变分分析–基础理论与前沿进展

稳定性的变分准则

第2讲：约束系统与凸优化的稳定性

张立卫

2021年11月

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.

系统稳定性

系统的模型

约束优化问题 $\begin{cases} \min f(x) \\ s.t.\ G(x) \in K \end{cases}$ → 约束集 $G^{-1}(K) = S(u_0)$

$G(x, u_0) = G(x)$

考虑下述集值映射的度量正则性的刻画: $G^{-1}(K)$ 在什么条件是凸集?

$u \in \mathcal{U}$

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

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.

$G(x,u) = 0$, Kummer 隐函数定理

- $K = \{0\} \subset \mathcal{Y}$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.

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. 外半连续
- 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$x_1, x_2 \in X$，$t \in [0, 1]$，

$(G(x), x)$ - epi F

闭凸的

$$t\Psi(x_1) + (1-t)\Psi(x_2) \subset \Psi(\underbrace{tx_1 + (1-t)x_2}_{x_1 \text{ 与 } x_2 \text{ 的凸组合}}). \quad (1)$$

用途：什么是凸优化问题？

$\mathcal{F}_G(x) = G(x) - K$ 是闭凸的

1) $\min f(x) \quad s.t. \quad G(x) \in K$

2) $\min f(x) + F(G(x))$

$g(x) \in \mathbb{R}^m_-$

$\Phi = \{x \in \mathbb{R}^n : g(x) \le 0\}$, $g : \mathbb{R}^n \to \mathbb{R}^m$

举例说明

$\mathcal{F}_g(x) = g(x) - \mathbb{R}^m_-$ 是凸的 $\Longrightarrow$ Φ 是凸集

$\Longleftrightarrow$ g 关于 $\le$ 是凸的.

- 凸的向量值函数

$\forall x_1, x_2 \in \mathbb{R}, \forall t \in [0,1]$, $t\mathcal{F}_g(x_1) + (1-t)\mathcal{F}_g(x_2) \subseteq \mathcal{F}_g(tx_1 + (1-t)x_2)$

$\Longleftrightarrow$ $t[g(x_1) - \mathbb{R}^m_-] + (1-t)[g(x_2) - \mathbb{R}^m_-] \subseteq g(x_t) - \mathbb{R}^m_-$

- 凸的矩阵函数

$\Longleftrightarrow$ $tg(x_1) + (1-t)g(x_2) \subseteq g(x_t) - \mathbb{R}^m_-$

$\Longleftrightarrow$ $\exists \delta \in \mathbb{R}^m, \delta \geq 0$, $tg(x_1) + (1-t)g(x_2) = g(x_t) + \delta$

$\Longleftrightarrow$ $g_i(x_t) \le t g_i(x_1) + (1-t) g_i(x_2)$, $i \in [m]$

$\Longleftrightarrow$ g_i 是凸函数, $i \in [m]$.

$\Phi = \{x \in \mathbb{R}^n : g(x) \geq 0\}$ $\mathcal{F}_g(x) = g(x) - \mathbb{R}^m_+$ 是闭凸的

$\Longleftrightarrow$ g_i 是凹函数, $i \in [m]$.

凸的矩阵函数

$\Phi = \{ x \in \mathbb{R}^n : \underline{G(x) \succeq 0} \}$, $G: \mathbb{R}^n \to S^p$

$\mathcal{F}_G(x) = G(x) - S^p_+$, $\mathcal{F}_G$ 是凹的 $\quad (x_1, x_2 \in \mathbb{R}^n,\ t \in [0,1])$

$\Leftrightarrow \mathcal{F}_G(x_t) \supseteq t\,\mathcal{F}_G(x_1) + (1-t)\,\mathcal{F}_G(x_2)$

$\Leftrightarrow G(x_t) - S^p_+ \supseteq t\,(G(x_1) - S^p_+) + (1-t)\,(G(x_2) - S^p_+)$

$\Leftrightarrow G(x_t) - S^p_+ \ni t\,G(x_1) + (1-t)\,G(x_2)$

$\Leftrightarrow \exists\, A \in S^p_+ \ (A \succeq 0)$ 满足

$G(x_t) - A = t\,G(x_1) + (1-t)\,G(x_2)$

$x_t = t x_1 + (1-t) x_2$

$\Leftrightarrow G(x_t) \succeq t\,G(x_1) + (1-t)\,G(x_2)$ ✓ (凹)

$\Phi = \{ x \in \mathbb{R}^n : G(x) \preceq 0 \}$, $G(x_t) \preceq t\,G(x_1) + (1-t)\,G(x_2)$. (凸)

A 映上 (onto) $\Leftrightarrow$ $AX=Y$ 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 1.1

$\Psi(x_t) \supseteq (1-t)\Psi(x_1) + t\,\Psi(x_2)$, $t \in [0,1]$

(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$y \in int\ \Psi(B_X(x, r))$.

即: $\forall r>0$, $\exists \varepsilon>0$, "$y+\varepsilon\mathbb{B}_Y \subset \Psi(x+r\mathbb{B}_X)$"

与 Ψ 的 openness 互相联系

定义 1.1

$y_0 \in \Psi(x_0)$

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V满足对$\forall (x, y) \in gph(\Psi) \bigcap V$, $\forall t \in [0, t_{\max}]$, 下述包含关系成立:

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

命题 1.1

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

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

证明

显然，取$\nu = t_{\max}, \eta = \gamma t_{\max}$，由(2)可得(3)。相反地，设$\Psi$是凸的，且 (3)成立。不妨设$x_0 = 0, y_0 = 0$。取

$$V = \nu \mathbb{B}_X \times \tfrac{1}{2}\eta \mathbb{B}_Y, \tag{4}$$

令$(x, y) \in \operatorname{gph}\Psi \bigcap V$。由$\Psi$的凸性及(3),$\forall t \in [0, 1]$,

$$\begin{aligned}
\underbrace{y + \tfrac{1}{2} t\eta \mathbb{B}_Y}_{y t'} &= (1-t)y + t(y + \tfrac{1}{2}\eta \mathbb{B}_Y) && (3)\\
&\subset (1-t)y + t\eta \mathbb{B}_y && y_0 + \eta B_y \subseteq \Psi(x_0 + \nu B_x)\\
&\subset (1-t)\Psi(x) + t\Psi(\nu \mathbb{B}_X)\\
&\subset \Psi((1-t)x + t\nu \mathbb{B}_X) && \text{convexity}\\
&\subset \underbrace{\Psi(x + 2t\nu \mathbb{B}_X)}_{t'}.
\end{aligned}$$

$t' = 2\nu t$

$\Rightarrow \frac{1}{2} t\eta = \frac{\eta}{4\nu} t' = \gamma t'$

则令

$$\gamma = \frac{\eta}{4\nu}, \qquad t_{\max} = 2\nu,$$

(2) $y + \gamma t B_Y \subseteq \Psi(x + t B_X)$ $\forall t \in [0, t_{\max}]$ $(x, y) \in V$

由V的定义，可得(2)。

命题 1.2

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$y_0 \in int(\mathrm{range}\,\Psi)$.

证明. 注意到，若集值映射Ψ 是外半连续的，凸的，根据广义开映射定理1.1，由正则性条件$y_0 \in \mathrm{int}(\mathrm{rge}\Psi)$可推出存在$\eta$与$\nu$ 满足(3)，从而由命题1.1，Ψ 在点$(x_0, y_0) \in \mathrm{gph}\Psi$处为开的. 显然，相反的结论亦成立. ■

$y_0 + \eta \mathbb{B}_Y \subseteq \Psi(x_0 + \nu \mathbb{B}_X)$

开性等价于度量正则性

定义 1.2

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

$$\operatorname{dist}(x, \Psi^{-1}(y)) \leq c \operatorname{dist}(y, \Psi(x)). \tag{5}$$

定理 1.2

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.

$\Longleftrightarrow \exists V \in \mathcal{N}(x_0, y_0)$, $\exists t_{max} > 0$ (使 $\forall (x,y) \in V$, $\forall t \in [0, t_{max}]$

$y + \gamma t \mathbb{B}_Y \subseteq \Psi(x + t\mathbb{B}_X)$.

$\exists V \in \mathcal{N}(x_0, y_0)$, $\exists t_{max} > 0$, $\forall (x, y) \in V \cap \mathrm{gph}\, \Psi$,
$y + \gamma t \mathbb{B}_Y \subset \Psi(x + t\mathbb{B}_X)$, $\forall t \in [0, t_{max}]$

证明

设Ψ在(x_0, y_0)处是以率$\gamma > 0$的开映射. 令$t_{\max} > 0$与V为定义1.1给出的. 不失一般性, 可设V具有下述形式

$$V = \varepsilon_x \mathbb{B}_X \times \varepsilon_y \mathbb{B}_Y.$$

若有必要, 减小$t_{\max}$, 再设

$$t_{\max}\gamma \le \frac{1}{2}\varepsilon_y. \tag{6}$$

令(x, y)满足

$$\|x - x_0\| < \varepsilon_x', \ \|y - y_0\| < \varepsilon_y', \tag{7}$$

其中ε_x', ε_y' 是满足下式的正常数

$$\varepsilon_x' \le \varepsilon_x, \ \gamma\varepsilon_x' + \varepsilon_y' \le t_{\max}\gamma. \tag{8}$$

注意到上述关系表明$\varepsilon_y' \leq \frac{1}{2}\varepsilon_y$. 现在来证关系式

$$d(x, \Psi^{-1}(y)) \leq cd(y, \Psi(x)) \qquad (9)$$

在$c = \gamma^{-1}$时是成立的. 事实上, 由于$\varepsilon_y' \leq \varepsilon_y$
且Ψ在(x_0, y_0)处是开的,
由$y \in y_0 + \|y - y_0\|\mathbb{B}_Y \subset \Psi(x_0 + \gamma^{-1}\|y - y_0\|\mathbb{B}_X)$ 得到

$$\Psi^{-1}(y) \subset x_0 + \gamma^{-1}\|y - y_0\|\mathbb{B}_X.$$

存在$x^* \in \Psi^{-1}(y)$满足

$$\|x^* - x_0\| \leq \gamma^{-1}\|y - y_0\|.$$

从而有

$$d\left(x, \Psi^{-1}(y)\right) \leq \|x - x^*\| \leq \|x - x_0\| + \gamma^{-1}\|y - y_0\| \leq \varepsilon_x' + \gamma^{-1}\varepsilon_y'.$$

- 如果

$$d\left(y,\Psi(x)\right) \geq \gamma\varepsilon_x' + \varepsilon_y' = \gamma(\varepsilon_x' + \gamma^{-1}\varepsilon_y')$$

(尤其, 若$\Psi(x)=\emptyset$这是成立的), 则

$$d\left(x,\Psi^{-1}(y)\right) \leq \gamma d\left(y,\Psi(x)\right),$$

证得结论.

- 否则,$d\left(y,\Psi(x)\right) < \gamma\varepsilon_x' + \varepsilon_y'$, 由(8), 对充分小的$\alpha > 0$, 存在$y_\alpha \in \Psi(x)$, 满足

$$\|y - y_\alpha\| \leq d\left(y,\Psi(x)\right) + \alpha < \gamma\varepsilon_x' + \varepsilon_y' \leq t_{\max}\gamma. \quad (10)$$

则由(6)–(8)可得

$$\|y_\alpha - y_0\| \leq \|y_\alpha - y\| + \|y - y_0\| < t_{\max}\gamma + \varepsilon_y' \leq \varepsilon_y. \quad (11)$$

- 因此, $(x, y_\alpha) \in \mathrm{gph}\Psi \bigcap V$. 结合(11)与$\Psi$在$(x_0, y_0)$处的开性,存在 $x' \in \Psi^{-1}(y)$满足$\|x' - x\| \leq \gamma^{-1}\|y - y_\alpha\|$. 于是得到

$$\begin{aligned} d\left(x, \Psi^{-1}(y)\right) & \leq \|x' - x\| \leq \gamma^{-1}\|y - y_\alpha\| \\ & \leq \gamma^{-1} d\left(y, \Psi(x)\right) + \gamma^{-1}\alpha. \end{aligned}$$

由于$\alpha > 0$是任意的, 当$c = \gamma^{-1}$ 时, (9)是成立的.

- 相反地, 设Ψ在(x_0, y_0)处以$c > 0$为率是度量正则的. 令$(x, y) \in \mathrm{gph}\Psi$, $z \in Y$满足$\|y - z\| < tc^{-1}$. 则对充分接近于 (x_0, y_0)的(x, y) 与充分小的$t > 0$, 有

$$d\left(x, \Psi^{-1}(z)\right) \leq c\, d\left(z, \Psi(x)\right) \leq c\|z - y\| < t.$$

这意味着存在$w \in \Psi^{-1}(z)$满足$\|w - x\| < t$; 因此有$z \in \Psi(x + t\mathbb{B}_X)$. 这就证得定理. ■

闭凸集值映射的度量正则性

定理 1.3

(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$y_0 \in int(\text{range}\,\Psi)$成立. 更精确地, 设(3)成立, (x, y)满足

$$\|x - x_0\| < \frac{1}{2}\nu, \quad \|y - y_0\| < \frac{1}{8}\eta. \tag{12}$$

则$c = 4\nu/\eta$时的(5)成立. 度量正则性: $V = \mathbb{B}(x_0, \frac{1}{2}\nu) \times \mathbb{B}(y_0, \frac{1}{8}\eta)$

$c = 4\nu/\eta$

$y_0 \in int(\text{range}\,\Psi) \Leftrightarrow$ (3): $\exists \eta > 0$, $\exists \nu > 0$ 使

$y_0 + \eta \mathbb{B}_Y \subseteq \Psi(x_0 + \nu \mathbb{B}_X)$

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

$\mathcal{F}_G$ 是凸的 $\Rightarrow$ Φ 是凸集

$x_0 \in \Phi$, $G(x_0) \in K$

约束系统的度量正则性

$y_0 = 0$ 的情况

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

$(x_0, y_0) \in \mathrm{gph}\,\mathcal{F}_G$

$$\mathcal{F}_G(x) = G(x) - K. \tag{13}$$

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

$$d(x, \mathcal{F}_G^{-1}(y)) \leq c\, d(y, \mathcal{F}_G(x)). \tag{14}$$

$= c\,\mathrm{dist}(G(x)-y, K)$

闭凸的集值映射 Φ, Φ 在 (x_0, y_0) 的度量正则性 $\Leftrightarrow$ $y_0 \in \mathrm{int}(\mathrm{range}\,\Phi)$

$(x_0, y_0) \in \text{gph}\, \mathcal{F}_G$, $\mathcal{F}_G(x) = G(x) - K$, 如何刻画 $\mathcal{F}_G$ 在 (x_0, y_0) 处的度量正则性?

Lipschitz扰动下的度量正则性

定理 1.4 $\mathcal{F}_H$

([1, 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

$\text{gph}\, \mathcal{F}_H$

$$d(x, \mathcal{F}_H^{-1}(y)) \leq c(\kappa) d(y, \mathcal{F}_H(x)) \tag{15}$$

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

$G(x)$ Taylor $H(x) = G(x_0) + DG(x_0)(x - x_0)$

$\mathcal{F}^* = \mathcal{F}_H$

$$\mathcal{F}^*(x) = G(x_0) + DG(x_0)(x - x_0) - K. \tag{16}$$

由中值定理, 差函数

$$G(x) - [\underline{G(x_0) + DG(x_0)(x - x_0)}]$$

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1.4, 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　　$\mathcal{F}^*$ 闭凸, 它在 $(x_0, 0)$ 处度量正则性 $\Longleftrightarrow$ $0 \in \text{int}\,(\text{range}\,\mathcal{F}^*)$

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

根据命题1.2[1], 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:

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

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

[1] 命题1.2 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$y_0 \in \text{int}(\text{range}\,\Psi)$.

注: Robinson CQ $\quad 0 \in \operatorname{int}\{G(x_0) + DG(x_0)X - K\}$

$\Leftrightarrow$ Φ_G 在 $(x_0, 0)$ 处度量正则性

$\Leftrightarrow$ $\exists \varepsilon > 0$, $\exists \delta > 0$, $\exists c > 0$, 有

$$\begin{aligned} \operatorname{dist}(x, \Phi_G^{-1}(y)) &\le c \operatorname{dist}(y, \Phi_G(x)), \quad \forall (x, y) \in \mathbb{B}_\delta(x_0) \times \varepsilon\mathbb{B} \\ &= c \operatorname{dist}(y, G(x) - K) \\ &= c \operatorname{dist}(G(x) - y, K) \end{aligned}$$

$\Rightarrow \operatorname{dist}(x, \Phi) \le c \operatorname{dist}(G(x), K)$

结论: 设 G 在 x_0 处 C^1, Robinson CQ 成立, 那么

$$T_\Phi(x_0) = \{d \in X : DG(x_0)d \in T_K(G(x_0))\}$$

证: 左端 $\subseteq$ 右端 易证. 要证右端 $\subseteq$ 左端. $\forall d \in X : DG(x_0)d \in T_K(G(x_0))$

$\exists t_k \downarrow 0$ 使 $\operatorname{dist}(G(x_0) + t_k DG(x_0)d, K) = o(t_k)$.

$$T_{\Phi}(x_0) = \{ h \in X : \exists t_k \downarrow 0 \text{ s.t. } \operatorname{dist}(x_0 + t_k h, \Phi) = o(t_k) \}$$

度量正则

$$\operatorname{dist}(x_0 + t_k d, \Phi) \leq c \operatorname{dist}(G(x_0 + t_k d), K) \qquad O(t_k) = o(t_k)$$

$$= c \operatorname{dist}(G(x_0) + t_k DG(x_0) d + O(t_k), K)$$

$$\leq c \operatorname{dist}(G(x_0) + t_k DG(x_0) d, K)$$

$$+ c \| O(t_k) \|$$

$$= o(t_k)$$

$$\Rightarrow d \in T_{\Phi}(x_0).$$

凸优化的Aubin性质

$\Longleftrightarrow$ 强正则性

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:

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

$$\delta_Q(x) = \begin{cases} 0, & x \in Q, \\ +\infty, & x \notin Q. \end{cases}$$

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:

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

$(x,y)\in \mathrm{gph}\, T$ $y\in T(x)$

T极大单调 $\Leftrightarrow \langle x'-x, y'-y\rangle \geq 0, \forall (x',y')\in \mathrm{gph}\, T$

极大单调算子

定义 2.1

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$\mathrm{gph}\, T$不真包含在其他任何的单调算子$T':\Re^n \rightrightarrows \Re^n$的图$\mathrm{gph}\, T'$中.

$0\in T(\bar{x}) \Leftrightarrow \langle x'-\bar{x}, y'-0\rangle \geq 0, \forall y'\in T(x'), \forall x'\in \mathrm{dom}\, T$

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.

$T^{-1}(y) = \{x \in \mathrm{dom}\, T : y \in T(x)\}$

以凸规划问题为例

$0 \in T(x)$

$T_f = \partial f$

$T_g = \partial g$ （凹）

$T_l = \partial_x l \times \partial_y l$

Ordinary Lagrangian

$l(x,y)$

$$(P)\quad \begin{cases} \min f_0(x) \\ \text{s.t. } f_i(x) \le 0,\ i \in [m] \\ x \in C \end{cases}$$

$$l(x,y) = \begin{cases} f_0(x) + \sum_{i=1}^{m} y_i f_i(x), & x \in C,\ y \in \mathbb{R}^m_+ \\ -\infty, & x \in C,\ y \notin \mathbb{R}^m_+ \\ +\infty & x \notin C \end{cases}$$

T_f^{-1}

T_g^{-1}

T_l^{-1}

$$(P) \iff \min_{x \in \mathbb{R}^n} f(x) = \sup_y l(x,y)$$

$$(D)\quad \max_{y \ge 0} g_0(y) = \inf_{x \in C} f_0(x) + \sum_{i=1}^{m} y_i f_i(x) \iff \max g(y) = g_0(y) - \delta_{\mathbb{R}^m_+}(y)$$

$$g(y) = \inf_x l(x,y)$$

（鞍点） $\inf_x \sup_y l(x,y)$

满足Aubin性质的极大单调算子

将通过利用[2, 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 2.1

[2, 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.

$$F(x) \cap W \subseteq F(x') + \gamma \|x - x'\| \mathbb{B}_{X^*}, \quad x, x' \in U,$$

$$W \in \mathcal{N}(y_0), \quad U \subset \mathcal{N}(x_0)$$

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

$\langle h_k, \cdot \rangle$

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

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

$\|h_k\| = 1$

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

因此, 存在序列$u_k \in F(x_k + t_k h_k)$ 使得 $(x_k, z_k) \in \operatorname{gph} F$

$$\|u_k - y_k\| \le \gamma t_k. \quad (x_k + t_k h_k, u_k) \in \operatorname{gph} F \tag{5}$$

由F的单调性,

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

结合(3), 有 $\Leftrightarrow \langle u_k, h_k \rangle \ge \langle z_k, h_k \rangle$

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

进一步, 由(4)和(5), $\gamma t_k < b_k/2$

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

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 2.1

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2.1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.

■

多面集值映射的上Lipschitz连续性

定义 2.2

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[4].[2]

定理 2.2

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.

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

引理 2.1

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

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

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 2.2

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

$$d_x(z, G) \leq \eta d_y(z, G) \text{ and } d_y(z, G) \leq \xi d_x(z, G) \qquad (7)$$

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

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

and

$$\operatorname{dist}\Big(y_0, \{y' : By' \leq b\}\Big) \leq \beta\|(By_0 - b)^+\|. \tag{8}$$

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 (8) that

$$\begin{aligned} d_y(z, G) &= \mathrm{dist}\Big(y, \{y' : By' \leq c - Ax\}\Big) \\ &\leq \beta\|(Ax + By - c)^+\|. \end{aligned} \tag{9}$$

For $\tilde{x}$ closest to x in the set $\{x' : Ax' \leq c - By\}$[3] one has

$$\|(Ax + By - c)^+\| \leq \|(Ax + By - c) - (A\tilde{x} + By - c)\|, \tag{10}$$

which yields

$$\|(Ax + By - c)^+\| \leq \|A\|\|x - \tilde{x}\|. \tag{11}$$

[3]有$\|x - \tilde{x}\| = d_x(z, G)$

But $\|x - \tilde{x}\| = d_x(z, G)$ by construction and thus, combining (9), (10) and (11), 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 (7) is proven in the same way. □

定理 2.3

[?][4] *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 $x \in \operatorname{dom} P$.*

Proof. Let $G_i, i \in [k]$, be the components of P. With the constant ξ_i associated with G_i according to Lemma 2.2 we put

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

[4] 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 2.1 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

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

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. $\square$

- Bonnans J F and Shapiro A. *Perturbation Analysis of Optimization Problems*. New York: Springer-Verlag, 2000.
- Dontchev A L and Hager W W. *Implicit functions, Lipschitz maps, and stability in optimization*. Mathematics of Operations Research, 1994, **19**: 753-768.
- Klatte D and Kummer B. *Aubin property and uniqueness of solutions in cone constrained optimization*. Math. methods Oper. Res., 2013, **77**: 291-304.
- Robinson S M. *Some continuity properties of polyhedral multifunctions*. Mathematical Programming Study, 1981, **14**: 206-214.
- Rockafellar R T and Wets R J B. *Variational Analysis*. New York: Springer-Verlag, 1998.