第 1 章 拓扑空间与连续映射

1.1 度量空间与连续映射

1.1.1 度量结构

¶ 一些定义

在数学中，度量是“距离”这个概念的抽象化，最早由法国数学家 Fréchet [1]引入：

定义 1.1.（度量空间）

若 X 是一个集合，而映射

$$d: X \times X \to \mathbb{R}$$

满足如下条件：对于任意 $x, y, z \in X$，均有

(a) (**正定性**) $d(x,y) \geq 0$, 而且 $d(x,y) = 0 \iff x = y$,

(b) (**对称性**) $d(x,y) = d(y,x)$,

(c) (**三角不等式**) $d(x,y) + d(y,z) \geq d(x,z)$,

则我们称 (X,d) 为一个**度量空间**，且称 d 为 X 上的一个 **度量**。 ♣

注 1.2. 不难发现，$d(x,y) \geq 0$ 是其他几条公理的直接推论。另一方面，通过减弱上述条件中的某一个或某几个，也可以得到度量空间的某些合理推广。

我们可以把欧氏空间中的很多定义推广到抽象度量空间中去，比如

定义 1.3.（有界性）

设 (X,d) 是一个度量空间，而 $A \subset X$ 为一个子集。我们称

$$\operatorname{diam}(A) := \sup_{x,y \in A} d(x,y)$$

为集合 A 的**直径**。若 $\operatorname{diam}(A) < +\infty$，我们称 A 为**有界集**，否则称 A 为**无界集**。特别地，如果 $\operatorname{diam}(X) < +\infty$，我们称 (X,d) 为**有界度量空间**。 ♣

在度量空间 (X,d) 中，我们也可以自然定义开球、闭球和球面等几何概念：

定义 1.4.（球与球面）

以 (X,d) 中点 x_0 为中心，半径为 r 的**开球**、**闭球**和**球面**分别定义为

$$B(x_0, r) = \{x \in X \mid d(x, x_0) < r\},$$

$$\overline{B(x_0, r)} = \{x \in X \mid d(x, x_0) \leq r\},$$

$$S(x_0, r) = \{x \in X \mid d(x, x_0) = r\}.$$

♣

[1] 弗雷歇 (René Maurice Fréchet，1878–1973)，法国数学家，现代分析的奠基人之一，对概率率统计等领域也有重要贡献。他师承著名数学家 Jacques Hadamard，1906 年毕业于法国高等师范学院。在其博士论文中，弗雷歇提出了抽象度量空间的概念，开启了在抽象空间中研究分析学的新时代。

有了开球的概念，又可以如欧氏空间中一样定义开集和闭集的概念：

定义 1.5.（开集和闭集）

设 (X,d) 是一个度量空间，$U \subset X$. 如果对于任意 $x \in U$，均存在 $\varepsilon > 0$ 使得

$$B(x,\varepsilon) \subset U,$$

则我们称子集 $U \subset X$ 是一个 **开集**。如果子集 $F \subset X$ 的补集 $F^c = X \setminus F$ 是开集，则称 F 为一个**闭集**。♣

不难由定义验证度量空间中的开球都是开集，而闭球都是闭集。

¶ 一些例子

度量空间的概念来源于欧氏空间的距离结构。下述例子表明度量空间事实上广泛存在于各个不同的数学分支中：

例 1.6.

(1) **(离散度量)** 在任意集合 X 上，均可定义如下度量，称为离散度量：

$$d(x,y) = \begin{cases} 0 & x = y, \\ 1 & x \neq y. \end{cases}$$

不妨思考一下：离散度量空间 $(X, d_{discrete})$ 中的开球、闭球、球面分别是什么？

(2) **($\mathbb{R}^n$ 上的各种度量)** 在 $X = \mathbb{R}$ 上，我们不仅有如上定义的离散度量，还有

- 最简单的绝对值度量 $d(x,y) = |x-y|$.
- 有界度量 $\bar{d}(x,y) = \min\{|x-y|, 1\}$ (或 $\bar{d}(x,y) = \frac{|x-y|}{1+|x-y|}$).

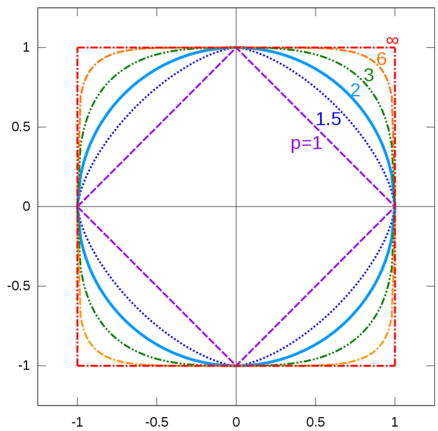
更一般地，在 $X = \mathbb{R}^n$ 上，我们有

- (通常的欧氏度量) $d_2(x,y) = \sqrt{(x_1-y_1)^2 + \cdots + (x_n-y_n)^2}$.
- (l^1 度量) $d_1(x,y) = |x_1-y_1| + \cdots + |x_n-y_n|$.
- (l^∞ 度量) $d_\infty(x,y) = \sup\{|x_1-y_1|, \cdots, |x_n-y_n|\}$.

事实上，这些度量都是如下 l^p $(1 \le p \le \infty)$ 度量的特例：

$$d_p(x,y) := (|x_1-y_1|^p + \cdots + |x_n-y_n|^p)^{1/p}.$$

对于不同的 p，下图展示了度量空间 $(\mathbb{R}^2, l^p)$ 中的单位球的形状（来自维基百科）



(3) (**$\mathbb{R}^{\mathbb{N}}$ 上的各种度量**) 在无限笛卡儿积

$$X=\mathbb{R}^{\mathbb{N}}:=\{(x_1,x_2,\cdots,x_n,\cdots)\mid x_n\in\mathbb{R}\}^{2}$$

上，我们不能像上面一样定义“l^p 度量”，因为涉及到的求和有可能是发散的。但是，我们可以用 $\mathbb{R}$ 上的有界度量 $\bar{d}$ 来解决收敛性问题。比如，我们有

- (一致度量)
$$d((x_n)_{n\in\mathbb{N}},(y_n)_{n\in\mathbb{N}}):=\sup_{n\in\mathbb{N}}\bar{d}(x_n,y_n).$$
- (无穷乘积度量):
$$d((x_n)_{n\in\mathbb{N}},(y_n)_{n\in\mathbb{N}}):=\sum_{n=1}^{\infty}2^{-n}\bar{d}(x_n,y_n).$$

此外，在特定问题中我们需要使用无穷序列之间的 l^p 度量，此时我们可以通过“把 l^p 度量限制在合适的子集上”来解决收敛性问题：

- (l^p 空间, $1\le p\le\infty$) 考虑子空间
$$X=l^p(\mathbb{R}):=\left\{(x_n)_{n\in\mathbb{N}}\;\middle|\;\|x\|_p:=\left(\sum_n|x_n|^p\right)^{1/p}<+\infty\right\}\subset\mathbb{R}^{\mathbb{N}},$$
现在我们可以如下定义 l^p 度量：
$$d((x_n)_{n\in\mathbb{N}},(y_n)_{n\in\mathbb{N}})=\|x-y\|_p.$$
- (Hilbert 立方体) 取
$$X=\prod_n[0,1/n]\subset\mathbb{R}^{\mathbb{N}}.$$
这是 $l^2(\mathbb{R})$ 的子集，所以自然继承一个 l^2 度量。

(4) (**函数空间上的度量**) 在区间 $[a,b]$ 上全体连续函数空间 $C([a,b])$ 上，我们有

- (L^1 度量)
$$d(f,g)=\int_a^b|f(x)-g(x)|dx.$$
- (L^∞ 度量)
$$d(f,g)=\sup_{x\in[a,b]}|f(x)-g(x)|.$$
- (L^2 度量)
$$d(f,g)=\left(\int_a^b|f(x)-g(x)|^2\mathrm{d}x\right)^{1/2}.$$

它们都是 L^p 度量 ($1\le p\le\infty$) 的特例：

$$d(f,g)=\left(\int_a^b|f(x)-g(x)|^p\mathrm{d}x\right)^{1/p}.$$

更进一步地，我们可以在 $[a,b]$ 上的全部 k 阶连续可导函数所构成的集合上，定义

[2]和集合论中一样，我们使用记号 $Y^X=\{f:X\to Y\}$. 注意集合 $Y^{\mathbb{N}}$ 与 Y 中的点列全体构成的集合是相同的。有些作者更喜欢使用记号 $\mathbb{R}^\omega$, 而不是 $\mathbb{R}^{\mathbb{N}}$.

$W^{k,p}$ 度量

$$d(f,g)=\left(\sum_{i=0}^{k}\int_a^b |f^{(i)}(x)-g^{(i)}(x)|^p \mathrm{d}x\right)^{1/p}.$$

这些度量在偏微分方程理论中有着重要价值。

(5) **(词度量)** 设 G 是一个群，$S\subset G$ 是一个生成子集 [3]，则 S 诱导的**词度量**为

$$d(g_1,g_2)=\min\{n:\exists s_1,\cdots,s_n\in S\cup S^{-1}\text{使得}g_1\cdot s_1\cdots s_n=g_2\}.$$

(6) **(图度量)** 设 $G=(V,E)$ 是一个连通图，则在顶点集 V 上可定义如下的图度量：

$$d(v_1,v_2)=\min\{n:G\text{中存在长度为}n\text{的道路连接}v_1\text{和}v_2\}.$$

事实上，(5) 中定义的词度量跟相应的 Caylay 图 $\Gamma(G,S)$ 上的图度量是一致的。

(7) (p **进度量**) 设 p 是一个素数. 那么，任意 $0\neq x\in\mathbb{Q}$ 可以唯一地表示为

$$x=p^n\frac{r}{s}$$

其中 $n,r,s\in\mathbb{Z}$ 且 $s>0$. 由此我们可以定义 $\mathbb{Q}$ 上的 p 进范数

$$|x|_p=p^{-n}\quad(\text{令}|0|_p=0),$$

而 $\mathbb{Q}$ 上的 p **进度量**则定义为

$$d(x_1,x_2):=|x_1-x_2|_p.$$

该度量在算术几何中有着重要意义。

(8) **(Hausdorff 度量)** 在 $\mathbb{R}^n$ 中的全体有界闭集构成的集合上，定义 Hausdorff 度量如下，

$$d(A,B)=\inf\{\varepsilon\geq 0:A\subset B_\varepsilon\text{ 且}B\subset A_\varepsilon\},$$

其中 $A_\varepsilon=\{y\in\mathbb{R}^n:\exists x\in A\text{ 使得}|x-y|\leq\varepsilon\}$ 是 A 的 ε-邻域。

¶ 从已有的度量空间构造新空间

我们也可以从已有的度量空间构造新的度量空间，最常见的构造是子集继承原空间度量而得到的子度量空间，以及在乘积空间上通过合适的方法定义的乘积度量空间。

命题 1.7.（子空间度量）

设 (X,d) 是度量空间，$Y\subset X$ 为子集，则

$$d_Y:=d|_{Y\times Y}$$

是 Y 上的一个度量。

证明 这是很显然的：对于任意 $y_1,y_2,y_3\in Y\subset X$，我们有

- $d_Y(y_1,y_2)=0\iff y_1=y_2$.
- $d_Y(y_1,y_2)=d(y_1,y_2)=d(y_2,y_1)=d_Y(y_2,y_1)$.
- $d_Y(y_1,y_3)=d(y_1,y_3)\leq d(y_1,y_2)+d(y_2,y_3)=d_Y(y_1,y_2)+d_Y(y_2,y_3)$.

□

[3]如果 G 中的任何元素都可以写成 $S\cup S^{-1}$ 中有限多个元素的乘积，则子集 $S\subset G$ 称为**生成子集**。

更一般地，任意给定一个单射 $f: Y \to X$，那么我们可以将 Y 等同于子集 $f(Y) \subset X$，从而可以通过 X 上的度量 d_X 给出在 Y 上的诱导度量

$$d(y_1, y_2) := d_X(f(y_1), f(y_2)).$$

下面我们解释如何在笛卡尔积上构造合理的度量：

命题 1.8.（乘积度量）

如果 $(X_1, d_1), (X_2, d_2)$ 是度量空间，那么

$$d((x_1, x_2), (y_1, y_2)) := d_1(x_1, y_1) + d_2(x_2, y_2)$$

是 $X_1 \times X_2$ 上的度量.

证明 验证如下：对于 $X_1 \times X_2$ 中的任意点 $(x_1, x_2), (y_1, y_2)$ 和 (z_1, z_2)，有

- $$\begin{aligned} d((x_1, x_2), (y_1, y_2)) = 0 &\Longleftrightarrow d_1(x_1, y_1) = 0 \text{ 且} d_2(x_2, y_2) = 0 \\ &\Longleftrightarrow x_1 = y_1 \text{且} x_2 = y_2. \end{aligned}$$

- $$\begin{aligned} d((x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2) &= d_1(y_1, x_1) + d_2(y_2, x_2) \\ &= d((y_1, y_2), (x_1, x_2)). \end{aligned}$$

- $$\begin{aligned} d((x_1, x_2), (z_1, z_2)) &= d_1(x_1, z_1) + d_2(x_2, z_2) \\ &\leq d_1(x_1, y_1) + d_1(y_1, z_1) + d_2(x_2, y_2) + d_2(y_2, z_2) \\ &= d((x_1, x_2), (y_1, y_2)) + d((y_1, y_2), (z_1, z_2)). \end{aligned}$$

□

注 1.9. 不同于子集上的子空间度量，对于度量空间的笛卡尔积我们可以用许多不同的方法给赋予“乘积度量”。例如，我们也可以定义

$$d((x_1, x_2), (y_1, y_2)) := \sqrt{d_1(x_1, y_1)^2 + d_2(x_2, y_2)^2}$$

可以验证这是 $X_1 \times X_2$ 上的一个度量. 更一般地，若 $(X_1, d_1), \cdots, (X_n, d_n)$ 都是度量空间，则对于任意 $1 \leq p \leq \infty$，我们可以用下式定义 $X_1 \times \cdots \times X_n$ 上的 l^p **型乘积度量**：

$$d_p((x_1, \cdots, x_n), (y_1, \cdots, y_n)) := \Big(d_1(x_1, y_1)^p + \cdots + d_n(x_n, y_n)^p\Big)^{1/p}.$$

为了在可数多个度量空间 (X_n, d_n) 的笛卡尔积 $\prod_{n=1}^{\infty} X_n$ 上显式地给出“乘积度量”，我们可以如例1.6(3) 一样，首先将每个 d_n 转换为有界度量

$$\bar{d}_n(x, y) = \min\{d_n(x, y), 1\}.$$

然后在笛卡尔积 $\prod_{n=1}^{\infty} X_n$ 上定义“l^∞ 乘积度量”

$$d_u((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \sup_{n\in\mathbb{N}} \bar{d}_n(x_n, y_n)$$

该度量被称为乘积空间 $\prod_{n=1}^{\infty} X_n$ 上的**一致度量**。

¶ 等距同构、嵌入和 Lipschitz 映射

正如我们提到过的，对于给定结构的集合，我们希望研究集合之间保结构的映射：

定义 1.10.（等距同构）

设 (X,d_X) 和 (Y,d_Y) 为度量空间。如果存在双射 $f:(X,d_X)\to(Y,d_Y)$，使得对于任意 $x_1,x_2\in X$，都有

$$d_X(f(x_1),f(x_2))=d_Y(x_1,x_2),$$

则我们称 f 为一个 **等距同构**，并称度量空间 (X,d_X) 和 (Y,d_Y) 是 **等距同构的**. ♣

因为等距同构的度量空间具有完全相同的度量性质，我们视等距同构的度量空间视为 **相同的** 度量空间。当然，绝大部分度量空间不是等距同构的。以下两个概念稍微放宽了限制：

定义 1.11.（等距嵌入）

如果单射 $f:(X,d_X)\to(Y,d_Y)$ 满足：对于任意 $x_1,x_2\in X$，均有

$$d_X(f(x_1),f(x_2))=d_Y(x_1,x_2),$$

则我们称 f 为一个（度量）**等距嵌入**。♣

显然，如果 $f:(X,d_X)\to(Y,d_Y)$ 是一个等距嵌入，那么 (X,d_X) 与 (Y,d_Y) 的子空间 $(f(X),d_Y)$ 是等距同构的。

定义 1.12.（Lipschit 映射）

我们称映射 $f:(X,d_X)\to(Y,d_Y)$ 为一个 Lipschitz 常数为 L 的 **Lipschit 映射**，如果对于任意 $x_1,x_2\in X$，均有

$$d_X(f(x_1),f(x_2))\le Ld_Y(x_1,x_2).$$ ♣

注 1.13. 注意以上概念均强烈依赖于空间上所给定的度量。例如，考虑恒等映射

$$\mathrm{Id}:\mathbb{R}\to\mathbb{R},\quad x\mapsto x.$$

现在考虑 $\mathbb{R}$ 上的两个度量，标准度量 $d(x,y)=|x-y|$ 和有界度量 $\bar{d}(x,y)=\min\{1,d(x,y)\}$，那么作为度量空间之间的映射，

$$\mathrm{Id}:(\mathbb{R},d)\to(\mathbb{R},d)\quad 和\quad \mathrm{Id}:(\mathbb{R},\bar{d})\to(\mathbb{R},\bar{d})$$

都是等距同构，

$$\mathrm{Id}:(\mathbb{R},d)\to(\mathbb{R},\bar{d})$$

是一个 Lipschitz 映射但不是等距同构，而

$$\mathrm{Id}:(\mathbb{R},\bar{d})\to(\mathbb{R},d)$$

不是一个 Lipschitz 映射。

1.1.2 度量空间之间的连续映射

¶ 收敛性和连续性

本课程的主要目标之一是研究连续函数和更一般的连续映射。对于度量空间之间的映射，定义连续性并不难：我们可以像在欧氏空间中一样，首先定义点列收敛的概念，然后通过收敛性来定义连续性。

定义 1.14.（收敛性）

设 (X,d) 是一个度量空间，(x_n) 为 X 中的一个点列。如果存在 X 中的点 x_0，满足：对于任意 $\varepsilon>0$，均存在 $N\in\mathbb{N}$ 使得对于所有 $n>N$，均有

$$d(x_n,x_0)<\varepsilon,$$

则我们称点列 (x_n)（关于度量 d）**收敛**到点 x_0，记为 $x_n\xrightarrow{d}x_0$. ♣

有了收敛性，我们可以自然定义

定义 1.15.（连续性）

设 (X,d_X) 和 (Y,d_Y) 是两个度量空间，$f:X\to Y$ 为一个映射。

(1) 若对于在 X 中收敛到 x_0 的任何点列 (x_n)，像点点列 $(f(x_n))$ 都收敛到 Y 中的点 $f(x_0)$，则我们说映射 $f:X\to Y$ 在 $x_0\in X$ 处是 **连续的**。

(2) 如果映射 f 在每个点 $x_0\in X$ 处都是连续的，我们就称 f 是一个 **连续映射**. ♣

注 1.16. 当我们讨论函数 $f:(X,d_X)\to\mathbb{R}$ 的连续性时，除非另外说明，总是赋予 $\mathbb{R}$ 标准度量。

我们先给出下列映射在一点处连续的等价刻画，其证明跟数学分析中所学的完全一致，故而略去：

命题 1.17.（一点处连续的等价刻画）

$$\begin{aligned}
&\text{映射} f:(X,d_X)\to(Y,d_Y)\ \text{在} x_0\in X\ \text{处连续}\\
\iff&\forall\varepsilon>0,\exists\delta>0\ \text{使得}\forall x\in X, d_X(x,x_0)<\delta,\ \text{我们有} d_Y(f(x),f(x_0))<\varepsilon\\
\iff&\forall\varepsilon>0,\exists\delta>0\ \text{使得} f(B(x_0,\delta))\subset B(f(x_0),\varepsilon)\\
\iff&\forall\varepsilon>0,\exists\delta>0\ \text{使得} B(x_0,\delta)\subset f^{-1}\big(B(f(x_0),\varepsilon)\big).
\end{aligned}$$

♠

由此不难看出度量空间之间的任何 Lipschitz 映射都是连续的。

¶ 连续映射的例子

度量空间之间映射连续性的概念是欧氏空间之间连续映射的自然推广。为了更好地理解度量空间中连续性的含义，让我们讨论一些简单的例子。

例 1.18.

(1) 设 (X,d) 是任何度量空间。

- 对于任何固定的 $\bar{x} \in X$, 函数
$$d_{\bar{x}} : X \to \mathbb{R}, x \mapsto d_{\bar{x}}(y) := d(x, \bar{x})$$
是连续的。
证明 对于 $\forall \varepsilon > 0, \forall x_0 \in X$, 和 $\forall x \in X$ 满足 $d(x, x_0) < \varepsilon$, 我们有
$$|d_{\bar{x}}(x) - d_{\bar{x}}(x_0)| = |d(x, \bar{x}) - d(x_0, \bar{x})| \leq d(x, x_0) < \varepsilon.$$
(所以函数 $d_{\bar{x}}$ 实际上是一个 Lipschitz 常数为 1 的 **Lipschitz 映射**.) □
- 更一般地，对于任何子集 $A \subset X$，我们可以定义
$$d_A : X \to \mathbb{R},\ x \mapsto d_A(x) := \inf\{d(x, y) : y \in A\}.$$
可以证明 d_A 是连续的。
证明概要 首先利用三角不等式证明
$$|d_A(x) - d_A(y)| \leq d(x, y).$$
然后可得要证结论。细节留作练习。 □
- 如果我们赋予 $X \times X$ 命题 1.8中的“乘积度量”$d_{X\times X}$, 那么函数 $d : X \times X \to \mathbb{R}$ 就是 $(X \times X, d_{X\times X})$ 上的连续函数。证明留作习题。

(2) 在空间 $X = C([a, b])$ 上赋予 l^∞ 度量
$$d(f, g) := \sup_{x\in[a,b]} |f(x) - g(x)|.$$
那么“积分映射”
$$\int : X \to \mathbb{R},\ f \mapsto \int_a^b f(x)\mathrm{d}x$$
是连续的，因为
$$\left| \int_a^b f(x)\mathrm{d}x - \int_a^b g(x)\mathrm{d}x \right| \leq \int_a^b |f(x) - g(x)|\mathrm{d}x \leq (b - a) \cdot d_X(f, g).$$

(3) 设 X 为任意集合，d_X 为 X 上的离散度量。设 (Y, d_Y) 是任意度量空间。
- 任何映射 $f : X \to Y$ 都是连续的。
证明 对于任意 $\varepsilon > 0$，我们取 $\delta = 1$ 即可：若 $x, x_0 \in X$ 满足 $d_X(x, x_0) < 1$，根据离散度量的定义，我们有 $x = x_0$，从而 $d_Y(f(x), f(x_0)) = 0 < \varepsilon$. □
- 问：哪些映射 $f : Y \to X$ 是连续的?
答：只有“局部常值映射” [4]是连续的。
证明 显然，如果 f 是局部常值的，那么它是连续的。
反之，假设 $f : Y \to X$ 在 y_0 处连续，那么存在 $\delta > 0$ 使得对于任意满足 $d_Y(y, y_0) < \delta$ 的 $y \in Y$，我们有 $d_X(f(y), f(y_0)) < 1$。但 d_X 是离散度量，所以 $f(y) = f(y_0)$，即 f 在 y_0 附近是常值映射。 □

[4]映射 $f : Y \to X$ 是 **局部常值**是指：对于任意 $y_0 \in Y$，都存在 $\delta > 0$，使得对于所有满足 $d(y_0, y) < \delta$ 的 y，都有 $f(y) = f(y_0)$.