arXiv:2303.09237v2 [math.CA] 5 Jan 2024

广义拉格朗日定理

KAROLINA ZAJĄC

摘要。本文致力于探讨经典拉格朗日中值定理的可能推广。我们考虑一个多变量实值函数，并且只假设它连续。核心思想是用所谓的双序列切锥取代导数的概念。我们首先证明罗尔型和拉格朗日型结果，随后在 Lipschitz 函数的情形下，将该切锥与克拉克次微分进行比较。我们还探讨了使用法锥的方法。

引言

对于单值可微实函数，经典形式的拉格朗日中值定理是数学分析中最重要的结论之一，具有大量重要应用。

一个自然且有趣的问题是，中值定理是否也能适用于仅连续而不一定可微的函数。要解决这个问题，就必须找到函数导数或其图像切空间的替代概念。

关于上述主题已有许多研究。G. Lebourg 于 [7] 发表了一项非常有趣的结果，其中运用了广义梯度。另一种基于近似雅可比矩阵的方法见 [6]。T. Ważewski 和 W. Mlak 分别于 [11] 和 [8] 得到了一些运用迪尼导数的结果。 [4] 中也介绍了一些可能的推广。关于向量值映射，另请参见 [5] 及其中的参考文献。

本文将介绍另一种针对多值实连续函数的方法，该方法运用副切锥，本文称之为双序列切锥（*BTC*）（定义 1.3）。BTC 最早分别由 F. Severi 和 G. Bouligand 在文章 [10] 和 [2] 中提出，随后 Yuntong W. 在 [13] 中对其进行了详细研究。其他一些资料也提到了它，例如 [9] 或 [12], ，但几乎没有提供进一步的信息。第一节将介绍若干基本例子以及这一对象的主要实用性质，随后给出中值定理的一个推广版本，即罗尔型结果（定理 1.12）和拉格朗日型结果（定理 1.13）。

第二个思路是探究法锥的性质。在这种情况下，几何解释截然不同。我们不再寻找与给定割线平行的切向量，而是可以在中值点找到法锥中的一个非零向量，使其垂直于割线，见定理 2.7。

最后，如果将所考虑的函数限制为局部利普希茨函数，则可以用克拉克次微分作为导数的自然对应物。在

日期：2024年1月8日。2020 数学主题分类。26A24。关键词和短语。中值定理；切锥；克拉克次微分。

第三节中，我们回顾 G. Lebourg 的定理，随后在定理 3.11 中展示克拉克次微分与双序列切锥之间的对应关系。这带来了对 Lebourg的结果一种非常优雅的几何解释。

我谨向导师 Maciej Denkowski 致谢，感谢他提出这个问题。

1. 双序列切锥

让我们回顾一下基本的中值定理：

定理 1.1 (拉格朗日)**.** 设 $a, b \in \mathbb{R}, a < b$，且令 $f\colon [a,b] \to \mathbb{R}$ 为连续函数，在开区间 (a, b)上可微。那么

$$\exists c \in (a,b) : f'(c) = \frac{f(b) - f(a)}{b - a}.$$

由于我们希望在函数仅连续的情形下得到上述定理的自然对应形式，因此可以用对图像的切向量作适当理解来替代函数的导数。

最直观的想法是考虑 Peano切锥：

定义 1.2。 如果存在序列$\{s_n\}_{n=1}^{+\infty} \subset S$ 和 $\{t_n\}_{n=1}^{+\infty} \subset \mathbb{R}_+$ 使得：则称向量 $v \in \mathbb{R}^n$ 在 S 处 a 切向于

(1) $s_n \to a$ $(n \to +\infty)$, (2) t_n $(s_n - a) \to v$ $(n \to +\infty)$。

集合 $C_a(S) = \{v \in \mathbb{R}^n$ 在 S 处 $a\}$ 切向于称为（*Peano*）切锥 S 在 a处。

遗憾的是，上述概念事实证明并不充分。举例来说，

$$f : [-1, 1] \ni x \mapsto |x| \in \mathbb{R}.$$

我们看到，水平线 $y = 0$ 不包含在图像上任何一点处的切锥中，这给出了一个反例。因此，有必要对经典定义稍作修改。于是，我们定义给定集合 X 在点 $x \in X$ 处的双序列切锥（BTC）（参见例如 [12]）。

定义 1.3。 若存在点的双序列$\{(a_n, b_n)\}_{n=1}^{+\infty} \subset X^2$ 以及正数序列 $\{t_n\}_{n=1}^{+\infty} \subset \mathbb{R}_+$，使得：则称向量 $v \in \mathbb{R}^n$ 属于的双序列切锥在 x 处（记作 $v \in B_x(X)$）。

(1) $a_n, b_n \to x$ $(n \to +\infty)$, (2) t_n $(a_n - b_n) \to v$ $(n \to +\infty)$.

等价地，若 $\|v\| = 1$，则 $v \in B_x(X)$，当且仅当存在点的双序列$\{(a_n, b_n)\}_{n=1}^{+\infty} \subset X^2$，使得

$$\frac{a_n - b_n}{\|a_n - b_n\|} \to v, (n \to +\infty).$$

上述定义有清晰的几何解释： X 在 x 处的切线由割线方向的极限描述，这些割线的两端都趋近于 x。

现在我们来考察双序列切锥（BTC）的一些性质。

注 1.4. 总是存在 $0 \in B_x(X)$。我们还注意到，除非 x 是 X 中的孤立点，否则 $B_x(X) \neq \{0\}$；如果 $v \in B_x(X)$，那么双序列切锥包含一条完整直线$\{tv : t \in \mathbb{R}\}$，因此它是一个全锥。

以下我们假设，余维为 1 的（非垂直）线性子空间可以表示为线性函数 L 的图像：$\mathbb{R}^{n-1} \to \mathbb{R}$。

定义 1.5。我们称上述线性子空间与集合 $X \in \mathbb{R}^n$在 $a \in X$处相切，如果 $(x, L(x)) \in B_a(X)\ \forall x \in \mathbb{R}^{n-1}$.

从现在起，我们考虑集合 X 是连续函数$f\colon A \to \mathbb{R}, A \subset \mathbb{R}^n$的图像，记作 Γ_f。对于线性映射

$$L : \mathbb{R}^n \ni (x_1, \dots, x_n) \mapsto \alpha_1 x_1 + \dots + \alpha_n x_n \in \mathbb{R}$$

使得 $\Gamma_L \subset B_{(a,f(a))}(\Gamma_f)$，我们将使用简化记号 $B_{a,\ f(a)}(\Gamma_f) =: B_a(f)$ 和 $(\alpha_1, \dots, \alpha_n) \in B_a(f)$ 或 $L \in B_a(f)$，其中 $B_a(f)$ 表示 $\Gamma_f \subset \mathbb{R}^n$ 在给定点 a处的切子空间方向集合，称为***f*** 在 ***a*** 处的 ***BTC*** 次微分（另见 [13] 定义 4.1）。

稍后将对这个名称作出恰当的解释，因为我们会比较 Lipschitz 函数的 BTC 次微分与克拉克次微分（见定理 3.11）。

注 1.6. 我们还需要处理一个竖直切子空间，它是 $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ 中由下式给出的超曲面

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 0,\ \alpha_1^2 + \dots + \alpha_n^2 \neq 0.$$

如果 BTC $B_a(f)$ 包含这样的超曲面，我们就记作 $\infty \in B_a(f)$。

基本例子

(1) $f\colon A \to \mathbb{R}$ 是一个常值函数，$a \in A$。于是 $B_a(f) = \{0\}$。稍后我们将看到，如果函数 f 在 a 的某个邻域内属于 C^1 类，且 $f'(a) = 0$，那么在 $a \in A$ 中，切方向的集合会缩减为 $\{0\}$。(2) $f\colon \mathbb{R}^n \to \mathbb{R}$ 是一个线性函数，$A \subset \mathbb{R}^n, g = f\mid_A, a \in A$。于是我们有$B_a(g) = \{\Gamma_f\}$。(3) $A \subset B \subset \mathbb{R}^n$, $f\colon A \to \mathbb{R}, g\colon B \to \mathbb{R}$, $f = g\mid_A$, $a \in A$。于是我们有$B_a(f) \subset B_a(g)$。(4) 如果 $f\colon A \to \mathbb{R}$ 在 $a \in intA$处可微，那么总有 $\nabla f(a) \in B_a(f)$。然而，BTC 次微分有可能不会缩减为$\nabla f(a)$，见例 1.8。(5) $f\colon \mathbb{R} \ni x \mapsto |x| \in \mathbb{R}$。取一个双序列

$$\left\{\left(\left(\frac{1}{n}, \frac{1}{n}\right), \left(\frac{-t}{n}, \frac{t}{n}\right)\right)\right\}_{n=1}^{+\infty}$$

对于某个 $t \in \mathbb{R}_{\geq 0}$，我们得到$\frac{1-t}{1+t} \in \widehat{B}_0(f)$，因此 $\widehat{B}_0(f) = [-1, 1]$。(6) $f\colon \mathbb{R} \ni x \mapsto x^{\frac{2}{3}} \in \mathbb{R}$。取映射

$$g : \mathbb{R} \ni x \mapsto x^{\frac{2}{3}} - \alpha x \in \mathbb{R}, \alpha \in \mathbb{R}$$

以及一个双序列

$$\{((a_n, f(a_n)), (b_n, f(b_n)))\}_{n=1}^{+\infty} \subset \mathbb{R}^2,$$

使得 $a_n, b_n \to 0\,(n \to +\infty)$ 且 $g(a_n) = g(b_n)$ $a_n \neq b_n\ \forall n \in \mathbb{N}$。于是

$$\frac{f(a_n) - f(b_n)}{a_n - b_n} = \frac{a_n^{\frac{2}{3}} - b_n^{\frac{2}{3}}}{a_n - b_n} = \frac{(a_n^{\frac{2}{3}} - \alpha a_n) - (b_n^{\frac{2}{3}} - \alpha b_n) + \alpha a_n - \alpha b_n}{a_n - b_n} = \alpha,$$

$\widehat{\quad}$ $\widehat{\quad}$所以 $\alpha \in B_0(f)$。

但由于 α 是任意选取的，因此 $\mathbb{R} \subset B_0(f)$。此外，通过取一个双序列

$$\left\{\left(\left(\frac{1}{2^{3n}}, \frac{1}{2^{2n}}\right), \left(\frac{1}{2^{3n+3}}, \frac{1}{2^{2n+2}}\right)\right)\right\}_{n=1}^{+\infty},$$

我们得到

$$\frac{\frac{1}{2^{2n}} - \frac{1}{2^{2n+2}}}{\frac{1}{2^{3n}} - \frac{1}{2^{3n+3}}} = \frac{2^{n+3} - 2^{n+1}}{2^3 - 1} = 2^n \cdot \frac{6}{7} \to \infty \ (n \to +\infty).$$

因此，$\widehat{B}_0(f) = \mathbb{R} \cup \{\infty\}$。双序列切锥等于整个空间：$B_0(f) = \mathbb{R}^2$。

注记 1.7。对于单变量函数 $f\colon A \to \mathbb{R}$ 我们可以将$\widehat{B}_a(f)$ 对于 $a \in A$ 视为极限 $\lim_{n\to+\infty} {}^{f_{(a_n)-f(b_n)}\, a_n - b_n}$ 的集合，其中

$$\{(a_n, b_n)\}_{n=1}^{+\infty} \subset A^2, a_n, b_n \to x \ (n \to +\infty), a_n \neq b_n, \forall n \in \mathbb{N}.$$

注记 1.8。 f 在点 a 处的可微性并不足以保证$\frown B_a(f) = \{\nabla f(a)\}$。考虑一个著名的例子

$$f : \mathbb{R} \ni x \mapsto \begin{cases} x^2 \sin(\frac{1}{x}), & x \in \mathbb{R} \setminus \{0\} \\ 0, & x = 0 \end{cases} \in \mathbb{R}.$$

该函数在 0 处可微，且 $f'(0) = 0$，然而其导数在原点处不连续$\widehat{\ }$。因此我们得到 $B_0(f) = [-1, 1]$。

为了找到 BTC 次微分的任何应用，需要具备一些基本性质。下面的命题对于克拉克次微分而言是众所周知的（例如参见 [9], 定理 9.8 或 [13], 命题 4.3）。我们根据 BTC 次微分的定义给出一个直接证明。

命题 1.9. 若 A 是 $\mathbb{R}^n$ 中的开集，且 $f\colon A \to \mathbb{R}$ 可微，则对于任意点 $a \in A$，以下条件等价：

(1) f 在 a的某个邻域内属于 C^1 类，$\widehat{\ }$(2) B_a

$(f) = \{\nabla f(a)\}$。

证明。

(1) $\Rightarrow$(2) 假设存在 $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R})$、 $L \neq f'(a)$，使得我们$\widehat{\ }$有 $\nabla L \in B_a(f)$。取满足 $L.b \neq f'(a).b$的 $b \in \mathbb{R}^n$，并令 $\varepsilon > 0$，使得$|L.b - f'(a).b| > 3\varepsilon$。然后选取一个双序列 $\{(a_n, b_n)\}_{n=1}^{+\infty} \subset A^2$，使得$a_n$、 $b_n \to a$ $(n \to +\infty)$，并且

$$t_n(a_n - b_n, f(a_n) - f(b_n)) \to (b, L.b) \ (n \to +\infty)$$

对于某个序列 $\{t_n\}_{n=1}^{+\infty} \subset \mathbb{R}$，存在 N_1、 $N_2 \in \mathbb{N}$，使得

$$|t_n(f(a_n) - f(b_n)) - L.b| < \varepsilon, \quad \forall n \geq N_1,$$
$$|L.(t_n(a_n - b_n)) - L.b| < \varepsilon, \quad \forall n \geq N_2.$$

设 $N = max\{N_1, N_2\}$。我们注意到，对于 $n > N$

$$\begin{aligned} &|t_n(f(a_n) - f(b_n)) - L.(t_n(a_n - b_n))| \\ &\leq |t_n(f(a_n) - f(b_n)) - L.b| + |L.(t_n(a_n - b_n)) - L.b| < 2\varepsilon. \end{aligned}$$

函数 $f\mid_{[a_n,b_n]}$ 对于 $n > N$ 可以通过投影与一个单变量函数相对应，因此可以应用基本的中值定理。我们在线段 $[a_n, b_n]$ 上找到一点 c_n，使得

$$f'(c_n).(t_n(a_n - b_n)) = (f\mid_{[a_n,b_n]})'(c_n).(t_n(a_n - b_n)) = t_n(f(a_n) - f(b_n)).$$

于是我们得到

$$\begin{aligned}
&|f'(c_n).(t_n(a_n - b_n)) - f'(a).(t_n(a_n - b_n))| \\
&= |t_n(f(a_n) - f(b_n)) - f'(a).(t_n(a_n - b_n))| \\
&\geq |L.(t_n(a_n - b_n)) - f'(a).(t_n(a_n - b_n))| \\
&\quad - |t_n(f(a_n) - f(b_n)) - L.(t_n(a_n - b_n))| \\
&> |L.(t_n(a_n - b_n)) - f'(a).(t_n(a_n - b_n))| - 2\varepsilon.
\end{aligned}$$

当 $n \to +\infty$ 时，对两边取极限，得到

$$0 = |f'(a).b - f'(a).b| \geq |L.b - f'(a).b| - 2\varepsilon > 3\varepsilon - 2\varepsilon = \varepsilon,$$

这就导出了矛盾。

(2) $\Rightarrow$ (1) 固定 $\varepsilon > 0$。那么，根据假设，Γ_f中两端都位于 $(a, f(a))$ 的一个足够小邻域内的割线方向，都接近由导数 $f'(a)$ 所确定的方向。换言之：

$$\exists r_a > 0 : (\|x - a\| < r_a, \|y - a\| < r_a) \Rightarrow (|(f(x) - f(y)) - f'(a).(x - y)| \leq \varepsilon)$$

对于 $x, y \in A$。现在令 $b \in A, |b - a| < r_a$，并以类似方式定义 r_b。那么对于x, $y \in \mathbb{B}(a, r_a) \cap \mathbb{B}(b, r_b)$，我们有：

$$\begin{aligned}
&|f'(b).(x - y) - f'(a).(x - y)| \leq \\
&\qquad |f'(b).(x - y) - (f(x) - f(y))| + |(f(x) - f(y)) - f'(a).(x - y)| \leq 2\varepsilon,
\end{aligned}$$

由此可知 f' 连续，因此 f 在 a 附近属于 C^1 类。 □

另一项有用性质涉及两个函数之和的 BTC 次微分的行为（参见 [9], 习题 8.8）。

命题1.10。设 $A \subset \mathbb{R}^n$, $a \in A$。对于两个连续函数 f, g： $A \to \mathbb{R}$，假设 g 在 a 附近属于 C^1 类。那么 $\widehat{B}_a(f + g) = \widehat{B}_a(f) + \nabla g(a)$。

证明。包含关系“$\supset$”可直接得到（使用 1.9）。然后将上述规则应用于 $f + g$ 和 $-g$, 即可得到“$\subset$”。□

注 1.11。我们必须假设其中一个函数属于 C^1 类。举例如下：

$$\begin{aligned}
f &: [-1, 1] \ni x \mapsto |x| \in \mathbb{R}, \\
g &: [-1, 1] \ni x \mapsto -|x| \in \mathbb{R}.
\end{aligned}$$

我们看到 $1 \in B_0(f)$ 且 $1 \widehat{\in} B_0(g)$。然而， $2 \notin \widehat{B}_0(f + g) = \{0\}$。

现在我们给出双序列切锥的广义罗尔定理和拉格朗日定理。

定理 1.12 (广义罗尔)。设 K 是 $\mathbb{R}^n$ 的紧子集，满足 $intK \neq \emptyset$，并设 f： $K \to \mathbb{R}$ 是一个连续函数， $f \equiv C$ 在 ∂K上某个 $C \in \mathbb{R}$处成立。那么

$$\exists c \in intK : \ 0 \in \widehat{B}_c(f).$$

证明。若 f 是常函数，则定理显然成立。否则，根据魏尔斯特拉斯定理，f 在 K 上达到两个不同的极值。不失一般性，我们可以假设

$$\exists c \in intK : f(c) = \inf_{x \in K} f(x),$$

我们考虑两种情况。

(1) 情况 $n = 1$。

我们分别考虑这样一种情形：存在一个由两两不同的点组成的序列 $\{c_n\}_{n=1}^{+\infty} \subset K \setminus \{c\}$，收敛于 c，且$f(c_n) = f(c)\ \forall n \in \mathbb{N}$。此时，双序列 $\{(c_n, c_{n+1})\}_{n=1}^{+\infty}$与一个数列 $\{\frac{1}{c_n - c_{n+1}}\}_{n=1}^{+\infty}$足以得到 $(1, 0) \in B_c(f)$。

否则，由于 f 连续，根据达布定理，我们可以构造一个双序列 $\{(a_n, b_n)\}_{n=1}^{+\infty} \subset K^2$，使得

$$a_n \to c_+, b_n \to c_- (n \to +\infty),$$

$$f(a_n) = f(b_n) = f(c) + \frac{C - f(c)}{n}, \forall n \in \mathbb{N}.$$

取数列 $\{\frac{1}{a_n - b_n}\}_{n=1}^{+\infty}$，我们得到

$$\frac{1}{a_n - b_n}(a_n - b_n, f(a_n) - f(b_n)) = (1, 0).$$

（2）情形 $n > 1$。

固定 $x \in \mathbb{R}^n$。我们将证明 $(x, 0) \in B_c(f)$。我们选择 $a, b \in \partial K$，使得 $c \in [a, b] \subset K$，且点 $c + x$ 位于一条经过 a 和 b的直线上。于是，我们可以将 $f|_{[a,b]}$ 视为一个单变量函数，其余证明与前一种情形相同，只需将数列乘以 $\pm\|x\|$。

□

定理 1.13 (广义拉格朗日定理)**.** 设 K 是 $\mathbb{R}^n$中的一个紧子集，$intK \neq \emptyset$，并设 f: $K \to \mathbb{R}$ 是一个连续函数，$f(x) = L.x + C, \forall x \in \partial K$，其中某个 $L \in \mathcal{L}(\mathbb{R}^n; \mathbb{R})\ C \in \mathbb{R}$。那么

$$\exists\ c \in intK :\ L \in \widehat{B}_c(f).$$

证明。考虑映射 g: $K \ni x \mapsto f(x) - L.x \in \mathbb{R}$。它在 ∂K,$\widehat{\ }$ $\widehat{\ }$上为常数，因此根据 1.12 $\exists c \in intK :\ 0 \in B_c(f)$。显然，$B_c(L) = \{\nabla L\}$，对 g和 L 应用 1.10 即可完成证明。□

注释 1.14. 必须假设 K 是紧的。以函数 $f(x, y) = x^2 y$和 $K = \{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 1\}$为例。我们有$f(x, y) = y$在 ∂K上成立，但 $\widehat{B}_{(x,y)}(f) = (2xy, x^2) \neq (0, 1)$ 位于 K内部。

2. 法锥

设 L 是 $\mathbb{R}^n$的一个线性子空间。任意向量 $v \in \mathbb{R}^n$ 都可以唯一地表示为 $v = v_1 + v_2$的形式，其中 $v_1 \in L, v_2 \in L^\perp$ 且 $\|v - v_1\| = \inf\{\|v - w\| :\ w \in L\}$。映射 P_L： $\mathbb{R}^n \ni v \mapsto v_1 \in L$ 是到子空间 L上的正交投影。

现在我们考虑以下情形。令 $S \subset \mathbb{R}^n$。对于 $a \in S$，我们回顾 a处 S 的法锥。

定义 2.1。集合 S 在 a 处的法锥是集合

$$N_a(S) = \{v \in \mathbb{R}^n : \langle v, w\rangle \le 0, \forall w \in C_a(S)\}.$$

若我们所考虑的集合 S 是函数 $f\colon A \to \mathbb{R}$, $A \subset \mathbb{R}^{n-1}$的图像，那么对于$a \in A$，我们将使用以下记号：

$$C_a(f) := C_{(a,f(a))}(\Gamma_f),\ N_a(f) = N_{(a,f(a))}(\Gamma_f).$$

注 2.2. 如果函数 $f\colon A \to \mathbb{R}$, $A \subset \mathbb{R}^{n-1}$ 在 $a \in intA$处可微，那么 $C_a(f) = \Gamma_{f'(a)}$ 且 $N_a(f) = (\Gamma_{f'(a)})^{\perp}$。

为便于后续讨论，我们引入两个（非零）向量之间角的修正定义。

定义 2.3. 对于 $x, y \in \mathbb{R}^n$, $x, y \neq 0$，我们将 x 与 y 之间的角定义为

$$\angle(x, y) = \frac{\langle x, y\rangle}{\|x\|\|y\|} \in [-1, 1].$$

注 2.4. 根据上述定义，1.2 中的第二个条件与$\angle(s_n - a, v) \to 1 (n \to +\infty)$一致。集合 $N_a(S)$ 是 $\{v \in \mathbb{R}^n : \angle(v, w) \le 0, \forall w \in C_a(S)\}$。

我们将利用法锥证明针对连续函数陈述的中值定理。下面这个已知结果将起到关键作用。

命题 2.5. 设 $a \in S \subset \mathbb{R}^n$，且 $x \in \mathbb{R}^n$ 是一个点，使得从 S 到 $x \in \mathbb{R}^n$ 的欧几里得距离在 a处取得，即

$$\|x - a\| = \inf\{\|x - y\| : y \in S\}.$$

那么 $x - a \in N_a(S)$。

证明。 不妨设 $a = 0$。若 $x \notin N_0(S)$，则根据定义，存在一点 $v \in C_0(S)$ $v \neq 0$，使得 $\langle x, v\rangle > 0$。不失一般性，设 $v \in \partial\mathbb{B}(x, \|x\|)$。根据 1.2,

$$\exists\{(a_n)\}_{n=1}^{+\infty} \subset S^2, \{t_n\}_{n=1}^{+\infty} \subset \mathbb{R}_+ : a_n \to 0\ (n \to +\infty), t_n a_n \to v\ (n \to +\infty).$$

那么 $\angle(a_n, x) \to \angle(v, x) > 0 (n \to +\infty)$，因此

$$\exists N \in \mathbb{N} : \angle(a_n, x) > 0, \forall n \ge N,$$

从几何上看，这可以解释为：经过 0 和 a_n 的直线与 $\partial\mathbb{B}(x, \|x\|)$ 相交于某些点 $b_n \neq 0$，对所有 $n \ge N$均如此。我们看到$b_n \to v$ $(n \to +\infty)$，因此当 n 充分大时，点 a_n 必定位于线段 $[0, b_n]$内。否则，我们将有

$$0 = \lim_{n\to+\infty} \|a_n\| \ge \lim_{n\to+\infty} \|b_n\| = \|v\| > 0.$$

但这样一来，$\|x - a_n\| < \|x\|$ 对于某个 $n \in \mathbb{N}$成立。这个矛盾完成了证明。 □

注 2.6。 反向蕴含并非总是成立。以 $f(x) = ||x| - 1|$ 的图像为例。点 (0, 7) 属于 $N_0(f)$，但它到该图像的欧几里得距离并非在 (0, 1) 处取得。

现在来看主定理。

定理 2.7。设 K 是 $\mathbb{R}^n$ 的一个紧子集，$intK \neq \emptyset$，并设 $f\colon\ K \to \mathbb{R}$ 是一个连续函数，使得 $f(x) = L.x + C$，$\forall x \in \partial K$ 对某个 $L \in \mathcal{L}(\mathbb{R}^n;\ \mathbb{R})C \in \mathbb{R}$ 成立。那么

$$\exists c \in intK : (N_c(f) \cap (\Gamma_L)^{\perp}) \setminus \{0\} \neq \emptyset.$$

换言之，集合 K 内某点 c 处的法锥包含一个非零向量，且该向量垂直于 Γ_L。

证明。不失一般性，设 $C = 0$。我们考虑映射

$$g : K \ni x \mapsto \|P_{(\Gamma_L)^{\perp}}((x, f(x)))\| \in \mathbb{R}_{\geq 0}.$$

它是连续函数的复合，因此是连续的；又因其定义在一个紧集上，所以它能达到其极限。根据假设，我们知道 $g\,|_{\partial K} \equiv 0$。如果 $g \equiv 0$，那么 f 是仿射的，且 $N_x(f) = L^{\perp}$，$\forall x \in intK$。否则，我们选取一点 $c \in intK$，使得 $g(c) = \sup_{x \in K} g(x)$，并考虑向量

$$v = P_{(\Gamma_L)^{\perp}}((c, f(c))).$$

那么 v 垂直于 Γ_L，而且

$$\forall x \in K : \|P_{(\Gamma_L)^{\perp}}((x, f(x)))\| \leq \|v\|.$$

我们将证明 $v \in N_c(f)$。只需验证 $v + (c, f(c))$ 到 f的图像的欧几里得距离是在 $(c, f(c))$ 处取得的。固定 $x \in K \setminus \{c\}$，并注意到

$$\begin{aligned} \|v + (c, f(c)) - (x, f(x))\| &\geq \|P_{(\Gamma_L)^{\perp}}(v + (c, f(c)) - (x, f(x)))\| \\ &= \|P_{(\Gamma_L)^{\perp}}(v) + P_{(\Gamma_L)^{\perp}}((c, f(c))) - P_{(\Gamma_L)^{\perp}}(x, f(x)))\| \\ &= \|2v - P_{(\Gamma_L)^{\perp}}((x, f(x)))\| \\ &\geq 2\|v\| - \|P_{(\Gamma_L)^{\perp}}((x, f(x)))\| \geq \|v\|. \end{aligned}$$

于是，根据命题 2.5，我们有 $v \in N_c(f)$，这正是要证明的结论。 □

3. Clarke subdifferential

本章将讨论函数满足局部利普希茨条件的情形。

定义 3.1。设 $A \subset \mathbb{R}^n$。若函数 $f\colon A \to \mathbb{R}$ 满足以下条件，则称其为局部利普希茨 函数：对任意点 $a \in A$，都存在 $M > 0, \varepsilon > 0$，使得

$$\forall x, y \in A :\ (x, y \in K(a, \varepsilon)) \Rightarrow (|f(x) - f(y)| \leq M\|x - y\|).$$

本节后续讨论的函数 f 均为按照 3.1 中定义的局部利普希茨函数。局部 一词可以省略，因为集合 $A \subset \mathbb{R}^n$可以缩小为点 $a \in A$的一个经过适当调整的邻域。

为方便读者，我们先回顾几个基本概念：函数 f 在点 a 沿方向 $v \in \mathbb{R}^n$ 的广义方向导数（见 [3],）。

定义 3.2. 设 $f\colon A \to \mathbb{R}$ 是一个 Lipschitz 函数，$A \subset \mathbb{R}^n$。函数 f 在点 $a \in A$ 沿方向$v \in \mathbb{R}^n$ 的广义方向导数 是以下极限

$$f^{\circ}(a; v) = \limsup_{x \to a, t \to 0^+} \frac{f(x + tv) - f(x)}{t}.$$

命题 3.3. $f^\circ(a;\ v)$ 作为单变量函数 v 有界、正齐次且次可加。此外，

$$|f^\circ(a;v)| \le M\|v\| \ \forall a \in A, v \in \mathbb{R}^n,$$

其中， M 是 f的*Lipschitz*常数。

我们可以对函数 $f^\circ(a;\ v)$应用哈恩-巴拿赫定理，从而找到一个线性泛函 ξ : $\mathbb{R}^n \to \mathbb{R}$ 使得

$$f^\circ(a;v) \ge \xi(v), \forall a \in A, v \in \mathbb{R}^n.$$

定义 3.4。 集合 $\partial f(a) = \{\xi \in \mathbb{R}^n \colon\ f^\circ(a;\ v) \ge \langle \xi, v\rangle \ \forall v \in \mathbb{R}^n\}$ 称为 f 在 $a \in A$处的广义梯度。

现在我们来关注广义梯度的基本性质，这些性质见 [3]。

命题 3.5。对于任意 $\lambda \in \mathbb{R}$ 我们有 $\partial\left(\lambda f\right)\ (a)\ = \lambda \partial f\ (a)$ 。

命题 3.6。 如果函数 f 在 a 处取得局部极值，则 $0 \in \partial f(a)$。

命题 3.7。 对于 *Lipschitz* 函数 $f_i \colon\ A_i \to \mathbb{R},\ i = 1, \ldots, n,\ A_i \subset \mathbb{R}^n,$ $a \in \bigcap_{i=1}^n A_i$ ， 有

$$\partial\left(\sum_{i=1}^n f_i\right)(a) \subset \sum_{i=1}^n \partial f_i(a),$$

右侧的集合由形如以下形式的元素组成

$$\sum_{i=1}^n \xi_i : \xi_i \in \partial f_i,\ i = 1, \ldots, n.$$

我们可以利用广义梯度，将中值定理表述为适用于 Lipschitz 函数的形式。在 [3] 中可以找到以下事实，称为 Lebourg 定理。

定理 3.8 (Lebourg)。 设 $U \subset \mathbb{R}^n$ 为一个开集，且 $x, y \in U$ 满足 $[x, y] \subset U$。那么对于 *Lipschitz* 函数 $f \colon\ U \to \mathbb{R}$ ，存在一点 $c \in (x, y)$ 使得 $f(y) - f(x) \in \langle \partial f(c)\ y - x\rangle$.

次微分还有其他几种定义。在一些资料中（例如参见 [1]或 [9]） ，可以找到以下定义：

定义 3.9。 设 U 是 $\mathbb{R}^n$ 的一个非空开集，且 $f \colon U \to \mathbb{R}$ 是一个局部利普希茨函数。对于 $x \in U$ ，定义：

$\widehat{}$ （1）弗雷歇次微分 $\partial f\ (x)$ 在 f 处 x 定义为

$$\widehat{\partial} f(x) = \left\{ v \in \mathbb{R}^n : \liminf_{y \to x, y \ne x} \frac{f(y) - f(x) - \langle v, y - x\rangle}{\|y - x\|} \ge 0 \right\},$$

（2）极限次微分 $\partial' f\ (x)$ 在 f 处 x 定义为

$$v \in \partial' f(x) \Leftrightarrow \exists \{x_n\}_{n=1}^{+\infty} \subset U, \{v_n\}_{n=1}^{+\infty} \subset \mathbb{R}^n :$$
$$x_n \to x, v_n \to v\ (n \to +\infty), v_n \in \widehat{\partial} f(x_n), \forall n \in \mathbb{N},$$

(3) 将 f 在 x 处的 克拉克次微分 $\partial^\circ f(x)$ 表示为 $\partial' f(x)$ （$\partial^\circ f(x) = cvx\ \partial' f(x)$）的凸包。

由此立即可得 $\partial f(x) \subset \partial' f(x) \subset \partial^\circ f(x)$。

在 [9], 中，所有不同次微分概念之间的联系都得到了清晰的阐释。特别是，众所周知，对于 f为局部利普希茨函数的情形，有 $\partial f(x) = \partial^\circ f(x)$。

注 3.10. 弗雷歇次微分具有一种便于理解的几何解释。对于 $x \in U$，考虑切锥 $C_x(f)$，并将所有收敛到 x 的序列$\{x_n\}_{n=1}^{+\infty} \subset U$ 都纳入考虑，使得

$$\exists (v, w) \in C_x(f) : \frac{x_n - x}{\|x_n - x\|} \to v, \ \frac{f(x_n) - f(x)}{\|x_n - x\|} \to w \ (n \to +\infty)$$

且 $x_n \neq x \ \forall n \in \mathbb{Z}_+$。那么， $\partial f(x)$可以看作由满足以下条件的向量 $\xi \in \mathbb{R}^n$ 构成的集合：对于任意序列$\{x_n\}_{n=1}^{+\infty}$，都有

$$\lim_{n \to +\infty} \frac{f(x_n) - f(x) - \langle \xi, x_n - x \rangle}{\|x_n - x\|} \geq 0$$

或者 $w \geq \langle \xi, \ v \rangle, \ \forall (v, w) \in C_x(f)$。

现在我们将看到 Clarke 次微分与第一节中讨论的 BTC 次微分之间的联系。

定理 3.11。设 U 是 $\mathbb{R}^n$中的开集， $a \in U$ 并设 $f: \ U \to \mathbb{R}$ 是一个局部利普希茨函数。那么我们有 $\partial^\circ f(a) = B_a(f)$。

证明。不失一般性，我们可以假设 $a = f(a) = 0$。固定$v \in \partial^\circ f(a)$。首先，我们将证明

$$(x, \langle v, x \rangle) \in B_a(f), \ \forall x \in \mathbb{R}^n,$$

这将分三步完成。

步骤 1。 $v \in \partial f(0)$。

固定 $x \in \mathbb{R}^n, x \neq 0$。令 Γ_x 表示函数图像在 $[0, x]$线段上方的截面。我们取满足以下条件的最小数 $y_1, y_2 \in \mathbb{R}$

$$(-x, y_1), (x, y_2) \in C_0(\Gamma_x) \subset C_0(f),$$

由 Lipschitz条件可知，这些数必定存在。

如果 $y_1 < -y_2$，那么根据 3.10，$\partial f(0) = \emptyset$，且无需证明任何内容。假设$y_1 \geq -y_2$。于是我们找到序列 $\{a_n\}_{n=1}^{+\infty} \subset (0, -x]$, $\{b_n\}_{n=1}^{+\infty} \subset (0, x]$ 满足

$$a_n \to 0, \ b_n \to 0 \ (n \to +\infty),$$

$$\left| \frac{f(a_n)}{\|a_n\|} - \frac{y_1}{\|x\|} \right| \leq \frac{1}{2^n}, \ \left| \frac{f(b_n)}{\|b_n\|} - \frac{y_2}{\|x\|} \right| \leq \frac{1}{2^n}, \forall n \in \mathbb{N}.$$

不失一般性，假设 $|a_n| |b_n| < \frac{1}{2^n}$ 对 $n \in \mathbb{N}$成立；如有必要，可选取适当的子序列。现在我们希望构造一个双序列 $\{(c_n, d_n)\}_{n=1}^{+\infty} \subset U^2$，使得

$$c_n, d_n \to 0, \ \frac{f(c_n) - f(d_n)}{\|c_n - d_n\|} \to \frac{\langle v, x \rangle}{\|x\|} \ (n \to +\infty).$$

固定 $N \in \mathbb{Z}_+$。如果

$$\left| \frac{f(a_N) - f(b_N)}{\|a_N - b_N\|} - \frac{\langle v, x \rangle}{\|x\|} \right| \leq \frac{1}{2^N},$$

那么我们取 $c_N = a_N, d_N = b_N$。否则，如果

$$\frac{f(a_N) - f(b_N)}{\|a_N - b_N\|} < \frac{\langle v, x\rangle}{\|x\|} - \frac{1}{2^N},$$

然后我们固定 $c_N = a_N$，并考虑映射

$$g : U \ni y \mapsto \frac{f(a_N) - f(y)}{\|a_N - y\|} \in \mathbb{R}.$$

它是连续函数的复合，因此连续；此外

$$g(b_N) < \frac{\langle v, x\rangle}{\|x\|} - \frac{1}{2^N},\ g(0) \geq \frac{y_2}{\|x\|} - \frac{1}{2^N} \geq \frac{\langle v, x\rangle}{\|x\|} - \frac{1}{2^N}.$$

因此，我们可以找到一个中值点 $\xi \in [0, b_N]$，使得

$$g(\xi) = \frac{\langle v, x\rangle}{\|x\|} - \frac{1}{2^N}$$

并选择 $d_N = \xi$。如果

$$\frac{f(a_N) - f(b_N)}{\|a_N - b_N\|} > \frac{\langle v, x\rangle}{\|x\|} + \frac{1}{2^N},$$

那么我们令 $d_N = b_N$，并以类似的方式找到一个中值点 $c_N \in [a_N, 0]$，使得

$$\frac{f(c_N) - f(d_N)}{\|c_N - d_N\|} = \frac{\langle v, x\rangle}{\|x\|} + \frac{1}{2^N}.$$

可以很容易看出，这个双序列满足所需条件。此外，所有线段 $[c_i, d_i],\ i \in \mathbb{Z}_+$ 都平行于线段 $[0, x]$。

步骤 2. $v \in \partial' f(0)$。

设 $\{x_n\}_{n=1}^{+\infty} \subset U$ 和 $\{v_n\}_{n=1}^{+\infty} \subset \mathbb{R}^n$ 为满足以下条件的序列：

$$x_n \to 0\ (n \to +\infty),\ v_n \in \widehat{\partial} f(x_n)\ \forall n \in \mathbb{N},\ \|v_n - v\| \leq \frac{1}{2^n}, \forall n \in \mathbb{N}.$$

固定一点 $x \in \mathbb{R}^n$ 后，对于所有的数对 $(x_i, v_i),\ i \in \mathbb{Z}_+$，我们按第一步中的方法构造双序列$\{(c_k^{(i)}, d_k^{(i)})\}_{k=1}^{+\infty} \subset U^2$。那么，对于任意 $n \in \mathbb{N}$，有

$$\begin{aligned}\left|\frac{f(c_n^{(n)}) - f(d_n^{(n)})}{\|c_n^{(n)} - d_n^{(n)}\|} - \frac{\langle v, x\rangle}{\|x\|}\right| &\leq \left|\frac{f(c_n^{(n)}) - f(d_n^{(n)})}{\|c_n^{(n)} - d_n^{(n)}\|} - \frac{\langle v_n, x\rangle}{\|x\|}\right| + \left|\frac{\langle v_n, x\rangle}{\|x\|} - \frac{\langle v, x\rangle}{\|x\|}\right| \\ &\leq \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}}.\end{aligned}$$

因此，对于双序列 $\{(c_n^{(n)}, d_n^{(n)})\}_{n=1}^{+\infty}$，有

$$\frac{f(c_n^{(n)}) - f(d_n^{(n)})}{\|c_n^{(n)} - d_n^{(n)}\|} \to \frac{\langle v, x\rangle}{\|x\|}\ (n \to +\infty),$$

且 $c_n^{(n)}, d_n^{(n)} \to 0\,(n \to +\infty)$。此外，每条线段 $[c_i^{(i)}, d_i^{(i)}],\ i \in \mathbb{Z}_+$ 都平行于线段 $[0, x]$。

步骤 3. $v \in \partial^\circ f(0)$。

如果集合 $\partial' f(0)$ 是凸的，那么无需证明。否则，我们选取一个向量 $v \in \partial^\circ f(0) \setminus \partial' f(0)$，并固定 $v_1, v_2 \in \partial' f(0)$，使得 $v = Tv_1 + (1 - T)v_2$

对某个 $T \in (0,1)$。类似于第二步，对于给定的 $x \in \mathbb{R}^n$，我们可以构造双序列 $\{(a_n,b_n)\}_{n=1}^{+\infty}, \{(c_n,d_n)\}_{n=1}^{+\infty} \subset U^2$，使得

$$a_n,b_n,c_n,d_n \to 0,\ \frac{f(a_n)-f(b_n)}{\|a_n-b_n\|} \to \frac{\langle v_1,x\rangle}{\|x\|},\ \frac{f(c_n)-f(d_n)}{\|c_n-d_n\|} \to \frac{\langle v_2,x\rangle}{\|x\|}\ (n\to+\infty).$$

现在我们可以寻找一个双序列 $\{(e_n,f_n)\}_{n=1}^{+\infty} \subset U^2$，使得

$$e_n,f_n \to 0,\ \frac{f(e_n)-f(f_n)}{\|e_n-f_n\|} \to \frac{\langle v,x\rangle}{\|x\|}\ (n\to+\infty).$$

不失一般性，我们可以假设 $\langle v_1,x\rangle < \langle v_2,x\rangle$。于是

$$\exists\ N_0 \in \mathbb{Z}_+ : \frac{\langle v_1,x\rangle}{\|x\|} + \frac{1}{2^{N_0}} \le \frac{\langle v,x\rangle}{\|x\|} \le \frac{\langle v_2,x\rangle}{\|x\|} - \frac{1}{2^{N_0}}.$$

对于所有 $i \in \mathbb{Z}_+$，我们考虑一个连续映射

$$g_i : [0,1] \ni t \mapsto \frac{f((1-t)a_i+tc_i)-f((1-t)b_i+td_i)}{\|(1-t)(a_i-b_i)+t(c_i-d_i)\|} \in \mathbb{R}.$$

现在令 $N \in \mathbb{Z}_+, N > N_0$。根据第二步，我们有

$$\begin{aligned} g_N(0) = \frac{f(a_N)-f(b_N)}{\|a_N-b_N\|} &\le \frac{\langle v_1,x\rangle}{\|x\|} + \frac{1}{2^{N-1}} \le \frac{\langle v,x\rangle}{\|x\|} \\ &\le \frac{\langle v_2,x\rangle}{\|x\|} - \frac{1}{2^{N-1}} \le \frac{f(c_N)-f(d_N)}{\|c_N-d_N\|} = g_N(1). \end{aligned}$$

因此，我们可以找到一个中值点 $t_N \in [0,1]$，使得

$$g_N(t_N) = \frac{\langle v,x\rangle}{\|x\|}.$$

现在只需取 $e_{N-N_0} = (1-t_N)a_N + t_N c_N$, $f_{N-N_0} = (1-t_N)b_N + t_N d_N$ 即可。这个双序列满足我们所需的条件。

因此我们有 $\partial^\circ f(a) = B_a(f)$。反过来，假设

$$(x,\langle v,x\rangle) \in B_a(f), \forall x \in \mathbb{R}^n$$

对某个 $v \in \mathbb{R}^n$ 成立。固定 $x_0 \in \mathbb{R}^n, x_0 \neq 0$。不失一般性，令 $\|x_0\| = 1$。那么存在一个双序列 $\{a_n,b_n\}_{n=1}^{+\infty}$，使得

$$a_n,b_n \to a,\ \frac{a_n-b_n}{\|a_n-b_n\|} \to x_0,\ \frac{f(a_n)-f(b_n)}{\|a_n-b_n\|} \to \langle v,x_0\rangle\ (n\to+\infty).$$

令 $v_n = a_n - b_n$。那么 $\|v_n\| \to 0_+$ 且 $\frac{v_n}{\|v_n\|} \to x_0$。此外，

$$\begin{aligned} \langle v,x_0\rangle &= \lim_{n\to+\infty} \frac{f(a_n)-f(b_n)}{\|a_n-b_n\|} \\ &= \lim_{n\to+\infty} \frac{f(b_n+v_n)-f(b_n)}{\|v_n\|} \\ &= \lim_{n\to+\infty} \frac{f(b_n+\|v_n\|x_0)-f(b_n)}{\|v_n\|} + \lim_{n\to+\infty} \frac{f(b_n+v_n)-f(b_n+\|v_n\|x_0)}{\|v_n\|}. \end{aligned}$$

对于上述求和的第一分量，我们有如下近似：

$$\lim_{n\to+\infty} \frac{f(b_n+\|v_n\|x_0)-f(b_n)}{\|v_n\|} \le \limsup_{b\to a, t\to 0_+} \frac{f(b+tx_0)-f(b)}{t} = f^\circ(a;x_0).$$

对于第二分量，根据Lipschitz条件：

$$\begin{aligned}\lim_{n\to+\infty}\frac{|f(b_n+v_n)-f(b_n+\|v_n\|x_0)|}{\|v_n\|} &\leq \lim_{n\to+\infty}\frac{L\|(b_n+v_n)-(b_n+\|v_n\|x_0)\|}{\|v_n\|}\\ &= \lim_{n\to+\infty}\frac{L\|v_n-\|v_n\|x_0\|}{\|v_n\|}\\ &= \lim_{n\to+\infty}L\left\|\frac{v_n}{\|v_n\|}-x_0\right\| = 0,\end{aligned}$$

其中 L 表示 f的Lipschitz常数。最后，$\langle v, x_0\rangle \leq f^{\circ}(a;\ x_0)$，由 x_0 的任意性可得 $v \in \partial f(a)$。□

作为上述定理的一个推论，我们得到了Lebourg定理一个非常优雅的几何解释。

推论 3.12. 设 $U \subset \mathbb{R}^n$ 为开集，且 $x, y \in U$满足 $[x, y] \subset U$。那么，对于*Lipschitz* 函数 $f:\ U \to \mathbb{R}$，存在一点 $c \in (x, y)$ 使得$\frown f(y) - f(x) \in \langle B_c(f)\ y - x\rangle$.

参考文献

[1] J. Bolte, A. Daniilidis, A. Lewis, M. Shiota, 亚解析*Lipschitz*连续函数的*Clarke*临界值, Ann. Polon. Math., 1987; [2] G. Bouligand, 关于一阶限制拓扑的若干问题, Bull. Soc. Math. France, 56:407–420, 1928;[3] F. H. Clarke, 《优化与非光滑分析》, John Wiley and Sons, 纽约, 1983; [4] J. B. Diaz, R. Výborný, 微分学的广义中值定理, J. Austral. Math. Soc. 20 (Series A), 290-300, 1975; [5] J. -B. Hiriart-Urruty, 非光滑优化中向量值映射的中值定理, Numer. Funct. Anal. Optim., 2, 1–30, 1980;[6] V. Jeyakumar, D. T. Luc, 非光滑连续映射与*C1*优化的近似雅可比矩阵。《*SIAM*控制与优化杂志》, 36(5), 1815–1832, 1998; [7] G. Lebourg, 广义梯度的中值 ´定 ´理 ´, Comptes Rendus de l'Acad ´ émie des Sciences de Paris, 281, 795–797, 1975; [8] W. Mlak, 关于中值定理的注记, Ann. Pol. Math. 3, 29-31, 1956; [9] R. T. Rockafellar, R. J.-B. Wets, 变分分析, Springer-Verlag, 1998; [10] F. Severi, 代数几何讲座, Reggia Universit'a di Roma (Anno 1927-1928), 1928;[11] T. Ważewski, 有限增量定理的推广, Ann. Polon. Math. 24, 132-147, 1953; [12] H. Whitney, 《复解析簇》, Addison-Wesley Publishing Company, 普林斯顿, 209-212, 1972;[13] Yuntong, W. 副切锥与*P*-次导数的一种刻画及其在非光滑分析中的应用, 《数学学报》7, 181–192, 1991.

Karolina Zając - 雅盖隆大学，数学与计算机科学学院，数学研究所，Łojasiewicza 6，30-348 克拉科夫，波兰。电子邮件地址：karolina99.zajac@student.uj.edu.pl