核范数函数的广义变分不等式问题的稳定性

赵亚莉¹, 张语乐^{2*}, 张立卫²

1. 渤海大学数学科学学院, 锦州 121013;

2. 大连理工大学数学科学学院运筹学与控制论研究所, 大连 116024

E-mail: yalizhao2000@163.com, zyl@mail.dlut.edu.cn, lwzhang@dlut.edu.cn

收稿日期: 2018-03-08; 接受日期: 2020-05-01; 网络出版日期: 2020-12-22; * 通信作者

国家自然科学基金 (批准号: 11371070, 11971089 和 11731013)、辽宁省教育厅项目 (批准号: LJ2019011) 和辽宁省自然科学基金计划指导项目 (批准号: 2019-ZD-0502) 资助项目

摘要 本文主要研究凸函数为线性矩阵映射核范数的广义变分不等式 (generalized variational inequality, GVI) 问题的稳定性, 提出 GVI 问题的严格 Robinson 约束规范和二阶充分最优性条件, 并证明它们是 KKT (Karush-Kuhn-Tucker) 映射逆映射孤立平稳性的充分条件; 还给出 GVI 问题的约束非退化条件和强二阶充分最优性条件, 并证明它们是 KKT 系统强正则性的充分条件.

关键词 广义变分不等式 核矩阵范数 稳定性 孤立平稳性 强正则性 严格 Robinson 约束规范

MSC (2020) 主题分类 90C30

1 引言

一类广义变分不等式问题可以描述为: 求 $x \in K$ 满足

$$\langle F(x), y - x \rangle + \phi(y) - \phi(x) \geq 0, \quad \forall y \in K, \quad (1.1)$$

其中 $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是连续可微映射, $\phi: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ 是正常下半连续凸函数, $K \subset \mathbb{R}^n$ 是闭凸集. 本文仅考虑 ϕ 是定义在 $\mathbb{R}^n$ 上的仿射对称矩阵映射的核范数、 K 为凸约束集合的情形, 即

$$\phi(x) = \left\| B_0 + \sum_{i=1}^n x_i B_i \right\|_*, \quad (1.2)$$

且 K 具有下述形式:

$$K = \{y \in \mathbb{R}^n : Ay = b, g(y) \leq 0\}, \quad (1.3)$$

其中 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ 是映射. 记 $g = (g_1, \dots, g_p)^T$, 并且假设每个 g_i 是二次连续可微凸函数.

英文引用格式: Zhao Y L, Zhang Y L, Zhang L W. The stability of a generalized variational inequality problem with the nuclear norm function (in Chinese). Sci Sin Math, 2022, 52: 85–104, doi: 10.1360/SCM-2018-0188

广义变分不等式 (1.1) 有许多重要应用. 例如, 如果 $F(x) = \nabla f(x)$, 其中 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是连续可微凸函数, 则 (1.1) 与下面的凸最优化问题密切相关:

$$\begin{aligned} \min \quad & f(x) + \phi(x) \\ \text{s.t.} \quad & Ax = b, \\ & g(x) \leq 0. \end{aligned} \quad (1.4)$$

例如, 当存在向量 $x_0 \in \mathbb{R}^n$ 满足

$$x_0 \in \text{ri dom } \phi, \quad Ax_0 = b, \quad g(x_0) < 0, \quad (1.5)$$

其中 ri 表示相对内部, $\text{dom } \phi$ 表示函数 ϕ 的有效域, 即 $\text{dom } \phi = \{x : \phi(x) < +\infty\}$, 则问题 (1.4) 的最优性条件就是变分不等式 (1.1). 事实上, 问题 (1.4) 可表示为

$$\min \hat{f}(x) = f(x) + \phi(x) + \delta_K(x), \quad (1.6)$$

其中 δ_K 是 K 的指示函数. 显然, $\hat{f}$ 是正常的下半连续凸函数, 如果 $\bar{x}$ 是问题 (1.6) 的解, 则

$$0 \in \partial \hat{f}(\bar{x}). \quad (1.7)$$

注意到

$$x_0 \in \text{ri dom } f \cap \text{ri dom } \phi \cap \text{ri dom } \delta_K,$$

由凸函数和的次微分性质, 可得

$$\partial \hat{f}(\bar{x}) = \nabla f(\bar{x}) + \partial \phi(\bar{x}) + N_K(\bar{x}), \quad (1.8)$$

其中 $N_K(\bar{x})$ 是 K 在 $\bar{x}$ 的法锥, 定义如下:

$$N_K(\bar{x}) = \{v : \langle v, x - \bar{x} \rangle \leq 0, \forall x \in K\},$$

这里 $\bar{x} \in K$. 注意到

$$\partial[\phi + \delta_K](\bar{x}) = \partial \phi(\bar{x}) + N_K(\bar{x}),$$

由 (1.7) 和 (1.8) 得

$$-F(\bar{x}) \in \partial[\phi + \delta_K](\bar{x}).$$

根据次微分的定义, 上式蕴涵着

$$0 \leq \langle F(\bar{x}), y - \bar{x} \rangle + [\phi(y) + \delta_K(y)] - [\phi(\bar{x}) + \delta_K(\bar{x})],$$

这恰好是当 $x = \bar{x}$ 时 (1.1) 的变形形式.

凸优化问题模型 (1.4) 在矩阵补全、图像处理和统计学习方面有许多应用. 文献中不乏求解问题 (1.4) 的有效数值算法. 流行的数值算法包括 Rockafellar^[1] 的临近点算法和被大家广泛研究的交替乘子方向方法. 这两种算法的收敛率的研究通常要求解映射的误差界条件, 参见文献 [1-4]. 问题模型

(1.1) 比问题模型 (1.4) 更一般, 问题 (1.1) 数值算法收敛率的研究也要求误差界条件成立, 这很自然促使对问题 (1.1) 的稳定性进行研究.

最优化问题和变分不等式的稳定性是最优化理论中的重要课题, 与这个课题相关的文献有很多. 对非线性规划而言, 稳定性理论非常完善. 例如, KKT 系统的强正则性由 Robinson^[5] 和 Jongen 等^[6] 给出了完整的刻画, Dontchev 和 Rockafellar^[7] 建立了凸多面体集上的仿射变分不等式的强正则性与 Aubin 性质的等价性, Robinson^[8]、Dontchev 和 Rockafellar^[9] 则刻画了非线性规划问题的 KKT 映射的上 Lipschitz 连续性. 关于广义非线性规划问题的上 Lipschitz 连续性, 可参见文献 [10,11]. 对非线性二阶锥最优化, KKT 系统的强正则性刻画由 Bonnans 和 Ramírez C^[12]、Wang 和 Zhang^[13] 给出. Sun^[14] 刻画了非线性半定规划 (semidefinite programming, SDP) 的 KKT 系统的强正则性, Chan 和 Sun^[15] 证明了线性 SDP 问题的强二阶条件与对偶问题的约束非退化条件的等价性.

值得关注的是, 近些年来, 矩阵优化的稳定性分析取得重要的进展. 例如, Zhang 和 Zhang^[16] 证明了非线性 SDP 问题在严格 Robinson 约束规范 (strict Robinson constraint qualification, SRCQ) 和二阶充分最优性条件下 KKT 映射的孤立平稳性. Han 等^[4] 给出了凸复合二次半定规划问题的 KKT 映射的孤立平稳性. Liu 和 Pan^[17] 把文献 [4,16] 中的结果推广到了矩阵最优化问题的 KKT 映射的孤立平稳性, 涉及 Ky Fan k -范数的上图锥. 对一类锥规划问题, Ding 等^[18] 在 Robinson 约束规范下证明了 KKT 解映射是鲁棒孤立平稳的当且仅当严格 Robinson 约束规范和二阶充分条件都成立. 最近, Cui 和 Sun^[19] 给出了核范数正则凸最优化问题的鲁棒孤立平稳性的刻画. 特别地, 文献 [19] 建立了在稳定性分析中起着重要作用的非对称矩阵的核范数的变分几何.

本文主要讨论问题 (1.1) 的稳定性 (其中 ϕ 由 (1.2) 定义, K 由 (1.3) 定义), 包括 KKT 映射的逆映射的孤立平稳性和 KKT 系统的强正则性. 实际上, 如果考虑最优化问题 (1.4), 可先将它转化为锥规划问题, 从而文献 [19] 中的理论可应用到问题 (1.4) 的稳定性. 因此, 对稳定性分析, 模型 (1.1) 的重要性在于 F 不是某个光滑函数的梯度映射的情形.

本文余下内容的结构如下: 第 2 节给出核范数的邻近映射的一些变分性质, 这些性质将在孤立平稳性的分析中用到. 第 3 节提出并证明关于广义变分不等式问题的 KKT 映射的逆映射的孤立平稳性的充分条件. 第 4 节研究广义变分不等式问题 KKT 系统的强正则性. 第 5 节给出本文的结论.

2 预备知识

本节首先回顾与集值映射相关的一些定义, 然后建立核范数的邻近映射的变分性质, 它们将在随后的讨论中用到.

设 $\mathcal{Y}$ 是有限维实 Euclid 空间, $S: \mathbb{R}^n \rightrightarrows \mathcal{Y}$ 是集值映射, $(\bar{x}, \bar{y}) \in \text{gph } S$, $\text{gph } S = \{(x, y) \in \mathbb{R}^n \times \mathcal{Y} : y \in S(x)\}$.

定义 2.1 集值映射 S 在 $\bar{x}$ 关于 $\bar{y}$ 称为是孤立平稳的, 如果存在常数 $\kappa > 0$ 、 $\bar{x}$ 的邻域 $\mathcal{U}$ 和 $\bar{y}$ 的邻域 $\mathcal{V}$, 使得

$$S(x) \cap \mathcal{V} \subset \{\bar{y}\} + \kappa\|x - \bar{x}\|\mathcal{B}_{\mathcal{Y}}, \quad \forall x \in \mathcal{U},$$

其中 $\mathcal{B}_{\mathcal{Y}} \subseteq \mathcal{Y}$ 是 $\mathcal{Y}$ 中的单位球.

图导数是研究孤立平稳性的一个合适的工具.

定义 2.2 S 在 $\bar{x}$ 关于 $\bar{y}$ 的图导数是集值映射 $DS(\bar{x} | \bar{y}): \mathbb{R}^n \rightrightarrows \mathcal{Y}$, 定义为

$$DS(\bar{x} | \bar{y})(u) := \{v \in \mathcal{Y} \mid (u, v) \in \mathcal{T}_{\text{gph } S}(\bar{x}, \bar{y})\}.$$

在孤立平稳性主要定理的证明中, 将用到集值映射 S 在 $\bar{x}$ 关于 $\bar{y}$ 孤立平稳性的下面的基本刻画.

引理 2.1 ^[20, 21] 设 $(\bar{x}, \bar{y}) \in \text{gph } S$, 则 S 在 $\bar{x}$ 关于 $\bar{y}$ 是孤立平稳的当且仅当

$$\{0\} = DS(\bar{x} \mid \bar{y})(0).$$

$\mathbb{S}^q$ 中对称矩阵的核范数是关于对称函数

$$\varsigma(z) = \sum_{i=1}^q |z_i|, \quad z = (z_1, \dots, z_q)^T \in \mathbb{R}^q$$

的谱分解函数, 即 $\|X\|_* = [\varsigma \circ \lambda](X)$. 由文献 [22, 第 121 页] 知, 对 $\bar{X} \in \mathbb{S}^q$, 有 $Y \in \partial[\varsigma \circ \lambda](\bar{X})$ 当且仅当存在 $w \in \partial\varsigma(\lambda(\bar{X}))$, 使得

$$Y = P \text{Diag}(w) P^T,$$

其中 $\text{Diag}(w)$ 是对角线元素为 w 分量的对角矩阵, 并且 X 有谱分解 $\bar{X} = P \text{Diag}(\lambda(\bar{X})) P^T$. 定义 3 个指标集如下:

$$\alpha = \{i : \lambda_i(\bar{X}) > 0\}, \quad \beta = \{i : \lambda_i(\bar{X}) = 0\}, \quad \gamma = \{i : \lambda_i(\bar{X}) < 0\},$$

则 $w \in \partial\varsigma(\lambda(\bar{X}))$ 具有下面的性质:

$$w_\alpha = \mathbf{1}_{|\alpha|}, \quad w_\gamma = -\mathbf{1}_{|\gamma|} \quad \text{和} \quad -\mathbf{1}_{|\beta|} \leq w_\beta \leq \mathbf{1}_{|\beta|},$$

以及

$$Y = P \text{Diag}(w) P^T = P_\alpha P_\alpha^T + P_\beta \text{Diag}(w_\beta) P_\beta^T - P_\gamma P_\gamma^T,$$

且 Y 的上述表示也可由文献 [23] 得到. 设 $\theta(X) = \|X\|_*$, 将其 Moreau-Yosida 正则化, 记为 $[e_\tau \theta](X)$, 定义为

$$[e_\tau \theta](X) = \inf_{X' \in \mathbb{S}^q} \left\{ \theta(X') + \frac{1}{\tau} \|X' - X\|^2 \right\},$$

θ 的邻近映射定义为

$$[\mathcal{P}_\tau \theta](X) = \underset{X' \in \mathbb{S}^q}{\text{argmin}} \left\{ \theta(X') + \frac{1}{\tau} \|X' - X\|^2 \right\}.$$

为简单起见, 用 $\mathcal{P}\theta$ 代表 $\mathcal{P}_\tau \theta$.

于是, $[e_\tau \theta](X)$ 是相应于 Moreau-Yosida 正则化 $e_{\tau\varsigma}$ 的谱分解函数, 即

$$[e_\tau \theta](X) = [e_{\tau\varsigma} \circ \lambda](X).$$

由文献 [24, 25] 得

$$[\mathcal{P}_\tau \theta](\bar{X}) = P \text{Diag}([\mathcal{P}_{\tau\varsigma}](\lambda(\bar{X}))) P^T,$$

或者 $[\mathcal{P}_\tau \theta](\bar{X})$ 是与 $p_\tau(t) = [t - \tau]_+ - [-t - \tau]_+$ 相联系的 Löwner 算子, 即

$$[\mathcal{P}_\tau \theta](\bar{X}) = P \text{Diag}(p_\tau(\lambda_1(\bar{X})), \dots, p_\tau(\lambda_q(\bar{X}))) P^T.$$

设 $\bar{X}$ 有 r 个不同的特征值, 其中有 r_1 个不同的正特征值、 $r - r_1$ 个不同的负特征值和零特征值:

$$\mu_1 > \mu_2 > \cdots > \mu_{r_1} > \mu_{r_1+1} = 0 > \mu_{r_1+2} > \cdots > \mu_r.$$

定义

$$\alpha_k = \{i : \lambda_i(\bar{X}) = \mu_k\}, \quad k = 1, \dots, r.$$

定义 $\bar{X}$ 沿 $H \in \mathbb{S}^q$ 的一阶差分矩阵如下: 对 $k, l = 1, \dots, r$,

$$(p_\tau^{[1]}(\Lambda(\bar{X}), P^T H P))_{\alpha_k \alpha_l} := \begin{cases} \frac{p_\tau(\mu_k) - p_\tau(\mu_l)}{\mu_k - \mu_l} P_{\alpha_k}^T H P_{\alpha_l}, & \text{当 } k \neq l \text{ 时,} \\ \Psi_k(P_{\alpha_k}^T H P_{\alpha_k}), & \text{当 } k = l \text{ 时,} \end{cases} \quad (2.1)$$

其中 $\Psi_k(\cdot)$ 是相对于 $\psi_k(\cdot) = p'_\tau(\mu_k; \cdot)$ 的 Löwner 算子, 则由文献 [6] 得 $\mathcal{P}_\tau \theta$ 在 $\bar{X}$ 沿 $H \in \mathbb{S}^q$ 的方向导数可表示为

$$[\mathcal{P}_\tau \theta]'(\bar{X}; H) = P[p_\tau^{[1]}(\Lambda(\bar{X}), P^T H P)]P^T. \quad (2.2)$$

设 $A \in \mathbb{S}^q$, 假设它有谱分解 $A = P \text{Diag}(\lambda) P^T$, 并且定义下面 3 个指标集:

$$\alpha = \{i : \lambda_i > 0\}, \quad \beta = \{i : \lambda_i = 0\}, \quad \gamma = \{i : \lambda_i < 0\},$$

则 $B \in \partial\theta(A)$ 当且仅当存在 $w \in \partial\zeta(\lambda(A))$, 使得

$$B = P \text{Diag}(w) P^T,$$

且 $w \in \partial\zeta(\lambda)$ 有下面的性质:

$$w_\alpha = \mathbf{1}_{|\alpha|}, \quad w_\gamma = -\mathbf{1}_{|\gamma|} \quad \text{和} \quad -\mathbf{1}_{|\beta|} \leq w_\beta \leq \mathbf{1}_{|\beta|}.$$

对指标集 β , 我们给它一个分划如下: $\beta = \beta_L \cup \beta_S \cup \beta_U$, 其中

$$\beta_L = \{i \in \beta : w_i = -1\}, \quad \beta_S = \{i \in \beta : -1 < w_i < 1\}, \quad \beta_U = \{i \in \beta : w_i = 1\}.$$

于是,

$$B = P \text{Diag}(w) P^T = P_{\alpha \cup \beta_U} P_{\alpha \cup \beta_U}^T + P_{\beta_S} \text{Diag}(w_{\beta_S}) P_{\beta_S}^T - P_{\gamma \cup \beta_L} P_{\gamma \cup \beta_L}^T,$$

且对 $C = A + B$, 有

$$C = [P_\alpha \ P_{\beta_U} \ P_{\beta_S} \ P_{\beta_L} \ P_\gamma] \begin{bmatrix} \Lambda_\alpha + I_{|\alpha|} & & & & \\ & I_{|\beta_U|} & & & \\ & & \text{Diag}(w_{\beta_S}) & & \\ & & & -I_{|\beta_L|} & \\ & & & & \Lambda_\gamma - \mathbf{1}_{|\gamma|} \end{bmatrix} \begin{bmatrix} P_\alpha^T \\ P_{\beta_U}^T \\ P_{\beta_S}^T \\ P_{\beta_L}^T \\ P_\gamma^T \end{bmatrix}. \quad (2.3)$$

θ 在 C 与 $B \in \partial\theta(A)$ 相联系的临界锥定义为

$$\mathcal{C}_\theta(C) = \left\{ \Delta A \in \mathbb{S}^q : \begin{array}{l} P_{\beta_S}^\top(\Delta A)[P_\beta] = 0, \\ P_{\beta_U}^\top(\Delta A)[P_{\beta_L}] = 0, \\ P_{\beta_U}^\top(\Delta A)P_{\beta_U} \in \mathbb{S}_+^{|\beta_U|}, \\ P_{\beta_L}^\top(\Delta A)P_{\beta_L} \in \mathbb{S}_-^{|\beta_L|} \end{array} \right\}. \quad (2.4)$$

定义 $\Upsilon \in \mathbb{S}^q$ 是对称矩阵, 满足

$$\Upsilon_{ij} = \begin{cases} \frac{2}{\lambda_i}, & i \in \alpha, \quad j \in \beta_L, \\ \frac{1-w_j}{\lambda_i}, & i \in \alpha, \quad j \in \beta_S, \\ \frac{2}{\lambda_i - \lambda_j}, & i \in \alpha, \quad j \in \gamma, \\ \frac{-2}{\lambda_i}, & i \in \gamma, \quad j \in \beta_U, \\ \frac{-1-w_j}{\lambda_i}, & i \in \gamma, \quad j \in \beta_S, \\ 0, & \text{否则.} \end{cases} \quad (2.5)$$

引理 2.2 设 $A, B, C \in \mathbb{S}^q$, $C = A + B$ 满足 $B \in \partial\theta(A)$. 设 ΔA 和 ΔB 是 $\mathbb{S}^q$ 中的两个矩阵, 则

$$\Delta A - [\mathcal{P}\theta]'(C; \Delta A + \Delta B) = 0 \quad (2.6)$$

当且仅当

$$\begin{aligned} P_\alpha^\top(\Delta B)[P_\alpha \ P_{\beta_U}] &= 0, \\ P_\gamma^\top(\Delta B)[P_\gamma \ P_{\beta_L}] &= 0, \\ P_{\beta_S}^\top(\Delta A)[P_\beta] &= 0, \\ P_{\beta_U}^\top(\Delta A)[P_{\beta_L}] &= 0, \\ P_{\beta_U}^\top(\Delta A)P_{\beta_U} &= \Pi_{\mathbb{S}_+^{|\beta_U|}}(P_{\beta_U}^\top(\Delta A + \Delta B)P_{\beta_U}), \\ P_{\beta_L}^\top(\Delta A)P_{\beta_L} &= \Pi_{\mathbb{S}_-^{|\beta_L|}}(P_{\beta_L}^\top(\Delta A + \Delta B)P_{\beta_L}), \\ P_\alpha^\top(\Delta B)P_{\beta_S \cup \beta_L \cup \gamma} &= P_\alpha^\top(\Delta A)P_{\beta_S \cup \beta_L \cup \gamma} \circ \Upsilon_{\alpha, \beta_S \cup \beta_L \cup \gamma}, \\ P_\gamma^\top(\Delta B)P_{\beta_U \cup \beta_S} &= P_\gamma^\top(\Delta A)P_{\beta_U \cup \beta_S} \circ \Upsilon_{\gamma, \beta_U \cup \beta_S}. \end{aligned} \quad (2.7)$$

而且, (2.7) 中的关系蕴涵

$$\Delta A \in \mathcal{C}_\theta(C) \quad \text{和} \quad \langle \Delta A, \Delta B \rangle = 2\langle P^\top[\Delta A]P, \Upsilon \circ P^\top[\Delta A]P \rangle.$$

证明 注意到 $\tau = 1$ 和 $\tau = -1$ 是 $p_1(\cdot) = [\mathcal{P}_1\zeta](\cdot)$ 的两个非光滑点, 对 $\hat{H} = P^T H P$, 由 (2.2) 得

$$P^T[\mathcal{P}\theta]'(C; H)P = \begin{bmatrix} \hat{H}_{\alpha\alpha} & \hat{H}_{\alpha\beta_U} & \hat{H}_{\alpha\beta_S} \circ \Omega_{\alpha\beta_S} & \hat{H}_{\alpha\beta_L} \circ \Omega_{\alpha\beta_L} & \hat{H}_{\alpha\gamma} \circ \Omega_{\alpha\gamma} \\ \hat{H}_{\beta_U\alpha} & \Pi_{\mathbb{S}_+^{|\beta_U|}}(\hat{H}_{\beta_U\beta_U}) & 0 & 0 & \hat{H}_{\beta_U\gamma} \circ \Omega_{\beta_U\gamma} \\ \hat{H}_{\beta_S\alpha} \circ \Omega_{\beta_S\alpha} & 0 & 0 & 0 & \hat{H}_{\beta_S\gamma} \circ \Omega_{\beta_S\gamma} \\ \hat{H}_{\beta_L\alpha} \circ \Omega_{\beta_L\alpha} & 0 & 0 & \Pi_{\mathbb{S}_-^{|\beta_L|}}(\hat{H}_{\beta_L\beta_L}) & \hat{H}_{\beta_L\gamma} \\ \hat{H}_{\gamma\alpha} \circ \Omega_{\gamma\alpha} & \hat{H}_{\gamma\beta_U} \circ \Omega_{\gamma\beta_U} & \hat{H}_{\gamma\beta_S} \circ \Omega_{\gamma\beta_S} & \hat{H}_{\gamma\beta_L} & \hat{H}_{\gamma\gamma} \end{bmatrix}, \quad (2.8)$$

其中

$$\Omega_{ij} = [p_1^{[1]}(\Lambda(C))]_{ij}, \quad (i, j) \in \alpha \times [\beta_S \cup \beta_L \cup \gamma] \quad \text{或} \quad (i, j) \in \gamma \times [\beta_U \cup \beta_S].$$

由 (2.8) 得 (2.6) 与 (2.7) 等价. 由 $\mathcal{C}_{\mathcal{P}\theta}(C)$ 的定义, 可得 $\Delta A \in \mathcal{C}_{\mathcal{P}\theta}(C)$. 由

$$P_{\beta_U}^T(\Delta A)P_{\beta_U} = \Pi_{\mathbb{S}_+^{|\beta_U|}}(P_{\beta_U}^T(\Delta A + \Delta B)P_{\beta_U})$$

和

$$P_{\beta_L}^T(\Delta A)P_{\beta_L} = \Pi_{\mathbb{S}_-^{|\beta_L|}}(P_{\beta_L}^T(\Delta A + \Delta B)P_{\beta_L}),$$

我们得

$$\langle P_{\beta_U}^T(\Delta A)P_{\beta_U}, P_{\beta_U}^T(\Delta B)P_{\beta_U} \rangle = 0, \quad \langle P_{\beta_L}^T(\Delta A)P_{\beta_L}, P_{\beta_L}^T(\Delta B)P_{\beta_L} \rangle = 0.$$

因此, 由 (2.7) 得

$$\begin{aligned} \langle \Delta A, \Delta B \rangle &= \langle P^T[\Delta A]P, P^T[\Delta B]P \rangle \\ &= 2\langle P_{\alpha}^T[\Delta A]P_{\beta_S \cup \beta_L \cup \gamma}, P_{\alpha}^T[\Delta B]P_{\beta_S \cup \beta_L \cup \gamma} \rangle + 2\langle P_{\gamma}^T[\Delta A]P_{\beta_U \cup \beta_S}, P_{\gamma}^T[\Delta B]P_{\beta_U \cup \beta_S} \rangle \\ &= 2\langle P^T[\Delta A]P, \Upsilon \circ P^T[\Delta A]P \rangle. \end{aligned}$$

证毕. □

下面的命题是对 $\partial\theta(A)$ 中元素的刻画, 其证明可参见文献 [26].

命题 2.1 设 $A, B, C \in \mathbb{S}^q$, $C = A + B$ 满足 $B \in \partial\theta(A)$, 则 $\mathcal{P}\theta$ 是强半光滑的, 并且对 $W \in \partial\mathcal{P}\theta(C)$, 存在 $W_{\beta_U} \in \partial\Pi_{\mathbb{S}_+^{|\beta_U|}}(0)$ 和 $W_{\beta_L} \in \partial\Pi_{\mathbb{S}_-^{|\beta_L|}}(0)$ 满足

$$W(H) = P \begin{bmatrix} \hat{H}_{\alpha\alpha} & \hat{H}_{\alpha\beta_U} & \hat{H}_{\alpha\beta_S} \circ \Omega_{\alpha\beta_S} & \hat{H}_{\alpha\beta_L} \circ \Omega_{\alpha\beta_L} & \hat{H}_{\alpha\gamma} \circ \Omega_{\alpha\gamma} \\ \hat{H}_{\beta_U\alpha} & W_{\beta_U}(\hat{H}_{\beta_U\beta_U}) & 0 & 0 & \hat{H}_{\beta_U\gamma} \circ \Omega_{\beta_U\gamma} \\ \hat{H}_{\beta_S\alpha} \circ \Omega_{\beta_S\alpha} & 0 & 0 & 0 & \hat{H}_{\beta_S\gamma} \circ \Omega_{\beta_S\gamma} \\ \hat{H}_{\beta_L\alpha} \circ \Omega_{\beta_L\alpha} & 0 & 0 & W_{\beta_L}(\hat{H}_{\beta_L\beta_L}) & \hat{H}_{\beta_L\gamma} \\ \hat{H}_{\gamma\alpha} \circ \Omega_{\gamma\alpha} & \hat{H}_{\gamma\beta_U} \circ \Omega_{\gamma\beta_U} & \hat{H}_{\gamma\beta_S} \circ \Omega_{\gamma\beta_S} & \hat{H}_{\gamma\beta_L} & \hat{H}_{\gamma\gamma} \end{bmatrix} P^T, \quad (2.9)$$

其中 $\hat{H} = P^T H P$,

$$\Omega_{ij} = [p_1^{[1]}(\Lambda(C))]_{ij}, \quad (i, j) \in \alpha \times [\beta_S \cup \beta_L \cup \gamma] \quad \text{或} \quad (i, j) \in \gamma \times [\beta_U \cup \beta_S].$$

下面的引理利用 $\theta(\cdot) = \|\cdot\|_*$ 的方向导数给出了临界锥 $\mathcal{C}_\theta$ 的刻画.

引理 2.3 令 $A, B, C \in \mathbb{S}^q$, $C = A + B$ 满足 $B \in \partial\theta(A)$, 则 $\Delta A \in \mathcal{C}_\theta(C)$ 当且仅当

$$\theta'(A; \Delta A) = \langle B, \Delta A \rangle. \quad (2.10)$$

证明 注意到

$$\partial\theta(A) = \{P_\alpha P_\alpha^T + P_\beta W_\beta P_\beta^T - P_\gamma P_\gamma^T : W_\beta \in \mathbb{S}^{|\beta|}, \|W_\beta\|_2 \leq 1\},$$

其中 $\|W_\beta\|_2$ 表示 W_β 的谱范数. θ 在 A 沿 ΔA 的方向导数是

$$\theta'(A; \Delta A) = \langle P_\alpha P_\alpha^T - P_\gamma P_\gamma^T, \Delta A \rangle + \|P_\beta^T [\Delta A] P_\beta\|_*.$$

注意到 B 有表示

$$B = P_\alpha P_\alpha^T + P_\beta \text{Diag}(w_\beta) P_\beta^T - P_\gamma P_\gamma^T,$$

其中 $w_\beta \in \mathbb{R}^{|\beta|}$ 满足 $\|w_\beta\|_\infty \leq 1$, 则 $\theta'(A; \Delta A) = \langle B, \Delta A \rangle$ 等价于

$$\|P_\beta^T [\Delta A] P_\beta\|_* = \langle P_\beta^T [\Delta A] P_\beta, \text{Diag}(w_\beta) \rangle. \quad (2.11)$$

由 Fan^[27] 的不等式, 可得

$$\langle P_\beta^T [\Delta A] P_\beta, \text{Diag}(w_\beta) \rangle \leq \lambda(P_\beta^T [\Delta A] P_\beta)^T w_\beta.$$

由 (2.11), 对

$$\lambda(P_\beta^T [\Delta A] P_\beta) = (\lambda_1(P_\beta^T [\Delta A] P_\beta), \dots, \lambda_{|\beta|}(P_\beta^T [\Delta A] P_\beta))^T$$

且

$$\lambda_1(P_\beta^T [\Delta A] P_\beta) \geq \dots \geq \lambda_{|\beta|}(P_\beta^T [\Delta A] P_\beta),$$

上式蕴涵着

$$\langle P_\beta^T [\Delta A] P_\beta, \text{Diag}(w_\beta) \rangle = \lambda(P_\beta^T [\Delta A] P_\beta)^T w_\beta = \|P_\beta^T [\Delta A] P_\beta\|_*,$$

则 $P_\beta^T [\Delta A] P_\beta$ 和 $\text{Diag}(w_\beta)$ 有相同阶特征值分解, 因此可验证 ΔA 满足

$$\begin{aligned} P_{\beta_S}^T (\Delta A) [P_\beta] &= 0, \\ P_{\beta_U}^T (\Delta A) [P_{\beta_L}] &= 0, \\ P_{\beta_U}^T (\Delta A) P_{\beta_U} &\in \mathbb{S}_+^{|\beta_U|}, \\ P_{\beta_L}^T (\Delta A) P_{\beta_L} &\in \mathbb{S}_-^{|\beta_L|}, \end{aligned}$$

即 $\Delta A \in \mathcal{C}_\theta(C)$. 显然容易验证 $\Delta A \in \mathcal{C}_\theta(C)$ 蕴涵 $\theta'(A; \Delta A) = \langle B, \Delta A \rangle$. 证毕. $\square$

引理 2.4 设 $A, B, C \in \mathbb{S}^q$, $C = A + B$ 满足 $B \in \partial\theta(A)$, 则

$$[\mathcal{P}\theta]'(C; V) = 0, \quad V \in \mathbb{S}^q \Leftrightarrow V \in \mathcal{C}_\theta(C)^\circ. \quad (2.12)$$

证明 此结论直接证明可得. $\square$

注 2.1 引理 2.2 和 2.3 中的结果可由文献 [19, 命题 4.1] 得到, 该命题处理了非对称矩阵的核范数. 为完整起见, 在这里给出了引理 2.2 的详细证明. 关系 (2.12) 也出现在文献 [19, 命题 4.2] 中, 用了不同的表示.

引理 2.5 设 $\mathcal{A}: \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 是线性算子且 $(c_1, c_2, C) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$, $c_1 = 0$, $C = A + B$, $B \in \partial\theta(A)$, 则 $v := (v_1, v_2, V) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 是方程组

$$\begin{aligned} \mathcal{A}^* v &= 0, \\ \Pi'_{\mathbb{R}^p_-}(c_2; v_2) &= 0, \\ [\mathcal{P}\theta]'(C; V) &= 0 \end{aligned} \quad (2.13)$$

的解当且仅当

$$v \in \left[\mathcal{A}\mathbb{R}^n + \begin{pmatrix} \{0_q\} \\ \mathcal{T}_{\mathbb{R}^p_-}([c_2]_-) \cap ([c_2]_+)^{\perp} \\ \mathcal{C}_\theta(C) \end{pmatrix} \right]^\circ, \quad (2.14)$$

其中

$$[c_2]_- := \Pi_{\mathbb{R}^p_-}(c), \quad [c_2]_+ = c_2 - [c_2]_-.$$

证明 由文献 [4, 引理 4.2] 得

$$\Pi'_{\mathbb{R}^p_-}(c_2; v_2) = 0 \Leftrightarrow v_2 \in [\mathcal{T}_{\mathbb{R}^p_-}([c_2]_+) \cap [c_2]_-^{\perp}]^\circ.$$

由引理 2.4 得

$$[\mathcal{P}\theta]'(C; V) = 0 \Leftrightarrow V \in \mathcal{C}_\theta(C)^\circ.$$

因此, v 满足 (2.13) 当且仅当

$$\mathcal{A}^* v = 0, \quad v \in \left[\begin{pmatrix} \{0_q\} \\ \mathcal{T}_{\mathbb{R}^p_-}([c_2]_-) \cap ([c_2]_+)^{\perp} \\ \mathcal{C}_\theta(C) \end{pmatrix} \right]^\circ,$$

它等价于 (2.14) 成立. 证毕. $\square$

3 GVI 的孤立平稳性

本节主要讨论广义变分不等式问题的 KKT 映射的逆映射的孤立平稳性, 描述主要的定理, 它给出孤立平稳性的充分条件.

设 $\mathcal{B} : \mathbb{R}^n \rightarrow \mathbb{S}^q$ 是线性算子, 定义为

$$\mathcal{B}x = \sum_{j=1}^n x_j B_j,$$

则 $\phi(x) = \|B_0 + \mathcal{B}x\|_*$, $\mathcal{B}$ 的伴随算子记为 $\mathcal{B}^* : \mathbb{S}^q \rightarrow \mathbb{R}^n$, 满足

$$\mathcal{B}^*Z = \begin{bmatrix} \langle B_1, Z \rangle \\ \vdots \\ \langle B_n, Z \rangle \end{bmatrix}, \quad \text{其中 } Z \in \mathbb{S}^q.$$

在问题 (1.1) 中, 函数 ϕ 是复合函数, 这使得 ϕ 的邻近映射难以表达. 我们引入一个新的变量 $Z \in \mathbb{S}^q$ 来表达问题 (1.1), 即找到 (x, Z) 使得

$$\langle F(x), x' - x \rangle + \theta(Z') - \theta(Z) \geq 0, \quad \forall (x', Z') \in \Phi, \quad (3.1)$$

其中 $\theta(Z) = \|Z\|_*$ 是 $Z \in \mathbb{S}^q$ 的核范数, Φ 定义为

$$\Phi = \left\{ (x, Z) \in \mathbb{R}^n \times \mathbb{S}^q : Z = B_0 + \sum_{i=1}^n x_i B_i, Ax = b, g(x) \leq 0 \right\}. \quad (3.2)$$

为建立问题 (1.1) 的最优性条件, 记问题 (1.1) 为 $(\text{GVI}(F, \phi, K))$. 首先考虑问题 (3.1) 的最优性条件, 问题 (3.1) 记为 $(\text{GVI}(F, \theta, \Phi))$. 为得到 $(\text{GVI}(F, \theta, \Phi))$ 方程形式的最优性条件, 我们需要下面的 Slater 条件.

假设 3.1 存在点 $x_0 \in \mathbb{R}^n$ 满足

$$Ax_0 = b, \quad g(x_0) < 0. \quad (3.3)$$

令

$$\hat{\phi}(x, Z) = \theta(Z) + \delta_{\Phi}(x, Z),$$

则 $(\text{GVI}(F, \theta, \Phi))$ 可写为

$$\langle F(x), x' - x \rangle + \hat{\phi}(x', Z') - \hat{\phi}(x, Z) \geq 0, \quad \forall (x', Z') \in \mathbb{R}^n \times \mathbb{S}^q. \quad (3.4)$$

问题 (3.4) 等价于

$$0 \in \begin{bmatrix} F(x) \\ \{0\} \end{bmatrix} + \partial \hat{\phi}(x, Z). \quad (3.5)$$

如果 Slater 条件, 即假设 3.1 成立, 则 (注意 $\text{dom } \theta = \mathbb{S}^q$) 对 $(x, Z) \in \Phi$,

$$\partial \hat{\phi}(x, Z) = \begin{bmatrix} \{0\} \\ \partial \theta(Z) \end{bmatrix} + N_{\Phi}(x, Z)$$

$$= \begin{bmatrix} \{0\} \\ \partial\theta(Z) \end{bmatrix} + \begin{bmatrix} \mathcal{B}^* \\ -\mathcal{I} \end{bmatrix} \mathbb{S}^q + \begin{bmatrix} A^T \mathbb{R}^m + \{\mathcal{J}g(x)^T \lambda : 0 \leq \lambda \perp g(x) \leq 0\} \\ \{0\} \end{bmatrix}. \quad (3.6)$$

因此, 包含关系 (3.5) 变为

$$\begin{aligned} 0 &= F(x) + A^T \mu + \mathcal{J}g(x)^T \lambda + \mathcal{B}^* Y, \\ Ax &= b, \\ Z &= B_0 + \sum_{i=1}^n x_i B_i, \\ 0 &\leq \lambda \perp g(x) \leq 0, \\ 0 &\in \partial\theta(Z) - Y, \end{aligned} \quad (3.7)$$

其中 $\mu \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^p$, $Y \in \mathbb{S}^q$. 因此, 我们得到 (GVI(F, ϕ, K)) 最优性条件的方程形式.

命题 3.1 假设 $\bar{x} \in \mathbb{R}^n$ 是 (GVI(F, ϕ, K)) 的解, 且假设 (1.1) 成立, 则存在 $\bar{\mu} \in \mathbb{R}^m$, $\bar{\lambda} \in \mathbb{R}^p$, $\bar{Y} \in \mathbb{S}^q$ 使得 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})$ 是下述非光滑方程:

$$\begin{bmatrix} L(x, \mu, \lambda, Y) \\ Ax - b \\ g(x) - \Pi_{\mathbb{R}_+^p}(g(x) + \lambda) \\ B_0 + \mathcal{B}x - [\mathcal{P}\theta](B_0 + \mathcal{B}x + Y) \end{bmatrix} = 0 \quad (3.8)$$

的解, 其中

$$L(x, \mu, \lambda, Y) = F(x) + A^T \mu + \mathcal{J}g(x)^T \lambda + \mathcal{B}^* Y$$

是 (GVI(F, ϕ, K)) 的标准 Lagrange 函数.

满足 $G(\bar{x}, \mu, \lambda, Y) = 0$ 的所有 (μ, λ, Y) 构成的集合记为 $\Lambda(\bar{x})$, 称为在 $\bar{x}$ 的乘子集合. 映射 $G(x, \mu, \lambda, Y)$ 称为 (GVI(F, ϕ, K)) 的 KKT 映射.

注 3.1 如果 $\bar{x}$ 是 (GVI(F, ϕ, K)) 的解且假设 3.1 满足, 则易证 $\Lambda(\bar{x})$ 是非空凸紧集.

映射 $G(x, \mu, \lambda, Y)$ 称为问题 (GVI(F, ϕ, K)) 的 KKT 映射. 显然, 问题 (1.1) 的解集是 $G^{-1}(0)$. 设 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}) \in G^{-1}(0)$. 讨论 G^{-1} 在 $(0, (\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}))$ 的孤立平稳性.

类似于文献 [4, 引理 4.4], 可以建立刻画 G^{-1} 孤立平稳性的下述结果.

引理 3.1 映射 G^{-1} 在 $(0, (\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}))$ 是孤立平稳的当且仅当

$$G'((\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}); (d_x, d_\mu, d_\lambda, d_Y)) = 0 \Rightarrow d = (d_x, d_\mu, d_\lambda, d_Y) = 0. \quad (3.9)$$

对 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}) \in G^{-1}(0)$, 令 $\bar{Z} = B_0 + \mathcal{B}\bar{x}$, 则 $\bar{Y} \in \partial\theta(\bar{Z})$. 为记号简单起见, 记

$$A = \bar{Z}, \quad B = \bar{Y}, \quad C = A + B.$$

下面的分析采用第 2 节中的所有记号.

设

$$\mathcal{G}(x) = \begin{bmatrix} Ax - b \\ g(x) \\ B_0 + \mathcal{B}x \end{bmatrix}. \quad (3.10)$$

定义 3.1 (GVI(F, ϕ, K)) 在 $\bar{x}$ 相对于 $\bar{\chi} = (\bar{\mu}, \bar{\lambda}, \bar{Y})$ 的临界锥定义为

$$\mathcal{C}(\bar{x}, \bar{\chi}) = \{d_x \in \mathbb{R}^n : Ad_x = 0, \mathcal{J}g(\bar{x})d_x \in \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x})) \cap \bar{\lambda}^\perp, \mathcal{B}d_x \in \mathcal{C}_\theta(\bar{Z} + \bar{Y})\}. \quad (3.11)$$

定义 3.2 (GVI(F, ϕ, K)) 在 $\bar{x}$ 相对于 $(\bar{\mu}, \bar{\lambda}, \bar{Y})$ 的严格 Robinson 约束规范定义为

$$D\mathcal{G}(\bar{x})\mathbb{R}^n + \begin{bmatrix} \{0_m\} \\ \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x})) \cap \bar{\lambda}^\perp \\ \mathcal{C}_\theta(\bar{Z} + \bar{Y}) \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathbb{R}^p \\ \mathbb{S}^q \end{bmatrix}. \quad (3.12)$$

定义 3.3 称二阶充分最优性条件在 $\bar{x}$ 相对于 $\bar{\chi} = (\bar{\mu}, \bar{\lambda}, \bar{Y})$ 成立, 若对任意 $d_x \in \mathcal{C}(\bar{x}, \bar{\chi}) \setminus \{0\}$, 有

$$\langle \mathcal{J}L(\bar{x}, \bar{\mu}, \bar{Y})d_x, d_x \rangle + 2\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle > 0, \quad (3.13)$$

其中 $\Upsilon \in \mathbb{S}^q$ 由 (2.5) 定义且 $\lambda_i = \lambda_i(\bar{Z})$, $i \in \alpha \cup \gamma$, $w_i = \lambda_i(\bar{Y})$, $i \in \beta_S$.

定理 3.1 设 $\bar{x}$ 是问题 (GVI(F, ϕ, K)) 的解, 且 $\bar{\chi} = (\bar{\mu}, \bar{\lambda}, \bar{Y}) \in \Lambda(\bar{x})$, 则有下列的结果:

(i) (GVI(F, ϕ, K)) 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的二阶充分最优性条件和严格 Robinson 约束规范成立, 则 G^{-1} 在 $(0, (\bar{x}, \bar{\chi}))$ 是孤立平稳的;

(ii) 如果 G^{-1} 在 $(0, (\bar{x}, \bar{\chi}))$ 是孤立平稳的, 则 (GVI(F, ϕ, K)) 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的严格 Robinson 约束规范成立;

(iii) 如果 G^{-1} 在 $(0, (\bar{x}, \bar{\chi}))$ 是孤立平稳的, 并且二次型

$$\mathcal{Q} : (d_x, d_x) \rightarrow \langle d_x, \mathcal{J}_x L(\bar{x}, \bar{\chi})d_x \rangle + 2\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle$$

满足

$$\mathcal{Q}(d_x, d_x) \geq 0, \quad \forall d_x \in \mathcal{C}(\bar{x}, \bar{\chi}), \quad \mathcal{Q}(d_x, d_x) = 0, \quad d_x \in \mathcal{C}(\bar{x}, \bar{\chi}) \Rightarrow \mathcal{J}_x L(\bar{x}, \bar{\chi})d_x = 0,$$

则 (GVI(F, ϕ, K)) 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的二阶充分最优性条件成立.

证明 首先证明结论 (i) 成立. 设 $(d_x, d_\mu, d_\lambda, d_Y) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 是任意选取的, 满足

$$G'((\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}); (d_x, d_\mu, d_\lambda, d_Y)) = 0.$$

可以把临界锥 $\mathcal{C}(\bar{x}, \bar{\chi})$ 写为

$$\mathcal{C}(\bar{x}, \bar{\chi}) = \mathcal{C}_1(\bar{x}, \bar{\lambda}) \cap \mathcal{C}_2(\bar{x}, \bar{Y}),$$

其中

$$\mathcal{C}_1(\bar{x}, \bar{\lambda}) = \{d_x \in \mathbb{R}^n : Ad_x = 0, \mathcal{J}g(\bar{x})d_x \in \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x})) \cap \bar{\lambda}^\perp\},$$

$$\mathcal{C}_2(\bar{x}, \bar{Y}) = \{d_x \in \mathbb{R}^n : \mathcal{B}d_x \in \mathcal{C}_\theta(\bar{Z} + \bar{Y})\}.$$

由 $G'((\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}); (d_x, d_\mu, d_\lambda, d_Y)) = 0$, 可得

$$\begin{aligned} \mathcal{J}_x L(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})d_x + A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y &= 0, \\ Ad_x &= 0, \\ \mathcal{J}g(\bar{x})d_x - \Pi'_{\mathbb{R}^p_+}(g(\bar{x}) + \bar{\lambda}; \mathcal{J}g(\bar{x})d_x + d_\lambda) &= 0, \\ \mathcal{B}d_x - [\mathcal{P}\theta]'(\bar{Z} + \bar{Y}; \mathcal{B}d_x + d_Y) &= 0. \end{aligned} \quad (3.14)$$

根据文献 [4, 引理 4.2], 由 (3.14) 中的第 3 个方程得

$$\mathcal{J}g(\bar{x})d_x \in \mathcal{T}_{\mathbb{R}^p_+}(g(\bar{x})) \cap \bar{\lambda}^\perp, \quad \langle \mathcal{J}g(\bar{x})d_x, d_\lambda \rangle = 0,$$

上式与 (3.14) 中的第 2 个方程一起蕴涵着 $d_x \in \mathcal{C}_1(\bar{x})$. 由引理 2.1, (3.14) 中的第 4 个方程蕴涵着

$$\mathcal{B}d_x \in \mathcal{C}_\theta(\bar{Z} + \bar{Y}) \quad \text{或} \quad d_x \in \mathcal{C}_2(\bar{x}), \quad \langle \mathcal{B}d_x, d_Y \rangle = 2\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle.$$

因此, $d_x \in \mathcal{C}(\bar{x})$. 通过对 (3.14) 中第一个方程两边与 d_x 作内积, 得

$$\langle d_x, \mathcal{J}_x L(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})d_x \rangle + \langle d_x, A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y \rangle = 0.$$

因此,

$$\langle d_x, \mathcal{J}_x L(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})d_x \rangle + 2\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle = 0.$$

于是, 由 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $(\bar{\mu}, \bar{\lambda}, \bar{Y})$ 的二阶充分最优性条件得 $d_x = 0$. 因此, (3.14) 简化为

$$\begin{aligned} \text{DG}(\bar{x})^*(d_\mu, d_\lambda, d_Y) &= 0, \\ \Pi'_{\mathbb{R}^p_+}(g(\bar{x}) + \bar{\lambda}; d_\lambda) &= 0, \\ [\mathcal{P}\theta]'(\bar{Z} + \bar{Y}; d_Y) &= 0. \end{aligned}$$

根据引理 2.5, 有

$$(d_\mu, d_\lambda, d_Y) \in \left[\text{DG}(\bar{x})\mathbb{R}^n + \begin{pmatrix} \{0_m\} \\ \mathcal{T}_{\mathbb{R}^p_+}(g(\bar{x})) \cap \bar{\lambda}^\perp \\ \mathcal{C}_\theta(\bar{Z} + \bar{Y}) \end{pmatrix} \right]^\circ.$$

由所假设的 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的严格 Robinson 约束规范成立知, 上式蕴涵着

$$(d_\mu, d_\lambda, d_Y) = 0.$$

因此 $(d_x, d_\mu, d_\lambda, d_Y) = 0$. 从而由引理 3.1 知 G^{-1} 在 $(0, (\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}))$ 是孤立平稳的. 因此, 由引理 3.1 得 $(\hat{\mu}, \hat{\lambda}, \hat{Y}) = 0$. 这个矛盾表明结论 (ii) 成立.

最后, 用反证法证明结论 (iii) 成立. 假设 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的二阶充分最优性条件不成立, 由于 G^{-1} 在原点关于 $(\bar{x}, \bar{\chi})$ 是孤立平稳的, 所以有 $\Lambda(\bar{x}) = \{\bar{\chi}\}$. 因此, 存在向量 $0 \neq d_x \in \mathcal{C}(\bar{x}, \bar{\chi})$ 满足 $\mathcal{Q}(d_x, d_x) = 0$. 由 (iii) 中的条件知 $\mathcal{J}_x L(\bar{x}, \bar{\chi})d_x = 0$, 因此 $\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle = 0$. 这蕴涵着

$$P_\alpha^T(\mathcal{B}d_x)P_{\beta_S \cup \beta_L \cup \gamma} = 0, \quad P_\gamma^T(\mathcal{B}d_x)P_{\beta_U \cup \beta_S} = 0.$$

又由 $C_1(\bar{x}, \bar{\lambda})$ 的定义、 $C_2(\bar{x}, \bar{Y})$ 的定义和文献 [9, 引理 4.2 和 2.1], 有

$$\begin{aligned} \mathcal{B}d_x &= 0, \\ \mathcal{J}g(\bar{x})d_x - \Pi'_{\mathbb{R}^p}(g(\bar{x}) + \bar{\lambda}; \mathcal{J}g(\bar{x})d_x) &= 0, \\ \mathcal{B}d_x - [\mathcal{P}\theta]'(\bar{Z} + \bar{Y}; \mathcal{B}d_x) &= 0. \end{aligned} \quad (3.15)$$

通过利用 $\mathcal{J}_x L(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})d_x = 0$ 、(3.15) 和 G 在 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})$ 的方向导数的表达, 得

$$G'((\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y}); (d_x, 0, 0, 0)) = 0 \quad \text{且} \quad d_x \neq 0.$$

于是, 由引理 3.1 知, 上式与 G^{-1} 在点关于 $(\bar{\mu}, \bar{\lambda}, \bar{Y})$ 的孤立平稳性相矛盾. 因此, 必有

$$\mathcal{Q}(d_x, d_x) > 0, \quad d_x \in \mathcal{C}(\bar{x}, \bar{\chi}) \setminus \{0\},$$

即 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $\bar{\chi}$ 的二阶充分最优性条件成立. 证毕. $\square$

由定理 3.1, 对问题 $(\text{GVI}(F, \phi, K))$, 当 F 是单调映射时, 可以得到 G^{-1} 的孤立平稳性的下面的完整刻画.

推论 3.1 设 F 是定义在 $\mathbb{R}^n$ 上的连续可微单调映射, $\bar{x}$ 是问题 $(\text{GVI}(F, \phi, K))$ 的解且 $\Lambda(\bar{x}) \neq \emptyset$, 则 G^{-1} 在点关于 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})$ 是孤立平稳的, 且 $(\bar{\mu}, \bar{\lambda}, \bar{Y}) \in \Lambda(\bar{x})$ 当且仅当 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $(\bar{\mu}, \bar{\lambda}, \bar{Y})$ 的二阶充分最优性条件和严格 Robinson 约束规范都成立.

证明 注意到, 如果 F 是 $\mathbb{R}^n$ 上的连续可微单调映射, 则 $\mathcal{J}F(x)$ 是半正定的. 结合 $\langle \mathcal{B}d_x, \Upsilon \circ \mathcal{B}d_x \rangle$ 是非负的事实, 我们得定理 3.1(iii) 中关于 $\mathcal{Q}(d_x, d_x)$ 的条件满足. 因此, 由定理 3.1 可得这个推论的结果成立. 证毕. $\square$

4 GVI 的强正则性

本节主要讨论广义变分不等式问题的 KKT 系统的强正则性. 为定义强二阶充分最优性条件, 需要刻画临界锥 $\mathcal{C}(\bar{x}, \bar{\chi})$ 的仿射包. 这是不容易的, 因为 $C_2(\bar{x}, \bar{Y})$ 不是多面体锥. 定义一个替代的锥如下:

$$\text{app}\mathcal{C}(\bar{x}, \bar{\chi}) = \{d_x \in \mathbb{R}^n : d_x \in \text{aff } C_1(\bar{x}, \bar{\lambda}), \mathcal{B}d_x \in \text{aff } C_\theta(\bar{Z} + \bar{Y})\} \quad (4.1)$$

(其中 aff 表示集合的仿射包, 即包含集合的最小的仿射集合), 即

$$\text{app}\mathcal{C}(\bar{x}, \bar{\chi}) = \left\{ d_x \in \mathbb{R}^n : \begin{aligned} &Ad_x = 0, \nabla g_i(\bar{x})^T d_x = 0, i \in I_+(\bar{x}, \bar{\lambda}), \\ &P_{\beta_S}^T(\mathcal{B}d_x)[P_\beta] = 0, P_{\beta_U}^T(\mathcal{B}d_x)[P_{\beta_L}] = 0 \end{aligned} \right\}, \quad (4.2)$$

其中

$$I_+(\bar{x}, \bar{\lambda}) = \{i : g_i(\bar{x}) = 0, \bar{\lambda}_i > 0, i = 1, \dots, p\}.$$

定义 4.1 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 的约束非退化约束规范定义为

$$\text{DG}(\bar{x})\mathbb{R}^n + \text{lin} \begin{bmatrix} \{0_m\} \\ \mathcal{T}_{\mathbb{R}^p}(g(\bar{x})) \\ \mathcal{C}_\theta(\bar{Z}) \end{bmatrix} = \begin{bmatrix} \mathbb{R}^m \\ \mathbb{R}^p \\ \mathbb{S}^q \end{bmatrix}, \quad (4.3)$$

其中

$$\lim C_\theta(\bar{Z}) = \{\Delta A \in \mathbb{S}^q : P_\beta^T \Delta A P_\beta = 0\}. \quad (4.4)$$

命题 4.1 设 $\bar{x}$ 是 $(\text{GVI}(F, \phi, K))$ 的解, 且在该点约束非退化约束规范成立, 则 $\Lambda(\bar{x})$ 是单点集.

证明 显然, 条件 (4.3) 蕴涵着 K 在 $\bar{x}$ 的 Robinson 约束规范成立, 因此 $\Lambda(\bar{x}) \neq \emptyset$. 假设有两个乘子 $\xi_1 = (\mu_1, \lambda_1, Y_1) \in \Lambda(\bar{x})$ 和 $\xi_2 = (\mu_2, \lambda_2, Y_2) \in \Lambda(\bar{x})$ 且 $\xi_1 \neq \xi_2$. 注意到 $Y_i \in \partial\theta(\bar{Z})$, 由表达式

$$\partial\theta(\bar{Z}) = \{P_\alpha P_\alpha^T + P_\beta W_\beta P_\beta^T - P_\gamma P_\gamma^T : W_\beta \in \mathbb{S}^{|\beta|}, \|W_\beta\|_2 \leq 1\},$$

得 $Y_1 - Y_2 \in \lim C_\theta(\bar{Z})^\perp$. 易验证 $\lambda_1 - \lambda_2 \in [\lim \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x}))]^\perp$. 因此,

$$\xi_1 - \xi_2 \in \text{Ker DG}(\bar{x})^* \cap \left\{ \lim \begin{bmatrix} \{0_m\} \\ \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x})) \\ \mathcal{C}_\theta(\bar{Z}) \end{bmatrix} \right\}^\perp,$$

这与 (4.3) 矛盾, 因为假设了 $\xi_1 \neq \xi_2$ 成立. 证毕. $\square$

定义 4.2 称强二阶充分最优性条件在 $\bar{x}$ 相对于 $\bar{\chi} = (\bar{\mu}, \bar{\lambda}, \bar{Y}) \in \Lambda(\bar{x})$ 成立, 如果对任意的 $d_x \in \text{app } C(\bar{x}, \bar{\chi}) \setminus \{0\}$, 有

$$\langle \mathcal{J}_x L(\bar{x}, \bar{\mu}, \bar{Y}) d_x, d_x \rangle + 2 \langle \mathcal{B} d_x, \Upsilon \circ \mathcal{B} d_x \rangle > 0, \quad (4.5)$$

其中 $\Upsilon \in \mathbb{S}^q$ 由 (2.5) 定义, 且对 $i \in \alpha \cup \gamma$ 有 $\lambda_i = \lambda_i(\bar{Z})$, 对 $i \in \beta_S$ 有

$$w_i = \lambda_i(\bar{Y}).$$

把系统 (3.7) 表示为下面的广义方程的形式:

$$0 \in \mathcal{F}(x, \mu, \lambda, Y) + \mathcal{N}(x, \mu, \lambda, Y), \quad (4.6)$$

其中

$$\begin{aligned} \mathcal{F}(x, \mu, \lambda, Y) &= \begin{bmatrix} L(x, \mu, \lambda, Y) \\ Ax - b \\ -g(x) \\ -Y \end{bmatrix}, \\ \mathcal{N}(x, \mu, \lambda, Y) &= \begin{bmatrix} 0 \\ 0 \\ \mathcal{N}_{\mathbb{R}_+^p}(\lambda) \\ \partial\theta(B_0 + \sum_{i=1}^n x_i B_i) \end{bmatrix}. \end{aligned} \quad (4.7)$$

现在描述并证明系统 (4.6) 的强正则性的充分条件.

定理 4.1 设 $\bar{x}$ 是问题 $(\text{GVI}(F, \phi, K))$ 的解且 $\Lambda(\bar{x}) \neq \emptyset$. 考虑下面 3 个性质:

(i) $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 的约束非退化约束规范和 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $\bar{\chi} \in \Lambda(\bar{x})$ 的强二阶充分最优性条件都成立;

(ii) 任意的 $V \in \partial G(\bar{x}, \bar{\xi})$ 是非奇异的;

(iii) 广义方程 (4.6) 在 $(\bar{x}, \bar{\xi})$ 是强正则的,

则下面的蕴涵关系成立:

$$(i) \Rightarrow (ii) \Rightarrow (iii).$$

证明 先证 (i) $\Rightarrow$ (ii). 假设 (i) 中的条件成立. 对任意的 $V \in \partial G(\bar{x}, \bar{\xi})$, 只需证 $Vd = 0$ 蕴涵 $d = 0$. 考虑 $Vd = 0$, 则存在 $\text{Diag}(v) \in \partial \Pi_{\mathbb{R}^p} g(\bar{x} + \bar{\lambda})$ 和 $W \in \partial[\mathcal{P}\theta](\bar{Z} + \bar{Y})$ 满足

$$Vd = \begin{bmatrix} \mathcal{J}_x L(\bar{x}, \bar{\xi})d_x + A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y \\ Ad_x \\ \mathcal{J}g(\bar{x})d_x - \text{Diag}(v)(\mathcal{J}g(\bar{x})d_x + d_\lambda) \\ \mathcal{B}d_x - W(\mathcal{B}d_x + d_Y) \end{bmatrix} = 0. \quad (4.8)$$

由 (4.8) 中的第 2 个方程, 可得

$$\begin{cases} [d_\lambda]_i = 0, & i \in I_-, \\ (1 - v_i)\nabla g_i(\bar{x})^T d_x - v_i[d_\lambda]_i = 0, & i \in I_0, \\ \nabla g_i(\bar{x})^T d_x = 0, & i \in I_+, \end{cases} \quad (4.9)$$

其中

$$I_- = \{i : g_i(\bar{x}) < 0, \bar{\lambda}_i = 0\}, \quad I_0 = \{i : g_i(\bar{x}) = \bar{\lambda}_i = 0\}, \quad I_+ = \{i : g_i(\bar{x}) = 0, \bar{\lambda}_i > 0\},$$

$v_i \in [0, 1]$, $i \in I_0$. 显然有

$$\langle d_x, \mathcal{J}g(\bar{x})^T d_\lambda \rangle = \sum_{i \in I_0} [d_\lambda]_i \nabla g_i(\bar{x})^T d_x \geq 0. \quad (4.10)$$

对 $W \in \partial[\mathcal{P}\theta](\bar{Z} + \bar{Y})$, 存在 $W_{\beta_U} \in \partial \Pi_{\mathbb{S}_+^{|\beta_U|}}(0)$ 和 $W_{\beta_L} \in \partial \Pi_{\mathbb{S}_-^{|\beta_L|}}(0)$ 使得 (4.8) 中的第 4 个方程变为

$$P^T(\mathcal{B}d_x)P = P \begin{bmatrix} \hat{H}_{\alpha\alpha} & \hat{H}_{\alpha\beta_U} & \hat{H}_{\alpha\beta_S} \circ \Omega_{\alpha\beta_S} & \hat{H}_{\alpha\beta_L} \circ \Omega_{\alpha\beta_L} & \hat{H}_{\alpha\gamma} \circ \Omega_{\alpha\gamma} \\ \hat{H}_{\beta_U\alpha} & W_{\beta_U}(\hat{H}_{\beta_U\beta_U}) & 0 & 0 & \hat{H}_{\beta_U\gamma} \circ \Omega_{\beta_U\gamma} \\ \hat{H}_{\beta_S\alpha} \circ \Omega_{\beta_S\alpha} & 0 & 0 & 0 & \hat{H}_{\beta_S\gamma} \circ \Omega_{\beta_S\gamma} \\ \hat{H}_{\beta_L\alpha} \circ \Omega_{\beta_L\alpha} & 0 & 0 & W_{\beta_L}(\hat{H}_{\beta_L\beta_L}) & \hat{H}_{\beta_L\gamma} \\ \hat{H}_{\gamma\alpha} \circ \Omega_{\gamma\alpha} & \hat{H}_{\gamma\beta_U} \circ \Omega_{\gamma\beta_U} & \hat{H}_{\gamma\beta_S} \circ \Omega_{\gamma\beta_S} & \hat{H}_{\gamma\beta_L} & \hat{H}_{\gamma\gamma} \end{bmatrix} P^T, \quad (4.11)$$

其中 $\hat{H} = P^T(\mathcal{B}d_x + d_Y)P$,

$$\Omega_{ij} = [p_1^{[1]}(\Lambda(\bar{Z} + \bar{Y}))]_{ij}, \quad (i, j) \in \alpha \times [\beta_S \cup \beta_L \cup \gamma] \quad \text{或} \quad (i, j) \in \gamma \times [\beta_U \cup \beta_S].$$

由 (4.11) 得

$$\begin{aligned} P_\alpha^T(d_Y)[P_\alpha \ P_{\beta_U}] &= 0, \\ P_\gamma^T(d_Y)[P_\gamma \ P_{\beta_L}] &= 0, \\ P_{\beta_S}^T(\mathcal{B}d_x)[P_\beta] &= 0, \quad P_{\beta_U}^T(\mathcal{B}d_x)[P_{\beta_L}] = 0, \\ P_{\beta_U}^T(\mathcal{B}d_x)P_{\beta_U} &= W_{\beta_U}(P_{\beta_U}^T(\mathcal{B}d_x + d_Y)P_{\beta_U}), \\ P_{\beta_L}^T(\mathcal{B}d_x)P_{\beta_L} &= W_{\beta_L}(P_{\beta_L}^T(\mathcal{B}d_x + d_Y)P_{\beta_L}), \\ P_\alpha^T(d_Y)P_{\beta_S \cup \beta_L \cup \gamma} &= P_\alpha^T(\mathcal{B}d_x)P_{\beta_S \cup \beta_L \cup \gamma} \circ \Upsilon_{\alpha, \beta_S \cup \beta_L \cup \gamma}, \\ P_\gamma^T(d_Y)P_{\beta_U \cup \beta_S} &= P_\gamma^T(\mathcal{B}d_x)P_{\beta_U \cup \beta_S} \circ \Upsilon_{\gamma, \beta_U \cup \beta_S}. \end{aligned} \quad (4.12)$$

注意到, 对任意 $d_{\beta_U} \in \mathbb{R}^{|\beta_U|}$ 和 $d_{\beta_L} \in \mathbb{R}^{|\beta_L|}$, 有

$$\|W_{\beta_U}d_{\beta_U}\|^2 \leq \langle W_{\beta_U}d_{\beta_U}, d_{\beta_U} \rangle, \quad \|W_{\beta_L}d_{\beta_L}\|^2 \leq \langle W_{\beta_L}d_{\beta_L}, d_{\beta_L} \rangle,$$

从 (4.12) 中的第 4 和 5 个方程可得

$$\langle P_{\beta_U}^T(\mathcal{B}d_x)P_{\beta_U}, P_{\beta_U}^T(d_Y)P_{\beta_U} \rangle \geq 0, \quad \langle P_{\beta_L}^T(\mathcal{B}d_x)P_{\beta_L}, P_{\beta_L}^T(d_Y)P_{\beta_L} \rangle \geq 0.$$

因此, (4.12) 中的关系蕴涵着

$$\begin{aligned} \langle \mathcal{B}d_x, d_Y \rangle &= \langle P_{\beta_U}^T(\mathcal{B}d_x)P_{\beta_U}, P_{\beta_U}^T(d_Y)P_{\beta_U} \rangle + \langle P_{\beta_L}^T(\mathcal{B}d_x)P_{\beta_L}, P_{\beta_L}^T(d_Y)P_{\beta_L} \rangle \\ &\quad + 2\langle P^T[\mathcal{B}d_x]P, \Upsilon \circ P^T[\mathcal{B}d_x]P \rangle \\ &\geq 2\langle P^T[\mathcal{B}d_x]P, \Upsilon \circ P^T[\mathcal{B}d_x]P \rangle. \end{aligned} \quad (4.13)$$

由 (4.2) 中 $\text{app}\mathcal{C}(\bar{x}, \bar{\xi})$ 的定义, 并由 (4.9) 和 (4.12) 得 $d_x \in \text{app}\mathcal{C}(\bar{x}, \bar{\xi})$. 通过 d_x 与 (4.8) 中的第一个等式两边同时取内积, 得 $Ad_x = 0$, 并由 (4.10) 和 (4.13) 得

$$\begin{aligned} 0 &= \langle d_x, [\mathcal{J}_x L(\bar{x}, \bar{\xi})d_x + A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y] \rangle \\ &\geq \langle d_x, \mathcal{J}_x L(\bar{x}, \bar{\xi})d_x \rangle + 2\langle P^T[\mathcal{B}d_x]P, \Upsilon \circ P^T[\mathcal{B}d_x]P \rangle. \end{aligned}$$

因此, 由于 $(\text{GVI}(F, \phi, K))$ 在 $\bar{x}$ 相对于 $\bar{\chi} \in \Lambda(\bar{x})$ 的强二阶充分最优性条件成立, 因此, $d_x = 0$. 由于 $d_x = 0$, 由 (4.8) 得

$$\begin{bmatrix} A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y \\ \text{Diag}(v)d_\lambda \\ W(d_Y) \end{bmatrix} = 0. \quad (4.14)$$

由 (4.3) 知, 存在 $\Delta x \in \mathbb{R}^n$ 、向量 $u \in \text{lin } \mathcal{T}_{\mathbb{R}^p_-}(g(\bar{x}))$ 和 $S \in \mathcal{T}_\theta(\bar{Z})$ 满足

$$\mathcal{J}h(\bar{x})\Delta x = d_\mu, \quad \mathcal{J}g(\bar{x})\Delta x + u = d_\lambda, \quad \mathcal{B}\Delta x + S = d_Y.$$

于是,

$$\begin{aligned}
& \langle d_\mu, d_\mu \rangle + \langle d_\lambda, d_\lambda \rangle + \langle d_Y, d_Y \rangle \\
&= \langle \mathcal{J}h(\bar{x})\Delta x, d_\mu \rangle + \langle \mathcal{J}g(\bar{x})\Delta x + u, d_\lambda \rangle + \langle \mathcal{B}\Delta x + S, d_Y \rangle \\
&= \langle \Delta x, A^T d_\mu + \mathcal{J}g(\bar{x})^T d_\lambda + \mathcal{B}^* d_Y \rangle + \langle u, d_\lambda \rangle + \langle S, d_Y \rangle.
\end{aligned}$$

由 $\text{Diag}(v)d_\lambda = 0$ 得

$$[d_\lambda]_{I_-} = 0,$$

由 $\lim_{\mathbb{T}_{\mathbb{R}^p}}(g(\bar{x}))$ 的定义得 $u_{I_0 \cup I_+} = 0$, 因此有 $\langle u, d_\lambda \rangle = 0$. 对 $(\omega, \omega') \neq (\beta, \beta)$, 由 $W(d_Y) = 0$ 得

$$P_\omega^T(d_Y)P_{\omega'} = 0.$$

又由 $\mathcal{T}_\theta(\bar{Z})$ 的定义得 $P_\beta^T S P_\beta = 0$. 因此 $\langle S, d_Y \rangle = 0$. 从而有 $\|(d_\mu; d_\lambda; d_Y)\| = 0$. 结合 $d_x = 0$, 可得 V 是非奇异的.

(ii) $\Rightarrow$ (iii). 现在证明, 在 (ii) 的条件下, (iii) 中的结论成立, 这意味着对 $\delta = (\delta_1; \delta_2; \delta_3; \delta_4) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 趋于 $0 \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$,

$$\delta \in \mathcal{F}(x, \mu, \lambda, Y) + \mathcal{N}(x, \mu, \lambda, Y) \quad (4.15)$$

在 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})$ 附近有单值解 $(x(\delta), \mu(\delta), \lambda(\delta), Y(\delta))$, 且在 $0 \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 附近是局部 Lipschitz 连续的. 方程 (4.15) 等价于

$$\Psi(x, \mu, \lambda, Y, \delta) := G(x, \mu, \lambda + \delta_3, Y + \delta_4) + \begin{bmatrix} -\delta_1 - \mathcal{J}g(x)^T \delta_3 - \mathcal{B}^* \delta_4 \\ -\delta_2 \\ \delta_3 \\ 0 \end{bmatrix} = 0.$$

由于 $\Psi(\bar{x}, \bar{\xi}, 0) = 0$, 而且 $\partial_{x, \mu, \lambda, Y} \Psi(\bar{x}, \bar{\xi}, 0) = \partial G(\bar{x}, \bar{\xi})$ 中的任意元素都是非奇异的, 因此, 由文献 [28] 中的 Clarke 隐函数定理知 $\Psi(x, \mu, \lambda, Y, \delta) = 0$ 决定着在 $(\bar{x}, \bar{\mu}, \bar{\lambda}, \bar{Y})$ 附近有单值解 $(x(\delta), \mu(\delta), \lambda(\delta), Y(\delta))$, 它在 $0 \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{S}^q$ 附近是局部 Lipschitz 连续的. 证毕. $\square$

5 结论

本文研究了一类广义变分不等式问题的稳定性, 这类变分不等式问题在最优化及其他领域有广泛应用. 本文主要研究两个扰动性质, 一个是 KKT 映射的逆映射的孤立平稳性, 另一个是 KKT 系统的强正则性. 受 Han 等 [4] 和 Sun [14] 的工作的启发, 本文提出了严格 Robinson 约束规范 (SRCQ) 和约束非退化条件、二阶充分最优性条件 (second-order sufficient optimality condition, SOSOC) 和强二阶充分最优性条件 (strong second-order sufficient optimality condition, SSOSOC), 并且证明了, 如果 SOSOC 和 SRCQ 成立, 那么 KKT 映射的逆映射是孤立平稳的; 如果 SSOSOC 和约束非退化条件成立, 则 KKT 系统是强正则的. 然而, 当 F 不是光滑函数的梯度时, 对这类广义变分不等式还有一些未解决的稳定性课题. 一个是 SOSOC 和 SRCQ 对孤立平稳性而言是否是必要条件; 类似地, 另一个问

题是 SSOSOC 和约束非退化条件对强正则性而言是否是必要条件. 具有矩阵映射或矩阵变量的其他变分不等式问题的稳定性也值得研究.

致谢 作者非常感谢审稿人的许多有益建议, 这些建议帮助作者改进了论文的质量. 特别感谢其中一位审稿人提供了重要参考文献 [19], 该文献给出了核函数的变分几何的完整刻画.

参考文献

- 1 Rockafellar R T. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Math Oper Res*, 1976, 1: 97–116
- 2 Han D R, Yuan X M. Local linear convergence of the alternating direction method of multipliers for quadratic programs. *SIAM J Numer Anal*, 2013, 51: 3446–3457
- 3 Hong M, Luo Z Q. On the linear convergence of the alternating direction method of multipliers. *Math Program*, 2017, 162: 165–199
- 4 Han D R, Sun D F, Zhang L W. Linear rate convergence of the alternating direction method of multipliers for convex composite programming. *Math Oper Res*, 2018, 43: 622–637
- 5 Robinson S M. Strongly regular generalized equations. *Math Oper Res*, 1980, 5: 43–62
- 6 Jongen H Th, Klatte D, Tammer K. Implicit functions and sensitivity of stationary points. *Math Program*, 1990, 49: 123–138
- 7 Dontchev A L, Rockafellar R T. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J Optim*, 1996, 6: 1087–1105
- 8 Robinson M. Generalized equations and their solutions, part II: Applications to nonlinear programming. In: *Optimality and Stability in Mathematical Programming*. Mathematical Programming Studies, vol. 19. Berlin-Heidelberg: Springer, 1982, 19: 200–221
- 9 Dontchev A L, Rockafellar R T. Characterizations of Lipschitzian stability in nonlinear programming. In: *Mathematical Programming with Data Perturbations*. New York: Marcel Dekker, 1997, 65–82
- 10 Klatte D. Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ programs. *Math Program*, 2000, 88: 285–311
- 11 Izmailov A F, Pogosyan A L, Solodov M V. Semismooth Newton method for the lifted reformulation of mathematical programs with complementarity constraints. *Comput Optim Appl*, 2012, 51: 199–221
- 12 Bonnans J F, Ramírez C H. Perturbation analysis of second-order cone programming problems. *Math Program*, 2005, 104: 205–227
- 13 Wang Y, Zhang L W. Properties of equation reformulation of the Karush-Kuhn-Tucker condition for nonlinear second order cone optimization problems. *Math Methods Oper Res*, 2009, 70: 195–218
- 14 Sun D F. The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. *Math Oper Res*, 2006, 31: 761–776
- 15 Chan Z X, Sun D F. Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. *SIAM J Optim*, 2008, 19: 370–396
- 16 Zhang Y L, Zhang L W. On the upper Lipschitz property of the KKT mapping for nonlinear semidefinite optimization. *Oper Res Lett*, 2016, 44: 474–478
- 17 Liu Y L, Pan S H. Locally upper Lipschitz of the perturbed KKT system of Ky Fan k -norm matrix conic optimization problems. *arXiv:1509.00681*, 2015
- 18 Ding C, Sun D F, Zhang L W. Characterization of the robust isolated calmness for a class of conic programming problems. *SIAM J Optim*, 2017, 27: 67–90
- 19 Cui Y, Sun D F. A complete characterization of the robust isolated calmness of nuclear norm regularized convex optimization problems. *J Comput Math*, 2018, 36: 441–458
- 20 King A J, Rockafellar R T. Sensitivity analysis for nonsmooth generalized equations. *Math Program*, 1992, 55: 193–212
- 21 Levy A B. Implicit multifunction theorems for the sensitivity analysis of variational conditions. *Math Program*, 1996, 74: 333–350
- 22 Borwein J M, Lewis A S. *Convex Analysis and Nonlinear Optimization*. New York: Springer, 2006
- 23 Watson G A. Characterization of the subdifferential of some matrix norms. *Linear Algebra Appl*, 1992, 170: 33–45
- 24 Lewis A S. The convex analysis of unitarily invariant matrix functions. *J Convex Anal*, 1995, 2: 173–183
- 25 Wu B, Ding C, Sun D F, et al. On the Moreau-Yosida regularization of the vector k -norm related functions. *SIAM J Optim*, 2014, 24: 766–794
- 26 Ding C. An introduction to a class of matrix optimization problems. PhD Thesis. Singapore: National University of Singapore, 2012

- 27 Fan K. Maximum properties and inequalities for the eigenvalues of completely continuous operators. Proc Natl Acad Sci USA, 1951, 37: 760–766
- 28 Clarke F H. Optimization and Nonsmooth Analysis. New York: John Wiley & Sons, 1983

The stability of a generalized variational inequality problem with the nuclear norm function

Yali Zhao, Yule Zhang & Liwei Zhang

Abstract This paper is to study the stability of the generalized variational inequality (GVI) problem whose convex function is the nuclear norm of a linear matrix mapping. The strict Robinson constraint qualification and the second-order sufficient optimality conditions for the GVI problem are proposed and are demonstrated as sufficient conditions for the isolated calmness of the inverse of the KKT (Karush-Kuhn-Tucker) mapping. The constraint non-degeneracy condition and the strong second-order sufficient optimality conditions for the GVI problem are proposed and are proven to be sufficient conditions for the strong regularity of the KKT system.

Keywords generalized variational inequality, nuclear matrix norm, stability, isolated calmness, strong regularity, strict Robinson constraint qualification

MSC(2020) 90C30

doi: 10.1360/SCM-2018-0188