probability theory

李博

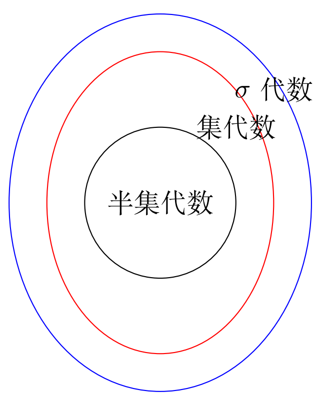
July 3, 2019

1 基本定义

定义 1.1 (半集代数). 如果 Ω 的子集类 $\mathscr{L}$ 满足
(1)$\Omega, \emptyset \in \mathscr{L}$
(2)$A, B \in \mathscr{L} \Rightarrow A \bigcap B \in \mathscr{L}$
(3)$A, B \in \mathscr{L} \Rightarrow \exists A_i \in \mathscr{L} A_i \quad A - B = A_i$

定义 1.2 (集代数). 如果 Ω 的子集类 $\mathscr{L}$ 满足
(1)$\Omega \in \mathscr{L}$
(2)$A, B \in \mathscr{L} \Rightarrow A - B \in \mathscr{L}$

定义 1.3 (σ 代数). 如果 Ω 的子集类 $\mathscr{L}$ 满足
(1)$\Omega \in \mathscr{L}$
(2) 若 $A \in \mathscr{L}, A^c \in \mathscr{L}$
(3) 可数的 $A_i \in \mathscr{L}, \Rightarrow \bigcup A_i \in \mathscr{L}$
将 σ 代数条件整理为 π 系 λ 系



定义 1.4 (borelσ 代数). 由开集生成的 σ 代数

命题 1.1. 半集代数需要添加有限集合的并变成集代数，集代数可数集合交和并变为 σ 代数 *(也称为单调类定理)*

定义 1.5 (乘积空间与乘积 σ 代数). $\mathscr{A}_i$ 是 Ω_i 的 σ 代数,$\mathscr{L} = \{\mathscr{A}_i \times ...\}$ 是半集代数,$\sigma(\mathscr{L})$ 称为乘积 σ 代数

定义 1.6. 如果集函数 Φ 具有 σ 可加性，为符号测度，如果非负，为测度，如果 $\Phi(\Omega)=1$，为概率测度。若测度还有限，则为有限可加测度。

定义 1.7 (lebesgue 测度)**.** 在 R^n 空间的所有子集上定义外侧度 *(不是测度)*，但是这个外测度不满足可数可加性，为此选出 *lebesgue* 可测集：$\forall T\in R^n, m*T=m*(T\bigcap E)+m*\left(T\bigcap E^C\right)$，定义在 *lebesgue* 可测集上的外侧度是 *lebesgue* 测度 *(满足外侧度的性质和完全可加性)*

定义 1.8. 设 $(\Omega,\mathscr{A},P)$ 是概率空间，$\xi:\Omega\to R$ 是实函数，如果 $\{\omega:\xi(\omega)<x\}\in\mathscr{A}$，则称 ξ 实随机变量。$F(x)=P(\xi<x)$ 为分布函数。[1]

定义 1.9 (可测映射)**.** $(\Omega,\mathscr{A})$ 与 $(E,\mathscr{B})$ 是两个可测空间，$f:\Omega\to E$，若 $\forall B\in\mathscr{B}, f^-(B)\in\mathscr{A}$，称 f 是可测映射。

定义 1.10 (几乎处处收敛)**.** $\{f_n\}$ 是可测函数列，若 $E\in\mathscr{A}$ 是非零测集，$f_n\to f$，则称 f_n 几乎处处收敛 f。[2]

定义 1.11 (依测度收敛)**.** 如果 $\forall\epsilon>0,\mu(|f_n-f|\geq\epsilon)\to 0\,(n\to\infty)$，称 f_n 依测度 μ 收敛于 f. [3]

命题 1.2 (控制收敛定理)**.** 设 g 是可积函数，$|f_n|\leq g$，如果 $f_n\xrightarrow{a.e}f$ 或者 $f_n\xrightarrow{\mu}f$，则 $\int_\Omega f_n d\mu\to\int_\Omega f d\mu$[4]

定义 1.12 (勒贝格积分)**.** $\int_\Omega f d\mu=sup\{\int_\Omega g d\mu:0\leq g\leq f, g$ 为简单函数$\}$，若上述确界存在则称可积。

定理 1.1 (Hahn 分解定理)**.** $(\Omega,\mathscr{A},\mu)$，则存在 $D\in\mathscr{A}, st.\mu(D)=\inf\limits_{A\in\mathscr{A}}\mu(A)$. 令 $\mu^+(A)=\mu(A\bigcap D^c), \mu^-(A)=-\mu(A\bigcap D)$，则 μ^-,μ^+ 均为测度，且 $\mu=\mu^+-\mu^-$

定理 1.2 (Fubini 定理)**.** $(\Omega_1,\mathscr{A}_1,\mu_1),(\Omega_2,\mathscr{A}_2,\mu_2)$ 是两个测度空间，$(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2,\mu=\mu_1\times\mu_2)$ 是乘积测度。若 f 是其上可积函数，则 $\int_{\Omega_1\times\Omega_2}f d\mu_1\times\mu_2=\int_{\Omega_1}\left(\int_{\Omega_2}f d\mu_2\right)d\mu_1$[5]

定义 1.13. $(R^n,\mathscr{B}^n,\mu)$，称 $f(t)=\int_{R^n}e^{i<t,x>}d\mu$ 为 μ 的特征函数。[6]

定理 1.3. 设 f 为 μ 的特征函数，若 $\forall x\in R^n,\mu(x)=0$，则 $\mu([a,b])=\lim\limits_{T\to\infty}\frac{1}{(2\pi)^n}\int_{-T}^{T}\cdots\int_{-T}^{T}\prod\limits_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}f(t_1...t_n)\,dt_1...dt_n$

Proof. 右端 $\overset{Fubini}{=}$

[1] 由分布函数生成的测度称为 L-S 测度
[2] 当 f 可积时，若以 $\mu(f*g)$ 作为”配合”，即引入对偶空间，则上述定义可以描述为弱收敛。
[3] 处处收敛必依测度收敛，以绝对值的积分当作 1 范数，测度收敛为 1 范数收敛，前提自然是 f 可积
[4] 此定理说明某收敛情况下，函数数列的极限在积分号内外相等
[5] 若 μ 不能表示为乘积测度，可以利用转移测度化为累次积分
[6] 特征函数类似展成傅立叶的一族正交基系数

$$
\begin{aligned}
&\int_{R^n} d\mu \frac{1}{(2\pi)^n} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{k=1}^{n} \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} f(t_1 \ldots t_n)\, dt_1 \ldots dt_n \\
&= \int_{R^n} d\mu \frac{1}{(2\pi)^n} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{k=1}^{n} \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} e^{it_k x_k} dt_1 \ldots dt_n \\
&= \int_{R^n} d\mu \frac{1}{(\pi)^n} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{k=1}^{n} \frac{\sin t_k (x_k - a_k) - \sin t_k (x_k - b_k)}{t_k} dt_1 \ldots dt_n
\end{aligned}
$$

由 $\int_{-\infty}^{+\infty} \frac{\sin t}{t} dt = \pi$ 可得结论 □

2 wassertein 距离与最优传输

定义 2.1. 在 *borel* 可测空间 $(E, \mathscr{B})$ 上，以 $\mathscr{B}_b$ 表示有界可测函数全体，以 C_b 表示有界连续函数全体。

- $\forall f \in \mathscr{B}_b, \mu_n(f) \to \mu(f)$，则称 μ_n 强收敛到 μ[7]
- $\forall f \in C_b, \mu_n(f) \to \mu(f)$，则称 μ_n 弱收敛到 μ[8]

命题 2.1. 在此定义下，测度空间是弱紧空间

Proof. https://wujilingfeng.top/2019/07/03/ □

倘若对于空间 E 中两个概率测度 u,v 如果用全变差度量其距离,则 $(u-v)^{+} E = (u-v)^{-} E = \frac{1}{2} |u-v|_{var}$

命题 2.2. 我们考虑以下最优传输模型，E 的 σ 代数 $\mathscr{A}$，若有 $a_i, a_j \in \mathscr{A}$，则代价函数 $cost(a_i, a_j) = 1 - \delta_{ij}$，可分析得出它的最优传输 (某个 *wassertein* 耦合)π，满足 $\pi(E) = (u-v)^{-} E$

命题 2.3. 最优 *wassertein* 耦合是距离正定性，对称性，三角不等式性，前两个条件需要限制代价函数的对称性，非负性及对角线为 *0* 的。最后一个三角不等式由最优性保证

Proof. d □

[7]等价全变差作为范数的收敛, 也等价与 $\forall A \in \mathscr{B}, \mu_n(A) \to \mu(A)$

[8]我不知道这个弱收敛跟泛函中由对偶空间定义相容，但是以下结论是成立的