随机分析中的几种导数

姓名：　张竣淞

学号：　2024233014

时间：　二〇二六年一月十一日

On Several Types of Derivatives in Stochastic Analysis

Author: Zhang Junsong

Completion date: January 11, 2026

摘要

本文从泛函分析中的 Frechet 导数和 Gateaux 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，Frechet 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；Ito 公式

ABSTRACT

This note commences with the Frechet derivative and Gateaux 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: Frechet derivative, L-derivative, extrinsic derivative, intrinsic derivative, Dupire's derivative, Malliavin derivative, Malliavin divergence, Ito'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 Frechet 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 $[a, b] := \{ta + (1-t)b : 0 \leq t \leq 1\}$

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 Frechet 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 Frechet 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

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

令 $R \to \infty$ 即可. ■

命题 2.2　若 $f : \Omega \to Y$ 在 $x \in \Omega$ 处 Frechet 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 $X_i, 1 \le i \le n$ 为 Banach 空间, 对 $x = (x_1, \cdots, x_n) \in X$, 定义 $\|x\|_\infty = \max_{1 \le i \le n} \|x_i\|_{X_i}$. 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 $f_i : \Omega \to Y, 1 \le i \le m$ 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 $1 \le i \le m$,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:

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

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

$$\begin{aligned} f(a_1+h_1,a_2) &= f(a_1,a_2)+f'(a)(h_1,0)+\|(h_1,0)\|_\infty \delta(h_1,0) \\ &= f(a_1,a_2)+A_1h_1+\|h_1\|_{X_1}\delta(h_1,0). \end{aligned} \tag{2.8}$$

且 $\|A_1h_1\|_Y = \|f'(a)(h_1,0)\|_Y \leq \|f'(a)\|_{\mathscr{L}(X;Y)}\|h_1\|_{X_1}$, 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　Gateaux 导数的定义

Frechet 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 Gateaux 导数, 是指: 极限

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

在 Y 中存在,$\partial_h f(a)$ 即为 f 在 a 处沿方向 h 的 Gateaux 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 Gateaux 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 Gateaux 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 $\|B(h_1,h_2)\| \le \|B\|\,\|h_1\|\,\|h_2\|$, 所以

$$\begin{aligned}
&B'(a_1,a_2)(h_1,h_2)=B(h_1,a_2)+B(a_1,h_2),\ (h_1,h_2)\in X_1\times X_2,\\
&\partial_1 B(a_1,a_2)(h_1)=B(h_1,a_2),\ h_1\in X_1,\\
&\partial_2 B(a_1,a_2)(h_2)=B(a_1,h_2),\ h_2\in X_2.
\end{aligned}$$

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,$M_n(\mathbb{R})\xrightarrow{\mathrm{tr}}\mathbb{R}$ 和 $M_n(\mathbb{R})\xrightarrow{\det}\mathbb{R}$ 分别表示取迹和行列式运算.

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

另一方面, 若 $W\in U_n(\mathbb{R})$, 则

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

所以 $(\det)'(W)(H)=\mathrm{tr}((\mathrm{Adj}(W))H)=\langle \mathrm{Adj}(W)^T,H\rangle$, 其中 $\mathrm{Adj}(W)$ 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 Frechet 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

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

由 f,g 的可微性可得,

$$f(a+h)=f(a)+f'(a)h+\|h\|\,\delta(h),\lim_{h\to0}\delta(h)=0$$

$$g(b+k)=g(b)+g'(b)k+\|k\|\,\eta(k),\lim_{k\to0}\eta(k)=0.$$

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 $h\to0$ 时,

$$\frac{\|h\|\,g'(f(a))\delta(h)+\|k(h)\|\,\eta(k(h))}{\|h\|}\to0,$$

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.

■

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 Frechet 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

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

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,

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

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

第 3 章　$\mathscr{P}_2$ 上的函数的 Lions 导数

本章内容主要参考引用自 (Carmona et al., 2018) 的 5.2 节.

记 $\mathscr{P}$ 表示 $\mathbb{R}^d$ 上全体概率测度,$\mathscr{P}_p$ 表示全体 p 阶矩有限概率测度, 即

$$\mathscr{P}_p = \{\mu \in \mathscr{P} : \int_{\mathbb{R}^d} |x|^p \mu(\mathrm{d}x) < \infty\}.$$

$\mathscr{P}_p$ 在 $p-$Wasserstein 距离 $W_p(\cdot,\cdot)$ 下是一个完备度量空间,$W_p(\cdot,\cdot)$ 定义为

$$W_p(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \pi(\mathrm{d}x,\mathrm{d}y)\Big)^{\frac{1}{p}},$$

其中 $\Pi(\mu,\nu)$ 表示所有边缘分布为 μ,ν 的联合分布. 本章中如无特别说明, 总默认 $\mathscr{P}_2$ 配备 $W_2(\cdot,\cdot)$ 距离及其诱导的度量拓扑. 称 $\mathscr{P}_2$ 的子集 K 是有界的, 是指: 存在 $R>0$, 使得 $\forall \mu \in K$, $\int_{\mathbb{R}^d} |x|^2 \mathrm{d}\mu(x) \le R$.

对一个函数 $f : \mathscr{P}_2 \to \mathbb{R}$, 我们希望考察 $\mu \in \mathscr{P}_2$ 处的局部性态, 但由于 $\mathscr{P}_2$ 不是 Banach 空间, 甚至不是向量空间, 所以不能直接使用第一章中的 Frechet 导数和 Gateaux 导数.Lions 导数的想法是不直接考察概率测度 μ 附近函数的性态, 而是考虑将 f 先" 提升" 为定义在 $L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d)$ 上的函数 $\tilde f$, 考察分布为 μ 的随机变量 $X \in L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d)$ 附近 $\tilde f$ 的性态, 而 $L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d)$ 不仅是 Banach 空间, 也是 Hilbert 空间, 所以可以应用 Frechet 可微性的概念. 一个自然的提升是令 $\tilde f = f \circ \mathscr{L}$, 其中 $\mathscr{L}(X)$ 表示 $\mathbb{R}^d$ 上由随机变量 X 诱导的概率分布, 也即 $\mathscr{L}(X)(A) = \mathbb{P}(X^{-1}(A)), \forall A \in \mathscr{B}(\mathbb{R}^d)$.

$$\begin{tikzcd}
X \arrow[rr, mapsto] & & \mu = \mathscr{L}(X) \arrow[dd, mapsto] \\
L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d) \arrow[r, "\mathscr{L}"] \arrow[dr, dashed, "\tilde f = f\circ\mathscr{L}"'] & B \arrow[d, "f"] & \\
& \mathbb{R} & f(\mu)
\end{tikzcd}$$

由 (Bogachev, 2007) 的命题 9.1.11, 对于任意的无原子的概率空间, 任意分布 $\mu \in \mathscr{P}$, 总存在一个随机变量 X 使得 $\mathscr{L}(X) = \mu$, 这一事实保证了提升的合理性. 本章讨论提升时, 无原子的概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 总是固定的, 简记 $L^2(\Omega;\mathbb{R}^d) = L^2(\Omega,\mathscr{F},\mathbb{P};\mathbb{R}^d)$. 对任意 $A \in \mathscr{F}$, 记 $\mathscr{F}|_A := \{E \cap A : E \in \mathscr{F}\}$. 另外, 在本章中, 对于一个 Banach 空间 B 上的函数 F, 不加区分的使用 $F'(a)$ 和 $DF(a)$ 表示 F 在

$a \in B$ 处的 Frechet 导数, 同样地, 若有定义, 不加区分地使用 F' 和 DF 表示 F 的导数映射.

3.1　L-可微性

3.1.1　定义

定义 3.1　给定一个函数 $f : \mathscr{P}_2 \to \mathbb{R}$, 称 f 在 $\mu_0 \in \mathscr{P}_2$ 是 L-可微的, 是指: 存在一个分布为 μ_0 的随机变量 X_0, 使得 f 的提升 $\tilde{f} = f \circ \mathscr{L} : L^2(\Omega; \mathbb{R}^d) \to \mathbb{R}$ 在 X_0 处是 Frechet 可微的.

这个定义的表述是自然的, 但良定性并不显然, 具体地说, 对于同一个概率测度, 可能有多个随机变量以其为分布, 那么就需要说明这些随机变量和其 Frechet 导数有某种不变量或者在某种等价关系下属于同一等价类, 使得定义在一定的标准下具有唯一性. 为此, 我们先陈述以下引理.

引理 3.1　设 $X, Y \in L^2(\Omega; \mathbb{R}^d)$ 分布相同. 则对任意 $\epsilon > 0$, 存在保测映射[①] $S, T : \Omega \to \Omega$ 使得

$$\mathbb{P}(\{\omega \in \Omega : (S \circ T)(\omega) = (T \circ S)(\omega) = \omega\}) = 1$$

且

$$\mathbb{P}(|Y - X \circ T| \le \epsilon) = 1.$$

证明梗概　记 $\mathscr{F}^*, \mathbb{P}^*$ 分别为 $\mathscr{F}, \mathbb{P}$ 的完备化扩张.

对给定的 ϵ, 设 $\{A_n\}_{n=1}^{\infty} \subset \mathscr{B}(\mathbb{R}^d)$ 是 $\mathbb{R}^d$ 的一个剖分, 且任意 A_n 直径不超过 ϵ(例如边长为 $\frac{\epsilon}{\sqrt{d}}$ 的半开半闭的方体). 令 $B_n = X^{-1}(A_n), C_n = Y^{-1}(A_n)$, 由于 X, Y 同分布, 所以 $\mathbb{P}(B_n) = \mathbb{P}(C_n)$. 不妨设 $\mathbb{P}(B_n) > 0$. 由于 $(\Omega, \mathscr{F}, \mathbb{P})$ 无原子, 所以存在 $M_n \subset \tilde{M}_n \subset B_n, N_n \subset \tilde{N}_n \subset C_n$,$\mathbb{P}(\tilde{M}_n) = \mathbb{P}(\tilde{N}_n) = 0$.

记 $\Omega_n^X = B_n - M_n, \mathscr{F}_n^X = \{A \cap \Omega_n^X : A \in \mathscr{F}^*\}, \mathbb{P}_n^X(E) = \frac{\mathbb{P}^*(E)}{\mathbb{P}(B_n)}, E \in \mathscr{F}_n^X$ 以及 $\Omega_n^Y = C_n - N_n, \mathscr{F}_n^Y = \{A \cap \Omega_n^Y : A \in \mathscr{F}^*\}, \mathbb{P}_n^Y(E) = \frac{\mathbb{P}^*(F)}{\mathbb{P}(C_n)}, F \in \mathscr{F}_n^Y$.

由 (Bogachev, 2007) 的推论 6.6.7 和定理 9.2.2 可知, 存在保测的双射

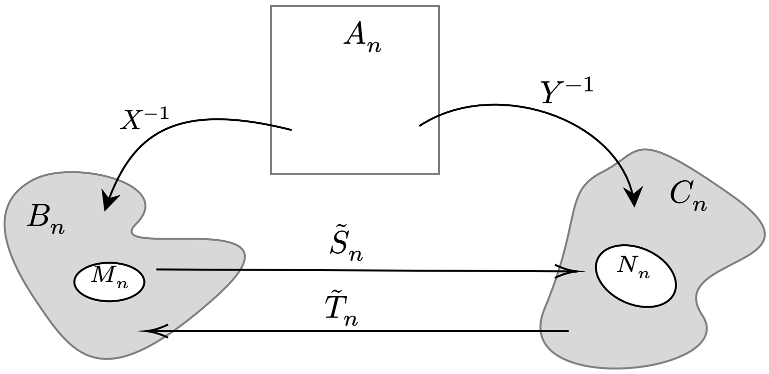
$$\tilde{S}_n : (\Omega_n^X, \mathscr{F}_n^X, \mathbb{P}_n^X) \to (\Omega_n^Y, \mathscr{F}_n^Y, \mathbb{P}_n^Y).$$

和

$$\tilde{T}_n : (\Omega_n^Y, \mathscr{F}_n^Y, \mathbb{P}_n^Y) \to (\Omega_n^X, \mathscr{F}_n^X, \mathbb{P}_n^X)$$

且 $\tilde{S}_n \circ \tilde{T}_n = \mathrm{id}_{\Omega_n^X}, \tilde{T}_n \circ \tilde{S}_n = \mathrm{id}_{\Omega_n^Y}$.

[①] 设 $(X_1, \mathscr{F}_1, \mu_1), (X_2, \mathscr{F}_2, \mu_2)$ 是两个测度空间, 称 $M : (X_1, \mathscr{F}_1, \mu_1) \to (X_2, \mathscr{F}_2, \mu_2)$ 是一个保测映射, 是指:M 可测且任意 $A \in \mathscr{F}_2, \mu_2(A) = \mu_1(M^{-1}(A))$.



下面把 $\tilde{S}_n, \tilde{T}_n$ 从 $B_n - \tilde{M}_n, C_n - \tilde{N}_n$ 延拓到 B_n, C_n 上, 记为 S_n, T_n. 只需补充在 $\tilde{M}_n, \tilde{N}_n$ 上的定义, 具体地说, 任取 $x_n \in M_n, y_n \in N_n$, 令 $\forall x \in \tilde{M}_n, S_n(x) = y_n, \forall y \in \tilde{N}_n, T_n(y) = x_n$ 即可. 容易验证, 延拓后的映射 S_n 是 $(B_n, \mathscr{F}|_{B_n})$ 到 $(C_n, \mathscr{F}|_{C_n})$ 的保测映射.T_n 类似.

当 $\mathbb{P}(B_n) = 0$ 时, 取 $M_n = B_n, N_n = C_n$.

令 $S = \sum_{n\geq 1} S_n I_{B_n}, T = \sum_{n\geq 1} T_n I_{C_n}$, 则 S, T 是 $(\Omega, \mathscr{F}, \mathbb{P})$ 到自身的保测映射. 且以概率 1 互为逆.

最后, 对几乎处处 ω, 存在 n, 使得 $\omega \in C_n$, 则 $T(\omega) \in B_n$, 所以 $X(T(\omega)), Y(\omega) \in A_n$. ■

因为 $L^2(\Omega; \mathbb{R}^d)$ 是一个 Hilbert 空间, 所以对函数 $F : L^2(\Omega; \mathbb{R}^d) \to \mathbb{R}$ 在一点 X 处的 Frechet 导数 $DF(X)$ 作为一个 Hilbert 空间的对偶空间的元素, 总可以视为 $L^2(\Omega; \mathbb{R}^d)$ 中的某个随机变量, 对任意 $Y \in L^2(\Omega; \mathbb{R}^d)$ 的作用以 $DF(X)(Y) = \langle DF(X), Y\rangle = \mathbb{E}[DF(X) \cdot Y]$ 的形式表现.

定理 3.2　设函数 $f : \mathscr{P}_2 \to \mathbb{R}$ 在 $\mu_0 \in \mathscr{P}_2$ 处 L-可微. 则对于任意分布为 μ_0 的随机变量 $X \in L^2(\Omega; \mathbb{R}^d)$,$f$ 的提升 $\tilde{f}$ 在 X 处是 Frechet 可微的, 且 $(X, D\tilde{f}(X))$ 的联合分布不依赖于 X 的选取.

证明　由定义, 存在分布为 μ 的随机变量 $X_0 \in L^2(\Omega; \mathbb{R}^d)$,$\tilde{f}$ 在 X_0 处是 Frechet 可微的. 对任意分布为 μ 的随机变量 $X \in L^2(\Omega; \mathbb{R}^d)$, 任意 $\epsilon > 0$, 存在 Ω 到自身的保测映射 S_ϵ 和 T_ϵ 满足

$$\mathbb{P}(\{\omega \in \Omega : S(T(\omega)) = T(S(\omega)) = \omega\}) = 1$$

和

$$\mathbb{P}(|X_0 - X \circ S| \leq \epsilon) = 1.$$

注意到 $\tilde{f}$ 作为 f 的提升, 其函数值只依赖于随机变量的分布, 所以对任意 $Y \in$

$L^2(\Omega;\mathbb{R}^d)$,

$$
\begin{aligned}
\tilde{f}(X+Y) &= \tilde{f}((X+Y)\circ S_\epsilon) \\
&= \tilde{f}(X_0 + (X\circ S_\epsilon - X_0) + Y\circ S_\epsilon) \\
&= \tilde{f}(X_0) + (\tilde{f})'(X_0)((X\circ S_\epsilon - X_0) + Y\circ S_\epsilon) \\
&\quad + \|(X\circ S_\epsilon - X_0) + Y\circ S_\epsilon\|_{L^2}\,\delta((X\circ S_\epsilon - X_0) + Y\circ S_\epsilon) \\
&= \tilde{f}(X_0) + \mathbb{E}[D\tilde{f}(X_0)\cdot((X\circ S_\epsilon - X_0) + Y\circ S_\epsilon)] \\
&\quad + \|(X\circ S_\epsilon - X_0) + Y\circ S_\epsilon\|_{L^2}\,\delta((X\circ S_\epsilon - X_0) + Y\circ S_\epsilon) \\
&= \tilde{f}(X) + \mathbb{E}[(D\tilde{f}(X_0)\circ T_\epsilon)\cdot Y] + \mathbb{E}[D\tilde{f}(X_0)\cdot(X\circ S_\epsilon - X_0)] \\
&\quad + \|(X\circ S_\epsilon - X_0) + Y\circ S_\epsilon\|_{L^2}\,\delta((X\circ S_\epsilon - X_0) + Y\circ S_\epsilon). \qquad (3.1)
\end{aligned}
$$

其中 $\delta(\cdot)$ 与 ϵ 无关且 $\lim_{\|Y\|_{L^2}\to 0}\delta(Y)=0$, 第 1 个等号和第 5 个等号分别使用了 S_ϵ, T_ϵ 的保测性.

取 $\epsilon=\frac{1}{n}, n\in\mathbb{N}$, 记 $Z_n = D\tilde{f}(X_0)\circ T_{\frac{1}{n}}$. 由式 (3.1) 可得

$$
\mathbb{E}[(Z_n - Z - m)\cdot Y] \le C(|\frac{1}{n}-\frac{1}{m}| + o(\|Y\|_{L^2})),
$$

所以 $\{Z_n\}$ 会在 $L^2(\Omega;\mathbb{R}^d))$ 中收敛到某个随机变量, 最后在式 (3.1) 中令 $\epsilon\to 0$ 即可. ∎

L-可微性的定义是将定义域从一个概率测度组成的空间转换到一个随机变量空间上讨论的, 但应该注意到 $f:\mathscr{P}_2\to\mathbb{R}$ 在一点处的局部性态应该只和该点和其周围的概率测度决定, 下面的定理将说明 L-可微性确实是函数的内蕴性质, 并对 Lions 导数给出一个稍微独立于随机变量的表示. 在此之前, 首先引入连续 L-可微的概念, 这是连续 Frechet 可微的自然延伸, 称 $f:\mathscr{P}_2\to\mathbb{R}$ 是连续 L-可微的, 是指:f 在任意 $\mu\in\mathscr{P}_2$ 处是 L-可微的, 且 $D\tilde{f}(X)$ 是以 X 为自变量的从 $L^2(\Omega;\mathbb{R}_d)$ 到自身的连续映射.

定理 3.3　设 $f:\mathscr{P}_2\to\mathbb{R}$ 是连续 L-可微的. 则对任意 $\mu\in\mathscr{P}_2$, 存在可测函数 $\xi:\mathbb{R}_d\to\mathbb{R}_d$, 使得任意以 μ 为分布的随机变量 X, 都有 $D\tilde{f}(X)=\xi(X)$.

证明　对给定的 X, 首先证明 $D\tilde{f}(X)$ 关于 X 生成的 σ 代数 $\sigma(X)$ 可测. 不妨设 f 有界. 证明分为以下几步.

(1) 先考虑 μ 关于 Lebsgue 测度绝对连续, 且对 $\forall q>4$,$\mu(|\cdot|^q)<\infty$. 对 $\epsilon>0$, 定义

$$
\Psi(Y) = \tilde{f}(Y) + \frac{1}{2\epsilon}\mathbb{E}[|X-Y|^2] + \mathbb{E}[|Y|^4], \quad Y\in L^4(\Omega;\mathbb{R}^d),
$$

则 Ψ 在 $L^4(\Omega;\mathbb{R}^d)$ 是 Frechet 可微的, 且任意 $Y\in L^4(\Omega;\mathbb{R}^d)$,

$$
\Psi'(Y) = D\tilde{(f)}(Y) + \frac{1}{\epsilon}(Y-X) + 4\mathbb{E}[|Y|^2]Y.
$$

由于 Ψ 有下界, 从而一定有下确界, 取 $\{Z_n\} \subset L^4(\Omega;\mathbb{R}^d)$ 为 Ψ 的极小化序列, 即 $\lim_{n\to\infty}\Psi(Z_n) = \inf_{Z\in L^4(\Omega;\mathbb{R}^d)}\Psi(Z)$ 以及 $\nu_n = \mathscr{L}(\nu_n)$. 由 Brenier 定理 (陈述及证明可见 (Figalli et al., 2021) 的定理 2.5.10), 存在一列 $\mathbb{R}^d$ 上的凸函数 $\{\psi_n\}$ 满足 $\mathscr{L}(\nabla\psi(X_n)) = \nu_n$ 且是 μ—a.e. 可微的, 此外还有 $W_2(\mu,\nu_n) = \mathbb{E}[|X-\nabla\psi(X_n)|^2]$. 记 $Y_n = \nabla\psi(X_n)$, 则

$$\Psi(Y_n) = \tilde{f}(Y_n) + \frac{1}{2\epsilon}W_2(\mu,\nu_n)^2 + \mathbb{E}[|Y|^4] \le \Psi(Z_n), \tag{3.2}$$

从而 $\{Y_n\}$ 也是 Ψ 的一个极小化序列.

而 $\sup_{n\ge1}\nu_n(|\cdot|^4)\infty$, 结合 Chebyshev 不等式可得 $\{\nu_n\}$ 是胎紧的. 不妨设 ν_n 在 W_2 距离和弱收敛意义下收敛到 ν, 所以

$$\begin{aligned} f(\nu) &= \lim_{n\to\infty} f(\nu), \\ W_2(\mu,\nu) &= \lim_{n\to\infty} W_2(\mu,\nu_n). \end{aligned} \tag{3.3}$$

由 Fatou 引理的测度收敛版本 (见 (Bogachev, 2018) 的推论 2.2.6) 可得

$$\int_{\mathbb{R}^d} |x|^4\nu(\mathrm{d}x) \le \liminf_{n\to\infty}\mathbb{E}[|Y|^4] < \infty.$$

再次利用 Brenier 定理可得, 存在凸函数 $\psi : \mathbb{R}^d \to \mathbb{R}$ 使得 $Y := \nabla\psi(X)$ 分布为 ν 且 $W_2(\mu,\nu) = \mathbb{E}[|X-Y|^2]$. 因为 Ψ 在 Y 处取得极小值, 由 (Ciarlet, 2017b) 的定理 7.1-5 知 $D\Psi(Y) = 0$. 即

$$D\tilde{f}(Y) = -\frac{1}{\epsilon}(Y-X) - 4|Y|^2Y,$$

这说明 $D\tilde{f}(Y) \in \sigma(X)$.

注意到这里的 Y 是依赖于 ϵ 的, 不妨记为 Y_ϵ, 由于 Y_ϵ 极小化 Ψ_ϵ, 所以对任意 ϵ,

$$\mathbb{E}[|X-Y_\epsilon|^2] \le 2(\tilde{f}(X) + \mathbb{E}[|X|^4] - (\tilde{f}(Y_\epsilon) + \mathbb{E}[|Y_\epsilon|^4])),$$

所以令 $\epsilon \to 0, Y_\epsilon \to X$. 这就证得了 $D\tilde{f}(X) \in \sigma(X)$.

(2) 假设 $\mu \in \mathscr{P}_2$ 关于 Lebsgue 测度绝对连续但不一定有 (1) 中的矩条件, 取分布为 μ 的随机变量 $X \in L^2(\Omega;\mathbb{R}^d)$, 令 $X_n = \frac{nX}{\sqrt{n^2+X^2}}$, 再令 $n\to\infty$ 即可.

(3) 若 $\mu \in \mathscr{P}_2$ 不是关于 Lebsgue 测度绝对连续的, 取分布为 μ 的随机变量 $X \in L^2(\Omega;\mathbb{R}^d)$ 和 Ω 上两个独立的标准正态分布的随机变量 N_1, N_2, 且 X 独立于 N_1, N_2. 令 $X_{n,i} = X + \frac{1}{n}N_i, n\in\mathbb{N}, i = 1,2$, 则 $X_{n,i}$ 的分布总是关于 Lebsgue 测度绝对连续的, 从而可得 $D\tilde{f}(X) \in \sigma(X, N_i)$, 而由 N_1, N_2 的独立性, 可得 $D\tilde{f}(X) \in \sigma(X)$.

最后我们说明 ξ 不依赖于 X 的选取. 假设对两个分布为 μ 的二阶矩有限随机变量 X_1, X_2 分别有对应的 ξ_1, ξ_2, 任取 $A \in \mathscr{B}(\mathbb{R}^d)$, 易见 $\mu(\xi_1 I_A) = \mu(\xi_2 I_A)$, 所以 $\xi_1 = \xi_2, \mu$—a.e.. ■

有了以上定理, 我们可以借助概率空间和随机变量给出 L-可微函数 $f : \mathscr{P}_2 \to \mathbb{R}$ 在 μ_0 处的局部性态:

$$f(\mu) = f(\mu_0) + \mathbb{E}[D\tilde{f}(X_0) \cdot (X - X_0)] + o(\|X - X_0\|_{L^2}),$$

其中 $\mathscr{L}(X) = \mu, \mathscr{L}(X_0) = \mu_0$.

另一方面, 为了突出 L-可微的内蕴性, 我们将定理3.3中的 ξ 记为 $\partial_\mu f(\mu_0)$, 称为 f 在 μ_0 处的 Lions 导数或 L-导数. 此时有

$$f(\mu) = f(\mu_0) + \mathbb{E}[\partial_\mu f(\mu_0)(X_0) \cdot (X - X_0)] + o(\|X - X_0\|_{L^2}), \tag{3.4}$$

其中 $\mathscr{L}(X) = \mu, \mathscr{L}(X_0) = \mu_0$.

作为以上微分公式的一个应用, 给出如下命题, 这是 (Bao et al., 2021) 的定理 2.1 的特殊情形.

推论 3.4　设函数 $f : \mathscr{P}_2 \to \mathbb{R}$ 连续, 在一个无原子的概率测度 μ_0 处 L-可微, 存在常数 C 使得

$$|\partial_\mu f(\mu_0)(x)| \le C(1 + |x|).$$

$(X_s)_{s\in[0,1]} \subset L^2(\Omega; \mathbb{R}^d)$ 是一族随机变量,$\mathscr{L}(X_0) = \mu_0$,X_s 关于 s 连续, 且 $\dot{X}_0 := \lim_{s\downarrow} \frac{X_s - X_0}{s}$ 在 $L^2(\Omega; \mathbb{R}^d)$ 中存在. 则

$$\lim_{s\downarrow 0} \frac{f(\mathscr{L}(X_s)) - f(\mu_0)}{s} = \mathbb{E}[\langle \partial_\mu f(\mu_0)(X_0), \dot{X}_0 \rangle].$$

3.1.2　若干例子

例 3.1 (线性函数)　我们首先计算积分形式给出的线性函数. 给定光滑函数 $h : \mathbb{R}^d \to \mathbb{R}$, 且 $|\nabla h(x)| \le C(1 + |x|)$. 令 $f(\mu) = \int_{\mathbb{R}^d} h(x)\mu(\mathrm{d}x)$,$f$ 的提升为 $\tilde{f}(X) = \mathbb{E}[h(X)]$, 其中 $\mathscr{L}(X) = \mu$. 计算 $\tilde{f}$ 的 Frechet 导数, 首先由中值定理,

$$h(x + y) = h(x) + \int_0^1 (\nabla h(x + ty) \cdot y)\mathrm{d}t, \tag{3.5}$$

$$\begin{aligned} \mathbb{E}[h(X + Y)] &= \mathbb{E}[h(X) + \int_0^1 (\nabla h(X + tY) \cdot Y)\mathrm{d}t] \\ &= \mathbb{E}[h(X)] + \mathbb{E}[\nabla h(X) \cdot Y] + \mathbb{E}[\int_0^1 (\nabla h(X + tY) \cdot Y - \nabla h(X) \cdot Y)\mathrm{d}t], \end{aligned} \tag{3.6}$$

其中,

$$\begin{aligned}
&|\mathbb{E}[\int_0^1 ((\nabla h(X+tY)-\nabla h(X))\cdot Y)\,\mathrm{d}t]| \\
\le &\int_0^1 \mathbb{E}[(((\nabla h(X+tY)-\nabla h(X))\cdot Y))\, I_{\{|Y|\le \|Y\|_{L^2}^{1/2}\}}]\mathrm{d}t \\
&\quad +\int_0^1 \mathbb{E}[(((\nabla h(X+tY)-\nabla h(X))\cdot Y))\, I_{\{|Y|> \|Y\|_{L^2}^{1/2}\}}]\mathrm{d}t \\
\le &\mathbb{E}[\sup_{y\le \|Y\|_{L^2}^{1/2}} (\nabla h(X+y)-\nabla h(X))]\,\|Y\|_{L^2} \\
&\quad + C'\mathbb{E}[(1+|X|+|Y|)|Y|] \\
\le &\mathbb{E}[\sup_{y\le \|Y\|_{L^2}^{1/2}} (\nabla h(X+y)-\nabla h(X))]\,\|Y\|_{L^2} \\
&\quad + C'\,\|Y\|_{L^2}\,(\|Y\|_{L^2}+\|Y\|_{L^2}^{1/2}+\sup_{\mathbb{P}(A)\le \|\|_{L^2}Y}\mathbb{E}[|X|^2 I_A]^{1/2})
\end{aligned} \tag{3.7}$$

所以 $\partial_\mu f(\mu)=\nabla h, \forall \mu\in\mathscr{P}_2$.

这个例子虽然简单, 但也说明了 Lions 导数和古典导数以及 Frechet 导数的一些区别: 尽管对线性函数, 其 L-导数仍然是固定的, 但并非 h, 而是其梯度.

例 3.2 (续例3.1, 卷积函数)　接下来考虑一个稍微复杂的函数, 令 $h*\mu(x)=\int_{\mathbb{R}^d} h(x-y)\mu(\mathrm{d}y)$,$g(\mu)=f(h*\mu)$. 易见 $\tilde{g}(X)=\int_{\mathbb{R}^d}\mathbb{E}[h(x-X)]\mathrm{d}\mathscr{L}(X)(x)$, 由第一章相关讨论可知, 当 $\tilde{g}$ 在一点 Frechet 可微时, 为了确定该点处的 Frechet 导数, 只需计算所有的 Gateaux 导数, 所以先计算 X 处沿 Y 方向的 Gateaux 导数,

$$\begin{aligned}
&\tilde{g}(X+\epsilon Y)-\tilde{g}(X) \\
=&\int_{\mathbb{R}^d}\mathbb{E}[h(x-(X+\epsilon Y))]\mathrm{d}\mathscr{L}(X+\epsilon Y)(x)-\int_{\mathbb{R}^d}\mathbb{E}[h(x-X)]\mathrm{d}\mathscr{L}(X)(x) \\
=&\left(\int_{\mathbb{R}^d}\mathbb{E}[h(x-X-\epsilon Y)]\mathrm{d}\mathscr{L}(X+\epsilon Y)(x)-\int_{\mathbb{R}^d}\mathbb{E}[h(x-X)]\mathrm{d}\mathscr{L}(X+\epsilon Y)(x)\right) \\
&\quad+\left(\int_{\mathbb{R}^d}\mathbb{E}[h(x-X)]\mathrm{d}\mathscr{L}(X+\epsilon Y)(x)-\int_{\mathbb{R}^d}\mathbb{E}[h(x-X)]\mathrm{d}\mathscr{L}(X)(x)\right) \\
\triangleq& I_1+I_2,
\end{aligned} \tag{3.8}$$

对于第二项, 结合 Fubini 定理交换积分次序可得

$$\lim_{\epsilon\to 0}\frac{I_2}{\epsilon}=\int_{\mathbb{R}^d}\mathbb{E}[\nabla h(x-X)\cdot Y]\mathrm{d}\mathscr{L}(X)(x),$$

而对于第一项, 取 (X,Y) 的一个独立复制 $(\hat{X},\hat{Y})$, 则

$$\begin{aligned}I_1 &= \mathbb{E}[h((\hat{X}+\epsilon\hat{Y})-(X+\epsilon Y))]-\mathbb{E}[h(\hat{X}+\epsilon\hat{Y}-X)]\\ &= \mathbb{E}[h((\hat{X}+\epsilon\hat{Y})-(X+\epsilon Y))]-\mathbb{E}[h(\hat{X}-X)]\\ &\quad+\mathbb{E}[h(\hat{X}-X)]-\mathbb{E}[h(\hat{X}+\epsilon\hat{Y}-X)],\end{aligned} \tag{3.9}$$

所以

$$\lim_{\epsilon\to 0}\frac{I_2}{\epsilon}=\mathbb{E}[\nabla h(\hat{X}-X)\cdot(\hat{Y}-Y)]-\mathbb{E}[\nabla h(\hat{X}-X)\cdot(\hat{Y})]=-\mathbb{E}[\nabla h(\hat{X}-X)\cdot Y].$$

综上可得,$\partial_\mu g(\mu)=(\nabla(h+\bar{h}))*\mu$, 其中 $\bar{h}(x):=h(-x)$.

例 3.3　考虑连续函数 $v:\mathbb{R}^d\times\to\mathscr{P}_2$ 满足:

(1) $\forall\mu\in\mathscr{P}_2$,$v(x,\mu)$ 关于 x 可微, 记 $\nabla v(x,\mu)$ 表示 v 在 (x,μ) 处关于 x 的梯度. 任意有界集 $K\subset\mathscr{P}_2$, 存在常数 C_1, 使得任意 $\mu\in K,|\nabla v(x,\mu)|\le C_1(1+|x|)$.

(2) 任意 $\mu\in\mathscr{P}_2,x\in\mathbb{R}^d$, 使得 $(x,y)\mapsto$ 可测, 且 $\partial_\mu v(x,\mu)(y)$ 关于 y 连续, 并且对任意有界集 $K\subset\mathscr{P}_2$, 存在常数 C_2, 使得任意 $\mu\in K$,$\partial_\mu v(x,\mu)(y)\le C_2(1+|y|)$.

令 $u(\mu)=\int_{\mathbb{R}^d}v(x,\mu)\mathrm{d}\mu(x)$, 计算 u 的 Lions 导数, 计算中由于涉及到多个随机变量, 因此在合适的时候应该回到函数的定义以免造成混淆. 首先注意到 u 的提升 $\tilde{u}(X)=\mathbb{E}[v(X,\mathscr{L}(X))]$,

$$\begin{aligned}&\tilde{u}(X+Y)-\tilde{u}(X)\\ =&\mathbb{E}[v(X+Y,\mathscr{L}(X+Y))]-\mathbb{E}[v(X,\mathscr{L}(X))]\\ =&\mathbb{E}[v(X+Y,\mathscr{L}(X+Y))]-\mathbb{E}[v(X,\mathscr{L}(X+Y))]\\ &+\mathbb{E}[v(X,\mathscr{L}(X+Y))]-\mathbb{E}[v(X,\mathscr{L}(X))]\\ \triangleq& I_1+I_2,\end{aligned} \tag{3.10}$$

由中值定理可得,

$$I_1=\mathbb{E}[\int_0^1\nabla v(X+tY,\mathscr{L}(X+Y))\cdot Y\mathrm{d}t)],$$

而

$$\begin{aligned}I_2&=\int_{\mathbb{R}^d}(v(x,\mathscr{L}(X+Y))-v(x,\mathscr{L}(X)))\mathrm{d}\mathscr{L}(X)(x)\\ &=\int_{\mathbb{R}^d}\int_0^1\mathbb{E}[\partial_\mu v(x,\mathscr{L}(X+tY))(X+tY)\cdot Y]\mathrm{d}t\mathrm{d}\mathscr{L}(X)(x),\end{aligned}$$

所以可以合理推测 $\partial_\mu f(\mu)(\cdot)=\nabla v(\cdot,\mu)+\int_{\mathbb{R}^d}\partial_\mu f(y,\mu)(\cdot)\mathrm{d}\mu(y)$, 事实上根据对 v 的假设, 采用和例3.1中类似的讨论方法可知,f 在 μ 处的 L-导数确实为上式.

例 3.4 (续3.3)　设 $K:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ 为一可测映射,$\nabla_1 K,\nabla_2 K$ 分别表示对第 1 个和第 2 个自变量的梯度, 且都是至多线性增长的, 令 $v(x,\mu)=$

$\int_{\mathbb{R}^d} K(x,y)\mathrm{d}\mu(y)$, 则由上例结果可得

$$\partial_\mu f(\mu)(x) = \int_{\mathbb{R}^d} (\nabla_1 K(x,y) + \nabla_1 K(y,x))\mathrm{d}\mu(y).$$

3.2　链式法则和 Itô 公式

给定一个完备的带流概率空间 $(\Omega, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ 和其上的一个 m 维的 $\mathscr{F}_t-$ 布朗运动 $(B_s)_{s\geq 0}$, 设 $\mathscr{F}_t$ 是右连续的, 即 $\mathscr{F}_{t+} := \cap_{s>t}\mathscr{F}_s = \mathscr{F}_t$,$(b_t)_{t\geq 0}, (\sigma_t)_{t\geq 0}$ 分别为 Ω 上 $\mathbb{R}^d$ 和 $\mathbb{R}^{d\times m}$ 值的循序可测过程,$X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; \mathbb{R}^d)$, 且对任意 $T > 0$ 都有

$$\mathbb{E}[\int_0^T (|b_t|^2 + \|\sigma_t\|_{HS}^4)\mathrm{d}t] < \infty,$$

其中 $\|\sigma_s\|_{HS} := \sqrt{\mathrm{tr}(a_s)} := \sqrt{\mathrm{tr}(\sigma_s\sigma_s^*)}$. 令

$$X_t = X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s, \quad t \geq 0. \tag{3.11}$$

对于式 (3.10) 给出的 Ito 过程和任一 $\mathbb{R}^d$ 上的二次连续可微函数 f, 都有 Ito 公式

$$f(X_t) = f(X_0) + \int_0^t \nabla f(X_s)\cdot b_s \mathrm{d}s + \int_0^t \nabla f(X_s)\sigma_s \mathrm{d}B_s + \frac{1}{2}\int_0^t \mathrm{tr}(a_s \nabla^2 f(X_s))\mathrm{d}s \tag{3.12}$$

我们希望对式 (3.10) 给出的过程的时间边缘分布 $\boldsymbol{\mu} := (\mu_t)_{t\geq 0} := (\mathscr{L}(X_t))_{t\geq 0}$ 和一类性质足够好的 $f : \mathscr{P}_2 \to \mathbb{R}$ 也能建立类似的公式表示 $f(\mu_t)$.

在此之前,, 我们先考虑一类较简单的复合函数及求导. 首先注意到任意概率测度的凸组合仍为概率测度, 所以对函数 $f : \mathscr{P}_2 \to \mathbb{R}^d$ 可以定义它的经验投影

$$\begin{aligned} f^N : \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{N\text{个}} = (\mathbb{R}^d)^N \quad &\longrightarrow \quad \mathbb{R} \\ (x^1, \cdots, x^N) \qquad\qquad &\longmapsto \quad f(\frac{1}{N}\sum_{l=1}^N \delta_{x^i}), \end{aligned} \tag{3.13}$$

其中 $N \in \mathbb{N}$,$\delta_x(A) = I_A(x), x \in \mathbb{R}^d, A \in \mathscr{B}(\mathbb{R}^d)$ 表示 x 处的 Dirac 质量.

对于经验投影的可微性和求导公式有以下命题.

定理 3.5　设 $f : \mathscr{P}_2 \to \mathbb{R}$ 是连续 L-可微的, 则经验投影 f^N 是 $(\mathbb{R}^d))^N$ 上的可微函数, 且

$$\partial_{x^i} f^N(x^1, \cdots, x^N) = \frac{1}{N}\partial_\mu f(\frac{1}{N}\sum_{l=1}^N \delta_{x^i})(x^i), i = 1, \cdots, N.$$

注　此处的 ∂_{x^i} 表示沿 x^i 的方向导数算子.

证明　设 $\theta : \Omega \to \{1, \cdots, N\}$ 为均匀分布. 则给定 $\boldsymbol{x} = (x^1, \cdots, x^N) \in (\mathbb{R}^d)^N$, 和常值随机变量 $X_i^{\boldsymbol{x}} = x^i, i = 1, \cdots, N$,$X_\theta^{\boldsymbol{x}}$ 是一个取值于 $\{x^1, \cdots, x^N\}$ 的随机变

量, 且对任意 $A \in \mathscr{B}(\mathbb{R}^d)$,

$$\begin{aligned}\mathbb{P}(X_\theta^{\boldsymbol{x}} \in A) &= \sum_{l=1}^N \mathbb{P}((X_\theta^{\boldsymbol{x}} \in A, \theta = l) \\ &= \sum_{l=1}^N \mathbb{P}(X_\theta^{\boldsymbol{x}} \in A | \theta = l)\mathbb{P}(\theta = l) \\ &= \frac{1}{N}\sum_{l=1}^N \mathbb{P}(X_l^{\boldsymbol{x}} \in A) \\ &= \frac{1}{N}\sum_{l=1}^N \delta_{x^i}(A),\end{aligned} \tag{3.14}$$

也即 $\mathscr{L}(X_\theta^{\boldsymbol{x}}) = \frac{1}{N}\sum_{l=1}^N \delta_{x^i}$. 所以有定义,

$$\begin{aligned}f^N(\boldsymbol{x}+\boldsymbol{h}) &= f(\frac{1}{N}\sum_{l=1}^N \delta_{x^l+h^l}) \\ &= \tilde{f}(X_\theta^{\boldsymbol{x}} + X_\theta^{\boldsymbol{h}}) \\ &= \tilde{f}(X_\theta^{\boldsymbol{x}}) + \mathbb{E}[D\tilde{f}(X_\theta^{\boldsymbol{x}}) \cdot X_\theta^{\boldsymbol{h}}] + o(\left\|X_\theta^{\boldsymbol{h}}\right\|_{L^2}) \\ &= f^N(\boldsymbol{x}) + \mathbb{E}[\partial_\mu f(\frac{1}{N}\sum_{l=1}^N \delta_{x^i})(X_\theta^{\boldsymbol{x}}) \cdot X_\theta^{\boldsymbol{h}}] + o(\left\|X_\theta^{\boldsymbol{h}}\right\|_{L^2}),\end{aligned} \tag{3.15}$$

注意到 $\partial_\mu f(\frac{1}{N}\sum_{l=1}^N \delta_{x^i})(X_\theta^{\boldsymbol{x}}) \cdot X_\theta^{\boldsymbol{h}}$ 是只取有限个值的随机变量, 所以其期望是容易计算的, 计算后即可得到定理中的导数公式. ■

3.2.1　完全 C^2 正则性

由定理3.3, 当 $f : \mathscr{P}_2 \to \mathbb{R}$ 是连续 L-可微时, 对任一 $\mu \in \mathscr{P}_2$ 存在一个可测映射 $\partial_\mu f(\mu)(\cdot) : \mathbb{R}^d \to \mathbb{R}^d$, 使得当 $\mathscr{L}(X) = \mu, D\tilde{f}(X) = \partial_\mu f(\mu)(X)$, 且该可测映射在 μ—a.e. 意义下是唯一的. 我们希望 $(\mu, x) \mapsto \partial_\mu f(\mu)(x)$ 是一个从 $\mathscr{P}_2 \times \mathbb{R}^d$ 上的性质充分好的函数, 为此我们引入以下假设.

完全 C^2 正则性 (Full C^2 Regularity) 假设

(A1) $f : \mathscr{P}_2 \to \mathbb{R}$ 是连续 L-可微的, 即 $\tilde{f}$ 是连续 Frechet 可微的, 且存在合适版本的一阶 L-导数使得映射 $(\mu, x) \mapsto \partial_\mu f(\mu)(x)$ 是 $\mathscr{P}_2 \times \mathbb{R}^d$ 到 $\mathbb{R}^d$ 的连续映射;

(A2) 对于 (A1) 中的一阶 L-导数版本, 给定任一 $\mu \in \mathscr{P}_2$, 映射 $x \mapsto \partial_\mu f(\mu)(x)$ 是可微的, 记 $\partial_\mu f(\mu)(x)$ 关于空间变量 x 的微分为 $\partial_x\partial_\mu f(\mu)(x)$, 映射 $(\mu, x) \mapsto \partial_x\partial_\mu f(\mu)(x)$ 是 $\mathscr{P}_2 \times \mathbb{R}^d$ 到 $\mathbb{R}^{d\times d}$ 的连续映射;

(A3) 对于 (A1) 中的一阶 L-导数版本, 给定任意 $x \in \mathbb{R}^d$, 映射 $\mu \mapsto \partial_\mu f(\mu)(x)$ 的各分量映射是 L-可微的, 也称映射 $\mu \mapsto \partial_\mu f(\mu)(x)$ 是 L-可微的, 记为 $\partial_\mu^2 f(\mu)(x, \cdot) :== \partial_\mu^2 f(\mu)(x)(\cdot)$, 且映射 $(\mu, x, y) \mapsto \partial_\mu^2 f(\mu)(x, y)$ 是 $\mathscr{P}_2 \times \mathbb{R}^d \times \mathbb{R}^d$ 上的连续映射.

■

注

(1) 粗略地讲, 完全 C^2 正则性对应于多元微积分中的 C^2 可微概念, 即要求二阶导数存在并连续;

(2) 完全 C^2 正则性假设中满足条件的映射 $\partial_\mu f(\mu)(x), \partial_x\partial_\mu f(\mu)(x), \partial_\mu^2 f(\mu)(x,y)$ 是存在且唯一的, 具体证明参见 (Carmona et al., 2018) 的注 5.82;

(3) 与多元微积分中的梯度,Jacobi 矩阵等保持一致, 对 $\mu \in \mathscr{P}_2(\mathbb{R}^d)$,$x, y \in \mathbb{R}^d$, 将 $\partial_x\partial_\mu f(\mu)(x)$ 表示为 d 维方阵, 第 i 行, 第 j 列元素为

$$(\partial_x\partial_\mu f(\mu)(x))_{i,j} = \partial_{x_j}(\partial_\mu f(\mu))^{(i)}(x),$$

以及将 $\partial_\mu^2 f(\mu)(x,y)$ 表示为 d 维方阵, 第 i 行, 第 j 列元素为

$$(\partial_\mu^2 f(\mu)(x,y))_{i,j} = (\partial_\mu((\partial_\mu f(\mu))(x))^{(i)}(y))^{(j)}.$$

在将 $\mathbb{R}^d$ 中的元素 z 看作列向量的情况下,$\partial_x\partial_\mu f(\mu)(x)z, \partial_\mu^2 f(\mu)(x,y)z$ 均为由通常矩阵乘法的 $d\times 1$ 列向量; 而对于 d 维方阵 A,$\langle \partial_x\partial_\mu f(\mu)(x), A\rangle = \partial_x\partial_\mu f(\mu)(x)\cdot A = \mathrm{tr}(\partial_x\partial_\mu f(\mu)(x)A^T)$ 表示矩阵内积,$\partial_\mu^2 f(\mu)(x,y)\cdot A$ 类似.

(4) 在无特殊说明下, 无前缀的 C^2 正则性是指完全 C^2 正则性.

(5) 若函数 $f : \mathscr{P}_2 \to \mathbb{R}^d$ 满足完全 C^2 正则性假设中的 (A1),(A2) 和 (A3), 则称 f 是 (完全)C^2 正则的或 f 具有 (完全)C^2 正则性.

对于 C^2 正则的函数 f, 在此陈述一些相关性质.

首先是定理3.5中经验投影的可微性的加强, 有如下定理.

定理 3.6　设 $f : \mathscr{P}_2 \to \mathbb{R}$ 是完全 C^2 正则的. 则对任意 $N \in \mathbb{N}$, 经验投影 $f^N : (\mathbb{R}^d)^N \to \mathbb{R}$ 是经典意义下 C^2 的, 且对任意 $\boldsymbol{x} = (x^1, \cdots, x^N) \in (\mathbb{R}^d)^N$, 和 $1 \le i, j \le N$,

$$\partial^2_{x^ix^j} f^N(\boldsymbol{x}) = \frac{1}{N}\partial_x\partial_\mu f(\frac{1}{N}\sum_{l=1}^N \delta_{x^l})(x_i)I_{i=j} + \frac{1}{N^2}\partial_\mu^2 f(\frac{1}{N}\sum_{l=1}^N \delta_{x^l})(x^i, x^j).$$

在实变函数的情形下, 一个函数与一个性质充分好的函数做卷积后会继承部分好的性质, 对于测度变量的函数也可以有类似操作, 但这里不是通过卷积来平滑, 而是通过拉回映射来实现, 具体陈述如下.

引理 3.7　设 $f : \mathscr{P}_2 \to \mathbb{R}$ 是 C^2 正则的,$\rho : \mathbb{R}^d \to \mathbb{R}^d$ 为光滑紧支映射. 则函数

$$\begin{aligned} f^*\rho : \mathscr{P}_2 &\to \mathbb{R} \\ \mu &\mapsto f(\mu\circ\rho) \end{aligned} \tag{3.16}$$

是完全 C^2 正则的, 且 $f^*\rho$ 及其一阶和二阶导数都是有界和一致连续的.

证明梗概　首先 $f^*\rho$ 的提升为 $\widetilde{f^*\rho}(X)=\tilde{f}(\rho\circ X), X\in L^2(\Omega;\mathbb{R}^d)$, 这是因为若 $\mathscr{L}(X)$, 则 $\mathscr{L}(\rho\circ X)=\mu\circ\rho^{-1}$. 考虑 $\mathscr{L}(X)=\mu\in\mathscr{P}_2$,

$$\begin{aligned}
&\widetilde{f^*\rho}(X+Y)-\widetilde{f^*\rho}(X)\\
=&\tilde{f}(\rho\circ(X+Y))-\tilde{f}(\rho\circ X)\\
=&\mathbb{E}[D\tilde{f}(\rho\circ X)\cdot(\rho\circ(X+Y)-\rho\circ X)]+o(\|\rho\circ(X+Y)-\rho\circ X\|_{L^2})\\
=&\mathbb{E}[D\tilde{f}(\rho\circ X)\cdot((D\rho\circ X)Y+o(|Y|))]+o(\|\rho\circ(X+Y)-\rho\circ X\|_{L^2})\\
=&\mathbb{E}[D\tilde{f}(\rho\circ X)\cdot((D\rho\circ X)Y)]+o(\|Y\|_{L^2})\\
=&\mathbb{E}[(D\rho\circ X)D\tilde{f}(\rho\circ X)\cdot Y]+o(\|Y\|_{L^2}),
\end{aligned}\tag{3.17}$$

其中第四个等号是因为紧支光滑映射的任意阶微分仍是紧支光滑的, 特别还是有界的, 所以可以把小量统一吸收到 $o(\|Y\|_{L^2}$. 从而

$$\partial_\mu(f^*\rho)(\mu)(x)=D\rho(x)\partial_\mu f(\mu\circ\rho^{-1})(\rho(x))=(\sum_{l=1}^d(\partial_\mu f(\mu\circ\rho^{-1})(\rho(x)))^{(l)}\frac{\partial\rho_i}{\partial x_l}(x))_{1\le i\le d}$$

并将其视为一个 $d\times 1$ 的列向量.

二阶可微性的验证类似, 需要注意的是对测度变量的二阶导数计算中使用公式3.4会更方便. 最后给出 $f^*\rho$ 的导数公式,

$$\begin{aligned}
&\partial_\mu^2(f^*\rho)(\mu)(x,y)=D\rho(y)D\rho(x)\partial_\mu^2 f(\mu\circ\rho^{-1})(\rho(x),\rho(y)),\\
&\partial_x\partial_\mu(f^*\rho)(\mu)(x)=D^2\rho(x)\partial_\mu f(\mu\circ\rho^{-1})(\rho(x))+D\rho(x)\partial_x\partial_\mu f(\mu\circ\rho^{-1})(\rho(x)).
\end{aligned}\tag{3.18}$$

■

3.2.2　Ito 公式

对于3.11定义的 Ito 过程和性质足够好的函数 $f:\mathscr{P}_2\to\mathbb{R}$, 下面的定理给出了其时间边缘分布满足的 Ito 公式.

定理 3.8　设 $f:\mathscr{P}_2\to\mathbb{R}$ 是 C^2 正则的, 且对任意 $\mathscr{P}$ 的紧子集 K, 存在常数 $C=C(K)$, 使得任意 $\mu\in K$,

$$\int_{\mathbb{R}^d}|\partial_x\partial_\mu f(\mu)(x)|^2\mathrm{d}\mu(x)<C.$$

则对任意 $t\ge 0$,

$$f(\mu_t)=f(\mu_0)+\int_0^t\mathbb{E}[\partial_\mu f(\mu_s)(X_s)\cdot b_s]\mathrm{d}s+\frac{1}{2}\int_0^t\mathbb{E}[\partial_x\partial_\mu f(\mu_s)(X_s)\cdot a_s]\mathrm{d}s.\tag{3.19}$$

证明　证明主要分为两步.

第一步:

设 f, 其一阶导数和其二阶导数都是有界且一致连续的. 且过程 $(b_s),(\sigma_s)$ 是有界的.

设 $((b_s^l,\sigma_s^l,B_s^l),X_0^l)_(l\in\mathbb{N})$ 是 $((b_s,\sigma_s,B_s),X_0)$ 的一列独立复制, 则

$$X_t^l=X_0^l+\int_0^t b_s^l\mathrm{d}s+\int_0^t\sigma_s^l\mathrm{d}B_s^l,\quad t\geq 0,l\geq 1,$$

则 $(X_t),(X_t^1),(X_t^2),\cdots$ 独立同分布.

考虑 f 到 $(\mathbb{R}^d)^N$ 的经验投影, 由定理3.6可知 f^N 是 C^2 的, 所以可以应用 Ito 公式3.12, 简记 $\boldsymbol{X}_t=(X_t^1,\cdots,X_t^N)$ 以及 $\bar{\mu}_t^N=\frac{1}{N}\sum_{l=1}^N\delta_{X_t^l}$; 此外, 对两个独立的布朗运动 $B^i=(B_s^i),B^j=(B_s^j)$ 是正交的, 即 $\langle B^i,B^j\rangle=0$, 所以

$$\begin{aligned}
&f^N(\boldsymbol{X}_t)-f^N(\boldsymbol{X}_0)\\
=&\sum_{l=1}^N\int_0^t\partial_{x^i}f^N(\boldsymbol{X}_s)\cdot b_s^l\mathrm{d}s+\sum_{l=1}^N\int_0^t\partial_{x^i}f^N(\boldsymbol{X}_s)\cdot(\sigma_s^l\mathrm{d}B_s^l)\\
&+\sum_{l=1}^N\partial^2_{x^lx^l}f^N(\boldsymbol{X}_s)\cdot a_s^l\mathrm{d}s\\
=&\frac{1}{N}\sum_{l=1}^N\int_0^t\partial_\mu f(\bar{\mu}_t^N)(X_s^l)\cdot b_s^l\mathrm{d}s+\sum_{l=1}^N\int_0^t\partial_\mu f(\bar{\mu}_t^N)(X_s^l)\cdot(\sigma_s^l\mathrm{d}B_s^l)\\
&+\frac{1}{2N}\sum_{l=1}^N\int_0^t\partial_x\partial_\mu f(\bar{\mu}_s^N)(X_s^l)\cdot a_s^l\mathrm{d}s\\
&+\frac{1}{2N^2}\int_0^t\partial_\mu^2 f(\bar{\mu}_s^N)(X_s^l,X_s^l)\cdot a_s^l\mathrm{d}s,
\end{aligned}\tag{3.20}$$

注意到 $\bar{\mu}_t^N$ 实际上是一个随机概率测度, 所以上式应理解为在一个 $\mathbb{P}-$ 零测集外逐点意义下成立.

对上式左右两边同时取期望, 并结合独立同分布的条件,

$$\begin{aligned}
&\mathbb{E}[f(\bar{\mu}_t^N)]-\mathbb{E}[f(\bar{\mu}_0^N)]\\
=&\int_0^t\mathbb{E}[\partial_\mu f(\bar{\mu}_t^N)(X_s)\cdot b_s]\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t\mathbb{E}[\partial_x\partial_\mu f(\bar{\mu}_s^N)(X_s)\cdot a_s]\mathrm{d}s\\
&\quad+\frac{1}{2N}\int_0^t\mathbb{E}[\partial_\mu^2 f(\bar{\mu}_s^N)(X_s,X_s)\cdot a_s]\mathrm{d}s\\
\triangleq& I_1+I_2+I_3,
\end{aligned}\tag{3.21}$$

根据有界性的假设, 当 $N\to\infty$ 时,$I_3\to 0$; 同样是由于有界性, 应用 (Fournier et al., 2015) 的定理 1 以及强大数定律, 可得任意 $s\in[0.t]$,$W_2(\mu_s^N,\mu_s)$ 以概率 1 收敛到 0. 由 f 的 C^2 正则性, 具体地说是 f 和一阶, 二阶导数的连续性, 结合有界收敛定理, 可得当 $N\to\infty$,$\mathbb{E}[f(\bar{\mu}_s^N)]\to f(\mu)$.$I_1,I_2$ 的收敛性类似.

第二步:

去掉第一步中对 f 及其导数的有界性和一致连续性假设. 取一列紧支光滑映射 $\{\rho_n:\mathbb{R}^d\to\mathbb{R}^d\}$ 使得在紧集上,$(\rho_n(x),D\rho_n(x),D^2\rho_n(x))$ 一致收敛到 $(x,I_d,0)$.

不妨设存在常数 C 使得任意 $n\geq 1, x\in\mathbb{R}^d$,

$$
\begin{aligned}
&|\rho_n(x)|\leq C|x|,\\
&|D\rho_n(x)|\leq C,\\
&|D^2\rho_n(x)|\leq C.
\end{aligned}
$$

以及当 $\rho_n(x)I_{B(0,n)}(x)=x$. 从而对任意 $\mu\in\mathscr{P}_2$ 以及以 μ 为分布的随机变量 X, 都有

$$
\begin{aligned}
W_2(\mu\circ\rho_n^{-1},\mu)^2 &\leq \mathbb{E}[|\rho_n(X)-X|^2]\\
&=\mathbb{E}[|\rho_n(X)-X|^2 I_{\{|X|\geq n\}}]\\
&\leq (C+1)^2\mathbb{E}[|X|^2 I_{\{|X|\geq n\}}],
\end{aligned}
\tag{3.22}
$$

由于 X 二阶矩有限, 所以当 $n\to\infty$,$W_2(\mu\circ\rho_n^{-1},\mu)\to 0$. 结合连续性可得以下收敛在 $\mathbb{P}-$a.s. 意义下成立:

$$
\begin{aligned}
&f^*\rho_n(\mu)\to mu,\\
&\partial_\mu(f^*\rho_n)(\mu)(X)\to\partial_\mu f(\mu)(X),\\
&\partial_x\partial_\mu(f^*\rho_n)(\mu)(X)\to\partial_x\partial_\mu f(\mu)(X).
\end{aligned}
\tag{3.23}
$$

由于

$$
\sup_{n\geq 1}\mathbb{E}[|\partial_\mu(f^*\rho_n)(\mu)(X)|^2+|\partial_x\partial_\mu(f^*\rho_n)(\mu)(X)|^2]<\infty,
$$

结合 (Bogachev, 2007) 的定理 4.5.9, 可得 $\{\partial_\mu(f^*\rho_n)(\mu)(X)\},\{\partial_x\partial_\mu(f^*\rho_n)(\mu)(X)\}$ 都是一致可积的, 所以可以应用 Vitalli 收敛定理得到对任意 $t\geq 0$ 和任意 $s\in[0,t]$

$$
\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}[\partial_\mu(f^*\rho_n)(\mu)(X)\cdot b_s]=\mathbb{E}[\partial_\mu f(\mu)(X)\cdot b_s],\\
&\lim_{n\to\infty}\mathbb{E}[\partial_x\partial_\mu(f^*\rho_n)(\mu)(X)\cdot a_s]=\mathbb{E}[\partial_x\partial_\mu f(\mu)(X)\cdot a_s].
\end{aligned}
\tag{3.24}
$$

对每个 n 应用 Ito 公式, 可得

$$
\begin{aligned}
&(f^*\rho_n)(\mu_t)-(f^*\rho_n)(\mu_0)\\
&=\int_0^t\mathbb{E}[\partial_\mu(f^*\rho_n)(\mu_s)(X_s)\cdot b_s]\mathrm{d}s+\frac{1}{2}\int_0^t\mathbb{E}[\partial_x\partial_\mu(f^*\rho_n)(\mu)(X_S)\cdot a_s]\mathrm{d}s\\
&\triangleq\int_0^t g_n(s)\mathrm{d}s+\frac{1}{2}\int_0^t h_n(s)\mathrm{d}s,
\end{aligned}
\tag{3.25}
$$

其中左边收敛到 $f(\mu_t)-f(\mu_0)$, 右边则有被积函数 $(g_n(s),h_n(s))$ 逐点收敛到 $(g(s),h(s))=(\mathbb{E}[\partial_\mu f(\mu_s)(X_s)\cdot b_s],\mathbb{E}[\partial_x\partial_\mu f(\mu_s)(X_s)\cdot a_s])$, 下面只需说明 $\{g_n\},\{h_n\}$ 是一致可积的. 事实上, 由 $\{b_s\},\{\sigma_s\}$ 的矩条件可知映射 $s\mapsto\mu_s$ 是连续的, 所以 $\{\mu_s\}_{0\leq s\leq t}$ 作为紧集 $[0,t]$ 的连续映射下的像仍是紧的, 所以

$$
\sup_{n\geq 1}\sup_{s\in[0,t]}(|h_n(s)|^2+|g_n(x)|^2)<\infty,
$$

从而一致可积性成立, 也就完成了证明.

■

在这一节的最后, 要说明的是虽然完全 C^2 正则性假设中要求 $\partial_\mu^2 f(\mu)(x,y)$ 的存在性和连续性, 但在 Ito 公式中并未出现这一项, 实际上对于它的要求确实可以减弱, 从而将完全 C^2 正则性假设的要求降低到部分 C^2 正则性 ((Partial C^2 Regularity) 假设, 同时其他设定不变, 保持 Ito 公式仍然成立; 此外, 还可以将 $f:\mathscr{P}_2 \to \mathbb{R}$ 扩展到 $f:[0,T]\times\mathbb{R}^d\times\mathscr{P}_2\to\mathbb{R}$, 给出

$$f(t,X_t,\mathscr{L}(X_t)-f(0,X_0,\mathscr{L}(X_0)))$$

的联合 Ito 公式. 具体可见 (Carmona et al., 2018) 的 5.4.5 节和 5.5.5 节相关内容

第 4 章　外在导数, 内蕴导数和线性泛函导数

本章内容主要参考引用自 (Ren et al., 2021),(王凤雨 等, 2020),(Bao et al., 2021) 和 (Wang et al., 2024).

Lions 导数可以看作是 Frechet 导数的延伸, 这一章则主要介绍 Gateaux 导数或方向导数的概念在测度变量的函数上的应用. 参考文献是在 Riemann 流形上对一般情况定义了几种导数, 为简单起见, 此处只考虑欧氏空间 $\mathbb{R}^d$ 的情形.$\mathscr{M}$ 表示 $\mathbb{R}^d$ 上有限测度全体, 记 $\mathscr{M}_p = \{\mu \in \mathscr{M} : \mu(|\cdot|^p) < \infty\}$ 表示 $\mathbb{R}^d$ 上 p 阶矩有限测度的全体,$p \in [0, \infty)$, 配备通常的加法, 减法和数乘. 对 $p > 0$, 定义 Wasserstein 距离

$$W_p(\mu, \nu) = (\mu(\mathbb{R}^d) \wedge \nu(\mathbb{R}^d))W_p(\frac{\mu}{\mu(\mathbb{R}^d)}, \frac{\nu}{\nu(\mathbb{R}^d)}) + |\mu(\mathbb{R}^d) - \nu(\mathbb{R}^d)|,$$

当 $p = 0$ 时, 定义 Prokhorov 度量

$$W_0(\mu, \nu) = \inf\{\epsilon > 0 : \mu(A) \le \epsilon + \nu(A^\epsilon), \nu(A) \le \epsilon + \mu(A^\epsilon), A \in \mathscr{B}(\mathbb{R}^d)\},$$

其中 $A^\epsilon = \{x \in \mathbb{R}^d : d(x, A) < \epsilon\}$.

对于任意 $p \in [0, \infty)$,$(\mathscr{M}_p, W_p)$ 是一个可分完备度量空间.

4.1　定义

我们首先给出外在导数的定义.

定义 4.1　设 $p \ge 0$. 称 $f : \mathscr{M}_p \to \mathbb{R}$ 在 $\mathscr{M}_p$ 上是外在可导的, 是指: 任意 $(x, \eta) \in \mathbb{R}^d \times \mathscr{M}_p$ 极限

$$D^E f(\eta)(x) := \lim_{\epsilon \downarrow 0} \frac{f(\eta + \epsilon\delta_x) - f(\eta)}{\epsilon}$$

存在, 其中 δ_x 表示 $x \in \mathbb{R}^d$ 处的 Diarc 质量.$D^E f$ 称为 f 的外在导数.

注　在内蕴导数的定义的基础上, 定义若干常用的函数类.

称函数 $f : \mathscr{M}_p \to \mathbb{R} \in C^{E,1}(\mathscr{M}_p)$, 是指:$f$ 在 $\mathscr{M}_p$ 上是外在可导的, 且映射 $(x, \eta) \mapsto D^E f(\eta)(x)$ 是 $\mathbb{R}^d \times \mathscr{M}_p$ 上的连续函数.

称函数 $f : \mathscr{M}_p \to \mathbb{R} \in C_K^{E,1}(\mathscr{M}_p)$, 是指:$f \in C^{E,1}(\mathscr{M}_p)$, 且对任意紧集 $K \subset \mathscr{M}_p$, 存在常数 C, 使得任意 $\eta \in K$,$|D^E f(\eta)(x)| \le C(1 + |x|^p)$.

称函数 $f : \mathscr{M}_p \to \mathbb{R} \in C^{E,1,1}(\mathscr{M}_p)$, 是指:$f \in C^{E,1}(\mathscr{M}_p)$,$D^E f(\eta)(x)$ 关于 x 可微, 且 $\nabla D^E f(\eta)(x)$ 是 $\mathbb{R}^d \times \mathscr{M}_p$ 上的连续函数.

此外,$C_b^{E,1}(\mathscr{M}_p), C_b^{E,1,1}(\mathscr{M}_p)$ 分别表示 $C^{E,1}(\mathscr{M}_p), C^{E,1,1}(\mathscr{M}_p)$ 和 $\mathscr{M}_p$ 上有界函数的交集.

容易看出, 外在导数的定义不能直接限制在 p 阶矩有限的的概率测度组成的集合 $\mathscr{P}_p \subset \mathscr{M}_p$ 上, 因此对以上定义稍作修改, 引入凸外在导数的概念.

定义 4.2　设 $p \geq 0$. 称函数 $f : \mathscr{P}_p \to \mathbb{R}$ 凸外在可导, 是指: 对任意 $(x, \mu) \in \mathbb{R}^d \times \mathscr{P}_p$, 极限

$$\tilde{D}^E f(\mu)(x) := \lim_{s \downarrow 0} \frac{f((1-s)\mu + s\delta_x)}{s}$$

存在,$\tilde{D}^E f$ 称为 f 的凸外在导数.

注　类比外在导数, 同样可以类似的函数空间.

称函数 $f : \mathscr{P}_p \to \mathbb{R} \in C^{E,1}(\mathscr{P}_p)$, 是指:$f$ 在 $\mathscr{P}_p$ 上是外在可导的, 且映射 $(x, \eta) \mapsto \tilde{D}^E f(\eta)(x)$ 是 $\mathbb{R}^d \times \mathscr{P}_p$ 上的连续函数.

称函数 $f : \mathscr{P}_p \to \mathbb{R} \in C_K^{E,1}(\mathscr{P}_p)$, 是指:$f \in C^{E,1}(\mathscr{P}_p)$, 且对任意紧集 $K \subset \mathscr{P}_p$, 存在常数 C, 使得任意 $\eta \in K$,$|\tilde{D}^E f(\eta)(x)| \leq C(1 + |x|^p)$.

称函数 $f : \mathscr{P}_p \to \mathbb{R} \in C^{E,1,1}(\mathscr{P}_p)$, 是指:$f \in C^{E,1}(\mathscr{P}_p)$,$\tilde{D}^E f(\eta)(x)$ 关于 x 可微, 且 $\nabla \tilde{D}^E f(\eta)(x)$ 是 $\mathbb{R}^d \times \mathscr{P}_p$ 上的连续函数.

此外,$C_b^{E,1}(\mathscr{P}_p), C_b^{E,1,1}(\mathscr{P}_p)$ 分别表示 $C^{E,1}(\mathscr{P}_p), C^{E,1,1}(\mathscr{P}_p)$ 和 $\mathscr{P}_p$ 上有界函数的交集.

虽然凸外在导数和外在导数定义上稍有差别, 但存在着紧密关联: 若函数 $f : \mathscr{M}_p \to \mathbb{R} \in C_b^{E,1}(\mathscr{M}_p)$, 则 $f|_{\mathscr{P}_p} \in C^{E,1}(\mathscr{P}_p)$, 且对任意 $\mu \in \mathscr{P}_p$,

$$\tilde{D}^E f|_{\mathscr{P}_p}(\mu) = D^E f(\mu) - \mu(D^E f(\mu)),$$

在这个意义下, 可以将凸外在导数看作外在导数的中心化.

给定向量场 $\phi : \mathbb{R}^d \to \mathbb{R}^d \in \mathscr{B}(\mathbb{R}^d \to \mathbb{R}^d)$, 定义

$$\phi_s(x) := x + s\phi(x), s \geq 0,$$

也即 $\phi_s = \mathrm{id}_{\mathbb{R}^d} + s\phi, s \geq 0$.

对于任意测度 $\eta \in \mathscr{M}_p, p \in (0, 2]$, 定义 η 处的切空间为 $L^2(\mathscr{B}(\mathbb{R}^d \to \mathbb{R}^d); \eta)$, 这是一个 Hilbert 空间, 可以应用 Riesz 表示定理, 将其上的连续线性泛函视为其中一个元素.

定义 4.3　设 $p \in [0, 2]$. 称函数 $f : \mathscr{M}_p \to \mathbb{R}$ 在 $\eta \in \mathscr{M}_p$ 内蕴可导, 是指: 对于任意 $\phi \in L^2(\mathscr{B}(\mathbb{R}^d \to \mathbb{R}^d); \eta)$, 极限

$$D_\phi^I f(\eta) := \lim_{s \downarrow 0} \frac{f(\eta \circ \phi_s^{-1}) - f(\eta)}{s}$$

存在, 且是 ϕ 的连续线性泛函. 此时存在唯一的 $D^I f(\eta) \in L^2(\mathscr{B}(\mathbb{R}^d \to \mathbb{R}^d); \eta)$, 使得

$$D_\phi^I f(\eta) = \langle D^I f(\eta), \phi \rangle_{L^2(\mathscr{B}(\mathbb{R}^d \to \mathbb{R}^d); \eta)} = \int_{\mathbb{R}^d} \langle D^I f(\eta), \phi \rangle \mathrm{d}\eta,$$

称 $D^I f(\eta)$ 为 f 在 η 处的内蕴导数, 若对任意 $\mu \in \mathscr{M}_p$,f 在 μ 处都是内蕴可导的, 则称 f 在 $\mathscr{M}_p$ 上内蕴可导或简称内蕴可导.

注　由于内蕴导数的定义不依赖于 $\mathscr{M}_p$ 的线性结构, 因此对 $\mathscr{P}_p$ 上的实值函数有完全相同的定义.

作为比较, 这里给出 Lions 导数的另一种定义. 在第二章中, 是将测度变量函数 $f:\mathscr{P}_2\to\mathbb{R}$ 提升为以随机变量为变量的函数 $\tilde{f}=f\circ\mathscr{L}:L^2(\Omega;\mathbb{R}^d)\to\mathbb{R}$, 然后用随机变量空间上的 Frechet 导数来刻画 f 的局部性态. 在考虑差 $\tilde{f}(X+Y)-\tilde{f}(X)$ 以及 $\|Y\|_{L^2}$ 逐渐趋于 0 时, 如果固定一个方向取 $Y=s\phi(X)$, 则

$$\begin{aligned}\tilde{f}(X+Y)-\tilde{f}(X)&=\tilde{f}(X+s\Phi(X))-\tilde{f}(X)\\&=f(\mathscr{L}(X)\circ(\mathrm{id}+s\phi)^{-1})-f(\mathscr{L}(X))\\&=f(\mathscr{L}(X)\circ\phi_s^{-1})-f(\mathscr{L}(X)),\end{aligned}$$

这就是内蕴导数的形式. 而 Lions 导数本质是 Frechet 导数, 应该考虑任意方向, 所以在令 Y 的二阶矩趋于 0 的过程不应该通过给一个固定的函数做伸缩来实现, 而应该让函数本身充分” 小”, 用严格的数学语言来叙述即如下定义.

定义 4.4　给定 $p\in[0,2]$. 称函数 $f:\mathscr{M}_p\to\mathbb{R}$ 在 $\eta\in\mathscr{M}_p$L-可导, 是指:f 是内蕴可导的, 且

$$\lim_{\|\phi\|_{L^p(\eta)}\downarrow 0}\frac{|f(\eta\circ(\mathrm{id}+\phi)^{-1})-f(\eta)-D^I_\phi f(\eta)|}{\|\phi\|_{L^p(\eta)}}=0,$$

若 f 在任意 $\eta\in\mathscr{M}_p$ 都 L-可导, 则称 f 是 L-可导的, 此时将 D^If 记为 D^Lf.

注　根据第二章的相关讨论, 在这个定义中, 要求 f 内蕴可导是自然的; 另外, 在这个 L-可导的定义下,L-可导显然强于内蕴可导, 而当 fL-可导时, 其 L-导数就是内蕴导数. 以上定义对 $\mathscr{M}_p$ 上的函数同样成立.

称 $f\in C^{L,1}(\mathscr{M}_p)$, 是指:$f$ 是 L-可导的, 且存在 η 版本的 $D^Lf(\eta)(\cdot)$ 使得 $(x,\eta)\mapsto D^Lf(\eta)(x)$ 是 $\mathbb{R}^d\times\mathscr{M}_p$ 上的连续映射; 若在此基础上,D^Lf 还是有界的, 则记 $f\in C_b^{L,1}(\mathscr{M}_p)$.

以上关于 L-导数的定义和函数类均可完全限制在 $\mathscr{P}_p$ 的情形下给出对应的定义和记号.

在第一章中, 对于 Banach 空间之间的映射如果是可微的并满足合适的条件, 则有中值定理成立. 对于 $\mathscr{M}_p$ 或 $\mathscr{P}_p$, 则可以用” 中值定理” 来定义一点处的导数, 从而一定程度上在形式上与 Banach 空间上的 Frechet 导数保持一致.

定义 4.5　设 $p\in[0,\infty)$. 给定 $f:\mathscr{M}_p\to\mathbb{R}$, 称可测函数 $D^Ff(\eta):\mathbb{R}^d\to\mathbb{R}$ 是 f 在 η 处的线性泛函导数, 是指: 对任意 $L>0$, 存在 C_L 使得

$$\sup_{\eta(|\cdot|^p)\le L}|D^Ff(\eta)(y)|\le C_L(1+|y|^p),y\in\mathbb{R}^d,$$

且对任意 $\mu,\nu\in\mathscr{M}_p$,

$$f(\mu)-f(\nu)=\int_0^1\int_{\mathbb{R}^d}D^Ff(\nu+t(\mu-\nu))(y)(\mu-\nu)(\mathrm{d}y)\mathrm{d}t.$$

注　由于只涉及到凸组合, 所以线性泛函导数的定义对 $\mathscr{P}_p$ 上的实值函数完全成立.

以上定义的几种导数并不是完全孤立的, 在合适的条件下, 它们之间可以相互表达. 将外在导数和凸外在导数看作一类, 内蕴导数和 L-导数看作一类, 线性泛函导数看作一类, 这三类导数之间都存在一定的联系.

由于 L-导数可以视为更强的内蕴导数, 所以这里只讨论外在导数和 L-导数的关系. 从定义的形式上来看, 外在导数的定义要比 L-导数的计算" 简单" 一些, 所以希望能找到它们间的关联, 从而通过计算外在导数来简化 L-导数的计算. 事实上, 有以下定理.

定理 4.1　设 $p \in [0,2]$. 若 $f \in C_b^{E,1,1}(\mathscr{M}_p)$, 则 $f \in C_b^{L,1}(\mathscr{M}_p)$, 且

$$D^L f(\eta) = \nabla[D^E f(\eta)], \eta \in \mathscr{M}_p$$

证明思路和细节参见 (王凤雨 等, 2020) 的定理 2.1 和 (Ren et al., 2021) 的定理 2.1(3).

实际上, 上述定理还说明在一定条件下, 测度变量的函数在一点处的 L-导数可以视为某标量函数的梯度; 反过来, 如果已知 L-导数, 则可以通过积分或中值定理计算, 在相差一个常数的意义下得到外在导数, 即

$$D^E f(\eta)(x) = D^E f(\eta)(0) + \int_0^1 D^L f(\eta)(x) \cdot x \mathrm{d}x.$$

而对于线性泛函导数, 在多数情况下, 如果只根据定义只能验证而非计算一个测度变量函数的线性泛函导数, 但以下定理说明对一大类函数, 外在导数若存在, 则外在导数就是线性泛函导数.

定理 4.2　若函数 $f \in C_K^{E,1}(\mathscr{M}_p)$, 则 f 的线性泛函导数存在, 且 $D^F f = D^E f$.

4.2　链式法则

由于在一定条件下, 三类导数可以相互表示, 所以这里只给出 L-导数的链式法则, 也是推论3.4的一般情形.

定理 4.3　设 $p \geq 1. f : \mathscr{P}_p \to \mathbb{R}$ 是连续函数,$(X_s)_{s\in[0,1]}$ 是一族 $\mathbb{R}^d$ 值的随机变量,X_s 在 L^p 意义下关于 s 连续, $\dot{X}_0 := \lim_{s\downarrow 0} \frac{X_s - X_0}{s}$ 在 $L^p(\Omega;\mathbb{R}^d)$ 中存在. 下列两个条件至少满足其一:

(1) $\mu_0 = \mathscr{L}(X_0)$ 无原子,f 是 L-可微的且 $D^L f(\mu_0)$ 存在一个连续版本以及常数 C 使得

$$|D^L f(\mu_0)(x)| \leq C(1 + |x|^{p-1}),$$

(2) f 在 μ_0 的一个邻域 O 内是 L-可微的,$D^L f$ 有一个版本使得 $\mathbb{R}^d \times O \ni$

$(x,\mu)\mapsto D^Lf(\mu)(x)$ 连续, 且存在常数 C 使得

$$|D^Lf(\mu_0)(x)|\le C(1+|x|^{p-1}).$$

则

$$\lim_{s\downarrow 0}\frac{f(\mathscr{L}(X_s))-f(\mu_0)}{s}=\mathbb{E}[\langle D^Lf(\mu_0)(X_0),\dot{X}_0\rangle]. \tag{4.1}$$

证明

(1) 当 μ_0 无原子时, 完全类比于第二章引理3.1的讨论可得, 对于任意 $n\in\mathbb{N}$, 存在可测映射 $S_n:\Omega\to\mathbb{R}^d$ 和 $T_n:\mathbb{R}^d\to\Omega$, 使得

$$\mathbb{P}(T_n\circ S_n=\mathrm{id}_\Omega)=\mu_0(S_n\circ T_n=\mathrm{id}_{\mathbb{R}^d})=1, \tag{4.2}$$

$$\mathbb{P}=\mu_0\circ T_n^{-1},\mu_0=\mathbb{P}\circ S_n^{-1}, \tag{4.3}$$

$$\|X_0-S_n\|_{L^\infty(\mathbb{P})}+\|\mathrm{id}_{\mathbb{R}^d}-X_0\circ T_n\|_{L^\infty(\mu_0)}\le\frac{1}{n}. \tag{4.4}$$

由 L-可导的定义4.4可知, 存在一个非增的函数 h 满足当 $s\to 0$,$h(s)\to 0$, 使得

$$\sup_{\|\phi\|_{L^p(\mu_0)}\le r}|f(\mu_0\circ(\mathrm{id}_{\mathbb{R}^d}+\phi)^{-1})-f(\mu_0)-D^L_\phi f(\mu_0)|\le rh(r),$$

由 $\mathscr{L}(X_s-X_0)\in\mathscr{P}_p$ 和式4.3可得 $\phi_{n,s}:=(X_s-X_0)\circ T_n\in L^p(\mu_0)$, 这是因为 $\left\|\phi_{n,s}\right\|_{L^p(\mu_0)}=\|X_s-X_0\|_{L^p(\mathbb{P})}$, 对任意的 $A\in\mathscr{B}(\mathbb{R}^d)$

$$\begin{aligned}\mathscr{L}(S_n+X_s-X_0)(A)&=\mathbb{P}((S_n+X_s-X_0)^{-1}(A))\\&=(\mu_0\circ T_n^{-1})((S_n+X_s-X_0)^{-1}(A))\\&=\mu_0(T_n^{-1}((S_n+X_s-X_0)^{-1}(A)))\\&=\mu_0(((S_n+X_s-X_0)\circ T_n)^{-1(A)})\\&=\mu_0((\mathrm{id}_{\mathbb{R}^d}+(X_s-X_0)\circ T_n)^{-1}(A))\\&=\mu_0((\mathrm{id}_{\mathbb{R}^d}+\phi_{n,s})^{-1}(A))\\&=(\mu_0\circ(\mathrm{id}_{\mathbb{R}^d}+\phi_{n,s})^{-1})(A).\end{aligned} \tag{4.5}$$

由条件可知, 存在 $\delta\in[0,1]$ 使得任意 $s\in(0,\delta)$,

$$\left\|\frac{X_s-X_0}{s}-\dot{X}_0\right\|_{L^p(\mathbb{P})}\le\left\|\dot{X}_0\right\|_{L^2(\mathbb{P})}\implies\|X_s-X_0\|_{L^2(\mathbb{P})}\le s(2\left\|\dot{X}_0\right\|_{L^2(\mathbb{P})}\vee 1)=cs,$$

另一方面,

$$\begin{aligned}D^L_{\phi_{n,s}}f(\mu_0)&=\langle D^Lf(\mu_0),\phi_{n,s}\rangle_{L^2(\mu_0)}\\&=\int_{\mathbb{R}^d}\langle D^Lf(\mu_0)(x),\phi_{n,s}(x)\rangle\mathrm{d}\mu_0(x)\\&=\int_{\mathbb{R}^d}\langle D^Lf(\mu_0)(x),\phi_{n,s}(x)\rangle\mathrm{d}(\mathbb{P}\circ S_n^{-1})(x)\\&=\mathbb{E}[\langle D^Lf(\mu_0)(S_n),\phi_{n,s}(S_n)\rangle]\\&=\mathbb{E}[\langle D^Lf(\mu_0)(S_n),X_s-X_0\rangle].\end{aligned} \tag{4.6}$$

综上可得,

$$\begin{aligned}
&|f(\mathscr{L}(S_n+X_s-X_0))-f(\mathscr{L}(X_0))-\mathbb{E}[\langle D^L f(\mu_0)(S_n),X_s-X_0\rangle]|\\
=&|f(\mu_0\circ(\mathrm{id}_{\mathbb{R}^d}+\phi_{n,s})^{-1})-f(\mu_0)-D^L_{\phi_{n,s}}f(\mu_0)|\\
\le&\|\phi_{n,s}\|_{L^p(\mu_0)}h(\|\phi_{n,s}\|_{L^p(\mu_0)})\\
=&\|X_s-X_0\|_{L^p(\mathbb{P})}h(\|X_s-X_0\|_{L^p(\mathbb{P})}),s\in[0,\frac{1}{c}]
\end{aligned}\tag{4.7}$$

由式4.4可知, 当 $n\to\infty$,S_n 几乎处处收敛到 X_0, 且由控制收敛定理加上 f 及其导数的连续性可得

$$|f(\mathscr{L}(X_s)-f(\mu_0)-\mathbb{E}[\langle D^L f(\mu_0)(X_0),X_s-X_0\rangle]|\le\|X_s-X_0\|_{L^p(\mathbb{P})}h(\|X_s-X_0\|_{L^p(\mathbb{P})}),$$

两边同时除以 s 再令 $s\to 0$ 即可.

(2) 若 μ_0 不是无原子的, 取一个独立于 (X_s) 且分布无原子的随机变量 X, 则随机变量 $X_0+\epsilon X+r(X_s-X_0)$ 的分布总是无原子的. 当 δ 充分小时, 应用中值定理可得,

$$\begin{aligned}
&f(\mathscr{L}(X_s+\epsilon X))-f(\mathscr{L}(X_0+sX))\\
=&\int_0^1\frac{\mathrm{d}}{\mathrm{d}\delta}|_{\delta=0}f(\mathscr{L}(X_0+\epsilon X+(r+\delta)(X_s-X_0)))\mathrm{d}r\\
=&\int_0^1\mathbb{E}[\langle D^L f(\mathscr{L}(X_0+\epsilon X+(r+\delta)(X_s-X_0)))(X_0+\epsilon X+(r+\delta)(X_s-X_0)),X_s-X_0\rangle],
\end{aligned}\tag{4.8}$$

令 $\epsilon\to 0$ 即可. ■

4.3　线性泛函导数和内蕴导数的补充

本节内容取自 (Cardaliaguet et al., 2019). 本节主要是补充对线性泛函导数和内蕴导数的其他观点, 空间取为 d 维环面 $\mathbb{T}^d$,$\mathscr{P}(\mathbb{T}^d)$ 表示 $\mathbb{T}^d$ 上全体概率测度.

定义 4.6　称函数 $U:\mathscr{P}(\mathbb{T}^d)\to\mathbb{R}$ 是 C^1 的, 是指: 存在连续映射 $\frac{\delta U}{\delta m}:\mathscr{P}(\mathbb{T}^d)\times\mathbb{T}^d\to\mathbb{R}$, 使得对任意 m,m' 都有

$$\int_{\mathbb{T}^d}\frac{\delta U}{\delta m}(m,y)\mathrm{d}m(y)=0$$

且

$$\lim_{s\downarrow 0}\frac{U((1-s)m+sm')-U(m)}{s}=\int_{\mathbb{T}^d}\frac{\delta U}{\delta m}(m,y)\mathrm{d}(m-m')(y)$$

注　在该定义下, 同样有

$$U(m')-U(m)=\int_0^1\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\frac{\delta U}{\delta m}((1-s)m+sm',y)\mathrm{d}(m-m')(y)\mathrm{d}s,m,m'\in\mathscr{P}(\mathbb{T}^d),$$

这一点与之前定义的线性泛函导数是一致的.

定义 4.7　设 $U : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ 是 C^1 的, 且 $\frac{\delta U}{\delta m}$ 关于第二个变量是连续可微的, 则定义 U 的内蕴导数为

$$D_m U(m,y) = \nabla_y \frac{\delta U}{\delta m}(m,y) : \mathscr{P}(\mathbb{T}^d) \times \mathbb{T}^d \to \mathbb{R}^d$$

注　可以证明, 该内蕴导数的定义与之前也是一致的. 实际上, 在上一节已经指出: 在一定条件下, 内蕴导数就是 L 导数, 线性泛函导数就是外在导数, 而 L 导数又是外在导数关于空间变量的梯度, 因此这些不同定义可以看作用不同性质刻画" 同一个对象", 所以这些一致性的保持是自然的.

依照同样的观点, 在一定条件下, 当固定空间变量 y 时, 可以对线性泛函导数 $\frac{\delta U}{\delta m}$ 再求一次线性泛函导数, 这样就可以归纳定义高阶线性泛函导数和高阶内蕴导数. 这里只考虑二阶的情形.

$\frac{\delta^2 U}{\delta m^2}$ 是一个从 $\mathscr{P}(\mathbb{T}^d) \times \mathbb{T}^d \times \mathbb{T}^d$ 到 $\mathbb{R}$ 的连续映射, 而 $D^2_{mm}U(m,y,y') = D^2_{y,y'}\frac{\delta^2 U}{\delta m^2}(m,y,y')$ 则是从 $\mathscr{P}(\mathbb{T}^d) \times \mathbb{T}^d \times \mathbb{T}^d$ 到 $\mathbb{R}^{d\times d}$ 的映射. 假设 U 的一阶和二阶线性泛函导数关于空间变量都是连续可微的, 二阶线性泛函导数和 $D_y \frac{\delta^2 U}{\delta m^2}$ 关于所有变量联合连续, 则有以下关系成立:

(1) $\frac{\delta^2 U}{\delta m^2}(m,y,y') - \frac{\delta^2 U}{\delta m^2}(m,y) = \frac{\delta U}{\delta m}(m,y',y) - \frac{\delta U}{\delta m}(m,y')$;

(2) $D_y \frac{\delta^2 U}{\delta m^2}(m,y,y') = \frac{\delta}{\delta m}(D_m(m,y))(y')$;

(3) $D_m(D_m U(\cdot,y))(m,y') = D^2_{mm}U(m,y,y')$.

需要说明的是, 上面的第三个等式并不是" 天然" 成立的, 因为二阶内蕴导数并不是归纳定义得到的.

设 $(B_t),(W_t)$ 是一个完备带流概率空间上的两个布朗运动. 考虑方程

$$dX_t = -\beta_t(X_t)\mathrm{d}t + \sqrt{2}(\mathrm{d}B_t + \mathrm{d}W_t), t \in [0,T],$$

$(m_t) = (\mathscr{L}(X_t))$. 映射 $U : [0,T] \times \mathbb{T}^d \times \mathscr{P}(\mathbb{T}^d)$ 满足一系列正则性条件 (具体见 (Cardaliaguet et al., 2019) 的定义 2.4.4), 则存在一族随机变量 $(\epsilon_{s,t})_{t\in[0,T],s\leq t}$, 使得

如下局部 Ito-Taylor 展开成立:

$$
\begin{aligned}
&\frac{1}{h}\Big(\mathbb{E}[U(t+h,x+\sqrt{2}(W_{t+h}),m_{t+h})-U(t+h,x+\sqrt{2}(W_{t+h}),m_t)\mid\mathscr{F}_t]\Big)\\
=&\Delta_x U(t,x+\sqrt{2}(W_t),m_t)\\
&+2\int_{\mathbb{T}^d}\mathrm{div}_y[D_mU](t,x+\sqrt{2}(W_t),m_t,y)dm_t(y)\\
&-\int_{\mathbb{T}^d}D_mU(t,x+\sqrt{2}(W_t),m_t,y)\cdot\beta_t(y)dm_t(y)\\
&+2\int_{\mathbb{T}^d}\mathrm{div}_xD_mU(t,x+\sqrt{2}(W_t),m_t,y)dm_t(y)\\
&+\int_{[\mathbb{T}^d]^2}\mathrm{tr}\Big(D^2_{mm}U(t,x+\sqrt{2}(W_t),m_t,y,y')\Big)dm_t(y)dm_t(y')\\
&+\epsilon_{t,t+h}.
\end{aligned}
$$

上述公式的完整陈述及证明见 (Cardaliaguet et al., 2019) 的附录 A.3.

4.4　若干例子

对上一节定义的几种导数, 我们介绍几个简单可计算的例子, 并体现它们之间的联系.

称函数 $f:\mathscr{M}_p\to\mathbb{R}$ 为 C_b^1- 柱函数或 $f\in\mathscr{F}C_b^1(\mathscr{M}_p)$, 是指: 存在 $n\in\mathbb{N}$,$g\in C^2(\mathbb{R}^n)$ 和 $h_i:1\le i\le n\subset C_b^1(\mathbb{R}^d)$, 使得

$$f(\mu)=g(\mu(h_1),\cdots,\mu(h_n)),\mu\in\mathscr{P}_p.$$

柱函数是应用较广的一类函数, 例如取 $n=1,g(x)=x$ 以及适当的 $h:\mathbb{R}^d\to\mathbb{R}$, 此时 $f(\mu)=\mu(h)$ 就是一个线性函数, 当 $\mu\in\mathscr{M}_p$ 或 $\mathscr{P}_p$ 时, 容易计算

$$
\begin{aligned}
&D^Ef(\mu)(x)=h(x),\\
&\tilde{D}^Ef(x)=h(x)-\mu(h),
\end{aligned}
\tag{4.9}
$$

并且可以验证 D^Ef 就是线性泛函导数, 且根据第三章计算知, 当 h 连续可微时,$D^Lf(\mu)(x)=\nabla h(x)$.

对于任意光滑函数 $g:\mathbb{R}\to\mathbb{R}$,

$$
\begin{aligned}
&f(\mu+s\delta_x)-f(\mu)\\
=&g((\mu+s\delta_x)(h))-g(\mu(h))\\
=&\int_0^1 g'(\mu(h)+t((\mu+s\delta_x)(h)-\mu(h)))((\mu+s\delta_x)(h)-\mu(h))\mathrm{d}t\\
=&\int_0^1 g'(\mu(h)+sth(x))sh(x)\mathrm{d}t,
\end{aligned}
\tag{4.10}
$$

观察并验证可得,$D^E f(\mu)(x) = g'(\mu(h))h(x)$. 对于 f 的内蕴导数, 取光滑向量场 ϕ,

$$\begin{aligned}
&f(\mu \circ (\mathrm{id} + s\phi)^{-1}) - f(\mu) \\
=&g(\mu(h \circ (\mathrm{id} + s\phi))) - g(\mu(h)) \\
=&\int_0^1 g'(\mu(h) + t(\mu(h \circ (\mathrm{id} + s\phi)) - \mu(h)))(\mu(h \circ (\mathrm{id} + s\phi)) - \mu(h)))\mathrm{d}t,
\end{aligned} \tag{4.11}$$

设 h 是紧支光滑的, 则 $\lim_{s\downarrow 0} \frac{h(x+s\phi(x))-h(x)}{s} = \langle \nabla h(x), \phi(x) \rangle$, 观察并验证可得 $D^I_\phi f(\mu) = g'(\mu(h))\langle \nabla, \phi \rangle$, 也即 $D^I f(\mu) = g'(\mu(h))\nabla h$. 以上讨论均可推至一般的柱函数 $f(\mu) = g(\mu(h_1), \cdots, \mu(h_n))$,

$$\begin{aligned}
D^E f(\mu) &= \sum_{i=1}^n \partial_i g(\mu(h_1), \cdots, \mu(h_n))h_i, \\
D^L f(\mu) &= \sum_{i=1}^n \partial_i g(\mu(h_1), \cdots, \mu(h_n))\nabla h_i.
\end{aligned} \tag{4.12}$$

下面考虑卷积函数 $f(\mu) = \mu(h * \mu) = \int_{\mathbb{R}^d} (h * \mu)(x)\mathrm{d}x = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(x - y)\mathrm{d}\mu(y)\mathrm{d}\mu(x)$, 则 $f(\mu + s\delta_z) = f(\mu) + s\mu(h) + s((h + \bar{h}) * \mu)(z) + s^2 h(0)$, 其中 $\bar{h}(x) = h(-x)$, 即 $D^E f(\mu)(z) = ((h + \bar{h}) * \mu)(z)$, 这也与第三章的计算相对应.

第 5 章　水平导数和垂直导数

本章主要参考引用自 (Cont et al., 2013) 和 (Peng et al., 2023).

考虑一个带流的概率空间 $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, 其中滤流 $(\mathscr{F}_t)$ 满足通常条件, 即 $(\mathscr{F}_t)$ 右连续: 对任意 $t \geq 0, \mathscr{F}_{t+} := \cap_{s>t}\mathscr{F}_s = \mathscr{F}_t$ 且 $\mathscr{F}_0$ 包含所有的 $\mathbb{P}-$ 零测集. 记 $\mathcal{P}, \mathcal{O}$ 分别表示 $\mathbb{R}_+ \times \Omega$ 上由全体左连续过程和全体右连左极[①]过程生成的 σ 代数, 分别称为可料 σ 代数和可选 σ 代数. $X : [0,T] \times \Omega \to \mathbb{R}^d$ 是一个连续半鞅,$\mathscr{F}_t^X$ 表示 X 生成的自然 σ 代数流.$\langle X\rangle := (\langle X^i, X^j\rangle)_{1\leq i,j\leq d}$ 表示 X 的二次变差, 是一个取值于 S_d^+ 的过程, 其中 S_d^+ 表示全体 d 维正定方阵. 我们总假设存在一个取值于 S_d^+ 的右连左极过程 (A_s), 使得

$$\langle X\rangle_t = \int_0^t A_s \mathrm{d}s.$$

记 $D([0,T];\mathbb{R}^d)$ 表示 $[0,T]$ 上全体右连左极的 $\mathbb{R}^d$ 值函数,$D([0,T];S_d^+)$ 表示 $[0,T]$ 上全体右连左极的 S_d^+ 值函数 $C_0([0,T];\mathbb{R}^d)$ 则表示 $[0,T]$ 上全体在 0 时刻取值为 0 的连续函数. 对于以上函数空间中的元素或其他正半轴及其子集上的函数 x, 不加区分地使用 x_t 或 $x(t)$ 表示函数 x 在 t 处的取值. 而对于 Ω 上的随机过程 $Z = (Z_t)$,Z_t 或 $Z(t,\cdot)$ 表示一个随机变量, 而 $Z_\cdot(\omega)$ 或 $Z(\cdot,\omega)$ 表示一个样本路径.

5.1　水平延拓和垂直扰动

对于一个右连左极函数 $x \in D([0,T];\mathbb{R}^d)$ 或 $x \in D([0,T];S_d^+)$, 它在 t 时刻的截断记为 $x^t = (x_s, 0 \leq s \leq t) \in D([0,t];\mathbb{R}^d)$.(类似地对于一个随机过程 X, 定义在 t 时刻的截断为 $X^t = (X_s, 0 \leq s \leq t)$)

对 $r > 0$, 定义 x^t 到 $[0,t+r]$ 的水平延拓为 $x^{t,\to r} \in D([0,t+r];\mathbb{R}^d)$,

$$x_s^{t,\to r} = \begin{cases} x_s, & s \in [0,t], \\ x_t, & s \in (t,t+r]. \end{cases} \tag{5.1}$$

对 $v \in \mathbb{R}^d$, 定义 x^t 沿方向 v 的垂直扰动为 $x^{t,\uparrow v} \in D([0,t];\mathbb{R}^d)$,

$$x_s^{t,\uparrow v} = \begin{cases} x_s, & s \in [0,t) \\ x_t + v, & s = t. \end{cases} \tag{5.2}$$

[①] 对 $[0,\infty]$ 或连通子集上取值于一个度量空间的映射来说, 右连左极是指处处右连续且左极限存在; 对随机过程则是指样本路径作为函数是右连左极的. 左连右极类似, 但较少使用.

注 水平延拓和垂直扰动分别满足交换律和结合律, 这继承自欧式空间的交换律和结合律, 也即对 $\forall x \in D([0,T];\mathbb{R}^d)$, $\forall t < T$,

$$\begin{aligned}&\forall r,s \in [0,T-t]\text{且}r+s \in [0,T-t], (x^{t,\to r})^{t+r,\to s} = (x^{t,\to s})^{t+s,\to r} = x^{t,\to r+s},\\&\forall v,w \in \mathbb{R}^d, (x^{t,\uparrow v})^{t,\uparrow w} = (x^{t,\uparrow w})^{t,\uparrow v} = x^{t,\uparrow v+w},\end{aligned} \tag{5.3}$$

但水平延拓和垂直扰动两种操作一般不能交换顺序.

注 直观上看, 如果将一个右连左极函数看作粒子运动轨迹的话, 水平延拓就是用当前时刻粒子的位置来近似预测往后一小段时间的位置, 而垂直扰动则是修改当前时刻粒子位置, 使其沿某一方向做微小平移.

此外, 对于 $x \in D([0,T];\mathbb{R}^d)$(或 $x \in D([0,T];S_d^+)$), 定义 $x^{t-} \in D([0,t];\mathbb{R}^d)$(或 $x \in D([0,t];S_d^+)$),

$$x_s^{t-} = \begin{cases} x_s, & s \in [0,t) \\ x_{t-}, & s = t. \end{cases} \tag{5.4}$$

即将 t 时刻的函数值修改为左极限.(一个右连左极函数将任意点的值修改为该点的左极限是可行的, 修改后的函数不再是右连左极而是左连右极的.)

考虑停止路径空间 $\Gamma = \{(t,\omega^{t,\to T-t}) : \omega \in D([0,T];\mathbb{R}^d) \times D([0,T];S_d^+)\}$, 配备一致度量 d_∞,

$$d_\infty((t,x),(s,y)) = |t-s| + \sup_{r\in[0,T]} |x_r^{t,\to T-t} - y_r^{s,\to T-s}|,$$

注 在 (Peng et al., 2023) 第 3 节的定义中, 对 $\mathbb{R}_+ \times D([0,T];\mathbb{R}_d)$ 使用的距离是

$$\tilde{d}_\infty((t,x),(s,y)) = |t-s|^{\frac{1}{2}} + \sup_{r\in[0,T]} |x_r^{t,\to T-t} - y_r^{s,\to T-s}|,$$

但没有本质上的差异.

在不致混淆的前提下, 对 Γ 中的元素 $(t,(x^{t,\to T-t}, a^{t,\to T-t}))$, 简记为 (x^t, a^t). 对记 $\mathcal{R}_t = D([0,t];\mathbb{R}^d, \mathcal{S}_t = D([0,t];S_d^+)$ 以及 $\mathcal{D}_t = D([0,T];\mathbb{R}^d) \times D([0,T];S_d^+)$. 以下总假设 $F = (F_t)_{t\in[0,T]}$ 满足:

(1) $\forall t \in [0,T]$,F_t 是 $\mathcal{D}_t$ 上的泛函,$F(t,x,a) = F_t(x^t,a^t)$;

(2) $\forall t \in [0,T]$,F_t 是关于 $\mathscr{B}(\mathcal{R}_t) \otimes \mathscr{B}(\mathcal{S}_t)$ 可测, 其中路径空间上的拓扑取为 (林正炎 等, 2023) 第七章第 5 节中度量 d_0 诱导生成的 Skorohod 拓扑; 而任意 $t \in [0,T]$, 映射 $\mathbb{R}_+ \times \mathcal{R}_T \times \mathcal{S}_T \ni (s,x,a) \mapsto F(s,x,a) = F_s(x^s,a^s), s \in [0,t]$ 是关于 $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{B}(\mathcal{R}_t) \otimes \mathscr{B}(\mathcal{S}_t)$ 可测的;

(3) $F_t(x,a) = F_t(x,a^{t-}), (x,a) \in \mathcal{D}_t$.

注 粗略地讲, 泛函 F_t 对第一个变量要求依赖于 t 时刻及之前的行为, 而对第二个变量的要求不依赖于 t 时刻; 另外, 在讨论 $F(t,x,a)$ 或 $F_t(x,a)$ 时, 默认通

过截断或水平延拓的方法使得 (x,a) 取于对应的定义域中.

5.2　泛函的连续性

考虑一个泛函 $F=(F_t)$, 它可以看作是 Γ 上的泛函, 而 Γ 已经配备了一个距离, 所以定义泛函的连续性是可行的.

定义 5.1　称泛函 $F=(F_t)_{t\in[0,T]}$ 在固定时间连续, 是指: 任意 $t\in[0,T]$,$\forall\epsilon>0$,$\forall(x,a)\in\mathcal{D}_t$, 存在 $\delta>0$, 使得若 $(x',a')\in\mathcal{D}_t$ 满足

$$d_\infty((t,(x,a)),(t,(x',a')))<\delta,$$

则

$$|F_t(x,a)-F_t(x',a')|<\epsilon.$$

定义 5.2　称泛函 $F=(F_t)_{t\in[0,T]}$ 在 $(x,a)\in\mathcal{D}_t$ 处连续, 是指: $\forall\epsilon>0$, 存在 $\delta>0$, 使得若 $(s,(y^s,b^s))\in\Gamma$ 满足

$$d_\infty((t,(x,a)),(s,(y^s,b^s)))<\delta,$$

则

$$|F_t(x,a)-F_s(y^s,b^s)|<\epsilon.$$

定义 5.3　称泛函 $F=(F_t)_{t\in[0,T]}$ 左连续, 是指:$\forall t\in[0,T],\forall\epsilon>0,\forall(x,a)\in\mathcal{D}_t$, 存在 $\delta>0$, 使得任意 $h>0$,$\forall(x',a')\in\mathcal{D}_{t-h}$, 若满足

$$d_\infty((t,(x,a)),(t-h,(x',a')))<\delta,$$

则

$$|F_t(x,a)-F_{t-h}(x',a')|<\epsilon.$$

类似于左连续, 还可以定义右连续. 将 Γ 上全体连续泛函记为 $C^{0,0}([0,T])$, 全体左连续泛函记为 $C_l^{0,0}([0,T])$.

命题 5.1　设 $F\in C_l^{0,0}([0,T])$. 则对任意 $(x,a)\in\mathcal{D}_T$, 映射 $t\mapsto F_t(x^{t-},a^{t-})$ 是左连续的.

由于修改后的函数是左连右极的, 所以这个性质是自然的.

定义 5.4　称泛函 $F=(F_t)$ 保持有界, 是指: 对任意 $\mathbb{R}^d$ 中的紧集 K, 任意 $R>0,t<T$, 都存在 $C>0$, 使得 $\forall s<t,\forall(x,a)\in D([0,T];K)\times\mathcal{S}_t$, 蕴含关系

$$\sup_{s\in[0,t]}|a_s|<R\implies|F_t(x,a)|<C$$

成立.

5.3　路径导数的定义和例子

下面给出水平导数和垂直导数的定义, 这两种导数统称为 Dupire 导数 (见 (Dupire, 2009)) 或路径导数.

定义 5.5　称泛函 $F=(F_t)_{t\in[0,T]}$ 在 $(x,a)\in\mathcal{D}_t$ 处水平可导, 是指: 极限

$$D_t^H F(x,a):=\lim_{r\downarrow 0}\frac{F_{t+r}(x^{t,\to r},a^{t,\to r})-F_t(x,a)}{r}$$

存在.

如果 F 在所有 $(x,a)\in\Gamma$ 处水平可导, 则可以定义映射

$$\begin{aligned}D_t^H F:\mathcal{D}_t&\to\mathbb{R}\\(x,a)&\mapsto D_t^H F(x,a),\end{aligned}\tag{5.5}$$

以及泛函 $D^H F=(D_t^H F)_{t\in[0,T]}$, 称为 F 的水平导数.

注　在 (Peng et al., 2023) 中, 对 $u:\mathbb{R}_+\times\mathcal{R}_T$ 的水平导数定义中, 考虑的是 $(u(t+r,x^t)-u(t,x^t))/r$, 这是一点细微的差别.

定义 5.6　称泛函 $F=(F_t)_{t\in[0,T]}$ 在 $(x,a)\in\mathcal{D}_t$ 处垂直可导, 是指: 存在 $p\in\mathbb{R}^d$, 使得任意

$$F_t(x^{t,\uparrow v},a^t)=F_t(x^t,a^t)+\langle p,v\rangle+o(|v|), \text{当} v\to 0_{\mathbb{R}^d},$$

p 称为 F 在 (x,a) 处的垂直导数, 记为 $D_t^V F(x,a)$, 类似地有映射 $D_t^V F$ 和泛函 $D^V F=(D_t^V F)_{t\in[0,T]}$.

注　类似于多元微积分, 以上垂直导数可以看作一阶的, 并可以定义高阶的垂直导数, 例如二阶垂直导数在计算中可以视为一阶垂直导数逐分量求导, 同构于一个 d 维对称方阵,k 阶垂直导数相应地记为 $D_t^{V,k}F$ 和 $D^{V,k}F$.

水平导数可以看作是关于时间的导数, 而垂直导数则是关于空间的导数, 在理论和应用中, 一般更关心对空间的正则性, 而关于时间的正则性要求则不会太高, 因此我们定义以下泛函类: 称 $F\in C^{1,k}([0,T])$, 是指:

(1) F 水平可导, 且 $D^H f$ 在固定时刻连续 (见定义5.1);

(2) F 是 k 次垂直可导的, 且任意 $1\le i\le k$,$D^{V,k}F\in C_l^{0,0}([0,T])$.

5.4　若干例子

这里给出几种常见泛函的路径导数的例子.

例 5.1　设 $F_t(x)=f(t,x_t),t\in[0,T],x\in\mathcal{R}_T$, 其中 $f\in C^1(\mathbb{R}^{1+d})$, 记 $\partial_1 f$ 表示 f 关于第一个分量的偏导数,$\partial_2 f$ 表示对第二个分量的梯度, 则

$$\begin{aligned}&F_{t+r}(x^{t,\to r})-F_t(x^t)\\=&f(t+r,x_{t+r}^{t,\to r})-f(t,x_t^t)\\=&f(t+r,x_t)-f(t,x_t),\end{aligned}\tag{5.6}$$

所以 F 的水平导数为 $D_t^H F(x) = \partial_1 f(t, x_t)$.

而

$$F_t(x^{t,\uparrow he_i}) - F_t(x^t) = f(t, x_t + he_i) - f(t, x_t),$$

所以 F 的垂直导数为 $D_t^V F(x) = \partial_2 f(t, x_t)$.

可以看出,f 的对第二个变量的光滑性越好,F 的垂直光滑性也越好.

例 5.2　给定 $h \in C^k(\mathbb{R}^d)$,$g \in C(\mathbb{R}^n)$ 以及 $\{t_i, i = 1, \cdots, n\} \subset [0, T]$, 设

$$F_t(x) = h(x_t - x_{t_n-}) I_{t \geq t_n} g(x(t_1-), \cdots, x(t_n-)).$$

注意到对 $\forall t < t_n$,$F_t \equiv 0$, 而对于 $t \geq t_n$,

$$\begin{aligned} F_{t+r}(x^{t,\to r}) &= h(x_t^{t,\to r} - x_t^{t,\to r}) g(x_{t_1-}^{t,\to r}, \cdots, x_{t_n-}^{t,\to r}) \\ &= h(x_t - x_{t_n-}) I_{t \geq t_n} g(x(t_1-), \cdots, x(t_n-)) \\ &= F_t(x^t), \end{aligned} \tag{5.7}$$

所以水平导数总是 0. 而垂直导数为

$$D_t^V F(x) = \nabla h(x_t - x_{t_n-}) I_{t \geq t_n} g(x(t_1-), \cdots, x(t_n-)).$$

注意对古典微积分或 Banach 空间中的 Frechet 导数, 如果一个函数的导数在一个连通开集上恒为 0, 则可以得到该函数在这个开集上取常值, 但这个例子则说明, 水平导数一般没有这样的性质, 这是因为在水平导数的定义中, 路径的变化实际上是通过水平延拓实现的, 但水平延拓只是扩展了函数的定义域, 没有真正带来更多” 信息”, 这也是 (Cont et al., 2013) 和 (Peng et al., 2023) 中水平导数的定义稍有差异但很多时候没有本质区别的原因之一.

例 5.3　给定 $g \in C(\mathbb{R}^d)$, 令 $F_t(x, a) = \int_0^t g(x_s) a_s \mathrm{d}s$. 上面两个例子中的泛函只依赖于路径在有限个时刻的取值, 这个例子则是依赖 t 时刻及之前所有时间的行为.

首先,

$$\begin{aligned} F_{t+r}(x^{t,\to r}, a^{t,\to r}) &= \int_0^{t+r} g(x_s^{t,\to r}) a_s^{t,\to r} \mathrm{d}s \\ &= F_t(x^t, a^t) + \int_t^{t+r} g(x_t) a_t \mathrm{d}s \\ &= F_t(x^t, a^t) + r g(x_t) a_t, \end{aligned} \tag{5.8}$$

所有水平导数为 $D_t^H F(x, a) = g(x_t) a_t$, 而由于单点值的改变不影响积分, 所以垂直导数为 0, 进而 F 还是任意阶垂直可导的.

例 5.4　给定 $\epsilon > 0$, 令 $F_t(x) = x(t - \epsilon)$, 显然 F 是任意阶垂直可导的且垂直导数总为 0, 但对充分小的 $r > 0$,

$$F_{t+r}(x^{t,\to r}) - F_t(x^t) = x^{t,\to r}(t + r - \epsilon) - x^t(t - \epsilon) = x(t + r - \epsilon) - x(t - \epsilon),$$

由于 x 只是右连左极的, 所以 F 在" 绝大多数" 点处都不是水平可导的.

这一例子说明水平可导和垂直可导之间不必有关联.

例 5.5　下面考虑一类简单的复合. 给定 $F\in C^{1,2}([0,T]), x\in\mathcal{D}_T$ 和 $0\le t<T$, 定义映射

$$\begin{aligned}\phi:\ [0,T-t] &\longrightarrow \mathbb{R}\\ s &\longmapsto F_{t+s}(x^{t,\to s}),\\ \psi:\ \mathbb{R}^d &\longrightarrow \mathbb{R}\\ v &\longmapsto F_t(x^{t,\uparrow v}),\end{aligned}$$

则 ϕ 在 $(0,T-t)$ 上是连续可导的, 这是因为, 对任意 $s\in(0,T-t)$,

$$\begin{aligned}\lim_{r\downarrow 0}\frac{\phi(s+r)-\phi(s)}{r}&=\lim_{r\downarrow 0}\frac{F_{t+s+r}(x^{t,\to(s+r)})-F_{t+s}(x^{t,\to r})}{r}\\&=\lim_{r\downarrow 0}\frac{F_{t+s+r}((x^{t,\to s})^{t+s,\to r})-F_{t+s}(x^{t,\to s})}{r}\\&=D^H_{t+s}F(x^{t,\to s}),\end{aligned}$$

类似可得 $\nabla\psi(v)=D^V_tF(x^{t,\uparrow v})$ 和 $\nabla^2\psi(v)=D^{V,2}_tF(x^{t,\uparrow v})$

5.5　Ito 公式

下面给出路径导数的 Ito 公式.

定理 5.2　设泛函 $F\in C^{1,2}([0,T])$. 则对任意 $0\le t<T$,

$$\begin{aligned}&F_t(X^t,A^t)-F_0(X^0,A^0)\\&=\int_0^t D^H_sF(X^s,A^s)\mathrm{d}s+\int_0^t D^V_sF(X^s,A^s)\mathrm{d}s\\&\quad+\frac{1}{2}\int_0^t \mathrm{tr}(D^{V,2}_sF(X^s,A^s)A_s)\mathrm{d}s\end{aligned}\tag{5.9}$$

证明　首先假设 X 的支集包含在一个紧集 K 内, 且 $\|A\|_\infty\le R$.

对于任意的 $n\in\mathbb{N}$, 将区间 $[0,t]$ 做 2^n 等分, 端点集 $P_n=\{\frac{kt}{2^n}:k=0,1,\cdots,2^n-1\}$. 考虑一列停时 $\{\tau^n_k,k=1,\cdots,k_n\}$, 其中 $\tau^n_0=0$,

$$\tau^n_k=t\wedge\inf\{s>\tau^n_{k-1}:2^ns\in\mathbb{N}\text{或}|A_s-A_{s-}|>\frac{1}{n}\},k\ge 1,$$

直观上看, 停时 τ^n_1 就是第一个 2^n 分端点或第一次发生较大跳的时间, 其余类似. 且根据定义和 A 的一致有界性, 一定会在有限个停时到达 t, 因此 $2^n\le k_n<\infty$, 从而 $\{\tau^n_k,k=1,\cdots,k_n\}$ 给出了区间 $[0,t]$ 的一个有限随机剖分, 且两个相邻停时之间的间隔 $h^n_k=\tau^n_{k+1}-\tau^n_k\le\frac{1}{2^n}$. 此外, 定义

$$\eta_n=\frac{t}{2^n}+\sup_{1\le k\le k_n}\sup_{s\in[\tau^n_k,\tau^n_{k+1})}\{|A_s-A_{\tau^n_k}|+|X_s-X_{\tau^n_k}|\},$$

由于右连左极的性质, 当 $n \to \infty, \eta_n \to 0$.

考虑 (X_t, A_t) 的分段近似, 对 $n \in \mathbb{N}$, 令

$$\begin{aligned} X_s^{(n)} &= X_t I_{\{t\}}(s) + \sum_{k=0}^{k_n} X_{\tau_k^n} I_{[\tau_k^n, \tau_{k+1}^n)}(s), \\ A_s^{(n)} &= A_t I_{\{t\}}(s) + \sum_{k=0}^{k_n} A_{\tau_k^n} I_{[\tau_k^n, \tau_{k+1}^n)}(s), \end{aligned} \tag{5.10}$$

注意到如果某 $\tau_k^n = t$, 则 $\tau_{k+1}^n = t$, 从而 $[\tau_k^n, \tau_{k+1}^n) = \emptyset$, 所以不会出现重复加和的情况.

考虑

$$\begin{aligned} & F_{\tau_{k+1}^n}((X^{(n)})^{\tau_{k+1}^n -}, (A^{(n)})^{\tau_{k+1}^n -}) - F_{\tau_k^n}((X^{(n)})^{\tau_k^n -}, (A^{(n)})^{\tau_k^n -}) \\ =& F_{\tau_{k+1}^n}((X^{(n)})^{\tau_{k+1}^n -}, (A^{(n)})^{\tau_k^n, \to h_k^n}) - F_{\tau_k^n}((X^{(n)})^{\tau_k^n}, (A^{(n)})^{\tau_k^n}) \\ & + F_{\tau_k^n}((X^{(n)})^{\tau_k^n}, (A^{(n)})^{\tau_k^n -}) - F_{\tau_k^n}((X^{(n)})^{\tau_k^n -}, (A^{(n)})^{\tau_k^n -}) \\ \triangleq & I_1 + I_2. \end{aligned} \tag{5.11}$$

结合性质5.1和 F 水平导数的连续性可得,

$$\begin{aligned} I_1 &= F_{\tau_{k+1}^n}((X^{(n)})^{\tau_{k+1}^n -}, (A^{(n)})^{\tau_k^n, \to h_k^n}) - F_{\tau_k^n}((X^{(n)})^{\tau_k^n}, (A^{(n)})^{\tau_k^n}) \\ &= F_{\tau_{k+1}^n}((X^{(n)})^{\tau_k^n, \to h_k^n}, (A^{(n)})^{\tau_k^n, \to h_k^n}) - F_{\tau_k^n}((X^{(n)})^{\tau_k^n}, (A^{(n)})^{\tau_k^n}) \\ &= \int_0^{h_k^n} D^H_{\tau_k^n + s} F((X^{(n)})^{\tau_k^n, \to s}, (A^{(n)})^{\tau_k^n, \to s}) \mathrm{d}s, \end{aligned} \tag{5.12}$$

这对应于5.9的第一项.

定义 $\phi(v) = F_{\tau_k^n}((X^{(n)})^{\tau_k^n -, \uparrow (v - X_{\tau_k^n})}, (A^{(n)})^{\tau_k^n})$, 这是一个从 $\mathbb{R}^d$ 到 $\mathbb{R}$ 的函数. 易见 $\phi(X_{\tau_k^n}) = F_{\tau_k^n}((X^{(n)})^{\tau_k^n -}, (A^{(n)})^{\tau_k^n -})$, 而

$$\begin{aligned} \phi(X_{\tau_{k+1}^n}) &= F_{\tau_k^n}((X^{(n)})^{\tau_k^n -, \uparrow (X_{\tau_{k+1}^n} - X_{\tau_k^n})}, (A^{(n)})^{\tau_k^n}) \\ &= F_{\tau_k^n}((X^{(n)})^{\tau_k^n}, (A^{(n)})^{\tau_k^n -}) \end{aligned} \tag{5.13}$$

另一方面, 由于 $F \in C_b^{1,2}([0,T])$, 所以 $\phi \in C^2(\mathbb{R}^d)$, 且 $\nabla\phi(v) = D^V_{\tau_k^n} F((X^{(n)})^{\tau_k^n -, \uparrow v}, (A^{(n)})^{\tau_k^n})$, $\nabla^2\phi(v) = D^{V,2}_{\tau_k^n}(X^{(n)})^{\tau_k^n -, \uparrow v}, (A^{(n)})^{\tau_k^n}$, 所以可以应

用经典的 Ito 公式

$$
\begin{aligned}
I_2 &= \phi(X_{\tau_{k+1}^n}) - \phi(X_{\tau_k^n}) \\
&= \int_{\tau_k^n}^{\tau_{k+1}^n} \nabla\phi(X_s)\cdot \mathrm{d}X_s \\
&\quad + \frac{1}{2}\int_{\tau_k^n}^{\tau_{k+1}^n} \mathrm{tr}(\nabla^2\phi(X_s)\mathrm{d}\langle X\rangle_s) \\
&= \int_{\tau_k^n}^{\tau_{k+1}^n} D^V_{\tau_k^n} F((X^{(n)})^{\tau_k^n-,\uparrow(X_s-X_{\tau_k^n})},(A^{(n)})^{\tau_k^n-})\cdot \mathrm{d}X_s \\
&\quad + \frac{1}{2}\int_{\tau_k^n}^{\tau_{k+1}^n} \mathrm{tr}(D^{V,2}_{\tau_k^n} F((X^{(n)})^{\tau_k^n-,\uparrow(X_s-X_{\tau_k^n})},(A^{(n)})^{\tau_k^n-})\mathrm{d}\langle X\rangle_s)
\end{aligned}
\tag{5.14}
$$

综合以上,

$$
\begin{aligned}
&F_t((X^{(n)})^t,(A^{(n)})^t) - F_0(X^0,A^0) \\
&= \int_0^t D^H_s F((X^{(n)})^{\tau_{k(s)}^n,\to(s-\tau_{k(s)}^n)},(A^{(n)})^{\tau_{k(s)}^n,\to(s-\tau_{k(s)}^n)})\mathrm{d}s \\
&\quad + \int_0^t D^V_{\tau_{k(s)}^n} F((X^{(n)})^{\tau_{k(s)}^n-,\uparrow(X_s-X_{\tau_{k(s)}^n})},(A^{(n)})^{\tau_{k(s)}^n-})\cdot \mathrm{d}X_s \\
&\quad + \frac{1}{2}\int_0^t \mathrm{tr}(D^{V,2}_{\tau_{k(s)}^n} F((X^{(n)})^{\tau_{k(s)}^n-,\uparrow(X_s-X_{\tau_{k(s)}^n})},(A^{(n)})^{\tau_{k(s)}^n-})\mathrm{d}\langle X\rangle_s)
\end{aligned}
$$

令 $n\to\infty$ 即得5.9成立.

对于一般的 X, A, 取 $\overline{B_n} := \{v\in\mathbb{R}^d : |x|\le n\}$. 定义停时

$$\tau_n = t\wedge\{s : X_s\notin B_n \text{或} |A_s|>n\}.$$

对 $(X^{t\wedge\tau_n}, A^{t\wedge\tau_n})$ 应用公式5.9, 再令 $n\to\infty$ 即可. ∎

第 6 章　Malliavin 分析入门

本章主要内容参考引用自 (Nualart, 2006).

6.1　Hilbert 空间相关预备

由于 Malliavin 分析理论与 Hilbert 空间紧密关联, 所以先在此陈述一些可能用到的 Banach 空间,Hilbert 空间和线性代数相关内容. 本节关于泛函分析的内容主要取自 (Brezis, 2010),(Kunze, 2013) 和 (Reed et al., 1980), 线性代数相关内容主要取自 (李文威, 2025)

如无特别说明, 本章涉及的向量空间 (自然包括 Hilbert 空间) 总是实线性的,Hilbert 空间总是可分的, 即有可数稠密子集.

6.1.1　直和与 Hilbert 和

定义 6.1　设 $\mathbb{H}$ 是一个 Hilbert 空间. 称 $\{e_n\}_{n\in\mathbb{N}} \subset \mathbb{H}$ 是 $\mathbb{H}$ 的一个规范正交基 (orthonormal basis, 有时简记为 O.N.B.)[①], 是指:

(1) $|e_n| = 1, \forall n \geq 1$ 且当 $n \neq m, \langle e_n, e_m\rangle = 0$;

(2) $\overline{\text{span}\{e_1, e_2, \cdots\}} = \mathbb{H}$, 其中 $\overline{\text{span}\{e_1, e_2, \cdots\}}$ 表示 $\{e_n\}_{n\in\mathbb{N}}$ 所有可能的线性组合.

定义 6.2　称 Hilbert 空间 $\mathbb{H}$ 的子集 T 为全子集 (total set), 是指:$\overline{\text{span}(T)} = \mathbb{H}$.

命题 6.1　设 Hilbert 空间 $\mathbb{H}$ 可分. 则 $\mathbb{H}$ 有规范正交基.

对于任意一族集合, 可以定义它们的直积和直和, 特别地考虑一族向量空间 $(V_i)_{i\in I}$, 其直积为

$$\Pi_{i\in I} V_i = \{(v_i)_{i\in I} : \forall i, v_i \in V_i\},$$

加法和数乘定义为逐分量地继承自各分量原来所在的向量空间. 直和作为直积的子空间, 定义为

$$\oplus_{i\in I} V_i = \{(v_i)_{i\in I} \in \Pi_{i\in I} V_i : \text{仅有有限分量非}0\}.$$

如果所有 $V_i = V$, 记直积为 V^I, 直和为 $V^{\oplus I}$.

给定一个一个向量空间 V, 和一族子空间 $(V_i)_{i\in I}$, 定义这族子空间的和为

$$\sum_{i\in I} V_i = \{\sum_i v_i : \forall i, v_i \in V_i \text{且仅有有限} i, v_i \neq 0\},$$

[①] 在有些教材或文献中会使用完备规范正交系 (complete orthonormal system) 一词来表示线性张成的正交补为零空间的集合, 而 O.N.B. 表示有限或无限线性组合张成整个空间的集合, 二者是等价的, 所以不作区分, 具体见 (张恭庆 等, 2021) 的 1.6 节

若 $\forall i, V_i \cap (\sum_{j\neq i} V_j) = \{0\}$, 则称该族子空间的和为直和, 记为 $\oplus_{i\in I} V_i$. 在 Hilbert 空间的情形下, 在子空间直和的基础上有 Hilbert 和的概念.

定义 6.3　给定 Hilbert 空间 $\mathbb{H}$ 和一列闭子空间 $\{E_n\}_{n\in\mathbb{N}}$. 称 $\mathbb{H}$ 是 $\{E_n\}$ 的 Hilbert 和, 是指:

(1) $\{E_n\}$ 两两正交, 即: 若 $n \neq m$, 则 $\forall v_n \in E_n, \forall v_m \in E_m, \langle v_n, v_m\rangle = 0$;

(2) $\overline{\text{span}(\cup_{n\geq 1} E_n)} = \mathbb{H}$. 也记为 $\mathbb{H} = \oplus_n E_n$

对任一 $\mathbb{H}$ 的闭子空间 M, 对应的正交投影算子记为 Proj_M. 在一定意义下,Hilbert 和给出了 $\mathbb{H}$ 中元素的分解或逼近方式, 即如下定理.

定理 6.2　设 Hilbert 空间 $\mathbb{H} = \oplus_{n\geq 1} E_n$. 给定 $u \in \mathbb{H}$, 定义 $u_n = \text{Proj}_{E_n} u$, 以及 $S_n = \sum_{i=1}^n u_n$. 则 $\lim_{n\to\infty} S_n = u$ 且

$$|u|^2 = \sum_{n=1}^{\infty} |u_n|^2$$

6.1.2　张量积

先从代数角度给出向量空间的张量积. 在这一部分, 尽管可以给出具体的向量空间构造, 但更多关心的是泛性质. 首先是线性映射定义的扩展.

定义 6.4　设 $V_1, \cdots, V_n, W$ 是向量空间. 称映射 $M : V_1 \times \cdots \times V_n \to W$ 是 n 重线性映射, 是指:M 对任一分量都是线性的, 也即: 任意 $1 \leq i \leq n, \forall a, b \in \mathbb{R}, \forall \alpha, \beta \in V_i$, 都有

$$M(\cdots, a\alpha + b\beta, \cdots) = aM(\cdots, \alpha, \cdots) + bM(\cdots, \beta, \cdots).$$

特别地, 当 $W = \mathbb{R}$, 称 M 为 n 重线性函数.

先考虑一种简单的情形, 即只有两个向量空间.

定理 6.3 (张量积的存在性和唯一性)　给定两个向量空间 V, W.

(1) 存在向量空间 L_u 和双线性映射 $B_u : V \times W \to L_u$, 使得对所有向量空间 L 和双线性映射 $B : V \times W \to L$, 都存在唯一的线性映射 $\phi : L_u \to L$, 使得下列图标交换:

$$\begin{tikzcd} V \times W \arrow[r, "B_u"] \arrow[dr, "B"'] & L_u \arrow[d, "\phi"] \\ & L \end{tikzcd}$$

(2) 若向量空间和双线性映射 $(L_u, B_u), (L'_u, B'_u)$ 都满足 (1) 中性质, 则存在唯一的线性同构 $\tilde{\phi} : L_u \to L'_u$ 使得下列图标交换:

$$
\begin{tikzcd}
V \times W \arrow[r, "B_u"] \arrow[dr, "B'_u"'] & L_u \arrow[d, "\tilde{\phi}"] \\
& L'_u
\end{tikzcd}
$$

注　我们将 (L_u, B_u) 称为张量积, 由于 L_u 在同构意义下唯一, 因此记 $V \otimes W = L_u$, 其中的元素 $v \otimes w = B_u(v, w)$, 一般也直接称 $V \otimes W$ 为 V, W 的张量积.$V \otimes W$ 可以具体实现为 $V \times W$ 上自由向量空间模去一个等价关系得到的商空间, 具体参见 (李文威, 2025) 的命题 15.1.1.

对于任意有限个向量空间也可以讨论张量积, 一种想法是在 $n=2$ 的基础上归纳定义, 下面通过多重线性映射给出直接的定义, 但可以证明这与归纳定义实际上是一致的.

定义 6.5　给定向量空间 $V_1, \cdots, V_n$. 存在向量空间 M_u 和 n 重线性映射 $C_u : V_1 \times \cdots \times V_n \to M_u$, 使得对任意 n 重线性映射 $C : V_1 \times \cdots \times V_n \to M$, 都存在唯一的线性映射 $\phi : M_u \to M$, 使得下列图表交换:

$$
\begin{tikzcd}
V_1 \times \cdots \times V_n \arrow[r, "C_u"] \arrow[dr, "C"'] & M_u \arrow[d, "\phi"] \\
& M
\end{tikzcd}
$$

类似于二元情形, 将 (M_u, c_u) 称为 $V_1, \cdots, V_n$ 的张量积, 其中 M_u 记为 $V_1 \otimes \cdots \otimes V_n$. 根据以上泛性质, 实际上给出了双射

$$
\begin{gathered}
\operatorname{Hom}(V_1 \otimes \cdots \otimes V_n; M) \xrightarrow{\sim} \operatorname{Mul}(V_1 \times \cdots \times V_n; M) \\
\phi \mapsto \phi \circ C_u
\end{gathered}
$$

我们还可以定义多重线性函数的张量积.

定理 6.4　设 $T : V_1 \times \cdots \times V_k \to \mathbb{R}$ 和 $S : V_{k+1} \times \cdots \times V_{k+l} \to \mathbb{R}$ 是两个多重线性函数, 则其张量积定义为如下函数:

$$
\begin{aligned}
T \otimes S : V_1 \times \cdots \times V_k \times V_{k+1} \times \cdots \times V_{k+l} \to \mathbb{R} & \\
(v_1, \cdots, v_k, v_{k+1}, \cdots, v_{k+l}) \mapsto T(v_1, \cdots, v_k) S(v_{k+1}, \cdots, v_{k+l}) &
\end{aligned}
\tag{6.1}
$$

注　容易验证,$T \otimes S$ 可以看作 $(V_1 \times \cdots \times V_k) \times (V_{k+1} \times \cdots \times V_{k+l})$ 上的双线性函数, 且满足结合律, 即 $(T \otimes S) \otimes R = T \otimes (S \otimes R)$.

我们来说明多元张量积在同构意义下的唯一性, 这基于以下命题.

定理 6.5　设 V_1, V_2, V_3 为向量空间. 则存在以下同构:

(1) (幺约束)

$$\begin{array}{ccccc} \mathbb{R}\otimes V_1 & \overset{\sim}{\longleftrightarrow} & V_1 & \overset{\sim}{\longleftrightarrow} & V_1\otimes\mathbb{R} \\ t\otimes v & \longmapsto & tv & \longleftarrow\!\shortmid & v\otimes t \end{array} \tag{6.2}$$

(2) (结合约束)

$$\begin{array}{ccccc} V_1\otimes(V_2\otimes V_3) & \overset{\sim}{\longleftrightarrow} & V_1\otimes V_2\otimes V_3 & \overset{\sim}{\longleftrightarrow} & (V_1\otimes V_2)\otimes V_3 \\ v_1\otimes(v_2\otimes v_3) & \longmapsto & v_1\otimes v_2\otimes v_3 & \longleftarrow\!\shortmid & (v_1\otimes v_2)\otimes v_3 \end{array} \tag{6.3}$$

(3) (交换约束)

$$\begin{array}{ccc} V_1\otimes V_2 & \overset{\sim}{\longleftrightarrow} & V_2\otimes V_1 \\ v_1\otimes v_2 & \longmapsto & v_2\otimes v_1 \end{array} \tag{6.4}$$

下面考虑 Hilbert 空间的情形. 设 H, V 为两个一般的 Hilbert 空间 (不假设是可分的),$(h_\alpha)_{\alpha\in A}, (v_\beta)_{\beta\in B}$ 分别是 H, V 的规范正交基. 定义 $H\otimes V$ 为所有形如 $\sum c_{\alpha,\beta}h_\alpha\otimes v_\beta, (c_{\alpha,\beta})\in l^2(A\times B)$ 的全体, 其中 $l^2(A\times B)$ 表示计数测度下平方可积函数全体, 这里的 $h_\alpha\otimes v_\beta$ 暂时只是形式上的记号, 无其他实际含义. 在 $H\otimes V$ 上定义内积如下:

$$\langle \sum c_{\alpha,\beta}h_\alpha\otimes v_\beta, \sum d_{\alpha,\beta}h_\alpha\otimes v_\beta\rangle := \sum c_{\alpha,\beta}d_{\alpha,\beta}.$$

从而 $H\otimes V$ 成为一个 Hilbert 空间, 一个规范正交基为 $(h_\alpha\otimes v_\beta)_{(\alpha,\beta)\in A\times B}$. 易见, 若 H, V 都是可分的, 则 $H\otimes V$ 也是可分的.

类似于向量空间的张量积, 对 Hilbert 空间的张量积也有泛性质. 首先定义映射

$$\begin{aligned} \rho: H\times V &\to H\otimes V \\ (\sum a_\alpha h_\alpha, \sum b_\beta v_\beta) &\mapsto \sum a_\alpha b_\beta h_\alpha\otimes v_\beta. \end{aligned} \tag{6.5}$$

容易验证,ρ 是一个连续双线性映射.

定理 6.6　给定 Hilbert 空间 H, V, 以及如上定义的 $H\otimes V$ 和连续双线性映射 ρ. 对任意的 Hilbert 空间 W 和连续双线性映射 $\eta: H\times V\to W$, 都存在唯一的连续线性映射 $T_\eta: H\otimes V\to W$, 使得下列图表交换:

$$\begin{tikzcd} H\times V \arrow[r, "\rho"] \arrow[dr, "\eta"'] & H\otimes V \arrow[d, "T_\eta"] \\ & W \end{tikzcd}$$

张量积是一个比较抽象的概念, 但对于一些特殊的 Hilbert 空间, 可以利用泛性质说明张量积同构于更加具体的空间. 以下命题说明两个空间上的平方可积空间的张量积同构于乘积空间上的平方可积空间.

定理 6.7　设 $(\Omega_i, \mathscr{F}_i, \mu_i)$ 是两个 σ 有限的测度空间. 则有以下等距同构:

$$\begin{aligned} L^2(\Omega_1,\mathscr{F}_1,\mu_1)\otimes L^2(\Omega_2,\mathscr{F}_2,\mu_2) &\overset{\sim}{\longrightarrow} L^2(\Omega_1\times\Omega_2,\mathscr{F}_1\otimes\mathscr{F}_2,\mu_1\times\mu_2) \\ f\otimes g &\mapsto ((x,y)\mapsto f(x)g(y)) \end{aligned}$$

下面给出可分 Hilbert 张量积的同构刻画, 为此先引入 HS 算子 (Hilbert-Schmidt operator) 的概念.

定义 6.6　给定可分 Hilbert 空间 H, V. 称算子 $T \in \mathscr{L}(H;V)$ 是一个 HS 算子, 是指: 存在一个规范正交基 $(e_n)_{n\in\mathbb{N}}$, 使得

$$\sum_{n=1}^{\infty} \|Te_n\|_V^2 < \infty.$$

若 T 是 HS 算子, 定义其 HS 范数为 $\|T\|_{\mathrm{HS}} := (\sum_{n=1}^{\infty} \|Te_n\|_V^2)^{\frac{1}{2}}$.

注　实际上,HS 算子 T 的 HS 范数不依赖于规范正交基的选取, 因而是标准的. 易见, 全体 HS 算子在通常的加法和数乘之下是一个向量空间, 记为 $\mathscr{L}_{\mathrm{HS}}(H;V)$ 或 $\mathscr{H}(H;V)$, 并配备 HS 范数.

注　设 T 是一个 HS 算子, 若

$$\|T\|_1 := \sup \sum_n |\langle Tf_n, g_n\rangle_V| < \infty,$$

其中上确界是取遍所有 H 的规范正交基 $\{f_n\}$ 和 V 的规范正交基 $\{g_n\}$, 则称 T 为迹算子,$\|T\|_1$ 称为迹范数.

特别地, 若 $H = V$, 则 $\sum_n |\langle Tf_n, f_n\rangle|$ 不依赖于基的选取,$\mathrm{Tr}(T) := \sum_n |\langle Tf_n, f_n\rangle|$ 称为算子 T 的迹.

定理 6.8　给定可分 Hilbert 空间 H, V,$(e_n), (v_m)$ 分别是 H, V 的规范正交基. 则有以下等距同构:

$$H \otimes V \xrightarrow{\sim} \mathscr{L}(H;V)$$
$$\sum_{n,m} a_{n,m} h_n \otimes v_m \longmapsto \sum_{n,m} a_{n,m} \langle \cdot, h_n\rangle v_m$$

以上同构已经给出了一种比较具体的视角, 我们希望对多元张量积也有上述定理中的同构, 从而给出更便于操作和理解的实现. 以下观点和构造来自于 (Reed et al., 1980) 和 (黄志远 等, 1997), 在本章之后内容中, 总是采取这种观点.

考虑 Hilbert 空间 $H_1, , H_2$, 对任意 $h_i \in H_i, i = 1, 2$, 定义 $h_1 \otimes h_2$ 为乘积空间 $H_1 \times H_2$ 上的双线性函数:

$$(h_1 \otimes h_2)(x_1, \cdots, x_2) = \Pi_{i=1}^2 \langle h_i, x_i\rangle_{H_i}, x_i \in H_i,$$

记 $\mathscr{E}$ 为所有形如 $h_1 \otimes h_2$ 的线性张成, 在其上定义内积:

$$\langle h_1 \otimes h_2, l_1 \otimes l_2\rangle_{\mathscr{E}} = \Pi_{i=1}^2 \langle h_i, l_i\rangle_{H_i},$$

而 $(H_1 \otimes H_2, \langle\cdot,\cdot\rangle_{H_1\otimes H_2})$ 则定义为 $(\mathscr{E}, \langle\cdot,\cdot\rangle_{\mathscr{E}})$ 的完备化, 可以证明若 $\{e_m^{(i)}\}$ 为 H_i 的规范正交基, 则 $\{e_{m_1}^{(1)} \otimes e_{m_2}^{(2)}\}$ 为 $H_1 \otimes H_2$ 的规范正交基. 对于多元情形则有如下定义.

定义 6.7　给定 Hilbert 空间 $H_1,\cdots,H_n$,$\{e_m^{(i)}\}$ 为 H_i 的规范正交基. 对乘积空间 $H_1\times\cdots\times H_n$ 上的任一 n 重线性函数 F, 定义其 HS 范数为 $\|F\|_{\mathrm{HS}}$,

$$\|F\|_{\mathrm{HS}}=\sum_{k_1,\cdots,k_n\in\mathbb{N}}|F(e_{k_1}^{(1)},\cdots,e_{k_n}^{(n)})|^2,$$

$H_1,\cdots,H_n$ 的张量积即为全体 HS 范数有限的 n 重线性函数.

特别地, 当 $H_1=\cdots=H_n=H$ 时, 记其张量积为 $H^{\otimes n}$.

这个定义与二元情形是统一的. 任取 $h\otimes v$, 它的 HS 范数为

$$\begin{aligned}\|h\otimes v\|_{\mathrm{HS}}^2&=\sum_{i,j}|(h\otimes v)(e_i^{(1)},e_j^{(2)})|^2\\&=\sum_{i,j}|\langle h,e_i^{(1)}\rangle\langle v,e_j^{(2)}\rangle|^2\\&=(\sum_i\langle h,e_i^{(1)}\rangle^2)(\sum_j\langle v,e_j^{(2)}\rangle^2),\end{aligned}$$

由 Parsaval 等式知,$\|h\otimes v\|_{\mathrm{HS}}=\|h\|_{H_1}\|v\|_{H_2}$.

设 S_n 为 n 阶置换群, 设 $\sigma\in S_n$, 对简单张量定义算子 $U_\sigma(h_1\otimes\cdots\otimes h_n)=h_{\sigma(1)}\otimes\cdots\otimes h_{\sigma(n)}$, 并将该算子延拓到全空间, 仍记为 U_σ. 然后定义对称化算子

$$\mathrm{Sym}=\frac{1}{n!}\sum_{\sigma\in S_n}U_\sigma,$$

和反对称化算子

$$\mathrm{Alt}=\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)U_\sigma$$

算子 Sym 和 Alt 的值域中的元素分别称为对称张量和交错张量. 记 $H^{\odot n}=\mathrm{Sym}(H^{\otimes n}),H^{\wedge n}=\mathrm{Alt}(H^{\otimes n})$, 它们上的内积取为张量积空间上内积的 $n!$ 倍.

6.2　Hermite 多项式与 Wiener 混沌分解

如无特别说明, 以下讨论的随机变量都是定义在一个完备的概率空间 $(\Omega,\mathscr{F},\mathbb{P})$ 上.

定义 6.8　给定 Hilbert 空间 $\mathbb{H}$. 称随机过程 $W=\{W(h):h\in\mathbb{H}\}$ 是一个等距 Gauss 过程 (或 $\mathbb{H}$ 上的 Gauss 过程), 是指:

(1) $\forall h\in\mathbb{H},\mathbb{E}[W(h)]=0$;

(2) $\forall n\in\mathbb{N},h_1,\cdots,h_n\in\mathbb{H},(W(h_1),\cdots,W(h_n))$ 的联合分布是多元正态分布;

(3) $\forall h,g\in\mathbb{H},\mathbb{E}[W(h)W(g)]=\langle h,g\rangle_{\mathbb{H}}$. W 生成的 σ 代数 $\sigma(W)$ 记为 $\mathscr{G}$.

注　显然, 等距 Gauss 过程中的单个随机变量 $W(h)$ 二阶矩和方差为 $\mathbb{E}[W(h)^2]=\|h\|_{\mathbb{H}}^2$, 且映射 $h\mapsto W(h)$ 是线性的, 也即: 任意 $\lambda,\mu\in\mathbb{R},h,g\in\mathbb{H}$, 都有:

$$W(\lambda h+\mu g)=\lambda W(h)+\mu W(g),\mathbb{P}-\text{a.s.}.$$

以下命题在后续将经常用到.

命题 6.9　$\{e^{W(h)}: h \in \mathbb{H}\}$ 是 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 的一个全子集.

定义 6.9　定义 n 次 Hermite 多项式为

$$H_n(x) = \frac{(-1)^n}{n!} e^{\frac{x^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}}, n \geq 1,$$

特别地, 规定 $H_0(x) = 1$.

注　令二元函数 $F(x,t) = e^{tx - \frac{t^2}{2}}$, 对变量 t 做 Taylor 展开, 则

$$F(x,t) = \sum_{n=1}^{\infty} t^n H_n(x),$$

由以上表示容易得到 Hermite 多项式的几个基本性质:

(1) $H_n'(x) = H_{n-1}(x)$;

(2) $(n+1)H_{n+1}(x) = xH_n(x) - H_{n-1}(x)$;

(3) $H_n(-x) = (-1)^n H_n(x)$;

(4) 若 n 为奇数, 则 $H_n(0) = 0$; 若 $n = 2k$ 为偶数, 则 $H_n(0) = \frac{(-1)^k}{2^k k!}$;

(5) $H_n(x)$ 的最高次项为 $\frac{x^n}{n!}$.

根据以上性质, 实际可以归纳计算 $H_n(x)$, 如 $H_1(x) = x, H_2(x) = \frac{1}{2}(x^2 - 1)$ 等.

Hermite 多项式不是随意定义的, 它和正态分布关系密切, 具体体现为以下命题.

命题 6.10　设 (X,Y) 是二元正态的,$\mathbb{E}[X] = \mathbb{E}[Y] = 0, \mathbb{E}[X^2] = \mathbb{E}[Y^2] = 1$. 则任意 $n, m \in \mathbb{N}$,

$$\mathbb{E}[H_n(X)H_m(Y)] = \begin{cases} 0, & n \neq m \\ \frac{(\mathbb{E}[XY])^n}{n!}, & n = m. \end{cases}$$

命题的证明只需考虑 $\mathbb{E}[F(X,s)F(Y,t)]$ 和零均值的多元正态分布矩母函数为 $M(t) = \exp(\frac{1}{2}t^T \Sigma t)$, 其中 Σ 为协方差矩阵.

借助 Hermite 多项式, 可以给出 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 的直和分解. 首先定义 $\mathscr{H}_0$ 表示常值随机变量全体,$\mathscr{H}_n$ 为 $\{H_n(W(h)) : h \in \mathbb{H}, \|h\|_{\mathbb{H}} = 1\}$ 生成的闭线性子空间,$n \geq 1$, 由前述命题, 当 $n \neq m$,$\mathscr{H}_n \perp \mathscr{H}_m$. 称 $\mathscr{H}_n$ 为 n 阶 Wiener 混沌 (Wiener chaos),J_n 为 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 到 $\mathscr{H}_n$ 的正交投影.

定理 6.11 (Wiener 混沌分解)

$$L^2(\Omega, \mathscr{G}, \mathbb{P}) = \oplus_{n=1}^{\infty} \mathscr{H}_n.$$

证明　设 $X \in L^2(\Omega, \mathscr{G}, \mathbb{P})$, 任意 n,$X \perp \mathscr{H}_n$, 往证 X=0,

注意到多项式 x^n 可以表示为有限个 Hermite 多项式的线性组合, 所以 $\mathbb{E}[X(W(h))^n] = 0$, 从而 $\mathbb{E}[Xe^{tW(h)}] = 0$, $\forall t \in \mathbb{R}, h \in \mathbb{H}, \|h\| = 1$, 这等价于 $\mathbb{E}[Xe^{W(h)}] = 0, \forall h \in \mathbb{H}$. 由于 $\{e^{W(h)}\}$ 是 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 的全子集, 所以 $X = 0$.　■

根据直和分解的性质, 任意 $X \in L^2(\Omega, \mathscr{H}, \mathbb{P})$,$\sum_{k=1}^{n} J_k X \to X, n \to \infty$ 且

$$\mathbb{E}[X^2] = \sum_{n=1}^{\infty} \mathbb{E}[|J_k X|^2].$$

记 $\mathcal{P}_n^0$ 为形如 $p(W(h_1), \cdots, W(h_n))$ 的随机变量全体, 其中 $k \in \mathbb{N}$,$h_1, \cdots, h_n \in \mathbb{H}$, p 是一个次数不高于 n 的多元多项式,$\mathcal{P}$ 表示 $\mathcal{P}_n^0$ 在 L^2 中的闭包, 则 $\mathcal{P}_n = \mathscr{H}_0 \oplus \mathscr{H}_1 \oplus \cdots \oplus \mathscr{H}_n$.

下面考虑一个简单的例子.

例 6.1　令 $(\Omega, \mathscr{F}, \mathbb{P}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}), \gamma)$, 其中 γ 为标准正态分布, 任意 $A \in \mathscr{B}(\mathbb{R})$,

$$\gamma(A) = \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x,$$

Hilbert 空间 $\mathbb{H}$ 取为 $\mathbb{R}$, 等距 Gauss 过程取为 $W(h)(x) = x$. 注意到 $\mathbb{R}$ 中模长为 1 的只有 $1, -1$, 且 $H_n(x) = (-1)^n H_n(-x)$, 所以 $\mathscr{H}_n$ 是 $H_n(x)$ 生成的一维线性子空间, 这实际上也说明 Hermite 多项式是 $L^2(\mathbb{R}, \gamma)$ 中的一个规范正交基.

设 $\mathbb{H}$ 是无穷维的 Hilbert 空间,$\{e_i\}$ 是一个规范正交基. 记 $\Lambda = \oplus_{n=1}^{\infty} \mathbb{N}$. 对 $a \in \Lambda$, 定义 $a! = \Pi_{n=1}^{\infty}(a_i!)$, $|a| = \sum_{n=1}^{\infty} a_i$, 根据直和定义, 这两个量分别是有限积和有限和, 因而是良定的. 定义扩展的 Hermite 多项式为

$$H_a(x) = \Pi_{n=1}^{\infty} H_{a_i}(x_i), x \in \mathbb{R}^{\mathbb{N}}, a \in \Lambda.$$

以及

$$\Phi_a = \sqrt{(}a!) \Pi_{n=1}^{\infty} H_{a_i}(W(e_i)), a \in \Lambda.$$

而 $\{\Phi_a : a \in \Lambda\}$ 是一个规范正交系, 即元素都为 1 且两两正交. 特别地,$\{\Phi_a : |a| = n\}$ 还是 $\mathscr{H}_n$ 的一个规范正交基.

$\mathbb{H}$ 上的 n 次对称张量空间 $(H^{\odot n}, \sqrt{n!}\|\cdot\|_{\mathbb{H}^{\otimes n}})$ 和 $\mathscr{H}_n$ 间有一个自然的同构:

$$\mathrm{Sym}(e^{\otimes a}) \mapsto \sqrt{a!} \Phi_a,$$

其中 $e^{\otimes a} := \otimes_{n=1}^{\infty} e_i^{\otimes a_i}, a = (a_1, a_2, \cdots) \in \Lambda, |a| = n$

注　对任意 Hilbert 空间 V,$L^2(\Omega; V)$ 也有类似的分解:

$$L^2(\Omega; V) = \oplus_{n=1}^{\infty} \mathscr{H}_n(V),$$

其中 $\mathscr{H}_n(V)$ 是由形如 $\sum_{i=1}^{n} F_i v_i, F_i \in \mathscr{H}_n . v_i \in V$ 的 V 值随机变量生成的闭子空间, 并且容易得到 $\mathbb{H}^{\odot n} \otimes V$ 和 $\mathscr{H}_n(V)$ 之间是同构的.

6.3　Malliavin 导数算子及其相关性质

在本节和下一节中取 $\mathscr{F} = \mathscr{G}$.

记 $C_p^\infty(\mathbb{R}^n)$ 表示 $\mathbb{R}^n$ 上本身及其各阶导数都只有至多多项式增长的无穷可微函数. 记 $\mathrm{Srv}(\mathbb{R})$ 为具有以下形式的随机变量 F 全体: 存在 $n \in \mathbb{N}, h_1, \cdots, h_n, f \in C_p^\infty(\mathbb{R}^n)$, 使得

$$F = f(W(h_1), \cdots, W(h_n)).$$

Srv 中的元素称为光滑随机变量. 若 f 及其各阶偏导数还是有界的, 则记 $F \in \mathrm{Srv}_b(\mathbb{R})$; 若 f 还有紧支集, 则记 $F \in \mathrm{Srv}_0(\mathbb{R})$. 若 f 是多项式, 记 $F \in \mathcal{P}$. 当只出现 Srv 时, 就是指 $\mathrm{Srv}(\mathbb{R})$, Srv_b, Srv_0 同理.$\mathrm{Srv}_0, \mathcal{P}$ 都是 $L^2(\Omega)$ 的稠密子集.

首先对光滑随机变量定义导数.

定义 6.10 设 $h_1, \cdots, h_n \in \mathbb{H}, f \in C_p^\infty(\mathbb{R}^n)$,$F = f(W(h_1), \cdots, W(h_n))$ 的导数定义为如下 $\mathbb{H}$ 值随机变量:

$$DF = \sum_{i=1}^n \partial_i f(W(h_1), \cdots, W(h_n)) h_i.$$

注 同一个光滑随机变量可能有不同的表示形式, 最简单的例如 $W(h) = \frac{1}{2}W(2h), W(h_1) + W(h_2) = W(h_1 + h_2)$, 这两个例子展示的表示的不同是由线性性带来的, 我们推至一般的情形. 将 $h_1, \cdots, h_n$ 视为列向量 $(h_1, \cdots, h_n)^T$. 记 $\vec{h} = (h_1, \cdots, h_n)^T$, 借用矩阵乘法的记号和规则,$DF$ 可记为 $(\nabla f)\vec{h}$.

假设存在矩阵 $A \in \mathbb{R}^{n\times m}$ 和 $\vec{l} = (l_1, \cdots, l_m)$ 使得 $\vec{h} = A\vec{l}$. 记 $\tilde{f} = f \circ A \in C_p^\infty(\mathbb{R}^m)$. 且 $F = f(W(h_1), \cdots, W(h_n)) = \tilde{f}(W(l_1), \cdots, W(l_m))$ 给出了两种不同的表示, 但实际上

$$\begin{aligned}(\nabla \tilde{f}(W(l_1), \cdots, W(l_m)))\vec{l} &= (\nabla f(A(W(l_1, \cdots, W(l_n)))) A\vec{l} \\ &= (\nabla f(W(h_1), \cdots, W(h_n)))\vec{h},\end{aligned}$$

其中第二个等号利用了映射 $h \mapsto W(h)$ 的线性性. 另外, 若对线性无关的 $h_1, \cdots, h_n$, 有 $F = f(W(h_1), \cdots, W(h_n)) = g(W(h_1), \cdots, W(h_n))$, 则 $f = g$, 这是因为当 $h_1, \cdots, h_n$ 线性无关时,$W(h_1), \cdots, W(h_n)$ 诱导了非退化的多元正态分布.

由以上注记, 还可以给出光滑随机变量的乘积的导数公式. 对 $F = f(W(h_1), \cdots, W(h_n)), G = g(W(l_1), \cdots, W(l_k))$, 定义 $\mathbb{R}^{n+k}$ 上的函数 $\tilde{f}$ 和 $\tilde{g}$ 如下:

$$\begin{aligned}\tilde{f}(x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+k}) &= f(x_1, \cdots, x_n) \\ \tilde{g}(x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+k}) &= g(x_{n+1}, \cdots, x_{n+k}),\end{aligned}$$

则

$$FG = \tilde{f}\tilde{g}(W(h_1), \cdots, W(h_n), W(l_1), \cdots, W(l_k)),$$

计算即得 $D(FG) = F(DG) + G(DF)$.

依定义, 对于光滑随机变量 $F=f(W(h_1),\cdots,W(h_n))$, 导数 DF 是 $\mathbb{H}$ 值的随机变量, 自然关心它和 $\mathbb{H}$ 中的元素做内积, 对任意 $h\in\mathbb{H}$,

$$\begin{aligned}\langle DF,h\rangle&=\langle\sum_{i=1}^n\partial_i f(W(h_1),\cdots,W(h_n))h_i,h\rangle\\&=\sum_{i=1}^n\partial_i f(W(h_1),\cdots,W(h_n))\langle h_i,h\rangle,\end{aligned}$$

所以 $\langle DF,h\rangle$ 是 f 在 $(W(h_1),\cdots,W(h_n))$ 处沿 $(\langle h_1,h\rangle),\cdots,\langle h_n,h\rangle)$ 的方向导数. 并且对该方向导数还有以下定理.

定理 6.12　设 $F\in$ Srv,$h\in\mathbb{H}$. 则

$$\mathbb{E}[\langle DF,h\rangle]=\mathbb{E}[FW(h)].$$

证明　不妨设 $\|h\|=1$,$F=f(W(h_1),\cdots,W(h_n))$. 而 $h_1,\cdots,h_n,h$ 总能通过正交化得到一个极大线性无关的规范正交系, 记为 $h,e_2,\cdots,e_k$, 再由线性变换可得 $F=g(W(h),W(e_2),\cdots,W(e_k))$, 这样,

$$DF=\partial_1 g(W(h),W(e_2),\cdots,W(e_k))h+\sum_{i=2}^k\partial_i g(W(h),W(e_2),\cdots,W(e_k))e_i,$$

从而 $\langle DF,h\rangle=\partial_1 g(W(h),W(e_2),\cdots,W(e_k))$.

记 $\phi_k(x)=\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{|x|^2}{2}}$ 表示 k 元标准正态分布的概率密度函数. 则有

$$\begin{aligned}\mathbb{E}[\langle DF,h\rangle]&=\int_{\mathbb{R}^k}\partial_1 g(x)\phi_k(x)\mathrm{d}x\\&=\int_{\mathbb{R}^k}x_1 f(x)\phi_k(x)\\&=\mathbb{E}[FW(h)].\end{aligned}$$

■

.

用 FG 代替上述定理的 F, 就得到了常用的分部积分公式:

$$\mathbb{E}[FGW(h)]=\mathbb{E}[F\langle DG,h\rangle]+\mathbb{E}[G\langle DF,h\rangle],F,G\in\mathrm{Srv},h\in\mathbb{H}.\tag{6.6}$$

下面将到导数算子的定义域从 Srv” 延拓” 到更大的集合上. 为此, 先介绍以下概念.(见 (Yosida, 2022) 第二章第 6 节)

定义 6.11　给定两个 Banach 空间 X,Y, 在乘积空间上定义范数

$$\|(x,y)\|=(\|x\|^2+\|y\|^2)^{\frac{1}{2}},$$

则乘积空间在此内积下是一个 Banach 空间.

给定线性映射 $A:X\to Y$,$\Gamma(A)=\{(x,Ax):x\in\mathrm{Dom}(A)\}\subset X\times Y$ 称为算子 A 的图.

称 A 为闭算子, 是指:$\Gamma(A)$ 是 $X \times Y$ 的线性闭子空间; 称 A 为可闭算子, 是指:$\Gamma(A)$ 在 $X \times Y$ 中的闭包 $\overline{\Gamma(A)}$ 是某线性算子的图, 也即 A 有一个闭延拓, 此时, 将该延拓记为 $\bar{A}$.

对于闭算子和可闭算子的判别有以下充要条件.

定理 6.13

(1) 算子 A 是闭算子, 当且仅当以下蕴含关系成立:

$$(\{x_n\} \subset \mathrm{Dom}(A), x_n \to x, Ax_n \to y) \implies (x \in \mathrm{Dom}(A), Ax = y).$$

(2) 算子 A 是可闭的, 当且仅当以下蕴含关系成立:

$$(\{x_n\} \subset \mathrm{Dom}(A), x_n \to 0, Ax_n \to y) \implies (y = 0)$$

定理 6.14　导数算子 D 是 $L^p(\Omega)$ 到 $L^p(\Omega; \mathbb{H})$ 的可闭算子.

注　在证明之前对该定理做出一点补充说明. 由于正态分布有任意阶矩, 所以任一 $F \in \mathrm{Srv}$, 总有 $F \in L^p(\Omega), \forall p \geq 1$. 而其导数 DF 作为一个 $\mathbb{H}$ 值随机变量,$\|DF\|_{\mathbb{H}}$ 是一个实值随机变量也有 p 阶矩, 只考虑固定的 $p \geq 1$, 在该定理之前, 导数算子 D 的定义域只有 Srv, 而该定理说明 D 作为 $\mathrm{Srv} \subset L^p(\Omega)$ 到 $L^p(\Omega; \mathbb{H})$ 的线性算子存在一个闭延拓, 使之成为一个闭算子, 而闭图像定理又说明闭线性算子总是有界的.

证明　为证 D 为可闭算子, 任取 $\{F_n\} \subset \mathrm{Srv}$, 满足 $F_n \overset{L^p}{\to} 0$ 且存在 $\eta \in L^(\Omega; \mathbb{H})$ 使得 $DF_n \overset{L^p}{\to} \eta$. 往证 $\eta = 0$.

对任意 $h \in \mathbb{H}$ 和光滑随机变量 F,

$$\begin{aligned}\mathbb{E}[\langle \eta, h \rangle F] &= \lim_{n\to\infty} \mathbb{E}[\langle DF_n, h \rangle F] \\ &= \lim_{n\to\infty} \mathbb{E}[-F_n \langle F, h \rangle + F_n F W(h)] \\ &= 0,\end{aligned}$$

由光滑随机变量的稠密性可得,$\eta = 0$. 也就证明了 D 是可闭的. ∎

对不同的 $p \geq 1$, 算子 D 的闭延拓的定义域是不同的. 对于固定的 $p \geq 1$,D 的闭延拓仍记为 D, 定义域记为 $\mathbb{D}^{1,p}$, 实际上,$\mathbb{D}^{1,p}$ 是 Srv$\subset L^P(\Omega)$ 在范数 $\|\cdot\|_{1,p}$ 下的闭包, 其中

$$\|F\|_{1,p} := (\mathbb{E}[|F^p|] + \mathbb{E}[\|DF\|_{\mathbb{H}}^p])^{\frac{1}{p}},$$

1 表示导数的阶数,p 表示算子从哪个空间出发. 更具体地说,$\mathbb{D}^{1,p}$ 是满足以下条件的子空间: 若 $F_n \overset{L^p}{\to} F$, 则 DF_n 在 $L^p(\Omega; \mathbb{H})$ 收敛到某 $\mathbb{H}$ 值的随机变量.

特别地, 当 $p = 2$,$\mathbb{D}^{1,2}$ 在内积 $\langle \cdot, \cdot \rangle_{1,2}$ 下成为一个 Hilbert 空间, 其中

$$\langle F, G \rangle := \mathbb{E}[FG] + \mathbb{E}[\langle DF, DG \rangle], F, G \in \mathbb{D}^{1,2}.$$

由此立见 $\mathbb{D}^{1,2}$ 是自反的. 实际上, 任意 $p \in (1,\infty)$,$\mathbb{D}^{1,p}$ 同构于 $L^p(\Omega) \times L^p(\Omega;\mathbb{H})$ 的一个闭子空间, 由 Pettis 定理知,$\mathbb{D}^{1,p}$ 也是自反的.

对于光滑随机变量 $F = f(W(h_1),\cdots,W(h_n)) \in \mathrm{Srv}$, 可以定义其高阶 Malliavin 导数.

定义 6.12　给定 $F = f(W(h_1),\cdots,W(h_n)) \in \mathrm{Srv}$, 其 k 阶 Malliavin 导数定义为如下 $\mathbb{H}^{\otimes k}$ 值随机变量

$$D^kF := \sum_{1\le i_1,\cdots,i_k\le n} \frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}} F(W(h_1),\cdots,W(h_n))h_{i_1}\otimes h_{i_k},$$

注　在多元微积分中, 由于光滑函数偏导算子的可交换性, 常使用多重指标简化记号, 但由于一般张量不是对称的, 所以有必要说明先后次序.

类似于一阶导数算子, 高阶导数算子也是可闭的, 记 $\mathbb{D}^{k,p}$ 表示 p 阶导数算子 D^k 在 $L^p(\Omega)$ 中的闭延拓的定义域, 也即 Srv 关于范数 $\|\cdot\|_{k,p}$ 的完备化, 其中

$$\|F\|_{k,p} := (\mathbb{E}[|F|^p] + \sum_{i=1}^{k} \mathbb{E}[\|D^iF\|_{\mathbb{H}^{\otimes i}}])^{\frac{1}{p}}$$

容易验证 $\mathbb{D}^{k+1,p} \subset \mathbb{D}^{k,p} \subset \cdots \subset \mathbb{D}^{1,p} \subset \mathbb{D}^{0,p} = L^p(\Omega)$①,$\forall p \ge 1$.

对于给定的 $h \in \mathbb{H}$, 定义算子 $D_hF := \langle DF, h\rangle$, 这也是 $L^p(\Omega)$ 上的可闭泛函, 记 $\mathbb{D}_h^p$ 表示其在 $L^p(\Omega)$ 中的闭延拓的定义域.

$\mathbb{D}^{1,2}$ 作为 Srv 关于范数 $\|\cdot\|_{1,2}$ 的闭包, 是比较抽象的, 不像光滑随机变量有显式的表达式, 所以希望找到一些充分条件或必要条件来辅助判断一个随机变量是否属于 $\mathbb{D}^{1,2}$, 也即希望找到平方可积随机变量的可微判别准则. 事实上有以下定理给出 $L^2(\Omega)$ 中元素可微的充要条件.

定理 6.15 (平方可积随机变量可微的充要条件)　设 $F \in L^2(\Omega)$, 其 Wiener 混沌分解为 $F = \sum_{n=1}^{\infty} J_nF$. 则 $F \in \mathbb{D}^{1,2}$, 当且仅当

$$\sum_{n=1}^{\infty} n\mathbb{E}[|J_nF|^2] < \infty. \tag{6.7}$$

此时,$\mathbb{E}[\|DF\|_{\mathbb{H}}^2] = \sum_{n=1}^{\infty} n\mathbb{E}[|J_nF|^2] < \infty.D(J_nF) = J_{n-1}(DF)$.

证明　证明需要多次使用空间的完备性和 Cauchy 列的性质.

先对固定的 $n, a \in \Lambda, |a| = n$ 考虑 $\Phi_a = \sqrt{a!}\prod_{i=1}^{\infty} H_{a_i}(W(e_i))$. 注意 Φ_a 只是有限个单变量多项式的乘积, 所以

$$D\Phi_a = \sum_{j=1}^{\infty} (H_{a_{j-1}}(W(e_j)) \prod_{i\ne j} H_{a_i}(W(e_i)))e_i,$$

且当 $a, b \in \Lambda, a \ne b$,$D\Phi_a$□$D\Phi_b$ 在 $L^2(\Omega;\mathbb{H})$ 中是正交的, 即 $\mathbb{E}[\langle D\Phi_a, D\Phi_b\rangle] = 0$.

①规定 $\|\cdot\|_{0,p} = \|\cdot\|_{L^p}$

另外有限 $\{e_1, \cdots, e_l\}$ 的正交性等价于 $W(e_1), \cdots, W(e_l)$ 的独立性, 所以可以计算

$$\mathbb{E}[\|D\Phi_a\|_{\mathbb{H}}^2] = \sum_{j=1}^{\infty} \left(\mathbb{E}[(H_{a_{j-1}}(W(e_j)))^2] \prod_{i \neq j} \mathbb{E}[(H_{a_i}(W(e_i)))^2]\right) = |a| = n.$$

下面说明: 若随机变量 $G \in \mathscr{H}_n \cap \mathbb{D}^{1,2}$, 则导数算子和极限可交换, 具体表述如下:

$$DG = \sum_{a \in \Lambda, |a|=n} \langle G, \Phi_a \rangle D\Phi_a.$$

由于 $\{\Phi_a : a \in \Lambda, |a| = n\}$ 是可数的, 为了便于讨论, 重排为 $\{\Phi_1, \Phi_2, \cdots\}$, 则 $G = \sum_{i=1}^{\infty} \langle G, \Phi_i \rangle \Phi_i$. 记 $S_k = \sum_{i=1}^{k} \langle G, \Phi_i \rangle \Phi_i$, 则

$$\begin{aligned}
&\left\| \sum_{i=m}^{k} \langle G, \Phi_i \rangle D\Phi_i \right\|_{\mathbb{H}}^2 \\
\leq &\sum_{i=m}^{k} \langle G, \Phi_k \rangle^2 \|D\Phi_i\|_{\mathbb{H}} \\
= &n \sum_{i=m}^{k} \langle G, \Phi_k \rangle^2
\end{aligned}$$

由 $\{S_k\}$ 为 Cauchy 列,$\{DS_k\}$ 也是 Cauchy 列, 再结合导数算子的连续性, 即得 $DG = \sum_{i=1}^{\infty} DS_k$, 且 $\mathbb{E}[\|DG\|_{\mathbb{H}}^2] = n\mathbb{E}[|G|^2]$.

由以上说明可知,J_nF 是可微的, 且 $D(J_nF) \in \mathscr{H}_{n-1}(\mathbb{H})$.

记 $T_k = \sum_{i=1}^{k} D(J_iF)$,

$$\begin{aligned}
\mathbb{E}[\|T_k - T_m\|_{\mathbb{H}}^2] &= \mathbb{E}[\left\| \sum_{i=m+1}^{k} D(J_iF) \right\|_{\mathbb{H}}^2] \\
&\leq \sum_{i=m+1}^{k} \mathbb{E}[\|D(J_iF)\|_{\mathbb{H}}^2] \\
&= \sum_{i=m+1}^{k} n\mathbb{E}[|J_iF|^2]
\end{aligned}$$

由条件可得,$\{T_k\}$ 是一个 $L^2(\Omega; H)$ 中的 Cauchy 列, 结合闭算子的连续性即得

$$DF = \sum_{n=1}^{\infty} D(J_nF),$$

以及

$$\mathbb{E}[\|DF\|_{\mathbb{H}}^2] = \sum_{n=1}^{\infty} n\mathbb{E}[|J_nF|^2].$$

■

注　以上讨论可以逐次反复进行, 从而得到 $F \in \mathbb{D}^{k,2}$ 的充要条件为:

$$\sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \mathbb{E}[|J_nF|^2] < \infty.$$

对于复合函数的 Malliavin 导数有类似多元微积分的链式法则, 具体陈述如下.

定理 6.16　设 $p \geq 1, \phi : \mathbb{R}^m \to \mathbb{R}$ 连续可微, 且偏导数有界, $F = (F^1, \cdots, F^m), F^i \in \mathbb{D}^{1,p}, i = 1, \cdots, m$. 则 $\phi(F) \in \mathbb{D}^{1,p}$ 且

$$D\phi(F) = \sum_{i=1}^{m} \partial_i \phi(F) DF^i.$$

证明　先考虑 $\phi : \mathbb{R} \to \mathbb{R}$ 是光滑函数, $F = f(W(h_1), \cdots, W(h_n)) \in \mathrm{Srv}$. 此时, $\phi \circ f : \mathbb{R}^n \to R$ 属于 C_p^∞, 所以 $\phi(F) = (\phi \circ f)(W(h_1), \cdots W(h_n)) \in \mathrm{Srv}$, 且

$$\begin{aligned}
D\phi(F) &= \sum_{i=1}^{n} \partial_i (\phi \circ f)(W(h_1), \cdots W(h_n)) h_i \\
&= \sum_{i=1}^{n} \phi'(F) \partial_i f(W(h_1), \cdots W(h_n)) h_i \\
&= \phi'(F) DF.
\end{aligned}$$

设 $F \in \mathbb{D}^{1,p}$, 由 Srv 的稠密性, 存在一列光滑随机变量 $\{F_n\}$ 使得 F_n 在 $\|\cdot\|_{1,p}$ 下收敛到 F, 也即

$$\|F_n - F\|_{1,p} = (\mathbb{E}[|F_n - F|^p] + \mathbb{E}[\|DF_n - DF\|_{\mathbb{H}}^p])^{1/p} \to 0,$$

由此, $\phi(F_n) \xrightarrow{L^p} \phi$ 以及 $\{D\phi(F_n)\}$ 是 Cauchy 列, 结合算子 D 的闭性, 可得

$$\phi(F) \in \mathbb{D}^{1,p}, D\phi(F_n) \xrightarrow{L^p} D\phi(F),$$

$$\begin{aligned}
&\mathbb{E}[\|D\phi(F) - \phi'(F)DF\|_{\mathbb{H}}^p] \\
\leq &C_p(\mathbb{E}[\|D\phi(F) - D\phi(F_n)\|_{\mathbb{H}}^p] + \mathbb{E}[\|\phi'(F_n) - \phi'(F)DF\|_{\mathbb{H}}^p]) \\
\leq &C_p(\mathbb{E}[\|D\phi(F) - D\phi(F_n)\|_{\mathbb{H}}^p] + \mathbb{E}[\|\phi'(F_n)DF_n - \phi'(F_n)DF\|_{\mathbb{H}}^p] + \mathbb{E}[\|\phi'(F_n)DF - \phi'(F)DF\|_{\mathbb{H}}^p])
\end{aligned}$$

令 $n \to \infty$, 结合之前的收敛性, 即得 $D\phi(F) = \phi'(F)DF$.

去掉 ϕ 光滑的假设, 考虑函数 $\rho(x) = Ce^{\frac{1}{|x|^2-1}} I_{(0,1)}(x)$, 其中常数 C 使得 ρ 在整个实轴上积分为 1, 以及 $\rho_\epsilon(x) = \frac{1}{\epsilon}\rho(\frac{x}{\epsilon})$. 则 $\rho_\epsilon * \phi$ 是光滑的, 且导数有界, 由以上结论可得

$$D((\rho_\epsilon * \phi)(F)) = (\rho_\epsilon * \phi)'(F)DF = (\rho_\epsilon * \phi')(F)DF,$$

由于导数有界, 所以 $|\phi(x)| = |\phi(0) + \int_0^1 \phi'(tx)x\mathrm{d}t| \leq C(1 + |x|)$, 而磨光子的性质保证了逐点收敛, 由控制收敛定理可得 $(\rho_\epsilon * \phi)(F) \xrightarrow{L^p} \phi(F)$, 类似可得 $(\rho_\epsilon * \phi')(F)DF \xrightarrow{L^p} \phi'(F)DF$, 最后由算子 D 的闭性可得 $\phi(F) \in \mathbb{D}^{1,p}$, 且 $D\phi(F) = \phi'(F)DF$. 对于 ϕ 是多元函数的情形, 论证是类似的. ■

注　Malliavin 导数的链式法则虽然形式上与经典微积分中的链式法则相似, 但还是有一定局限性, 例如不能直接取 $\phi(x, y) = xy$ 来计算随机变量乘积的导

数, 尽管形式上看起来十分合理, 实际上一般的随机变量乘积未必是 Malliavin 可导的, 还需要合适的可积性条件. 其中一个原因是, 经典微积分中的导数定义是局部的, 固定一点处的导数只取决于局部性态, 而太远处的行为则影响不大, 但 Malliavin 导数则不然, 它依赖于全空间上的积分.

另外, 要求导数有界, 只是为了是 $\phi(F)$ 和 F 的导数有相同的可积性, 实际上可以考虑导数具有多项式增长, 依然证 $\phi(F)$ 可导, 但要求的导数 ϕ' 增长速度越快,$D\phi(F)$ 可积性越低

事实上, 注意到证明中对 ϕ 的关键要求是导数有界, 所以对 ϕ 的要求可以减弱到仅是 Lipschitz 连续的. 为了证明 Lipschitz 函数的链式法则, 需要以下引理.

引理 6.17　设 $p>1$,$\{F_n\}\subset\mathbb{D}^{1,p}$,$F_n\overset{L^p(\Omega)}{\longrightarrow}F$, 且

$$\sup_n\mathbb{E}[\|DF_n\|_{\mathbb{H}}^p]<\infty.$$

则 $F\in\mathbb{D}^{1,p}$ 且 DF_n 弱收敛到 DF.

证明　首先由于 F_n 在 $L^p(\Omega)$ 中按 L^p 范数收敛, 自然也是按范数有界的, 而当 $p\in(0,\infty)$,$\mathbb{D}^{1,p}$ 是自反的, 所以存在 $\{F_n\}$ 的一个子列 $\{F_{n_k}\}$ 弱收敛到 $\mathbb{D}^{1,p}$ 中某元素, 暂记为 G. 事实上,F_{n_k} 作为 $L^p(\Omega)$ 中的点列也弱收敛到 G, 从而 $F=G$. 也就有了 $F\in\mathbb{D}^{1,p}$.

对 $\{F_n\}$ 的任意子列应用以上讨论, 可得该子列有弱收敛的子列, 这也就说明 F_n 在 $\mathbb{D}^{1,p}$ 中弱收敛到 F, 也就证明了结论. ∎

定理 6.18　设对函数 $\phi:\mathbb{R}^m\to\mathbb{R}$, 存在常数 K, 使得对所有 $x,y\in\mathbb{R}^m$,

$$|\phi(x)-\phi(y)|\le K|x-y|.$$

$F=(F^1,\cdots,F^m),F^i\in\mathbb{D}^{1,p},i=1,\cdots,m.$

则 $\phi(F)\in\mathbb{D}^{1,p}$ 且存在一个有界的随机向量 $G=(G^1,\cdots,G^m)$, 使得

$$D\phi(F)=\sum_{i=1}^m G_iDF^i$$

注　一个 Lipschitz 函数几乎处处可导, 所以定理中的 G_i 几乎可以视为 $\partial_i\phi(F)$.

证明　为简单起见, 依然假设 $m=1$.

记 $\phi_n=p_{\frac{1}{n}}*\phi,n\ge1$, 则 ϕ_n 是光滑函数, 导数 $|\phi_n'|\le K$, 且在整个实轴上, $\phi(x)$ 一致收敛到 $\phi(x)$.

根据光滑函数的链式法则有 $D\phi_n(F)=\phi_n'(F)DF$.

先考虑等式左边. 由于 ϕ 是 Lipschitz 函数, 所以 $|\phi(x)|\le|\phi(0)|+|\phi(x)-\phi(0)|\le|\phi(0)|+K|x|$, 由此可得 $\phi_n(F)\overset{L^p}{\longrightarrow}\phi(F)$, 而 $\phi_n'(F)$ 的有界性给出了

$D\phi_n(F)$ 矩的一致有界性, 由以上引理可得 $\phi(F) \in \mathbb{D}^{1,2}$, 且 $D\phi_n(F)$ 在 $L^2(\Omega; H)$ 中弱收敛到 $D\phi(F)$.

再考虑等式右边. 导数的有界性给出了弱收敛和弱极限的存在性, 将此极限记为 G, 就得到了结论.

多元情形类似. ■

下面说明 Malliavin 导数与经典导数中" 一定条件下, 若导数为 0, 则函数为常值函数" 的平行对应结果. 为此, 先陈述以下技术性引理.

引理 6.19　随机变量族 $\{W(h)G - D_hG : h \in \mathbb{H}, G \in \mathrm{Srv}_b\} \cup \{1\}$ 是 $L^2(\Omega)$ 的全子集.

注　注意到 $W(h)G - D_hG$ 的形式是由分部积分公式自然得到的.

定理 6.20　设 $F \in \mathbb{D}^{1,1}$,$DF = 0$. 则 $F = \mathbb{E}[F]$.

证明　若 $F \in \mathbb{D}^{1,2}$, 由定理6.15立得结论成立.

取有界光滑函数 ψ_n 满足以下条件:

$$\psi_n(x) = \begin{cases} x, x \in [-n, n], \\ 0, x \notin (-n-1, n+1), \\ *, n < |x| < n+1, \end{cases}$$

以及光滑随机变量 $F_n \xrightarrow{\mathbb{D}^{1,1}} F$. 对任意 $h \in \mathbb{H}$,$G \in \mathrm{Srv}_b$,

$$\mathbb{E}[\psi_n(F_m)(W(h)G - D_hG)] = \mathbb{E}[GD_h(\psi_n(F_m))],$$

令 $n \to \infty$ 并结合上一引理即可. ■

作为以上定理的推论, 可以给出 $\sigma(W)$ 中的 0-1 律.

推论 6.21　设 $A \in \sigma(W)$. 则

$$I_A \in \mathbb{D}^{1,1} \Leftrightarrow \mathbb{P}(A) = 0 或 1.$$

下面考虑任意可分 Hilbert 空间 V 以及取值于 V 的随机变量的 Malliavin 导数. 首先定义 $\mathrm{Srv}(V)$ 为具有以下形式的 V 值随机变量 F 全体:

$$F = \sum_{i=1}^n F_i v_i, n \in \mathbb{N}, F_i \in \mathrm{Srv}(\mathbb{R}), v_i \in V,$$

对这样的随机变量, 定义其 k 阶导数为 $\mathbb{H}^{\otimes k} \otimes V$ 值的随机变量

$$D^kF = \sum_{i=1}^n (D^kF_i) \otimes v_i,$$

类似于实值情形, 可以证明 D^k 是 $L^p(\Omega; V)$ 到 $L^p(\Omega; \mathbb{H}^{\otimes k} \otimes V)$ 的可闭算子, 记其闭延拓的定义域为 $D^{k,p}(V)$, 范数 $\|\cdot\|_{k,p,V}$ 定义为:

$$\|F\|_{k,p,V} = (\mathbb{E}[\|F\|_V^p] + \sum_{j=1}^k \mathbb{E}[\|D^jF\|^p_{\mathbb{H}^{\otimes j} \otimes V}])^{\frac{1}{p}}.$$

6.4　导数算子的伴随–散度算子

定义 6.13　称 (无界) 算子 $\delta : L^2(\Omega;\mathbb{H}) \supset \mathrm{Dom}(\delta) \to L^2(\Omega)$ 为导数算子 D 的伴随, 是指:

(1) $u \in \mathrm{Dom}(\delta)$, 当且仅当存在常数 $C = C(u)$, 使得对任意 $F \in \mathbb{D}^{1,2}$,

$$\mathbb{E}[\langle DF, u\rangle_{\mathbb{H}}]^2 \le C\mathbb{E}[F^2];$$

(2) 若 $u \in \mathrm{Dom}(\delta)$, 则 $\delta(u)$ 由

$$\mathbb{E}[F\delta(u)] = \mathbb{E}[\langle DF, u\rangle_{\mathbb{H}}], F \in \mathbb{D}^{1,2}$$

唯一确定.

称 δ 为 (Malliavin) 散度算子.

注　散度算子是闭线性算子, 且易见 $\mathbb{E}[\delta(u)] = 0, \forall u \in \mathrm{Dom}(\delta)$. 下面使用记号 $\delta(u)$ 时, 总默认 $u \in \mathrm{Dom}(\delta)$.

先对 $L^2(\Omega;\mathbb{H})$ 中形式较为简单的元素进行讨论. 设 $u = \sum_{i=1}^n F_i h_i \in \mathrm{Srv}(\mathbb{H})$, 由导数算子的分部积分公式可得

$$\delta(u) = \sum_{i=1}^n F_i W(h_i) - \sum_{i=1}^n \langle DF_i, h_i\rangle_{\mathbb{H}},$$

实际上, 有一些文献也以上式为出发点来定义散度算子.

对于 $u \in \mathbb{D}^{1,2}(\mathbb{H})$, 其导数 Du 是一个 $\mathbb{H}\otimes\mathbb{H}$ 值的随机变量, 而 $\mathbb{H}\otimes\mathbb{H}$ 中的元素可以视为 HS 算子, 其同构由

$$e_i \otimes e_j \mapsto \langle \cdot, e_i\rangle e_j$$

给出. 在这种观点下, 对 $u = \sum_{i=1}^n F_i h_i \in \mathrm{Srv}(\mathbb{H})$,

$$Du = \sum_{i=1}^n DF_i \otimes h_i,$$

Du 作为 HS 算子作用在 $h \in \mathbb{H}$ 则是

$$\langle Du, h\rangle = \sum_{i=1}^n \langle DF_i, h\rangle_{\mathbb{H}} h_i,$$

等号左边的 $\langle Du, h\rangle$ 只是泛函分析中" 约定" 记号, 不代表真正的内积, 也记 $D_h u = \langle Du, h\rangle$. 而二阶 Malliavin 导数也是 $\mathbb{H}\otimes\mathbb{H}$ 值的随机变量: 对一个光滑随机变量 $G = g(W(v_1), \cdots, W(v_n))$,

$$D^2G = \sum_{i,j} \partial_{ij} g(W(v_1), \cdots, W(v_n)) v_i \otimes v_j,$$

D^2G 作为 HS 算子作用在 $h \in \mathbb{H}$ 上式可记为

$$\langle D^2G, h\rangle = \sum_{i,j} \partial_{ij} g(W(v_1), \cdots, W(v_n)) \langle h, v_i\rangle_{\mathbb{H}} v_j,$$

将其与 $l\in\mathbb{H}$ 做内积可得

$$\langle\langle D^2G,h\rangle,l\rangle_{\mathbb{H}}=\sum_{i,j}\partial_{ij}g(W(v_1),\cdots,W(v_n))\langle h,v_i\rangle_{\mathbb{H}}\langle l,v_j\rangle_{\mathbb{H}},$$

注意光滑函数二阶偏导的对称性, 实际上有 $\langle\langle D^2G,h\rangle,l\rangle_{\mathbb{H}}=\langle\langle D^2G,l\rangle,h\rangle_{\mathbb{H}}$. 结合以上讨论和分部积分公式计算可得,

$$D_h(\delta(u))-\delta(D_hu)=\langle u,h\rangle_{\mathbb{H}},\tag{6.8}$$

借用换位子的记号, 上式可记为 $\langle u,h\rangle_{\mathbb{H}}=[D_h,\delta](u)$, 其中 $[D_h,\delta]=D_h\circ\delta-\delta\circ D_h$, 将上式简称为对易关系. 由该对易关系, 可得以下重要结论, 它说明了 $\mathrm{Dom}(\delta)$ 包含了一类重要的 $\mathbb{H}$ 值随机变量.

定理 6.22　$\mathbb{D}^{1,2}(\mathbb{H})\subset\mathrm{Dom}(\delta)$. 若 $u,v\in\mathbb{D}^{1,2}(\mathbb{H})$, 则

$$\mathbb{E}[\delta(u)\delta(v)]=\mathbb{E}[\langle u,v\rangle_{\mathbb{H}}]+\mathbb{E}[\mathrm{tr}(Du\circ Dv)].$$

证明　若 $u,v\in\mathrm{Srv}(\mathbb{H})$, 由分部积分公式和对易关系可计算定理陈述中的等式成立, 并且给出了以下估计:

$$\mathbb{E}[\delta(u)^2]\le\|u\|_{1,2,\mathbb{H}}^2.$$

对任一 $u\in\mathbb{D}^{1,2}(\mathbb{H})$, 存在一列 $\{u^n\}\subset\mathrm{Srv}(\mathbb{H})$, 使得 $u^n\stackrel{L^2(\Omega;\mathbb{H})}{\longrightarrow}u$ 且 $Du^n\stackrel{L^2(\Omega;\mathbb{H}\otimes\mathbb{H})}{\longrightarrow}Du$, 所以由以上估计可得 $\delta(u^n)$ 在 $L^2(\Omega)$ 中收敛到某极限, 并且可以验证该极限就是 $\delta(u)$. ∎

前述对易关系只是对光滑随机变量而言, 实际上可以推至更一般的情形. 陈述如下.

命题 6.23　设 $u\in\mathbb{D}^{1,2}$ 且 $D_hu\in\mathrm{Dom}(\delta)$. 则 $\delta(u)\in\mathbb{D}_h^2$, 且成立对易关系

$$\langle u,h\rangle_{\mathbb{H}}=[D_h,\delta](u).$$

该命题可看作以下引理的推论.

引理 6.24　设 $G\in L^2(\Omega)$, 存在 $Y\in L^2(\Omega)$, 使得对所有的 $F\in\mathbb{D}^{1,2}$,

$$\mathbb{E}[G\delta(hF)]=\mathbb{E}[YF].$$

则 $G\in\mathbb{D}_h^2, D_hG=Y$.

注　引理的证明只需考虑对 Y,G 做 Wiener 混沌分解. 需要注意的是, 引理中出现了 $\delta(hF)$, 但 F 不一定是光滑随机变量, 所以暂时无法计算, 但下一命题说明在一定条件下可以像光滑随机变量那样作运算.

命题 6.25　设 $F\in\mathbb{D}^{1,2},u\in\mathrm{Dom}(\delta)$ 使得 $Fu\in L^2(\Omega;\mathbb{H})$. 则当 $F\delta(u)-\langle DF,u\rangle_{\mathbb{H}}\in L^2(\Omega)$ 时, $Fu\in\mathrm{Dom}(\delta)$, 且

$$\delta(Fu)=F\delta(u)-\langle DF,u\rangle_{\mathbb{H}}.$$

6.5　白噪声情形的导数算子和散度算子

前两节是在一般框架下定义了导数算子和梯度算子, 这节集中讨论一种重要模型, 即 $\mathbb{H}$ 取为一个测度空间上的平方可积函数的全体.

给定一个完备概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$. 设 $(T, \mathscr{B}, \mu)$ 是一个无原子的 σ 有限的测度空间,W 为 $L^2(T, \mathscr{B}, \mu)$ 上的等距 Gauss 过程, 此时称 W 为强度为 μ 的 Gauss 白噪声, 简称白噪声. 对任意 $f, g \in L^2(T)$, 有等距同构

$$\mathbb{E}[W(f)W(g)] = \int_T fg\mathrm{d}\mu.$$

若 $A \in \mathscr{B}, \mu(A) < \infty$, 则记 $W(A) = W(I_A)$. 特别地, 若 $A, B \in \mathscr{B}$ 测度有限且 $A \cap B = \emptyset$, 则 $\mathbb{E}[W(A)W(B)] = 0$, 即 $W(A), W(B)$ 为独立的 Gauss 随机变量. 对于一般的 $f \in L^2(T)$, 也将 $W(f)$ 形式地记为 $\int_T f\mathrm{d}W$. 另外, 当使用记号 $W(A)$ 时, 默认 $A \in \mathscr{B}, \mu(A) < \infty$.

6.5.1　多元 Wiener 积分

设 $m \geq 1, \mathscr{B}_0 = \{A \in \mathscr{B} : \mu(A) < \infty\}$. 首先定义乘积测度空间 $(T^m, \mathscr{B}^m, \mu^m)$ 上平方可积函数 $f \in L^2(T^m) = L^2(T^m, \mathscr{B}^m, \mu^m)$ 的 Wiener 积分. 类似于 Lebsgue 积分的做法, 先在一类简单但稠密的函数上定义. 令 $\mathscr{E}_m$ 为具有以下形式的简单函数 f 的集合

$$f(t_1, \cdots, t_m) = \sum_{\sigma \in S_m} a_\sigma I_{A_{\sigma(1)} \times \cdots \times A_{\sigma(m)}}(t_1, \cdots, t_m),$$

其中 $a_\sigma \in \mathbb{R}, A_1, \cdots, A_m \in \mathscr{B}_0$ 两两不交.

对于简单函数 $f = \sum_{\sigma \in S_m} a_\sigma I_{A_{\sigma(1)} \times \cdots \times A_{\sigma(m)}}$, 定义其 Wiener 积分为

$$I_m(f) = \sum_{\sigma \in S_m} a_\sigma W(A_1) \cdots W(A_m)$$

算子 I_m 有以下性质:

(1) I_m 是线性的;

(2) $I_m(f) = I_m(\mathrm{Sym}(f))$, 其中 Sym(f) 表示函数 f 的对称化:

$$\mathrm{Sym}(f)(t_1, \cdots, t_m) = \frac{1}{m!} \sum_{\tau \in S_m} f(t_{\tau(1)}, \cdots, t_{\tau(m)});$$

(3) 对 $f \in \mathscr{E}_m, g \in \mathscr{E}_q$,

$$\mathbb{E}[I_m(f)I_q(g)] = \begin{cases} 0, m \neq q, \\ m! \displaystyle\int_{T^m} \mathrm{Sym}(f)\mathrm{Sym}(g)\mathrm{d}\mu^m, m = q \end{cases}$$

对以上性质不做证明, 仅举一个简单的例子.

例 6.2　考虑正半轴上 Lebsgue 测度. 取 $A = [0, 1], B = [2, 3], f = I_{A \times B} + 2I_{B \times A}$. 首先, 对任意的 $x \geq 0, f(x, x) = 0$.

注意到 $f(1,2)=1$, 而 $f(2,1)=2$, 说明 f 不是对称的,f 的对称化 $\mathrm{Sym}(f)$ 为

$$\begin{aligned}\mathrm{Sym}(f)(x,y) &= \frac{1}{2}(f(x,y)+f(y,x)) \\ &= \frac{1}{2}(I_A(x)I_B(y)+2I_B(x)I_A(y)+I_A(y)I_B(x)+2I_B(y)I_A(x)) \\ &= \frac{3}{2}(I_{A\times B}(x,y)+I_{B\times A}(x,y))\end{aligned}$$

f 的 Wiener 积分为 $I_2(f)=W(A)W(B)+2W(B)W(A)=3W(A)W(B)=I_2(\mathrm{Sym}(f))$.

注　尽管这个例子十分简单, 但高维情形基本一致, 只是涉及到的组合更复杂. 从这个例子也可以看出, 简单函数的定义中实际已经有了对称化的操作, 其不对称性来源于系数, 所以简单函数的对称化, 实际上就是各加项系数的平均化.

简单函数在 L^2 中是稠密的, 证明中要用到 μ 无原子的条件. 无原子的好处之一是对任意的测度有限集 A, 以及任意的 $a\in(0,\mu(A))$, 都存在一个 A 的子集可测且其测度正好为 a. 在正半轴和 Lebsgue 测度的情形下较为直观.

由 Fubini 定理和 Minkowski 定理可得,

$$\mathbb{E}[I_m(f)^2]=m!\,\|\mathrm{Sym}(f)\|^2_{L^2(T^m)}\leq\|f\|^2_{L^2(T^m)}.$$

从而算子 I_m 可以延拓到整个 $L^2(T^m)$ 上.

对于 $f\in L^2(T^p), g\in L^2(T^q)$, 定义其 r-缩并为 $f\otimes_r g\in L^2(p+q-2r)$

$$f\otimes_r g(t_1,\cdots,t_{p-r},s_1,\cdots,s_{q-r})=\int_{T^r}f(t_1,\cdots,t_{p-r},x)g(s_1,\cdots,s_{q-r},x)\mu^r(\mathrm{d}x).$$

记 $f\tilde{\otimes}g=\mathrm{Sym}(f\otimes g), f\tilde{\otimes}_r g=\mathrm{Sym}(f\otimes_r g)$, 以及 $L^2_S(T^m)=\mathrm{Sym}(L^2(T^m))$.

借助缩并, 可以对任意维中平方可积函数的 Wiener 积分的乘积表示为若干 Wiener 积分的线性组合.

定理 6.26　设 $f\in L^2_S(T^p), g\in L^2_S(T^q)$. 则

$$I_p(f)I_q(g)=\sum_{r=0}^{p\wedge q}\binom{p}{r}\binom{q}{r}I_{p+q-2r}(f\otimes_r g).$$

上述定理一个重要且常用的特例是取 $q=1$, 此时有

$$I_p(f)I_1(g)=I_{p+1}(f\otimes g)+pI_{p-1}(f\otimes_1 g).$$

由该公式, 可以给出一类特殊的对称函数的递推公式,

$$I_{m+1}(f^{\otimes(m+1)})=I_m(f^{\otimes m})I_1(f)-(m-1)\|f\|_{L^2(T)}I_{m-1}(f^{\otimes(m-1)})$$

根据这个递推公式以及归纳法可以得到以下命题.

命题 6.27　设 $f\in L^2(T)$ 且 $\|f\|_{L^2}=1$. 则

$$H_m(W(f))=\frac{1}{m!}I_m(f^{\otimes m}).$$

记 $\overline{L_S^2(T^m)}$ 为 $L^2(T^m)$ 中对称函数生成的闭子空间. 由以上命题,n 阶 Wiener 混沌 $\mathscr{H}_n \subset I_n(\overline{L_S^2(T^n)})$, 再结合不同阶的 Wiener 积分的正交性可得 $\mathscr{H}_n = I_n(\overline{L_S^2(T^n)})$. 从而有以下定理.

定理 6.28　设 $F \in L^2(\Omega, \sigma(W), \mathbb{P})$. 则存在 $f_n \in L_S^2(T^n), n \geq 1$, 使得

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} I_n(f_n)$$

6.5.2　作为随机过程的导数

考虑简单情形 $F = W(f)^k, f \in L^2(T)$, 则 $DF = kW(f)^{k-1}f$,$W(f)$ 是一个实值随机变量, 而 f 是一个 T 上的函数. DF 作为 $L^2(\Omega; L^2(T))$ 中的元素是说: 将任意 $\omega \in \Omega$ 映为一个平方可积函数 $k(W(f)(\omega))^{k-1}f$, 该函数在 $t \in T$ 处取值为 $k(W(f)(\omega))^{k-1}f_t$. 另一方面, 也可以说对任一 $t \in T$, 由导数确定了一个随机变量 $kf_tW(f)^{k-1}$, 记为 D_tF. 两种观点只是先后顺序的差异. 对 DF 再求导得 $D^2F = k(k-1)W(f)^{k-2}f \otimes f$, 这是一个 $L^2(T^2)$ 值随机变量, 也就是说需要给定两个参数 $t_1, t_2 \in T$ 才能确定一个实值随机变量 $k(k-1)f_{t_1}f_{t_2}W(f)^{k-2}$, 记为 $D^2_{t_1,t_2}F$, 反过来说, 随机变量族 $\{D^2_{t_1,t_2}F\}$ 也完全确定了二阶导数 D^2F.

推至一般情形, 将 $F \in \mathbb{D}^{k,2}$ 的 k 阶导数

$$D^kF = \{D^k_{t_1,\cdots,t_k}F : t_1, \cdots, t_k \in T\}$$

视为 $T^k \times \Omega$ 上的可测函数.

例 6.3 ([0,1)　区间上的布朗运动] 取 $\Omega = C_0([0,1;\mathbb{R}^d])$,$\mathbb{H} = L^2([0,1];\mathbb{R}^d)$. 定义

$$H^1 = \{x \in \mathrm{AC}([0,1];\mathbb{R}^d) : \dot{x} \in \mathbb{H}\}$$

这里的 $\dot{x}$ 表示绝对连续函数的导数. 子空间 H^1 被称为 Cameron-Martin 空间. 在 H^1 上定义内积

$$\langle x, y\rangle_{H^1} = \langle \dot{x}, \dot{y}\rangle_{\mathbb{H}},$$

在此内积下,H^1 是一个 Hilbert 空间. 并且嵌入映射 $H^1 \hookrightarrow \Omega$ 是连续的.

为便于讨论, 设 $d = 1$. 对任意 $t \in [0,1]$, 记 $W(t) = W(I_{[0,t]})$. 考虑光滑随机变量 $F = f(W(t_1), \cdots, W(t_n))$,$g \in L^2([0,1])$,

$$\begin{aligned}\langle DF, g\rangle &= \sum_{i=1}^{n} \partial_i f(W(t_1), \cdots, W(t_n))\langle I_{[0,t_i]}, g\rangle \\ &= \sum_{i=1}^{n} \partial_i f(W(t_1), \cdots, W(t_n)) \int_0^{t_i} g(s)\mathrm{d}s,\end{aligned}$$

另一方面, 固定 ω, 记 $\tilde{g}(t)=\int_0^t g(s)\mathrm{d}s$,

$$
\begin{aligned}
&\lim_{\epsilon\downarrow 0}\frac{F(x+\epsilon\tilde{g})-F(x)}{\epsilon}\\
=&\lim_{\epsilon\downarrow 0}\frac{f(W(t_1)(x+\epsilon\tilde{g}),\cdots,W(t_n)(x+\epsilon\tilde{g}))-F(\omega)}{\epsilon}\\
=&\lim_{\epsilon\downarrow 0}\frac{f(x(t_1)+\epsilon\tilde{g}(t_1),\cdots,x(t_n)+\epsilon\tilde{g}(t_n))-f(x(t_1),\cdots,x(t_n))}{\epsilon}\\
=&\sum_{i=1}^n\partial_i f(x(t_1),\cdots,x(t_n))\tilde{g}(t_i)\\
=&\sum_{i=1}^n\partial_i f(W(t_1)(x),\cdots,W(t_n)(x))\int_0^{t_i}g(s)\mathrm{d}s,
\end{aligned}
$$

所以 $\langle DF,g\rangle(x)=\frac{\mathrm{d}}{\mathrm{d}\epsilon}|_{\epsilon=0}F(x+\epsilon\int_0^{\cdot}g(s)\mathrm{d}s)$. 这个等式左边是在 Malliavin 导数框架下做内积得到的随机变量, 而右边则是将 F 看作样本空间 Ω 上的泛函, 在给定的样本点处, 计算 F 的方向导数, 当然这个方向不是任意的, 暂时只是取自 H^1, 但实际上, 为了保证一定的合理性, 也只能取于 H^1.

考虑 $F\in L^2(\Omega)$ 有如下表示:

$$F=\mathbb{E}[F]+\sum_{n=1}^{\infty}I_n(f_n),f_n\in L_S^2(T^n).$$

则导数算子和无穷求和可交换, 也即如下定理.

定理 6.29　设 $F=\mathbb{E}[F]+\sum_{n=1}^{\infty}I_n(f_n),f_n\in L_S^2(T^n)\in\mathbb{D}^{1,2}$. 则 $D_tF=\sum_{n=1}^{\infty}nI_{n-1}(f(\cdot,t))$.

下面介绍和条件期望相关的性质. 对 $A\in\mathscr{B}$, 记 $\mathscr{F}_A=\sigma(W(B):B\subset A,\mu(B)<\infty)$. 首先有以下引理.

引理 6.30　设 $F=\mathbb{E}[F]+\sum_{n=1}^{\infty}I_n(f_n),f_n\in L_S^2(T^n)\in L^2(\Omega)$. 则

$$\mathbb{E}(F|\mathscr{F}_A)=\sum_{n=0}^{\infty}I_n(f_n(I_A)^{\otimes n})$$

①

由以上引理可得以下命题.

命题 6.31　设 $F\in\mathbb{D}^{1,2}$,$A\in\mathscr{B}$. 则 $\mathbb{E}(F|\mathscr{F}_A)\in\mathbb{D}^{1,2}$. 且

$$D_t\mathbb{E}(F|\mathscr{F}_A)=\mathbb{E}(D_tF|\mathscr{F}_A)I_A(t).$$

特别地还有如下推论.

推论 6.32　设 $A\in\mathscr{B}$,$F\in\mathbb{D}^{1,2}$ 是 $\mathscr{F}_A$ 可测的. 则 DF 在 $T\times\Omega$ 上几乎处处为 0.

① $I_0=\mathrm{id}_{\mathbb{R}}$

6.5.3　Skorohod 随机积分

在白噪声情形下, 散度算子也称为 Skorohod 随机积分. 对 $u \in \mathrm{Dom}(\delta)$, 将 $\delta(u)$ 形式地记为 $\int_T u\mathrm{d}W$ 或 $\int_T u_t \mathrm{d}W_t$. $u \in L^2(\Omega; L^2(T)) \xrightarrow{\sim} L^2(T \times \Omega)$, 对任意 $t \in T$,u_t 作为一个平方可积的随机变量有 Wiener 混沌分解:

$$u_t = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)),$$

其中 $f_n \in L^2(T^{n+1}), n \geq 1$ 对前 n 个变量是对称的, 但不必对所有变量对称. 在如上分解的基础上, 可以给出 $u \in \mathrm{Dom}()$

命题 6.33　设 $(u_t) = \big(\sum_{n=0}^{\infty} I_{n+1}(f_n(\cdot, t))\big) \in L^2(T \times \Omega)$. 则 $u \in \mathrm{Dom}(\delta)$, 当且仅当级数

$$\sum_{n=0}^{\infty} I_{n+1}(\mathrm{Sym}(f_n)).$$

在 $L^2(\Omega)$ 中收敛.

证明　设 $g \in L_S^2(T^n)$. 计算可得

$$\begin{aligned}
\mathbb{E}[\langle u, DG\rangle] &= \mathbb{E}[\int_T u_t D_t G \mu(\mathrm{d}t)] \\
&= \mathbb{E}[\int_T u_t(nI_{n-1}(g(\cdot, t)))\mu(\mathrm{d}t)] \\
&= \mathbb{E}[\int_T (\sum_{m=0}^{\infty} I_m(f_m(\cdot, t)))(nI_{n-1}(g(\cdot, t)))\mu(\mathrm{d}t)] \\
&= n\int_T \mathbb{E}[I_{n-1}(f_{n-1}(\cdot, t))I_{n-1}(g(\cdot, t))]\mu(\mathrm{d}t) \\
&= n(n-1)!\int_T \langle I_{n-1}(f_{n-1}(\cdot, t)), I_{n-1}(g(\cdot, t))\rangle \mu(\mathrm{d}t) \\
&= \mathbb{E}[I_{n-1}(\mathrm{Sym}(f_{n-1}))Y].
\end{aligned}$$

若 $u \in \mathrm{Dom}(\delta)$, 则上式蕴含对任意 $G \in \mathscr{H}_n$,

$$\mathbb{E}[\delta(u)Y] = \mathbb{E}[\langle u, DY\rangle] = \mathbb{E}[I_{n-1}(\mathrm{Sym}(f)_{n-1})Y],$$

从而说明 $J_n(\delta(u)) = I_{n-1}(\mathrm{Sym}(f)_{n-1})$, 进而级数

$$\sum_{n=0}^{\infty} I_{n+1}(\mathrm{Sym}(f)_n)$$

收敛且在 $L^2(\Omega)$ 中收敛到 $\delta(u)$.

反过来, 若级数收敛, 记其收敛到 Z, 则

$$\mathbb{E}[\langle u, DY\rangle] = \mathbb{E}[ZY], \forall Y \in \mathscr{H}_n,$$

线性性保证了对 $\forall Y \in \oplus_{n=1}^N \mathscr{H}_n$ 同样有上式成立, 再由连续性即得 $Z = \delta(u)$.　■

记 $\mathbb{L}^{1,2} = \mathbb{D}^{1,2}(L^2(T))$, 则 $\mathbb{L}^{1,2}$ 同构于 $L^2(T;\mathbb{D}^{1,2})$. 依定理6.22, $\mathbb{L}^{1,2} \subset \mathrm{Dom}(\delta)$. 若 $u, v \in \mathbb{L}^{1,2}$, 则6.22中的等式在白噪声情形下的对应为

$$\mathbb{E}[\delta(u)\delta(v)] = \int_T \mathbb{E}[u_t v_t]\mu(\mathrm{d}t) + \int_T\int_T \mathbb{E}[(D_s u_t)(D_t u_s)]\mu(\mathrm{d}t)\mu(\mathrm{d}s).$$

在白噪声情形下, 对易关系6.23表述如下.

定理 6.34　设 $u \in \mathbb{L}^{1,2}$. 对几乎处处 $t \in T$,$\{D_t u_s : s \in T\} \in \mathrm{Dom}(\delta)$ 且存在一个版本使得 $\{\int_T D_t u_s \mathrm{d}W_s : t \in T\} \in L^2(T\Omega)$. 则 $\delta(u) \in \mathbb{D}^{1,2}$ 且

$$D_t(\delta(u)) = u_t + \delta(D_t u) = u_t + \int_T D_t u_s \mathrm{d}W_s.$$

6.5.4　Ito 随机积分

下面将 Ito 随机积分作为 Skorohod 随机积分的特例来讨论.

取 $T = (0,\infty)$,$\mathscr{B}$ 为 Lebsgue 可测集全体,μ 为 Lebsgue 测度. 令 $B_t = W(t) = W(I_{(0,t]})$ 为布朗运动.$(\mathscr{F}_t)$ 为布朗运动生成的自然 σ 代数流, 并注意到 $\mathscr{F}_t = \mathscr{F}_{(0,t]}$.

对适应过程 $u \in \mathbb{L}^{1,2}$,u_t 是 $\mathscr{F}_t$ 可测的, 所以由6.32可得 $D_s u_t = 0, s > t$, 所以对易关系蕴含 Ito 等距:

$$\mathbb{E}[\delta(u)^2] = \mathbb{E}[\int_0^\infty u_t^2 \mathrm{d}t].$$

考虑具有如下形式的初等随机过程 u:

$$u = \sum_{i=1}^n X_i I_{(t_i,t_{i+1}]},$$

其中 $0 \le t_0 < t_1 < \cdots < t_n$,$X_i \in L^2(\Omega, \mathscr{F}_{t_i})$, 可得 $\delta(u) = \sum_{i=1}^n X_i(B_{t_i} - B_{t_{i-1}})$. 由标准的逼近方法和散度算子的闭性可得: 全体初等随机过程在 $L^2((0,\infty)\times\Omega)$ 中的闭包, 记为 $L^2_{\mathscr{F}}((0,\infty)\times\Omega)$, 中的任一元素 v 都是适应的, 且 $v \in \mathrm{Dom}(\delta)$. 下面给出扩散过程的一个重要结论: 对随机积分 $X_t = \int_0^t u_s \mathrm{d}B_s$, 其 Malliavin 可微性取决于" 被积函数"u_s 的可微性, 反过来 X_t 的可微性也蕴含了 u_s 的可微性.

在此之前, 先对适应性做出一点简单的说明. 考虑 $F = f(W(t_1),\cdots,W(t_n))$ 是 $\mathscr{F}_T$ 可测的,$t_1 \le \cdots t_n \le t$. 则 $DF = \sum_{i=1}^n \partial_i f(W(t_1),\cdots,W(t_n))I_{[0,t_i]}$ 是一个初等随机过程, 依6.32有 $D_s F = 0, s \ge t$. 而当 $s < t$ 时, $D_t F = \sum_{i=1}^n \partial_i f(W(t_1),\cdots,W(t_n))I_{[0,t_i]}(t)$. 定义初等随机过程 $u = FI_{(S,R]}, S \ge T$. 则 $D_t u_s = \sum_{i=1}^n \partial_i f(W(t_1),\cdots,W(t_n))I_{[0,t_i]}(t)I_{(S,R]}(s)$. 固定 t, 考虑 $s > t$. 只有在 $s > S >\ge t_n$ 时, $D_t u_s \ne 0$ 才可能成立, 此时一定有 $D_t u_s \mathscr{F}_s$, 从而 $\{D_t u_s : s \ge t\}$ 是适应的. 由标准的逼近讨论可知对任一 $u \in L^2_{\mathscr{F}}$, $\{D_t u_s : s \ge t\}$ 都是适应的.

定理 6.35　设 $u \in L^2_{\mathscr{F}}((0,T)\times\Omega)$,$X = \int_0^T u_s \mathrm{d}B_s$. 则 $X \in \mathbb{D}^{1,2}$, 当且仅当 $u \in \mathbb{L}^{1,2}$.

此时,$\{D_t u_s : t \in (0,T)\} \in L^2_{\mathscr{F}}$ 且

$$D_t X = u(t) + \int_t^T D_t u_s \mathrm{d}B_s.$$

证明　若 $u \in L^2_{\mathscr{F}}$, 则 $D_t u_s$ 几乎处处存在且平方可积, 且 $\{D_t u_s : s \leq t\}$ 是适应和平方可积的, 所以 $D_t u_\cdot \in \mathrm{Dom}(\delta)$, 由对易关系可得,

$$D_t X = u_t + \int_t^T D_t u_s \mathrm{d}B_s.$$

反过来, 若 $X \in \mathbb{D}^{1,2}$. 记 u_t^n 为 u_t 在 $\mathscr{P}_n = \mathscr{H}_0 \oplus \cdots \oplus \mathscr{H}_n$ 上的正交投影. 令 $X^n = \int_0^T u_t^n \mathrm{d}B_s$, 则 X^n 恰是 X 在 $\mathscr{P}_{n+1}$ 上的正交投影, 且在 $\mathbb{D}^{1,2}$ 中收敛到 X. 从而由以下估计,

$$\begin{aligned}
\mathbb{E}[\|Du^n\|_{L^2((0,T)^2)}] &= \mathbb{E}[\int_0^T \int_0^s |D_t u_s^n|^2 \mathrm{d}t\mathrm{d}s] \\
&\leq \mathbb{E}[\int_0^T \int_0^s |D_t u_s^n|^2 \mathrm{d}t\mathrm{d}s] + \mathbb{E}[\int_0^t |u_t|^2 \mathrm{d}t] \\
&\leq \mathbb{E}[\int_0^T |D_t X^n| \mathrm{d}t].
\end{aligned}$$

由该有界性结合引理 6.17 即可得 $u \in \mathbb{L}^{1,2}$.

■

下面的定理说明如何已知 X 计算被积函数 u_s.

定理 6.36 (Clark-Ocone 公式)　设 $X \in \mathbb{D}^{1,2}$. 则

$$X = \mathbb{E}[X] + \int_0^t \mathbb{E}(D_s X | \mathscr{F}_s) \mathrm{d}B_s.$$

6.6　Ornstein-Uhlenbeck 半群

Ornstein-Uhlenbeck 半群相关内容十分丰富, 在此仅介绍一些重要的性质. 具体内容和证明参见 (Nualart, 2006).

定义 6.14　定义算子 $T_t, t \geq 0$ 为

$$T_t(F) := \sum_{n=0}^{\infty} e^{-nt} J_n F, F \in L^2(\Omega).$$

T_t 形成的单参数算子半群 $\{T_t : t \geq 0\}$ 称为 Ornstein-Uhlenbeck 半群, 简称为 OU 半群.

记 L 为 OU 半群的无穷小生成元, 也即:

$$LF = \lim_{t \downarrow 0} \frac{T_t F - F}{t},$$

其中 $F \in \mathrm{Dom}(L) = \{F \in \mathbb{D}^{1,2} : \text{极限} \lim_{t\downarrow 0} \frac{T_t F - F}{t} \text{在强收敛意义下存在}\}$.

注　OU 半群的无穷小生成元 L 有相对具体的表达式:$LF=\sum_{n=0}^{\infty}-nJ_nF$, 并且 L 具有一定的对称性, 即

$$\mathbb{E}[G(LF)]=\mathbb{E}[F(LG)], L, G\in \mathrm{Dom}(L).$$

算子 L 和导数算子, 散度算子有紧密关联, 具体体现在如下命题, 这一命题也为实际的计算和应用提供许多便利.

命题 6.37　设 $F\in L^2(\Omega)$. 则 $F\in \mathrm{Dom}(L)$, 当且仅当 $F\in \mathbb{D}^{1,2}$ 且 $DF\in \mathrm{Dom}(\delta)$, 此时

$$\delta(DF)=-LF.$$

作为以上命题的应用, 可以对 $F=f(W(h_1),\cdots,W(h_n))\in \mathrm{Srv}$ 计算

$$LF=\sum_{i,j}\partial_{ij}f(W(h_1),\cdots,W(h_n))\langle h_i,h_j\rangle_{\mathbb{H}}-\sum_{i=1}^{n}\partial_i f(W(h_1),\cdots,W(h_n))W(h_i)$$

更一般地, 类似于链式法则有如下定理.

定理 6.38　设 $F=(F^1,\cdots,F^m), F_i\in \mathbb{D}^{2,4}, i=1,\cdots,m,\phi\in C^2(\mathbb{R}^m)$ 且各一阶, 二阶偏导数有界. 则 $\phi(F)\in \mathrm{Dom}(L)$,

$$L(\phi(F))=\sum_{i,j}\partial_{ij}\phi(F)\langle DF^i,DF^j\rangle_{\mathbb{H}}+\sum_i\partial_i\phi(F)LF^i$$

OU 半群最重要的性质之一即超压缩性.

定理 6.39 (超压缩性)　设 $p>1$, 定义 $q(t)=(p-1)e^{2t}+1, t>0$. 则对任意 $F\in L^p(\Omega)$,

$$\|T_tF\|_{q(t)}\leq \|F\|_p$$

注　称为超压缩性的原因之一是 $q(t)$ 严格大于 p, 而高阶矩可以控制低阶矩.

6.7　具有 Lipschitz 系数的 SDE 的 Malliavin 可微性

本章最后介绍随机微分方程解的 Malliavin 可微性. 考虑方程

$$\mathrm{d}X_t=b_t(X_t)\mathrm{d}t+\sigma_t(X_t)\mathrm{d}W_t, X_0=x_0, t\in[0,T] \tag{6.9}$$

其中 $(W_t)_{t\geq 0}=(W_t^1,\cdots,W_t^d)$ 是定义在一个完备带流概率空间 $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\geq 0},\mathbb{P})$ 上的 d 维布朗运动.

$$b:[0,T]\times\mathbb{R}^m\to\mathbb{R}^m$$
$$\sigma:[0,T]\times\mathbb{R}^m\to\mathbb{R}^{m\times d}$$

为可测映射, 分别称为漂移系数和扩散系数. 记 $b=(b^1,\cdots,b^m)^T$,$\sigma=(\sigma^{i,j})_{m\times d}$ 作为一个矩阵值映射, 记其第 j 列为 $A^j=(\sigma^{j,1},\cdots,\sigma^{j,m})^T$, 所以方程6.9还可以写为

$$X_t = x_0 + \int_0^t b_s(X_s)\mathrm{d}s + \sum_{j=1}^d \int_0^t A_t^j \mathrm{d}W_t^j, t \in [0,T]. \tag{6.10}$$

本节讨论的随机微分方程总假设漂移系数和扩散系数满足 Lipschitz 连续性和有界性假设, 简记为方程6.9或6.10满足假设 (A), 其中假设 (A) 具体如下.

A　存在常数 $K > 0$ 使得

(A1) $\forall x, y \in \mathbb{R}^m, \forall t \in [0,T]$,

$$|b_t(x) - b_t(y)| + \sum_{j=1}^d |A_t^j(x) - A_t^j(y)| \le K|x-y|.$$

(A2) 映射 $t \mapsto A_t^j(0), t \mapsto b_t(0)$ 是 $[0,T]$ 上的有界映射. ■

若方程6.9满足假设 (A), 则6.9有唯一强解, 记该唯一强解为 $(X_t)_{t\in[0,T]} = (X_t^1, \cdots, X_t^m)_{t\in[0,T]}$, 具体可见 (Le Gall, 2016).

与第二节中的记号对应, 取 $\Omega = C_0([0,T]; \mathbb{R}^d)$ 为 $[0,1]$ 上在 0 处取值为 0 的 $\mathbb{R}^d$ 值连续函数全体,$\mathbb{P}$ 为 Wiener 测度, $\mathscr{F}$ 为 $\mathscr{B}(\Omega)$ 关于 $\mathbb{P}$ 的完备化,$\mathscr{F}_t$ 为布朗运动产生的自然流,$\mathbb{H}$ 为 $L^2([0,T]; \mathbb{R}^d)$. 另记 $\mathbb{D}^{1,\infty} = \cap_{p\ge 1}\mathbb{D}^{1,p}$. 先陈述本节的主要定理如下.

定理 6.40 (SDE 解的 Malliavin 可微性)　对任意 $t \in [0,T], i = 1, \cdots, m, X_t^i \in \mathbb{D}^{1,\infty}$. 进一步地,

$$\sup_{0\le r\le t} \mathbb{E}[\sup_{r\le s\le T} |D_r^j X_s^i|^p] < \infty,$$

且存在一致有界的 m 维适应过程 $\bar{A}_s^{k,l}, \bar{B}_s^k, k \in \{1, \cdots, m\}, l \in \{1, \cdots, d\}$, 使得

$$D_r^j X_t = A_r^j(X_r) + \sum_{k=1}^m \sum_{l=1}^d \int_r^t \bar{A}_s^{k,l} D_r^j X_s^k \mathrm{d}W_s^l + \sum_{k=1}^m \int_r^t \bar{B}_s^k D_r^j X_s^k \mathrm{d}s, r \le t,$$

以及 $D_r^j X_t, r > t$.

在证明该定理之前, 先介绍一个技术性引理, 具体证明见 (Nualart, 2006) 的命题 1.5.5.

引理 6.41　设 $p > \alpha > 1, F \in \mathbb{D}^{1,\alpha}$. 若 $DF \in L^p(\Omega; \mathbb{H})$, 则 $F \in L^p(\Omega)$.

定理6.40的证明　考虑 Picard 迭代, 令 $X_0(t) \equiv x_0$, 归纳定义

$$X_{n+1}(t) = x_0 + \int_0^t b(s, X_n(s))\mathrm{d}s + \sum_{j=1}^d \int_0^t A^j(s, X_n(s))\mathrm{d}W_s,$$

记 $\psi_n(t) = \sup_{0\le r\le t} \mathbb{E}[\sup_{r\le s\le T} |D_r X_n(s)|^p], p \ge 2$.

下面使用归纳法证明 $X_n^k(t) \in \mathbb{D}^{1,\infty}, i = 1, \cdots, m, t \in [0,T], \psi_n(t) < \infty$ 以及存在常数 c_1, c_2 使得

$$\psi_{n+1}(t) \le c_1 + c_2 \int_0^t \psi_n(s)\mathrm{d}s.$$

当 $n=0$ 时显然成立, 假设对 $k=1,\cdots,n$ 都有以上结论成立, 考虑 $k=n+1$.

首先由于系数的 Lipschitz 连续性, 固定 s, 对 $\sigma_s^{i,j}, b_s^i$, 应用链式法则, 可得存在一致有界的随机变量 $\bar{A}_s^{n,i,j,l}, \bar{B}_s^{n,i,l}$ 使得

$$D\sigma_s^{i,j}(X_n(s)) = \sum_{l=1}^{m} \bar{A}_s^{n,i,j,l} DX_n^i(s),$$

$$Db_s^i(X_n(s)) = \sum_{l=1}^{n} \bar{B}_s^{n,i,l} DX_n^i(s)$$

根据归纳假设,$DX_n^i(s) \in \mathbb{D}^{1,\infty}$, 而系数又是一致有界的, 所以 $D\sigma_s^{i,j}(X_n(s)), Db_s^i(X_n(s)) \in \mathbb{D}^{1,\infty}$.

由于" 被积函数"$D\sigma_s^{i,j}(X_n(s)) \in \mathbb{D}^{1,2}$, 所以 Ito 随机积分

$$\sum_{j=1}^{m} \int_0^t \sigma_s^{i,j}(X_n(s)) \mathrm{d}W_s^j \in \mathbb{D}^{1,2},$$

且

$$D_r^k \left(\sum_{j=1}^{m} \int_0^t \sigma_s^{i,j}(X_n(s)) \mathrm{d}W_s^j \right) = \sigma_r^{i,k}(X_n(r)) + \sum_{j=1}^{m} \int_r^t D_r^k(\sigma_s^{i,j}(X_n(s))) \mathrm{d}W_s^j.$$

对于上面等式右边, 由 Lipchitz 连续性可知,$\sigma_r^{i,k}(X_n(r))$ 和 $X_n(r)$ 有相同的可积性, 对于积分项的控制则由 BDG 不等式和 $\psi_n(t)$ 的有界性得到:

$$\begin{aligned} &\mathbb{E}[|\int_r^t D_r^k(\sigma_s^{i,j}(X_n(s))) \mathrm{d}W_s^j|^p] \\ \leq &\mathbb{E}[(|\int_r^t (D_r^k(\sigma_s^{i,j}(X_n(s))))^2 \mathrm{d}s|)^{p/2}] \end{aligned}$$

结合引理6.41, 可得

$$\sum_{j=1}^{m} \int_0^t \sigma_s^{i,j}(X_n(s)) \mathrm{d}W_s^j \in \mathbb{D}^{1,\infty}.$$

对于确定性积分部分, 可以将其看作级数的极限, 由于导数算子的闭性,$\int_0^t b_s^i(X_n(s)) \mathrm{d}s \in \mathbb{D}^{1,2}$, 以及

$$D_r^k \left(\int_0^t b_s^i(X_n(s)) \mathrm{d}s \right) = \int_0^t D_r^k b_s^i(X_n(s)) \mathrm{d}s.$$

下面估计 $\mathbb{E}[\sup_{r\le s\le T}|D_rX_{n+1}(s)|^p]$.

$$
\begin{aligned}
&\mathbb{E}[\sup_{r\le s\le T}|D_r^kX_{n+1}(s)|^p]\\
=&\mathbb{E}[\sup_{r\le s\le T}|\int_r^s D_rb_s^k(X_n(u))\mathrm{d}u+\sum_{l=1}^d(\sigma_r^{k,l}(X_n(r))+\int_r^s D_r(\sigma_u^{k,l}(X_n(u)))\mathrm{d}W_u^l)|^p]\\
=&\mathbb{E}[\sup_{r\le s\le T}|\int_r^s\sum_{i=1}^m\bar{B}_u^{n,k,i}D_rX_n^i(u)\mathrm{d}u+\sum_{l=1}^d((\sigma_r^{k,l}(X_n(r))+\int_r^s\sum_{i=1}^m\bar{A}_u^{n,k,l,i}D_rX_n^i(u)\mathrm{d}W_u|^p)]\\
\le&C_p\mathbb{E}[\sup_{r\le s\le T}\left(K^p|\int_r^s\sum_{i=1}^m D_rX_n^i(u)\mathrm{d}u|^p+dK^p|\int_r^s\sum_{i=1}^m D_rX_n^i(u)\mathrm{d}W_u^i|^p+\sum_{l=1}^d\sigma_r^{k,l}(X_n(r))\right)]\\
\le&C_pK^p\mathbb{E}[\sup_{r\le s\le T}||\int_r^s\sum_{i=1}^m D_rX_n^i(u)\mathrm{d}u|^p|]+dC_pK^p\sum_{i=1}^m\mathbb{E}[|\int_r^u|D_rX_n^i(u)|^2\mathrm{d}u|^{p/2}]+C_p\gamma_p.
\end{aligned}
$$

其中 $\gamma_p=\sup_{n,k}\mathbb{E}[\sup_{0\le s\le T}|A_s^k(X_n(s))|^p]<\infty$. 先估计第一个期望. 由 Holder 不等式可得,

$$
\begin{aligned}
&\mathbb{E}[\sup_{r\le s\le T}|\int_r^s D_rX_n^i(u)\mathrm{d}u|^p]\\
\le&(s-r)^{p-1}\mathbb{E}[\sup_{r\le s\le T}\int_r^s|D_rX_n^i(u)|^p\mathrm{d}u]\\
\le&T^{p-1}\mathbb{E}[\int_r^T|D_rX_n^i(u)|^p\mathrm{d}u].
\end{aligned}
$$

再估计第二个期望. 由 BDG 不等式和 Holder 不等式可得,

$$
\begin{aligned}
&\mathbb{E}[\sup_{r\le s\le T}|\int_r^s D_rX_n^i(u)\mathrm{d}W_u^i|^p]\\
\le&C_p\mathbb{E}[\left(\int_r^T|D_rX_n^i(u)|^2\mathrm{d}u\right)^{p/2}]\\
\le&C_p(T-r)^{\frac{p}{2}-1}\mathbb{E}[\int_r^T|D_rX_n^i(u)|^p\mathrm{d}u].
\end{aligned}
$$

结合以上可得 $\psi_{n+1}(t)\le c_1+c_2\int_0^t\psi_n(s)\mathrm{d}s$ 成立.

由归纳计算可得 $\psi_n(t)\le c_1\sum_{i=1}^n\frac{(c_2t)^i}{i!}\le c_1e^{c_2t}$, 也即 X_n^i 的导数在 $L^p(\Omega;\mathbb{H})$ 中是一致有界的, 结合6.41就说明了 $X^i(t)\in\mathbb{D}^{1,\infty}$. 证明完成. ■

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, 17

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, 54

张恭庆, 林源渠, 2021. 泛函分析讲义 (上册)[M]. 2 版. 北京大学出版社. 45

李文威, 2025. 代数学讲义[EB/OL]. https://wwli.asia/downloads/books/EAlg-Notes.pdf. 45, 47

林正炎, 陆传荣, 苏中根, 2023. 概率极限理论基础[M]. 3 版. 高等教育出版社. 38

王凤雨, 任盼盼, 2020. 关于测度值过程的随机分析[J]. 中国科学（数学）, 50(2): 231-252. 28, 31

黄志远, 严加安, 1997. 无穷维随机分析引论[M]. 科学出版社. 49

Bao J, Ren P, Wang F Y, 2021. Bismut formula for lions derivative of distribution-path dependent SDEs[J]. Journal of Differential Equations, 282: 285-329. 18, 28

Bogachev, 2007. Measure theory[M]. Springer Berlin, Heidelberg. 13, 14, 26

Bogachev V I, 2018. Weak convergence of measures[M]. Providence, Rhode Island : American Mathematical Society. 17

Brezis H, 2010. Functional analysis, sobolev spaces and partial differential equations[M]. Springer New York, NY. 45

Cardaliaguet P, Delarue F, Lasry J M, et al., 2019. Annals of mathematics studies: the master equation and the convergence problem in mean field games[M]. Princeton University Press. 33, 34, 35

Carmona R, Delarue F, 2018. Probabilistic theory of mean field games with applications I[M]. Springer Cham. 13, 23, 27

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

Cont R, Fournié D A, 2013. Functional Itô calculus and stochastic integral representation of martingales[J]. The Annals of Probability, 41(1): 109-133. 37, 41

Dupire B, 2009. Functional Itˆo calculus[J]. Portfolio Research Paper. 40

Figalli A, Glaudo F, 2021. An invitation to optimal transport, wasserstein distances, and gradient flows[M]. EMS Press. 17

Fournier N, Guillin A, 2015. On the rate of convergence in wasserstein distance of the empirical measure[J]. Probability Theory and Related Fields. 25

Kunze M, 2013. An introduction to malliavin calculus[EB/OL]. https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.020/kunze/malliavin/Malliavin_skript.pdf. 45

Le Gall J F, 2016. Stochastic differential equations[M]//Brownian Motion, Martingales, and Stochastic Calculus. Cham: Springer International Publishing: 209-233. 71

Nualart D, 2006. The malliavin calculus and related topics[M]. 2nd ed. Springer Berlin, Heidelberg. 45, 69, 71

Peng S, Song Y, Wang F, 2023. Survey on path-dependent PDEs[J]. Chinese Annals of Mathematics, Series B, 44(6): 837-856. 37, 38, 40, 41

Reed M, Simon B, 1980. Methods of modern mathematical physics: Vol. 1[M]. Academic Press. 45, 49

Ren P, Wang F Y, 2021. Derivative formulas in measure on riemannian manifolds[J]. Bulletin of the London Mathematical Society, 53(6): 1786-1800. 28, 31

Wang F Y, Ren P, 2024. Distribution dependent stochastic differential equations[M]. WORLD SCIENTIFIC. 28