^{q} \lambda_j \nabla h_j(\boldsymbol{x}) + \sum_{k\in\alpha\beta(\boldsymbol{x})} \lambda_k \nabla G_k(\boldsymbol{x}) + \sum_{l\in\beta\gamma(\boldsymbol{x})} \lambda_l \nabla H_l(\boldsymbol{x})\Big],$$

其中 $$\sum_{i\in I_g(\boldsymbol{x})} \lambda_i + \sum_{i=1}^{q} \bar{\lambda}_i^{'+''} + \sum_{k\in\alpha\beta(\boldsymbol{x})} (\bar{\lambda}_k^{'+''}) + \sum_{l\in\beta\gamma(\boldsymbol{x})} (\bar{\lambda}_l^{'+''}) = 1 - \lambda_0 \leqslant 1.$$

令 $\boldsymbol{x} = \boldsymbol{x}^*$. 则有

$$\partial F(\boldsymbol{x}^*, \mu) = \nabla f(\boldsymbol{x}^*) + \mu[\lambda_0 \cdot 0 + \lambda_g^{\mathrm{T}} \nabla g_i(\boldsymbol{x}^*) + \bar{\lambda}_h^{'-''\mathrm{T}} \nabla h(\boldsymbol{x}^*) + \bar{\lambda}_{\alpha\beta}^{'-''\mathrm{T}} \nabla G(\boldsymbol{x}^*) + \bar{\lambda}_{\beta\gamma}^{'-''\mathrm{T}} \nabla H(\boldsymbol{x}^*)].$$

于是, 定理 1.1 中的 1) 成立. 因为 λ_0 可以取为非零常数, 所以有 $\nabla f(\boldsymbol{x}^*) \in \partial F(\boldsymbol{x}^*, \mu)$. 由此可得,

$$\begin{aligned}&\mathrm{lin}(\partial F(\boldsymbol{x}^*, \mu) - \nabla f(\boldsymbol{x}^*)) = \\ &\quad \mathrm{lin}\left\{\boldsymbol{0}, \nabla \boldsymbol{g}(\boldsymbol{x}^*), \nabla \boldsymbol{h}(\boldsymbol{x}^*), \nabla \boldsymbol{G}(\boldsymbol{x}^*), \nabla \boldsymbol{H}(\boldsymbol{x}^*)\right\} = \\ &\quad \mathrm{lin}\left\{\nabla \boldsymbol{g}(\boldsymbol{x}^*), \nabla \boldsymbol{h}(\boldsymbol{x}^*), \nabla \boldsymbol{G}(\boldsymbol{x}^*), \nabla \boldsymbol{H}(\boldsymbol{x}^*)\right\},\end{aligned}$$

上式中 $\boldsymbol{0}$、$\boldsymbol{g}$、$\boldsymbol{h}$、$\boldsymbol{G}$、$\boldsymbol{H}$ 表示分量的集合. 令 $\mathcal{U} = \mathcal{V}^{\perp}$. 则显然 2) 的第 2 个公式成立.

若有 $\mathrm{int}\,\mathcal{V} = \varnothing$ 与 $\mathrm{ri}\,\mathcal{V} \neq \varnothing$, 或 $\mathrm{int}\,\mathcal{U} = \varnothing$ 与 $\mathrm{ri}\,\mathcal{U} \neq R^n$ 成立, 则称 $\mathcal{UV}$- 分解是正常的.

推论 1.1 定理 1.1 中的 $\mathcal{UV}$- 分解是正常的.

证明 由定理 1.1 的证明, 可知 $\partial F(\boldsymbol{x}^*, \mu) \neq \varnothing$. 换言之, 至少有 1 个乘子是非零的, 如取

$$\lambda_0 = 1, \ \lambda_g = \boldsymbol{0}, \ \lambda_h = \boldsymbol{0}, \ \bar{\lambda}_{\alpha\beta}^{'-''} = \boldsymbol{0}, \ \bar{\lambda}_{\beta\gamma}^{'-''} = \boldsymbol{0},$$

即 $\nabla f(\boldsymbol{x}^*) \in \partial F(\boldsymbol{x}^*, \mu)$. 另一方面, $\mathrm{int}\,\partial F(\boldsymbol{x}^*, \mu) = \varnothing$. 因此, 存在 $\boldsymbol{v}$ 和 γ 使得

$$\boldsymbol{v} \in \mathrm{lin}(\partial F(\boldsymbol{x}^*, \mu) - \gamma), \quad \gamma \in \mathrm{ri}\,\partial F(\boldsymbol{x}^*, \mu).$$

假设 $\boldsymbol{0} \in \partial F(\boldsymbol{x}, \mu)$. 根据文献[4] 和[2], 有 F 的 $\mathcal{U}$ - Lagrange 函数为

$$L_0(\boldsymbol{u}) = \inf_{\boldsymbol{v}\in\mathcal{V}} \left\{F(\boldsymbol{x} + \boldsymbol{u} \oplus \boldsymbol{v}, \mu)\right\}.$$

其相应的最优解集为

$$\boldsymbol{W}(\boldsymbol{u}) = \arg\inf_{\boldsymbol{v}\in\mathcal{V}} \left\{F(\boldsymbol{x} + \boldsymbol{u} \oplus \boldsymbol{v}, \mu)\right\}.$$

任一 $\boldsymbol{x} \in R^n = \mathcal{U} \oplus \mathcal{V}$ 可表示为

$$\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{u} \oplus \boldsymbol{v} = \boldsymbol{x} + \boldsymbol{U}\boldsymbol{u} + \boldsymbol{V}\boldsymbol{v},$$

其中 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 分别是子空间 $\mathcal{U}$ 和 $\mathcal{V}$ 的基矩阵. 关于 F 的 $\mathcal{U}$ - Lagrange 函数的某些结论将在下节中用到, 参阅文献[4].

结论

P1 若 $\boldsymbol{0} \in \mathrm{ri}\,\partial F(\boldsymbol{x}, \mu)$, 则有 $\boldsymbol{W}(\boldsymbol{0}) = \left\{\boldsymbol{0}\right\}$. 根据文献[4] 的定理 3.3, 有

$$\partial L_g(\boldsymbol{u}) = \left\{\boldsymbol{g}_u \mid \boldsymbol{g}_u \oplus \boldsymbol{g}_v \in \partial F(\boldsymbol{x} + \boldsymbol{u} \oplus \boldsymbol{w}, \mu)\right\},$$

这个结论可以由令 $\boldsymbol{g} = \boldsymbol{0}$ 推出;

P2 若 $\boldsymbol{W}(\boldsymbol{u}) \neq \varnothing$, 则有

$$\partial L_g(\boldsymbol{u}) = \left\{\boldsymbol{g}_u \mid \boldsymbol{g}_u \oplus \boldsymbol{g}_v \in \partial F(\boldsymbol{x} + \boldsymbol{u} \oplus \boldsymbol{w}, \mu)\right\},$$

若 $\boldsymbol{w} \in \boldsymbol{W}(\boldsymbol{u})$, 则

$$\partial L_0(\boldsymbol{u}) = \left\{\boldsymbol{g}_u, \boldsymbol{g}_u \oplus \boldsymbol{0} \in \partial F(\boldsymbol{x}+\boldsymbol{u}\oplus \boldsymbol{w}, \mu)\right\};$$

P3 L_0 在 $\boldsymbol{0}$ 处可微且 $\nabla L_0(\boldsymbol{0}) = \boldsymbol{0}$. 由于 $\boldsymbol{W}(\boldsymbol{0}) = \{\boldsymbol{0}\}$, 则有

$$\partial L_0(\boldsymbol{0}) = \left\{\boldsymbol{g}_u, \boldsymbol{g}_u \oplus \boldsymbol{0} \in \partial F(\boldsymbol{x}+\boldsymbol{0}\oplus \boldsymbol{0}, \mu)\right\} = \left\{\boldsymbol{g}_u, \boldsymbol{g}_u \oplus \boldsymbol{0} \in \partial F(\boldsymbol{x}, \mu)\right\},$$

根据 $\mathcal{U}$ 和 $\mathcal{V}$- 空间的定义, 有 $\nabla L_0(\boldsymbol{0}) = \boldsymbol{0}$;

P4 若 $\boldsymbol{0} \in \mathrm{ri}\partial F(\boldsymbol{x}, \mu)$, 则有 $\boldsymbol{w}(\boldsymbol{u}) = o(\boldsymbol{u})$;

P5 若 F 在 $\boldsymbol{x}$ 处有 $\mathcal{U}$-Hessian 阵 $\boldsymbol{H}_{\mathcal{U}}F(\cdot, \mu)$, 则下述等式和包含关系成立:

$$F(\boldsymbol{x}+\boldsymbol{u}\oplus \boldsymbol{w}, \mu) = F(\boldsymbol{x}, \mu) + \frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{H}_{\mathcal{U}}F(\boldsymbol{x}, \mu)\boldsymbol{u} + o(\|\boldsymbol{u}\|),$$

$$\left\{\boldsymbol{g}_u \mid \boldsymbol{g}_u \oplus \boldsymbol{0} \in \partial F(\boldsymbol{x}+\boldsymbol{u}\oplus \boldsymbol{w}, \mu)\right\} \subset \boldsymbol{H}_{\mathcal{U}}F(\boldsymbol{x}, \mu)\boldsymbol{u} + B(\boldsymbol{0}, o(\|\boldsymbol{u}\|)).$$

2 算法和收敛性

基于第 1 节中所给出的结果, 本节将给出一种求解有线性互补约束的(MPEC) 问题的算法格式, 它是通过求解(TMPEC) 问题得到的. 令 $\boldsymbol{x}$ 是(TMPEC) 的解.

算法 2.1 解(TMPEC).

步 0 初始化. 给定充分接近 $\boldsymbol{x}$ 的一个初始点 $\boldsymbol{x}^{(0)}$, 置 $k = 0$.

步 1 确定约束的积极指标集. 确定指标集 I_g、α、β 和 γ.

步 2 在 $\boldsymbol{x}$ 处进行 $\mathcal{UV}$- 分解. 作 $\mathcal{UV}$- 空间分解 $R^n = \mathcal{U} \oplus \mathcal{V}$.

计算 $\mathcal{U}$-Hessian 阵 $\boldsymbol{H}_{\mathcal{U}}F(\boldsymbol{x}, \mu) = \nabla^2 L_0(\boldsymbol{0})$.

步 3 执行 $\mathcal{V}$- 步. 计算 $\delta_{\mathcal{V}}^{(k)} \in \arg\min\left\{F(\boldsymbol{x}^{(k)} + \boldsymbol{0} \oplus \delta_{\mathcal{V}}, \mu): \delta_{\mathcal{V}} \in \mathcal{V}\right\}$.

置 $x^{(k)} = x^{(k)} + \boldsymbol{0} \oplus \delta_{\mathcal{V}}^{(k)}$.

步 4 执行 $\mathcal{U}$- 步. 由方程

$$\boldsymbol{g}_{\mathcal{U}}^{(k)} + \boldsymbol{H}_{\mathcal{U}}F(\boldsymbol{x}, \mu)\delta_{\mathcal{U}} = \boldsymbol{0}, \tag{9}$$

计算 $\delta_{\mathcal{U}}^{(k)}$, 其中 $\boldsymbol{g}^{(k)} \in \partial F(\boldsymbol{x}^{(k)}, \mu)$ 且 $\boldsymbol{g}_{\mathcal{V}}^{(k)} = \boldsymbol{0}$, $(\boldsymbol{g}_{\mathcal{U}}^{(k)} \in \partial L_0((\boldsymbol{x}^{(k)} - \boldsymbol{x})_{\mathcal{U}}))$.

步 5 重置 $\boldsymbol{x}^{(k)}$. 令

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \delta_{\mathcal{U}}^{(k)} \oplus \boldsymbol{0} = \boldsymbol{x}^{(k)} + \delta_{\mathcal{U}}^{(k)} \oplus \delta_{\mathcal{V}}^{(k)}.$$

令 $k = k+1$, 转步 3.

下面给出算法的超线性收敛定理.

定理 2.1[收敛定理] 假设 $\boldsymbol{g} = \boldsymbol{0} \in \mathrm{ri}\partial F(\boldsymbol{x}, \mu)$, $\nabla^2 L_0(\boldsymbol{0}) \succ \boldsymbol{0}$. 则由算法产生的迭代点列满足 $\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}\| = o(\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|)$.

证明 令 $\boldsymbol{u}^{(k)} = (\boldsymbol{x}^{(k)} - \boldsymbol{x})_{\mathcal{U}}$, $\boldsymbol{v} = (\boldsymbol{x}^{(k)} - \boldsymbol{x})_{\mathcal{V}} + \delta_{\mathcal{V}}$, $\boldsymbol{v}^{(k)} = (\boldsymbol{x}^{(k)} - \boldsymbol{x})_{\mathcal{V}} + \delta_{\mathcal{V}}^{(k)}$. 则

$$\boldsymbol{x} + \boldsymbol{u}^{(k)} \oplus \boldsymbol{v} = \boldsymbol{x}^{(k)} + \boldsymbol{0} \oplus \delta_{\mathcal{V}},$$

并且 $\delta_{\mathcal{V}}^{(k)} \in \arg\min\left\{F(\boldsymbol{x}^{(k)} + \boldsymbol{0} \oplus \delta_{\mathcal{V}})\right\} = \arg\min F(\boldsymbol{x} + \boldsymbol{u}^{(k)} \oplus \boldsymbol{v})$. 因此, 有 $\boldsymbol{v}^{(k)} \in \boldsymbol{W}(\boldsymbol{u}^{(k)})$. 由 P4, 知

$$\|(\boldsymbol{x}^{(k+1)} - \boldsymbol{x})_{\mathcal{V}}\| = \|(\boldsymbol{x}^{(k)} - \boldsymbol{x})_{\mathcal{V}}\| = o(\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|) = o(\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|). \tag{10}$$

由于 $\nabla^2 L_0(\boldsymbol{0})$ 存在且 $\nabla L_0(\boldsymbol{0}) = \boldsymbol{0}$, 我们由 $\mathcal{U}$-Hessian 阵的定义有

$$\partial L_0(\boldsymbol{u}^{(k)}) \in \boldsymbol{g}_{\mathcal{U}}^{(k)} = \boldsymbol{0} + \nabla^2 L_0(\boldsymbol{0})\boldsymbol{u}^{(k)} + o(\|\boldsymbol{u}^{(k)}\|).$$

由式(9), 有 $\nabla^2 L_0(\boldsymbol{0})(\boldsymbol{u}^{(k)} + \delta_{\mathcal{U}}^{(k)}) = o(\|\boldsymbol{u}^{(k)}\|)$. 由假设 $\nabla^2 L_0(\boldsymbol{0}) \succ \boldsymbol{0}$ 知, $\nabla^2 L_0(\boldsymbol{0})$ 是可逆的, 因此, 有 $\|\boldsymbol{u}^{(k)} + \delta_{\mathcal{U}}^{(k)}\| = o(\|\boldsymbol{u}^{(k)}\|)$. 于是, 有

$$(x^{(k+1)}-x)_{\mathcal{U}}=(x^{(k+1)}-x^{(k)})_{\mathcal{U}}+(x^{(k)}-x^{(k)})_{\mathcal{U}}+(x^{(k)}-x)_{\mathcal{U}}=$$
$$u^{(k)}+\delta_{\mathcal{U}}^{(k)}=o(\|u^{(k)}\|_{\mathcal{U}})=o(\|x^{(k)}-x\|). \quad (11)$$

由式(10)和(11)定理得证.

[参 考 文 献]

[1] Ye J J, Zhu D L, Zhu Q J. Exact penalization and necessary optimality conditions for generalized bilevel programming problems[J]. SIAM Journal on Optimization, 1997, **7**(2): 481-507.

[2] Luo Z Q, Pang J S, Ralph D. Mathematical Programs With Equilibrium Constraints [M]. Cambridge: Cambridge University Press, 1996.

[3] Outrata J V, Kočvara M, Zowe J. Nonsmooth Approach to Optimization Problem With Equilibrium Constraints: Theory, Application and Numerical Results [M]. Dordrecht, Netherlands: Kluwer, 1998.

[4] Lemarchal C, Oustry F, Sagastizâbal C. The $\mathcal{U}$-Lagrangian of a convex function[J]. Transactions of the American Mathematical Society, 2000, **352**(2): 711-729.

[5] Lemarchal C, Sagastizâbal C. More than first-order developments of convex function: primal-dual relations[J]. Journal of Convex Analysis, 1996, **3**(2): 1-14.

[6] Rockafellar R T. Convex Analysis [M]. NJ: Princeton University Press, 1970.

A $\mathcal{UV}$-Decomposed Method for Solving an MPEC Problem

SHAN Feng[1], PANG Li-ping[2], ZHU Li-mei[1], XIA Zun-quan[2]

(1. School of Science Courses, Shenyang Institute of Aeronautical Engineering, Shenyang 110136, P. R. China;

2. CORA, Department of Applied Mathematics; Dalian University of Technology, Dalian 116024, P. R. China)

Abstract: A $\mathcal{UV}$-decomposition method for solving an MPEC problem with linear complementarity constraints is presented. First of all the problem was converted into a nonlinear programming one, and the structure of subdifferential of a corresponding penalty function and results of its $\mathcal{UV}$-decomposition were given. Then a conceptual algorithm for solving this problem with a superlinear convergence rate was constructed in terms of the results obtained.

Key words: nonsmooth optimization; nonlinear programming; subdifferential; $\mathcal{UV}$-decomposition; $\mathcal{U}$-Lagrangian; MPEC problem