变分分析–基础理论与前沿进展

稳定性的变分准则

第3讲：NLP的稳定性

张立卫

2021年11月

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.

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. 上Lipschitz连续
- 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).

KKT GE 强正则于 $(\bar{x}, \bar{u}, \bar{\lambda})$

$\Longleftrightarrow$ • LICQ在 $\bar{x}$ 处成立

• 强二阶充分条件

$$\min f(x) \quad \text{s.t.} \quad h(x) = 0,\ g(x) \le 0$$

$$\begin{cases} \nabla_x L(x, \mu, \lambda) = 0 \\ h(x) = 0 \\ 0 \le \lambda \perp g(x) \le 0 \end{cases}$$

KKT 系统

GE 形式

$$\delta \in \begin{bmatrix} \nabla_x L(x, \mu, \lambda) \\ -h(x) \\ -g(x) \end{bmatrix} + N_{\mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^p_+}(x, \mu, \lambda)$$

$\delta \mapsto (x(\delta), \mu(\delta), \lambda(\delta))$ Lipschitz 连续

强正则性

$\Lambda(\bar{x}) \neq \phi$

$\forall\, d \in C(\bar{x}) = \left\{ d \in \mathbb{R}^n : \begin{array}{l} \nabla h(\bar{x}) d = 0 \\ \nabla g_i(\bar{x})^T d \le 0,\ i \in I(\bar{x}), \end{array} \nabla f(\bar{x})^T d \le 0 \right\},\quad d \neq 0$

$$\sup_{(\mu,\lambda)\in\Lambda(\bar{x})} d^T \nabla^2 L(\bar{x},\mu,\lambda) d > 0. \quad \Rightarrow \text{二阶严格局部最优}$$

↑ 二阶充分性条件

强二阶充分性条件

$\forall\, d \in \text{aff}\, C(\bar{x}),\ d \neq 0$

$$\sup_{(\mu,\lambda)\in\Lambda(\bar{x})} d^T \nabla^2 L(\bar{x},\mu,\lambda) d > 0.$$

对非多面体锥不对 $\quad G(x) \in K$

- 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.

严格MF: $\quad DG(x)X + T_K(G(x)) \cap \lambda^{\perp} = Y$

$$S_{KKT}(\delta) = \{ \delta \in G(u) + N_D(u) \}$$

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 1.1

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[15].[1]

定理 1.1

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.

$S(\bar{x}) \neq \phi$

$S(x) \subseteq S(\bar{x}) + \kappa \|x - \bar{x}\| \mathbb{B}$

$\forall x \in \mathbb{B}_\delta(\bar{x})$

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

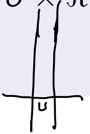
引理 1.1

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*

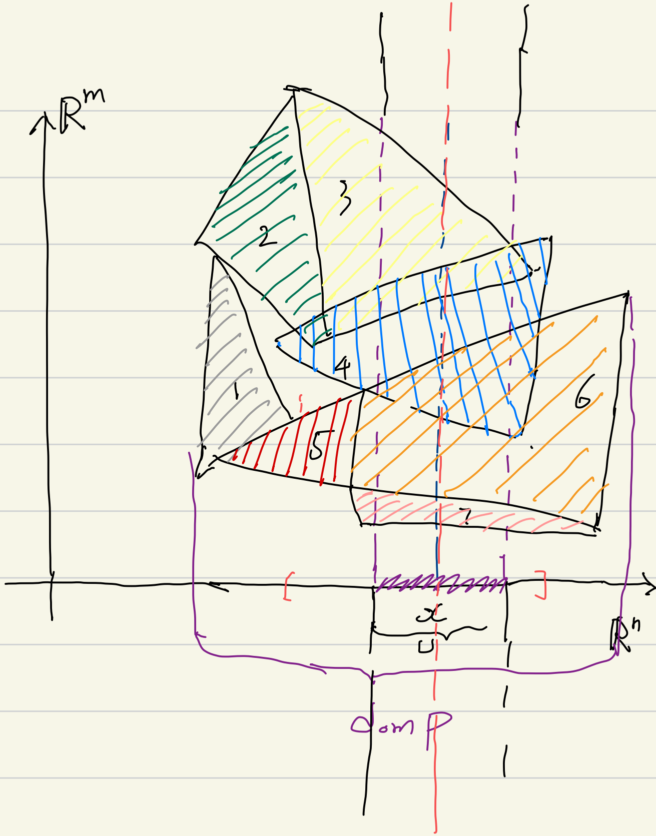
$$J(x) = \{i \in [k] : x \in \pi_1(G_i)\},$$

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*

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



Outrata et. al (1998)

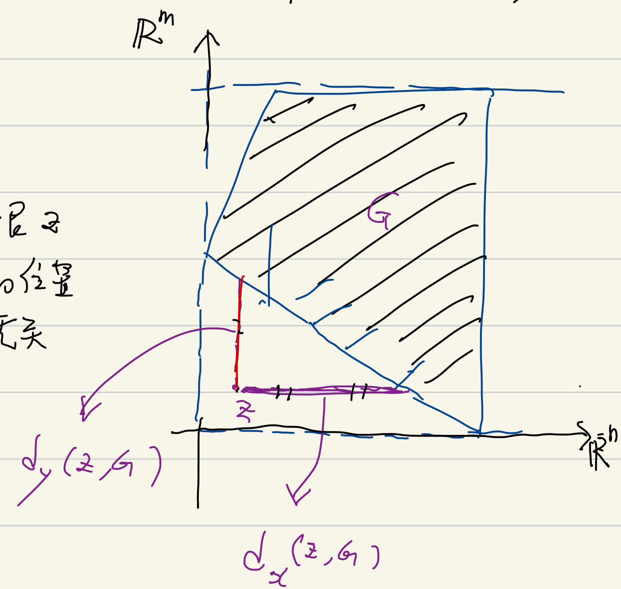


$J(x) = \{3, 4, 6, 7\}$

$1, 2, 5 \notin J(x)$

$\pi_1(G) \times \pi_2(G)$

距z
的位置
无关



$d_y(z, G) \leq x, \; d_x(z, G)$

$x \in \operatorname{dom} P \Leftrightarrow P(x) \neq \phi$

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. □

引理 1.2

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

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

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

$S^* = \{x : Ex + d \leq 0\}$, dist$(x, S^*) \leq \kappa [Ex + d]_+$

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 $\{x' : Ax' \leq c - By\}$[2] one has

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

which yields

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

$\|x - \tilde{x}\| = d_x(z, G)$

[2] 有$\|x - \tilde{x}\| = d_x(z, G)$

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 1.2

[13][3] *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 1.2 we put

$d_y(z, G_i) \leq \xi_i d_x(z, G_i)$, $\quad \lambda = \max\{\xi_1, \ldots, \xi_k\}. \quad d_y(z, G) \leq \xi d_x(z, G)$

[3] 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 1.1 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

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

$(x', y') \in G_i$

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. □

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

$\lambda \in N_{\mathbb{R}^p_-}(g(x)) \Leftrightarrow -g(x) + \Pi_{\mathbb{R}^p_-}(g(x)+\lambda) = 0$

$\Leftrightarrow g(x) \in N_{\mathbb{R}^p_+}(\lambda)$

$\Leftrightarrow \lambda - \Pi_{\mathbb{R}^p_+}(\lambda + g(x)) = 0$

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

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

$\Updownarrow$

$0 \le \lambda \perp g(x) \le 0$

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

Natural mapping

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

✓

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

GE

KKT条件(7)也可以等价地表示为下述的广义方程

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

$$\delta \in H(\bar{x},\bar{\zeta},\bar{\lambda}) + DH(\bar{x},\bar{\zeta},\bar{\lambda}) \begin{pmatrix} x-\bar{x} \\ \zeta-\bar{\zeta} \\ \lambda-\bar{\lambda} \end{pmatrix} + N_K(x,\zeta,\lambda)$$

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

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

广义方程的法映射

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 1.3

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.

$$0 \in \begin{bmatrix} \nabla_x L(x, \zeta, \lambda) \\ -h(x) \\ -g(x) \end{bmatrix} + N_{\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p_+}(x, \zeta, \lambda) \qquad (9)$$

$\lambda(\delta)$

$$\begin{cases} \min & f(x) - \langle \delta_f, x \rangle \\ \text{s.t.} & h(x) + \delta_h = 0 \\ & g(x) + \delta_g \le 0 \end{cases}$$

标准扰动

$$\delta = \mathcal{F}(x, \zeta, y), \quad \lambda(\delta) = \Pi_{\mathbb{R}^p_+}(y(\delta))$$

$$\Rightarrow (x(\delta), \zeta(\delta), y(\delta))$$

命题 1.1

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. $\bar{y} = \bar{\lambda} + g(\bar{x})$

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

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

"$V \in \partial \mathcal{F}(\bar{x}, \bar{\zeta}, \bar{y})$, $Vh = 0 \Rightarrow h = 0$"

一致二阶增长条件

引理 1.4

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 1.3

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)[4] showed that the strong regularity of S is equivalent to Aubin property of S around a point $(z_0, w_0, x_0) \in \operatorname{ghp} S$.

- 非线性规划KKT系统的强正则性等价于Aubin性质.

[4] 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

S 在 $(\bar{x},\bar{y})$ 处孤立平稳的 $\iff DS(\bar{x}|\bar{y})(0)=\{0\}$

内容取自Dontchev and Rockafellar(1997)[3][5]. 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.

[5] 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.

C^2参数化扰动问题

考虑下述参数非线性规划问题

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

其中$C(u, w)$表示下列约束:

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

其中$g_i : \Re^d \times \Re^n \to \Re$, $i = 0, 1, \ldots, m$是二次连续可微函数, 向量$w \in \Re^d$, $v \in \Re^n$与$u = (u_1, \ldots, u_m)^T \in \Re^m$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$S_{\mathrm{KKT}}(p)$, 称映射$p\mapsto S_{\mathrm{KKT}}(p)$为KKT映射.
- 记$X_{\mathrm{KKT}}(p)$为稳定点集, 即

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

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

与$(v_0,u_0,w_0,x_0,y_0)\in \mathrm{gph} S_{\mathrm{KKT}}(p)$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$\nabla_x g_i(w_0, x_0)^T z < 0$.

$$Y = DG(x_0)X + T_K(G(x_0)) \cap x^{\perp},$$

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(20)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$\mathcal{J}_x f(w_0, x_0)$. 考虑$f$的线性化映射:

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

定理 1.4

[2][6] 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) $\mathcal{L}$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.

[6] 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 1.1

假设定理1.4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

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

(ii) 映射Σ在(p_0, x_0)处是孤立平稳的.

证明

映射$\mathcal{L} = [f(w_0, x_0) + \mathcal{J}_x f(w_0, x_0)(\cdot - x_0) + F(\cdot)]^{-1}$是多面体，因此由[15]知$\mathcal{L}$ 在$\Re^m$上是平稳的(是局部上Lipschitz连续的)，则(i)可推出$\mathcal{L}$在(y_0, x_0)处是孤立平稳的. 应用定理1.4得Σ在(p_0, x_0)处是孤立平稳的. 再应用定理1.4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1.1可得结论:

推论 1.2

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 1.5

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.

孤立局部极小点 + (p_0, x_0)处的MF约束条件成立

$\Longrightarrow$ 映射X在p_0处下半连续

证明

由[12, 推论4.5][7]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$\mathbb{B}(x_0, \alpha) \subset U$.

[7] 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

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

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$\|x_p - x\| \le \gamma\|p - p_0\|$. 则有

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

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[8], (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.

---

[8]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$p\in \mathbf{B}(p_0,\eta')$, 任何解$x\in X_\alpha(p)$满足$\|x-x_0\|\le\delta'<\alpha+\gamma\|p-p_0\|$. 因此, 对$p\in \mathbf{B}(p_0,\eta')$约束$\|x-x_0\|\le\alpha+\gamma\|p-p_0\|$在问题(19)中是无效的. 所以任何$p\in \mathbf{B}(p_0,\eta')$, 有

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

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 1.5

下述条件等价:

(i) 映射S_{KKT}在$(p_0, x_0, y_0) \in \operatorname{gph} S_{\mathrm{KKT}}$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$Y_{\mathrm{KKT}}(x_0, p_0)$中的孤立点. 注意到$Y_{\mathrm{KKT}}(x_0, p_0)$ 是凸的, 则有$Y_{\mathrm{KKT}}(x_0, p_0) = \{y_0\}$. 因此严格MF约束条件成立[9]. 进一步, 由推论1.2, 不存在(x_0, y_0)附近的(x, y)满足$(x, y) \in L_{\mathrm{KKT}}(p_0)$.[10] 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.

---

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

[10] 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.

不存在(x_0, y_0)附近的(x, y)满足$(x, y) \in L_{\text{KKT}}(p_0)$

- $p_0 = (v_0, u_0, w_0)$, (x_0, y_0)满足KKT条件(14).
  记$A = \nabla^2_{xx} L(w_0, x_0, y_0)$, $B = \nabla^2_{yx} L(w_0, x_0, y_0)$.
- 系统

$$
\begin{cases}
\nabla_x L(w_0, x_0, y_0) + A(x - x_0) + B^T(y - y_0) = 0, \\
g(w_0, x_0) + B(x - x_0) \in N_Y(y).
\end{cases}
\tag{20}
$$

有唯一解(x^0, y^0).

那么$(x,y)=(0,0)$为下述变分系统的孤立解:

$$
\begin{aligned}
&Ax+B^Ty=0,\\
&B_1x=0,\\
&B_2x\le 0, y_i\ge 0, y_i(Bx)_i=0, i\in[m_1+1,m_2].
\end{aligned}
\tag{21}
$$

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(21)的解集是一个锥, 则$(0,0)$为(21)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(21)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1.5，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$S_{\mathrm{KKT}}(p) \cap U \neq \emptyset$.[11]

[11] 稳健性

如果$I_1 = \emptyset$, 那么$y_0 = 0 = Y_{\mathrm{KKT}}(x_0, p_0)$. 由引理1.5, 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1.2,[12] (21)[13]有非零解(x', y')且该解可与$(0, 0)$无限接近.

---

[12] 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.

[13] 即系统

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

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(21) 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. ∎

定理 1.6

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

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

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.

证明

由Levy和Rockafellar[11][14]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[8]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. ■

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

[14] 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.

乘子映射的稳健孤立平稳性

基于引理1.5和定理1.6可以得到下述推论:

推论 1.3

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1.6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.

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

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

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.

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

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

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

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.