中国科学: 数学 2020 年 第 50 卷 第 12 期: 1869～1890
SCIENTIA SINICA Mathematica
论 文

一类变分包含正则间隙函数的可微性与解的局部唯一性

献给吴从炘教授 85 华诞

宋文

哈尔滨师范大学数学科学学院, 哈尔滨 150025
E-mail: wsong@hrbnu.edu.cn

收稿日期: 2020-05-23; 接受日期: 2020-07-17; 网络出版日期: 2020-12-15
国家自然科学基金 (批准号: 11871182) 资助项目

摘要 本文借助于邻近正则函数的 Moreau 包络及其邻近映射的性质, 建立一类非凸变分包含正则间隙函数及其解映射的 Lipschitz 连续性和可微性; 利用二阶变分分析理论得到了变分包含解的局部唯一性及关于参数的连续性. 同时也概述了变分不等式和优化问题正则间隙函数的发展过程及邻近正则函数的变分性质.

关键词 正则间隙函数 Moreau 包络 邻近映射 邻近正则 上图可微

MSC (2010) 主题分类 90C31, 49J52, 46C05

1 引言

标准的变分不等式是, 寻找 $x \in K$ 使得

$$\langle F(x), y - x\rangle \geqslant 0, \quad \forall\, y \in K, \tag{1.1}$$

其中 $F\colon \mathbb{R}^n \to \mathbb{R}^n$ 是一个映射, $K \subset \mathbb{R}^n$ 一个闭凸子集. 求解变分不等式 (1.1) 的一个常用方法是将变分不等式 (1.1) 转化为一个等价的优化问题. 例如, Auslender [1] 考虑了一个等价的优化问题, 即 $\max_{x\in K} \omega(x)$, 其中

$$\omega(x) := \inf_{y\in K} \langle F(x), y - x\rangle$$

表示变分不等式 (1.1) 的间隙函数.

英文引用格式: Song W. Differentiability of regularized gap functions for a kind of variational inclusions and local uniqueness of solutions (in Chinese). Sci Sin Math, 2020, 50: 1869–1890, doi: 10.1360/SSM-2020-0158

© 2020《中国科学》杂志社

然而, 除非 K 是紧集时, 间隙函数 ω 是广义实值函数, 且一般是非可微的. 为了克服函数 ω 的非可微性, Auchmuty [2] 和 Fukushima [3] 独立引进了下面可微的正则间隙函数:

$$\omega_\tau(x) = \inf_{y\in K}\left\{\langle F(x), y-x\rangle + \frac{1}{2\tau}\|x-y\|^2\right\}. \tag{1.2}$$

Fukushima [3] 证明了点 $\bar{x}\in K$ 是 (1.1) 的解当且仅当 $\omega_\tau(\bar{x})=0$ 且 $\bar{x}$ 是优化问题 $\max_{x\in K}\omega_r(x)$ 的最优解, 并设计了一个基于此优化问题的求解变分不等式 (1.1) 的可行下降方法. 文献 [4–7] 继续研究了 D- 间隙函数 (即两个正则间隙函数 (1.2) 关于不同参数 τ 的差). 关于变分不等式 (1.1) 间隙函数更多的讨论参见文献 [8]. 自此, 正则间隙函数的方法被推广到拟变分不等式 [9]、广义 Nash 均衡问题 [10]、非光滑变分不等式 [11] 和 K 为非凸情形 [12] 等.

另一方面, Hearn [13] 对线性约束的凸优化问题

$$\min f(x) \quad \text{s.t. } x\in K=\{x\colon Ax=b, x\geqslant 0\} \tag{1.3}$$

引入并使用了间隙函数这个术语, 其定义如 $\omega(x)=\inf_{y\in K}\langle\nabla f(x), y-x\rangle$. Hearn 证明了这个间隙函数恰为优化问题 (1.3) 最优值与其 Lagrange 对偶问题最优值的差, 同时解释了该间隙函数与求解凸优化问题 (1.3) 的 Frank-Wolfe 算法的关系 (Frank-Wolfe 最早考虑的是二次凸优化问题, 这个算法在 Bertsekas 的文献 [14] 中也被称为条件梯度算法). 自此, 间隙函数成为优化问题的对偶理论与方法中的重要概念. Migdalas [15] 利用正则化子问题

$$\min_{y\in K}\langle\nabla f(x), y-x\rangle + \frac{1}{2\tau}\|x-y\|^2 \tag{1.4}$$

生成下降方向, 发展了优化问题 (1.3) 的一类正则 Frank-Wolfe 算法. 事实上, 各种梯度投影法生成的可行方向都与优化子问题 (1.4) 紧密相关 (参见文献 [14,16,17]). 它的解集可以表示成如下形式:

$$\begin{aligned}&\operatorname*{argmin}_{y\in K}\langle\nabla f(x), y-x\rangle + \frac{1}{2\tau}\|x-y\|^2\\ &\quad= \operatorname*{argmin}_{y\in K}\frac{1}{2\tau}\|x-\tau\nabla f(x)-y\|^2\\ &\quad= P_K(x-\tau\nabla f(x)).\end{aligned} \tag{1.5}$$

近年来, 下面更一般的复合优化问题

$$\min\phi(x) := f(x)+g(x), \tag{1.6}$$

其中 $f\colon H\to\mathbb{R}$ 是定义在 Hilbert 空间 H 上的连续可微 (可能非凸) 函数, $g\colon H\to\bar{\mathbb{R}}$ 是下半连续真凸函数, 因其在信号处理、图像恢复和机器学习等方面的广泛应用, 无论在理论还是算法方面均得到了广泛且深入研究 (参见文献 [18,19,21–26]).

邻近梯度算法 (也称为前后分裂算法) 是求解问题 (1.6) 的最典型的方法之一 (参见文献 [18,20,22–24,26–28]). 这种方法是经典梯度方法和梯度投影方法到一般复合优化 (包含非光滑项) 的推广. 基于 Nesterov 的工作, 邻近梯度算法的各种加速形式 (或称为最优一阶算法) 也被广泛研究 (参见文献 [17,26]). 复合优化问题 (1.6) 的各种邻近梯度算法均与下面的正则子优化问题及其解映射有关:

$$\gamma_\tau(x) = \inf_{y\in H}\left\{\langle\nabla f(x), y-x\rangle + \frac{1}{2\tau}\|x-y\|^2 + g(y)\right\}$$

$$= \inf_{y\in H}\left\{g(y)+\frac{1}{2\tau}\|x-\tau\nabla f(x)-y\|^2\right\}, \tag{1.7}$$

$$P_\tau(x) = \operatorname*{argmin}_{y}\left\{\langle\nabla f(x), y-x\rangle+\frac{1}{2\tau}\|x-y\|^2+g(y)\right\}$$

$$= \operatorname*{argmin}_{y\in H}\left\{g(y)+\frac{1}{2\tau}\|x-\tau\nabla f(x)-y\|^2\right\}, \tag{1.8}$$

其中函数 $\gamma_\tau(x)$ 被称为复合优化问题 (1.6) 的正则间隙函数 (或者称为 $\phi(x)$ 的前后包络函数), 算子 $P_\tau(x)$ 被称为前后包络算子 (参见文献 [24,27]). 文献 [27] 在 $H=\mathbb{R}^n$ 空间的情形给出了函数 $\gamma_\tau(x)+f(x)$ 的可微性质、与 ϕ 的不等式关系, 以及它的极小点与问题 (1.6) 最优解的关系. 文献 [24] 说明了, 当参数 λ 介于 0 与 ∇f 的 Lipschitz 常数的倒数之间时, 邻近梯度算法和前后包络算子的不动点迭代及函数 $f(x)+g(x)$ 梯度流的前后分裂算法是一致的. 显然, 前后包络函数和前后包络算子建立在所谓的凸函数 Moreau 包络 (也称为 Moreau-Yosida 逼近, 事实上, 一个凸函数 Moreau 包络的 Fréchet 导数是该凸函数次微分的 Yosida 逼近)

$$e_\tau g(x) = \inf_{w\in H}\left\{g(w)+\frac{1}{2\tau}\|x-w\|^2\right\} \tag{1.9}$$

及其邻近映射 (proximal mapping)

$$\operatorname{prox}_{\tau g}(x) := \operatorname*{argmin}_{w\in H}\left\{g(w)+\frac{1}{2\tau}\|w-x\|^2\right\} \tag{1.10}$$

的基础上. 众所周知, 凸函数的包络函数及相关邻近映射在提供优化问题的正则性和基本的计算方法方面起到了基础性的作用. 因此, 无论在理论还是算法方面均受到了人们的广泛关注 (参见文献 [24,28–30]).

一类重要的函数称为邻近正则函数 [31], 在变分分析理论及应用的许多方面都显示出其应用价值, 它为光滑分析和凸分析提供了一个完美的统一框架. 一方面, 邻近正则性的结构能保证该类函数的包络和邻近映射具有许多类似于凸函数的良好性质; 另一方面, 它的范围足以包括许多重要的函数类, 如下半连续凸函数、可微函数和强 amenable 函数 (即, 函数能局部地表示成一个下半连续真凸函数与一个二次连续可微映射的复合形式且满足某些约束品性), 或者仿凸函数 (即, 函数能局部地表示成一个连续凸函数与范数平方的正常数倍之差的形式) 等 (参见文献 [30,32–38]).

本文将讨论下面变分包含问题:

$$0\in F(x)+\partial g(x) \tag{1.11}$$

的正则间隙函数及邻近解映射的连续性和可微性, 以及变分包含解的局部唯一性及关于参数的连续性, 其中 $F\colon H\to H$ 是定义在 Hilbert 空间 H 上的映射, $g\colon H\to\bar{\mathbb{R}}$ 是定义在 H 上真的邻近正则函数, ∂g 代表函数 g 的 Mordukhovich (或极限) 次微分映射. 具体地, 考虑如下定义的正则间隙函数及其解映射的性质:

$$\gamma_\tau(x) := \inf_{y\in H}\left\{\langle F(x), y-x\rangle+\frac{1}{2\tau}\|x-y\|^2+g(y)\right\}, \tag{1.12}$$

$$P_\tau(x) = \operatorname*{argmin}_{y}\left\{\langle F(x), y-x\rangle+\frac{1}{2\tau}\|x-y\|^2+g(y)\right\}. \tag{1.13}$$

简单的计算表明,

$$P_\tau(x) = \mathrm{prox}_{\tau g}(x - \tau F(x)). \tag{1.14}$$

特别地, 当 $g(\cdot) = \delta_K(\cdot)$ 是 H 中某个非空闭集的指示函数时, 变分包含问题 (1.11) 就退化为变分不等式

$$-F(X) \in N_K(x) \tag{1.15}$$

的形式, 其中 $N_K(x)$ 表示 K 在 x 点的 Mordukhovich 法锥. Shapiro [12] 在 K 是邻近正则集时, 引进了变分不等式 (1.15) 的局部正则间隙函数, 并讨论了它的可微性和变分不等式 (1.15) 解的局部唯一性及关于参数的连续性. 当映射 F 是某个可微函数 f 的梯度时, 变分包含问题 (1.11) 就是复合优化问题 (1.6) 的必要最优性条件, 进一步在 f 和 g 都是凸函数时, (1.11) 也是复合优化问题 (1.6) 的充分最优性条件.

本文余下内容的结构如下: 第 2 节回顾映射可微的定义、广义实值函数次微分、邻近有界和邻近正则的定义及关于 Moreau 包络以及邻近映射的某些结果, 利用局部正则间隙函数给出变分包含解存在性的等价刻画; 第 3 节利用邻近正则函数的 Moreau 包络和邻近映射的可微性, 得到变分包含的正则间隙函数及邻近解映射的一阶、二阶可微性; 第 4 节讨论变分包含解的局部唯一性及对参数的连续性; 第 5 节将前面的结果应用到特殊的变分不等式情形, 所获得的结果改进了文献 [12] 中的某些结果.

2 基本结果

首先回顾映射的各种可微性, 可参见文献 [12,28,39]. 令 $F\colon H \to Y$ 为 Hilbert 空间 H 到另一个 Hilbert 空间 Y 的映射.

定义 2.1 令 $x, h \in H$, 如果极限

$$F'(x,h) := \lim_{t\downarrow 0} \frac{F(x+th) - F(x)}{t} \tag{2.1}$$

存在, 则称 F 在 x 点沿方向 $h \in H$ 是方向可微的. 如果 F 在 x 点沿任意方向 $h \in H$ 都是可微的, 则称 F 在 x 点是方向可微的. 如果 F 在 x 点是方向可微的, 且方向导数 $F'(x,h)$ 是 h 的连续线性映射, 记作 $DF(x)h$, 则称 F 在 x 点是 Gâteaux 可微的.

定义 2.2 如果对任意的 $h \in H$, 极限

$$F'(x,h) = \lim_{\substack{h' \to h \\ t\downarrow 0}} \frac{F(x+th') - F(x)}{t}$$

存在, 或者对任意的点列 $h_n \to h$ 和任意的非负数列 $t_n \downarrow 0$, 极限

$$F'(x,h) = \lim_{n\to\infty} \frac{F(x+t_n h_n) - F(x)}{t_n}$$

存在, 则称 F 在 x 点沿方向 h 是 Hadamard 方向可微的. 如果 F 在 x 点沿任意方向 $h \in H$ 都是 Hadamard 方向可微的, 则称 F 在 x 点是 Hadamard 方向可微的.

显然, 如果 F 在 x 点是 Hadamard 方向可微的, 则它的方向导函数 $F'(x,\cdot)$ 在 H 上是连续的. 此外, 如果 $F'(x,h)$ 关于 h 是线性的, 则称 F 在 x 点是 Hadamard 可微的. 如果 F 在 x 点是局部 Lipschitz 的, 则 F 在 x 点是 Hadamard 方向可微的当且仅当它在 x 点是方向可微的.

定义 2.3 如果 F 在 x 点是方向可微的, 且

$$F(x+h)=F(x)+F'(x,h)+o(\|h\|), \tag{2.2}$$

则称 F 在 x 点是 Fréchet 方向可微的. 此外, 如果 $F'(x,\cdot)$ 是线性连续的, 则称 F 在 x 点是 Fréchet 可微的.

如果 F 在 H 的某个开子集 O 上是 Gâteaux 可微的, 且相应的导算子 $DF(x)$ 在 O 上是连续的, 则称 F 在 O 上是连续可微的 (简称 C^1). 如果 $DF(x)$ 在 O 上是 Lipschitz 连续的, 则称 F 在 O 上是 C^{1+}. 易知, 如果 F 在 O 上是连续可微的, 则 F 在 O 的每个点 x 是 Fréchet 可微的.

令 $g\colon H\to(-\infty,+\infty]$ 为定义在 H 上的广义实值函数, g 的有效域为 $\mathrm{dom}g=\{x\in H\colon g(x)<\infty\}$. 当 g 是下半连续真凸函数时, 它的 Moreau 包络函数是凸连续可微的, 相应的邻近映射是单值 Lipschitz 连续的, 并且可用于构造求解函数 g 极小点的邻近点算法 (proximal point algorithm). 正是由于这些丰富的理论及计算性质使得它们在凸分析及优化领域有非常重要的应用. 对于非凸情形, Poliquin 和 Rockafellar [31] 引进了邻近正则函数, 并证明了这类函数的包络函数具有类似于凸函数的性质, 例如, 是连续局部仿凸的. 自此, 这类函数在变分分析及应用等许多方面显示出了其重要意义. 关于邻近映射在凸和邻近正则情形的计算可参见文献 [17,34].

称函数 g 为邻近有界的, 如果存在 $r>0$ 及某个 x 使得 $e_{\frac{1}{r}}g(x)>-\infty$, 等价地, 存在 $r>0$ 使得 $g+\frac{r}{2}\|\cdot\|^2$ 是下有界的. 所有这些常数 r 的下确界, 记为 r_g, 称为 g 的邻近界的阈值. 当 $\tau<(\max\{r_g,0\})^{-1}$ 时, $e_\tau g(x)>-\infty$ 并且 $\mathrm{prox}_{\tau g}(x)\neq 0$.

定义 2.4 (参见文献 [40] 和 [41, 第 240 页]) 考虑 $g\colon H\to(-\infty,\infty]$ 和 $\bar{x}\in\mathrm{dom}g$.

(i) 如果存在 $\rho\geqslant 0$ 和 $\delta>0$ 使得对任意的 $x\in B_\delta(\bar{x})$, 有

$$g(x)\geqslant g(\bar{x})+\langle u,x-\bar{x}\rangle-\frac{\rho}{2}\|x-\bar{x}\|^2, \tag{2.3}$$

则称 $u\in H$ 是 g 在 $\bar{x}$ 点的邻近次梯度 (proximal subgradient). g 在 $\bar{x}$ 点的所有邻近次梯度组成的集合称为 g 在 $\bar{x}$ 点的邻近次微分, 记作 $\partial_p g(\bar{x})$.

(ii) 如果对任意的 $\epsilon>0$, 存在 $\delta>0$ 使得对任意的 $x\in B_\delta(\bar{x})$, 有

$$g(x)\geqslant g(\bar{x})+\langle u,x-\bar{x}\rangle-\epsilon\|x-\bar{x}\|, \tag{2.4}$$

则称 $u\in H$ 是 g 在 $\bar{x}$ 点的 Fréchet 次梯度. g 在 $\bar{x}$ 点的所有 Fréchet 次梯度组成的集合, 记作 $\partial_F g(\bar{x})$, 称为 g 在 $\bar{x}$ 点的 Fréchet 次微分.

(iii) 集合

$$\partial g(\bar{x})=\operatorname*{Lim\,sup}_{x\to^g\bar{x}}\partial_p g(x):=\{w=w-\lim w_n,w_n\in\partial_p g(x_n),\exists\, x_n\to^g\bar{x}\}$$

称为 g 在 $\bar{x}$ 点的极限 (或 Mordukhovich) 次微分, 其中 $x_n\to^g\bar{x}$ 指的是 $\|x_n-\bar{x}\|\to 0$ 且 $g(x_n)\to g(\bar{x})$.

我们知道 (参见文献 [38,42]) 极限次微分也等于 Fréchet 次微分的弱序列上极限集, 即

$$\partial g(\bar{x})=\operatorname*{Lim\,sup}_{x\to^g\bar{x}}\partial_F g(x).$$

下面结果说明, 如果 g 是邻近有界的, 则适当放大 (2.3) 中的 ρ, 可以使得不等式 (2.3) 全局成立. 有限维情形参见文献 [30, 命题 8.46], Hilbert 空间情形参见文献 [37, 引理 4.1] 或 [32, 引理 3.7].

命题 2.1 若 g 是邻近有界的, 则 $v \in \partial_p g(\bar{x})$ 当且仅当存在某个充分小的 $\lambda > 0$, $\bar{w} \in \mathrm{prox}_{\lambda g}^{-1}(\bar{x})$ 使得 $v = \lambda^{-1}(\bar{w} - \bar{x})$ 且满足

$$g(x) \geqslant g(\bar{x}) + \langle v, x - \bar{x}\rangle - \frac{1}{2\lambda}\|x - \bar{x}\|^2, \quad \forall\, x \in H.$$

如果存在 $\epsilon > 0$, $\alpha > g(\bar{x})$ 使得 g 相对于集合 $\{x\colon \|x - \bar{x}\| < \epsilon; g(x) < \alpha\}$ 是下半连续的, 则称 g 在 $\bar{x}$ 点是局部下半连续的. 这等价于上图 epi$g$ 在 $(\bar{x}; g(\bar{x}))$ 点是局部闭的. 下面的邻近正则函数首先由 Poliquin 和 Rockafellar [31] 在 $\mathbb{R}^n$ 空间中引入, 继而由文献 [32,33,37,38] 推广到 Hilbert 空间, 并给出许多无限维空间的基本结果.

定义 2.5 一个函数 $g\colon H \to (-\infty, \infty]$ 称为在 $\bar{x} \in \mathrm{dom} g$ 点关于 $\bar{v} \in \partial g(\bar{x})$ 是邻近正则的, 如果 g 在 $\bar{x}$ 点是局部下半连续的, 且存在 $\rho > 0$ 和 $r > 0$ 使得对任意 x 和 $v \in \partial g(x)$ 满足 $\|x - \bar{x}\| < \rho$, $|g(x) - g(\bar{x})| < \rho$, $\|v - \bar{v}\| < \rho$, 有

$$g(x') > g(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2, \quad \forall\, x' \in B_\rho(\bar{x}). \tag{2.5}$$

如果对任意 $\bar{v} \in \partial g(\bar{x})$, 上面的式子都成立, 则称 g 在 x 点是邻近正则的.

定义 2.6 一个函数 $g\colon H \to (-\infty, \infty]$ 称为在 $x_0 \in \mathrm{dom} g$ 点是局部一致邻近正则的, 如果存在 $\delta > 0$, $\rho > 0$ 和 $r > 0$ 使得对任意 $\bar{x} \in B_\delta(x_0)$ 和 $\bar{v} \in \partial g(\bar{x})$, 以及对任意 x 和 $v \in \partial g(x)$ 满足 $\|x - \bar{x}\| < \rho$, $|g(x) - g(\bar{x})| < \rho$, $\|v - \bar{v}\| < \rho$, 有

$$g(x') > g(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2, \quad \forall\, x' \in B_\rho(\bar{x}). \tag{2.6}$$

文献 [38, 命题 3.3] 表明, g 在 x_0 点是局部一致邻近正则的当且仅当存在 $\delta > 0$ 和 $r > 0$ 使得对任意的 $x, x' \in B_\delta(x_0)$ 和 $v \in \partial g(x)$, 有

$$g(x') \geqslant g(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2.$$

定义 2.7 一个函数 $g\colon H \to (-\infty, \infty]$ 称为在 $\bar{x} \in \mathrm{dom} g$ 点关于 $\bar{v} \in \partial g(\bar{x})$ 是次微分连续的, 如果对任意的 $\epsilon > 0$, 存在 $\delta > 0$ 使得对任意的 x 及 $v \in \partial g(x)$ 满足 $\|x - \bar{x}\| < \delta$, $\|v - \bar{v}\| < \delta$, 有 $|g(x) - g(\bar{x})| < \epsilon$.

邻近正则函数一般不是次微分连续的. 例如, $f(x)$ 是定义在 $\mathbb{R}$ 上的函数, 如果当 $x > 0$ 时函数值为 1; 当 $x \leqslant 0$ 时函数值为 0, 则 $f(x)$ 在 $\bar{x} = 0$ 关于 $\bar{v} = 0$ 是邻近正则, 但不是次微分连续的. Poliquin [35] 引入的 p.l.n. (primal-lower-nice) 函数类既是邻近正则的, 也是次微分连续的.

定义 2.8 一个函数 $g\colon H \to (-\infty, \infty]$ 称为在 $\bar{x} \in \mathrm{dom} g$ 点是 p.l.n. 的, 如果存在 $R > 0$, $c > 0$ 和 $\delta > 0$, 使得对任意的 $r > R$ 和 x 及 $v \in \partial g(x)$ 满足 $\|x - \bar{x}\| < \delta$ 和 $\|v\| < cr$, 下面不等式成立:

$$g(x') \geqslant g(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2, \quad \forall\, x' \in B_\delta(\bar{x}). \tag{2.7}$$

文献 [31,37] 分别在 $\mathbb{R}^n$ 和 Hilbert 空间证明了, 如果 g 在 $\bar{x}$ 是 p.l.n. 的, 则存在 $\bar{x}$ 的某个邻域 V 使得对任意的 $x \in V$, $v \in \partial g(x)$, g 在 x 点关于 v 是邻近正则和次微分连续的. 显然, 每个凸函数和局部仿凸函数是 p.l.n. 的. 从文献 [35,43,44] 可知, 定义在 $\mathbb{R}^n$ 空间上和 Hilbert 空间上的强 amenable (见下面的定义) 函数是 p.l.n. 的, 进而是邻近正则且次微分连续的. 而文献 [37] 证明了 Hilbert 空间上的凸函数与 $C^{1,+}$ 映射的复合函数在 Robinson 约束品性条件下 (即下面强 amenable 函数定义中的 G 是 $C^{1,+}$, 而不必是二阶连续可微的) 是邻近正则和次微分连续的, 文献 [45] 进一步证明了 Banach 空间中这类凸与 $C^{1,+}$ 复合函数也是 p.l.n. 的. 相关的定义及性质参见文献 [30,31,37,44–48].

定义 2.9 [30,31,35] 一个函数 $g\colon H\to\bar{\mathbb{R}}$ 称为在 $\bar{x}$ 点是 amenable 的, 如果 $g(\bar{x})$ 是有限的, 存在 $\bar{x}$ 的一个开邻域 V, 使得 g 在 V 上能局部地表示为 $\psi\circ G$ 的形式且满足 Robinson 约束品性

$$0\in\operatorname{int}[G(\bar{x})+\nabla G(\bar{x})(H)-\operatorname{dom}\psi], \tag{2.8}$$

其中 G 是从 V 到 Y 的连续可微映射, ψ 是定义在 Y 上的下半连续真凸函数. 如果 G 是二次连续可微映射, 则称 g 在 $\bar{x}$ 点是强 amenable 的. 在强 amenable 函数类中一类重要的函数是完全 amenable 函数, 即在上面定义中, ψ 不只是凸的而且是逐片线性二次函数.

例 2.1 [30] (i) H 上的每个真凸下半连续函数是强 amenable 的, 每个逐片线性二次凸函数是完全 amenable 的.

(ii) 令 $g_i\colon H\to\mathbb{R}\cup\{+\infty\}$, 如果 $g_i\in C^1$, $i=1,2,\ldots,m$, 则 $g=\max\{g_1,\ldots,g_m\}$ 是 amenable 的; 如果 $g_i\in C^2$, 则 g 是完全 amenable 的.

(iii) 令 ϕ 为定义在 H 上的真凸下半连续函数, 如果 $g\in C^1$, 则 $\phi+g$ 是 amenable 的; 如果 $g\in C^2$, 则 $\phi+g$ 是强 amenable 的; 此外, 如果 ϕ 是逐片线性二次的, 则 $\phi+g$ 是完全 amenable 的.

设 C 是 H 的一个子集 (不必是凸的), 令 $N_C^p(x)$ ($N_C(x)$) 表示 C 在 x 点的邻近法锥 (极限法锥), 即 $N_C^p(x)=\partial_P\delta_C(x)$ ($N_C(x)=\partial\delta_C(x)$).

定义 2.10 [31,49] 一个集合 $C\subset H$ 称为在 $\bar{x}\in C$ 点关于 $\bar{v}\in N_C(\bar{x})$ 是邻近正则的, 如果 C 在 $\bar{x}$ 点是局部闭的, 且存在 $\rho>0$ 和 $r>0$ 使得对 $x\in C$ 和 $v\in N_C(x)$ 满足 $\|x-\bar{x}\|<\rho$ 和 $\|v-\bar{v}\|<\rho$, x 是 $x+r^{-1}v$ 在集合 $\{x'\in C\colon\|x'-\bar{x}\|<\rho\}$ 中的唯一最近点. 如果这个性质对每个 $v\in N_C(\bar{x})$ 都成立, 则称 C 在 $\bar{x}$ 点是邻近正则的.

显然, 每个闭凸集是邻近正则的. 文献 [31,49] 证明了 C 在 $\bar{x}\in C$ 点关于 $\bar{v}\in N_C(\bar{x})$ 是邻近正则的当且仅当 δ_C 在 $\bar{x}$ 点关于 $\bar{v}$ 邻近正则的. 文献 [49] 还证明了闭集 C 在 $\bar{x}$ 点是邻近正则的当且仅当对所有的 (或者存在) $\sigma>0$, 函数 $d(\cdot,C)^2+\sigma\|\cdot\|^2$ 在 $\bar{x}$ 点的某凸邻域上是凸的, 当且仅当存在 $\bar{x}$ 的某个邻域 V, 使得对任意的 $x\in C\cap V$ 有 $N_C^p(x)\neq\emptyset$ (或者 $N^F(x)\neq\emptyset$); 同时也证明了 C 在 $\bar{x}$ 点是邻近正则的当且仅当 δ_C 在 $\bar{x}$ 点是 p.l.n. 的, 当且仅当 C 在 $\bar{x}$ 点有所谓的 Shapiro 性质, 即存在常数 $k>0$ 和 $\delta>0$ 使得

$$d(y-x,T_C(x))\leqslant k\|y-x\|^2,\quad \text{对所有的 } x,y\in C\cap B_\delta(\bar{x}). \tag{2.9}$$

称集值映射

$$T_\epsilon(x)=\begin{cases}\{v\in\partial g(x)\colon\|v-\bar{v}\|<\epsilon\}, & \text{如果 } \|x-\bar{x}\|<\epsilon,\quad |g(x)-g(\bar{x})|<\epsilon,\\ \emptyset, & \text{否则}\end{cases} \tag{2.10}$$

为次微分映射 ∂g 在 $(\bar{x},\bar{v})$ 点的 g 相关 ϵ- 局部化. 当 g 在 $\bar{x}$ 点关于 $\bar{v}$ 是次微分连续时, 我们仅需考虑通常的 ϵ- 局部化 (不必考虑 $g(x)$ 的不等式).

邻近正则函数的一个最关键的性质是下面的等价刻画.

定理 2.1 (参见文献 [50, 定理 3.2] 和 [37, 定理 3.4]) 令 g 在 $\bar{x}\in\operatorname{dom}g$ 点为局部下半连续. g 在 $\bar{x}$ 点关于 $\bar{v}\in\partial g(\bar{x})$ 是邻近正则的当且仅当 $\bar{v}\in\partial_p g(\bar{x})$ 且存在 $\epsilon>0$ 和 $r>0$ 使得对 ∂g 的任意 g 相关 ϵ- 局部化 T_ϵ, $T_\epsilon+rI$ 是单调的, 也就是说, 对任意的 x_i 和 $v_i\in\partial g(x_i)$ $(i=1,2)$ 满足 $\|x_i-\bar{x}\|<\epsilon$, $|g(x_i)-g(\bar{x})|<\epsilon$, $\|v_i-\bar{v}\|<\epsilon$, 下面不等式成立:

$$\langle v_1-v_2,x_1-x_2\rangle\geqslant -r\|x_1-x_2\|^2.$$

当 g 是定义在 $\mathbb{R}^n$ 上的函数、$\bar{v}=0$ 时, 下面关于 Moreau 包络及邻近映射的结果首先是在文献 [31] 中给出的, 文献 [32,37,38] 将这些结果推广到 Hilbert 空间 H 和任意次梯度 $\bar{v}$ 的情形.

定理 2.2 假设 $g\colon H\to(-\infty,\infty]$ 为邻近有界的, 在 $\bar{x}\in\mathrm{dom}g$ 点关于 $\bar{v}\in\partial g(\bar{x})$ 是邻近正则的 (正则常数, 即定义 (2.6) 的常数, $r>0$), 则

(i) 存在次梯度映射 ∂g 在 $(\bar{x},\bar{v})$ 点的一个 g 相关 ϵ- 局部化 T_ϵ, 使得对任意的 $\tau\in(0,\frac{1}{r})$,

$$\mathrm{prox}_{\tau g}=(I+\tau T_\epsilon)^{-1}$$

在 $\bar{x}+\tau\bar{v}$ 的某个邻域内是非空单值 Lipschitz 的, 且 Lipschitz 常数为 $\frac{1}{1-\tau r}$, 特别地, $\mathrm{prox}_{\tau g}(\bar{x}+\tau\bar{v})=\bar{x}$;

(ii) 对任意的 $\tau\in(0,\frac{1}{r})$, 函数 $e_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 点是局部一致邻近正则的 (正则常数为 $\frac{r}{1-\tau r}$), 特别地, 在 $\bar{x}+\tau\bar{v}$ 的某个邻域内是仿凸且 $C^{1,+}$ 的, 具体来说, $e_\tau g+\frac{r}{2(1-\tau r)}\|\cdot\|^2$ 是凸函数;

$$\nabla e_\tau g=\frac{1}{\tau}(I-\mathrm{prox}_{\tau g})=\frac{1}{\tau}[I-(I+\tau T_\epsilon)^{-1}]$$

在 $\bar{x}+\tau\bar{v}$ 点是局部 Lipschitz 的, 特别地, $\nabla e_\tau g(\bar{x}+\tau\bar{v})=\bar{v}$.

注 2.1 注意到, 如果 g 在 $\bar{x}$ 点是 p.l.n. 的, 则存在 $\bar{x}$ 的某个邻域 V, 使得对任意的 $x\in V$, g 在 x 点关于任意的 $v\in\partial g(x)$ 是邻近正则、且次微分连续的. 于是, 在这种情形, 当 $(\bar{x},\bar{v})$ 被任意的 (x,v) (其中, $x\in V$, $v\in\partial g(x)$) 取代时, 上面的结论仍然成立.

给定常数 $\tau>0$ 和 $\bar{x}\in\mathrm{dom}g$ 的某个邻域 V, 考虑一个局部化的正则间隙函数 (3.1), 即

$$\gamma_\tau^l(x):=\inf_{y\in V}\left\{\langle F(x),y-x\rangle+\frac{1}{2\tau}\|x-y\|^2+g(y)\right\}. \tag{2.11}$$

注意到这个间隙函数的定义依赖于某个邻域 V, 当 g 是凸函数时, 定义 (2.11) 中的极小化问题就是凸优化问题, 于是在考虑间隙函数及其解映射的性质时不必限制在邻域 V 内.

命题 2.2 假设 g 在 $\bar{x}\in\mathrm{dom}g$ 点是 p.l.n. 的, F 在 $\bar{x}$ 点是局部 Lipschitz 的, 则存在 $\bar{x}$ 的某个邻域 V 和常数 $\tau>0$, 使得对任意的 $x\in V$, 有 $0\in F(x)+\partial g(x)$ 当且仅当 $\gamma_\tau^l(x)=g(x)$.

证明 由于 g 在 $\bar{x}\in\mathrm{dom}g$ 点是 p.l.n. 的, 则存在 $R>0$, $c>0$ 和 $\rho>0$, 使得对任意的 $r>R$、x 和 $v\in\partial g(x)$ 满足 $\|x-\bar{x}\|<\rho$ 和 $\|v\|<cr$, 下式成立:

$$g(y)\geqslant g(x)+\langle v,y-x\rangle-\frac{r}{2}\|y-x\|^2,\quad\forall\,y\in B_\rho(\bar{x}). \tag{2.12}$$

F 在 $\bar{x}$ 的局部 Lipschitz 性质蕴涵存在 $\bar{x}$ 的某个邻域 $V\subset B_\rho(\bar{x})$ 和常数 $L>0$, 使得对所有的 $x\in V$, 有 $\|F(x)\|\leqslant L$. 令 $0<\tau<\frac{1}{\max\{R,\frac{L}{c}\}}$, 则对任意的 $x\in V$, 如果 $-F(x)\in\partial g(x)$, 则有

$$g(y)+\langle F(x),y-x\rangle+\frac{1}{2\tau}\|y-x\|^2\geqslant g(x),\quad\forall\,y\in V. \tag{2.13}$$

于是, $\gamma_\tau^l(x)\geqslant g(x)$. 因为相反的不等式自然成立, 所以, $\gamma_\tau^l(x)=g(x)$. 反之, 如果 $\gamma_\tau^l(x)=g(x)$, 则 (2.11) 右端的下确界在 x 点可达. 于是, 最优性条件蕴涵 $0\in F(x)+\partial g(x)$. □

当 g 为 $\mathbb{R}^n$ 中某个凸集或邻近正则集的指示函数时, 上面结论分别在文献 [2,3,8,12] 被证明. 从命题 2.2 的证明可以看出, 如果 $0\in F(\bar{x})+\partial g(\bar{x})$, g 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是邻近正则且次微分连续的, 而且 F 在 $\bar{x}$ 点是局部 Lipschitz 连续的, 则命题 2.2 的结论仍然成立.

3 正则间隙函数的可微性

本节假设 $g\colon H\to(-\infty,\infty]$ 是真下半连续、邻近有界的, 且 $\bar{x}\in\mathrm{dom}g$, $0\in F(\bar{x})+\partial g(\bar{x})$. 假设 g 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是邻近正则的, $F(x)$ 在 $\bar{x}$ 点是局部 Lipschitz 连续的. 由于邻近正则性是在 $\bar{x}$ 点附近的局部性质, 所以可以将 g 限制在 $\bar{x}$ 的某个邻域上而不影响其邻近正则性. 换句话说, 令 V 为 $\bar{x}$ 的某个凸邻域, g 和 $g+I_V$ 在 x 点的邻近正则性是一致的. 所以, 不失一般性, 我们可以去掉定义 (2.11) 对邻域 V 的限制, 即当 x 充分接近 $\bar{x}$ 时, 局部化正则间隙函数可以表示成

$$\begin{aligned}\gamma_\tau(x)&=\inf_{y\in H}\left\{\langle F(x),y-x\rangle+\frac{1}{2\tau}\|x-y\|^2+g(y)\right\}\\&=e_\tau g(x-\tau F(x))-\frac{\tau}{2}\|F(x)\|^2.\end{aligned}\tag{3.1}$$

于是, 当 x 充分接近 $\bar{x}$ 时, 定义 (2.11) 中优化问题的解映射 P_τ 可以表示为 $P_\tau(x)=\mathrm{prox}_{\tau g}(x-\tau F(x))$ (参见文献 [32,51]).

3.1 正则间隙函数的一阶可微性

作为定理 2.2 的直接应用, 我们可以得到下面结果.

定理 3.1 假设 g 在 $\bar{x}$ 点关于 $-F(x)$ 是邻近正则的 (正则常数为 r), 则

(i) 对任意的 $\tau\in(0,\frac{1}{r})$, 对所有 x 充分接近 $\bar{x}$, 极小化问题 (3.1) 有唯一解 $P_\tau(x)=\mathrm{prox}_{\tau g}(x-\tau F(x))$, 且 $P_\tau(x)$ 在 $\bar{x}$ 点是局部 Lipschitz 连续的, $\gamma_\tau(\bar{x})=g(\bar{x})$;

(ii) 如果 F 在 $\bar{x}$ 点的某个邻域上是 C^1 (或者 $C^{1,+}$) 的, 则对任意的 $\tau\in(0,\frac{1}{r})$, 函数 γ_τ 在 $\bar{x}$ 的某个邻域内是 C^1 (或者 $C^{1,+}$) 的, 且

$$\nabla\gamma_\tau(x)=\left(\frac{1}{\tau}I-DF(x)\right)(x-\mathrm{prox}_{\tau g}(x-\tau F(x)))-F(x)$$

对所有 x 充分接近 $\bar{x}$ 都成立.

当函数 g 是 $\mathbb{R}^n$ 上的凸函数或 $\mathbb{R}^n$ 中某个邻近正则集合的指示函数时, γ_τ 的可微性已经分别在文献 [3,12,27,52] 中给出. 当 F 在 $\bar{x}$ 点是局部 Lipschitz 的情形下, 我们也能证明 γ_τ 在 $\bar{x}$ 点是 Fréchet 可微的.

定理 3.2 假设 g 在 $\bar{x}$ 点关于 $-F(x)$ 是邻近正则的 (正则常数为 r), 那么对任意的 $\tau\in(0,\frac{1}{r})$, 函数 γ_τ 在 $\bar{x}$ 点是 Fréchet 可微的, 且 $\nabla\gamma_\tau(\bar{x})=-F(\bar{x})$.

证明 对固定的 $\tau\in(0,\frac{1}{r})$, 由于 $e_\tau g$ 在 $\bar{x}-\tau F(\bar{x})$ 附近是 $C^{1,+}$ 的, 存在某个常数 $C>0$, 使得当 $\|h\|$ 充分小时, 有

$$\begin{aligned}&\|e_\tau g(\bar{x}-\tau F(\bar{x})+h)-e_\tau g(\bar{x}-\tau F(\bar{x}))-\nabla e_\tau g(\bar{x}-\tau F(\bar{x}))h\|\\&\leqslant\|h\|\int_0^1\|\nabla e_\tau g(\bar{x}-\tau F(\bar{x})+th)-\nabla e_\tau g(\bar{x}-\tau F(\bar{x}))\|dt\\&\leqslant C\|h\|^2.\end{aligned}$$

注意到 $\nabla e_\tau g(\bar{x}-\tau F(\bar{x}))=-F(\bar{x})$ 以及 F 在 $\bar{x}$ 点是局部 Lipschitz 连续的, 于是, 对 ξ 充分接近 0, 有

$$e_\tau g(\bar{x}+\xi-\tau F(\bar{x}+\xi))-e_\tau g(\bar{x}-\tau F(\bar{x}))$$

$$= \langle -F(\bar{x}), \xi - \tau(F(\bar{x}+\xi) + F(\bar{x}))\rangle + O(\|\xi - \tau(F(\bar{x}+\xi) + F(\bar{x}))\|^2).$$

进而由 (3.1) 得, 对 ξ 充分接近 0, 有

$$\begin{aligned}
&\gamma_\tau(\bar{x}+\xi) - \gamma_\tau(\bar{x}) \\
&\quad = \langle -F(\bar{x}), \xi\rangle - \frac{\tau}{2}\|F(\bar{x}+\xi) - F(\bar{x}))\|^2 + O(\|\xi - \tau(F(\bar{x}+\xi) - F(\bar{x}))\|^2) \\
&\quad = \langle -F(\bar{x}), \xi\rangle + O(\|\xi\|^2).
\end{aligned}$$

这蕴涵 γ_τ 在 $\bar{x}$ 点是 Fréchet 可微的, 且 $\nabla\gamma_\tau(\bar{x}) = -F(\bar{x})$. □

3.2 正则间隙函数的二阶可微性

本小节首先回顾广义实值函数的二阶上图导数的定义及性质, 讨论邻近正则函数 Moreau 包络函数的二阶可微性及邻近映射的 proto- 可微性, 进而探讨正则间隙函数的二阶可微性.

令 $\{S_t\}_{t>0} \subset H$ 为一族集合, 称

$$\begin{aligned}
&\operatorname{Lim\,sup} S_t := \{\xi \mid \exists \text{ 序列 } t_k \downarrow 0,\ \xi_k \to \xi, \text{ 满足 } \xi_k \in S_{t_k}, k = 1, 2, \ldots\}, \\
&\operatorname{Lim\,inf} S_t := \{\xi \mid \forall\, t_k \downarrow 0, \exists\, \xi_k \to \xi \text{ 满足 } \xi_k \in S_{t_k}, k = 1, 2, \ldots\}
\end{aligned}$$

为集合族 S_t 在 $t \downarrow 0$ 时的上极限集和下极限集. 若 $S = \operatorname{Lim\,sup}_{t\downarrow 0} S_t = \operatorname{Lim\,inf}_{t\downarrow 0} S_t$, 则称 S_t 在 Painlevé-Kuratowski 意义下收敛于集合 S, 记作 $S = \operatorname{Lim}_{t\downarrow 0} S_t$.

考虑一族广义实值函数 $\varphi, \varphi_t \colon H \to \bar{\mathbb{R}}$. 如果 $\operatorname{epi}\varphi = \operatorname{Lim}_{t\downarrow 0} \operatorname{epi}\varphi_t$, 则称当 $t \downarrow 0$ 时, φ_t 上图收敛于函数 φ, 记作 $\varphi = \text{epi-}\lim_{t\downarrow 0} \varphi_t$.

函数的上图极限可以借助函数值的极限来表示 (参见文献 [30]). 例如, φ_t 上图收敛于 φ 当且仅当对任意的 $t_k \downarrow 0$ 和任意 $\xi \in H$, 有

$$\begin{aligned}
&\liminf_{k\to\infty} \varphi_{t_k}(\xi_k) \geqslant \varphi(\xi), \quad \text{对任意的序列 } \xi_k \to \xi, \\
&\limsup_{k\to\infty} \varphi_{t_k}(\xi_k) \leqslant \varphi(\xi), \quad \text{对某个序列 } \xi_k \to \xi.
\end{aligned}$$

上面的不等式事实上蕴涵至少对某个序列 $\xi_k \to \xi$ 有 $\varphi_{t_k}(\xi_k)$ 收敛于 $\varphi(\xi)$. 关于上图收敛更详细的性质可以参见文献 [29, 30]. 下面上图导数的定义首先是 Rockafellar [47] 给出的. 令 $g \colon H \to \bar{\mathbb{R}}$ 为广义实值函数, $\bar{x} \in \operatorname{dom} g$ 为给定的点且 $g(\bar{x})$ 是有限数.

定义 3.1 称函数 g 在 x 点是上图可微的, 如果一阶差商函数

$$\varphi_{x,t}(\xi) = \frac{g(x+t\xi) - g(x)}{t}, \quad t > 0$$

上图收敛于某个函数 $g'_x \colon H \to \bar{\mathbb{R}}$, 即 $g'_x = \text{epi-}\lim_{t\downarrow 0} \varphi_{x,t}$ 且 $g'_x(0) > -\infty$. 称 $g'_x(\xi)$ 为 g 在 x 点的一阶上图导数. 进一步, 如果向量 $v \in H$ 满足

$$g'_x(\xi) \geqslant \langle v, \xi\rangle, \quad \text{对所有的 } \xi \in H,$$

则称其为 g 在 x 点的上图梯度.

定义 3.2 称函数 g 在 x 点关于 $v \in H$ 是二阶上图可微的, 如果 g 在 x 点是一阶上图可微的, 二阶差商函数

$$\varphi_{x,v,t}(\xi) = \frac{g(x+t\xi) - g(x) - t\langle \xi; v\rangle}{\frac{1}{2}t^2}$$

上图收敛于某个函数 $g''_{x,v}\colon H \to \bar{\mathbb{R}}$, 即 $g''_{x,v} = \text{epi-}\lim_{t\downarrow 0} \varphi_{x,v,t}$ 且 $g''_{x,v}(0) > -\infty$. 称 $g''_{x,v}(\xi)$ 为 g 在 x 点关于 v 的二阶上图方向导数. 如果 g 在 x 点关于任意的上图梯度 v 是二阶上图可微的 (前提是, g 在 x 点是上图可微的且至少存在一个上图梯度), 则称函数 g 在 x 点是二阶上图可微的.

函数的一阶和二阶上图导数的性质及很多上图可微的函数的例子可参见文献 [46, 47]. 下面仅列出几个在后面经常使用的性质.

命题 3.1 [47] 一阶上图导数 g'_x (若存在) 是下半连续正齐次的, 二阶上图导数 $g''_{x,v}$ (若存在) 是正二次齐次的. 显然, $g'_x(0) > -\infty$ 等价于 $g'_x(\xi) > -\infty$, $\forall \xi \in \mathbb{R}^n$, $g'_x(0) = 0$; $g''_{x,v}(0) > -\infty$ 蕴涵 $g''_{x,v}(\xi) > -\infty$, $\forall \xi \in \mathbb{R}^n$, $g''_{x,v}(0) = 0$.

注 3.1 从定义 3.1 和 3.2, 注意到, 如果 g 在 x 点关于次梯度 v 是二次上图可微的, 则

$$\{\xi \in H\colon g''_{x,v}(\xi) < \infty\} \subset \{\xi \in H\colon g'_x(\xi) = \langle v, \xi\rangle\} =: C_g(x, v).$$

命题 3.1 确保 $g''_{x,v}$ 在任何紧集上达到它的极小值. 从文献 [53, 定理 2.2] 知, 如果一个凸函数 g 在 x 点关于次梯度 v 是二阶上图可微的, 则 v 一定是 g 在 x 点在凸分析意义下的次梯度, 即

$$g(x') \geqslant g(x) + \langle v, x' - x\rangle, \quad \forall\, x' \in \mathbb{R}^n,$$

并且它的二阶上图导数 $g''_{x,v}$ 是非负下半连续真凸函数, 且满足 $g''_{x,v}(0) = 0$.

设 $g\colon H \to \mathbb{R}$ 在 x 点是 Gâteaux 可微的, 如果对任意的 $h \in H$, 极限

$$g''(x, h) = \lim_{\substack{t\downarrow 0\\ h'\to h}} \frac{g(x+th') - g(x) - t\langle \nabla g(x), h'\rangle}{\frac{1}{2}t^2}$$

存在, 则称 g 在 x 点是二阶 Hadamard 方向可微的. 从定义可知, 如果 $g''(x, h)$ 存在, 则它关于 h 一定是连续. 显然 g 在 x 点是二阶 Hadamard 方向可微的, 则它在 x 点关于 $\nabla g(x)$ 一定是二阶上图可微的.

设 $T\colon U \rightrightarrows H$ 是从 Hilbert 空间 U 到 H 的集值映射, $\bar{x} \in T(\bar{u})$. T 在 $\bar{u}$ 关于 $\bar{x}$ 的余切导数是一个从 U 到 H 的集值映射, 记作 $DT(\bar{u}, \bar{x})$, 定义为

$$\text{gph}DT(\bar{u}, \bar{x}) = \operatorname*{Lim\,sup}_{t\,\downarrow 0} \frac{\text{gph}T - (\bar{u}, \bar{x})}{t}.$$

如果集值映射

$$\Delta_{\bar{u},\bar{x},t}T\colon \xi \rightrightarrows \frac{T(\bar{u}+t\xi) - \bar{x}}{t} \tag{3.2}$$

当 $t \downarrow 0$ 时图收敛, 即极限

$$\operatorname*{Lim}_{t\downarrow 0} \text{gph}\Delta_{\bar{u},\bar{x},t}T = \operatorname*{Lim}_{t\downarrow 0} \frac{\text{gph}T - (\bar{u}, \bar{x})}{t}$$

存在, 则称 T 在 $\bar{u}$ 点关于 $\bar{x}$ 是 proto- 可微的, 称其极限, 记作 $T'_{\bar{u},\bar{x}}$, 为 T 在 $\bar{u}$ 关于 $\bar{x}$ 的 proto- 导数. 当 $T(\bar{u}) = \{\bar{x}\}$ 是单点集时, 简记为 $T'_{\bar{u}}(\xi)$, 且仅称 T 在 $\bar{u}$ 点是 proto- 可微的. 显然, 如果 T 在 $\bar{u}$ 点关于 $\bar{x}$ 是 proto- 可微的, 则 $DT(\bar{u}, \bar{x})(\cdot) = T'_{(\bar{u},\bar{x})}(\cdot)$. 如果 T 在 $\bar{u}$ 的某个邻域是单值的, 且 T 在 $\bar{u}$ 点是 Hadamard 方向可微的, 则 T 在 $\bar{u}$ 点一定是 proto- 可微的且 $T'_{\bar{u}}(\xi) = T'(\bar{u}, \xi)$, 反之也成立, 如果 T 在 $\bar{u}$ 点是局部 Lipschitz 连续的.

定理 3.3 (参见文献 [31, 定理 6.1]) 假设 T_ϵ 是 ∂g 在 $(\bar{x},\bar{v})$ 点的一个 g 相关 ϵ- 局部化, 则 g 在 $\bar{x}$ 点关于 $\bar{v}\in\partial g(\bar{x})$ 是二阶上图可微的当且仅当 T_ϵ 在 $\bar{x}$ 点关于 $\bar{v}$ 是 proto- 可微的, 此时有

$$(T_\epsilon)'_{\bar{x},\bar{v}}(\xi)=\frac{1}{2}\partial[g''_{\bar{x},\bar{v}}](\xi),\quad \forall\,\xi\in\mathbb{R}^n.$$

此外, 如果 g 在 $\bar{x}$ 点关于 $\bar{v}$ 是次微分连续的, 则上面结论对 $T_\epsilon=\partial g$ 也成立.

文献 [30, 推论 13.15 和 命题 13.32] 表明, 如果 g 在 $\bar{x}$ 点是完全 amenable 的, 则存在 $\bar{x}$ 点某个邻域 V 使得 g 在任意的 $x\in V\cap(\mathrm{dom}g)$ 关于任意的 $v\in\partial g(x)$ 都是邻近正则且次微分连续的. 文献 [47,48] 证明了完全 amenable 函数类是二阶上图可微的, 同时也给出了该类函数 g 的二阶上图导数及其次微分的公式, 并证明了该次微分等于 ∂g 的 proto- 导数.

文献 [50, 定理 3.5] 给出了下面引理在 $\bar{v}=0$ 情形的证明, 但其证明引用了许多其他文献的结论而且详细证明隐含在文献 [31, 定理 6.1] 中. 为了便于理解, 我们给出引理 3.1 的详细证明.

引理 3.1 假设 g 在 $\bar{x}\in\mathrm{dom}g$ 点关于 $\bar{v}\in\partial g(\bar{x})$ 是二阶上图可微且邻近正则的 (常数为 r), 则对任意的 $\tau\in(0,\frac{1}{r})$, 下列等价性质成立:

(i) $e_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 点关于 $\bar{v}$ 是二阶上图可微的;

(ii) $e_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 点关于 $\bar{v}$ 是二阶 Hadamard 方向可微的;

(iii) $\nabla e_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 点关于 $\bar{v}$ 是 proto- 可微的;

(iv) $\mathrm{prox}_{\tau g}$ 在 $\bar{x}+\tau\bar{v}$ 点关于 $\bar{v}$ 是 proto- 可微的.

在上面等价条件成立的情形下, 有

$$(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\xi)=e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}(\xi)=\inf_{h\in\mathbb{R}^n}\left\{g''_{\bar{x},\bar{v}}(h)+\frac{1}{\tau}\|\xi-h\|^2\right\}\tag{3.3}$$

且 $g''_{\bar{x},\bar{v}}(\cdot)+r\|\cdot\|^2$ 是非负凸函数, 这说明 $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\xi)$ 既是 C^{1+} 的, 也是仿凸的且满足

$$\nabla e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}=\frac{2}{\tau}\left[I-\left(I+\frac{\tau}{2}D\right)^{-1}\right],\quad \text{其中 } D:=\partial g''_{\bar{x},\bar{v}}.\tag{3.4}$$

进一步有下面展开式:

$$e_\tau g(\bar{x}+\tau\bar{v}+\xi)=e_\tau g(\bar{x}+\tau\bar{v})+\langle\bar{v},\xi\rangle+e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}(\xi)+o(\|\xi\|^2),\tag{3.5}$$

$$\nabla e_\tau g(\bar{x}+\tau\bar{v}+t\xi)=\bar{v}+\frac{1}{2}t\nabla e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}(\xi)+o(\|t\xi\|),\tag{3.6}$$

$$\mathrm{prox}_{\tau g}(\bar{x}+\tau\bar{v}+t\xi)=\bar{x}+t\left[I-\frac{1}{2}\tau\nabla e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}\right](\xi)+o(\|t\xi\|).\tag{3.7}$$

证明 根据定理 2.2, 有

$$e_\tau g(\bar{x}+\tau\bar{v})=g(\bar{x})+\frac{\tau}{2}\|\bar{v}\|^2\quad \text{且}\quad \nabla e_\tau g(\bar{x}+\tau\bar{v})=\bar{v}.\tag{3.8}$$

考虑下面的二阶差商函数:

$$\Delta_t^2 g(\xi):=\frac{g(\bar{x}+t\xi)-g(\bar{x})-t\langle\bar{v},\xi\rangle}{\frac{1}{2}t^2},$$

$$\Delta_t^2 e_\tau g(\xi):=\frac{e_\tau g(\bar{x}+\tau\bar{v}+t\xi)-e_\tau g(\bar{x}+\tau\bar{v})-t\langle\bar{v},\xi\rangle}{\frac{1}{2}t^2}.$$

容易证明 $\Delta_t^2 e_\tau g = e_{\frac{\tau}{2}}\Delta_t^2 g$ 为 $\Delta_t^2 g$ 的包络函数. 事实上, 根据包络函数的定义及 (3.8), 有

$$\begin{aligned}\Delta_t^2 e_\tau g(\xi) &= \frac{2}{t^2}\inf_{h\in\mathbb{R}^n}\left\{g(h)+\frac{1}{2\tau}\|\bar{x}+\tau\bar{v}+t\xi-h\|^2-\left(g(\bar{x})+\frac{\tau}{2}\|\bar{v}\|^2\right)-t\langle\bar{v},\xi\rangle\right\}\\ &= \frac{2}{t^2}\inf_{h\in\mathbb{R}^n}\left\{g(\bar{x}+th)+\frac{1}{2\tau}\|(\bar{x}+\tau\bar{v}+t\xi)-(\bar{x}+th)\|^2-\left[g(\bar{x})+\frac{\tau}{2}\|\bar{v}\|^2\right]-t\langle\bar{v},\xi\rangle\right\}\\ &= \inf_{h\in\mathbb{R}^n}\left\{\frac{g(\bar{x}+th)-g(\bar{x})-t\langle\bar{v},h\rangle}{\frac{1}{2}t^2}+\frac{1}{\tau}\|\xi-h\|^2\right\}.\end{aligned}$$

由命题 2.1 知, 存在充分小的 $\lambda_0>0$, 使得

$$\Delta_t^2 g(\xi)\geqslant -\frac{1}{\lambda_0}\|\xi\|^2,\quad \forall\,\xi\in\mathbb{R}^n,\quad t>0. \tag{3.9}$$

于是, g 在 $\bar{x}$ 点关于 $\bar{v}$ 是二阶上图可微的当且仅当 $t\to 0$ 时二阶差商函数 $\Delta_t^2 g(\cdot)$ 上图收敛于 $g''_{\bar{x},\bar{v}}(\cdot)$. 根据文献 [30, 定理 7.37], 一致邻近有界条件 (3.9) 及 $\Delta_t^2 g(\cdot)$ 上图收敛于 $g''_{\bar{x},\bar{v}}(\cdot)$ 等价于 $g''_{\bar{x},\bar{v}}$ 是邻近有界的, 当 $t\to 0$ 时 $\Delta_t^2 e_\tau g(\cdot)$ 在 $\mathbb{R}^n$ 上逐点收敛, 并且极限为 $e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}$; 同时也说明收敛不仅是逐点收敛而且是连续收敛的, 这蕴涵当 $t\to 0$ 时 $\Delta_t^2 e_\tau g(\cdot)$ 在 $\mathbb{R}^n$ 的有界子集上一致收敛于 $e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}$, 进而也是上图收敛的. 这说明 $e_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 点关于 $\bar{v}$ 是二阶 Hadamard 方向可微的, 也是二阶上图可微的且 $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\xi)=e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}$ 以及展开 (3.5) 成立. 显然, 在这种情形, (i) 与 (ii) 是等价的.

根据定理 2.2(ii) 可知, 对任意的 $\tau\in(0,\frac{1}{r})$, 函数 $\hat{e}_\tau g:=e_\tau g+\frac{r}{2(1-\tau r)}\|\cdot\|^2$ 在 $\bar{x}+\tau\bar{v}$ 的某个邻域上是凸和 C^{1+} 的, 并且 $\nabla\hat{e}_\tau g=\nabla e_\tau g+\frac{r}{(1-\tau r)}I$. 由定理 3.3 知, $\hat{e}_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 是二阶上图可微的当且仅当 $\nabla\hat{e}_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 是 proto- 可微的, 并且

$$(\nabla e_\tau g)'_{\bar{x}+\tau\bar{v}}(\xi)=\partial\left[\frac{1}{2}(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}\right](\xi)=\partial\left[\frac{1}{2}e_{\frac{\tau}{2}}(g''_{\bar{x},\bar{v}})\right](\xi),\quad \forall\,\xi\in\mathbb{R}^n.$$

这说明 (ii) 与 (iii) 是等价的, 并且 (3.6) 成立. 进一步, 由注 3.1 知, $\hat{e}_\tau g$ 在 $\bar{x}+\tau\bar{v}$ 的二阶上图导数 $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\cdot)+\frac{r}{(1-\tau r)}\|\cdot\|^2$ 是非负下半连续真凸的, 于是, $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\cdot)$ 是仿凸的, 进而由 $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\xi)=e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}$ 和文献 [32, 定理 2.3, 命题 2.4] 知, $(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\cdot)$ 是 $C^{1,+}$ 的. 因为当 $\tau\to 0$ 时,

$$(e_\tau g)''_{\bar{x}+\tau\bar{v},\bar{v}}(\cdot)+\frac{r}{(1-\tau r)}\|\cdot\|^2=e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}(\cdot)+\frac{r}{(1-\tau r)}\|\cdot\|^2$$

逐点递增收敛于 $g''_{\bar{x},\bar{v}}(\cdot)+r\|\cdot\|^2$, 所以, $g''_{\bar{x},\bar{v}}(\cdot)+r\|\cdot\|^2$ 也是非负凸的, 即 $g''_{\bar{x},\bar{v}}(\cdot)$ 是仿凸的.

再次应用定理 3.3, 有

$$\nabla e_{\frac{\tau}{2}}g''_{\bar{x},\bar{v}}=\frac{2}{\tau}\left[I-\left(I+\frac{\tau}{2}D\right)^{-1}\right],\quad \text{其中 } D:=\partial g''_{\bar{x},\bar{v}}.$$

于是, (iii) 和 (iv) 及 (3.6) 和 (3.7) 的等价性可以直接从 $\nabla e_\tau g$ 与 $\mathrm{prox}_{\tau g}$ 的关系得到. □

例 3.1 (参见文献 [30, 命题 13.33, 练习 7.28 和 8.31]) 假设 $g_i\colon\mathbb{R}^n\to\mathbb{R}$ 是 C^2 的, $i=1,2,\ldots,$ $g=\max\{g_1,\ldots,g_m\}$, 则 g 在 $\mathbb{R}^n$ 上的任一点 $\bar{x}$ 是邻近正则的、一阶 Hadamard 可微的, 且关于任意次梯度 $v\in\partial g(\bar{x})$ 是二阶上图可微的. 进一步, 令 $I(\bar{x}):=\{i\colon g_i(\bar{x})=g(x)\}$, 则有

$$g'_{\bar{x}}(\xi)=\max_{i\in I(\bar{x})}\{\langle\nabla g_i(\bar{x}),\xi\rangle\},\quad \partial g(\bar{x})=\mathrm{conv}\{\nabla g_i(\bar{x})\colon i\in I(\bar{x})\};$$

令 $I'(\bar{x},\xi):=\{i\in I(\bar{x})\colon\langle\nabla g_i(\bar{x}),\xi\rangle=g'_{\bar{x}}(\xi)\}$, 则有

$$g''_{\bar{x},v}(\xi)=I_{K(\bar{x},v)}(\xi)+\max\left\{\sum_{i\in I(\bar{x})}\alpha_i\langle\xi,\nabla^2 g_i(\bar{x})\xi\rangle\colon\alpha_i\geqslant 0,\ \sum_{i\in I(\bar{x})}\alpha_i=1,\ \sum_{i\in I(\bar{x})}\alpha_i\nabla g_i(\bar{x})=v\right\}$$

且 $\xi \in K(\bar{x}, v) \Leftrightarrow v \in \operatorname{conv}\{\nabla g_i(\bar{x}) \colon i \in I'(\bar{x}, \xi)\}$.

作为上面引理的直接应用, 我们得到正则间隙函数 γ_τ 的二阶可微性.

定理 3.4 假设 $F\colon \mathbb{R}^n \to \mathbb{R}^n$ 在 $\bar{x}$ 点是 Hadamard 方向可微的, 且 $-F(\bar{x}) \in \partial g(\bar{x})$, 假设 g 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是二阶上图可微且邻近正则的 (正则常数为 r), 则对任意的 $\tau \in (0, \frac{1}{r})$, 下列结论成立:

(i) γ_τ 在 $\bar{x}$ 点是二阶 Hadamard 方向可微的, 且二阶导数满足

$$\begin{aligned}\gamma_\tau''(\bar{x}, \xi) &= \inf_{h \in \mathbb{R}^n} \left\{ g''_{\bar{x}, -F(\bar{x})}(h) + 2\langle F'(\bar{x}, \xi), h - \xi \rangle + \frac{1}{\tau}\|h - \xi\|^2 \right\} \\ &= \inf_{h \in C(\bar{x}, F(\bar{x}))} \left\{ g''_{\bar{x}, -F(\bar{x})}(h) + 2\langle F'(\bar{x}, \xi), h - \xi \rangle + \frac{1}{\tau}\|h - \xi\|^2 \right\}; \end{aligned} \tag{3.10}$$

(ii) 邻近解映射 P_τ 在 $\bar{x}$ 点是 proto- 可微的, 其导算子满足

$$(P_\tau)'_{\bar{x}}(\xi) = \left[I + \frac{\tau}{2} \partial g''_{\bar{x}, -F(\bar{x})} \right]^{-1} (\xi - \tau F'(\bar{x}, \xi)).$$

证明 (i) 对任意 $\xi \in \mathbb{R}^n$ 和 $t > 0$, 根据展开 (3.5), 有

$$\begin{aligned} & e_\tau g(\bar{x} + t\xi - \tau F(\bar{x} + t\xi)) - e_\tau g(\bar{x} - \tau F(\bar{x})) \\ &= \langle F(\bar{x}), t\xi + \tau(F(\bar{x} + t\xi) - F(\bar{x})) \rangle + \frac{t^2}{2} e_{\frac{\tau}{2}} g''_{\bar{x}, \bar{v}} \left(\xi - \tau \frac{F(\bar{x} + t\xi) - F(\bar{x})}{t} \right) \\ &\quad + o\left(t^2 \left\| \xi - \tau \frac{F(\bar{x} + t\xi) + F(\bar{x})}{t} \right\|^2 \right). \end{aligned}$$

于是, 由 (3.1) 得

$$\begin{aligned} & \gamma_\tau(\bar{x} + t\xi) - \gamma_\tau(\bar{x}) - t\langle -F(\bar{x}), \xi \rangle \\ &= \frac{t^2}{2} e_{\frac{\tau}{2}} g''_{\bar{x}, -F(\bar{x})} \left(\xi - \tau \frac{F(\bar{x} + t\xi) + F(\bar{x})}{t} \right) - \frac{\tau}{2} \|F(\bar{x} + t\xi) - F(\bar{x})\|^2 \\ &\quad + o\left(t^2 \left\| \xi - \tau \frac{F(\bar{x} + t\xi) - F(\bar{x})}{t} \right\|^2 \right). \end{aligned} \tag{3.11}$$

注意到 $e_{\frac{\tau}{2}} g''_{\bar{x}, -F(\bar{x})}(\cdot)$ 是连续的, 而 F 在 $\bar{x}$ 点是 Hadamard 方向可微的, (3.11) 蕴涵 γ_τ 在 $\bar{x}$ 点是二阶 Hadamard 方向可微的, 并且

$$\begin{aligned} \gamma_\tau''(\bar{x}, \xi) &= e_{\frac{\tau}{2}} g''_{\bar{x}, -F(\bar{x})}(\xi - \tau F'(\bar{x}, \xi)) + \tau \|F'(\bar{x}, \xi)\|^2 \\ &= \inf_{h \in \mathbb{R}^n} \left\{ g''_{\bar{x}, -F(\bar{x})}(h) + 2\langle F'(\bar{x}, \xi), h - \xi \rangle + \frac{1}{\tau}\|h - \xi\|^2 \right\}. \end{aligned}$$

从注 3.1 知, $\operatorname{dom} g''_{\bar{x}, -F(\bar{x})} \subset C_g(\bar{x}, -F(\bar{x}))$. 于是 (3.10) 成立.

(ii) 由 (1.14) 知,

$$P_\tau(x) = \operatorname{prox}_{\tau g}(x - \tau F(x)).$$

根据 (3.7), 有

$$P_\tau(\bar{x} + t\xi) = \bar{x} + t \left[I - \frac{1}{2} \tau \nabla e_{\frac{\tau}{2}} g''_{\bar{x}, -F(\bar{x})} \right] \left(\xi - \tau \frac{F(\bar{x} + t\xi) + F(\bar{x})}{t} \right)$$

$$+o\left(t\left\|\xi-\tau\frac{F(\bar{x}+t\xi)-F(\bar{x})}{t}\right\|\right).$$

这说明 P_τ 在 $\bar{x}$ 点是 Hadarmad 方向可微的, 进而是 proto- 可微的, 并且

$$(P_\tau)'_{\bar{x}}(\xi)=\left[I-\frac{1}{2}\tau\nabla e_{\frac{\tau}{2}}g''_{\bar{x},-F(\bar{x})}\right](\xi-\tau F'(\bar{x},\xi))=\left[I+\frac{\tau}{2}\partial g''_{\bar{x},-F(\bar{x})}\right]^{-1}(\xi-\tau F'(\bar{x},\xi)).$$

证毕. □

4 变分包含解的局部唯一性及对参数的连续性

本节将讨论变分包含 (1.11) 解唯一的条件. 令 $g\colon\mathbb{R}^n\to(-\infty,+\infty]$ 为真下半连续邻近有界函数, $F\colon\mathbb{R}^n\to\mathbb{R}^n$ 为局部 Lipschitz 连续映射. 令 $\bar{x}$ 为 (1.11) 的一个给定的解, 即 $0\in F(\bar{x})+\partial g(\bar{x})$. 假设 g 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是邻近正则且次微分连续的且 F 在 $\bar{x}$ 是 Hadamard 方向可微的.

由命题 2.2 可知, $\bar{x}$ 是 (1.11) 的解当且仅当存在某个 $\tau>0$ 使得 $\bar{x}$ 是下面极小化问题:

$$\min\theta_\tau(x):=g(x)-\gamma_\tau(x)\tag{4.1}$$

的局部最优解且 $\theta_\tau(\bar{x})=0$.

称优化问题 (4.1) 在 $\bar{x}$ 处满足二阶增长条件, 如果存在 $c>0$ 和 $\bar{x}$ 的一个邻域 V 使得

$$\theta_\tau(x)\geqslant\theta_\tau(\bar{x})+c\|x-\bar{x}\|^2\tag{4.2}$$

对所有 $x\in V$ 成立. 显然, 如果 $\bar{x}$ 是 (1.11) 的一个解, 则二阶增长条件 (4.2) 蕴涵 $\bar{x}$ 是 (1.11) 的局部唯一解.

定理 4.1 令 $\bar{x}\in\mathrm{dom}g$ 为变分包含 (1.11) 的解. 假设 g 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是二阶上图可微的, 则对任意充分小的 $\tau>0$, 下面陈述等价, 且蕴涵 $\bar{x}$ 是 (1.11) 的局部唯一解.

(i) 二阶增长条件 (4.2) 成立;

(ii) $g''_{\bar{x},F(\bar{x})}(\xi)-\gamma''_\tau(\bar{x},\xi)>0,\forall\,\xi\in C_g(\bar{x},\bar{v})\setminus\{0\}$;

(iii) $\xi=0$ 是下面变分包含:

$$0\in F'(\bar{x},\xi)+(\partial g)'_{\bar{x},F(\bar{x})}(\xi)$$

的唯一解.

证明 (i) $\Leftrightarrow$ (ii). 显然 θ_τ 是真下半连续的. 因为 $\bar{x}\in\mathrm{dom}g$ 是 (1.11) 的解, 所以有 $0\in F(\bar{x})+\partial g(\bar{x})$. 根据定理 3.1 可知, γ_τ 在 $\bar{x}$ 点是连续可微的, 且 $\nabla\gamma_\tau(\bar{x})=-F(\bar{x})$. 于是, 根据文献 [41, 命题 1.107] 知, $0\in\partial g(\bar{x})-\nabla\gamma_\tau(\bar{x})=\partial\theta_\tau(\bar{x})$. 由于 γ 在 $\bar{x}$ 点是二阶 Hadamard 方向可微的, 根据文献 [47, 命题 2.10] (或者文献 [30, 定理 7.46]) 知, θ_τ 在 $\bar{x}$ 点是二阶上图可微的, 且 $(\theta_\tau)''_{\bar{x},0}(\cdot)=g''_{\bar{x},F(\bar{x})}(\cdot)-\gamma''_\tau(\bar{x},\cdot)$. 不难证明二阶增长条件 (4.2) 成立当且仅当

$$(\theta_\tau)''_{\bar{x},0}(\xi)\geqslant c\|\xi\|^2,\quad 对所有\ \xi\in\mathbb{R}^n.\tag{4.3}$$

由注 3.1 知, 上式与 (ii) 等价.

(ii) $\Leftrightarrow$ (iii). 由 (3.10) 知, $g''_{\bar{x},F(\bar{x})}(\xi)\geqslant\gamma''_\tau(\bar{x},\xi)$ 对所有的 $\xi\in\mathbb{R}^n$ 成立. 于是 (ii) 等价于 $\xi=0$ 是使得 $g''_{\bar{x},F(\bar{x})}(\xi)=\gamma''_\tau(\bar{x},\xi)$ 成立的唯一向量. 再由 (3.10), 这个等式成立当且仅当 (3.10) 极小化问题在

$h=\xi$ 达到最小值. 由引理 3.1 知函数 $\phi(h):=g''_{\bar{x},F(\bar{x})}(h)+2\langle F'(\bar{x},\xi),h-\xi\rangle+\frac{1}{\tau}\|\xi-h\|^2$ 是凸的. 于是, $h=\xi$ 是 $\phi(h)$ 极小点的充要条件是 $0\in\partial\phi(\xi)$, 即 $-F'(\bar{x},\xi)\in\partial(\frac{1}{2}g''_{\bar{x},F(\bar{x})})(\xi)=(\partial g)'_{\bar{x},F(\bar{x})}(\xi)$. 这证明了 (ii) $\Leftrightarrow$ (iii). □

下面假设映射 F 依赖某个参数 $u\in\mathbb{R}^m$, 即 $F\colon\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, 我们讨论参数变分包含

$$0\in F(x,u)+\partial g(x) \tag{4.4}$$

的解映射 $S(u):=\{x\in\mathbb{R}^n\colon 0\in F(x,u)+\partial g(x)\}$ 关于参数 u 的连续性. 假设 $-F(\bar{x},\bar{u})\in\partial g(\bar{x})$, $F(\cdot,\cdot)$ 在 $(\bar{x},\bar{u})$ 点是局部 Lipschitz 连续且方向可微的, g 在 $\bar{x}$ 关于 $-F(\bar{x},\bar{u})$ 是二阶上图可微的.

一个集值映射 $\Phi\colon\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ 称为在 $\bar{u}$ 点关于 $\bar{x}\in\Phi(\bar{u})$ 是局部上 Lipschitz 的, 如果存在一个正常数 c 以及 $\bar{u}$ 和 $\bar{x}$ 的邻域 U 和 V, 使得

$$\Phi(u)\cap V\subset\bar{x}+c\|u-\bar{u}\|B,\quad \text{所有的 } u\in U, \tag{4.5}$$

其中 B 代表 $\mathbb{R}^m$ 中的闭单位球. 上面的定义在文献 [54] 中也称 Φ 在 $\bar{u}$ 点关于 $\bar{x}$ 有孤立平稳性.

下面关于 Φ 的局部上 Lipschitz 性质的刻画在文献 [55,56] 中给出.

命题 4.1 集值映射 $\Phi\colon\mathbb{R}^m\rightrightarrows\mathbb{R}^n$ 在 $\bar{u}$ 点关于 $\bar{x}\in\Phi(\bar{u})$ 是局部上 Lipschitz 的当且仅当 $D\Phi(\bar{u},\bar{x})(0)=\{0\}$.

由文献 [56, 推论 3.3] 或者 [57, 定理 4.1] 知,

$$DS(\bar{u},\bar{x})(p)\subset\{\xi\colon 0\in -F'(\bar{x},\bar{u})(\xi,p)+D(\partial g)(\bar{x},F(\bar{x},\bar{u}))(\xi)\}. \tag{4.6}$$

这和定理 3.3 蕴涵

$$DS(\bar{u},\bar{x})(p)\subset\{\xi\colon 0\in -F'(\bar{x},\bar{u})(\xi,p)+(\partial g)'_{\bar{x},\bar{v}}(\xi)\}. \tag{4.7}$$

因此, 我们有下面结果.

定理 4.2 如果定理 4.1 中的等价条件 (i)–(iii) 之一成立, 则解映射 S 在 $\bar{u}$ 点关于 $\bar{x}$ 是局部上 Lipschitz 的.

证明 根据 (4.7), 如果定理 4.1(iii) 成立, 则 $DS(\bar{u},\bar{x})(0)=\{0\}$. 由命题 4.1 知, S 在 $\bar{u}$ 点关于 $\bar{x}$ 是局部上 Lipschitz 的. □

进一步, 如果考虑下面的解映射:

$$S(u,v):=\{x\in\mathbb{R}^n\colon v+F(x,u)\in\partial g(x)\}, \tag{4.8}$$

则根据文献 [57, 定理 4.1] 知, $S(u,v)$ 在 $(\bar{u},0,\bar{x})$ 是 proto- 可微的, 且

$$S'_{\bar{u},0,\bar{x}}(p,d)=\{\xi\colon d\in F'(\bar{x},\bar{u})(\xi,p)+(\partial g)'_{\bar{x},\bar{v}}(\xi)\}. \tag{4.9}$$

5 特殊的变分不等式 VI(K,F) 情形

本节将前面获得的结果应用到变分不等式情形. 令 $K\subset H$ 为给定的非空闭集, $x\in K$ 为给定的点. 首先回顾一些基本概念. K 在 x 点的余切锥 (bouligand)

$$T_K(x):=\operatorname*{Lim\,sup}_{t\downarrow 0}\frac{K-x}{t},$$

内切锥

$$T_K^i(x) := \operatorname*{Lim\,inf}_{t\downarrow 0} \frac{K-x}{t},$$

以及 K 在 x 点关于方向 h 的外、内二阶切集

$$T_K^2(x,h) := \operatorname*{Lim\,sup}_{t\downarrow 0} \frac{K-x-th}{\frac{1}{2}t^2}$$

和

$$T_K^{i,2}(x,h) := \operatorname*{Lim\,inf}_{t\downarrow 0} \frac{K-x-th}{\frac{1}{2}t^2}.$$

令 $\sigma(y,K) := \sup_{z\in K}\langle y,z\rangle$ 表示 K 的支撑函数, 它是指示函数 $\delta_K(\cdot)$ 的共轭函数.

当 $g(\cdot) := \delta_K(\cdot)$ 是非空闭集 K 的指示函数时, 变分包含问题 (1.11) 退化为下面的变分不等式问题:

$$-F(x) \in N_K(x). \tag{5.1}$$

特别当 K 是凸集时, (5.1) 恰好是经典变分不等式 (1.1). 变分不等式 (5.1) 的局部正则间隙函数变为

$$\gamma_\tau(x) := \inf_{y\in K\cap V}\left\{\langle F(x), x-y\rangle + \frac{1}{2\tau}\|x-y\|^2\right\}. \tag{5.2}$$

注意到由文献 [49] 可知, K 在 $\bar{x}\in K$ 点是邻近正则的当且仅当 $\delta_K(x)$ 在 $\bar{x}$ 点是 p.l.n. 的. 作为定理 3.1 和 3.2 的应用, 我们获得下面结果.

定理 5.1 令 $\bar{x}$ 为 (5.1) 的一个解. 假设 F 在 $\bar{x}$ 点是局部 Lipschitz 连续的, 假设 K 在 $\bar{x}$ 关于 $-F(\bar{x})$ 是邻近正则的, 且有正则常数 r, 则

(i) 对任意的 $\tau\in(0,\frac{1}{r})$, γ_τ 在 $\bar{x}$ 点是 Fréchet 可微的, 且 $\nabla\gamma_\tau(\bar{x}) = F(\bar{x})$. 对所有的 x 充分接近 $\bar{x}$, (5.2) 中的极小化问题有唯一解 $P_\tau(x) = P_K(x+\tau F(x))$, 并且 $P_\tau(x)$ 在 $\bar{x}$ 点是局部 Lipschitz 连续的, 且 $\gamma_\tau(\bar{x}) = g(\bar{x})$;

(ii) 如果 F 在 $\bar{x}$ 点的某个邻域上是 C^1 (或者 $C^{1,+}$) 的, 则对任意的 $\tau\in(0,\frac{1}{r})$, 函数 γ_τ 在 $\bar{x}$ 某个邻域内是 C^1 (或者 $C^{1,+}$) 的, 且

$$\nabla\gamma_\tau(x) = \frac{1}{\tau}(I+\tau(DF(x))^*)(I-P_K)(x+\tau F(x)) - \tau(DF(x))^*F(x)$$

对所有的 x 充分接近 $\bar{x}$ 成立.

定义 5.1 称集合 $K\subset H$ 在 $\bar{x}\in K$ 沿方向 $h\in T_K(\bar{x})$ 是二阶方向可微的, 如果 $T_K^i(\bar{x}) = T_K(\bar{x})$ 且 $T_K^{i,2}(\bar{x},h) = T_K^2(\bar{x},h)$.

定义 5.2 称 K 在 $\bar{x}\in K$ 点是二阶正则的, 如果 $h\in T_K(\bar{x})$, K 在 $\bar{x}$ 点沿方向 $h\in T_K(\bar{x})$ 是二阶方向可微的, 并且对任意的序列 $x_k := \bar{x}+t_kh+\frac{1}{2}t_k^2w_k\in K$, 其中 $t_k\downarrow 0$, $t_kw_k\to 0$, 有下面条件成立:

$$\lim_{k\to\infty}\operatorname{dist}(w_k, T_K^2(\bar{x},\xi)) = 0.$$

给定向量 $v\in H$, $[v]$ 表示 v 生成的子空间, $[v]^\perp$ 表示 $[v]$ 的直交补子空间. 由文献 [39, 命题 2.58] 知, 如果 $T_K^i(\bar{x}) = T_K(\bar{x})$, 则 $(\delta_K)'_{\bar{x}}(\cdot) = \delta_{T_K(\bar{x})}(\cdot)$. 下面的命题在文献 [58] 中给出.

命题 5.1 令 $\bar{x} \in K$, $\bar{u} \in N_K(\bar{x})$ 为给定的点. 假设 K 在 $\bar{x}$ 点是二阶正则的, 则 δ_K 在 $\bar{x}$ 点关于 $\bar{u}$ 是二阶上图可微的, 并且

$$(\delta_K)''_{\bar{x},\bar{u}}(h) = \begin{cases} -\sigma(\bar{u}, T_K^2(\bar{x}, h)), & \text{如果 } h \in T_K(\bar{x}) \cap [\bar{u}]^\perp, \\ +\infty, & \text{否则}. \end{cases} \tag{5.3}$$

在余下部分, 我们假设空间是有限维的, K 为 $\mathbb{R}^n$ 中的非空闭集. 作为定理 3.4、4.1 和命题 5.1 的应用, 我们获得下面结果.

定理 5.2 给定 (5.1) 的解 $\bar{x}$, 假设 K 在 $\bar{x}$ 点是邻近正则和二阶正则的, $F\colon \mathbb{R}^n \to \mathbb{R}^n$ 在 $\bar{x}$ 点 Hadamard 方向可微的, 则对任意充分小的 $\tau > 0$, 下列陈述成立:

(i) γ_τ 在 $\bar{x}$ 点是二阶 Hadamard 方向可微的, 且二阶导数满足

$$\gamma''_\tau(\bar{x}, \xi) = \inf_{h \in T_K(\bar{x}) \cap [F(\bar{x})]^\perp} \left\{ -\sigma(F(\bar{x}), T_K^2(\bar{x}, h)) + 2\langle F'(\bar{x}, \xi), \xi - h\rangle + \frac{1}{\tau}\|\xi - h\|^2 \right\};$$

(ii) 邻近解映射 P_τ 在 $\bar{x}$ 点是 proto- 可微的, 且

$$(P_\tau)'_{\bar{x}}(\xi) = [I + \tau(N_K^p)'_{\bar{x},F(\bar{x})}]^{-1}(\xi + \tau F'(\bar{x}, \xi)).$$

证明 命题 5.1(i) 恰为定理 4.1 在 $g = \delta_K$ 情形的相应的部分. 因为 K 在 $\bar{x}$ 点是邻近正则, 等价于它在 $\bar{x}$ 点是 p.l.n. 的. 由文献 [49, 推论 2.2] 知, 存在 $\bar{x}$ 的一个开邻域 V 使得对所有的 $x \in V$, 极限法锥 $N_K(x)$ 是闭凸的, 进而有 $N_K(x) = N_K^p(x)$. 根据定理 3.3, 邻近法锥 $N_K^p(\cdot)$ 在 $\bar{x}$ 点关于 $-F(\bar{x})$ 是 proto- 可微的, 并且

$$\frac{1}{2}(N_K^p)'_{\bar{x},F(\bar{x})} = \partial[(\delta_K)''_{\bar{x},F(\bar{x})}](\xi), \quad \text{对所有的 } \xi \in \mathbb{R}^n. \tag{5.4}$$

于是 (ii) 恰好为定理 3.4 在 $g = \delta_K$ 情形的相应的部分. □

当 F 在 $\bar{x}$ 点为局部 Lipschitz 连续且 Hadamard 方向可微时, 文献 [12] 使用不同的方法证明了定理 5.2(i). 由定理 5.1、5.2(ii) 和 4.1, 容易得到下面的定理 5.3. 这个定理改进了文献 [12, 定理 3.1], 那里要求 $K = G^{-1}(C)$ 是一类特殊的二阶正则且邻近正则集合, 其中 G 和 C 满足下面的 (A1)–(A3); 其证明是基于文献 [39, 命题 3.105] 中的二阶增长条件与抛物二阶导数的关系以及抛物二阶导数的计算公式.

定理 5.3 令 $\bar{x}$ 为 (5.1) 给定的解, 假设 K 在 $\bar{x}$ 点是邻近正则和二阶正则的, $F\colon \mathbb{R}^n \to \mathbb{R}^n$ 在 $\bar{x}$ 点是 Hadamard 方向可微的, 则对任意充分小的 $\tau > 0$, 下列陈述等价且蕴涵 $\bar{x}$ 是 (5.1) 的局部唯一解.

(i) 二阶增长条件成立, 即存在常数 $c > 0$ 和 $\bar{x}$ 的某个邻域 V, 使得

$$-\gamma_\tau(x) \geqslant -\gamma_\tau(\bar{x}) + c\|x - \bar{x}\|^2, \quad \text{对所有的 } x \in V \cap K; \tag{5.5}$$

(ii) $\gamma''_\tau(\bar{x}, \xi) + \sigma(F(\bar{x}), T_K^2(\bar{x}, \xi)) < 0$, 对所有的 $\xi \in T_K(\bar{x}) \cap [F(\bar{x})]^\perp) \setminus \{0\}$;

(iii) $\xi = 0$ 是

$$0 \in F'(\bar{x}, \xi) + (N_K^p)'_{\bar{x},F(\bar{x})}(\xi)$$

的唯一解.

下面将考虑 $K=\{x\in\mathbb{R}^n\colon G(x)\in C\}$ 的情形, 并且假设

(A1) $G\colon\mathbb{R}^n\to\mathbb{R}^m$ 是二阶连续可微的;

(A2) $C\subset\mathbb{R}^m$ 是闭凸的且在 $G(\bar{x})$ 点二阶正则;

(A3) Robinson 约束品性在 $\bar{x}$ 点成立, 即

$$DG(\bar{x})\mathbb{R}^n+T_C(G(\bar{x}))=\mathbb{R}^m.$$

根据文献 [39, 命题 3.88 和 3.33] 知, K 在 $\bar{x}$ 点是二阶正则的, 并且

$$\begin{aligned}&T_K(\bar{x})=DG(\bar{x})^{-1}(T_C(G(\bar{x}))),\\&T_K^2(x_0,h)=DG(x_0)^{-1}[T_C^2(G(x_0),DG(x_0)h)-D^2G(x_0)(h,h)].\end{aligned}$$

因此, 应用命题 5.1, 我们获得下列结果.

命题 5.2 令 $\bar{u}\in N_K(\bar{x})$ 为给定的法向量, 则 δ_K 在 $\bar{x}$ 点关于 $\bar{u}$ 是二阶上图可微的, 且

$$(\delta_K)''_{\bar{x},\bar{u}}(h)=\begin{cases}\sup\limits_{\lambda\in\Lambda(\bar{x})}\{\langle\lambda,D^2G(\bar{x})(h,h)\rangle-\sigma(\lambda,T_C^2(G(\bar{x}),DG(\bar{x})h))\}, & \text{如果 } h\in C(\bar{x},\bar{u}),\\ +\infty, & \text{否则},\end{cases}\tag{5.6}$$

其中 $C(\bar{x},\bar{u})=(DG(\bar{x}))^{-1}(T_C(G(\bar{x})))\cap[\bar{u}]^{\perp}$, $\Lambda(\bar{x})=\{\lambda\in\mathbb{R}^m\colon\bar{u}=DG(\bar{x})^*\lambda,\lambda\in N_C(G(\bar{x}))\}$.

注意到, 根据文献 [59, 第 134–135 页] 或者 [31, 推论 2.12], 如果条件 (A2) 和 (A3) 成立, 则 K 在 $\bar{x}$ 点是邻近正则的. 因为 Robinson 约束品性 (A3) 关于 x 在小扰动下是稳定的, 所以 K 在 $\bar{x}$ 的某邻域内的任一点 x 是邻近正则的. 于是, 在这种情形, $N_K(x)$ 是闭凸的, 且

$$N_K(x)=N_K^p(x)=(DG(x))^*N_C(G(x))$$

对所有的 x 充分接近 $\bar{x}$. 因此, 以上说明如果假设 (A1)–(A3) 成立, 则 $K=\{x\in\mathbb{R}^n\colon G(x)\in C\}$ 满足定理 5.2 和 5.3 的条件.

令 $\bar{y}:=F(\bar{x})\in N_K(G(\bar{x}))=(DG(\bar{x}))^*N_C(G(\bar{x}))$, 则存在某个 $\lambda\in N_C(G(\bar{x})$ 使得 $\bar{y}=DG(\bar{x})^*\lambda$. 令

$$\Lambda(\bar{x})=\{\lambda\in N_C(G(\bar{x}))\colon\bar{y}=DG(\bar{x})^*\lambda\}$$

表示 Lagrange 乘子集合. 我们说严格约束品性在 $\bar{\lambda}\in\Lambda(\bar{x})$ 点成立, 如果

$$DG(\bar{x})\mathbb{R}^n+T_{C_0}(G(\bar{x}))=\mathbb{R}^m,\tag{5.7}$$

其中

$$C_0:=\{y\in Q\colon\langle\bar{\lambda},y-G(\bar{x})\rangle=0\}.$$

显然, 严格约束品性 (5.7) 蕴涵 Robinson 条件 (A3) 成立, 而且也蕴涵 $\Lambda(\bar{x})$ 是单点集. 根据文献 [12, 引理 5.1], 在条件(A1)、(A2) 和 (5.7) 成立下, 有

$$(N_K^p)'_{\bar{x},F(\bar{x})}(\xi)\subset G(\bar{x})^*[(N_C)'_{G(\bar{x}),\bar{\lambda}}(DG(\bar{x})\xi)]+(D^2_{xx}G(\bar{x})\xi)^*\bar{\lambda}.\tag{5.8}$$

进一步, 当 $DG(\bar{x})$ 是满射时, 上面的包含变成等式.

因此, 如果假设严格约束品性成立代替 Robinson 条件 (A3), 并且 $\xi=0$ 是下面广义方程:

$$0\in F'(\bar{x},\xi)+G(\bar{x})^*[(N_C)'_{G(\bar{x}),\bar{\lambda}}(DG(\bar{x})\xi)]+(D^2_{xx}G(\bar{x})\xi)^*\bar{\lambda}\tag{5.9}$$

的唯一解, 则 $\bar{x}$ 是 (5.1) 的局部唯一解.

定义 5.3 [39] 称 $\mathbb{R}^n$ 中的一个闭凸子集 K 在 $\bar{x} \in K$ 点是多面的 (polyhedric), 若对任意的 $u \in N_K(\bar{x})$, 有

$$T_K(\bar{x}) \cap [\mathrm{Ker}\ u] = \mathrm{cl}\{R_K(\bar{x}) \cap [\mathrm{Ker}\ u]\}.$$

如果 K 在每一个 $x \in K$ 都是多面的, 则称它是多面的.

根据文献 [60, 例 2.10] 或者 [61, 定理 2.1], 函数 $\delta_K(\cdot)$ 是二阶上图可微的, 并且

$$(\delta_K)''_{\bar{x},\bar{u}} = \delta_{T_K(\bar{x})\cap[u]^\perp}. \tag{5.10}$$

在这种情形, 定理 5.2 中 γ_τ 的二阶 Hadamard 方向导数和邻近法锥 N_K^p 的 proto- 导数有下面简单形式:

$$\gamma''_\tau(\bar{x},\xi) = \inf_{h\in T_K(\bar{x})\cap[u]^\perp} \left\{2\langle F'(\bar{x},\xi), \xi - h\rangle + \frac{1}{\tau}\|\xi - h\|^2\right\},$$
$$(N_K^p)'_{\bar{x},F(\bar{x})}(\xi) = N_{T_K(\bar{x})\cap[u]^\perp}(\xi).$$

参考文献

1 Auslender A. Optimisation: Méthodes Numériques. Paris: Masson, 1976
2 Auchmuty G. Variational principles for variational inequalities. Numer Funct Anal Optim, 1989, 10: 863–874
3 Fukushima M. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math Program, 1992, 53: 99–110
4 Peng J M. Equivalence of variational inequality problems to unconstrained minimization. Math Program, 1997, 78: 347–355
5 Yamashita N, Taji K, Fukushima M. Unconstrained optimization reformulations of variational inequality problems. J Optim Theory Appl, 1997, 92: 439–456
6 Peng J M, Fukushima M. A hybrid Newton method for solving the variational inequality problem via the D-gap function. Math Program, 1999, 86: 367–386
7 Solodov M V, Tseng P. Some methods based on the D-gap function for solving monotone variational inequalities. Comput Optim Appl, 2000, 17: 255–277
8 Facchinei F, Pang J S. Finite-Dimensional Variational Inequalities and Complementarity Problems. New York: Springer, 2003
9 Harms N, Kanzow C, Stein O. Smoothness properties of a regularized gap function for quasi-variational inequalities. Optim Methods Softw, 2014, 29: 720–750
10 Kubota K, Fukushima M. Gap function approach to the generalized Nash equilibrium problem. J Optim Theory Appl, 2010, 144: 511–531
11 Ng K F, Tan L L. Error bounds of regularized gap functions for nonsmooth variational inequality problems. Math Program, 2007, 110: 405–429
12 Shapiro A. Sensitivity analysis of parameterized variational inequalities. Math Oper Res, 2005, 30: 109–126
13 Hearn D W. The gap function of a convex program. Oper Res Lett, 1982, 1: 67–71
14 Bertsekas D. Nonlinear Programming, 2nd ed. Belmont: Athena Scientific, 1999
15 Migdalas A. A regularization of the Frank-Wolfe method and unification of certain nonlinear programming methods. Math Program, 1994, 65: 331–345
16 Figueiredo M A T, Nowak R D, Wright S J. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. IEEE J Sel Top Signal Process, 2007, 1: 586–597
17 Nesterov Yu. Gradient methods for minimizing composite functions. Math Program, 2013, 140: 125–161
18 Lions P L, Mercier B. Splitting algorithms for the sum of two nonlinear operators. SIAM J Numer Anal, 1979, 16: 964–979
19 Chambolle A, Lions P L. Image recovery via total variation minimization and related problems. Numer Math, 1997, 76: 167–188
20 Chen G H G, Rockafellar R T. Convergence rates in forward-backward splitting. SIAM J Optim, 1997, 7: 421–444

21 Vese L A, Osher S J. Modeling textures with total variation minimization and oscillating patterns in image processing. J Sci Comput, 2003, 19: 553–572

22 Combettes P L, Wajs V R. Signal recovery by proximal forward-backward splitting. Multiscale Model Simul, 2005, 4: 1168–1200

23 Combettes P L, Pesquet J-C. Proximal splitting methods in signal processing. In: Fixed-Point Algorithms for Inverse Problems in Science and Engineering. New York-Dordrecht-Heidelberg-London: Springer, 2011, 185–212

24 Parikh N, Boyd S. Proximal algorithms. Found Trends Optim, 2013, 1: 123–231

25 Aujol J F, Aubert G, Blanc-Féraud L, et al. Image decomposition into a bounded variation component and an oscillating component. J Math Imaging Vision, 2005, 22: 71–88

26 Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imag Sci, 2009, 2: 183–202

27 Stella L, Themelis A, Patrinos P. Forward-backward quasi-Newton methods for nonsmooth optimization problems. Comput Optim Appl, 2017, 67: 443–487

28 Bauschke H H, Combettes P L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics. New York: Springer-Verlag, 2011

29 Attouch H. Variational Convergence for Functions and Operators. Boston: Pitman, 1984

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

31 Poliquin R A, Rockafellar R T. Prox-regular functions in variational analysis. Trans Amer Math Soc, 1996, 348: 1805–1838

32 Bačák M, Borwein J M, Eberhard A, et al. Infimal convolutions and Lipschitzian properties of subdifferentials for prox-regular functions in Hilbert spaces. J Convex Anal, 2010, 17: 737–763

33 Hare W L, Poliquin R A. Prox-regularity and stability of the proximal mapping. J Convex Anal, 2007, 14: 589–606

34 Hare W, Sagastizábal C. Computing proximal points of nonconvex functions. Math Program, 2009, 116: 221–258

35 Poliquin R A. Integration of subdifferentials of nonconvex functions. Nonlinear Anal, 1991, 17: 385–398

36 Shapiro A. On concepts of directional differentiability. J Optim Theory Appl, 1990, 66: 477–487

37 Bernard F, Thibault L. Prox-regular functions in Hilbert spaces. J Math Anal Appl, 2005, 303: 1–14

38 Bernard F, Thibault L. Uniform prox-regularity of functions and epigraphs in Hilbert spaces. Nonlinear Anal, 2005, 60: 187–207

39 Bonnans J F, Shapiro A. Perturbation Analysis of Optimization Problems. New York: Springer, 2000

40 Clarke F H, Ledyaev Y S, Stern R J, et al. Nonsmooth Analysis and Control Theory. Graduate Texts in Mathematics, vol. 178. New York: Springer, 1998

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

42 Ioffe A D. Proximal analysis and approximate subdifferentials. J Lond Math Soc (2), 1990, s2-41: 175–192

43 Thibault L, Zagrodny D. Integration of subdifferentials of lower semicontinuous functions on Banach spaces. J Math Anal Appl, 1995, 189: 33–58

44 Levy A B, Poliquin R A, Thibault L. Partial extensions of Attouch's theorem with applications to proto-derivatives of subgradient mappings. Trans Amer Math Soc, 1995, 347: 1269–1294

45 Combari C, Elhilali Alaoui A, Levy A, et al. Convex composite functions in Banach spaces and the primal lower-nice property. Proc Amer Math Soc, 1998, 126: 3701–3708

46 Rockafellar R T. Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives. Math Oper Res, 1989, 14: 462–484

47 Rockafellar R T. First- and second-order epi-differentiability in nonlinear programming. Trans Amer Math Soc, 1988, 307: 75–108

48 Poliquin R A, Rockafellar R T. A calculus of epi-derivatives applicable to optimization. Canad J Math, 1993, 45: 879–896

49 Poliquin R A, Rockafellar R T, Thibault L. Local differentiability of distance functions. Trans Amer Math Soc, 2000, 352: 5231–5249

50 Poliquin R A, Rockafellar R T. Generalized Hessian properties of regularized nonsmooth functions. SIAM J Optim, 1996, 6: 1121–1137

51 Mazade M, Thibault L. Primal lower nice functions and their Moreau envelopes. In: Computational and Analytical Mathematics. Springer Proceedings in Mathematics & Statistics, vol. 50. New York: Springer, 2013, 521–553

52 Li W, Peng J. Exact penalty functions for constrained minimization problems via regularized gap function for variational inequalities. J Global Optim, 2007, 37: 85–94

53 Rockafellar R T. Generalized second derivatives of convex functions and saddle functions. Trans Amer Math Soc, 1990, 322: 51–77
54 Dontchev A L, Rockafellar R T. Implicit Functions and Solution Mappings. Berlin: Springer, 2009
55 King A J, Rockafellar R T. Sensitivity analysis for nonsmooth generalized equations. Math Program, 1992, 55: 193–212
56 Levy A B. Implicit multifunction theorems for the sensitivity analysis of variational conditions. Math Program, 1996, 74: 333–350
57 Levy A B, Rockafellar R T. Sensitivity analysis of solutions to generalized equations. Trans Amer Math Soc, 1994, 345: 661–671
58 Kan C, Song W. Augmented Lagrangian duality for composite optimization problems. J Optim Theory Appl, 2015, 165: 763–784
59 Shapiro A. Existence and differentiability of metric projections in Hilbert spaces. SIAM J Optim, 1994, 4: 130–141
60 Do C N. Generalized second-order derivatives of convex functions in reflexive Banach spaces. Trans Amer Math Soc, 1992, 334: 281–301
61 Levy A B. Sensitivity of solutions to variational inequalities on Banach spaces. SIAM J Control Optim, 1999, 38: 50–60

Differentiability of regularized gap functions for a kind of variational inclusions and local uniqueness of solutions

Wen Song

Abstract In this paper, we mainly explore continuity and differentiability properties of regularized gap functions for a kind of nonconvex variational inclusions and its associated solution mapping by making use of the properties of the Moreau envelope functions of the prox-regular function and its proximal mapping. By applying the second order variational analysis, we obtain the local uniqueness and continuity on parameters of solutions for the variational inclusions. We also survey the developing process of regularized gap functions for the variational inequality and optimization, and variational properties of prox-regular functions.

Keywords **regularized gap function, Moreau envelope, proximal mapping, prox-regularity, epi-differentiability**

MSC(2010) **90C31, 49J52, 46C05**

doi: 10.1360/SSM-2020-0158