运筹学基础与发展论坛

DOOR#1 变分分析—基础理论与前沿进展

主办单位:
中国运筹学会数学规划分会
南方科技大学数学系

稳定性的变分准则

张立卫

2021年11月8日–11日

目录

素材基于

- Bonnans J. F. and Shapiro A., *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, 2000.
- Rockafellar R.T., *Convex Analysis*, Princeton University Press, 1970.
- Rockafellar R.T. and Wets R.J.-B, *Variational Analysis*, Springer-Verlag, New York, 1998.
- 张立卫,殷子然,最优化问题的稳定性分析,科学出版社,2020.

稳定性概念

定义 1.1

设U, Y是度量空间, 对于集值映射$S: U \rightrightarrows Y$,

(1) 称集值映射S在$u_0 \in U$处是Berge上半连续的, 若对每一个满足$S(u_0) \subset \mathcal{O}$的开集合$\mathcal{O}$, 存在$\delta > 0$, 满足

$$S(u) \subset \mathcal{O}, \forall u \in \mathbb{B}_\delta(u_0).$$

(2) 称集值映射S在$u_0 \in X$处是Berge下半连续的, 若对每一个满足$S(u_0) \cap \mathcal{O} \neq \emptyset$的开集合$\mathcal{O}$, 存在$\delta > 0$, 满足

$$S(u) \cap \mathcal{O} \neq \emptyset, \forall u \in \mathbb{B}_\delta(u_0).$$

Hausdorff半连续性

定义 1.2

设U, Y是度量空间, 对于集值映射$S : U \rightrightarrows Y$,

(1) 称集值映射S在$u_0 \in U$处是*Hausdorff* 意义下上半连续的, 若任意$\varepsilon > 0$, $\exists \delta > 0$, 满足

$$S(u) \subset S(u_0) + \varepsilon \mathbb{B}, \forall u \in \mathbb{B}_\delta(u_0).$$

(2) 称集值映射S在$u_0 \in U$处是*Hausdorff* 意义下下半连续的, 若任意 $\varepsilon > 0$, $\exists \delta > 0$, 满足

$$S(u_0) \subset S(u) + \varepsilon \mathbb{B}, \forall u \in \mathbb{B}_\delta(u_0).$$

一个例子

设$F:\Re^n\to\Re^m$在$x\in\Re^n$处局部Lipschitz连续, $\partial F(x)$是Clarke广义Jacobian

$$\partial F(x)=\mathrm{conv}\{\lim \mathcal{J}F(x_i): x_i\to x, x_i\notin\Omega_F\}.$$

由[6, 2.6.2 Proposition][1],

(a) $\partial F(x)$是$\Re^{m\times n}$的一非空闭紧致集合.

(b) ∂F在x处是闭的;即$x_t\to x$, $Z_t\in\partial F(x_t)$, $Z_t\to Z$, 则$Z\in\partial F(x)$.

(c) ∂F在x处是上半连续的:$\forall\epsilon>0$, $\exists\delta>0$满足

$$\partial F(y)\subset\partial F(x)+\epsilon\mathbf{B}_{m\times n},\ \forall y\in x+\delta\mathbf{B}.$$

[1] Clarke F H. *Optimization and Nonsmooth Analysis*. New York: John Wiley and Sons, 1983.

Berge与Huasdorff半连续性的关系

Berge意义下的连续性和 Hausdorff意义下的连续性有如下的关系,见[1].[2]

(a) 如果集值映射S在u_0处Hausdorff意义下上半连续且$S(u_0)$是紧致的, 则S在u_0处Berge意义下上半连续.

(b) 如果集值映射S在u_0处Berge 意义下上半连续, 则S在u_0处Hausdorff意义下上半连续.

(c) 如果集值映射S在u_0处Hausdorff意义下下半连续, 则S在u_0处Berge意义下下半连续.

(d) 如果集值映射S在u_0处Berge 意义下下半连续且$\mathrm{cl}S(u_0)$是紧致的, 则S在Hausdorff意义下是下半连续的.

[2] Bank B, Guddat J, Klatte D, Kummer B and Tammer K. *Nonlinear Parametric Optimization*. Berlin: Akademie-Verlag, 1982.

与外半连续性的关系

在局部有界的条件下, 集合的外半连续性等价于Berge意义下的上半连续的[29, 定理 5.19].[3] 对闭值的集值映射, Hausdorff意义下上半连续等价于外半连续性(闭性)[29, 命题5.12(a)].

[3]Rockafellar R T and Wets R J B. *Variational Analysis*. New York: Springer-Verlag, 1998.

定义 1.3

集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$(x^0,y^0)\in\operatorname{gph}F$处是以率$\kappa$为度量正则的，如果存在邻域$U\in\mathcal{N}(x^0)$, $V\in\mathcal{N}(y^0)$和常数$\kappa>0$，满足

$$d(x,F^{-1}(y))\leq\kappa d(y,F(x)),\ \forall(x,y)\in U\times V.$$

定义 1.4

集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$对$y^0$是度量次正则的，如果$(x^0,y^0)\in\operatorname{gph}F$，存在邻域$U\in\mathcal{N}(x^0)$, $V\in\mathcal{N}(y^0)$和常数$\kappa>0$，满足

$$d(x,F^{-1}(y^0))\leq\kappa d(y^0,F(x)\cap V),\ \forall x\in U.$$

一个例子

考虑约束集合

$$\Phi = \{x \in \mathcal{X} : G(x) \in K\},$$

其中$\mathcal{X}$与$\mathcal{Y}$是有限维Hilbert空间,$K \subset \mathcal{Y}$是非空闭凸集合. 定义

$$\mathcal{F}_G(x) = K - G(x),$$

则$\Phi = \mathcal{F}_G^{-1}(0)$. 设$x_0 \in \Phi$, 则 $\mathcal{F}_G$在$(x_0, 0)$处满足度量正则性意味着

$$\mathrm{dist}\,(x, \mathcal{F}_G^{-1}(y)) \leq \kappa \mathrm{dist}\,(y, \mathcal{F}_G(x)) = \kappa \mathrm{dist}\,(G(x) + y, K)$$

对$(x, y) \in \mathbf{B}_\delta(x_0) \times \delta\mathbf{B}$成立.

定义 1.5

(*Aubin*性质和图模). 称集值映射$S : \mathcal{X} \rightrightarrows \mathcal{Y}$相对于$X$在$x^0$点关于$u^0$具有*Aubin*性质, 其中$x^0 \in X$, $u^0 \in S(x^0)$, 若$\mathrm{gph}S$在(x^0, u^0)点处是局部闭的, 且存在邻域$V \in \mathcal{N}(x^0)$, $W \in \mathcal{N}(u^0)$ 和常数$\kappa \in \Re_+$, 满足

$$S(x') \cap W \subset S(x) + \kappa \|x' - x\| \mathbf{B}, \ \forall x, x' \in X \cap V. \tag{1}$$

若将上述条件中的$X \cap V$替换为V, 则称*Aubin*性质在x^0点关于u^0成立. 此时, S在x^0点关于u^0的图模*(graphical modulus)*为

$$\mathrm{lip}S(x^0|u^0) := \inf\{\kappa : \exists V \in \mathcal{N}(x^0), W \in \mathcal{N}(u^0), \text{ 满足 } S(x') \cap W \subset S(x) + \kappa \|x' - x\| \mathbf{B}, \ \forall x, x' \in V\}.$$

定义 1.6

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处关于$y^0$是平稳的(*calm*), 如果$y^0\in F(x^0)$, 存在一常数$\kappa>0$, x_0 的一邻域V 和 y^0 的一邻域W满足

$$F(x)\cap W\subseteq F(x^0)+\kappa\|x-x^0\|\mathbf{B},\quad \forall x\in V.$$

定义 1.7

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处关于$y^0$是稳健平稳的(*robustly calm*), 如果$y^0\in F(x^0)$, 存在一常数$\kappa>0$, x_0 的一邻域V 和 y^0 的一邻域W满足

$$\emptyset\neq F(x)\cap W\subseteq F(x^0)+\kappa\|x-x^0\|\mathbf{B},\quad \forall x\in V.$$

定义 1.8

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处关于$y^0$是孤立平稳的*(isolated calm)*, 如果$y^0\in F(x^0)$, 存在一常数$\kappa>0$, x_0 的一邻域V 和 y^0 的一邻域W满足

$$F(x)\cap W\subseteq\{y^0\}+\kappa\|x-x^0\|\mathbf{B},\quad\forall x\in V.$$

定义 1.9

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处关于$y^0$是稳健孤立平稳的*(robustly isolated calm)*, 如果$y^0\in F(x^0)$, 存在一常数$\kappa>0$, x_0 的一邻域V和 y^0的一邻域W 满足

$$\emptyset\neq F(x)\cap W\subseteq\{y^0\}+\kappa\|x-x^0\|\mathbf{B},\quad\forall x\in V.$$

定义 1.10

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处是上*Lipschitz* 的,如果存在一常数$\kappa>0$和$x_0$ 的一邻域V 满足

$$F(x)\subseteq F(x^0)+\kappa\|x-x^0\|\mathbf{B},\quad\forall x\in V.$$

定义 1.11

称集值映射$F:\mathcal{X}\rightrightarrows\mathcal{Y}$在$x^0$处关于$y^0$是线性开的,如果存在邻域$V\in\mathcal{N}(x^0)$, $W\in\mathcal{N}(y^0)$ 与常数 $\kappa\in\Re_+$, 满足

$$F(x+\kappa\varepsilon\mathbf{B})\supset[F(x)+\varepsilon\mathbf{B}]\cap W,\ \forall x\in V,\varepsilon>0;$$

定义 1.12

(集值映射的Lipschitz连续性). 称映射$S:\Re^n \rightrightarrows \Re^m$在$\Re^n$的子集$X$上是Lipschitz 连续的, 若它在$X$上是非空闭值的且存在$\kappa \in \Re_+$为Lipschitz 常数, 满足

$$\mathbf{d}_\infty(S(x'), S(x)) \leq \kappa\|x' - x\|, \ \forall x, x' \in X,$$

或等价地,

$$S(x') \subset S(x) + \kappa\|x' - x\|\boldsymbol{B}, \ \forall x, x' \in X.$$

定义 1.13

考虑广义方程
$$0 \in f(x, p) + F(x),$$
其中$F : \mathcal{X} \rightrightarrows \mathcal{Y}$ 是一集值映射. 定义
$$G(x) = f(x^0, p^0) + D_x f(x^0, p^0)(x - x^0) + F(x).$$
如果G^{-1}是从$0 \in \mathcal{Y}$的一个邻域到x^0的一邻域的单值*Lipschitz*连续映射, 则称广义方程在(x^0, p^0)处是强正则的.

Karush-Kuhn-Tucker系统

考虑约束优化问题

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & G(x) \in K, \end{aligned} \tag{2}$$

其中$f : X \to \Re$, $G : X \to Y$, $K \subset Y$是一闭凸锥, X, Y是有限维的Hilbert空间. Karush-Kuhn-Tucker条件

$$\mathrm{D}_x L(x, \lambda) = 0, \ \lambda \in N_K(G(x)).$$

写成广义方程形式

$$0 \in F(x, \lambda) + N_{\Re^n \times K^\circ}(x, \lambda), \ F(x, \lambda) = \begin{bmatrix} \mathrm{D}_x L(x, \lambda) \\ -G(x) \end{bmatrix}.$$

参数Karush-Kuhn-Tucker系统

考虑参数约束优化问题

$$\begin{aligned} \min_x \quad & f(x,u) \\ \text{s.t.} \quad & G(x,u) \in K, \end{aligned} \tag{3}$$

其中$f : X \times U \to \Re$, $G : X \times U \to Y$, $K \subset Y$是一闭凸锥, X, Y是有限维的Hilbert空间. Karush-Kuhn-Tucker条件

$$\mathrm{D}_x L(x,u,\lambda) = 0, \ \lambda \in N_K(G(x,u)).$$

写成广义方程形式

$$0 \in F(x,u,\lambda) + N_{\Re^n \times K^\circ}(x,\lambda), \ F(x,u,\lambda) = \begin{bmatrix} \mathrm{D}_x L(x,u,\lambda) \\ -G(x,u) \end{bmatrix}.$$

强正则性

$$\delta \in F(x_0, u_0, \lambda_0) + \mathrm{D}_x F(x_0, u_0, \lambda_0) \begin{bmatrix} (x - x_0) \\ (\lambda - \lambda_0) \end{bmatrix} + N_{\Re^n \times K^\circ}(x, \lambda)$$

在$(0, x_0, \lambda_0)$附近有唯一的Lipschitz连续的映射$(x(\delta), \lambda(\delta))$.

系统详细写成:

$$\begin{bmatrix} \delta_x \\ \delta_\lambda \end{bmatrix} \in \begin{bmatrix} \mathrm{D}_x L(x_0, u_0, \lambda_0) \\ -G(x_0, u_0) \end{bmatrix} + \begin{bmatrix} \mathrm{D}^2_{xx} L(x_0, u_0, \lambda_0) & \mathrm{D}_x G(x_0, u_0)^* \\ -\mathrm{D}_x G(x_0, u_0) & 0 \end{bmatrix} \begin{bmatrix} (x - x_0) \\ (\lambda - \lambda_0) \end{bmatrix} + N_{\Re^n \times K^\circ}(x, \lambda).$$

定义 1.14

(局部 *Lipschitz* 同胚). 称连续函数 $F:\mathcal{O}\subseteq\mathcal{X}\to\mathcal{X}$在 $x\in\mathcal{O}$ 处是局部 *Lipschitz* 可逆的, 如果存在x的一个开邻域 $\mathcal{N}\subseteq\mathcal{O}$使得限定在这个邻域上的映射 $F|_{\mathcal{N}}:\mathcal{N}\to F(\mathcal{N})$ 是双射并且它的逆函数是 *Lipschitz* 连续的. 称F在x附近是局部 *Lipschitz* 同胚的, 如果F在x附近是局部 *Lipschitz* 可逆的并且F在x处是局部 *Lipschitz* 连续的.

一般形式的约束优化

一般形式的约束优化问题

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & G(x) \in K, \end{aligned} \tag{4}$$

其中$f : X \to \Re$, $G : X \to Y$, $K \subset Y$是一闭凸集合, X, Y是有限维的Hilbert空间. 设U是一Banach空间, $f : X \times U \to \Re$, $G : X \times U \to Y$.

C^2-光滑参数化

称$(f(x,u),G(x,u))$, $u\in U$, 是问题(4)的一C^2-光滑参数化, 如果$f(\cdot,\cdot)$与$G(\cdot,\cdot)$是二次连续可微的, 且存在$\overline{u}\in U$满足$f(\cdot,\overline{u})=f(\cdot)$, $G(\cdot,\overline{u})=G(\cdot)$. 相对应的参数优化问题具有下述形式

$$\begin{array}{ll}\min\limits_{x} & f(x,u)\\ \text{s.t.} & G(x,u)\in K.\end{array}\tag{5}$$

称上述参数化是标准的(canonical), 如果$U:=X\times Y$, $\overline{u}=(0,0)\in X\times Y$, 且

$$\begin{aligned}(f(x,u),G(x,u))= & \ (f(x)-\langle u_1,x\rangle, G(x)+u_2),\\ & \ x\in X,\ u=(u_1,u_2)\in X\times Y.\end{aligned}$$

强稳定性

现在介绍[3, Definition 5.33][4] 中的稳定点的强稳定性的概念, 它在优化问题的灵敏度分析中起重要的作用.

定义 1.15

设x^*是问题(4)的稳定点. 称在x^*处关于$\mathcal{C}^2$-光滑参数化$(f(x,u), G(x,u))$是强稳定的(*strongly stable*), 如果存在 x^*的邻域$\mathcal{V}_X$与$\bar{u}$的邻域$\mathcal{V}_U \subset U$, 满足对任何$u \in \mathcal{V}_U$, 问题(5)存在唯一的稳定点$x(u) \in \mathcal{V}_X$, $x(\cdot)$在$\mathcal{V}_U$ 上连续. 如果这一性质对每一$\mathcal{C}^2$-光滑参数化均是成立的, 则称x^*是强稳定的.

[4] Bonnans J F and Shapiro A. *Perturbation Analysis of Optimization Problems*. New York: Springer-Verlag, 2000.

下述一致二阶增长条件的定义取自[3, Definition 5.16].

定义 1.16

设x^*是问题(4)的稳定点. 称在x^*处关于C^2-光滑参数化$(f(x,u), G(x,u))$的一致二阶增长条件成立, 如果存在$\alpha > 0$, x^*的邻域$\mathcal{V}_X$与$\bar{u}$的邻域$\mathcal{V}_U \subset U$, 满足对任何$u \in \mathcal{V}_U$与问题(5)的稳定点$x(u) \in \mathcal{V}_X$, 下述不等式成立:

$$f(x,u) \geq f(x(u),u) + \alpha\|x - x(u)\|^2, \ \forall x \in \mathcal{V}_X \text{满足} G(x,u) \in K. \quad (6)$$

称在x^*处的一致二阶增长条件成立, 如果(6)式对问题(4)的任何C^2-光滑参数化均是成立的.

定理 1.1

[29, Theorem 9.43][5] 若gph$S$在$(\bar{x},\bar{u}) \in$ gph$S$点处是局部闭的, 则下述条件等价:

(a) (逆Aubin性质): S^{-1} 在$\bar{u}$点关于$\bar{x}$具有Aubin性质;

(b) (度量正则性): $\exists V \in \mathcal{N}(\bar{x})$, $W \in \mathcal{N}(\bar{u})$, $\kappa \in \Re_+$, 满足

$$d(x, S^{-1}(u)) \le \kappa d(u, S(x)), \text{ 若} x \in V, u \in W;$$

(c) (线性开性): $\exists V \in \mathcal{N}(\bar{x})$, $W \in \mathcal{N}(\bar{u})$, $\kappa \in \Re_+$, 满足

$$S(x + \kappa\varepsilon \mathbf{B}) \supset [S(x) + \varepsilon \mathbf{B}] \cap W, \ \forall x \in V, \varepsilon > 0;$$

(d) (伴同导数非奇异性): 满足$0 \in D^*S(\bar{x}|\bar{u})(y)$的$y$只有$y = 0$.

[5] Rockafellar R T and Wets R J B. *Variational Analysis*. New York: Springer-Verlag, 1998.

下述定理表明集值映射的度量次正则性等价于逆映射的平稳性.

定理 1.2

[11, Theorem 3H.3][6] 设集值映射$F:\Re^n \rightrightarrows \Re^m$, $\bar{y} \in F(\bar{x})$. 那么F在$\bar{x}$处关于$\bar{y}$以常数$\kappa > 0$是度量次正则的当且仅当它的逆映射$F^{-1}:\Re^m \rightrightarrows \Re^n$在$\bar{y}$处关于$\bar{x}$以相同的常数$\kappa > 0$是平稳的, 并且有$\mathrm{clm}(F^{-1};\bar{y}|\bar{x}) = \mathrm{subreg}(F;\bar{x}|\bar{y})$.

[6] Dontchev A L and Rockafellar R T. *Implicit Functions and Solution Mappings*. New York: Springer, 2009.

稳定性的变分准则

非凸集合的切锥

设$C \subset \Re^n$是一闭集合.

- Radial cone/雷达锥

$$\mathcal{R}_C(\bar{x}) = \{d : \exists t^* > 0, \forall t \in [0, t^*], \bar{x} + td \in C\}.$$

- Tangent cone/切锥

$$T_C(\bar{x}) = \{d : \exists t_k \searrow 0, d^k \to d, \bar{x} + t_k d^k \in C\}.$$

- Regular tangent cone/正则切锥

$$\widehat{T}_C(\bar{x}) = \{d : \forall y^k \to \bar{x}, \forall t_k \searrow 0, \exists d^k \to d, y^k + t_k d^k \in C\}.$$

非凸集合的法锥

- Regular normal cone/正则法锥

$$\widehat{N}_C(\bar{x}) = \{v \in \Re^n : \langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|) \text{ for } x \in C\}.$$

- Normal cone/法锥

$$N_C(\bar{x}) = \left\{ v \in \Re^n : \begin{array}{c} \exists x^k \in C, \exists v^k \in \widehat{N}_C(x^k) \\ \text{such that } v^k \to v \end{array} \right\}.$$

- 切法共轭关系

Graphical derivatives and coderivatives

定义 2.1

The graphical derivative of S at $\bar{x}$ for any $\bar{u} \in S(\bar{x})$ is the mapping $DS(\bar{x}|\bar{u}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ defined by

$$z \in DS(\bar{x}|\bar{u})(w) \Longleftrightarrow (w, z) \in T_{gph\ S}(\bar{x}, \bar{u}),$$

whereas the coderivative is $D^*S(\bar{x}|\bar{u}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ defined by

$$v \in D^*S(\bar{x}|\bar{u})(y) \Longleftrightarrow (v, -y) \in N_{gph\ S}(\bar{x}, \bar{u}).$$

Regular graphical derivatives and coderivatives

定义 2.2

The regular derivative $\widehat{D}(\bar{x}|\bar{u}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ *and the regular coderivative* $\widehat{D}^*(\bar{x}|\bar{u}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ *are defined by*

$$z \in \widehat{D}S(\bar{x}|\bar{u})(w) \Longleftrightarrow (w, z) \in \widehat{T}_{gph\ S}(\bar{x}, \bar{u}),$$

$$v \in \widehat{D}^*S(\bar{x}|\bar{u})(y) \Longleftrightarrow (v, -y) \in \widehat{N}_{gph\ S}(\bar{x}, \bar{u}).$$

Strict graphical derivatives

For a mapping $S : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$, the strict derivative mapping $D_*S(\bar{x}|\bar{u}) : \mathbf{R}^n \rightrightarrows \mathbf{R}^m$ for S at $\bar{x}$ for $\bar{u}$, where $\bar{u} \in S(\bar{x})$, is defined by

$$D_*S(\bar{x}|\bar{u})(w) := \Big\{ z \mid \exists \tau^\nu \searrow 0, (x^\nu, u^\nu) \overset{\text{gph}\, S}{\to} (\bar{x}, \bar{u}), w^\nu \to w,$$
$$\text{with } z^\nu \in [S(x^\nu + \tau^\nu w^\nu) - u^\nu]/\tau^\nu, z^\nu \to z \Big\}$$

or in other words, for $\Delta_\tau S(x|u)(w) := \dfrac{S(x + \tau w) - u}{\tau}$

$$D_*S(\bar{x}|\bar{u}) := \operatorname{g-limsup}_{\tau \searrow 0, (x,u) \overset{\text{gph}\, S}{\to} (\bar{x}, \bar{u})} \Delta_\tau S(x|u).$$

Note that if F is strictly continuous, one then has the simpler formula

$$D_*F(\bar{x})(w) = \Big\{ z | \exists \tau^\nu \searrow 0, x^\nu \to \bar{x} \text{ with } \Delta_{\tau^\nu} F(x^\nu)(w) \to z \Big\}.$$

Mordukhovich准则

Mordukhovich准则提供了一种计算公式, 可用于判断集值映射是否具有Aubin性质.

定理 2.1

[29, Theorem 9.40] (*Mordukhovich准则*). 设$S : \Re^n \rightrightarrows \Re^m$, $\bar{x} \in \mathrm{dom} S$, $\bar{u} \in S(\bar{x})$. 设$\mathrm{gph} S$ 在$(\bar{x}, \bar{u})$ 点是局部闭的. 则S在$\bar{x}$点关于$\bar{u}$具有*Aubin*性质当且仅当

$$D^* S(\bar{x}|\bar{u})(0) = \{0\},$$

或等价的, $|D^* S(\bar{x}|\bar{u})|^+ < \infty$.

强正则性意味着什么?

考虑由参数广义方程定义的解映射

$$S(x) = \{z \in \Re^k : 0 \in C(x, z) + N_Q(z)\}, \tag{1}$$

其中$C : \mathcal{A} \times \Re^k \to \Re^k$是一连续可微映射，$\mathcal{A} \subset \Re^n$是一开集合，$Q \subset \Re^k$是一非空闭凸集合. 给定$(x_0, z_0) \in \text{gph}S$，$x_0 \in \mathcal{A}$. **要寻求条件，在这一条件下存在$x_0$的一个邻域$\mathcal{O}$和$z_0$的一个邻域$V$满足在$\mathcal{O}$上存在一个单值的Lipschitz连续映射$\sigma : \mathcal{O} \to V$使得**

$$\sigma(x_0) = z_0,\ \sigma(x) \in S(x),\ \forall x \in \mathcal{O}, \tag{2}$$

或者使得

$$\sigma(x) = S(x) \cap V,\ \forall x \in \mathcal{O}. \tag{3}$$

定义

$$\Sigma(\xi) = \{z \in \Re^k : \xi \in C(x_0, z_0) + \mathcal{J}_z C(x_0, z_0)(z - z_0) + N_Q(z)\} \tag{4}$$

与

$$r(x, z) = C(x_0, z_0) + \mathcal{J}_z C(x_0, z_0)(z - z_0) - C(x, z).$$

容易验证下述结论.

命题 2.1

下述关系成立

$$z \in S(x) \text{当且仅当 } z \in \Sigma(r(x, z)).$$

证明

根据Σ与r的定义，有

$$z \in S(x) \text{当且仅当 } 0 \in C(x,z) + N_Q(z) \tag{5}$$

与

$$\begin{aligned} & z \in \Sigma(r(x,z))\text{当且仅当} \\ & r(x,z) \in C(x_0,z_0) + \mathcal{J}_z C(x_0,z_0)(z-z_0) + N_Q(z). \end{aligned} \tag{6}$$

简单的计算可得(5)与(6)中的广义方程是等价的. ■

由于C在$\mathcal{A} \times \Re^k$上是连续可微的, 可以选取x_0的邻域$\widetilde{U}$,z_0的邻域$\widetilde{V}$与一正的实常数L满足

$$\|C(x_1, z) - C(x_2, z)\| \leq L\|x_1 - x_2\|, \ \ \forall x_1 \in \widetilde{U}, z \in \widetilde{V}. \qquad (7)$$

定理 2.2

(a) 设存在$0 \in \Re^k$的邻域W, 存在单值*Lipschitz*连续映射$\phi : W \to \Re^k$, 其*Lipschitz*常数为γ, 满足

$$\phi(0) = z_0, \ \phi(\xi) \in \Sigma(\xi), \ \forall \xi \in W. \tag{8}$$

则对每一$\varepsilon > 0$, 存在x_0的邻域U_ε与z_0的邻域V_ε, 以及一单值映射$\sigma : U_\varepsilon \to V_\varepsilon$满足

$$\sigma(x_0) = z_0, \ \sigma(x) \in S(x), \ \forall x \in U_\varepsilon, \tag{9}$$

且映射σ在U_ε上是*Lipschitz*连续的, *Lipschitz*常数为$(\gamma + \varepsilon)L$, 其中L由(7)定义.

(b) 如果还有,存在z_0的一邻域V满足

$$\phi(\xi) = \Sigma(\xi) \cap V, \ \forall \xi \in W, \tag{10}$$

则

$$\sigma(x) = S(x) \cap V_\varepsilon, \ \forall x \in U_\varepsilon. \tag{11}$$

证明 先证明(a). 对任意固定的$\varepsilon > 0$, 选取$\delta = \delta(\varepsilon) > 0$, $\rho = \rho(\varepsilon) > 0$与$x_0$的一邻域$U_\varepsilon$, 满足对于$V_\varepsilon = z_0 + \rho \mathbf{B}$, 有

$$\begin{aligned}
&\gamma\delta < \varepsilon/(\gamma+\varepsilon), \\
&r(x,z) \in W, && \forall (x,z) \in U_\varepsilon \times V_\varepsilon, \\
&\|\mathcal{J}_z C(x_0,z_0) - \mathcal{J}_z C(x,z)\| \le \delta, && \forall (x,z) \in U_\varepsilon \times V_\varepsilon, \\
&\|C(x_0,z_0) - C(x,z)\| \le (1-\gamma\delta)\rho/\gamma, && \forall x \in U_\varepsilon.
\end{aligned} \tag{12}$$

把(a)的证明分成两部分: (i) 构造σ; (ii) 验证σ 的Lipschitz连续性.
对每一固定的$\overline{x} \in U_\varepsilon$, 定义映射$\Phi_{\overline{x}} : \Re^k \to \Re^k$,

$$\Phi_{\overline{x}}(\cdot) := \phi(r(\overline{x}, \cdot)). \tag{13}$$

下面我们证明

$$\Phi_{\overline{x}} \text{ 是 } V_\varepsilon \text{ 上的一压缩映射, 它把} V_\varepsilon \text{映到} V_\varepsilon. \tag{14}$$

如果上述结论成立, 则由Banach不动点定理可得存在$\overline{z} \in V_\varepsilon$满足

$$\overline{z} = \Phi_{\overline{x}}(\overline{z}) = \phi(r(\overline{x}, \overline{z})).$$

于是根据(8),

$$\overline{z} \in \Sigma(r(\overline{x}, \overline{z})).$$

由命题2.1可得$\overline{z} \in S(\overline{x})$, 因为$\overline{x}$是$U_\varepsilon$中的任意点, 定义在$U_\varepsilon$上的映射

$$\sigma : x \to z \in S(x)$$

是存在的. 由

$$\Phi_{x_0}(z_0) = \phi(r(x_0, z_0)) = \phi(0) = z_0$$

可得$\sigma(x_0) = z_0$, 这证得(9). 下面只需验证(14)式.
为验证$\Phi_{\overline{x}}$的压缩性质, 对$z_1, z_2 \in V_\varepsilon$, 由$W$上定义的$\phi$的Lipschitz连续性质可得

$$\|\Phi_{\overline{x}}(z_1) - \Phi_{\overline{x}}(z_2)\| \leq \gamma\|r(\overline{x}, z_1) - r(\overline{x}, z_2)\|$$
$$\leq \gamma \cdot \sup\{\|\mathcal{J}_z r(\overline{x}, (1-\mu)z_1 + \mu z_2)\| : \mu \in (0,1)\} \cdot \|z_1 - z_2\|.$$

由于$\mathcal{J}_z r(\overline{x}, z) = \mathcal{J}_z C(x_0, z_0) - \mathcal{J}_z C(\overline{x}, z)$, 由 (12)可得

$$\|\Phi_{\overline{x}}(z_1) - \Phi_{\overline{x}}(z_2)\| \leq \gamma\delta\|z_1 - z_2\|, \ \forall z_1, z_2 \in V_\varepsilon. \tag{15}$$

由δ的选取有$\gamma\delta < 1$, $\Phi_{\overline{x}}$实际上是一压缩映射. 进一步,

$$\begin{aligned}
\|\Phi_{\overline{x}}(z_0) - z_0\| &= \|\phi(r(\overline{x}, z_0)) - \phi(0)\| \\
&\leq \gamma\|r(\overline{x}, z_0) - 0\| \\
&= \gamma\|C(x_0, z_0) - C(\overline{x}, z_0)\| \\
&\leq (1 - \gamma\delta)\rho.
\end{aligned}$$

这意味着对于$z \in V_\varepsilon (= z_0 + \rho\mathbf{B})$,

$$\begin{aligned}
\|\Phi_{\overline{x}}(z) - z_0\| &\leq \|\Phi_{\overline{x}}(z) - \Phi_{\overline{x}}(z_0)\| + \|\Phi_{\overline{x}}(z_0) - z_0\| \\
&\leq \gamma\delta\|z - z_0\| + (1 - \gamma\delta)\rho \leq \rho,
\end{aligned} \tag{16}$$

即$\Phi_{\overline{x}}$映V_ε 到自身.

不等式(15)与(16) 表明, 可以用Banach不动点定理, 从而保证映射σ的存在性.
现在证明σ在U_ε上是Lipschitz连续的, Lipschitz常数是$(\gamma+\varepsilon)L$. 不妨设$U_\varepsilon \times V_\varepsilon \subset \widetilde{U} \times \widetilde{V}$, 其中$\widetilde{U}, \widetilde{V}$由(7)定义, 则对任意$x_1, x_2 \in U_\varepsilon$,

$$\begin{aligned}\|\sigma(x_1)-\sigma(x_2)\| &= \|\Phi_{x_1}(\sigma(x_1))-\Phi_{x_2}(\sigma(x_2))\| \\ &\le \|\Phi_{x_1}(\sigma(x_1))-\Phi_{x_1}(\sigma(x_2))\|+\|\Phi_{x_1}(\sigma(x_2))-\Phi_{x_2}(\sigma(x_2))\|.\end{aligned}$$

由(15)可得

$$\|\Phi_{x_1}(\sigma(x_1))-\Phi_{x_1}(\sigma(x_2))\| \le \gamma\delta\|\sigma(x_1)-\sigma(x_2)\|.$$

由ϕ的Lipschitz连续性可得

$$\begin{aligned}\|\Phi_{x_1}(\sigma(x_2))-\Phi_{x_2}(\sigma(x_2))\| &= \|\phi(r(x_1,\sigma(x_2)))-\phi(r(x_2,\sigma(x_2)))\| \\ &\le \gamma\|C(x_1,\sigma(x_2))-C(x_2,\sigma(x_2))\|.\end{aligned}$$

结合这些估计和(7)得到

$$\begin{aligned}\|\sigma(x_1)-\sigma(x_2)\| &\leq \gamma\delta\|\sigma(x_1)-\sigma(x_2)\| \\ &\quad +\gamma\|C(x_1,\sigma(x_2))-C(x_2,\sigma(x_2))\| \\ &\leq \gamma\delta\|\sigma(x_1)-\sigma(x_2)\|+\gamma L\|x_1-x_2\|,\end{aligned}$$

由此可推出

$$\|\sigma(x_1)-\sigma(x_2)\| \leq \frac{\gamma L}{1-\gamma\delta}\|x_1-x_2\| < (\gamma+\varepsilon)L\|x_1-x_2\|,$$

即σ在U_ε上是Lipschitz连续的.

再来证明(b). 如果有必要, 可以选择(12)中的ρ充分小, 所以可以假设$V_\varepsilon \subset V$. 现在固定$x \in U_\varepsilon$, 令z是从$S(x) \cap V_\varepsilon$中任意选取的元素. 为证明(11), 只需证明$z = \sigma(x)$. 根据命题2.1有$z \in \Sigma(r(x,z)) \cap V_\varepsilon$. 由(12)可得$r(x,z) \in W$, 于是由假设(10)和定义式(13)有

$$z = \phi(r(x,z)) = \Phi_x(z).$$

因为$\Phi_x(\cdot)$在V_ε仅有一个不动点, z必是由(a)确定的唯一的不动点$\sigma(x)$, 这证得

$$\sigma(x) = S(x) \cap V_\varepsilon, \ \forall x \in U_\varepsilon.$$

■

定义 2.3

考虑广义方程

$$0 \in f(x, p) + F(x),$$

其中$F : \mathcal{X} \rightrightarrows \mathcal{Y}$是一集值映射. 定义

$$G(x) = f(\bar{x}, \bar{p}) + D_x f(\bar{x}, \bar{p})(x - \bar{x}) + F(x).$$

如果G^{-1}是从$0 \in \mathcal{Y}$的一个邻域到$\bar{x}$的一邻域的Lipschitz连续映射,则称广义方程在$(\bar{x}, \bar{p})$处是强正则的.

解映射

由参数广义方程定义的解映射

$$S(x) = \{z \in \mathbb{R}^k : 0 \in C(x, z) + N_Q(z)\}, \tag{17}$$

其中$C : \mathcal{A} \times \mathbb{R}^k \to \mathbb{R}^k$ 是一连续可微映射, $\mathcal{A} \subset \mathbb{R}^n$是一开集合, $Q \subset \mathbb{R}^k$是一非空闭凸集合.

定理 2.3

给定$(x_0, z_0) \in \mathrm{gph}\, S$, $x_0 \in \mathcal{A}$. 如果(x_0, z_0)是系统(17)的强正则解,则存在x_0的一个邻域$\mathcal{O}$和z_0的一个邻域V, 满足在$\mathcal{O}$上存在一个单值的*Lipschitz*连续映射$\sigma : \mathcal{O} \to V$使得

$$\sigma(x_0) = z_0,\ \sigma(x) \in S(x),\ \ \forall x \in \mathcal{O}, \tag{18}$$

或者使得

$$\sigma(x) = S(x) \cap V,\ \ \forall x \in \mathcal{O}. \tag{19}$$

严格图导数准则

- 针对集值映射的强正则性, Klatte and Kummer[17][7]利用严格图导数的性质给出了强正则性的刻画.
- 称集值映射$F: X \rightrightarrows Y$在$(\bar{x}, \bar{y})$处是强正则的, 若其逆映射F^{-1}在$(\bar{y}, \bar{x})$处具有Aubin性质, 且还分别存在$\bar{x}$的邻域U, $\bar{y}$的邻域V, 使得对$y \in V$, 有$U \cap F^{-1}(y)$是单值的.
- 称严格图导数$D_* F(\bar{x}|\bar{y})$是单射, 若有

$$0 \in D_* F(\bar{x}|\bar{y})(u) \Rightarrow u = 0.$$

[7] Klatte D and Kummer B. *Nonsmooth Equations in Optimization*. Kluwer Academic, 2002.

定理 2.4

[17, Lemma 3.1] (强正则性的严格图导数准则) 设集值映射$F : X \rightrightarrows Y$(赋范空间), $\bar{z} = (\bar{x}, \bar{y}) \in \mathrm{gph} F$. 则有

(a) 若F在$\bar{z}$处是强正则的, 那么$D_* F(\bar{z})$是单射;

(b) 若$X = \Re^n$, 则F在$\bar{z}$处是强正则的充要条件是$D_* F(\bar{z})$是单射且F^{-1} 在$(\bar{y}, \bar{x})$处是*Lipschitz*下半连续的.

S is called Lipschitz l.s.c. at (x, y) with rank L

$$\mathrm{dist}\,(y, S(x')) \le L d_{\mathcal{X}}(x, x')$$

for all x' in some neighborhood V of x.

证明

反证法. (a) 假设$D_* F(\bar{z})$[8] 不是单射, 则$\exists u \neq 0$, 使得$0 \in D_* F(\bar{z})(u)$. 由定义知, 存在$\eta^k \in F(x^k + t_k u^k)$, $t_k \downarrow 0$, $\mathrm{gph} F \ni (x^k, y^k) \to \bar{z}$, $u^k \to u$, 有$v^k = (\eta^k - y^k)/t_k \to 0$. 令$\xi^k = x^k + t_k u^k$, 则

$$\text{存在}(x^k, y^k), (\xi^k, \eta^k) \xrightarrow{\mathrm{gph} F} \bar{z}, \text{使得}\frac{d(\xi^k, x^k)}{d(\eta^k, y^k)} = \|u^k\| \|v^k\|^{-1} \to \infty, \tag{20}$$

这与F在$\bar{z}$处是强正则的矛盾.

[8] 回顾定义

$$D_* F(\bar{x}|\bar{y})(u) := \Big\{ v \mid \exists \tau^\nu \searrow 0, (x^\nu, y^\nu) \overset{\mathrm{gph} F}{\to} (\bar{x}, \bar{y}), u^\nu \to u, \\ \text{with } v^\nu \in [F(x^\nu + \tau^\nu u^\nu) - y^\nu]/\tau^\nu, v^\nu \to v \Big\}$$

(b) 设$X = \Re^n$. 假设F在$\bar{z}$处不是强正则的, 这等价于

$$\text{存在}y^k \to \bar{y}, \text{使得}d(\bar{x}, F^{-1}(y^k)) > kd(y^k, \bar{y}). \qquad (21)$$

或存在$(x^k, y^k) \in \operatorname{gph} F$, $(\xi^k, \eta^k) \in \operatorname{gph} F$它们都收敛到$\bar{z}$,满足

$$\frac{d(\xi^k, x^k)}{d(\eta^k, y^k)} > k. \qquad (22)$$

其中, (21)意味着F^{-1}在$(\bar{y}, \bar{x})$处不是Lipschitz下半连续的.
若(22)成立,(特别地, 当$F^{-1}(y^k)$在$\bar{x}$附近是多值时),
设$t_k = \|\xi^k - x^k\|$ 且$u^k = (\xi^k - x^k)/t_k$. (22) 等价于满足以下条件的这些序列的存在性:

$$\eta^k \in F(x^k + t_k u^k), t_k \downarrow 0, \operatorname{gph} F \ni (x^k, y^k) \to \bar{z},$$

$$\|u^k\| = 1, v := \lim(\eta^k - y^k)/t_k = 0.$$

由于$X=\Re^n$, 则$\exists u\neq 0$, 使得$u^k\to u$, 则有$\exists u\neq 0$, $v=\lim(\eta^k-y^k)/t_k=0$, 其中$\eta^k\in F(x^k+t_k u^k)$, $t_k\downarrow 0$, $\mathrm{gph}F\ni(x^k,y^k)\to\bar{z}$, $u^k\to u$,即$\exists u\neq 0$, 使$0\in D_*F(\bar{z})(u)$, 与单射条件矛盾.

反之, 假设$D_*F(\bar{z})$不是单射或F^{-1}在$(\bar{y},\bar{x})$处不是Lipschitz下半连续的, 由(a)知第一种情况与F在$\bar{z}$处是强正则的矛盾, 而第二种情况也与F在$\bar{z}$处是强正则的矛盾. ■

设X, Y是两个有限维的Hilbert空间, $(\bar{x},\bar{y})\in X\times Y$, 记

$$\pi_x\partial H(\bar{x},\bar{y})=\partial H(\bar{x},\bar{y})\text{ 到空间 }X\text{上的投影}\,.$$

下面的引论是需要的

引理 2.1

设 $H: X\times Y\to X$ 是$(\bar{x},\bar{y})\in X\times Y$的某一开邻域上的局部*Lipschitz*连续函数, $H(\bar{x},\bar{y})=0$. 如果$\pi_x\partial H(\bar{x},\bar{y})$中的每一元素均是非奇异的, 则存在$\bar{y}$的一开邻域 $\mathcal{O}_Y$与一局部*Lipschitz*连续函数 $x(\cdot):\mathcal{O}_Y\to X$ 满足$x(\bar{y})=\bar{x}$且对每一$y\in\mathcal{O}_Y$,

$$H(x(y),y)=0\,.$$

进一步, 如果H在$(\bar{x},\bar{y})$的开邻域中的每一点均是(强)半光滑的, 则$x(\cdot)$在$\mathcal{O}_Y$中的每一点均是(强)半光滑的.

证明

结合Clarke关于局部Lipschitz连续函数的隐函数定理[6, Section 7.1][9]可直接得到前半部分结论成立. 后半部分的证明由 [30, Corollary 2.1][10]和[34, Lemma 2][11]可知,如果H在$(\bar{x}, \bar{y})$的开邻域中的每一点均是(强)半光滑的, 则$x(\cdot)$在$\mathcal{O}_Y$中的每一点均是(强)半光滑的. ∎

---

[9]Clarke F H. *Optimization and Nonsmooth Analysis*. New York: John Wiley and Sons, 1983.

[10]Sun D F. *A further result on an implicit function theorem for locally Lipschitz functions*. Operations Research Letters, 2001, **28**: 193-198.

[11]Kummer, B. *Newton's Method Based on Generalized Derivatives for Nonsmooth Functions: Convergence Analysis*. In: Oettli, W., Pallaschke D. eds., Lecture Notes in Economics and Mathematical Systems 382; Advances in Optimization. Springer, Berlin, 1992, 171-194.

应用场景

变分不等式VI(G,K):求$x \in K$ 满足

$$\langle G(x), z-x\rangle \geq 0 \quad \forall z \in K.$$

它等价于

$$-G(x) \in N_K(x).$$

扰动问题

$$-G(x) - y \in N_K(x).$$

用Natural mapping表达成非光滑方程:

$$H(x,y) = x - \Pi_K(x - G(x) - y) = 0.$$

Clarke广义Jacobian

- H的 Clarke广义Jacobian

$$\partial H(x,y) = \left\{ [\mathcal{I} - (\mathcal{I} - DG(x)^*)V \; V] : V \in \partial \Pi_K(x - G(x) - y) \right\}.$$

- 到$\mathcal{X}$上的投影

$$\pi_x \partial H(x,y) = \left\{ \mathcal{I} - (\mathcal{I} - DG(x)^*)V : V \in \partial \Pi_K(x - G(x) - y) \right\}.$$

- 很多情况下Π_K是半光滑的,比如$K = P$,P是凸的多面体集合,$K = \mathbb{S}^p$, $K = Q_{m+1}$, $K = \mathrm{epi}\, \| \cdot \|$, $\| \cdot \|_2$是谱范数,$\| \cdot \|_*$是核范数.

定理 2.5

[29, Theorem 9.54] (单值局部化). 设$S:\Re^n \rightrightarrows \Re^m$, $\bar{u}\in S(\bar{x})$. 设S以下述含义在$\bar{x}$ 相对于$\bar{u}$是局部内半连续的, 即存在邻域$V\in\mathcal{N}(\bar{x})$与$W\in\mathcal{N}(\bar{u})$, 满足对任何$x\in V$与$\varepsilon>0$, 存在$\delta>0$,

$$\left.\begin{array}{l} S(x)\cap W\subset S(x')+\varepsilon \boldsymbol{B} \\ S(x')\cap W\subset S(x)+\varepsilon \boldsymbol{B} \end{array}\right\} \text{当 } x'\in V\cap \boldsymbol{B}(x,\delta).$$

(a) S在$\bar{x}$处相对于$\bar{u}$具有一*Lipschitz* 连续的单值局部化T当且仅当$D_*S(\bar{x}|\bar{u})(0)=\{0\}$.

(b) 如果S是凸值的, 则S^{-1}在$\bar{u}$处相对于$\bar{x}$有*Lipschitz*连续的单值局部化T当且仅当$D_*S(\bar{x}|\bar{u})^{-1}(0)=\{0\}$且$m=n$.

推论 2.1

[29, Corollary 9.55] (单值*Lipschitz*函数的可逆性).
令$\mathcal{O} \subset \Re^n$是一开集合, $F : \mathcal{O} \to \Re^n$是一连续映射. 对于$\bar{x} \in \mathcal{O}$, F^{-1}在$\bar{u} = F(\bar{x})$处具有一*Lipschitz*连续的单值局部化的充分必要条件是F满足非奇异严格导数条件:

$$D_* F(\bar{x})(w) = 0 \Longrightarrow w = 0.$$

证明. 本推论是将集值映射S替换为单值映射F后定理2.5(b)的具体化. ■

以下的内容取自Kummer (1991)[18].[12]
下面的定义可用于刻画局部Lipschitz可逆性质.

定义 2.4

[18, Definition 1.3]设 $F: \Re^n \to \Re^m$ 是一个连续函数, $x \in \Re^n$, 定义 $\Delta F(x)$为

$$\Delta F(x) = \left\{ z \;\middle|\; \begin{array}{l} \exists x^\nu \to x, y^\nu \to x, x^\nu \neq y^\nu \\ \text{使得 } \dfrac{F(y^\nu) - F(x^\nu)}{\|y^\nu - x^\nu\|} \to z \end{array} \right\}.$$

[12] Kummer B. *Lipschitzian inverse functions, directional derivatives, and applications in $C^{1,1}$-optimization*. Journal of Optimization Theory and Applications, 1991, **70**: 559-580.

引理 2.2

[18, Lemma 2.1] 连续函数 $F:\Re^n \to \Re^n$在x附近是局部*Lipschitz*可逆的当且仅当 $0 \notin \Delta F(x)$.

证明. $(\Rightarrow)$ 利用反证法. 假设连续函数F在x附近是局部Lipschitz可逆的, 但 $0 \in \Delta F(x)$, 则存在 $x^\nu \to x, y^\nu \to x\, , x^\nu \neq y^\nu$使得

$$[F(y^\nu) - F(x^\nu)]/\|y^\nu - x^\nu\| \to 0$$

成立. 而由F^{-1}在$F(x)$处局部Lipschitz连续可得, 存在正数κ使得

$$\|F^{-1}(F(y^\nu)) - F^{-1}(F(x^\nu))\| \leq \kappa \|F(y^\nu) - F(x^\nu)\|$$

成立. 这产生了矛盾.

$(\Leftarrow)$ 若$0 \notin \Delta F(x)$, 则存在正数ε和κ使得

$$\|F(x') - F(x'')\| \geq \kappa \|x' - x''\|, \ \forall \ x', x'' \in B(x, \varepsilon). \tag{23}$$

由(23)知F是一单射, 再注意它是连续的, 则对于开集合$\mathcal{O} \subset \Re^n$, 有$F(\mathcal{O})$是开集合. 于是存在$\delta > 0$使得$\mathbf{B}(F(x), \delta) \subset F(\mathbf{B}(x, \varepsilon))$. 所以由(23)可得$F(y) = z, y \in \mathbf{B}(x, \varepsilon)$存在唯一解$y = F^{-1}(z)$且 F^{-1}在$\mathbf{B}(F(x), \delta)$上是Lipschitz连续的. ■

引理 2.3

[18, Lemma 2.2][13] 若 $F \in \mathcal{C}^{0,1}(\Re^n, \Re^m)$, 则

$$\Delta F(x) = \bigcup_{\|u\|=1} D_* F(x)(u).$$

证明. 由定义可知, $\bigcup_{\|u\|=1} D_* F(x)(u) \subset \Delta F(x)$是显然的. 现证明相反的包含关系.

[13] Note that if F is strictly continuous, one then has the simpler formula

$$D_* F(\bar{x})(w) = \left\{ z | \exists \tau^\nu \searrow 0, x^\nu \to \bar{x} \text{ with } \Delta_{\tau^\nu} F(x^\nu)(w) \to z \right\}.$$

任取 $z \in \Delta F(x)$, 则存在 $x^\nu \to x, y^\nu \to x\,, x^\nu \neq y^\nu$ 使得

$$z^\nu = [F(y^\nu) - F(x^\nu)]/\|y^\nu - x^\nu\| \to z$$

成立. 令 $\lambda^\nu = \|y^\nu - x^\nu\|, u^\nu = (y^\nu - x^\nu)/\lambda^\nu$, 则 $\{u^\nu\}$是有界数列, 不妨设 $u^\nu \to u$, 则 $\|u\| = 1$. 定义

$$v^\nu = (F(x^\nu + \lambda^\nu u) - F(x^\nu))/\lambda^\nu,$$

则对充分大的 ν, 存在正数 κ 使得

$$\begin{aligned} \|z^\nu - v^\nu\| &= \|F(y^\nu) - F(x^\nu + \lambda^\nu u)\|/\lambda^\nu \\ &= \|F(x^\nu + \lambda^\nu u^\nu) - F(x^\nu + \lambda^\nu u)\|/\lambda^\nu \\ &\leq \kappa\|\lambda^\nu(u^\nu - u)\|/\lambda^\nu \end{aligned}$$

成立. 所以 $\lim_{\nu\to\infty} z^\nu = z = \lim_{\nu\to\infty} v^\nu \in D_*F(x)(u)$. ■

Kummer逆函数定理

基于上述引理可以得到逆函数定理.

定理 2.6

[18, Theorem 1.1](*Kummer*逆函数定理). 设函数 $F:\mathcal{O}\subset\mathcal{X}\to\mathcal{X}$ 在 $x\in\mathcal{O}$附近是局部*Lipschitz*连续的. 则 F在 $\bar{x}$ 附近是*Lipschitz*同胚的当且仅当下述非奇异条件成立:

$$0\notin D_*F(\bar{x})(u),\quad \forall\, 0\neq u\in X.$$

证明

由引理2.2和引理2.3可知，F 在$\bar{x}$附近是Lipschitz同胚的当且仅当

$$0 \notin \bigcup_{\|u\|=1} D_* F(\bar{x})(u).$$

而由$D_* F(x)$的正齐次性可得

$$0 \notin \bigcup_{\|u\|=1} D_* F(x)(u)$$

等价于 $0 \notin D_* F(\bar{x})(u), \forall\, 0 \neq u \in X.$ ■

孤立平稳性的图导数准则

结合[19, Proposition 4.1] [14]及[15, Proposition 2.1][15], 给出集值映射孤立平稳性的图导数准则, 即用集值映射的图导数刻画孤立平稳性.

定理 2.7

(孤立平稳性的图导数准则) 设 X, Y是两个有限维的*Hilbert*空间. 设集值映射$F: X \rightrightarrows Y$. 对$(\bar{x}, \bar{y}) \in \mathrm{gph} F$, F在$\bar{x}$处关于$\bar{y}$孤立平稳的充要条件是$\{0\} = DF(\bar{x}|\bar{y})(0)$.

[14]Levy A B. *Implicit multifunction theorems for the sensitivity analysis of variational conditions*. Math. Program., 1996, **74**: 333-350.

[15]King A and Rockafellar R T. *Sensitivity analysis for nonsmooth generalized equations*. Math Program, 1992, **55**: 341 - 364.

证明

必要性.[16]考虑任一$v \in DF(\bar{x}|\bar{y})(0)$, 则存在序列$v^k \to v$, $u^k \to 0$, $t_k \downarrow 0$, 满足对$\forall k$, 有$\bar{y} + t_k v^k \in F(\bar{x} + t_k u^k)$. 因为$F$在$\bar{x}$处关于$\bar{y}$孤立平稳的, 有$\bar{x}$的邻域$V$, $\bar{y}$ 的邻域W, 常数$\kappa > 0$满足$F(x) \cap W \subseteq \{\bar{y}\} + \kappa\|x - \bar{x}\|\mathbb{B}_Y$, $\forall x \in V$. 对充分大的k, 有$\bar{y} + t_k v^k \in \bar{y} + \kappa\|t_k u^k\|\mathbb{B}_Y$, 即$v^k$ 包含在半径为$\kappa\|u^k\|$ 的球中, 由于$\{u^k\} \to 0$, 则v必为0.

[16]回顾

$$\begin{aligned} DS(\bar{x}|\bar{u})(\bar{w}) &= \limsup_{\tau \searrow 0, w \to \bar{w}} \frac{S(\bar{x} + \tau w) - \bar{u}}{\tau} \\ &= \left\{ v : \exists t^\nu \searrow 0, w^\nu \to \bar{w}, \exists u^\nu \in S(\bar{x} + \tau^\nu w^\nu), \frac{u^\nu - \bar{u}}{\tau^\nu} \to v \right\}. \end{aligned}$$

充分性. 反证法. 假设F在$\bar{x}$处关于$\bar{y}$不是孤立平稳的, 则存在序列$x^k \to \bar{x}$, $\exists y^k \in F(x^k)$, 使得$y^k \notin \{\bar{y}\} + k\|x^k - \bar{x}\|\mathrm{B}_Y$, 则有$\|y^k - \bar{y}\| > k\|x^k - \bar{x}\|$. 令$t_k = \|y^k - \bar{y}\|$, $v^k = (y^k - \bar{y})/t_k$, 则$\|v^k\| = 1$, 由于$Y$有限维, $\exists v \neq 0$, 使得$v^k \to v$. 令$u^k = (x^k - \bar{x})/t_k$,则

$$\|u^k\| = \|x^k - \bar{x}\|/\|y^k - \bar{y}\| < 1/k \to 0,$$

即存在$t_k \downarrow 0$, $(u^k, v^k) \to (0, v)$满足$\bar{y} + t_k v^k \in F(\bar{x} + t_k u^k)$, 这意味着$0 \neq v \in DF(\bar{x}|\bar{y})(0)$, 与条件矛盾. ■

系统稳定性

考虑下述集值映射的度量正则性的刻画:

$$S(u) = \{x \in \mathcal{X} : G(x, u) \in K\}$$

其中$\mathcal{X}$是有限维Hilbert空间, $K \subset \mathcal{Y}$是非空闭凸集合, $\mathcal{Y}$是有限维Hilbert空间.

- $K = \{0\} \subset \mathcal{Y}$的情况,经典的隐函数定理可回答最简单的情况;
- K是一般的闭凸集合的情况,Robinson约束规范刻画度量正则性.

定理 3.1

[17] 设X是赋范向量空间,Y与Z是两个Banach空间, $\Omega \subset X \times Y$是包含点$(a,b)$的一开集合,$\phi \in \mathcal{C}(\Omega; Z)$ 满足

(i) $\phi(a,b) = 0$;

(ii) $\dfrac{\partial\phi}{\partial y}(x,y) \in \mathcal{L}(Y;Z)$对任何$(x,y) \in \Omega$ 存在,且 $\dfrac{\partial\phi}{\partial y} \in \mathcal{C}(\Omega; \mathcal{L}(Y;Z))$;

(iii) $\dfrac{\partial\phi}{\partial y}(a,b) \in \mathcal{L}(Y;Z)$是双射,$\left(\dfrac{\partial\phi}{\partial y}(a,b)\right)^{-1} \in \mathcal{L}(Z,Y)$.

[17] pp.548-549, Philippe G. Ciarlet, Linear and Nonlinear Functional Analysis with Applications,SIAM,Philadelphia,2013.

(a) 则存在X中的a的一开邻域V,存在Y中的b的一开邻域W与一隐函数$f \in \mathcal{C}(V; W)$满足

$$V \times W \subset \Omega, \left[\begin{array}{l} \{(x,y) \in V \times W : \phi(x,y) = 0\} \\ \quad = \{(x,y) \in V \times W : y = f(x)\}. \end{array} \right].$$

(b) 如果ϕ在$(a,b) \in \Omega$是可微的,那么f在a处是可微的,

$$f'(a) = -\left(\frac{\partial \phi}{\partial y}(a,b)\right)^{-1} \frac{\partial \phi}{\partial x}(a,b) \in \mathcal{L}(X; Y).$$

(c) 如果$\phi \in \mathcal{C}^m(\Omega; Z)$,其中$m \geq 1$是整数,则存在$X$中的$a$的一开邻域$\widetilde{V} \subset V$与$Y$中的$b$的一开邻域$\widetilde{W} \subset W$ 满足对任何$(x,y) \in \widetilde{V} \times \widetilde{W}$, $\dfrac{\partial \phi}{\partial y}(x,y) \in \mathcal{L}(Y; Z)$是双射,

$$\left(\frac{\partial \phi}{\partial y}(x,y)\right)^{-1} \in \mathcal{L}(Z, Y), f \in \mathcal{C}^m(\widetilde{V}; Y)$$

对任何$x \in \widetilde{V}$,

$$f'(x) = -\left(\frac{\partial \phi}{\partial y}(x, f(x))\right)^{-1} \frac{\partial \phi}{\partial x}(x, f(x)) \in \mathcal{L}(X; Y).$$

集值映射的闭与凸性

- 集值映射Ψ在$x \in X$处被称为是闭的，若$x_n \to x$，$y_n \in \Psi(x_n)$，且 $y_n \to y$，则$y \in \Psi(x)$. 称Ψ是闭的，若它在X中的每一点均是闭的.
- 注意到Ψ是闭的当且仅当它的图gph(Ψ)是$X \times Y$中的一闭子集.
- 称Ψ是凸的(convex)，若它的图gph(Ψ)是$X \times Y$中的一凸子集. 或等价地，Ψ是凸的充分必要条件是对任何$x_1, x_2 \in X$，$t \in [0,1]$，

$$t\Psi(x_1) + (1-t)\Psi(x_2) \subset \Psi(tx_1 + (1-t)x_2). \qquad (1)$$

广义开映射定理

若$A: X \to Y$是一连续的线性算子, A是映上的条件等价于条件 $0 \in \text{int}A(X)$. 可将开映射定理推广到具有闭凸图的集合值函数的情形.

定理 3.2

(广义开映射定理)设X与Y是$Banach$空间, $\Psi : X \to 2^Y$是闭的凸的集值函数. 令$y \in int(range\ \Psi)$. 则
对$x \in \Psi^{-1}(y)$及$\forall r > 0$有$y \in int\ \Psi(B_X(x, r))$.

定义 3.1

称多值函数$\Psi : X \to 2^Y$在$(x_0, y_0) \in gph(\Psi)$ 以线性率$\gamma > 0$为开的, 若存在$t_{\max} > 0$及(x_0, y_0)的邻域N满足对$\forall (x, y) \in gph(\Psi) \bigcap N$, $\forall t \in [0, t_{\max}]$, 下述包含关系成立:

$$y + t\gamma B_Y \subset \Psi(x + tB_X). \tag{2}$$

命题 3.1

若多值函数Ψ是凸的,则Ψ在点$(x_0, y_0) \in gph(\Psi)$处为开的充分必要条件是存在正数η, ν满足

$$y_0 + \eta B_Y \subset \Psi(x_0 + \nu B_X). \tag{3}$$

多值函数的闭凸性

注意到，若多值函数Ψ是闭凸的，则由广义开映射定理3.2，由正则性条件$y_0 \in \text{int}(\mathrm{range}\,\Psi)$可得，存在$\eta$与$\nu$满足(3)．显然，相反的结论亦成立．

命题 3.2

设多值函数$\Psi : X \to 2^Y$是闭的，凸的．则Ψ在(x_0, y_0)处是开的当且仅当$y_0 \in int(\mathrm{range}\,\Psi)$.

开性等价于度量正则性

定义 3.2

称多值函数$\Psi: X \to 2^Y$在$(x_0, y_0) \in gph\Psi$以率c度量正则的，若对(x_0, y_0)一邻域中的所有的(x, y)有

$$\operatorname{dist}(x, \Psi^{-1}(y)) \leq c \operatorname{dist}(y, \Psi(x)). \tag{4}$$

定理 3.3

多值函数$\Psi: X \to 2^Y$在$(x_0, y_0) \in gph\Psi$以率c为度量正则的当且仅当Ψ在(x_0, y_0)处以$\gamma = c^{-1}$为率是开的.

闭凸集值映射的度量正则性

定理 3.4

(Robinson–Ursescu稳定性定理)令$\Psi : X \to 2^Y$是闭凸多值函数. 则Ψ在$(x_0, y_0) \in gph(\Psi)$ 处度量正则的充要条件是正则性条件$y_0 \in int(\mathrm{range}\,\Psi)$成立. 更精确地, 设(3)成立, (x, y)满足

$$\|x - x_0\| < \frac{1}{2}\nu, \quad \|y - y_0\| < \frac{1}{8}\eta. \tag{5}$$

则$c = 4\nu/\eta$时的(4)成立.

约束系统的度量正则性

考虑连续映射$G : X \to Y$, 闭凸集$K \subset Y$, 与相应的集值映射

$$\mathcal{F}_G(x) = G(x) - K. \tag{6}$$

关系$y_0 \in \mathcal{F}_G(x_0)$意味着$G(x_0) - y_0 \in K$. 设 $y_0 \in \mathcal{F}_G(x_0)$, 若$\mathcal{F}_G$ 在(x_0, y_0)处是度量正则的, 即如果(x, y)在(x_0, y_0) 的一邻域中, 有

$$d(x, \mathcal{F}_G^{-1}(y)) \leq c\, d(y, \mathcal{F}_G(x)). \tag{7}$$

Lipschitz扰动下的度量正则性

定理 3.5

([3, Theorem 2.84]). 设$G : X \to Y$是一连续映射. 设相应的集值映射$\mathcal{F}_G$在(x_0, y_0)处以率$c > 0$ 度量正则, 差值映射$D(x) := G(x) - H(x)$在x_0的一邻域以模$\kappa < c^{-1}$ Lipschitz连续. 则集值映射$\mathcal{F}_H$ 在$(x_0, y_0 - D(x_0))$处以率$c(\kappa) := c(1 - c\kappa)^{-1}$ 度量正则, 即

$$d(x, \mathcal{F}_H^{-1}(y)) \leq c(\kappa) d(y, \mathcal{F}_H(x)) \tag{8}$$

对充分接近于$(x_0, y_0 - D(x_0))$的(x, y)成立.

Taylor展式对应的集值映射

设$G(x)$是可微的且$DG(x)$是关于x的连续映射. 考虑点$x_0 \in \Phi$和集值映射

$$\mathcal{F}^*(x) = G(x_0) + DG(x_0)(x - x_0) - K. \tag{9}$$

由中值定理, 差函数

$$G(x) - [G(x_0) + DG(x_0)(x - x_0)]$$

在x_0的邻域V内是Lipschitz连续的, 其相应的Lipschitz常数κ可以充分小. 结合定理3.5, 这可推出,若线性化集值映射$\mathcal{F}^*$在$(x_0, 0)$处是度量正则的, 则$\mathcal{F}_G$在$(x_0, 0)$处亦是度量正则的. 相反地,$\mathcal{F}_G$在$(x_0, 0)$处的度量正则性可推出$\mathcal{F}^*$的度量正则性.

关于Robionson约束规范

注意

$$\mathcal{F}^*(x) = G(x_0) + DG(x_0)(x - x_0) - K,$$

有

$$\mathrm{range}\,\mathcal{F}^* = G(x_0) + DG(x_0)X - K.$$

根据命题3.2[18], 线性化集值映射$\mathcal{F}^*$在$(x_0, 0)$处是度量正则的充分必要条件是Robinson约束规范成立:

$$0 \in \mathrm{int}\,\left(G(x_0) + DG(x_0)X - K\right).$$

$\mathcal{F}_G$ 在$(x_0, 0)$处的度量正则性等价于Robinson约束规范.

[18] 命题3.2 设多值函数$\Psi : X \to 2^Y$ 是闭的, 凸的. 则Ψ 在(x_0, y_0)处是开的当且仅当$y_0 \in \mathrm{int}(\mathrm{range}\,\Psi)$.

凸优化的Aubin性质

凸优化模型

考虑具有下述一般形式的凸优化问题:

$$(\mathrm{P}) \qquad \min f(x) \quad \text{s.t.} \quad x \in Q. \tag{1}$$

其中$f : \Re^n \to \Re$为$\mathcal{C}^2$的凸函数,$Q \subseteq \Re^n$ 为闭凸集.
可用指示函数将问题(P)等价的写成无约束优化问题:

$$\min_{x \in \Re^n} f(x) + \delta_Q(x).$$

由于$f(x) + \delta_Q(x)$为凸函数, 且f 二次连续可微, 则凸优化问题(P)的KKT系统可以写成下述广义方程的形式:

$$0 \in \partial(f(x) + \delta_Q(x)) = \nabla f(x) + N_Q(x). \tag{2}$$

定义 4.1

称映射$T : \Re^n \rightrightarrows \Re^n$ 是单调的(*monotone*), 如果对于$v_0 \in T(x_0)$, $v_1 \in T(x_1)$, 有

$$\langle v_1 - v_0, x_1 - x_0 \rangle \geq 0;$$

称映射T是严格单调的(*strictly monotone*), 如果对于$x_0 \neq x_1$, 上述不等式成为严格不等式. 映射T是极大单调的, 如果图gph$\,T$不真包含在其他任何的单调算子$T' : \Re^n \rightrightarrows \Re^n$的图gph$\,T'$中.

关于极大单调性, 已有许多熟知的结论, 例如,下半连续凸函数的次微分是极大单调的. 对于$\Re^n$ 中的任意闭凸集合$C \neq \emptyset$, 法锥映射N_C是极大单调的. 如果T具有下述形式:

$$T(x) = \begin{cases} T_0(x) + N_D(x), & x \in D, \\ \emptyset, & x \notin D, \end{cases}$$

其中$D \subset \Re^n$是一非空闭凸子集合, $T_0 : D \to \Re^n$是单值的单调的连续映射, 则这样的算子是极大单调的. 当然, 对于单调算子T, 其逆映射T^{-1}亦是单调的.

满足Aubin性质的极大单调算子

将通过利用[13, Proposition 5.1]来构建问题(P)的KKT系统的强正则性与Aubin性质的等价性.

命题 4.1

[13, Proposition 5.1] 设X是$Banach$空间, $F : X \rightrightarrows X^*$为单调映射且在$(x_0, y_0)$处具有$Aubin$ 性质, 那么F在x_0的一邻域内是单值的.

证明

假设F在x_0的任何邻域内均不是单值的. 那么存在序列$x_k \to x_0$ 使得对任一$y_k \in F(x_k)$, $k = 1, 2, \ldots,$ 存在$z_k \in F(x_k)$, $k = 1, 2, \ldots,$ 使得对所有k, 均有$z_k \neq y_k$. 因为F在(x_0, y_0)处具有Aubin性质, 可选取$y_k \in F(x_k)$ 满足$y_k \to y_0$. 对每一个k存在一个线性函数严格分离点y_k和z_k, 即对每个$k = 1, 2, \ldots,$ 存在$h_k \in X$, $\|h_k\| = 1$与常数$b_k > 0$, 使得

$$\langle z_k, h_k \rangle \geq b_k + \langle y_k, h_k \rangle. \tag{3}$$

设F以模γ及邻域U和W具有Aubin 性质. 取一数列t_k满足

$$t_k > 0,\ \ t_k \to 0,\ \ 且 t_k < b_k/2\gamma. \tag{4}$$

则对充分大的k, 有$x_k \in U$, $x_k + t_k h_k \in U$且$y_k \in W$. 根据F的Aubin性质, 有

$$y_k \in F(x_k) \cap W \subset F(x_k + t_k h_k) + \gamma t_k \mathbf{B}.$$

因此, 存在序列$u_k \in F(x_k + t_k h_k)$ 使得

$$\|u_k - y_k\| \leq \gamma t_k. \tag{5}$$

由F的单调性,

$$\langle u_k - z_k, x_k + t_k h_k - x_k \rangle \geq 0.$$

结合(3), 有

$$\langle u_k, h_k \rangle \geq \langle z_k, h_k \rangle \geq b_k + \langle y_k, h_k \rangle.$$

进一步, 由(4)和(5),

$$b_k + \langle y_k, h_k \rangle \leq \langle u_k, h_k \rangle \leq \langle y_k, h_k \rangle + \gamma t_k < b_k/2 + \langle y_k, h_k \rangle,$$

显然矛盾. 因此, 假设不成立, F在x_0的一邻域是单值的. ■

强正则性与Aubin性质的等价性

定理 4.1

对于凸优化问题(1), 广义方程(2)在x_0附近是强正则的(即x_0为$0 \in \nabla f(x) + N_Q(x)$ 的强正则解)的充分必要条件为T^{-1}在x_0处具有$Aubin$性质, 其中映射$T : \Re^n \rightrightarrows \Re^n$为
$T(x) := \nabla f(x_0) + \nabla^2 f(x_0)(x - x_0) + N_Q(x)$.

证明

必要性由强正则的定义即可得到.

充分性. 设T^{-1}在x_0处具有Aubin 性质. 由于凸函数f的Hessian阵是半正定的, 则有下式成立:

$$\langle \nabla^2 f(x_0)(x_1 - x_2), x_1 - x_2 \rangle \geq 0, \ \forall x_1, x_2 \in \Re^n. \tag{6}$$

因此, 映射$T'(x) := \nabla f(x_0) + \nabla^2 f(x_0)(x - x_0)$ 是单调的且单值的, 则映射T是单调的, 进而T^{-1}是单调的. 由命题4.1知, T^{-1}在x_0的一邻域内是单值的, 由强正则定义得结论成立.

■

NLP的稳定性

稳定性结果综述

- Robinson (1981): If the multi-valued mapping $F : \mathcal{X} \rightrightarrows \mathcal{Y}$ is piecewise polyhedral, then F is calm at x^0.
- Robinson (1980): showed that the strong second order sufficient condition and the LICQ imply the strong regularity of the solution to the KKT system. Interestingly, the converse is also true, see Jongen et al. (1990).

- Robinson (1982): showed that the second order sufficient condition and MFCQ imply the upper Lipschitz continuity of KKT solutions.
- Dontchev and Rockafellar (1997) showed that the strict MFCQ and the second-order sufficient optimality conditons are equivalent to the robust isolated calmness of the KKT system.

NLP稳定性的参考文献

1 S. M. Robinson, Some continuity properties of polyhedral multifunctions, Mathematical Programming Study, 14 (1981), 206-214.

2 S. M. Robinson, Strongly regular generalized equations, Mathematics of Operations Research 5(1980), 43 – 62.

3 H. Th. Jongen, J. Ruckmann, and K. Tammer, Implicit functions and sensitivity of stationary points, Mathematical Programming 49 (1990), 123 – 138.

4 S. M. Robinson, Generalized Equations and Their Solutions, Part II: Applications to Nonlinear Programming,Mathematical Programming Study 19 (1982), 200-221.

5 A.L. Dontchev and R.T. Rockafellar, Characterizations of Lipschitzian stability in nonlinear programming. in *Mathematical Programming With Data Perturbations*, A.V. Fiacco, ed.), Marcel Dekker, New York, 1997, 65–82.

多面集值映射的上Lipschitz连续性

定义 5.1

A set-valued mapping $\Gamma : \Re^n \rightrightarrows \Re^m$ *is called polyhedral, if its graph is the union of finitely many polyhedral sets, called components of* Γ.

下述关于多面集值映射的定理来源于文献[26].[19]

定理 5.1

设$S : \Re^n \rightrightarrows \Re^m$是一多面集值映射，则$S$在每个点$\bar{x} \in \mathrm{dom}\, S$处均是上*Lipschitz* 连续的.

[19] Robinson S M. *Some continuity properties of polyhedral multifunctions*. Mathematical Programming Study, 1981, **14**: 206-214.

引理 5.1

Let $P : \Re^n \rightrightarrows \Re^m$ *be a polyhedral set-valued mapping with components* $G_i, i = 1, \ldots, k$*. Suppose that* $x \in \mathrm{dom}\, P$ *and define the index set*

$$J(x) = \{i \in [k] : x \in \pi_1(G_i)\},$$

where π_1 *denotes the canonical projection of* $\Re^n \times \Re^m$ *onto* $\Re^n$*. Then there is a neighborhood* U *of* x *such that*

$$(U \times \Re^m) \cap \mathrm{ghp}\, P \subset \bigcup_{i \in J(x)} G_i.$$

Proof

The affine subspace $\{x\} \times \Re^m$ and the components G_i, $i \in [k]$, are nonempty polyhedral subsets of $\Re^n \times \Re^m$. If $j \notin J(x)$, the intersection of $\{x\} \times \Re^m$ and G_i is empty and these two sets can be strongly separated. Hence there are neighborhoods U_i of x such that

$$(U_i \times \Re^m) \cap G_i = \emptyset \text{ for } i \notin J(x).$$

Thus $U := \cap_{i \notin J(x)} U_i$ is also a neighborhood of x and

$$(U \times \Re^m) \cap \operatorname{gph} P \subset \left(\bigcup_{i=1}^{k} G_i \right) \setminus \left(\bigcup_{i \notin J(x)} G_i \right) \subset \bigcup_{i \in J(x)} G_i,$$

as required. $\square$

引理 5.2

Let G be a nonempty polyhedral set in $\Re^n \times \Re^m$. *For* $z = (x, y) \in \pi_1(G) \times \pi_2(G)$ *define*

$$d_x(z, G) = \min\{\|x' - x\| : (x', y) \in G\}$$

and

$$d_y(z, G) = \min\{\|y' - y\| : (x, y') \in G\}$$

the "horizontal" and the "vertical" distance of z to G, respectively. Then there exist nonnegative real numbers ξ,η *such that*

$$d_x(z, G) \leq \eta d_y(z, G) \text{ and } d_y(z, G) \leq \xi d_x(z, G) \qquad (1)$$

for all $z \in \pi_1(G) \times \pi_2(G)$.

Proof

The convex polyhedral G can be represented in the form

$$G = \{(x, y) \in \Re^n \times \Re^m : Ax + By \leq c\},$$

where $A \in \Re^{l \times n}$, $B \in \Re^{l \times m}$ and $c \in \Re^l$. By the standard form of Hoffman's theorem there are reals α and β such that for each $a \in \mathcal{R}(A) + \Re^l_+$, $b \in \mathcal{R}(B) + \Re^l_+$,$x_0 \in \Re^n$ and $y_0 \in \Re^m$ one has

$$\mathrm{dist}\Big(x_0, \{x' : Ax' \leq a\}\Big) \leq \alpha \|(Ax_0 - a)^+\|$$

and

$$\mathrm{dist}\Big(y_0, \{y' : By' \leq b\}\Big) \leq \beta \|(By_0 - b)^+\|. \tag{2}$$

Put $\xi := \beta\|A\|$,$\eta := \alpha\|B\|$ and choose any $z := (x, y) \in \pi_1(G) \times \pi_2(G)$. Then we get from (2) that

$$\begin{aligned} d_y(z, G) &= \operatorname{dist}\Big(y, \{y' : By' \leq c - Ax\}\Big) \\ &\leq \beta\|(Ax + By - c)^+\|. \end{aligned} \tag{3}$$

For $\tilde{x}$ closest to x in the set $\{x' : Ax' \leq c - By\}$ one has

$$\|(Ax + By - c)^+\| \leq \|(Ax + By - c) - (A\tilde{x} + By - c)\|, \tag{4}$$

which yields

$$\|(Ax + By - c)^+\| \leq \|A\|\|x - \tilde{x}\|. \tag{5}$$

But $\|x - \tilde{x}\| = d_x(z, G)$ by construction and thus, combining (3), (4) and (5), we get

$$d_y(z, G) \leq \beta\|(Ax + By - c)^+\| \leq \beta\|A\|\|x - \tilde{x}\| = \xi d_x(z, G).$$

The first inequality in (1) is proven in the same way. □

定理 5.2

[23][20] *Let $P : \Re^n \rightrightarrows \Re^m$ be a polyhedral set-valued mapping. Then there is a constant λ such that P is locally upper Lipschitz with modulus λ at each $x \in \mathrm{dom}\, P$.*

Proof. Let $G_i, i \in [k]$, be the components of P. With the constant ξ_i associated with G_i according to Lemma 5.2 we put

$$\lambda = \max\{\xi_1, \dots, \xi_k\}.$$

[20] Outrata J, Kočvara M and Zowe J. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Theory, Applications and Numerical Results. Kluwer Academic Publishers, 1998.

Now consider some arbitrary $x \in \operatorname{dom} P$ and the index set

$$J(x) = \{i \in [k] : x \in \pi_1(G_i)\}.$$

By Lemma 5.1 there is a neighborhood U of x such that

$$(U \times \Re^m) \cap \operatorname{ghp} P \subset \bigcup_{i \in J(x)} G_i.$$

For $x' \in U$ with $x' \notin \operatorname{dom} P$ nothing has to be shown. Hence let $x' \in \operatorname{dom} P$ and $y' \in P(x')$. Then we have

$$(x', y') \in [(U \times \Re^m) \cap \operatorname{ghp} P] \subset \bigcup_{i \in J(x)} G_i,$$

which implies $(x', y') \in G_i$ for some $i \in J(x)$.

For this i we get

$$\begin{aligned}
\operatorname{dist}(y', P(x)) &= \operatorname{dist}\left(y', \{v : (x, v) \in \operatorname{ghp} P\}\right) \\
&\leq \operatorname{dist}\left(y', \{v : (x, v) \in G_i\}\right) \\
&= d_y((x, y'), G_i) \leq \xi_i d_x((x, y'), G_i) \\
&= \xi_i \operatorname{dist}\left(x, \{u : (u, y') \in G_i\}\right) \\
&\leq \xi \|x' - x\| \leq \lambda \|x' - x\|.
\end{aligned}$$

Since $P(x)$ is closed and y' was arbitrary in $P(x')$, it follows that

$$P(x') \subset P(x) + \lambda \|x' - x\| \mathbf{B}$$

and we are done. □

NLP模型

非线性规划问题

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & h_i(x) = 0, i = 1, \ldots, m, \\ & g_i(x) \le 0, i = 1, \ldots, p, \end{aligned} \tag{6}$$

其中$f : \Re^n \to \Re$, $h_i : \Re^n \to \Re$, $i = 1, \ldots, m$, $g_i : \Re^n \to \Re$, $i = 1, \ldots, p$是二次连续可微函数.

NLP的KKT条件

问题(6)的Lagrange函数定义为

$$L(x,\zeta,\lambda) = f(x) + \langle \zeta, h(x)\rangle + \langle \lambda, g(x)\rangle.$$

设x是问题(6)的可行点，用$\mathcal{M}(x)$记x点处的乘子集合. 如果$\mathcal{M}(x) \neq \emptyset$, 则 $(\zeta,\lambda) \in \mathcal{M}(x)$意味着 (x,ζ,λ)满足KKT条件

$$\nabla_x L(x,\zeta,\lambda) = 0, \ \ -h(x) = 0, \ \ \lambda \in N_{\Re^p_-}(g(x)). \tag{7}$$

KKT系统的非光滑方程形式

KKT条件(7)可以等价地表示为下述非光滑方程组

$$F(x,\zeta,\lambda) = \begin{bmatrix} \nabla_x L(x,\zeta,\lambda) \\ -h(x) \\ -g(x) + \Pi_{\Re^p_-}(g(x)+\lambda) \end{bmatrix} = 0 \qquad (8)$$

或者

$$\begin{bmatrix} \nabla_x L(x,\zeta,\lambda) \\ -h(x) \\ \lambda - \Pi_{\Re^p_+}(g(x)+\lambda) \end{bmatrix} = 0.$$

KKT系统的广义方程形式

KKT条件(7)也可以等价地表示为下述的广义方程

$$0 \in \begin{bmatrix} \nabla_x L(x,\zeta,\lambda) \\ -h(x) \\ -g(x) \end{bmatrix} + \begin{bmatrix} N_{\Re}^{n}(x) \\ N_{\Re^m}(\zeta) \\ N_{\Re^p_+}(\lambda) \end{bmatrix}. \tag{9}$$

KKT映射S_{KKT}

令$Z = \Re^n \times \Re^m \times \Re^p$, $D = \Re^n \times \Re^m \times \Re^p_+$. 定义

$$\phi(z) = \begin{bmatrix} \nabla_x L(x, \zeta, \lambda) \\ -h(x) \\ -g(x) \end{bmatrix},$$

则广义方程(9)可表示为

$$0 \in \phi(z) + N_D(z).$$

对$\eta \in Z$, 定义

$$S_{\mathrm{KKT}}(\eta) = \{z \in Z : \eta \in \phi(z) + N_D(z)\}. \tag{10}$$

广义方程的法映射(normal map)定义为

$$\mathcal{F}(z) = \begin{bmatrix} \nabla_x L(x, \zeta, y - \Pi_{\Re^p_-}(y)) \\ -h(x) \\ -g(x) + \Pi_{\Re^p_-}(y) \end{bmatrix}. \qquad (11)$$

则$(\bar{x}, \bar{\zeta}, \bar{\lambda})$ 是广义方程(9)的解当且仅当

$$\mathcal{F}(\bar{x}, \bar{\zeta}, \bar{y}) = 0,$$

其中$\bar{y} = \bar{\lambda} + g(\bar{x})$, $\bar{\lambda} = \Pi_{\Re^p_+}(\bar{y})$.

Lipschitz 同胚

引理 5.3

点$(\bar{x}, \bar{\zeta}, \bar{\lambda})$是广义方程(9)的强正则解当且仅当$\mathcal{F}$在$(\bar{x}, \bar{\zeta}, \bar{y})$附近是*Lipschitz* 同胚的.

命题 5.1

设$\bar{x}$是问题(6)的可行点满足$\mathcal{M}(\bar{x}) \neq \emptyset$. 令$(\bar{\zeta}, \bar{\lambda}) \in \mathcal{M}(\bar{x})$, $\bar{y} = \bar{\lambda} + g(\bar{x})$. 考虑下述条件:

(a) 强二阶充分条件在$\bar{x}$成立, 且$\bar{x}$满足线性无关约束规范.

(b) $\partial \mathcal{F}(\bar{x}, \bar{\zeta}, \bar{y})$中的任何元素是非奇异的.

(c) KKT点$(\bar{x}, \bar{\zeta}, \bar{\lambda})$是广义方程(9)的强正则解.

则$(a) \Longrightarrow (b) \Longrightarrow (c)$.

一致二阶增长条件

引理 5.4

设$\bar{x}$是问题(6)的稳定点. 设*MF*约束规范在$\bar{x}$处成立. 如果在$\bar{x}$处关于标准参数化的一致二阶增长条件成立, 则强二阶充分条件在$\bar{x}$处成立.

定理 5.3

设$\bar{x}$是问题(6)的局部最优解. 设*MF*约束规范在$\bar{x}$成立, 从而$\bar{x}$为稳定点. 设$(\bar{\zeta}, \bar{\lambda}) \in \mathcal{M}(\bar{x})$, 那么$(\bar{\zeta}, \bar{\lambda})$满足问题(6)的*KKT*条件. 令$\bar{y} = g(\bar{x}) + \bar{\lambda}$. 则下述条件是等价的:

(a) 强二阶充分条件在$\bar{x}$ 成立且$\bar{x}$满足线性无关约束规范.

(b) $\partial \mathcal{F}(\bar{x}, \bar{\zeta}, \bar{y})$中的任何元素均是非奇异的.

(c) *KKT*点$(\bar{x}, \bar{\zeta}, \bar{\lambda})$是广义方程(9)的强正则解.

(e) 一致二阶增长条件在$\bar{x}$成立且$\bar{x}$满足线性无关约束规范.

Dontchev and Rockafellar 1996

- Consider

$$S(z,w) = \{x : 0 \in z + f(w,x) + N_C(x)\}$$

  where C is a polyhedral convex set. Dontchev and Rockafellar (1996)[21] showed that the strong regularity of S is equivalent to Aubin property of S around a point $(z_0, w_0, x_0) \in \mathrm{ghp}\, S$.
- 非线性规划KKT系统的强正则性等价于Aubin性质.

[21] A.L.Dontchev and R.T. Rockafellar, Characterizations of Strong Regularity for Variational Inequalities over Polyhedral Convex Sets, SIAM J. Optim. 6 (1996), 1087-1105.

KKT映射的稳健孤立平稳性

结果叙述

内容取自Dontchev and Rockafellar(1997)[9][22]. 主要结论为NLP问题的KKT映射的稳健孤立平稳性等价于严格MF约束规范与二阶充分性条件成立.

[22] Dontchev A L and Rockafellar R T. *Characterizations of Lipschitz stability in nonlinear programming*. In: Fiacco AV, editor. Mathematical programming with data perturbations. New York: Marcel Dekker, 1997: 65-82.

$\mathcal{C}^2$参数化扰动问题

考虑下述参数非线性规划问题

$$\min g_0(w,x) + \langle v, x\rangle \quad \text{s.t.} \quad x \in C(u,w), \tag{12}$$

其中$C(u,w)$表示下列约束:

$$g_i(w,x) - u_i \begin{cases} = 0 & i = 1,\dots,r, \\ \leq 0 & i = r+1,\dots,m, \end{cases} \tag{13}$$

其中$g_i : \Re^d \times \Re^n \to \Re$, $i = 0,1,\dots,m$是二次连续可微函数, 向量$w \in \Re^d$, $v \in \Re^n$与$u = (u_1,\dots,u_m)^T \in \Re^m$是参数. 将它们结合起来记为$p = (v,u,w)$, 记$X(p)$为(12)的局部最优解集, 称映射$p \mapsto X(p)$ 为解映射.

Kurash-Kuhn-Tucker条件

- 称$x \in X(p)$是孤立的如果在x的某个邻域U内有$X(p) \cap U = \{x\}$.
- 记$C(p)$为可行集, 称映射$p \mapsto C(p)$为约束映射.
- 定义Lagrange函数

$$L(w, x, y) = g_0(w, x) + \sum_{i=1}^{m} y_i g_i(w, x),$$

- 这一问题的Karush-Kuhn-Tucker条件为

$$\begin{cases} v + \nabla_x L(w, x, y) = 0, \\ -u + \nabla_y L(w, x, y) \in N_Y(y), \end{cases} \tag{14}$$

其中$Y = \Re^r \times \Re_+^{m-r}$.

- 对于给定的$p=(v,u,w)$, KKT系统的解集(x,y)记为$S_{\rm KKT}(p)$, 称映射$p\mapsto S_{\rm KKT}(p)$为KKT映射.
- 记$X_{\rm KKT}(p)$为稳定点集, 即

$$X_{\rm KKT}(p)=\{x|\exists y\ s.t.\ (x,y)\in S_{\rm KKT}(p)\},$$

  称映射$p\mapsto X_{\rm KKT}(p)$为稳定点映射.
- 关于x和p的Lagrange乘子集合记为$Y_{\rm KKT}(x,p)=\{y|(x,y)\in S_{\rm KKT}(p)\}$.

与$(v_0,u_0,w_0,x_0,y_0)\in {\rm gph} S_{\rm KKT}(p)$相联系的$\{1,2,\ldots,m\}$的指标集合$I_1, I_2$ 与I_3定义为

$$\begin{aligned} I_1 &= \{i\in\{r+1,\ldots,m\}\,|\,g_i(w_0,x_0)-u_{0i}=0, y_{0i}>0\}\cup[r],\\ I_2 &= \{i\in\{r+1,\ldots,m\}\,|\,g_i(w_0,x_0)-u_{0i}=0, y_{0i}=0\},\\ I_3 &= \{i\in\{r+1,\ldots,m\}\,|\,g_i(w_0,x_0)-u_{0i}<0, y_{0i}=0\}. \end{aligned}$$

严格Mangasarian-Fromovitz条件

称严格Mangasarian-Fromovitz (MF) 条件在(p_0, x_0)处成立如果存在Lagrange乘子$y_0 \in Y_{\mathrm{KKT}}(x_0, p_0)$使得:

(a) $i \in I_1$中的$\nabla_x g_i(w_0, x_0)$线性无关;

(b) 存在向量$z \in \Re^n$使得$i \in I_1$时$\nabla_x g_i(w_0, x_0)^T z = 0$,
$i \in I_2$时$\nabla_x g_i(w_0, x_0)^T z < 0$.

线性变分不等式

对给定的$p_0 = (v_0, u_0, w_0)$, 设(x_0, y_0)满足KKT条件(14).
记$A = \nabla^2_{xx} L(w_0, x_0, y_0)$, $B = \nabla^2_{yx} L(w_0, x_0, y_0)$,
(14)在$(v_0, u_0, w_0, x_0, y_0)$处的线性化表示为下述线性变分不等式:

$$\begin{cases} v + \nabla_x L(w_0, x_0, y_0) + A(x - x_0) + B^T(y - y_0) = 0, \\ -u + g(w_0, x_0) + B(x - x_0) \in N_Y(y). \end{cases} \tag{15}$$

对任何(u, v), 记所有满足(15)的(x, y)的集合为$L_{\mathrm{KKT}}(u, v)$.

解映射与线性化解映射

- 令$P = \Re^d \times \Re^m$, 考虑映射

$$\Sigma(p) = \{x \in \Re^n | y \in f(w,x) + F(w,x)\}, p = (w,y), \tag{16}$$

其中$f : \Re^d \times \Re^n \to \Re^m$, $F : \Re^d \times \Re^n \rightrightarrows \Re^m$.

- 设对$p_0 = (w_0, y_0) \in P$, $x_0 \in \Sigma(p_0)$,$f(w_0, \cdot)$在x_0处可微,Jacobian阵为$\nabla_x f(w_0, x_0)$. 考虑f的线性化映射:

$$\Lambda(p) = \{x \in \Re^n | y \in f(w_0,x_0) + \nabla_x f(w_0,x_0)(x - x_0) + F(w,x)\}. \tag{17}$$

定理 5.4

[8][23] 假设存在x_0的邻域U, w_0的邻域W与常数l使得对任何$x \in U$, $w \in W$有

$$\|f(w,x) - f(w_0,x)\| \le l\|w - w_0\|. \tag{18}$$

那么下述结论等价:

(i) Λ在(p_0, x_0)处是孤立平稳的;

(ii) Σ在(p_0, x_0)处是孤立平稳的.

[23] Dontchev A L. *Characterization of Lipschitz stability in optimization*. in Recent Developments in Well-Posed Variational Problems, Lucchetti R and Revalski J (eds), 1995, 95-116.

推论 5.1

假设定理5.4中的假设条件成立且设$F : \Re^n \to \Re^m$为多面体映射, 那么下述结论等价:

(i) 存在x_0的邻域U使得

$$[f(w_0, x_0) + \nabla f(w_0, x_0)(\cdot - x_0) + F(\cdot)]^{-1}(y_0) \cap U = \{x_0\};$$

(ii) 映射Σ在(p_0, x_0)处是孤立平稳的.

证明

映射$\Lambda = [f(w_0, x_0) + \nabla f(w_0, x_0)(\cdot - x_0) + F(\cdot)]^{-1}$是多面体，因此由[26]知$\Lambda$在$\Re^m$上是平稳的(是局部上Lipschitz连续的)，则(i)可推出Λ在(y_0, x_0)处是孤立平稳的. 应用定理5.4得Σ在(p_0, x_0)处是孤立平稳的. 再应用定理5.4可得(ii)可推出(i). ■

S_{KKT}的孤立平稳性

由于N_Y是多面体集,对KKT系统(14)应用推论5.1可得结论:

推论 5.2

下述结论等价:

(i) (x_0, y_0)为集合$L_{\mathrm{KKT}}(p_0)$的孤立点;

(ii) 映射S_{KKT}在$(p_0, x_0, y_0) \in \mathrm{gph} S_{\mathrm{KKT}}$处是孤立平稳的.

引理 5.5

假设x_0是当$p = p_0$时(12)的孤立局部极小点, 设(p_0, x_0)处的MF约束条件成立. 那么映射X在(p_0, x_0)处下半连续, 即对任何x_0的邻域U, 存在p_0的邻域V使得对任何$p \in V$, 集合$X(p) \cap U$ 非空.

证明

由[21, 推论4.5][24]可知约束映射C在(w_0, u_0, x_0)处具有Aubin性质当且仅当MF约束条件在(w_0, u_0, x_0) 处成立. 设a, b和γ为映射C的Aubin性质相关常数, 即对$p_1, p_2 \in \mathbb{B}(p_0, b)$,

$$C(p_1) \cap \mathbb{B}(x_0, a) \subset C(p_2) + \gamma(\|p_1 - p_2\|)\mathbb{B}.$$

令U为x_0的任意邻域. 选取$\alpha \in (0, a)$使得x_0是当$p = p_0$时(12)在$\mathbb{B}(x_0, \alpha)$中的唯一极小点且$\mathbb{B}(x_0, \alpha) \subset U$.

[24] Mordukhovich B S. *Lipschitzian stability of constraint systems and generalized equations*. Nonlinear analysis, 1994, **22**: 173-206.

对此固定的α和$p \in \mathbf{B}(p_0, b)$, 考虑映射

$$p \mapsto C_\alpha(p) = \{x \in C(p) : \|x - x_0\| \leq \alpha + \gamma \|p - p_0\|\}.$$

显然映射C_α在$p = p_0$处是上半连续的,**下证其也是下半连续的.** 选取$x \in C_\alpha(p_0) = C(p_0) \cap \mathbf{B}(x_0, \alpha)$. 由$C$的Aubin性质,对任何$p_0$附近的$p$,存在$x_p \in C(p)$使$\|x_p - x\| \leq \gamma \|p - p_0\|$. 则有

$$\|x_p - x_0\| \leq \|x_p - x\| + \|x - x_0\| \leq \alpha + \gamma \|p - p_0\|.$$

因此$x_p \in C_\alpha(p)$且当$p \to p_0$时, $x_p \to x$. 所以C_α在$p = p_0$处是下半连续的.

由于$C_\alpha(p)$是非空紧致的, 则问题

$$\min_x g_0(w,x) + \langle x, v\rangle \ s.t.\ x \in C_\alpha(p) \tag{19}$$

对任何p_0附近的p有解, 并且由α的选择可知x_0是$p = p_0$时此问题的唯一极小点. 由Berge定理[25], (19)的解映射X_α在$p = p_0$处上半连续; 换言之, 对任何$\delta > 0$, 存在$\eta \in (0, b)$使得对任何$p \in \mathrm{B}(p_0, \eta)$, (19)的(全局)最优解集是非空的且包含在$\mathrm{B}(x_0, \delta)$内.

[25] 考虑

$$\mathrm{val}\,(y) = \inf_x \{f(x,y) : x \in A(y)\},$$

$$S(y) = \mathrm{argmin}\{f(x,y) : x \in A(y)\},$$

其中A在y_0处上半连续,在y_0下半连续, $A(y_0)$是非空紧致集合, f在$A(y_0)\times\{y_0\}$的每一点处是连续的,则$\mathrm{val}(y)$在y_0处连续,$S(y)$在y_0处上半连续.

因为$X_\alpha(p_0)=\{x_0\}$, 映射X_α在p_0处连续. 设δ'满足$0<\delta'<\alpha$, 则存在$\eta'>0$ 使得对任何$p\in \mathrm{B}(p_0,\eta')$, 任何解$x\in X_\alpha(p)$满足$\|x-x_0\|\le\delta'<\alpha+\gamma\|p-p_0\|$. 因此, 对$p\in \mathrm{B}(p_0,\eta')$约束$\|x-x_0\|\le\alpha+\gamma\|p-p_0\|$在问题(19)中是无效的. 所以任何$p\in \mathrm{B}(p_0,\eta')$, 有

$$X_\alpha(p)\subset X(p)\cap \mathrm{B}(x_0,\delta').$$

证毕. □

二阶充分性条件

二阶充分性条件在$(p_0, x_0, y_0) \in \mathrm{gph}\, S_{\mathrm{KKT}}$处成立如果$\forall x' \in D\backslash\{0\}$有

$$\langle x', \nabla^2_{xx} L(w_0, x_0, y_0)x' \rangle > 0,$$

其中锥$D = \{x' | \nabla_x g_i(w_0, x_0)x' = 0, i \in I_1; \nabla_x g_i(w_0, x_0)x' \leq 0, i \in I_2\}$.

KKT系统的孤立平稳性

定理 5.5

下述条件等价:

(i) 映射S_{KKT}在$(p_0, x_0, y_0) \in \mathrm{gph}\, S_{\mathrm{KKT}}$处是稳健孤立平稳的, 且$x_0$是问题(12)关于$p_0$的局部最优解;

(ii) 严格MF约束条件和二阶充分性条件在(p_0, x_0, y_0)处成立.

证明

假设(i)成立, 则y_0是$Y_{\rm KKT}(x_0,p_0)$中的孤立点. 注意到$Y_{\rm KKT}(x_0,p_0)$ 是凸的, 则有$Y_{\rm KKT}(x_0,p_0)=\{y_0\}$. 因此严格MF约束条件成立[26]. 进一步, 由推论5.2, 不存在(x_0,y_0)附近的(x,y)满足$(x,y)\in L_{\rm KKT}(p_0)$.[27] 不失一般性, 假设$I_1=\{1,2,\cdots,m_1\}$, $I_2=\{m_1+1,\cdots,m_2\}$且分别记B_1和B_2为B对应指标集I_1和I_2的子矩阵.

---

[26] J. Kyparisis, On uniqueness of Kuhn-Tucker multipliers in nonlinear programming, Math. Programming 32 (1985), 242 – 246.

[27] 下述两个性质等价:

(i) (x_0,y_0)为集合$L_{\rm KKT}(p_0)$的孤立点;

(ii) 映射$S_{\rm KKT}$在$(p_0,x_0,y_0)\in {\rm gph}S_{\rm KKT}$处是孤立平稳的.

那么$(x,y)=(0,0)$为下述变分系统的孤立解:

$$\begin{aligned} &Ax + B^T y = 0, \\ &B_1 x = 0, \\ &B_2 x \le 0, y_i \ge 0, y_i(Bx)_i = 0, i \in [m_1+1, m_2]. \end{aligned} \tag{20}$$

观察到由于$y_{0i} > 0, i \in I_1$, 则对于$i \in I_1$中的y_i符号没有限制. 事实上, 因为(20)的解集是一个锥, 则$(0,0)$为(20)的唯一解. 由x_0处的二阶必要性条件, 可得

$$\langle x', Ax' \rangle \ge 0, \forall x' \in D \backslash \{0\}.$$

只需证上述不等式的等号始终不成立. 假设存在非零向量$x' \in D$使得$Ax' = 0$, 则非零向量$(x', 0)$也为(20)的解, 矛盾.

反之，假设(ii)成立，则x_0是(12)关于p_0的孤立局部解且y_0为相应的唯一乘子. 假设(p_0, x_0)相应的指标集I_1是非空的且$\mathcal{U}$和$\mathcal{W}$分别为x_0和w_0的邻域使得对所有的$x \in \mathcal{U}$, $w \in \mathcal{W}$有$\nabla_x g_i(w, x), i \in I_1$是线性无关的. 由引理5.5, 对$p_0$附近的$p$, $X(p) \cap \mathcal{U} \neq \emptyset$. 那么对所有$p_0$附近的$p$, x_0附近的$x(p) \in X(p)$, 存在接近$y_{0i}, i \in I_1$的$y_i(p), i \in I_1$使得

$$v + \nabla_x g_0(w, x(p)) + \sum_{i \in I_1} y_i(p) \nabla_x g_i(w, x(p)) = 0.$$

注意到$\forall i \in I_1$, $y_i(p) > 0$, 对$i \in I_2 \cup I_3$, 取$y_i(p) = 0$, 得到$y(p) = (y_1(p), \cdots, y_m(p))$是扰动问题的Lagrange乘子且接近$y_0$. 因此, 如果$U$为$(x_0, y_0)$的邻域且$p$充分接近$p_0$, 则有$S_{\mathrm{KKT}}(p) \cap U \neq \emptyset$.

如果$I_1 = \emptyset$, 那么$y_0 = 0 = Y_{\mathrm{KKT}}(x_0, p_0)$. 由引理5.5, 对$x_0$的任何邻域$\mathcal{U}$与充分接近$p_0$的$p$, 有$X(p) \cap \mathcal{U} \neq \emptyset$. 进一步, MF约束条件保证对$p_0$附近的$p$, x_0附近的x, Lagrange 乘子集$Y_{\mathrm{KKT}}(x, p)$非空有界. 假设存在$\alpha > 0$, 序列$p_k \to p_0$和$x_k \to x_0$使得$\forall y \in Y_{\mathrm{KKT}}(x_k, p_k)$, $k = 1, 2, \cdots$有$\|y\| \geq \alpha$. 选取序列$y_k \in Y_{\mathrm{KKT}}(x_k, p_k)$, 则该序列有界, 存在聚点$\bar{y} \neq 0$. 在KKT系统中对$k$取极限可得$\bar{y} \in Y_{\mathrm{KKT}}(x_0, p_0)$, 这表明$Y_{\mathrm{KKT}}(x_0, p_0)$不是单点集, 这与严格MF约束条件矛盾. 因此对$y_0 = 0$的任何邻域$\mathcal{Y}$, 当$p$充分接近$p_0$且$x \in X(p)$充分接近$x_0$时, 有$Y_{\mathrm{KKT}}(x, p) \cap \mathcal{Y} \neq \emptyset$. 那么, 对$(x_0, y_0)$的某邻域$U$及充分接近$p_0$的$p$, 也有$S_{\mathrm{KKT}}(x, p) \cap U \neq \emptyset$.

假设映射S_{KKT}在$(p_0, x_0, y_0) \in \mathrm{gph} S_{\mathrm{KKT}}$处不是孤立平稳的，那么由推论5.2,[28] (20)有非零解(x', y')且该解可与$(0, 0)$无限接近. 假设$y' \in \Re^m$且对$i \in I_3$有$y_i' = 0$. 如果$x' = 0$, 则$y' \neq 0$. 注意到如果对某些$i \in I_2$有$y_i' \neq 0$, 则$y_i' > 0$. 因为对$i \in I_1$有$y_{0i} > 0$, 且y'充分接近0, 向量$y_0 + y'$为关于x_0与p_0的Lagrange乘子. 这与严格MF约束条件矛盾. 因此, $x' \neq 0$, 但是$x' \in D$. 在(20)的第一个方程两边同时乘以x', 得到$\langle x', Ax' \rangle = 0$, 与二阶充分性条件矛盾. 证毕. ■

[28]下述两个性质等价:

(i) (x_0, y_0)为集合$L_{\mathrm{KKT}}(p_0)$的孤立点;

(ii) 映射S_{KKT}在$(p_0, x_0, y_0) \in \mathrm{gph} S_{\mathrm{KKT}}$处是孤立平稳的.

定理 5.6

设MF约束条件在$x_0 \in X_{\mathrm{KKT}}(p_0)$处成立,$p_0 = (v_0, u_0, w_0)$. 那么下述条件是映射$X_{\mathrm{KKT}}$在$(p_0, x_0)$处孤立平稳的充分必要条件: 不存在$x' \neq 0$与某一选择

$$y_0 \in \arg\max\{\langle x', \nabla^2_{xx} L(w_0, x_0, y)x'\rangle | y满足(x_0, y) \in S_{\mathrm{KKT}}(p_0)\}$$

满足目标函数为$h_0(x') = \langle x', \nabla^2_{xx} L(w_0, x_0, y_0)x'\rangle$, 约束条件为

$$\begin{cases} \langle \nabla_x g_0(w_0, x_0) - v_0, x'\rangle = 0, \\ \langle \nabla_x g_i(w_0, x_0), x'\rangle = 0, i \in [1, r], \\ \langle \nabla_x g_i(w_0, x_0), x'\rangle \leq 0, i \in [r+1, m], g_i(w_0, x_0) - u_{0i} = 0. \end{cases}$$

的子问题的KKT条件.

由Levy和Rockafellar[20][29]中的定理3.1和3.2, 满足子问题KKT条件的x'构成了0在与(p_0, x_0)处X_{KKT}相联系的图导数映射下的像. 应用[15]中的命题2.1, 可得该集合中这样的$x' \neq 0$的不存在性与映射X_{KKT}在(p_0, x_0)处是孤立平稳的等价. ■

$$DX_{\mathrm{KKT}}(p_0|x_0)(0) = \{0\}$$

[29] Levy A B and Rockafellar R T. *Sensitivity of Solutions in Nonlinear Programming Problems with Nonunique Multipliers*. in Recent Advances In Nonsmooth Optimization, 1995: 215-223.

基于引理5.5和定理5.6可以得到下述推论:

推论 5.3

设x_0是(12)的孤立局部极小点, 其中$p_0 = (v_0, u_0, w_0)$. 假设MF约束条件在(p_0, x_0)处成立且设定理5.6中的条件成立. 那么(12)的解映射在(p_0, x_0)处是稳健孤立平稳的.

锥优化的稳定性

二阶锥优化稳定性综述

- Bonnans J F, Ramírez (2005): the constraint nondegeneracy and the strong second order sufficient condition $\Longleftrightarrow$ the strong regularity of the solution to the KKT system.
- Wang and Zhang (2009): developed other equivalent conditions, including the nonsingularity of nonsmooth reformulation of the KKT system.
- Zhang et. al (2017): proved that the isolated calmness of the KKT system is equivalent to the second-order sufficiency optimality condition together with the strict Robinson CQ.

二阶锥优化稳定性的参考文献

1 Bonnans J F, Ramírez C H, Perturbation analysis of second order cone programming problems, Mathematical Programming, Series B, 104(2005), 205-227.

2 Wang Y, Zhang L W. Properties of Equation Reformulation of the Karush-Kuhn-Tucker Condition for Nonlinear Second Order Cone Optimization Problems. Mathematical Methods of Operations Research, Math Meth Oper. Res.70 (2009), 195-218.

3 Zhang Y, Zhang L W, Wu J and Wang K D. Characterizations of local upper Lipschitz property of perturbed solutions to nonlinear second-order cone programs. Optimization 66(2017),1079-1103.

SDP 优化的稳定性综述

- Sun (2006): established nine conditions equivalent to the strong regularity of Karush-Kuhn-Tucker (KKT) system.
- Chan and Sun (2008): proved that, for linear SDP, the constraint nondegeneracy for the dual problem is equivalent to the strong second order sufficient condition for the primal problem.
- 丁超,孙德锋,张立卫 (2017)基于孤立平稳性的图导数准则,给出一大类锥约束优化问题Karush-Kuhn-Tucker系统的稳健孤立平稳性的刻画,证明这一稳健孤立平稳性等价于严格Robinson约束规范和二阶充分性最优条件.

SDP稳定性文献

1 D. F. Sun, The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications, Math. Oper. Res., 31(2006), 761-776.

2 Chan Z X, Sun D F, Constraint nondegeneracy, strong regularity and nonsingularity in semidefinite programming, SIAM Journal on Optimization, 19(2008),370-396.

- Zhang Y L and Zhang L W. On the upper Lipschitz property of the KKT mapping for nonlinear semidefinite optimization. Operations Research Letter, 44 (2016), 474-478.
- Ding C, Sun D F, Zhang L W. Characterization of the robust isolated calmness for a class of conic programming problems, SIAM Journal on Optimization, 27(2017), 67-90.

线性SDP的强正则性的精彩结果

- Linear SDP problem:

$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \mathcal{A}X = b, \\ & X \in \mathbb{S}^n_+, \end{array} \tag{1}$$

- The dual of SDP problem (1):

$$\begin{array}{ll} \max & b^T y \\ \text{s.t.} & \mathcal{A}^* y + S = C, \\ & S \in \mathbb{S}^n_+. \end{array} \tag{2}$$

Constraint non-degeneracy

- The constraint non-degeneracy of the primal SDP (1)

$$\mathcal{A}\,\mathrm{lin}(T_{\mathbb{S}^n_+}(\overline{X})) = \Re^m, \tag{3}$$

- The dual problem (2) satisfies the constraint non-degeneracy at $(\overline{y}, \overline{S}) \in \Re^m \times \mathbb{S}^n_+$

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{I} \\ 0 & \mathcal{I} \end{bmatrix} \begin{pmatrix} \Re^m \\ \mathbb{S}^n \end{pmatrix} + \begin{bmatrix} \{0\} \\ \mathrm{lin}(T_{\mathbb{S}^n_+}(\overline{S})) \end{bmatrix} = \begin{bmatrix} \mathbb{S}^n \\ \mathbb{S}^n \end{bmatrix} \tag{4}$$

or equivalently,

$$\mathcal{A}^*\Re^m + \mathrm{lin}(T_{\mathbb{S}^n_+}(\overline{S})) = \mathbb{S}^n. \tag{5}$$

定义 6.1

Let $\overline{X} \in \mathbb{S}^n_+$ be an optimal solution to the SDP problem (1). The strong second-order sufficiency optimality holds at $\overline{X}$ if

$$\sup_{(y,S)\in\mathcal{M}(\overline{X})} \{-\Upsilon_{\overline{X}}(-S,H)\} > 0, \ \ \forall 0 \neq H \in \left\{ \bigcap_{(y,S)\in\mathcal{M}(\overline{X})} \mathrm{app}(y,S) \right\}.$$

Interesting result

命题 6.1

[30] *Let* $\overline{X} \in \mathbb{S}^n_+$ *be an optimal solution to the SDP problem (1). Under the condition* $\mathcal{M}(\overline{X}) = \{(\overline{y}, \overline{S})\}$*, the following two properties are equivalent:*

(i) *The strong second-order sufficiency optimality holds at* $\overline{X}$*;*

(ii) *Dual constraint non-degeneracy condition (5) holds at* $(\overline{y}, \overline{S})$*.*

[30] Chan Z X, Sun D F, Constraint nondegeneracy, strong regularity and nonsingularity in semidefinite programming, SIAM Journal on Optimization, 19(2008),370-396.

非线性SDP问题

非线性SDP问题[31]

Consider the upper Lipshitz continuity of KKT mapping for nonlinear SDP problem

$$\min\ f(x) \quad \text{s.t.} \quad x \in \Phi, \tag{6}$$

$$\Phi = \{x \in \Re^n : G(x) \preceq 0, h(x) = 0, g(x) \leq 0\},$$

where $f : \Re^n \to \Re, h : \Re^n \to \Re^l, g : \Re^n \to \Re^m, G : \Re^n \to \mathbb{S}^p$ are twice continuously differentiable.

[31] Main results are taken from Y.L. Zhang and L.W. Zhang. On the upper Lipschitz property of the KKT mapping for nonlinear semidefinite optimization. *Operations Research Letter* 44 (2016) 474–478.

分析的关键

If $G : \mathcal{X} \to \mathcal{Y}$ is a locally Lipschitz mapping near a given x_0, then for $S(x) = \{G(x)\}$, we write $\mathbf{d}G(x_0)(\cdot) := \mathbf{d}S(x_0; G(x_0))(\cdot)$. For $G \in C^{0,1}$, we have

$$DG(x_0)(u) = \left\{ v : \exists t_k \searrow 0,\, v = \lim_{k\to\infty} \frac{G(x_0 + t_k u) - G(x_0)}{t_k} \right\}.$$

If $G(x_0) = 0$, then for $T(\delta) = \{x : G(x) = \delta\}$,$(0, x_0) \in \mathrm{ghp}\, T$, and

$$\begin{array}{c} T \text{ is locally upper Lipschitz at } (0, x_0) \text{ if and only if} \\ 0 \in \mathbf{d}G(x_0)(u) \text{ implies } u = 0. \end{array} \tag{7}$$

Define

$$\mathcal{G}(x) = (h(x), g(x), G(x)),$$

$$\mathcal{K} = \{0_l\} \times \Re^m_- \times \mathbb{S}^p_-.$$

Then Problem (6) is equivalently expressed as

$$\begin{array}{ll} \min & f(x) \\ s.t. & \mathcal{G}(x) \in \mathcal{K}. \end{array} \tag{8}$$

If $\overline{x}$ is a local minimizer of (6) and Robinson constraint qualification holds at $\overline{x}$, then there exist $\overline{\mu} \in \Re^l, \overline{\lambda} \in \Re^m$ and $\overline{\Gamma} \in \mathbb{S}^p$ such that the following Krush-Kuhn-Tucker conditions are satisfied

$$\left\{ \begin{array}{l} \nabla_x L(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{\Gamma}) = 0_n, h(\overline{x}) = 0_l, \\ 0_m \geq g(\overline{x}) \perp \overline{\lambda} \geq 0_m, 0 \succeq G(\overline{x}) \perp \overline{\Gamma} \succeq 0. \end{array} \right. \tag{9}$$

For $\mathcal{Y} = \Re^l \times \Re^m \times \mathbb{S}^p$, the strict Robinson constraint qualification at $\overline{x}$ with respect to $\overline{\omega} = (\overline{\mu}, \overline{\lambda}, \overline{\Gamma})$ is defined by

$$\mathrm{D}\mathcal{G}(\overline{x})\Re^n + T_{\mathcal{K}}(\mathcal{G}(\overline{x})) \cap \overline{\omega}^{\perp} = \mathcal{Y}, \tag{10}$$

It follows from Bonnans and Shapiro [3] that if $\overline{x}$ is a local minimizer for Problem (6) at which the strict constraint qualification holds, then the set of Lagrange multipliers $\Lambda(\overline{x})$ is a singleton.

Let us introduce the notations:

$$I_+ = \{i : g_i(\overline{x}) = 0, \overline{\lambda_i} > 0, i = 1, \dots, m\},$$

$$I_0 = \{i : g_i(\overline{x}) = 0, \overline{\lambda_i} = 0, i = 1, \dots, m\},$$

$$I_- = \{i : g_i(\overline{x}) < 0, \overline{\lambda_i} = 0, i = 1, \dots, m\},$$

$$\alpha = \{i : \lambda_i(G(\overline{x})) = 0, \lambda_i(\overline{\Gamma}) > 0, i = 1, \dots, p\},$$

$$\beta = \{i : \lambda_i(G(\overline{x})) = 0, \lambda_i(\overline{\Gamma}) = 0, i = 1, \dots, p\},$$

$$\gamma = \{i : \lambda_i(G(\overline{x})) < 0, \lambda_i(\overline{\Gamma}) > 0, i = 1, \dots, p\}.$$

Let $\overline{Y} = G(\bar{x}) + \overline{\Gamma}$ have the following spectral decomposition

$$\overline{Y} = P\Lambda P^T,$$

where

$$\Lambda = \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0_\beta & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix},$$

in which $\Lambda_\alpha \succ 0$ and $\Lambda_\gamma \prec 0$. Denote

$$P = \begin{bmatrix} P_\alpha & P_\beta & P_\gamma \end{bmatrix},$$

then $G(\bar{x}) = P_\gamma \Lambda_\gamma P_\gamma^T$ and $\overline{\Gamma} = P_\alpha \Lambda_\alpha P_\alpha^T$.

引理 6.1

Let $\bar{x}$ be a feasible point of problem (6) at which the strict Robinson constraint qualification holds. Then

(i) *the set of vectors*

$$\nabla h_j(\bar{x}), j = 1, \ldots, l,$$

$$\nabla g_i(\bar{x}), i \in I_+,$$

and

$$\left\{ v_{ij} : \begin{array}{l} i \in \alpha, j \in \alpha, i \leq j \\ \text{or } i \in \alpha, j \in \gamma \end{array} \right\}$$

are linearly independent,

where

$$v_{ij} = \begin{bmatrix} p_i^T \dfrac{\partial G}{\partial x_1}(\overline{x})p_j \\ \vdots \\ p_i^T \dfrac{\partial G}{\partial x_n}(\overline{x})p_j \end{bmatrix}.$$

(ii) there exists a vector $d_0 \in \Re^n$ such that $\mathcal{J}h(\overline{x})d_0 = 0$, $\mathcal{J}g_{I_+}(\overline{x})d_0 = 0, \langle v_{ij}, d_0 \rangle = 0$, for $i \in \alpha, j \in \alpha, i \leq j$ or $i \in \alpha, j \in \beta$, and

$$\mathcal{J}g_{I_0}(\overline{x})d_0 < 0, \sum_{k=1}^{n} [d_0]_k P_\beta^T \frac{\partial G}{\partial x_k}(\overline{x})P_\beta < 0.$$

Let $\overline{x} \in \Phi$ be a feasible point. The critical cone of Problem (6) at $\overline{x}$ is defined by

$$\mathcal{C}(\overline{x}) = \{d \in \Re^n : \mathrm{D}\mathcal{G}(\overline{x})d \in T_{\mathcal{K}}(\mathcal{G}(\overline{x})), \nabla f(\overline{x})^T d \leq 0\}.$$

定义 6.2

(The second-order sufficient optimality conditions) Let $\overline{x}$ be a stationary point at which $\Lambda(\overline{x}) \neq \emptyset$. If, for any $d \in \mathcal{C}(\overline{x}) \setminus \{0\}$, one has for $H = \mathrm{D}G(\overline{x})d$,

$$\sup_{(\overline{\mu},\overline{\lambda},\overline{\Gamma})\in\Lambda(\overline{x})} \{\langle d, \nabla^2_{xx}L(\overline{x},\overline{\mu},\overline{\lambda},\overline{\Gamma})d\rangle - 2\langle\overline{\Gamma}, HG(\overline{x})^{\dagger}H\rangle\} > 0,$$

then we say that the second-order sufficient optimality conditions hold at $\overline{x}$.

Canonical perturbation

We consider the canonical perturbation of Problem (6)

$$(\mathrm{P}_\delta) \quad \begin{cases} \min & f(x) - \langle \delta_0, x \rangle \\ \text{s.t.} & G(\overline{x}) - \delta_G \preceq 0, \\ & h(x) - \delta_h = 0_l, \\ & g(x) - \delta_g \leq 0_m. \end{cases} \tag{11}$$

KKT mapping

The Karush-Kuhn-Tucker conditions for Problem (P_δ) are the following conditions

$$\left\{\begin{array}{l} \nabla_x L(x,\mu,\lambda,\Gamma) = \delta_0, \\ h(x) = \delta_h, \\ 0_m \geq g(x) - \delta_g \perp \lambda \geq 0, \\ 0 \preceq \Gamma \perp G(x) - \delta_G \preceq 0. \end{array}\right. \tag{12}$$

The set of all $(x,\mu,\lambda,\Gamma) \in \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p$ satisfying (12) are denoted by $S_{\mathrm{KKT}}(\delta)$ for $\delta = (\delta_0, \delta_h, \delta_g, \delta_G)$.

It is obvious that, if $\overline{x}$ is a local minimizer with $(\overline{\mu}, \overline{\lambda}, \overline{\Gamma}) \in \Lambda(\overline{x}) \neq \emptyset$ at $\overline{x}$, then

$$S_{\rm KKT}(0) = \{(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{\Gamma})\}.$$

Define $\overline{y} = \overline{\lambda} + g(\overline{x})$, $\overline{Y} = G(\overline{x} + \overline{\Gamma})$ and the Kojima mapping $F : \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p \to \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p$ by

$$F(x, \mu, y, Y) = \begin{bmatrix} \nabla_x L(x, \mu, y_+, Y_+) \\ h(x) \\ g(x) - y_- \\ G(x) - Y_- \end{bmatrix}, \tag{13}$$

where $y_+ = \max[y, 0], y_- = \min[y, 0], Y_+ = \Pi_{\mathbb{S}^p_+}(Y)$ and $Y_- = \Pi_{\mathbb{S}^p_-}(Y)$.

The mapping $F(x, \mu, y, Y)$ can be expressed as

$$F(x, \mu, y, Y) = M(x) \circ N(\mu, y, Y), \tag{14}$$

where

$$M(x) = \begin{bmatrix} \nabla f(x) & \nabla h(x) & \nabla g(x) & \mathrm{D}G(x)^* & 0 & 0 \\ h(x) & 0 & 0 & 0 & 0 & 0 \\ g(x) & 0 & 0 & 0 & -I_m & 0 \\ G(x) & 0 & 0 & 0 & 0 & -\mathcal{I} \end{bmatrix}$$

and

$$N(\mu, y, Y) = \begin{bmatrix} 1 \\ \mu \\ y_+ \\ Y_+ \\ y_- \\ Y_- \end{bmatrix}.$$

Then we have that $(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{\Gamma})$ is a KKT point of Problem (6) if and only if $F(\overline{x}, \overline{\mu}, \overline{y}, \overline{Y}) = 0$. And $(x, \mu, \lambda, \Gamma)$ is a KKT point of Problem (P_δ) if and only if $F(x, \mu, y, Y) = \delta$ with $\lambda = y_+$ and $\Gamma = Y_+$. In this notation,

$$S_{\mathrm{KKT}}(\delta) = \left\{ \begin{array}{l} (x, \mu, y_+, Y_+) \in \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p : \\ F(x, \mu, y, Y) = \delta \end{array} \right\}.$$

Upper Lipschitz property

定理 6.1

Let $\overline{x} \in \Phi$ *be a feasible point for Problem (6) at which* $\Lambda(\overline{x}) \neq \emptyset$, *the second-order sufficient optimality conditions hold and the strict Robinson constraint qualification holds for some* $(\overline{\mu}, \overline{\lambda}, \overline{\Gamma}) \in \Lambda(\overline{x})$. *Then* S_{KKT} *is upper Lipschitz continuous at* $\delta = (0, 0, 0, 0) \in \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p$.

Proof

Since $(x, \mu, y, Y) \to (x, \mu, y_+, Y_+)$ is Lipschitz continuous, we obtain that $S_{\rm KKT}$ is upper Lipschitz continuous at $\delta = (0,0,0,0)$ if $\widehat{S}_{\rm KKT}$ is upper Lipschitz continuous at $\delta = (0,0,0,0)$, where

$$\widehat{S}_{\rm KKT}(\delta) = \{(x, \mu, y, Y) \in \Re^n \times \Re^l \times \Re^m \times \mathbb{S}^p : F(x, \mu, y, Y) = \delta\}.$$

We use the relation (7) to establish the upper Lipschitz continuity of $\widehat{S}_{\rm KKT}$ at $\delta = (0,0,0,0)$, or we only need to check $\mathbf{d}\widehat{S}_{\rm KKT}(0)(0) = \{0\}$ under the given assumptions. It is easy to obtain

$$\begin{aligned} &\mathbf{d}\widehat{S}_{\rm KKT}(0)(0) \\ &= \{(\triangle x, \triangle \mu, \triangle y, \triangle Y) : 0 \in \mathbf{d}F(\overline{x}, \overline{\mu}, \overline{y}, \overline{Y})(\triangle x, \triangle \mu, \triangle y, \triangle Y)\}. \end{aligned} \tag{15}$$

Now we calculate the strict graphical derivative $\mathbf{d}F(\overline{x},\overline{\mu},\overline{y},\overline{Y})(\triangle x,\triangle\mu,\triangle y,\triangle Y)$. From the definition $F(x,\mu,y,Y)=M(x)\circ N(\mu,y,Y)$, we obtain

$$
\begin{aligned}
&\mathbf{d}F(\overline{x},\overline{\mu},\overline{y},\overline{Y})(\triangle x,\triangle\mu,\triangle y,\triangle Y)\\
&=\mathbf{d}M(\overline{x})(\triangle x)\circ N(\overline{\mu},\overline{y},\overline{Y})+M(\overline{x})\mathbf{d}N(\overline{\mu},\overline{y},\overline{Y})(\triangle\mu,\triangle y,\triangle Y)
\end{aligned}
$$

$$
=\left\{\begin{array}{l}
\nabla^2_{xx}L(\overline{x},\overline{\mu},\overline{\lambda},\overline{Y})(\Delta x)\\
+\nabla h(\overline{x})\triangle\mu+\nabla g(\overline{x})u+\mathrm{D}G(\overline{x})^*U\\
\mathcal{J}h(\overline{x})\triangle x,\ u\in\mathbf{d}[\overline{y}]_+(\triangle y),\ v\in\mathbf{d}[\overline{y}]_-(\triangle y)\\
\mathcal{J}g(\overline{x})\triangle x-v,\ U\in\mathbf{d}[\overline{Y}]_+(\triangle Y),\ V\in\mathbf{d}[\overline{Y}]_-(\triangle Y)\\
\mathrm{D}G(\overline{x})\triangle x-V
\end{array}\right\}.
$$

Since $y \mapsto [y]_+, y \mapsto [y]_-, Y \mapsto [Y]_+$ and $Y \mapsto [Y]_-$ are Lipschitz continuous and directionally differentiable, we obtain that

$$u \in \mathbf{d}[\overline{y}]_+(\triangle y) \text{ if and only if } u = \Pi'_{\Re^m_+}(\overline{y}; \triangle y),$$

$$v \in \mathbf{d}[\overline{y}]_-(\triangle y) \text{ if and only if } v = \Pi'_{\Re^m_-}(\overline{y}; \triangle y),$$

$$U \in \mathbf{d}[\overline{Y}]_+(\triangle Y) \text{ if and only if } U = \Pi'_{\mathbb{S}^p_+}(\overline{Y}; \triangle Y),$$

$$V \in \mathbf{d}[\overline{Y}]_-(\triangle Y) \text{ if and only if } V = \Pi'_{\mathbb{S}^p_-}(\overline{Y}; \triangle Y).$$

It is easy to get

$$u_i = \begin{cases} \triangle y_i, & i \in I_+, \\ [\triangle y_i]_+, & i \in I_0, \\ 0, & i \in I_-, \end{cases} \qquad v_i = \begin{cases} 0, & i \in I_+, \\ [\triangle y_i]_-, & i \in I_0, \\ \triangle y_i, & i \in I_-. \end{cases}$$

$$U = P \begin{bmatrix} P_\alpha^T \triangle Y P_\alpha & P_\alpha^T \triangle Y P_\beta & P_\alpha^T \triangle Y P_\gamma \circ \Omega_{\alpha\gamma} \\ P_\beta^T \triangle Y P_\alpha & \Pi_{S_+^{|\beta|}}(P_\beta^T \triangle Y P_\beta) & 0 \\ P_\gamma^T \triangle Y P_\alpha \circ \Omega_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^T$$

and

$$V = P \begin{bmatrix} 0 & 0 & P_\alpha^T \triangle Y P_\gamma \circ \overline{\Omega}_{\alpha\gamma} \\ 0 & \Pi_{S_-^P}(P_\beta^T \triangle Y P_\beta) & P_\beta^T \triangle Y P_\gamma \\ P_\gamma^T \triangle Y P_\alpha \circ \overline{\Omega}_{\gamma\alpha} & P_\gamma^T \triangle Y P_\beta & P_\gamma^T \triangle Y P_\gamma \end{bmatrix} P^T,$$

where

$$\Omega_{ij} = \frac{[\lambda_i]_+ + [\lambda_j]_+}{|\lambda_i| + |\lambda_j|}, i \in \alpha, j \in \gamma \text{ or } i \in \gamma, j \in \alpha,$$

$$\overline{\Omega}_{ij} = 1 - \Omega_{ij}, i \in \alpha, j \in \gamma \text{ or } i \in \gamma, j \in \alpha.$$

Noting that

$$\widehat{S}_{\mathrm{KKT}}(\delta_0, \delta_h, \delta_g, \delta_G)$$
$$= \{(x, \mu, y, Y) : F(x, \mu, y, Y) = (\delta_0, \delta_h, \delta_g, \delta_G)\},$$

one has that

$$\mathbf{d}\widehat{S}_{\mathrm{KKT}}(0)(0)$$
$$= \{(\Delta x, \Delta \mu, \Delta y, \Delta Y) : 0 \in \mathbf{d}F(\overline{x}, \overline{\mu}, \overline{y}, \overline{Y})(\Delta x, \Delta \mu, \Delta y, \Delta Y)\}.$$

Therefore, if $(\Delta x, \Delta\mu, \Delta y, \Delta Y) \in \mathbf{d}\widehat{S}_{\rm KKT}(0)(0)$, then

$$
\begin{cases}
0 = \nabla^2_{xx} L(\overline{x}, \overline{\mu}, \overline{y}, \overline{Y})(\Delta x) + \nabla h(\overline{x})\Delta\mu \\
\quad +\nabla g(\overline{x})\Pi'_{\Re^m_+}(\overline{y}; \Delta y) + {\rm D}G(\overline{x})^*\Pi'_{\mathbb{S}^p_+}(\overline{Y}; \Delta Y), \\
0 = \mathcal{J} h(\overline{x})\Delta x, \\
0 = \mathcal{J} g(\overline{x})\Delta x - \Pi'_{\Re^m_-}(\overline{y}; \Delta y), \\
0 = {\rm D}G(\overline{x})\Delta x - \Pi'_{\mathbb{S}^p_-}(\overline{Y}; \Delta Y).
\end{cases}
\tag{16}
$$

From the expressions of $\Pi'_{\Re^m_-}(\overline{y};\Delta y)$ and $\Pi'_{\mathbb{S}^p_-}(\overline{Y};\Delta Y)$, we have from the last two lines of (16) that

$$\begin{aligned} \nabla g_i(\overline{x})^T\Delta x &= 0, i \in I_+, \\ \nabla g_i(\overline{x})^T\Delta x &= [\Delta y_i]_-, i \in I_0, \\ \nabla g_i(\overline{x})^T\Delta x &= \Delta y_i, i \in I_-, \end{aligned} \tag{17}$$

and

$$\begin{aligned} & P^T(\mathrm{D}G(\overline{x})\triangle x)P \\ = & \begin{bmatrix} 0 & 0 & P_\alpha^T\triangle Y P_\gamma \circ \overline{\Omega}_{\alpha\gamma} \\ 0 & \Pi_{\mathbb{S}^{|\beta|}_-}(P_\beta^T\triangle Y P_\beta) & P_\beta^T\triangle Y P_\gamma \\ P_\gamma^T\triangle Y P_\alpha \circ \overline{\Omega}_{\gamma\alpha} & P_\gamma^T\triangle Y P_\beta & P_\gamma^T\triangle Y P_\gamma \end{bmatrix}. \end{aligned} \tag{18}$$

From (18), we have

$$
\begin{aligned}
&P_\alpha^T(\mathrm{D}G(\overline{x})\Delta x)[P_\alpha P_\beta] = 0, \\
&P_\beta^T(\mathrm{D}G(\overline{x})P_\beta = \Pi_{\mathbb{S}_-^{|\beta|}}(P_\beta^T \triangle Y P_\beta), \\
&P_\alpha^T(\mathrm{D}G(\overline{x})\Delta x)P_\gamma = P_\alpha^T \triangle Y P_\gamma \circ \overline{\Omega}_{\alpha\gamma}, \\
&P_\gamma^T(\mathrm{D}G(\overline{x})\Delta x)[P_\beta P_\gamma] = P_\gamma^T(\triangle Y)'[P_\beta P_\gamma],
\end{aligned} \tag{19}
$$

frow which and combining the second line of (16) as well as (17), we obtain

$$\begin{cases} \mathcal{J} h(\overline{x})\triangle x = 0, \\ \mathcal{J} g_{I_+}(\overline{x}) = 0, \\ \mathcal{J} g_{I_0}(\overline{x}) = 0, \\ P_\alpha^T[\mathrm{D}G(\overline{x})\triangle x][P_\alpha P_\beta] = 0, \\ P_\beta^T(\mathrm{D}G(\overline{x})\triangle x)P_\beta \preceq 0. \end{cases} \tag{20}$$

From (20), we obtain $\triangle x \in T_\Phi(\overline{x})$, because

$$
\begin{aligned}
& T_\Phi(\overline{x}) \\
& = \{d : \mathcal{J} h(\overline{x})d = 0, \mathcal{J} g_{I_+ \cup I_0}(\overline{x})d \leq 0, P_{\alpha\cup\beta}^T[\mathrm{D}G(\overline{x})d]P_{\alpha\cup\beta} \preceq 0\},
\end{aligned}
$$

under Robinson constraint qualification, which is implied by the strict Robinson constraint qualification.

Nothing that

$$\langle \overline{\lambda}, \mathcal{J} g(\overline{x}) \triangle x \rangle = 0,$$

$$\langle \overline{\Gamma}, \mathrm{D} G(\overline{x}) \triangle x \rangle = \left\langle \begin{bmatrix} \Lambda_\alpha & & \\ & 0 & \\ & & 0 \end{bmatrix}, P^T (\mathrm{D} G(\overline{x}) \triangle x) P \right\rangle = 0,$$

we obtain that $\triangle x \in \mathcal{C}(\overline{x})$, where $\mathcal{C}(\overline{x})$ is the critical cone is defined by

$$\mathcal{C}(\overline{x}) = \left\{ d \in \Re^n : \begin{array}{c} \mathcal{J} h(\overline{x}) d = 0, \mathcal{J} g(\overline{x}) d \in T_{\Re^m_-}(g(\overline{x})) \\ \mathrm{D} G(\overline{x}) d \in T_{\mathbb{S}^p_-}(G(\overline{x})), \langle \overline{\lambda}, g(\overline{x}) \rangle = 0 \\ \langle \overline{\Gamma}, \mathrm{D} G(\overline{x}) d \rangle = 0 \end{array} \right\}.$$

Premultiplying (16) by $\triangle x^T$, we obtain

$$
\begin{aligned}
0 = & \langle \triangle x, \nabla^2_{xx} L(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\
& + \langle \mathcal{J} g(\overline{x}) \triangle x, \Pi'_{\Re^m_+}(\overline{y}; \triangle y) \rangle \\
& + \langle \mathrm{D} G(\overline{x}) \triangle x, \Pi'_{\mathbb{S}^p_+}(\overline{Y}; \triangle Y) \rangle \\
= & \langle \triangle x, \nabla^2_{xx} L(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\
& + \langle \mathrm{D} G(\overline{x}) \triangle x, \Pi'_{\mathbb{S}^p_+}(\overline{Y}; \triangle Y) \rangle \\
= & \langle \triangle x, \nabla^2_{xx} L(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\
& + \langle \Pi_{\mathbb{S}^{|\beta|}_-}(P_\beta^T \triangle Y P_\beta), \Pi_{\mathbb{S}^{|\beta|}_+}(P_\beta^T \triangle Y P_\beta) \rangle \\
& + 2 \langle P_\alpha^T (\mathrm{D} G(\overline{x}) \triangle x) P_\gamma, P_\alpha^T \Pi'_{\mathbb{S}^p_+}(\overline{Y}; \triangle Y) P_\gamma \rangle \\
= & \langle \triangle x, \nabla^2_{xx} L(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\
& + 2 \langle P_\alpha^T (\mathrm{D} G(\overline{x}) \triangle x) P_\gamma, P_\alpha^T \triangle Y P_\gamma \circ \Omega_{\alpha\gamma} \rangle .
\end{aligned}
\tag{21}
$$

From the third line of (19), we obtain

$$P_\alpha^T \triangle Y P_\gamma = P_\alpha^T (\mathrm{D}G(\overline{x})\triangle x) P_\gamma \circ T_{\alpha\gamma},$$

where $T_{\alpha\gamma} = (T_{ij} : i \in \alpha, j \in \gamma)$ is defined by

$$T_{ij} = \frac{1}{\overline{\Omega}_{ij}} = \frac{|\lambda_i| + |\lambda_j|}{|\lambda_j|} = \frac{|\lambda_i| + |\lambda_j|}{-|\lambda_j|}, i \in \alpha, j \in \gamma.$$

Therefore, we obtain from (21) that

$$\begin{aligned} 0 &= \langle \triangle x, \nabla^2_{xx}(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\ &\quad +2\langle P_\alpha^T (\mathrm{D}G(\overline{x})\triangle x) P_\gamma, P_\alpha^T (\mathrm{D}G(\overline{x})\triangle x) P_\gamma \circ \Omega_{\alpha\gamma} \circ T\alpha\gamma \rangle \\ &= \langle \triangle x, \nabla^2_{xx}(\overline{x}, \overline{\mu}, \overline{\lambda}, \overline{Y}) \triangle x \rangle \\ &\quad -2\langle \overline{\Gamma}, (\mathrm{D}G(\overline{x})\triangle x) G(\overline{x})^\dagger (\mathrm{D}G(\overline{x})\triangle x) \rangle. \end{aligned}$$

Then it follows from the second-order sufficient optimality conditions that $\triangle x = 0$. And, from (16), we have

$$
\begin{aligned}
0 &= \nabla h(\bar{x})\nabla\mu + \nabla g(\bar{x})\Pi'_{\Re^m_+}(\bar{y};\triangle y) \\
&\quad + \mathrm{D}G(\bar{x})^*\Pi_{\mathbb{S}^p_+}(\overline{Y};\triangle Y), \\
0 &= \Pi'_{\Re^m_+}(\bar{y};\triangle y) - \triangle y, \\
0 &= \Pi'_{\mathbb{S}^p_+}(\overline{Y};\triangle Y) - \triangle Y.
\end{aligned}
\tag{22}
$$

Nothing $\triangle y = \Pi'_{\Re^m_+}(\overline{y}; \triangle y)$ implies $\triangle y_{I_-}$ and $\triangle y_{I_0} \geq 0$ and $\triangle Y = \Pi'_{\mathbb{S}^p_+}(\overline{Y}; \triangle Y)$ implies

$$P_\alpha^T \triangle Y P_\gamma = 0, P_\beta^T \triangle Y P_\gamma = 0, P_\gamma^T \triangle Y P_\gamma = 0, P_\beta^T \triangle Y P_\beta \succeq 0,$$

we have from (22) that

$$\begin{cases} \nabla h(\overline{x})\triangle\mu + \nabla g_{x_0}(\overline{x}) y_{I_0} + \nabla g_{I_+}(\overline{x}) y_{I_+} \\ +\mathrm{D}(P_\alpha^T G P_\alpha)(\overline{x})^*(P_\alpha^T \triangle Y P_\alpha) + 2\mathrm{D}(P_\alpha^T G P_\beta)(\overline{x})^*(P_\alpha^T \triangle Y P_\beta) \\ +\mathrm{D}(P_\beta^T G P_\beta)(\overline{x})^*(P_\beta^T \triangle Y P_\beta) = 0, \\ \triangle y_{I_0} \geq 0, P_\beta^T \triangle Y P_\beta \succeq 0. \end{cases} \tag{23}$$

From 6.1, we obtain from (23) that

$$\triangle\mu = 0, \triangle y_{I_0 \cup I_+} = 0, P_{\alpha\cup\beta}^T \triangle Y P_{\alpha\cup\beta} = 0,$$

combining with $\triangle x = 0$ and the third line of (17), the third line and the fourth line of (19), implying
$\triangle x = 0, \triangle\mu = 0, \triangle y = 0, \triangle Y = 0.$
Therefore, we obtain the equality as follows

$$\mathbf{d}\widehat{S}_{\mathrm{KKT}}(0)(0) = \{0\}.$$

The proof is completed. □

$\mathcal{C}^2$-简约锥优化问题

Problem[32]

Consider the following canonically perturbed optimization problem:

$$\begin{array}{ll} \min & f(x) - \langle a, x \rangle \\ \text{s.t.} & G(x) + b \in \mathcal{K}, \end{array} \qquad (24)$$

where $f : \mathcal{X} \to \Re$ and $G : \mathcal{X} \to \mathcal{Y}$ are twice continuously differentiable functions, $\mathcal{K} \subset \mathcal{Y}$ is a nonempty closed convex set, and $(a, b) \in \mathcal{X} \times \mathcal{Y}$ is the perturbation parameter.

[32] Results are taken from C. Ding, D. F. Sun, L. W. Zhang. Characterization of the robust isolated calmness for a class of conic programming problems, SIAM Journal on Optimization, 27, 67-90 (2017).

Notations

Let $(a, b) \in \mathcal{X} \times \mathcal{Y}$ be given.

- Let $L : \mathcal{X} \times \mathcal{Y} \to \Re$ be the Lagrangian function

$$L(x; y) := f(x) + \langle y, G(x) \rangle, \quad (x, y) \in \mathcal{X} \times \mathcal{Y}. \tag{25}$$

- $L'_x(x; y)$: the derivative of $L(\cdot; y)$ at $x \in \mathcal{X}$; $\nabla_x L(x; y)$: the adjoint of $L'_x(x; y)$.
- The Karush-Kuhn-Tucker (KKT) optimality condition:

$$\begin{cases} a = \nabla_x L(x; y), \\ b \in -G(x) + \partial \sigma(y, \mathcal{K}), \end{cases} \tag{26}$$

- $S_{\rm KKT}(a, b)$: the set of all solutions (x, y) to the KKT system (26).

SRCQ and constraint non-degeneracy

- The SRCQ is said to hold for problem (24) with $(a, b) = (0, 0)$ at $\bar{x}$ with respect to $\bar{y} \in M(\bar{x}, 0, 0) \neq \emptyset$ if

$$G'(\bar{x})\mathcal{X} + \mathcal{T}_{\mathcal{K}}(G(\bar{x})) \cap \bar{y}^{\perp} = \mathcal{Y}. \tag{27}$$

- The constraint non-degeneracy is said to hold at $\bar{x}$ if

$$G'(\bar{x})\mathcal{X} + \mathrm{lin}\left(\mathcal{T}_{\mathcal{K}}(G(\bar{x}))\right) = \mathcal{Y}. \tag{28}$$

C^2-cone reducibility

定义 6.3 (Definition 3.135[33])

The closed convex set $\mathcal{K}$ is said to be C^2-cone reducible at $\overline{A} \in \mathcal{K}$, if there exist an open neighborhood $\mathcal{W} \subset \mathcal{Y}$ of $\overline{A}$, a pointed closed convex cone $\mathcal{Q}$ in a finite dimensional space $\mathcal{Z}$ and a twice continuously differentiable mapping $\Xi : \mathcal{W} \to \mathcal{Z}$ such that: (i) $\Xi(\overline{A}) = 0 \in \mathcal{Z}$; (ii) the derivative mapping $\Xi'(\overline{A}) : \mathcal{Y} \to \mathcal{Z}$ is onto; (iii) $\mathcal{K} \cap \mathcal{W} = \{A \in \mathcal{W} \mid \Xi(A) \in \mathcal{Q}\}$. We say that $\mathcal{K}$ is C^2-cone reducible if $\mathcal{K}$ is C^2-cone reducible at every $\overline{A} \in \mathcal{K}$.

[33] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

Normal mapping $S_{\rm KKT}$

$$S_{\rm KKT}(a,b) = \big\{(x, z-\Pi_{\mathcal{K}}(z)) \in \mathcal{X}\times\mathcal{Y} \mid \Psi(x,z) = (a,-b)\big\},$$

where $\Psi : \mathcal{X}\times\mathcal{Y} \to \mathcal{X}\times\mathcal{Y}$ is Robinson's normal mapping defined by

$$\Psi(x,z) = \left[\begin{array}{c} \nabla f(x) + G'(x)^*(z-\Pi_{\mathcal{K}}(z)) \\ G(x)-\Pi_{\mathcal{K}}(z) \end{array}\right], \quad (x,z)\in\mathcal{X}\times\mathcal{Y}.$$

Natural mapping for KKT system

When $(a, b) = (0, 0)$, the KKT system (26) is equivalent to the following system of nonsmooth equations:

$$F(x, y) = 0,$$

where $F : \mathcal{X} \times \mathcal{Y} \to \mathcal{X} \times \mathcal{Y}$ is the natural mapping defined by

$$F(x, y) := \begin{bmatrix} \nabla f(x) + G'(x)^* y \\ G(x) - \Pi_{\mathcal{K}}(G(x) + y) \end{bmatrix}, \quad (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

Isolated calmness of S_{KKT} and F^{-1}

引理 6.2

Let $(0, 0, \bar{x}, \bar{y}) \in \operatorname{gph} S_{\mathrm{KKT}}$. The set-valued mapping S_{KKT} is isolated calm at the origin for $(\bar{x}, \bar{y})$ if and only if the set-valued mapping F^{-1} is isolated calm at the origin for $(\bar{x}, \bar{y})$.

The (robustly) isolated calmness of S_{KKT}

定理 6.2

Let $\bar{x}$ be a feasible solution to problem (24) with $(a,b)=(0,0)$. Suppose that the RCQ holds at $\bar{x}$. Assume that $\mathcal{K}$ is C^2-cone reducible and $\bar{y}\in M(\bar{x},0,0)\neq\emptyset$. Then the following statements are equivalent:

(i) *the SRCQ (27) holds at $\bar{x}$ with respect to $\bar{y}$ and the SOSC holds at $\bar{x}$ for problem (24) with $(a,b)=(0,0)$;*

(ii) *$\bar{x}$ is a locally optimal solution to problem (24) with $(a,b)=(0,0)$ and S_{KKT} is robustly isolated calm at the origin for $(\bar{x},\bar{y})$;*

(cont.)

(iii) $\bar{x}$ is a locally optimal solution to problem (24) with $(a, b) = (0, 0)$ and S_{KKT} is isolated calm at the origin for $(\bar{x}, \bar{y})$;

(iv) $\bar{x}$ is a locally optimal solution to problem (24) with $(a, b) = (0, 0)$ and F^{-1} is isolated calm at the origin for $(\bar{x}, \bar{y})$.

定理 6.3

(*About the three stability notions for C^2-reducible problems*)
Suppose that $\bar{x}$ is a locally optimal solution to problem (24) *with* $(a, b) = (0, 0)$ *and the RCQ holds at* $\bar{x}$. *Assume that* $\mathcal{K}$ *is* C^2*-cone reducible and* $\bar{y} \in M(\bar{x}, 0, 0)$. *Let* $\bar{z} = G(\bar{x}) + \bar{y}$. *Consider the following statements:*

(i) *The KKT point* $(\bar{x}, \bar{y})$ *is a strongly regular solution to the KKT system (26) with* $(a, b) = (0, 0)$;

(ii) *The mapping* Ψ^{-1} *has the Aubin property at the origin for* $(\bar{x}, \bar{z})$;

(cont.)

(iii) The KKT solution mapping S_{KKT} has the Aubin property at the origin for $(\bar{x}, \bar{y})$;

(iv) the constraint non-degeneracy (28) holds at $\bar{x}$ and the SOSC holds at $\bar{x}$ for problem (24) with $(a, b) = (0, 0)$;

(v) the SRCQ (27) holds at $\bar{x}$ with respect to $\bar{y}$ and the SOSC holds at $\bar{x}$ for problem (24) with $(a, b) = (0, 0)$;

(cont.)

(vi) The KKT solution mapping S_{KKT} is robustly isolated calm at the origin for $(\bar{x}, \bar{y})$;

(vii) The mapping F^{-1} is isolated calm at the origin for $(\bar{x}, \bar{y})$.

Then it holds that

$$(i) \Longrightarrow (ii) \Longleftrightarrow (iii) \Longrightarrow (iv) \Longrightarrow (v) \Longleftrightarrow (vi) \Longleftrightarrow (vii).$$

Thank You!

Bank B, Guddat J, Klatte D, Kummer B and Tammer K. *Nonlinear Parametric Optimization*. Berlin: Akademie-Verlag, 1982.

Bonnans J F and Ramírez C H. *Perturbation analysis of second order cone programming problems*. Mathematical Programming, Series B, 2005, **104**: 205-227.

Bonnans J F and Shapiro A. *Perturbation Analysis of Optimization Problems*. New York: Springer-Verlag, 2000.

Chan Z X and Sun D F. *Constraint nondegeneracy, strong regularity and nonsingularity in semidefinite programming*. SIAM Journal on Optimization, 2008, **19**: 370-396.

Clarke F H. *On the inverse function theorem*. Pacific Journal of Mathematics, 1976, **64**: 97-102.

Clarke F H. *Optimization and Nonsmooth Analysis*. New York: John Wiley and Sons, 1983.

Ding C, Sun D F and Zhang L W. *Characterization of the robust isolated calmness for a class of conic programming problems*. SIAM J. Optim., 2017, **27**: 67 – 90.

Dontchev A L. *Characterization of Lipschitz stability in optimization*. in Recent Developments in Well-Posed Variational Problems, Lucchetti R and Revalski J (eds), 1995, 95-116.

Dontchev A L and Rockafellar R T. *Characterizations of Lipschitz stability in nonlinear programming*. In: Fiacco AV, editor. Mathematical programming with data perturbations. New York: Marcel Dekker, 1997: 65-82.

- Dontchev A L and Rockafellar R T. *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*. SIAM J. Optim., 1996, **6**: 1087-1105.
- Dontchev A L and Rockafellar R T. *Implicit Functions and Solution Mappings*. New York: Springer, 2009.
- Dontchev A L and Rockafellar R T. *Regularity and Conditioning of Solution Mappings in Variational Analysis*. Set-Valued Analysis, 2004, **12**: 79-109.
- Dontchev A L and Hager W W. *Implicit functions, Lipschitz maps, and stability in optimization*. Mathematics of Operations Research, 1994, **19**: 753-768.
- Fusek P. *Isolated zeros of Lipschitzian metrically regular $\Re^n$-functions*. Optimization, 2001, **49**: 425-446.

King A and Rockafellar R T. *Sensitivity analysis for nonsmooth generalized equations*. Math Program, 1992, **55**: 341 – 364.

Klatte D and Kummer B. *Aubin property and uniqueness of solutions in cone constrained optimization*. Math. methods Oper. Res., 2013, **77**: 291-304.

Klatte D and Kummer B. *Nonsmooth Equations in Optimization*. KLUWER ACADEMIC, 2002.

Kummer B. *Lipschitzian inverse functions, directional derivatives, and applications in $C^{1,1}$-optimization*. Journal of Optimization Theory and Applications, 1991, **70**: 559-580.

Levy A B. *Implicit multifunction theorems for the sensitivity analysis of variational conditions*. Math. Program., 1996, **74**: 333-350.

Levy A B and Rockafellar R T. *Sensitivity of Solutions in Nonlinear Programming Problems with Nonunique Multipliers*. in Recent Advances In Nonsmooth Optimization, 1995: 215-223.

Mordukhovich B S. *Lipschitzian stability of constraint systems and generalized equations*. Nonlinear analysis, 1994, **22**: 173-206.

Mordukhovich B S. *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*. Berlin: Springer, 2006.

Outrata J, Kočvara M and Zowe J. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Theory, Applications and Numerical Results. Kluwer Academic Publishers, 1998.

Outrata J and Sun D F. *On the coderivative of the projection operator onto the second order cone*. Set-Valued Analysis, 2008, **16**: 999-1014.

Robinson S M. *Normal maps induced by linear transformations*. Math. of Oper. Res., 1992, **17**: 691-714.

Robinson S M. *Some continuity properties of polyhedral multifunctions*. Mathematical Programming Study, 1981, **14**: 206-214.

Robinson S M. *Strongly regular generalized equations*. Mathematics of Operations Research, 1980, **5**: 43-62.

Rockafellar R T. *Convex Analysis*. Princeton University Press, 1970.

Rockafellar R T and Wets R J B. *Variational Analysis*. New York: Springer-Verlag, 1998.

Sun D F. *A further result on an implicit function theorem for locally Lipschitz functions*. Operations Research Letters, 2001, **28**: 193-198.

Sun D F. *The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*. Mathematics of Operations Research, 2006, **31**: 761-776.

Walkup D W and Wets R J B. *A Lipschitzian characterization of convex polyhedra*. Proceedings of the American Mathematical Society, 1969, **23**: 167-173.

Zhang Y, Zhang L W, Wu J and Wang K D. *Characterizations of local upper Lipschitz property of perturbed solutions to nonlinear second-order cone programs*. Optimization, 2017, **66**: 1079-1103.

Kummer, B. *Newton's Method Based on Generalized Derivatives for Nonsmooth Functions: Convergence Analysis*. In: Oettli, W., Pallaschke D. eds., Lecture Notes in Economics and Mathematical Systems 382; Advances in Optimization. Springer, Berlin, 1992, 171-194.