≪ 测度论与概率论基础 ≫ 笔记-符号测度

QingShan

2021 年 4 月 12 日

在微积分中，如果函数 f 在区间 $[a,b]$ 上连续，那么

$$F(x)=\int_a^x f(y)dy$$

称为不定积分，而 f 是 F 的导数。设 f 是测度空间 $(X,\mathscr{F},\mu)$ 上积分存在的可测函数。我们自然把 $\mathscr{F}$ 上定义的集函数

$$\varphi(A)=\int_A fd\mu,\quad \forall A\in\mathscr{F}$$

也称之为不定积分，而 f 也叫做 φ 对测度 μ 的导数。而一般的一个 $\mathscr{F}$ 上的集函数，如果也像 φ 一样，除了非负性以外满足测度的所有其他性质，是否一定就有导数呢？本章将围绕这一中心问题展开讨论。

1 符号测度

考虑测度空间 $(X,\mathscr{F},\mu)$ 上积分存在的可测函数 f 的**不定积分**

$$\varphi(A)=\int_A fd\mu,\quad \forall A\in\mathscr{F} \tag{1}$$

根据积分的定义有 $\varphi(\varnothing)=0$。同时根据单调收敛定理（Levi 定理）可知 φ 是可列可加的。可以看出，除了非负性外，φ 满足测度的所有其他条件。我们把具有这样性质的集函数称为**符号测度**，正式定义为：

定义 1.1 设 $(X,\mathscr{F})$ 是一个可测空间。如果从 $\mathscr{F}$ 到 $\overline{\mathbb{R}}$ 的集函数 φ 满足：

(1) $\varphi(\varnothing)=0$

(2) 可列可加性：对任何两两不交的 $\{A_n, n=1,2,\cdots\} \subset \mathscr{F}$，有

$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \varphi(A_n)$$

两条性质，则称这样的 φ 为**符号测度**。

如果 φ 满足 $|\varphi(A)| < \infty, \forall A \in \mathscr{F}$，则称它是**有限的**；如果存在 X 的可测分割 $\{A_n, n=1,2,\cdots\}$ 使得 $|\varphi(A)| < \infty$ 对每个 $n=1,2,\cdots$ 成立，则称它是 σ **有限的**。

从定义的（1）（2）可以推出符号测度具有**有限可加性**。但是遗憾的是，依赖于测度非负性而推出的单调性和半可列可加性不再成立。然而由于上、下连续性证明中没有用到非负性，故对于有限符号测度来说**上、下连续性**依旧成立。

由于符号测度取值在 $[-\infty, +\infty]$ 中且满足可列可加性，容易发现其值域不可能同时包含 $-\infty$ 和 $+\infty$, 因此之后我们都讨论值域在 $(-\infty, +\infty]$ 中的符号测度。

命题 1.1 设 φ 是可测空间 $(X, \mathscr{F})$ 上的符号测度。如果 $A, B \in \mathscr{F}, A \supset B$ 且 $|\varphi(A)| < \infty$，则 $\varphi(B) < \infty$。

命题 1.2 设 φ 是一个符号测度。如果 $\{A_n, n=1,2,\cdots\} \subset \mathscr{F}$ 两两不交且满足

$$|\varphi\left(\bigcup_{n=1}^{\infty} A_n\right)| < \infty,$$

则

$$\sum_{n=1}^{\infty} |\varphi(A_n)| < \infty.$$

2 Hahn 分解和 Jordan 分解

如果我们能够找到 X 的可测分割 X^+, X^-，使得 X^+ 中所有可测集具有非负符号测度且 X^- 中所有可测集具有非正符号测度，那么我们称其为符号测度 φ 诱导出的 **Hahn 分解**。

对每个 $A \in \mathscr{F}$，令

$$\varphi^{\pm}(A) = \int_A f^{\pm} d\mu,$$

则 φ^+ 和 φ^- 都是测度且

$$\varphi = \varphi^+ - \varphi^- \tag{2}$$

成立。我们把 (2) 式称为 φ 的 **Jordan 分解**。

容易发现，把符号测度限制在 X^+ 上，那么它就会升级成为一个测度；把符号测度限制在 X^- 上并且取相反数，它也可以升级成为一个测度。可见 Hahn 分解将符号测度分解成为了两个测度，并且这两个符号测度的定义域刚好给出了 X 的一个可测分割。

容易看出，对于不定积分 $\varphi(A)=\int_A f d\mu$，其 Hahn 分解是显然的，取 $X^+=\{f\geq 0\}$,$X^-=\{f<0\}$ 即可满足要求。所以现在的问题是：对于一般的符号测度，Hahn 分解是否一定存在？答案也是肯定的！

定义一个新的函数

$$\varphi^*(A)=sup\{\varphi(B):B\subset A,B\in\mathscr{F}\}$$

直观地理解，其用于度量 A 中的正贡献元有多少。我们可以将 $\varphi^*(A)$ 理解为：将 A 中所有正贡献元拿出来，看看其测度有多大。同时不难看出其为 $\mathscr{F}$ 上**非负，单调且满足** $\varphi^*(\varnothing)=0$ **的集函数**。

引理 2.1 如果 $A\in\mathscr{F}$ 且满足 $\varphi(A)<\infty$，则对 $\forall\varepsilon>0$，$\exists A_\varepsilon\in\mathscr{F}$，使得

$$A_\varepsilon\subset A,\varphi(A_\varepsilon)\geq 0 \text{ 且 } \varphi^*(A-A_\varepsilon)\leq\varepsilon$$

此引理的意义在于：对于符号测度有限的可测集，我们总能够将其中所有的正贡献元 ($\varphi(A_\varepsilon)\geq 0$) 在“几乎”意义下 ($\varphi^*(A-A_\varepsilon)\leq\varepsilon$) 分离出来。

引理 2.2 对 $\forall A\in\mathscr{F}$，只要 $\varphi(A)<0$，就一定 $\exists A_0\in\mathscr{F}$ 使得

$$A_0\subset A,\varphi(A_0)<0 \text{ 且 } \varphi^*(A_0)=0$$

此引理的意义在于：对于符号测度为负的可测集，我们总能够将其中所有的负贡献元 ($\varphi(A_0)<0$) 完全的 ($\varphi^*(A_0)=0$) 分离出来。

定理 2.3(Hahn 分解) 对可测空间 $(X,\mathscr{F})$ 上的符号测度 φ,$\exists X^\pm\in\mathscr{F}$ 使得

$$X^+\cup X^-=X,\quad X^+\cap X^-=\varnothing$$

而且对每个 $A\in\mathscr{F}$ 有

$$\varphi(A\cap X^+)\geq 0\geq\varphi(A\cap X^-).$$

Hahn 分解 $\{X^+,X^-\}$ 在下列意义下是唯一的：如果存在两个分解 $\{X_1^+,X_1^-\}$ 和 $\{X_2^+,X_2^-\}$ 都符合上述两个要求，则

$$A\in\mathscr{F},\quad A\subset X_1^+\triangle X_2^+\Rightarrow\varphi(A)=0$$

$$B \in \mathscr{F}, \quad B \subset X_1^- \triangle X_2^- \Rightarrow \varphi(B) = 0$$

对于任意一个给定的符号测度 φ，Hahn 分解把空间 X 分解为两部分 X^+ 和 X^-：在 X^+ 上 φ 只取正值，在 X^- 上 φ 只取负值。于是，只要令

$$\varphi^+(A) = \varphi(A \cap X^+) \text{ 和 } \varphi^-(A) = -\varphi(A \cap X^-)$$

就得到了 Jordan 分解的形式，因此有以下定理。

定理 2.4(Jordan 分解)　对可测空间 $(X, \mathscr{F})$ 上的符号测度 φ，存在测度 φ^+ 和有限测度 φ^- 使得 Jordan 分解成立，并且

$$\varphi^+ = \varphi^*; \ \varphi^- = (-\varphi)^*$$

φ^- 有限的原因为我们只讨论定义域为 $(-\infty, +\infty]$ 的符号测度。我们把 φ^+ 和 φ^- 分别叫做 φ 的**上变差**和**下变差**，而 $|\varphi| \stackrel{\text{def}}{=} \varphi^+ + \varphi^-$ 叫做它的**全变差**。显然，符号测度的上变差、下变差和全变差都是测度。

将符号测度分解为两个测度之差其实并不新奇，因为这是很容易做到的，而且有很多种方式可以做到。但是通过上述过程得到的符号测度分解中，我们对于两个测度的结构都有着清晰的认识，这是可贵的！Jordan 分解的一个特殊之处在于其在符号测度的测度分解中具有极大性。即对于符号测度的任意测度分解 $\varphi = \mu + \nu$，总有 $\varphi^+ \geq \mu, \varphi^- \geq \nu$ 成立。

3　Radon-Nikodym 定理

设 φ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的符号测度，我们希望定义 φ 的导数。基本想法其实很简单：如果这个符号测度能够唯一地表示成不定积分的形式

$$\varphi(A) = \int_A f d\mu, \quad \forall A \in \mathscr{F} \tag{3}$$

的话，那么像初等微积分里面把不定积分看成它的导数的原函数那样，就认为 f 是符号测度 φ 对于 μ 的导数。

定义 3.1　设 φ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的符号测度。如果存在 a.e. 唯一的可测函数 f 使得 (3) 式成立，则称 f 为 φ 对于 μ 的 R-N(Radon-Nikodym) **导数**（或简称导数），记为 $\frac{d\varphi}{d\mu} \stackrel{\text{def}}{=} f$.

如微积分中并非所有函数都可以求导一样，也不是每一个符号测度都有 R-N 导数。那么什么样的符号测度才有 R-N 导数呢？下面将给出答案。首先定义绝对连续性。

定义 3.2 设 φ 和 μ 分别是可测空间 $(X,\mathscr{F})$ 上的符号测度和测度。如果对任何 $A\in\mathscr{F}$，均有

$$\mu(A)=0\Rightarrow\varphi(A)=0,$$

则称 φ **对** μ **绝对连续**，记做 $\varphi\ll\mu$.

有了绝对连续性的定义后，我们就可以给出 R-N 定理，R-N 定理告诉了我们什么样的符号测度才有 R-N 导数。

定理 3.1(Radon-Nikodym 定理) 设 φ 和 μ 分别是可测空间 $(X,\mathscr{F})$ 上的符号测度和 σ 有限测度。如果 $\varphi\ll\mu$，则存在 $(X,\mathscr{F},\mu)$ 上 a.e. 唯一的可测函数 f 使得

$$\int_X f^- d\mu<\infty \text{ 且 } \forall A\in\mathscr{F},\varphi(A)=\int_A f d\mu$$

特别的，当 φ 为 σ 有限时，则 $f<\infty$ a.e.。

Radon-Nikodym 定理的条件事实上很弱，只要求 μ 为一个 σ 有限测度即可保证任意符号测度对其的 R-N 导数存在。如果直接从这么弱的条件出发去证明这个定理是很难的，所以这个定理的证明是先对较强的条件证明，然后再逐步放宽条件进行推广。具体来说，遵循以下链条进行证明：

φ,μ 为有限测度 $\to$ φ 为有限符号测度，μ 为有限测度 $\to$ φ 为 σ 有限符号测度，μ 为有限测度 $\to$ φ 为符号测度，μ 为有限测度 $\to$ φ 为符号测度，μ 为 σ 有限测度。

最后，我们需要注意，定理中 μ 为 σ 有限的这个条件是必要的。如果把它去掉，定理的结论可能会不对。

4 Lebesgue 分解

在上节中，我们定义了符号测度关于测度的绝对连续性。即对于符号测度 φ 以及测度 μ 有 $\varphi\ll\mu$ 当且仅当对任意的 μ-零测集，其必为 φ-零测集。

作为推广，我们可以将绝对连续性的定义扩展到两个符号测度之间。我们可以定义对于两个符号测度 φ,ϕ，有 $\varphi\ll\phi$ 当且仅当对任意的 $|\phi|$-零测集，其必为 $|\varphi|$-零测集。(注意这里的 $|\,|$ 指的是全变差而不是绝对值！)

接下来，我们给出两个符号测度相互奇异的定义。

定义 4.1 对可测空间 $(X,\mathscr{F})$ 上的符号测度 φ 和 ϕ。若 $\exists N\in\mathscr{F}$ 使得 $|\varphi|(N)=|\phi|(N^c)=0$，则称这两个符号测度相互奇异，记为 $\varphi\perp\phi$.

奇异性是通过全变差定义的，但是也可以用两个符号测度自身来直接验证。具体来说，我们有以下等价条件。

引理 4.1 对任意的符号测度 φ 和 ϕ，$\varphi \perp \phi$ 当且仅当 $\exists N \in \mathscr{F}$ 使得 $\forall A \in \mathscr{F}$ 有

$$\varphi(A \cap N^c) = \phi(A \cap N) = 0.$$

即总能找到一个可测集，使得这两个符号测度中的其中一个在这个可测集内部无作用，而另一个在这个可测集外部无作用，他们恰好是交错的。

下面的引理说明了奇异性和绝对连续性是两个对立的概念。

引理 4.2 设 φ 和 ϕ 都是符号测度。如果 $\varphi \ll \phi$ 且 $\varphi \perp \phi$，则

$$\varphi \equiv 0.$$

接下来，我们再来说明绝对连续性和相互奇异性进一步的含义。

首先对于绝对连续性来说，我们能够大概的意识到绝对连续性的意义在于：对于充分小的 ε，总是能够找到一个固定的 δ，使得某个量小于 δ 时，总是有依赖于这个量的另一个量小于 ε。

在这里也是同样的。对于符号测度 $\varphi, \phi, \varphi \ll \phi$ 成立且 φ 有限的话，$\forall \varepsilon > 0, \exists \delta > 0$，使得对于可测集 A 如果有 $|\phi|(A) < \delta$，那么 $|\varphi|(A) < \varepsilon$。

另一种理解方式可以从 Radon-Nikodym 定理的角度出发。我们不妨进一步假定测度空间 $(X, \mathscr{F}, \mu)$ 是 σ 有限的。那么其上的符号测度都可以不定积分化。即

$$\forall A \in \mathscr{F}, \varphi(A) = \int_A f d\mu, \phi(A) = \int_A g d\mu$$

那么

$$\forall A \in \mathscr{F}, |\varphi|(A) = \int_A |f| d\mu, |\phi|(A) = \int_A |g| d\mu$$

从而可知绝对连续性告诉我们：如果 $\int_A |g| d\mu = 0$ 那么 $\int_A |f| d\mu = 0$，即如果 $g|_A = 0$ a.e. 那么 $f|_A = 0$ a.e. 。故可将绝对连续性理解为函数空间中的 dominance。

类似的，如果有相互奇异性 $\varphi \perp \phi$，同样进行上述不定积分化，那么 $\exists E \in \mathscr{F}$，使得 $\int_N |f| d\mu = \int_{N^c} |g| d\mu = 0$。故

$$f|_N = 0, g|_{N^c} = 0 \ a.e. \ fg = 0 \ a.e.$$

则 $\int_X fg d\mu = 0$。回忆 L_2 空间作为 Hilbert 空间上的内积，我们就容易理解为何相互奇异要用“垂直”符号来表示。因为这本质上就是函数空间中的正交性！

接下来,我们很自然的引入 Lebesgue 分解,其目的是在测度空间 $(X,\mathscr{F},\mu)$ 上将一个符号测度 φ 分解为两个测度之和，即

$$\varphi=\varphi_c+\varphi_s \text{ 使得 } \varphi_c\ll\mu,\varphi_s\perp\mu. \tag{4}$$

不难发现，Lebesgue 分解其实与函数空间中的正交分解相呼应，φ_c 可以理解为一个被 μ 控制住的符号测度（平行方向），φ_s 可以理解为一个与 μ 正交的符号测度（正交方向）。

接下来我们正式提出 Lebesgue 分解定理。

定理 4.1(Lebesgue 分解定理) 对于 σ 有限测度空间 $(X,\mathscr{F},\mu)$ 上的 σ 有限符号测度 φ，存在 σ 有限符号测度 φ_c 和 φ_s 使 (4) 式成立，而且这样的分解是唯一的。

要直接证明 Lebesgue 分解定理是困难的，故我们模仿 Radon-Nikodym 定理的证明，步步推进，有如下推广链：

φ,μ 为有限测度 $\to$ φ,μ 为 σ 有限测度 $\to$ φ 为 σ 有限符号测度，μ 为有限测度。

事实上，Lebesgue 分解定理的条件还能再弱化，如以下引理所述：

引理 4.3 对于测度空间 $(X,\mathscr{F})$ 上的 σ 有限符号测度 φ 和 ϕ，存在两个 σ 有限符号测度 φ_c 和 φ_s 使下式成立：

$$\varphi=\varphi_c+\varphi_s;\quad \varphi_c\ll\phi;\quad \varphi_s\perp\phi.$$

这样的分解是唯一的。

Lebesgue 分解定理的一个重要应用就是用来解释初等概率论中随机变量的分类问题（这是初等概率论中说不清楚的事情）。

考虑概率空间 $(X,\mathscr{F},P)$ 上的随机变量 f。其总是诱导了 $(\mathbb{R},\mathscr{R}_{\mathbb{R}})$ 上的概率测度 Pf^{-1}。注意到了 $(\mathbb{R},\mathscr{R}_{\mathbb{R}})$ 上又有 Lebesgue 测度 λ 为 σ 有限的，故我们可以将 Pf^{-1} 关于 λ 做 Lebesgue 分解得到：

$$Pf^{-1}=\mu_1+\mu_s,\mu_1\ll\lambda,\mu_s\perp\lambda.$$

对于有限测度 μ_s，考虑 $D=\{x\in X|\mu_s(\{x\})>0\}$，令：

$$\mu_2=\mu_s|_D;$$

$$\mu_3=\mu_s|_{D^c};$$

由此得到了测度 Pf^{-1} 的唯一分解式：

$$Pf^{-1} = \mu_1 + \mu_2 + \mu_3, \tag{5}$$

其中的右侧三者均为有限测度。

其中需要说明 D 为至多可数集，这是因为 $\mu_s(\{x\}) > 0 \Rightarrow \lambda(\{x\}) = 0 \Rightarrow \mu_1(\{x\}) = 0 \Rightarrow Pf^{-1}\{x\} > 0 \Rightarrow \lambda_F\{x\} > 0 \Rightarrow F(x) - F(x^-) > 0$。所以 D 中的点对应了 F 的间断点，而单调函数的间断点至多只有可数个。

接下来我们根据 Pf^{-1} 的分解式 (5) 来讨论随机变量的分类问题：

(1) 若 μ_2, μ_3 恒为零，则 $Pf^{-1} \ll \lambda$，此时由 Radon-Nikodym 定理，可得几乎处处非负的可测函数 p，使得 $Pf^{-1}(A) = \int_A p d\lambda$。我们发现，$p$ 就是**概率密度**！此时，随机变量 f 就是一个**连续性随机变量**！（同时我们得到的概率密度在几乎意义下是唯一的）。这时，随机变量函数的期望的计算公式为：

$$E(g \circ f) = \int_R g(x)p(x)dx.$$

(2) 若 μ_1, μ_3 恒为零，则 $Pf^{-1} = \mu_2 \perp \lambda$，且 $Pf^{-1} = \mu_s|_D$。这意味着：仅 D 中至多可数个元素处的取值就能够完全决定 Pf^{-1} 的整体结构。故不妨令 $D = \{x_1, x_2, \dots\}, p_i = Pf^{-1}(\{x_i\})$。则 $\{p_i\}$ 即为随机变量 f 的**概率分布列**！，此时随机变量为**离散型随机变量**！（其取值集合即为 D，D 至多可数，这与初等概率论吻合！）。这时，随机变量函数的期望的计算公式为：

$$E(g \circ f) = \sum_{n-1}^{\infty} g(x_n)p_n.$$

(3) 若 μ_1, μ_2 恒为零，则 $Pf^{-1} = \mu_s|_{D^c} = \mu_3 \perp \lambda$。此时比较奇特的一点是：任意至多可数集的 Pf^{-1} 测度均为 0，但其还是与 Lebesgue 测度相互奇异。这样的随机变量称为**奇异性随机变量**！

事实上，任意随机变量都是以上三种随机变量的组合。

接下来我们来解释为什么要对随机变量诱导出的概率密度 Pf^{-1} 进行分解。事实上，我们只需要注意到：$Pf^{-1} = \lambda_F(f \sim F)$ 就可以理解这一点。因为随机变量诱导出的概率密度本质上说与随机变量的分布函数诱导出的 Lebesgue-Stieljes 测度是一致的！由初等概率论的学习，我们知道对随机变量分类事实上就是观察随机变量的分布函数。如果分布函数在 a 点处发生了跳跃，那么意味着随机变量在这个点处有离散取值。故研究分布函数诱导出的 Lebesgue-Stieljes 测度的 Lebesgue 分解本质上等于研究随机变量的分类问题！

事实上，我们能够生成任意连续型、离散型、奇异型以及混合型随机变量。因为只要我们有分布函数 F，我们总能求出其左连续逆 $F^{\leftarrow}$。从而左连续逆复合上均匀分布的随机变量就是一个服从 F 分布的随机变量了!