

# 测度粘土——将只能加的东西粘起来

小飞舞

2023 | 10 | 04

人类自古以来就掌握加法了,我在这专门指朴素的加法,就像  $1+1=2$  那样. 此时的加法就像堆积木一般,将一个个元素放到一起,让他们紧挨着彼此,正如我们孩童时期所做的一般.

稍微大些,我们会想要更多积木. 无穷级数的发展大致来自于 14 世纪,此时我们终于不是将积木组来组去了,而是想办法思考如何将无穷个元素堆叠起来. 这样的想法非同小可:直到 Augustin-Louis Cauchy 的时代以前,人们都不能严谨地描述无穷求和.

那么,在后来我们能够加些什么东西呢?

## 第一节 测度,积分与不可数求和

正如我们标题所展现的那样,我们在此必然要引入测度. 我们先回顾下它的定义:

**定义** ( $\sigma$ -代数上的) 测度). 假定  $X$  是集合,  $\mathcal{P}(X)$  记为  $X$  子集的集合,即幂集,令  $\mathcal{M} \subset \mathcal{P}(X)$  满足

- $\{\emptyset, X\} \subset \mathcal{M}$ ;
- 若  $\{A_n\}_{n \geq 0} \subset \mathcal{M}$ , 则  $\bigcup_{n \geq 0} A_n \in \mathcal{M}$ ;
- 令  $A \in \mathcal{M}$ , 则  $A$  的补集  $A^c \in \mathcal{M}$ .

称  $\mathcal{M}$  是一个  $\sigma$ -代数, 一个  $\sigma$ -代数上的测度  $\mu$  是满足以下性质的函数:

- $\forall A \in \mathcal{M}, \mu(A) \geq 0$  或  $\mu(A) = \infty$ ;
- 若  $\{A_n\}_{n \geq 0} \subset \mathcal{M}$  两两不交, 则  $\mu(\bigcup_{n \geq 0} A_n) = \sum_{n \geq 0} \mu(A_n)$ ;
- $\mu(\emptyset) = 0$ .

┘

简而言之, 测度是对集合的测量, 由于种种原因 (比如我们想要考虑可数集, 如  $\mathbb{Q}$  的测度), 我们会让其满足可数可加性, 因此会用允许可数操作的  $\sigma$ -代数来承载它.

好, 现在有了测度, 大家都可以搞积分了. 此时如果有人问:

如何求和不可数个东西呢?

你会怎么回答他?

嗯.....一下子想到的回答真可能是积分: 换句话说, 即使是在 Isaac Newton 的时代, 积分不也是把无穷多个无穷小加起来得到一个正经的数吗?

但是稍微想想也知道, 不管使用 Bernhard Riemann 亦或是 Henri Lebesgue 对积分的定义, 其都是用有限个或者可数个东西加起来取极限得到的. 也就是说, 这本质上没有回答这个问题. 不过, 不可数个 1 加起来, 真的有意义吗?

一个暂时性的回答是, 利用上确界技术, 假定我们对不可数个正数求和:  $\sum_{x \in X} f(x)$ . 那么应该有:

$$\sum_{x \in X} f(x) := \sup \left\{ \sum_{\substack{F \subset X \\ F \text{ 有限}}} f(x) \mid \text{对所有这样的 } F \subset X \right\}.$$

其求和的关键指标是:

$$\text{card}\{x \in X \mid f(x) > 0\}.$$

是否是可数的. 如果是可数的, 那么就返回到可数求和, 如果是不可数的, (留作习题?) 容易证明, 其求和值是无限.

看起来问题好像解决了, 但是这不是又回到了可数求和上吗? 还是我们要求不可数求和本身就有问题? 但是平面图形的面积通过积分得到不就是对不可数个 0 加起来得到的吗?

或许到最后,我们可以认为,是“不可数”的要求太高了,或者说“基数太大了”:我们唯一一个能用合理方法解释不可数求和的,被叫做积分的东西,它不是正经的求和,而是将东西粘起来.换言之,不是离散地堆起来,而是将那一些元素泡开粘在一起.或许我们要求所有求和都是离散的本身就是一种任性吧.

从 Lebesgue 那篇划时代的论文起的往后几十年,测度的理论被推广到更抽象的地步.我们可以允许取值是复数,甚至是抽象向量的测度,这些处理都是为了更好地粘东西.回到标题上来,为什么标题叫做“测度黏土”,笔者认为,测度除了带来更完善的积分理论之外,本身就蕴含了一种单位分解的新技巧,这使得我们更精准地切割集合,粘黏集合变成了可能,应用最广泛的“粘”就是对测度的积分.

## 第二节 从多项式到.....?

自然,只讨论积分,或者是求和的“粘化”是没意思的,具体细节可以在任意一本实分析课本中找到,我们想要的是应用更广泛的粘化,除了手抓饼以外,我们还可以想想有什么东西是能加的.

笔者在这里打算讨论矩阵,或者是 Hilbert 空间上的有界线性算子(如果读者不熟悉的话可以暂时当矩阵看待).

给定算子  $T$ ,我们能“加”,事实上,我们可以做更多——我们可以将多项式作用在算子上,我们在线性代数课上应该学过这样的操作:  $P(x) \mapsto P(T)$ . 如果用一种比较花里胡哨的说法,这叫做算子的多项式演算.但即使是所有多项式组成的集合,它也总是给人一种“离散”的感觉,换言之,它不够“粘(稠)”.

我知道你想说什么, Karl Weierstrass 的定理阐明,在底空间是  $\mathbb{R}$  上的紧致子集时,多项式在连续函数空间(以  $\sup$  度量)内稠密.应此,在有多项式出现的场合,我们可以有以下直观:

如果“多项式”的行为只被  $\mathbb{R}$  上的某个紧致集中的行为确定的话,我们大致可以通过多项式来逼近连续函数,得到“连续函数”的行为.

考虑到“多项式演算”这个名称,我想,应该有种叫做“连续函数演算”的东西,以下是一个例子:

**例 (矩阵指数).** 读者可以将矩阵  $A$  视为  $\mathbb{R}^n$  或  $\mathbb{C}^n$  中的子集, 赋予矩阵范数  $\|A\| = \sup_{\|x\|=1} \|Ax\|$ . 这样, 我们可以承认幂级数:

$$\exp(A) = \sum_{n \geq 0} \frac{A^n}{n!}.$$

假定有微分方程:

$$y: \mathbb{R} \rightarrow \mathbb{R}^n, \quad y' = Ay, \quad y(0) = y_0.$$

那么实际上,  $y(t) = \exp(tA)y_0$  是这个方程的唯一解.

因此, 幂级数的确是一个解法, 我们叫做“幂级数”演算, 或者“全纯演算”. 不过对于  $x \mapsto \sqrt{x}$  这类不好展成幂级数的连续函数, 这样用幂级数的定义显然会非常困难.

在线性代数中, 我们会学到一个定理:

**定理 (自伴矩阵的平方根).** 令  $A$  是对称正定矩阵, 或者是自伴正定矩阵, 那么存在正定矩阵  $B_m$  满足:

$$B_m^m = A.$$

证明. 我们只需要知道对称 (自伴) 矩阵正交相似于一个对角矩阵  $\text{diag}(\lambda_1, \dots, \lambda_n)$ , 也即是:

$$A = T^{-1} \text{diag}(\lambda_1, \dots, \lambda_n) T.$$

那么实际上只需令

$$B_m = T^{-1} \text{diag}(\lambda_1^{1/m}, \dots, \lambda_n^{1/m}) T.$$

即可. 此时  $B$  的正定性是显然的, 顺带一提, 这样的正定矩阵  $B_m$  是唯一的.

□

这个例子本身是朴素的, 但是其给出了一个想法: 对矩阵的演算可以变成对其特征值的演算, 考虑到所有特征值组成的集合是紧致的, 正好也满足 Weierstrass 定理的需求. 因此, 借由上面的例子, 我们可以这样做:

**定义** (可对角化矩阵的函数演算 (暂行条例)). 给定一个可对角化的函数  $\mathbf{A}$  和复值函数  $f$ , 则定义

$$f(\mathbf{A}) := \mathbf{T}^{-1} \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) \mathbf{T}.$$

其中  $\mathbf{A} = \mathbf{T}^{-1} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{T}$ . 」

这样在直观上就定义了对可对角化矩阵的任意函数演算, 但是这样的定义并不是万能的: 以下是一个可能存在的问题:

**例** (坏例子). 考虑这样的函数  $f: \lambda \mapsto |\lambda|^2$ , 那么按照最直观的想法<sup>①</sup>,  $f(\mathbf{A})$  应该是  $\mathbf{A}$  的“模”, 也就是  $\mathbf{A}^* \mathbf{A}$ . 但是:

$$\mathbf{A}^* \mathbf{A} \neq \mathbf{A} \mathbf{A}^*, \quad \text{除非 } \mathbf{A} \text{ 正规}.$$

这样子, 我们的函数演算 (暂行款) 实际上会诱导出这样一个问题:  $f(\mathbf{A})g(\mathbf{A}) \neq (fg)(\mathbf{A})$ . 也就是说, 它没有乘性. 而这是万万不可的, 因为我们在 Hilbert 空间很难将算子对角化, 如果要用多项式逼近连续函数的话, 函数演算的乘性是必须的.

那么我们应该怎么办呢? 历史给出的回答是只考虑正规, 也就是  $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$  的算子. 这样我们的函数演算暂且可以保持乘性继续运行下去. 那么测度黏土在其中起到了什么作用呢?

我们知道, 对待矩阵这些有限维的东西, 我们可以把特征值直接加起来, 或者说, 特征值有一个正的“权”, 这个权就是特征子空间的维数, 但是对 Hilbert 空间上的正规算子, 仅仅用特征值是没办法完整描述它的, 我们会遇到一个“连续的特征值”, 或者是“带有粘性的特征值”, 我们称之为谱.

**定义** (谱). 假定  $T: \mathcal{H} \rightarrow \mathcal{H}$  是一个 Hilbert 空间上的线性算子, 那么其特征值 (点谱, 离散谱) 定义为:

$$\sigma_{\text{pt}}(T) := \{ \lambda \in \mathbb{C} \mid \ker(\lambda - T) \neq \{0\} \}.$$

一般的谱定义为:

$$\sigma(T) := \{ \lambda \in \mathbb{C} \mid \lambda - T \text{ 没有连续线性逆} \}.$$

---

<sup>①</sup> 经由物理系同学的验证 (雾).

事实上, 由 Banach 逆算子定理, 如果  $T$  是连续的,  $\lambda - T$  是双射, 那么  $(\lambda - T)^{-1}$  也一定是连续的, 故此时  $\lambda \notin \sigma(T)$ . 我们在此无意讨论谱的构成, 我们现在只需着眼于“点谱”这个名字上来. 事实上, 我们可以给出下面的直观:

点谱上的每个点是某个测度下的非零测集, 也就是说:

$$\mu(\{\lambda\}) \neq 0.$$

其中  $\mu$  是某一个特殊的测度,  $\lambda \in \sigma_d(T)$ , 这件事和  $\lambda$  有一个正维数的特征子空间直接相关.

**例** (一个没有特征值的算子). 考虑 Hilbert 空间  $\mathcal{L}^2([0, 1])$ . 以及算子  $Tf(x) = xf(x)$ . 这样的算子没有特征值, 但是它的谱是  $[0, 1]$ . 这可以这么验证: 若  $(T - \lambda)f = 1 \in \mathcal{L}^2([0, 1])$ ,  $\lambda \in [0, 1]$ , 则  $f(x) \stackrel{\text{a.e.}}{=} 1/(x - \lambda)$ , 而后者不是平方可积的, 因此  $Tf - \lambda$  不是满射, 更别谈逆了.

好, 由于涉及的定义比较多, 我们暂且在这里回顾一下.

一开始, 我们认为对可对角化矩阵的函数演算可以简化到对特征值的演算上来, 但是这样的演算并不一定具有乘性. 稍退一步, 我们现在只考虑正规矩阵的演算, 并寄希望其能满足乘性. 推广到 Hilbert 空间上的情形, 我们注意到“离散的特征值”只是有限维下面才会有的东西, 无限维的情形下, 特征值, 或者谱是粘在一块的. 我们并没有办法将谱离散地分隔出来然后再加到一起, 这也就意味着我们不能像有限维的情形下定义函数演算. 我们需要用一些函数演算“粘化”这样的方法了.

至少有三四种简单的方法可以用来定义函数演算, 笔者选取了测度介入的其中一种, 因为这样的演算可以在几乎同一种定义的前提下推广到无界算子和无界函数演算上来. 既然如标题所需我们需要涉及到测度, 因此, 我们最重要的想法是将测度还原出来. 我们的方法是这样的: 首先, 还是利用一般的多项式演算将函数演算推广到连续的情形, 接下来, 我们找到这些演算和测度的关系, 就可以用测度和积分理论去讨论无界的函数演算了.

## 第一步: 从多项式到连续函数

让我们用完整的形式叙述 Weierstrass 的多项式逼近定理及其推广”

**定理** (Weierstrass 多项式逼近定理). 令  $[a, b] \subset \mathbb{R}$ , 则  $\forall f \in C([a, b])$ , 存在多项式列  $P_n \Rightarrow f$ .

证明. 为了方便假定  $[a, b] = [0, 1]$ . 利用 Bernstein 多项式:

$$B_f(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

由于  $\forall x \in [0, 1]$ , 都有  $\sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = 1$ , 因此:

$$\begin{aligned} & |B_f(x) - f(x)| \\ & \leq \sum_{k=0}^n \left| f\left(\frac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k} \\ & = \left( \sum_{|f(k/n) - f(x)| > \varepsilon} + \sum_{|f(k/n) - f(x)| \leq \varepsilon} \right) \left| f\left(\frac{k}{n}\right) - f(x) \right| \binom{n}{k} x^k (1-x)^{n-k}. \end{aligned}$$

令两个求和分别是  $I_1$  和  $I_2$ . 则  $I_2 \leq \varepsilon \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = \varepsilon$ . 对于  $I_1$ , 我们有:

$$\begin{aligned} I_1 & \leq 2\|f\|_{\sup} \sum_{|f(k/n) - f(x)| > \varepsilon} \binom{n}{k} x^k (1-x)^{n-k} \\ & \leq \frac{2\|f\|_{\sup}}{\varepsilon^2} \sum_{k=0}^n \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k}. \end{aligned}$$

利用归纳法, 可以得到  $\sum_{k=0}^n (k/n - x)^2 \binom{n}{k} x^k (1-x)^{n-k} = x(1-x)/n \leq 1/4$ , 因此  $I_1 \leq \|f\|_{\sup}/2n\varepsilon^2$ . 因此只需  $n$  恰到好处的大即可.  $\square$