随机分析中的几种导数

姓名：　张竣淞

学号：　2024233014

时间：　二〇二六年二月十四日

On Several Types of Derivatives in Stochastic Analysis

Author: Zhang Junsong

Completion date: February 14, 2026

摘要

本文从泛函分析中的 Fréchet 导数和 Gâteaux 导数出发, 对随机分析中常见的几种导数及其性质进行简单的介绍, 包括定义, 例子和一些相关的计算等.

关键词：随机分析；Fréchet 导数；Lions 导数；内蕴导数；外在导数；线性泛函导数；水平导数；垂直导数；Malliavin 导数；Malliavin 散度；Itô 公式

ABSTRACT

This note commences with the Fréchet derivative and Gâteaux derivative in functional analysis, and provides a concise introduction to several common derivatives in stochastic analysis and their properties, including definitions, examples, and some relevant computations.

KEY WORDS: Fréchet derivative, L-derivative, extrinsic derivative, intrinsic derivative, Dupire's derivative, Malliavin derivative, Malliavin divergence, Itô's formula

目录

第 1 章　导论

在古典微积分中，对于一个性质足够好的函数

$$f : \mathbb{R} \to \mathbb{R},$$

可以从两种观点去看待它在 $x \in \mathbb{R}$ 处的局部性态, 一种是差商的极限, 称 f 在 x 处可导, 是指极限：

$$f'(x) := \lim_{y \to 0} \frac{f(x+y) - f(x)}{y},$$

存在. 另一种则是用线性映射局部逼近, 即称 f 在 $x \in \mathbb{R}$ 处可微, 是指: 存在一个连续线性映射

$$L_x : \mathbb{R} \to \mathbb{R},$$

使得当 $y \to 0$ 时,

$$\delta(y) = f(x+y) - f(y) - L_x(y) \to 0.$$

并且由于 $\mathbb{R}$ 上连续泛函全体 $\mathscr{L}(\mathbb{R};\mathbb{R})$ 同构于 $\mathbb{R}$, 所以这两种角度实际上是一致的.

当考虑 $f : \mathbb{R}^d \to \mathbb{R}, d \geq 2$ 时,$\mathbb{R}^d$ 中元素之间不存在天然的除法, 差商的极限的观点不能完全推广, 只能考虑方向导数, 即固定 $v \in \mathbb{R}^d$, 考察

$$\partial_v f(x) := \lim_{h \to 0 \in \mathbb{R}} \frac{f(x+hv) - f(x)}{h},$$

上面的极限如果存在, 则称为 f 沿 v 的方向导数. 而线性逼近的观点则可以完全推广, 定义完全一致, 即称 f 在 $x \in \mathbb{R}^d$ 处可微, 是指: 存在一个连续线性映射

$$L_x : \mathbb{R}^d \to \mathbb{R},$$

使得当 $|y| \to 0$ 时,

$$\delta(y) = f(x+y) - f(y) - L_x(y) \to 0.$$

由于 $\mathbb{R}^d$ 上连续泛函全体 $\mathscr{L}(\mathbb{R}^d;\mathbb{R})$ 同构于 $\mathbb{R}^d$, 该线性映射在 $\mathbb{R}$ 中的对应表示就是

$$\nabla f(x) := (\partial_{e_1} f(x), \cdots, \partial_{e_d} f(x)),$$

其中 e_i 表示 $\mathbb{R}^d$ 中第 i 个分量为 1, 其余分量为 0 的元素,$i = 1, \cdots, d$. 虽然这两种观点依然有紧密的联系, 但由于空间几何性质较之一维情形复杂很多, 所以两者并不完全等同, 具体地说: 如果 f 在一点可微, 则该点处各方向导数均存在; 但各方向导数存在并不能推出可微性, 需要对方向导数添加额外的连续性要求.

当考虑更一般的赋范向量空间之间的映射时, 可以继承这两种观点去考察映射的局部性态: 从线性逼近的角度推广得到 Fréchet 导数的概念, 而从差商极限或方向导数的角度推广得到 Gâteaux 导数的概念, 这是第二章的主要内容; 而在此基础上, 出于随机分析研究的需要, 第三章从 Fréchet 可微的角度, 引入 L-可微和 Lions 导数的概念; 第四章则是从方向导数的角度, 引入外在导数和内蕴导数, 并简要介绍线性泛函导数; 第五章是为了研究路径依赖随机微分方程等对象, 引入了关于路径的泛函和路径导数的概念. 第六章将重点介绍 Malliavin 导数和相关性质, 及其在随机微分方程中的初步应用.

最后在此声明, 本文作为一份学习笔记, 仅是参考文献列表中相关内容的收集, 整理和翻译, 偶尔穿插解释性论述, 几乎并无原创性内容, 每章开始部分, 会给出本章内容主要取用的来源. 对于笔者认为"标准"的内容, 包括基础的定义定理等, 不再特别说明; 为了便于理解或保持记号的一致, 对原文献中的一些记号或证明可能有细微的修改, 不再特别指出. 内容的取舍是完全主观的, 如有欠缺或冗余之处, 敬请谅解.

第 2 章　Banach 空间中的 Fréchet 导数和 Gâteaux 导数

本章内容主要参考引用自 (Ciarlet, 2025) 的 Chapter 9: Differential Calculus in Normed Vector Spaces 及其第一版的中译本 (Ciarlet, 2017b) 的第七章相关内容。

如无特殊说明, 本章涉及 X, Y 为 Banach 空间, 即完备赋范向量空间,$\|\cdot\|_X, \|\cdot\|_Y$ 分别为 X, Y 上的完备范数, 在不致混淆的前提下, 均简记为 $\|\cdot\|$;$0_X, 0_Y$ 分别表示 X, Y 中的零元, 在不致混淆的前提下均简记为 0;X 上拓扑一般为该完备范数诱导的拓扑, Ω 为 X 中在范数拓扑下的开集.$\mathscr{L}(X;Y)$ 表示 X 到 Y 的有界线性算子全体,$A \in \mathscr{L}(X;Y)$ 的算子范数为

$$\|A\|_{\mathscr{L}(X;Y)} := \sup_{x\in X:\|x\|_X>0} \frac{\|Ax\|_Y}{\|x\|_X}.$$

$\mathscr{L}_k(X;Y)$ 表示从 $\underbrace{X\times\cdots\times X}_{k\text{个}}$ 到 Y 的连续 $k-$ 线性映射全体. 对赋范向量空间 X 和给定的 $a, b \in X$, 定义开区间 $(a, b) := \{ta + (1-t)b : 0 < t < 1\}$ 和闭区间 $[a, b] := \{ta + (1-t)b : 0 \le t \le 1\}$

2.1　Fréchet 导数和 Gâteaux 导数的定义和例子

2.1.1　Fréchet 导数的定义

定义 2.1　给定映射 $f : \Omega \to Y, x \in \Omega$, 称 f 在 x 处 Fréchet 可微, 是指: 存在 $A \in \mathscr{L}(X;Y)$, 使得当 $h \to 0_X$, 时

$$\delta(h) := \frac{f(x+h) - f(h) - Ah}{\|h\|_X} \to 0_Y. \tag{2.1}$$

记此连续线性算子 A 为 $f'(x)$ 或 $df(x)$, 称为 f 在 x 处的 Fréchet 导数.

根据定义, 可以得到 Fréchet 导数以下两个基本性质.

命题 2.1　给定映射 $f : \Omega \to Y, x \in \Omega$, 若 $f'(x)$ 存在, 则 $f'(x)$ 是唯一的.

证明　假设存在 $A_1, A_2 \in \mathscr{L}(X;Y)$ 均满足定义2.1中的要求. 由于 Ω 为开集, 存在 $r > 0$ 使得 $B(x,r) \subset \Omega$. 对任意 $h \in B(0,r)$,

$$\begin{aligned} f(x+h) &= f(x) + A_1 h + \|h\|\,\delta_1(h) \\ &= f(x) + A_2 h + \|h\|\,\delta_2(h), \end{aligned} \tag{2.2}$$

所以对任意 $h \in B(0,r)$,

$$\|(A_1 - A_2)h\| = \|h\|\,\|\delta_1(h) - \delta_2(h)\|, \tag{2.3}$$

对任意 $u \in X$, 存在充分大的 $R > 0$ 使得 $\frac{1}{R}u \in B(0, r)$, 则

$$\begin{aligned}\|(A_1 - A_2)u\| &= R\left\|(A_1 - A_2)(\frac{u}{R})\right\| \\ &\leq R\left\|\frac{u}{R}\right\|\left\|\delta_1(\frac{u}{R}) - \delta_2(\frac{u}{R})\right\| \\ &= \|u\|\left\|\delta_1(\frac{u}{R}) - \delta_2(\frac{u}{R})\right\|. \end{aligned} \tag{2.4}$$

令 $R \to \infty$ 即可. ■

命题 2.2　若 $f : \Omega \to Y$ 在 $x \in \Omega$ 处 Fréchet 可微, 则 f 在 x 处连续.

定义 2.2　称 $f : \Omega \to Y$ 在 Ω 内可微, 是指: 任意 $x \in \Omega$,f 在 x 处可微. 此时定义映射

$$\begin{aligned} f' : \Omega &\to \mathscr{L}(X; Y) \\ x &\mapsto f'(x). \end{aligned}$$

若 f' 是连续的, 则称 f 在 Ω 内连续可微, 简称在 Ω 内是 C^1 的.

若 f 为单射,$f(\Omega)$ 为 Y 中开集且 $f^{-1} : f(\Omega) \to X$ 在 $f(\Omega)$ 内是 C^1 的, 则称 f 是一个 C^1 微分同胚.

注　记 $C^1(\Omega; Y) := \{f : \Omega \to Y : f在\Omega内是C^1的\}$, 特别地, 当 $Y = \mathbb{R}$ 时, 记 $C^1(\Omega) = C^1(\Omega; \mathbb{R})$. 容易验证,$C^1(\Omega; Y)$ 是向量空间.

类似于欧式空间, 考虑 X 或 Y 为有限个 Banach 空间的乘积空间的情形, 对乘积空间配备最大值范数, 即考虑 $X = X_1 \times \cdots \times X_n$, 其中 $X_i, 1 \leq i \leq n$ 为 Banach 空间, 对 $x = (x_1, \cdots, x_n) \in X$, 定义 $\|x\|_\infty = \max_{1\leq i\leq n} \|x_i\|_{X_i}$. 乘积空间在最大值范数下仍为 Banach 空间且最大值范数诱导了乘积拓扑 (见 (Ciarlet, 2017a) 的 2.2 节).

定理 2.3　设 $Y_1, \cdots, Y_m$ 为 Banach 空间,$Y = Y_1 \times \cdots \times Y_m$ 并配备最大值范数. 给定映射 $f_i : \Omega \to Y, 1 \leq i \leq m$ 和映射

$$\begin{aligned} f : \Omega &\to Y = Y_1 \cdots Y_m \\ x &\mapsto (f_1(x), \cdots, f_m(x)). \end{aligned} \tag{2.5}$$

则 f 在 $a \in \Omega$ 处可微等价于任意 $1 \leq i \leq m$,f_i 在 a 处可微, 且此时 $f'(a) \in \mathscr{L}(X; Y)$ 等同于 $(f_1'(a), \cdots, f_m'(a)) \in \mathscr{L}(X; Y_1) \times \cdots \times \mathscr{L}(X; Y_m)$.

证明　若 f 在 $a \in \Omega$ 处可微, 则

$$f(a + h) = f(a) + f'(a)h + \|h\|\,\delta(h),$$

其中当 $h \to 0_X$ 时,$\delta(h) \to 0_Y$.

写成分量形式即:

$$f_i(a + h) = f_i(a) + A_ih + \|h\|\,\delta_i(h), i = 1, \cdots, m,$$

其中 $A_i \in \mathscr{L}(X;Y_i)$ 是 $f \in \mathscr{L}(X;Y)$ 的第 i 个分量,$i=1,\cdots,m$. 由此即得 f_i 在 a 处的可微性以及 $f_i'(a)=A_i$.

反过来, 若任意 $1 \le i \le m$,f_i 在 $a \in \Omega$ 处可微, 即

$$f_i(a+h)=f_i(a)+f_i'(a)h+\|h\|\,\delta_i(h), i=1,\cdots,m,$$

定义线性映射

$$\begin{aligned} A : X &\to Y \\ x &\mapsto (f_1'(a)x,\cdots,f_m'(a)x), \end{aligned}$$

则有

$$f(a+h)=f(a)+Ah+\|h\|\,\delta(h), \tag{2.6}$$

其中 $\delta(h)=(\delta_1(h),\cdots,\delta_m(h))$. 由于

$$\begin{aligned} \|Ah\|_Y &= \max_{1\le i\le m}\|f_i'(a)h\|_{Y_i} \\ & \max_{1\le i\le m}\|f_i'(a)\|_{\mathscr{L}(X;Y_i)}\|h\|_X, \end{aligned} \tag{2.7}$$

所以 $A \in \mathscr{L}(X;Y)$, 从而 f 在 a 处可微且 $f'(a)=A$. ■

下面考虑 X 为乘积空间的情况.

定义 2.3　给定 $X_1,\cdots,X_n$ 为 Banach 空间,$X=X_1\times\cdots\times X_n$, 配备最大值范数,$\Omega$ 为 Banach 空间 X 中开集,$a\in\Omega$, 且任意 $1\le i\le n$, 存在 X_i 中开集 Ω_i, 使得 $\Omega_1\times\cdots\times\Omega_n\subset\Omega$. 若对某个 $1\le i\le n$, 映射

$$\begin{aligned} f(a_1,\cdots,a_{i-1},\cdot,a_{i+1},\cdots,a_n) : \Omega_i &\to Y \\ x &\mapsto f(a_1,\cdots,a_{i-1},x,a_{i+1},\cdots,a_n) \end{aligned}$$

在 a_i 处可微, 则称其导数为 f 在 a 处的第 i 个偏导数, 记为 $\partial_i f(a)$.

下面的定理给出了映射在一点的微分和其偏导数间的关系.

定理 2.4　给定 $X_1,\cdots,X_n$ 为 Banach 空间,$X=X_1\times\cdots\times X_n$. 若映射 $f:\Omega\subset X\to Y$ 在 $a\in\Omega$ 处可微, 则其各偏导数均存在, 且任意

$$f'(a)h=\sum_{i=1}^n \partial_i f(a)h_i,\ h=(h_1,\cdots,h_n)\in X.$$

证明　只证 $n=2$ 的情形,$n\ge 3$ 时证明没有本质区别.

定义线性映射

$$\begin{aligned} A_1 : X_1 &\to Y \\ h_1 &\mapsto f'(a)(h_1,0), \end{aligned}$$

和

$$\begin{aligned} A_2 : X_2 &\to Y \\ h_2 &\mapsto f'(a)(0, h_2). \end{aligned}$$

则

$$\begin{aligned} f(a_1 + h_1, a_2) &= f(a_1, a_2) + f'(a)(h_1, 0) + \|(h_1, 0)\|_\infty \delta(h_1, 0) \\ &= f(a_1, a_2) + A_1 h_1 + \|h_1\|_{X_1} \delta(h_1, 0). \end{aligned} \tag{2.8}$$

且 $\|A_1 h_1\|_Y = \|f'(a)(h_1, 0)\|_Y \le \|f'(a)\|_{\mathscr{L}(X;Y)} \|h_1\|_{X_1}$, 所以 $A_1 = \partial_1 f(a) \in \mathscr{L}(X_1; Y)$. 对第二个分量的证明类似. ■

2.1.2　Gâteaux 导数的定义

Fréchet 可微可以看作是欧氏空间之间映射局部线性逼近的自然推广, 下面考虑推广方向导数的概念.

定义 2.4　给定映射 $f : \Omega \to Y$,$a \in \Omega, h \in X$ 以及一个开区间 $I_h \in \mathbb{R}$, 其中 I_h 满足:

(1) $0 \in I_h$;

(2) $\forall \theta \in I_h$,$a + \theta h \in \Omega$.

称 f 在 a 处有沿方向 h 的 Gâteaux 导数, 是指: 极限

$$\partial_h f(a) := \lim_{\theta \to 0} \frac{f(a + \theta h) - f(a)}{\theta}$$

在 Y 中存在,$\partial_h f(a)$ 即为 f 在 a 处沿方向 h 的 Gâteaux 导数.

注　与欧式空间之间的映射相同, 若 f 在一点处可微, 则沿任意方向的 Gâteaux 导数均存在, 并且 $\partial_h f(a) = f'(a)h$; 反之则不成立. 此外, 定义2.3中的偏导数可以看作特殊方向上的 Gâteaux 导数.

2.1.3　若干例子

下面介绍几个简单的例子.

例 2.1 (仿射映射)　给定 $A \in \mathscr{L}(X; Y), b \in Y$, 定义映射

$$\begin{aligned} f : X &\to Y \\ x &\mapsto Ax + b. \end{aligned}$$

易见 $\forall x \in X, f'(x) = A$.

例 2.2 (双线性映射)　设 $B : X \times X \to Y$ 是连续双线性映射. 由于 B 是双线性的,

$$B(a_1 + h_1, a_2 + h_2) = B(a_1, a_2) + (B(h_1, a_2) + B(a_1, h_2)) + B(h_1, h_2),$$

且 $\|B(h_1,h_2)\| \leq \|B\|\,\|h_1\|\,\|h_2\|$, 所以

$$B'(a_1,a_2)(h_1,h_2)=B(h_1,a_2)+B(a_1,h_2),\ (h_1,h_2)\in X_1\times X_2,$$
$$\partial_1 B(a_1,a_2)(h_1)=B(h_1,a_2),\ h_1\in X_1,$$
$$\partial_2 B(a_1,a_2)(h_2)=B(a_1,h_2),\ h_2\in X_2.$$

注　$B'(a_1,a_2)$ 表示一个连续线性算子, 而 $B'(a_1,a_2)(h_1,h_2)$ 则表示该算子作用于 (h_1,h_2).

例 2.3 (续例2.2)　设 B 是对称的, 即 $B(x,y)=B(y,x),\forall x,y\in X$. 定义映射 $\tilde{B}(x):=B(x,x)$. 则 $\tilde{B}'(a)(h)=2B(a,h)$.

进一步地, 设 $M\in\mathscr{L}_k(X;Y)$ 是一个对称的 $k-$ 线性映射, 令 $\tilde{M}(x)=M(x,\cdots,x)$, 则 $(\tilde{M})'(a)(h)=kM(a,\cdots,a,h)$, 简记 $(\tilde{M})'(a)=Ma^{k-1}$.

例 2.4 (方阵的函数)　记 $M_n(\mathbb{R})$ 为 n 阶实方阵的全体,$U_n(\mathbb{R})$ 为 n 阶可逆实方阵的全体,$M_n(\mathbb{R})\xrightarrow{\text{tr}}\mathbb{R}$ 和 $M_n(\mathbb{R})\xrightarrow{\det}\mathbb{R}$ 分别表示取迹和行列式运算.

$\text{tr}(\cdot)$ 作为连续线性泛函,$\forall W\in M_n(\mathbb{R})$,$\text{tr}(\cdot)$ 在 W 处的 Fréchet 导数由 $(\text{tr})'(W)(H)=\text{tr}(H)=\langle I_n,H\rangle$ 决定, 其中 $\langle A,B\rangle:=\text{tr}(A^*B)$ 表示 n 阶实方阵上的内积.

另一方面, 若 $W\in U_n(\mathbb{R})$, 则

$$\begin{aligned}\det(W+H)&=\det(W)\det(I_n+W^{-1}H)\\&=\det(W)(1+\text{tr}(W^{-1}H)+o(\|W^{-1}H\|))\\&=\det(W)+\text{tr}((\det(W)W^{-1})H)+\det(W)o(\|W^{-1}H\|).\end{aligned}$$

所以 $(\det)'(W)(H)=\text{tr}((\text{Adj}(W))H)=\langle\text{Adj}(W)^T,H\rangle$, 其中 $\text{Adj}(W)$ 表示 W 的伴随矩阵.

例 2.5 (续例2.4)　考虑取逆映射

$$\begin{aligned}f:U_n(\mathbb{R})&\to U_n(\mathbb{R})\\A&\mapsto A^{-1}.\end{aligned}$$

由 (Ciarlet, 2017a) 的定理 3.6-3, 当 H 的矩阵范数充分小时

$$\begin{aligned}f(A+H)&=(A+H)^{-1}\\&=(I_n+A^{-1}H)^{-1}A^{-1}\\&=(I_n-A^{-1}H+o(H))A^{-1}\\f(A)&-A^{-1}HA^{-1}+o(H)A^{-1}.\end{aligned}$$

所以 $f'(A)(H)=-A^{-1}HA^{-1}$.

2.2　链式法则

在古典微积分中, 链式法则是计算导数的重要工具, 在一般的 Banach 空间中对 Fréchet 同样有链式法则.

定理 2.5 (链式法则)　给定 Banach 空间 X, Y, Z,U, V 分别为 X, Y 中开集. 映射 $f : U \to Y$ 在 $a \in U$ 处可微且 $f(U) \subset V$, 映射 $g : V \to Z$ 在 $f(a) \in V$ 处可微. 则映射 $g \circ f : U \to Z$ 在 a 处可微, 且

$$(g \circ f)'(a) = g'(f(a)) \circ f'(a).$$

此外, 若 $f \in C^1(U; Y), g \in C^1(V; Z)$, 则 $g \circ f \in C^1(U; Z)$.

证明　对任意 $a + h \in U$, 定义

$$b := f(a), k(h) = f(a + h) - f(a).$$

由 f, g 的可微性可得,

$$f(a + h) = f(a) + f'(a)h + \|h\| \delta(h), \lim_{h\to 0} \delta(h) = 0$$

$$g(b + k) = g(b) + g'(b)k + \|k\| \eta(k), \lim_{k\to 0} \eta(k) = 0.$$

所以

$$\begin{aligned}
&(g \circ f)(a + h) - (g \circ f)(a) \\
=&g(f(a + h)) - g(f(a)) \\
=&g'(f(a))(k(h)) + \|k(h)\| \eta(k(h)) \\
=&g'(f(a))(f'(a)h + \|h\| \delta(h)) + \|k(h)\| \eta(k(h)) \\
=&g'(f(a))(f'(a)h) + \|h\| g'(f(a))\delta(h) + \|k(h)\| \eta(k(h))
\end{aligned} \tag{2.9}$$

并且当 $h \to 0$ 时,

$$\frac{\|h\| g'(f(a))\delta(h) + \|k(h)\| \eta(k(h))}{\|h\|} \to 0,$$

另外连续线性映射的复合仍然是连续线性映射, 所以 $g \circ f$ 在 a 处可微且 $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$.

若 $f \in C^1(U; Y), g \in C^1(V; Z)$, 为证 $g \circ f \in C^1(U; Z)$, 只需证映射

$$x \mapsto (f'(x), g'(f(x))) = (f'(x), (g' \circ f)(x))$$

和

$$\mathscr{L}(X; Y) \times \mathscr{L}(Y; Z) \ni (A, B) \mapsto B \circ A \in \mathscr{L}(X; Z)$$

的连续性. 而这两个的映射的连续性则由 f', g' 以及映射复合保持连续性得到.

$\blacksquare$

2.3　高阶导数和 Taylor 展开

在古典微积分中, 如果一个函数 $f:\mathbb{R}\to\mathbb{R}$ 处处可导, 则自然地有一个导函数 $f':\mathbb{R}\to\mathbb{R}$, 则可以考察导函数的连续性和可微性. 而对于一般的向量空间, 若 X, Y 是 Banach 空间, 则 $\mathscr{L}(X;Y)$ 也是 Banach 空间, 所以对于 $f\in C^1(\Omega;Y)$, 同样可以考察 $f':\Omega\to\mathscr{L}(X;Y)$ 的 Fréchet 可微性, 也就是二阶导数.

定义 2.5　设 $f:\Omega\to Y$ 在 Ω 内可微. 称 f 在 $a\in\Omega$ 处二次可微, 是指: 映射

$$\begin{aligned} f':\Omega&\to\mathscr{L}(X;Y)\\ x&\mapsto f'(x) \end{aligned}\tag{2.10}$$

在 a 处 Frehcet 可微. 记 $f''(a):=(f')'(a)$ 为 f 处的二次导数.

若 f 在 Ω 内处处二次可微, 且映射

$$\begin{aligned} f'':\Omega&\to\mathscr{L}(X;\mathscr{L}(X;Y))\\ x&\mapsto f''(x) \end{aligned}\tag{2.11}$$

是连续的, 则称 f 在 Ω 内是 C^2 的,Ω 内到 Y 的 C^2 映射全体记为 $C^2(\Omega;Y)$, 特别地,$Y=\mathbb{R}$ 时, 记 $C^2(\Omega;Y)=C^2(\Omega)$

注　由 (Ciarlet, 2017a) 的定理 2.11-5,$\mathscr{L}(X;\mathscr{L}(X;Y))$ 同构于 $\mathscr{L}_2(X;Y)$, 所以对 $h,k\in X$, $(f''(a)h)k=f''(a)(h,k)$, 这里 $f''(a)$ 作为 $\mathscr{L}(X;\mathscr{L}(X;Y))$ 中的元素先作用于 h 再作用于 k, 但下面的定理说明这两步操作的顺序可以交换: 即先作用于 h 再作用于 k 和先作用于 k 再作用于 h 的结果是相同的, 也即 $f''(a)$ 作为一个双线性映射是对称的.

定理 2.6　设 $f:\Omega\to Y$ 在 $a\in\Omega$ 处二次可微, 则 $f''(a)\in\mathscr{L}_2(X;Y)$ 是对称的.

定理的证明见 (Ciarlet, 2017b) 的定理 7.8-1, 该证明依赖于如下的中值定理.

定理 2.7　给定映射 $f:\Omega\to Y$, 闭区间 $[a,b]\subset\Omega$,$\forall x\in[a,b]$,f 在 x 处连续,$\forall x\in(a,b)$,f 在 x 处可微. 则

$$\|f(b)-f(a)\|_Y\le\Big(\sup_{x\in(a,b)}\|f'(x)\|_{\mathscr{L}(X;Y)}\Big)\|b-a\|_X\,.$$

证明　不妨设 $M=\sup_{x\in(a,b)}\|f'(x)\|_{\mathscr{L}(X;Y)}<\infty$.

定义映射 $\phi(t):=f(ta+(1-t)b), t\in[0,1]$, 则 $\phi:[0,1]\to Y$ 连续, 在 $(0,1)$ 可微, 且

$$\phi'(t)=f'(ta+(1-t)b)(b-a), t\in(0,1)\implies\sup_{t\in(0,1)}\|\phi'(t)\|\le M\|b-a\|$$

以下讨论被称为“连续性方法”.

对任意 $\epsilon > 0$, 定义映射

$$
\begin{aligned}
F : [0,1] &\to \mathbb{R} \\
t &\mapsto \|\phi(t) - \phi(0)\| - (M\|b-a\| + \epsilon)t - \epsilon
\end{aligned}
$$

以及 $I(\epsilon) = F^{-1}((\infty, 0])$, 由于 $0 \in I(\epsilon)$, 所以 $I(\epsilon)$ 非空.

F 是由连续函数的复合和四则运算得到, 因此仍是连续的, 从而 $I(\epsilon)$ 是闭集, 所以 $t_0 := \sup(I(\epsilon) \cap [0,1]) \in I(\epsilon)$. 往证 $t_0 = 1$.

假设 $t_0 < 1$, 则对于充分小的 $\delta > 0$ 有 $t_0 + \delta < 1$,

$$
\phi(t_0+\delta) - \phi(t_0) = \phi(t_0) + \phi'(t_0)\delta + \delta\eta(\delta),\ \lim_{\delta\to 0}\eta(\delta) = 0.
$$

取充分小 δ_0 使得 $t_0 + \delta_0 < 1$ 且 $\eta(\delta) < \epsilon$, 则

$$
\begin{aligned}
\|\phi(t_0+\delta_0) - \phi(0)\| &\le \|\phi(t_0+\delta_0) - \phi(t_0)\| + \|\phi(t_0) - \phi(0)\| \\
&\le (M\|b-a\| + \epsilon)(t_0+\delta_0) + \epsilon. \qquad (2.12)
\end{aligned}
$$

矛盾, 所以 $t_0 = 1$. 这也就证明了定理.

■

上述中值定理只对单点处可微性有要求, 如果 $f \in C^1(\Omega; Y)$, 则还有更精确的表达.

定理 2.8　设 $f \in C^1(\Omega; Y)$, $[a,b] \subset \Omega$. 则

$$
f(b) - f(a) = \int_0^1 f'((1-t)a + tb)(b-a)\mathrm{d}t.
$$

注　关于向量值函数的积分或称 Bochner 积分的相关定义和性质见 (Yosida, 2022), 上述定理的证明见 (Ciarlet, 2017a).

类似于二阶导数, 可以归纳定义高阶导数, 为统一记号, 记 $f^{(0)} = f, f^{(1)} = f', f^{(2)} = f''$.

定义 2.6　设 $m \in \mathbb{N}$, 设映射 $f : \Omega \to Y$ 在 Ω 内 $(m-1)$ 次可微, 称 f 在 $a \in \Omega$ 处 m 次可微, 是指: 映射

$$
\begin{aligned}
f^{(m-1)} : \Omega &\to \mathscr{L}(X; \mathscr{L}_{m-1}(X;Y)) \\
x &\mapsto f^{(m-1)}(x)
\end{aligned}
$$

在 a 处可微,a 处的 m 阶导数记为 $f^{(m)}(a) = (f^{(m-1)})'(a)$. 若 m 阶导数映射连续, 则称 f 在 Ω 内是 C^m 的; 若 f 是任意次可微的, 则称 f 是光滑的.

注　类似 $C^1(\Omega; Y)$ 和 C^1 微分同胚, 可以定义 $C^m(\Omega; Y)$ 和 C^m 微分同胚,$m \in \mathbb{N} \cup \{\infty\}$; 另外, 简记 $f^{(m)}(a)h^m = f^{(m)}(a)(\underbrace{h, \cdots, h}_{m\text{个}})$

有了以上准备, 可以叙述并证明 Banach 空间中的 Taylor 公式.

定理 2.9 (Taylor 公式)　给定映射 $f:\Omega\to Y$,$[a,a+h]\subset\Omega$, 以及整数 $m\geq 1$.

(1) f 在 Ω 中 $(m-1)$ 次可微且在 a 处 m 次可微, 则

$$f(a+h)=f(a)+\sum_{i=1}^{m}\frac{1}{i!}f^{(i)}(a)h^i+\|h\|^m\delta(h).$$

(2) 若 $Y=\mathbb{R}$,f 在 Ω 中 $(m-1)$ 次连续可微, 在 $(a,a+h)$ 中 m 次可微, 则存在 $\theta\in(0,1)$ 使得

$$f(a+h)=f(a)+f'(a)h+\cdots+\frac{1}{(m-1)!}f^{(m-1)}(a)h^{m-1}+\frac{1}{m!}f^{(m)}(a+\theta h)h^m.$$

(3) 设 f 在 Ω 内 m 次连续可微, 则

$$f(a+h)=f(a)+f'(a)h+\cdots+\frac{1}{(m-1)!}f^{(m-1)}(a)h^{m-1}+\frac{1}{(m-1)!}\int_0^1(1-t)^{m-1}(f^m(a+th)h^m)\mathrm{d}t.$$

(4) 设 f 在 Ω 中 $(m-1)$ 次连续可微, 在 $(a,a+h)$ 中 m 次可微, 则

$$\left\|f(a+h)-f(a)+\sum_{i=1}^{m-1}\frac{1}{i!}f^{(i)}(a)h^i\right\|\leq\frac{1}{m!}(\sup_{x\in(a.a+h)}\left\|f^{(m)}(x)\right\|)\|h\|^m.$$

注　上述定理的 (1),(2),(3) 分别对应单变量微积分中带 Peano 余项,Lagrange 余项和积分余项的 Taylor 展开公式,(4) 则是中值定理2.3对高阶导数的推广.

证明

先证明 (1). 当 $m=1$ 时根据导数定义自动成立. 假设对 $m=1,\cdots,k-1$ 都成立. 存在 $r>0$ 使得 $B(a,r)\subset\Omega$ 且映射

$$g:B(a,r)\to Y$$

$$x\mapsto f(a+x)-f(a)-(\sum_{i=1}^{k}\frac{1}{i!}f^{(k)}(a)x^k)$$

在 $B(a,r)$ 内可微, 且

$$g'(x)=f'(a+x)-(\sum_{i=1}^{k}\frac{1}{(i-1)!}f^{(k)}(a)x^{k-1}).$$

根据归纳假设,

$$f'(a+x)=f'(a)+\cdots+\frac{1}{(k-1)!}f^{(k)}(a)x^{k-1}+\|x\|^{k-1}\delta(x)$$

且 $\lim_{x\to 0}\delta(x)=0$. 由中值定理2.3,

$$\begin{aligned}&\left\|f(a+h)-f(a)-(\sum_{i=1}^{m}\frac{1}{i!}f^{(i)}(a)h^i)\right\|\\&=\|g(h)-g(0)\|\\&\leq(\sup_{x\in(a,a+h)}\|g'(x)\|)\|h\|\\&\leq(\sup_{x\in(a,a+h)}\|x\|^{k-1}\|\delta(x)\|)\|h\|\\&\leq\|h\|^k\|\eta(h)\|.\end{aligned}\tag{2.13}$$

且 $\lim_{h\to 0}\eta(h)=0_Y$.

定义辅助函数 $\phi(t):=f(a+th), t\in(0,1)$, 则在 (2) 的条件下分别使用链式法则和单变量实函数的带 Lagrang 余项的 Taylor 展开式即可证得 (2).

而在 (3) 的条件下可知 ϕ 在一个包含 $[0,1]$ 的开区间上是 m 次连续可微的. 定义函数 $\psi(t):=\phi(t)+\sum_{i=1}^{m-1}(1-t)^i\phi^{(i)}(t)$, 再将定理2.8应用于

$$\psi(1)-\psi(0)=\int_0^1\psi'(t)\mathrm{d}t$$

即可证得 (3).

最后证明 (4). 当 $m=1$ 时由中值定理2.3成立, 假设对 $m=1,\cdots,k-1$ 均成立.

令 $u(t)=\phi(t)-(f(a)+\sum_{i=1}^{k-1}\frac{1}{i!}f^{(i)}(a)(th)^{i-1})$. 由 f 的可微性可得 u 在包含 $[0,1]$ 的一个开区间上可微.

对 f' 用归纳假设可得,

$$\begin{aligned}\|u'(t)\|&=\left\|(f'(a+th)-(f'(a)+\cdots+\frac{1}{(k-2)!}f^{(k-1)}(th)^{k-2})h)\right\|\\&\le\frac{1}{(k-1)!}(\sup_{x\in(a,a+h)}\left\|f^{(k)}(x)\right\|)t^{k-1}\left\|h\right\|^k, 0\le t\le 1.\end{aligned}\tag{2.14}$$

令 $\chi(t):=\frac{1}{(k)!}(\sup_{x\in(a,a+h)}\left\|f^{(k)}(x)\right\|)t^k\left\|h\right\|^k$. 则

$$\begin{aligned}\|u(1)-u(0)\|&\le\sum_{i=0}^{l-1}\left\|u(\frac{j+1}{l})-u(\frac{j}{l})\right\|\\&\le\frac{1}{l}\sum_{i=0}^{l-1}\sup_{t\in(\frac{j}{l},\frac{j+1}{l})}\|u'(t)\|\\&\le\frac{1}{l}\sum_{j=0}^{l-1}\sup_{t\in(\frac{j}{l},\frac{j+1}{l})}\chi'(t).\end{aligned}\tag{2.15}$$

右边是一个 Riemann 和的形式, 令 $l\to\infty$, 即得

$$\|u(1)-u(0)\|\le\chi(1)-\chi(0).$$

也即证得了结论. ■

参考文献

Ciarlet P G, 2017a. 线性与非线性泛函分析及其应用（上册）[M]. 秦铁虎, 童裕孙, 译. 北京: 高等教育出版社. 4, 7, 9, 10

Ciarlet P G, 2017b. 线性与非线性泛函分析及其应用（下册）[M]. 秦铁虎, 译. 北京: 高等教育出版社. 3, 9

Yosida K, 2022. 泛函分析[M]. 吴元恺, 孙顺华, 唐志远, 等译. 6 版. 北京: 高等教育出版社. 10

Ciarlet P G, 2025. Linear and nonlinear functional analysis with applications, second edition[M]. 2nd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics. 3