Functional Analysis 1.1 Introduction 01

Chen Gong

27 July 2020

目录

小编的主要课题是做人工智能的，人工智能中脱离不开泛函分析的知识，经常会由于泛函中的一些知识不熟悉而卡壳。一直就想系统的学习一个泛函分析了，小编终于开始好好看看泛函分析了。课程使用的是哔哩哔哩上内蒙古大学，孙炯老师的泛函分析课程。参考书籍为 Erwin Kreyszig 写的“泛函分析导论及应用”的译本。PDF 版本的地址为，`https://github.com/2019ChenGong/Functional-Analysis`。

1 Background

泛函分析这门课的主要目的是：**了解和掌握空间理论 (包括距离空间、赋范空间、内积空间) 和线性算子理论 (包括线性算子空间、线性算子谱分析) 中的基本概念和基本理论**。小编在机器学习中经常会遇到一些很装逼的名词，比如希尔伯特空间，紧集等，后来发现这些名词都是泛函分析，所以决定要好好学学泛函分析。

下面从分析、代数、微分方程中的一些例子展开讨论，1. 学习如何从问题中抽象出泛函分析中的一些基本概念和基本理论；2. 学习如何使用类比、联想等方法，从问题中归纳出一些基本的数学思想方法，进而去解决未知的问题。个人觉得这个数学思路非常的重要。其实我们之前看线性代数，解析几何，好像之间没有什么关系，实际上他们之间是有着潜在的联系。**泛函分析正是从这些类似的东西中探寻一般的、真正属于本质的东西，把它们抽象化并加以统一处理**。

1.1 泛函分析的研究对象和方法

我们的数学研究中，主要可以分成以下两个部分。

1. 函数 → 映射。函数可以表示为: $x \in X \rightarrow f(x) \in Y$。进一步可以表示为，一个空间 X 到另一个空间 Y 的映射。

2. 运算 (算子)。我们之前学的微分，积分都是运算。而且其都具有可加性，所以也可称为线性运算。运算实际就是一种映射，比如，通过以下运算，可以将一个函数映射为另一个函数。

$$\begin{aligned} \sin x &\underset{'}{\rightarrow} \cos x \\ \cos x &\underset{\int}{\rightarrow} \sin x \end{aligned} \tag{1}$$

所以，可以看到，**函数和运算其实都是一种映射**。这样我们就可以用统一的观点来看待函数和运算，$X \rightarrow Y$。所以，综上所述，**泛函研究的对象有两个，函数和运算**。函数用来描述一个对象，而运算则建立了函数之间的关系。

泛函分析研究的方法:

泛函分析是 20 世纪初从变分法、微分方程、积分方程、函数论、量子物理等研究中发展起来的一门数学分支学科。

1. 泛函分析综合分析、代数、几何的观点和方法来研究**无穷维空间**上的函数、算子和极限理论，处理和解决数学研究中最关心的一些基本问题。我们之前研究的问题，线性代数等都是位于有限维度的空间中，而泛函研究的问题都是位于无穷维度中。

2. 泛函分析的特点是它不但把古典分析的基本概念和方法一般化了，而且还把这些概念和方法几何化了。之所以几何化是把问题给具体化，将函数和图像结合起来，使问题变得更简单。随着笛卡尔

坐标系的建立，解析几何的创立，人们把代数问题几何化，把几何问题代数化，为初等数学的许多问题开辟了全新的研究模式。例如，$x^2+y^2=a^2$ 可以转换成，平面以原点为中心，以 a 为半径的圆。

1.2 泛函问题的建立过程

而在泛函分析中，往往将解决几何问题的模式进行推广，

1. 建立一个新的空间框架。包含，**空间中的元素，**$x \to f(x)$**；定义在函数上的运算** $X \to Y$。而且，以后特别要注意的是空间中的元素是什么，空间是什么样的结构 (距离、范数、内积)？因为元素不会是简单的堆积到一起，它肯定有结构，我理解就是元素之间关系。

而之后，我们可能建立一个空间，空间中的元素是函数，比如 $\sin x, \cos x, e^x$ 等等。而空间的结构就是这些函数之间的距离，范数，内积等等。我们参照的方法就是平面解析几何。

2. 在新的空间框架下，研究解决分析、代数、几何中的问题。(把分析中的问题结合几何、代数的方法加以处理。)

在解析几何、线性代数中，研究的是：有限维 (n 维) 空间中的运动和映射：从 n 维空间到 m 维空间的线性运算。

在泛函分析 (线性泛函分析) 中，主要研究的是：从无穷维空间到无穷维空间的线性运算。主要是泛函研究的集合中的元素是函数，函数的性质比较复杂，你不能用几个特征来描述一个函数，想想之前学的 Taylor 展开，我们使用无穷级数来近似函数，所以通常需要无穷维向量来描述一个函数。

于是大家想想，泛函中无穷维空间和有限维会是一样的吗？废话，肯定是不一样的，要是一样的还学什么泛函？它们肯定有一样的地方，也肯定有不一样的地方。我们希望可以类比一样的地方，并未解决不同的地方提供思路。

于是要特别关注，无穷维空间的性质，与有限维空间的区别：。

1. **无穷维空间的收敛性问题** (加法与无穷级数的区别)。比如三个数 $a+b+c=d$ 这必然是收敛的。而 $x_1+x_2+\cdots$ 这个是收敛的吗？真不好说。

2. 相同点在于其都是线性空间，大家不知道还记不记得线性空间的重要特征，就是可加性和数乘。

2 有限维空间的坐标分解和算子分解

我们通过一些熟悉的例子，研究和探讨如何类比地建立起这样的空间框架。我们刚刚泛泛的讲了，**泛函分析的目的是什么，是研究无穷空间中的元素，无穷空间中的运算**。我们希望借助过去的东西来学习，这种借鉴是非常重要的。

我们希望把有限维空间的研究方法和结论自然地推广到无穷维空间。从分析、代数中的问题出发，引出泛函分析研究的思想方法。这些例子都是我们熟悉的，我们希望从中领悟到数学处理问题的基本思路。进而把他们推广到我们未知的领域上去。数学研究中最忌讳的就是从一个概念钻到另一个概率，从而忽略本身需要解决的问题。

我们首先对简单的问题建立空间框架。之后，我们要对函数和运算建立空间框架，我们看看有限维空间框架的建立会不会对无限维建立有影响。

2.1 $\mathbb{R}^3$(三维实空间) 中的向量分解.

1. 在 $\mathbb{R}^3$ 维空间建立正交坐标系，

$$\vec{i}=(1,0,0),\vec{j}=(0,1,0),\vec{k}=(0,0,1)$$

2. 建立了空间中的结构，

$$a\cdot\vec{b}=(\vec{a},\vec{b})=a_1b_1+a_2b_2+a_3b_3$$

其中，$\vec{a}=(a_1,a_2,a_3)$, $\vec{b}=(b_1,b_2,b_3)$，且

$$\vec{a}\cdot\vec{b}=|\vec{a}||\vec{b}|\cdot\cos\theta$$

这样我们就定义出了角度的概率。于是对于空间中的任意一个向量 $a\in\mathbb{R}^3$，用三个有序的数字表示，

$$\vec{a}=(a_1,a_2,a_3) \tag{2}$$

其中，$a_1=(\vec{a},\vec{i}),a_2=(\vec{a},\vec{j}),a_3=(\vec{a},\vec{k})$。现在内积不仅仅表示一种运算，而且表示一种投影。**投影的概念非常的重要，我们以后要将其引申到一般的函数空间或者算子空间上去**。于是，$\vec{a}$ 可以写成，

$$\begin{aligned}\vec{a}&=a_1\vec{i}+a_2\vec{j}+a_3\vec{k}\\&=(\vec{a},\vec{i})\vec{i}+(\vec{a},\vec{j})\vec{j}+(\vec{a},\vec{k})\vec{k}\end{aligned} \tag{3}$$

并且其模长，或者说是范数等于投影的平方和。

$$|\vec{a}|^2=|(\vec{a},\vec{i})|^2+|(\vec{a},\vec{j})|^2+|(\vec{a},\vec{k})|^2 \tag{4}$$

那么扩展到 n 维欧几里得空间中，我们也可以得到类似的结果，

$$\begin{aligned}\vec{a}&=(a_1,a_2,\cdots,a_n)\\\vec{a}&=a_1\vec{e}_1+\cdots+a_n\vec{e}_n\\&=(\vec{a},\vec{e}_1)\,\vec{e}_1+\cdots+(\vec{a},\vec{e}_n)\,\vec{e}_n\\|\vec{a}|^2&=\sum_{i=1}^{n}|(\vec{a},\vec{e}_i)|^2\end{aligned} \tag{5}$$

其中，$\vec{a}\in\mathbb{R}^n$，而且 $(e_1,e_2,\cdots,e_n)$ 是 $\mathbb{R}^n$ 空间中的一组标准正交基。我们刚刚将空间中的向量投影到三个正交向量上去，实际上是使问题变简单了。可能大家还没有深刻的感受到问题哪里变简单了。这样的方法同样可以类推到线性变换 (映射) 中，大家就可以深刻感受一下了。

2.2 线性变换 A 按坐标分解

其中，A 是从 $\mathbb{R}^4$ 到 ^{4} 的线性变换。并且 A 是一个对称矩阵，

$$A=\begin{pmatrix}0&1&1&-1\\1&0&-1&1\\1&-1&0&1\\-1&1&1&0\end{pmatrix} \tag{6}$$

其中，$x, y \in \mathbb{R}^4$，

$$A : x \to Ax, \quad Ax = y$$

并且，A 具有以下几条良好的性质，

1. A 是对称的线性变换，

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2$$

   我们可以看到，A 具有分配律和数乘交换律等良好的性质，这在泛函分析里是非常重要的性质。

2. 对称矩阵的特征值都是实数。

3. A 的属于不同特征值的特征向量相互正交。

4. A 可以换成对角矩阵。对称矩阵一定正交相似于一个对角矩阵。

大家注意线性代数是什么，**线性代数是研究线性运算最基本的方式**。而特征值往往反映了矩阵运算最基本的特征。而对矩阵做一些相似的变换，矩阵的特征值是肯定不变的。而我们为什么做相似变换，其实就是希望把矩阵变得简单一些。

具体做法，

1. $|\lambda E - A| = (\lambda - 1)^3(\lambda + 3)$。$\lambda = -3, +1$ 是特征值，其中 $\lambda = +1$ 是三重特征值，$\lambda = -3$ 是单重特征值。

2. $\lambda = +1$ 时，求其基础解系如下:

$$\begin{aligned} a_1 &= (1, 1, 0, 0) \\ \alpha_2 &= (1, 0, 1, 0) \\ \alpha_3 &= (-1, 0, 0, 1) \end{aligned}$$

3. 该基础解系不正交，将其单位正交化 (施密特正交法):

$$\begin{aligned} \beta_1 &= \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0, 0\right) \\ \beta_2 &= \left(\tfrac{1}{\sqrt{6}}, -\tfrac{1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}, 0\right) \\ \beta_3 &= \left(-\tfrac{1}{\sqrt{12}}, \tfrac{1}{\sqrt{12}}, \tfrac{1}{\sqrt{12}}, \tfrac{3}{\sqrt{12}}\right) \end{aligned}$$

   当 $\lambda = -3$ 时，可得: $\beta_4 = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$。而 $\beta_1, \beta_2, \beta_3, \beta_4$ 就是 $\mathbb{R}^4$ 中的一组标准正交基了。在这个正交基下，A 就成对角矩阵了，

$$A \sim A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} \tag{7}$$

$$T = (\beta_1, \beta_2, \beta_3, \beta_4)_{4\times 4} \tag{8}$$

$$T^{-1}AT = A_1 = \begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} \tag{9}$$

那么原矩阵和这个对角矩阵是相似的，大家是不是感觉到，这个矩阵多简单呀。

注 1，在新的坐标系 $(\beta_1, \beta_2, \beta_3, \beta_4)$ 下，线性变换 A 有最简单的标准型。

注 2，在每一个特征子空间上 (新的坐标系对应的一维子空间上)，**A 作用的形式是最简单的（放大、缩小特征值的倍数)**。

$$A\beta_1 = \beta_1, A\beta_2 = \beta_2, A\beta_3 = \beta_3, A\beta_4 = -3\beta_4$$

详细讲到这里，大家对特征值的概念有了更加清晰的认识了。**特征值就是矩阵在基向量上放大或缩小的倍数，和二维空间中的在坐标系上的投影长度一样**。

4. 在空间中构建一组新的标准正交基为 $(\beta_1, \beta_2, \beta_3, \beta_4)$，则

$$\vec{x} = (a_1, a_2, a_3, a_4) = a_1\beta_1 + a_2\beta_2 + a_3\beta_3 + a_4\beta_4$$

其中，

$$\begin{aligned} a_1 = (x, \beta_1), a_2 = (x, \beta_2) \\ a_3 = (x, \beta_3), a_4 = (x, \beta_4) \end{aligned}$$

是 $\vec{x}$ 在 $(\beta_1, \beta_2, \beta_3, \beta_4)$ 上的投影。那么有

$$\begin{aligned} Ax &= A(a_1\beta_1 + a_2\beta_2 + a_3\beta_3 + q_4\beta_4) \\ &= a_1A\beta_1 + a_2A\beta_2 + a_3A\beta_3 + a_4A\beta_4 \\ &= a_1\beta_1 + a_2\beta_2 + a_3\beta_3 - 3a_4\beta_4 \\ &= y = (a_1, a_2, a_3, -3a_4) \end{aligned} \tag{10}$$

那么，我们就得到了新的向量在新空间中的表示方法。实际上和三维空间中是一样的，特征值就代表了在正交基向量上的长度。

5. 矩阵 A 确定了一组标准正交基 $(\beta_1, \beta_2, \beta_3, \beta_4)$。只要知道了 $\vec{x}$ 在 $(\beta_1, \beta_2, \beta_3, \beta_4)$ 上的投影 $(a_1, a_2, a_3, -3a_4)$，到 A 的作用方式就一目了然了。

$$\begin{aligned} Ax &= (\lambda_1 a_1, \lambda_2 a_2, \lambda_3 a_3, \lambda_4 a_4) \\ &= (a_1, a_2, a_3, -3a_4) \end{aligned}$$

假设 (P_1, P_2, P_3, P_4) 是在 $(\beta_1, \beta_2, \beta_3, \beta_4)$ 上的投影算子。投影向量合上放缩大小，就可以得到新坐标下的向量。则有，

$$\begin{aligned} A &= P_1 + P_2 + P_3 - 3P_4 \\ &= \lambda_1 P_2 + \lambda_2 P_2 + \lambda_3 P_3 + \lambda_4 P_4 \end{aligned}$$

其中，$\lambda_1 = \lambda_2 = \lambda_3 = 1, \lambda_4 = -3$。这样我们就将线性变换 A 分解成 4 个投影变换 (算子) 的线性组合。通过以上的步骤，我们就把一个很复杂的变换，在这个特征值特征向量下变得非常简单了。而**数学处理问题的原则就是把复杂的问题简单化**。

3 无穷维空间的类比和联想

泛函分析要研究的对象是函数、运算。微分、积分运算，它们作用的对象是函数。微分、积分运算与 $\mathbb{R}^n$ 空间中线性变换 A 相比较，

1. 相同之处: 线性运算;

2. 不同之处: A 把一个 n 维向量变成 n 维 (或 m 维) 向量。微 (积) 分把一个函数映射成另一个函数。

3. 函数数不能用有限个数刻画，可能可以用无穷多个数。而函数实际上可以看成是一个无穷维的向量。我们希望通过 "类比和联想"，把有限维空间处理问题的这种方式推广到更一般的空间 (无穷维空间) 中。

参照有限维空间中的情况，我们将考虑无穷维空间中需要具体讨论的问题。

3.1 无穷空间中的几何结构

1. 是否存在坐标系 $(e_1, \cdots, e_n, \cdots)$?

2. 是否具有正交性?

3. 无穷维空间中的向量 x 能不能分解?

$$\begin{gathered}\vec{x} = (a_1, a_2, a_3, a_4, \cdots) \\ = a_1 e_1 + a_2 e_2 + a_3 e_3 + a_4 e_4 + \cdots \\ \|x\|^2 = \sum_i |a_i|^2 ?\end{gathered} \tag{11}$$

其中，$a_i = (x, e_i), i = 1, 2, \cdots$。这些都还能不能成立呢? 这些都是问题了，不是我们已有的结论。

3.2 线性算子的特征和结构

1. 线性算子的性质，有没有对称算子? 比如积分微分什么样的叫对称的。

2. 线性算子 T 能不能分解? 有限维的可以分解，而无穷维的可不可以分解呢?

$$\begin{gathered}A = P_1 + P_2 + P_3 - 3P_4 \\ T ? = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3 + \cdots\end{gathered} \tag{12}$$

   其中，$P_1, P_2, \cdots$ 是在 $e_1, e_2, \cdots$ 上的投影算子。注由于式中有无穷项相加，于是存在是不是收敛的问题，如果收敛, 是在什么意义下的收敛? 这些问题都是我们泛函分析中要讨论的问题，但是这些问题的提出都是很自然的。

3.3 无穷维空间的坐标分解

为了考虑算子的分解，首先要研究函数的分解，函数可以用无穷多个数形成的数组来刻画。为什么考虑这个问题, 因为我们知道在之前讲的矩阵运算的分解, 我们可以用特征值和特征向量来分解。很自然我们想到无穷维空间是不是也可以用一样的方法来刻画。比如，Taylor 展开，

3.3.1 Taylor 展开

如果函数满足很好的性质，则在它的收敛半径内，有，

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+\cdots \tag{13}$$

即函数可以和一个可数无穷数列一一对应，也就是

$$f(x)\sim\left(f(0),\frac{f'(0)}{1!},\frac{f''(0)}{2!},\ldots,\frac{f^{(n)}(0)}{n!},\ldots\right)$$

实际上这和一个向量在 n 维空间的展开没有什么不同。最大的区别就在于，$(x^0,x^1,x^2,\cdots)$ 不是正交基。实际上泰勒展开是比较方便的，有一一对应的概念在里面，可以的是没有几何概念（垂直，正交这些类似的概念）在里面。那么计算起来就有一定的问题在里面。下面我们熟知的 Fourier 展开就是一种在正交系中的展开。

3.3.2 Fourier 级数展开

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kx+b_k\sin kx)$$

其中，$\frac{9y}{2}=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)dx, a_k=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos kx\mathrm{d}x\ b_k=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin kxdx$。那么，$f$ 就可以由无穷多个数来一一对应了，

$$f(x)\sim\left(\frac{a_0}{2},a_1,b_1,\cdots,a_n,b_n,\cdots\right)$$

之前在 Taylor 展开中很遗憾没有正交的几何关系，在这里我们希望引入。大家可以看到其坐标系为，

$$\begin{aligned}&e_0=\tfrac{1}{\sqrt{2\pi}},e_1=\tfrac{1}{\sqrt{\pi}}\cos x,e_2=\tfrac{1}{\sqrt{\pi}}\sin x,e_3=\tfrac{1}{\sqrt{\pi}}\cos 2x\\&e_4=\tfrac{1}{\sqrt{\pi}}\sin 2x,\cdots,e_{2k-1}=\tfrac{1}{\sqrt{\pi}}\cos kx,e_{2k}=\tfrac{1}{\sqrt{\pi}}\sin kx,\cdots\end{aligned}$$

而 $\left(\frac{a_0}{2},a_1,b_1,\cdots,a_n,b_n,\cdots\right)$ **为函数 f 在** $(e_0,e_1,\cdots,)$ **下的坐标**。类似于 $\mathbb{R}^n$ 我们可以在函数空间 $L^2(-\pi,\pi)$ 上定义内积，

$$(f,g)=\int_{-\pi}^{\pi}f(x)g(x)dx \tag{14}$$

这个是怎么来的呢，我在`https://zhuanlan.zhihu.com/p/92648889` 中将函数看成无穷维向量做出过详细的解读。什么是内积，其满足四条性质，1. 正定；2. 交换；3. 分配；4. 可数乘。大学数学学习一定要从一般化的东西里抽象出最基本的本质性的东西，再形成定义。

而且我们在数学分析里学了，

$$\begin{aligned}&\textstyle\int_{-\pi}^{\pi}\frac{1}{\sqrt{\pi}}\cos nx\cdot\frac{1}{\sqrt{\pi}}\cos mxdx=\begin{cases}0, & n\neq m\\1, & n=m\end{cases}\\&\textstyle\int_{-\pi}^{\pi}\frac{1}{\sqrt{\pi}}\sin nx\cdot\frac{1}{\sqrt{\pi}}\sin mxdx=\begin{cases}0, & n\neq m\\1, & n=m\end{cases}\\&\textstyle\int_{-\pi}^{\pi}\frac{1}{\sqrt{\pi}}\cos nx\cdot\frac{1}{\sqrt{\pi}}\sin mxdx=0\end{aligned} \tag{15}$$

即，

$$\int_{-\pi}^{\pi}e_ie_jdx=(e_i,e_j)=\begin{cases}0, & i\neq j\\1, & i=j\end{cases} \tag{16}$$

所以 $\{e_i\}$ 形成空间中的一组标准正交基。这些都是我们从之前知道的基础知识中推导得出的。

4 n 维空间与无穷维空间的对照

在 $\mathbb{R}^n$ 的空间中，$\forall x \in \mathbb{R}^n$，

$$\vec{x} = x_1 e_1 + x_2 e_2 + \cdots \tag{17}$$

其中，$x_1 = (x, e_1), x_2 = (x, e_2), \cdots, x_n = (x, e_n)$。对于函数 f，我们有，

$$f(x) = \frac{a_0}{2} e_0 + a_1 e_1 + b_1 e_2 + \cdots + a_k e_{2k-1} + b_k e_{2k} + \cdots \tag{18}$$

那么我们的下一个问题就是，

$$\begin{aligned}
&\frac{a_0}{2} = (f, e_0) \\
&a_1 = (f, e_1)\,, b_1 = (f, e_2)\,, \cdots \\
&a_k = (f, e_{2k-1})\,, b_k = (f, e_{2k})\,, \cdots
\end{aligned} \tag{19}$$

那么系数是否等于 (f, e_i)?

$$\begin{aligned}
\frac{a_0}{2} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{2\pi} (f(x), 1) \\
&= \frac{1}{\sqrt{2\pi}} \left(f(x), \frac{1}{\sqrt{2\pi}} \right) \\
a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{1}{\pi} (f(x), \cos kx) \\
&= \frac{1}{\sqrt{\pi}} \left(f(x), \frac{1}{\sqrt{\pi}} \cos kx \right) \\
b_k &= \frac{1}{\sqrt{\pi}} \left(f(x), \frac{1}{\sqrt{\pi}} \sin kx \right)
\end{aligned} \tag{20}$$

很显然是成立的，那么我们可以将傅里叶级数换个方式表达，

$$\begin{aligned}
f(x) &= \frac{a_0}{2} + \sum_{i=1}^{\infty} (a_k \cos kx + b_k \sin kx) \\
&= \left(f(x), \frac{1}{\sqrt{2\pi}} \right) \frac{1}{\sqrt{2\pi}} + \sum_{i=1}^{\infty} \left[\left(f(x), \frac{1}{\sqrt{\pi}} \cos kx \right) \frac{1}{\sqrt{\pi}} \cos kx + \right. \\
&\left. \left(f(x), \frac{1}{\sqrt{\pi}} \sin kx \right) \frac{1}{\sqrt{\pi}} \sin kx \right] \\
&= \sum_{i=0}^{\infty} (f, e_n)\, e_{n-}
\end{aligned} \tag{21}$$

其中，$(\cdot)$ 表示的就是内积。这样在函数空间建立了一个正交坐标系。每一个函数和一组 (可数的) 数一一对应。

$$f(x) \sum_{i=1}^{\infty} (f, e_i) e_i \tag{22}$$

其中，系数是 $f(x)$ 和 e_i 内积，即 $f(x)$ 是 e_i 上的投影。对照，

$$x = \sum_{i=1}^{n} (x, e_i) e_i \;\; x \in \mathbb{R}^n \tag{23}$$

$$|x|^2 = \sum_{k=1}^{n} |(x, e_k)|^2 \tag{24}$$

二者之间的区别是什么? $\mathbb{R}^n$ 是有限维空间，而函数空间是无穷维的。无穷维求和是一个求极限的过程。无穷维肯定和有限维的情况不一样，不然我们就没有研究的意义了。

$$\begin{array}{c} f(x) = (?) \lim_{n\to\infty} \sum_{k=1}^{n} (f, e_k)\, e_k \\ \|f(x)\|^2 = (?) \lim_{n\to\infty} \sum_{k=1}^{n} |(f, e_k)|^2 \end{array} \tag{25}$$

我们需要考虑函数项组数 (Fourier 组数) 是否收放的问题。如果收敛，在什么意义下收做？这都是我们要研究的问题。

$$f(x) = (?) \lim_{n\to\infty} \sum_{k=1}^{n} (f, e_k)\, e_k \tag{26}$$

根据数学分析中的结论有，$f(x)$ 逐段可微，则其 Fourier 级数收敛，且收敛到，

$$\frac{f(x+0)+f(x-0)}{2}$$

我们要在无穷维空间研究收敛性 (在什么意义下收做)，即要引进极限等概念。这里的极限将不再是三维空间中的简单极限了，而是函数的极限，算子的极限那些。

在 Fourier 级数中，有 Riemann 引理，

$$\begin{array}{c} \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)\cos nx \mathrm{d}x = 0 \\ \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)\sin nx \mathrm{d}x = 0 \\ \text{即:}\ \lim_{k\to\infty} (f(x), e_k) = 0, \forall f(x) \end{array} \tag{27}$$

黎曼引理是傅里叶级数中最重要的一部分，没有黎曼引理的话，没法证明那个收敛性。以后我们将看到这是 e_k 弱收敛到 0。所以，在泛函中，我们以后要讨论什么，强收敛，弱收敛，一致收敛等等。**所以，数学学习中很重要的一点就是，我们不仅仅要考虑问题本身，而且要考虑问题是从哪来的**。

5 泛函研究的重点

通过上面的步步描述，我们最终得到了泛函分析的重点:

1. 空间的概念 (无穷维空间);
2. 空间的结构: 距离，长度，内积;
3. 空间中的收做性 (强，弱，一致收敛等)。

这些问题的引出都是非常自然的，我们从简单的情况一步步向困难的问题中进行扩展。自然而然的引出了这些问题，大家一定有清楚问题从哪来的，我们为什么要解决它。这样会让大家对数学有一个更清晰的认识。

Functional Analysis 1.1 Introduction 02

Chen Gong

August 2020

目录

1 Introduction

在上一小节中，我们讲了泛函分析的一些重点的问题。以及泛函分析问题的主要来源。在数学研究中，非常重要的是将一类问题中的共同点抽象出来形成定义，从而把复杂的问题简单化。在上一小节中，讲了二维空间中的坐标分解，四维空间中的一个矩阵的分解，还有对函数的傅里叶函数分解。在这一小节中，我们再把无穷维空间的线性算子 (微分运算) 与有限维空间的线性算子相对照，进而研究线性算子的分解问题。

而我们上一小节中介绍的矩阵分解，对一个矩阵 $A = (a_{ij})_{n\times n}$ 可以求得特征值，进而求得特征向量。特征向量可以产生一个坐标系，而在这个坐标系下 A 成为对角矩阵。而不同的对称矩阵可以产生不同的坐标系。

而在上节中，我们利用 Fourier 级数可以展开到一个正交坐标系下，

$$(1, \cos x, \sin x, \cdots, \cos kx, \sin kx, \cdots)$$

但是 Fourier 级数只是求得了一个函数的正交坐标系。那么我们先想想是否也可以想矩阵的运算一样，将一些运算 (算子，比如微分运算) 按矩阵分解那样一样进行分解? 将函数变换到一个坐标系下，从而使得函数被分解得更加的简单。下面首先将用一个 Sturm Liouville 问题来进行描述。

2 Sturm Liouville 问题

2.1 Sturm Liouville 问题通解计算

$$\begin{cases} -y''(t) = \lambda y(t), \quad -\pi < t \leqslant \pi \\ y(-\pi) = y(\pi) \\ y'(-\pi) = y'(\pi) \end{cases} \tag{1}$$

这是一个二阶的常微分方程，加上两个边界条件 (周期边界条件) 限制。**微分是一种运算，边界条件给出了它的定义域**。我们可以把 $\mathcal{T}y = -y''$（比如令 $y = e^x$，$\mathcal{T}y = e^x$）看成一种运算，而边界条件对运算的定义域加以适当的限制，使之成为一个“对称”算子。为什么是对称的呢，这里的这个含义不做过多的研究（后面会讲）。而这里的对称和前面的对称矩阵对应起来，这个含义是基本类似的。

而对应到线性代数中，$Ax = y$ 中 A 是一个变换矩阵，作用到 x 上得到一个 y。这里 $\mathcal{T}$ 是一个算子，作用在函数 y 上得到一个新的函数。只不过前者是在有限维空间上的运算，而后者是在无限维空间的运算。

在绪论中我们希望可以比较这些类似的例子，然后抽象出本质的东西。**在数学研究中对于概念本质的认识是非常重要的，而认识不清晰很有可能会在后面将概念搞糊涂**。

于是，这样在 Sturm Lioville(S- L) 问题上 $\mathcal{T}y = \lambda y$ 与 $Ax = \lambda y$ 形式上相似。当然 $\mathcal{T}y = \lambda y$ 中的 λ 不再是特征值了，而是特征元素或者特征函数。这样我们将一个很简单的东西，而无穷维空间中函数的概念对应了起来，可以帮助我们去认识新的概念。当然不是所有的函数都有特征元素，就和线性代数里一样，要想矩阵由特征值，必须保证 $x \neq 0$。有解的那些 λ，称为 S-L 问题的特征值。那么下一步则是对这个问题进行求解。求出 $-y''(t) = \lambda y(t), \quad -\pi < t \leqslant \pi$ 的通解，

1. 当 $\lambda > 0$ 时，有 $y(t) = A\cos\sqrt{\lambda}t + B\sin\sqrt{\lambda}t$。这里是将复数用欧拉公式展开的。然后代入边界条件 $y(-\pi) = y(\pi)$ 可以得到，

$$\begin{gathered} A\cos\sqrt{\lambda}\pi - B\sin\sqrt{\lambda}\pi = A\cos\sqrt{\lambda}\pi + B\sin\sqrt{\lambda}\pi \\ \Rightarrow 2B\sin\sqrt{\lambda}\pi = 0 \end{gathered}$$

   而两边代入边界条件 $y'(-\pi) = y'(\pi)$，可以得到，

$$\begin{gathered} -A\sqrt{\lambda}\sin\sqrt{\lambda}\pi + B\sqrt{\lambda}\cos\sqrt{\lambda}\pi \\ = A\sqrt{\lambda}\sin\sqrt{\lambda}\pi + B\sqrt{\lambda}\cos\sqrt{\lambda}\pi \\ \Rightarrow A\sqrt{\lambda}\sin\sqrt{\lambda}\pi = 0 \end{gathered}$$

   由于 $\lambda > 0$, 其 A, B 不能同时为 0, 所有我们可以得到, $\sin\sqrt{\lambda}\pi = 0$。所以, $\lambda = n^2(n = 1, 2, \cdots,)$ 时，可以满足两个边界条件。

2. 而当 $\lambda = 0$ 时，可得 $-y'' = 0$。在满足边界条件的情况下，可得 $y = 1$。

3. 而当 $\lambda < 0$，比较容易可以解得 $y = Ce^{\pm\sqrt{-\lambda}t}$。但是这个解并不能满足边界条件。

综上所述，我们可以求解出特征元素为，$\{\lambda_n\} = \{0, 1^2, 2^2, \cdots, n^2, \cdots\}$。将此结果代入到第一步计算出的通解中进行求解，并且把 $\mathcal{T}y = -y''$ 看成是一种运算，可以得到其对应的特征函数为，

$$\{e_n\} = \{1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots\}$$

这样我们就得到一个特征函数系。然而，这恰好就是 **Fourier 展开中的正交坐标系（乘以系数可使之单位化）**。我们在这似乎找到了结合点，下面将线性代数中有限维算子的分解，和无穷维函数算子的分解做一个对比。

2.2 有限维算子和无限维算子对比

在这一小节中，我们将 $\mathcal{T}y = -y''$ 看成一种运算（自共轭（对称）算子）。再一小节中，将 $\mathcal{T}y_n = \lambda_n y_n$ 与 $Ay_n = \lambda_n y_n$ 进行对比。在线性代数中，我们可以将运算分解为在一个正交坐标系中 $(\beta_1, \beta_2, \cdots, \beta_n)$ 投影的形式，而我们下一个问题就是在无限维的函数算子上，是不是也可以得到类似的结论。这就是我们泛函分析的一个基本问题，而实际上是可以做如下的分解的，

$$\begin{gathered} A = \lambda_1 P_1 + \cdots + \lambda_n P_n; \quad (\text{有限维}) \\ \mathcal{T} = (?)\lambda_1 P_1 + \cdots + \lambda_n P_n + \cdots; \quad (\text{无限维}) \end{gathered} \tag{2}$$

其中 P 是其在特征元素上的投影算子，而也可以被称为是无穷维空间上线性算子的一种分解（谱分解）。而接下来的问题就是，这个无限维的东西加在一起是什么东西，它已经不再是一个数了。那么通过上述的分解，我们将两个看似无关的东西，联系在了一起，由简单的东西推广到复杂的东西，在数学研究中是非常重要的。

并且，我们希望在新的函数空间下，函数 f 可以分解为，

$$f(x) = \sum_{i=1}^{\infty} (f, e_i)\, e_i \tag{3}$$

其中，e_i 是 $\mathcal{T}$ 关于 λ 的特征函数，$\mathcal{T}e_i = \lambda_i e_i$，比如 $\mathcal{T}\sin 2x = 4\sin 2x$。所以 S-L 算子 $\mathcal{T}$ 作用在 f 上，是否可以有:

$$
\begin{aligned}
\mathcal{T}f &= \mathcal{T}\sum_{i=1}^{\infty}(f, e_i)\, e_i = (?)\sum_{i=1}^{\infty}(f, e_i)\, \mathcal{T}e_i \\
&= \sum_{i=1}^{\infty}(f, e_i)\, \lambda_i e_i
\end{aligned}
\tag{4}
$$

红色标记问号的地方，有一个运算次序是否可以交换的问题。

在有限维空间，可以有不同的正交系 (它们可以由不同的对称矩阵产生)，在无穷维空间是否也可以有不同的正交系，它们可以由不同的算子产生？答案是肯定的。

3 勒让德（Legendre）多项式

3.1 勒让德多项式特征值和特征函数分析

Legendre 多项式，考虑 Legendre 方程，

$$
\begin{cases}
-\left(1-x^2\right) y'' + 2xy' = \lambda y, \quad (-1 < x < 1) \\
y(1) < \infty \\
y(-1) < \infty
\end{cases}
\tag{5}
$$

而方程可以化简为，

$$
-\frac{\mathrm{d}}{\mathrm{d}x}\left(\left(1-x^2\right) y'\right) = \lambda y \tag{6}
$$

它是对称的微分算子，可以求出特征值为 $\lambda_n = n(n+1)$。而特征函数为：$y_n = \frac{d^2}{dx^2}\left(x^2-1\right)^n$。

这里的求解比较的复杂，在这里不做过多的解释，之后的学习中大家自然会遇到的。这里的特征函数是一个多项式，比之前的前面的三角函数为特征函数，显然多项式要简单很多。并且，我们定义内积运算，

$$
(y_n, y_m) = \int_{-1}^{1} y_n(x) y_m(x)\, dx
$$

其中，我们将此函数定义为内积。很显然这满足内积的定义，非负性，正定性，数乘性质和乘法分配率。很显然还是可以满足那四条，可以称之为内积运算。很显然，我们定义的特征函数可以满足，

$$
\int_{-1}^{1} y_n(x) y_m(x) dx = 0 (m \neq n) \tag{7}
$$

它们是 $L_2[-1,1]$ 上的正交系 (对称线性算子的特征函数系)。其中，

$$
\begin{aligned}
&y(x) = 1, \quad y_3(x) = \tfrac{1}{2}\left(5x^3 - 3x\right) \\
&y_1(x) = x, \quad y_3(x) = \tfrac{1}{8}\left(35x^4 - 30x + 3\right) \\
&y_2(x) = \tfrac{1}{2}\left(3x^2 - 1\right), \quad y_5(x) = \tfrac{1}{8}\left(63x^3 - 70x^3 + 15x\right)
\end{aligned}
\tag{8}
$$

那么他们可以称之为正交特征函数，可以类比思考为一个正交系。**这组特征多项式和 Fourier 级数展开得到的特征向量是同一种性质的东西，他们都是正交系**。而勒让德多项式特征函数图像，如下所示，

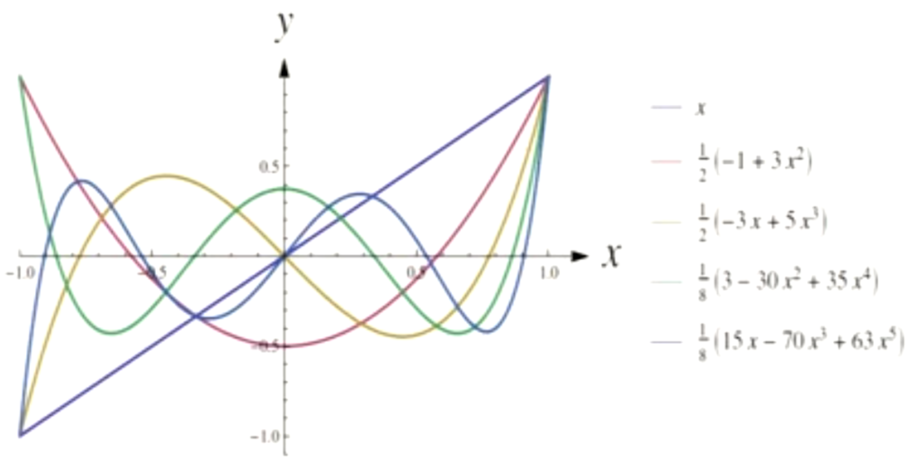

图 1: 勒让德多项式特征函数图像

通过上述几个例子，我们似乎找到了无穷维分解和有穷维分解直接的联系。**在数学研究中，我们经常这样来思考问题。通过特例来发现一些规律或者猜想，大胆的假设猜想是成立的，然后给出普适性的证明。**

3.2 Fourier 特征函数与特征多项式对比

不知道大家有没有注意到 Fourier 特征函数与特征多项式有没有什么相似的地方。大家发现，好像这个 n 值越大，他们和 x 轴的交点都是越来越多，这个零点越来越多也就意味着震动的频率越来越大。而且随着 n 的增大他们的函数图像很像。那么这又可以给我们怎样的启发呢?

不论我们是用什么样的方式去分解一个微分算子，得到其特征值和特征函数，而特征函数都有非常好的震动性。这都是我们观察到的，那他们到底对不对呢?

4 无穷维空间线性算子的性质上的区别

我们线性泛函分析主要研究两个问题，一个是空间，其中包括函数空间，算子的空间；还有一个就是线性算子，包括微分，积分这些东西。

微分和积分是高等数学研究的主要对象。

共同点：它们都是线性运算:

不同点：粗略地说微分可能把函数“放大”，积分可能把函数“变小”。

例如函数，$y = x^n, x \in [0,1]$，则有，

$$\begin{array}{l} \frac{d}{dx}\left(x^n\right) = nx^{n-1} \\ \int x^n dx = \frac{1}{n+1}x^{n+1} \end{array} \tag{9}$$

当然这只是我们直观的一种感觉形式。下面这个例子我们将可以看得更加清晰。

4.1 函数级数项分析

设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 满足：

1. $u_n(x)(n=1,2,\cdots)$ 在函数 $[a,b]$ 上连续。
2. $\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上一致收敛到 $S(x)$；

则 $S(x)=\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上可积，且满足

$$\int_0^b S(x)\cdot dx=\int_a^b \sum_{a=1}^{\infty} u_m(x)dx=\sum_{n=1}^{\infty}\int_0^4 u_a(x)dx \tag{10}$$

即在一致收敛的条件下，积分运算可以与无穷级数运算交换顺序。这个证明在数学分析中已经有了，这里不再做过多的说明。

1. $u_n(x)(n=1,2,\cdots)$ 在函数 $[a,b]$ 上连续可导；
2. $\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上点点收敛到 $S(x)$；
3. $\sum_{n=1}^{\infty} u_n'(x)$ 在 $[a,b]$ 上一致收敛到 $\sigma(x)$；

那么，在这三个条件下，我们可以得出 $S(x)=\sum_{n=1}^{\infty} u_n(x)$ 在 $[a,b]$ 上是可导的。并且，

$$\frac{d}{dx}S(x)=\frac{d}{dx}\sum_{n=1}^{\infty} u_n(x)=\sum_{n=1}^{\infty}\frac{d}{dx}u_n(x) \tag{11}$$

即求导运算可以与无穷级数运算交换顺序，但是微分与级数运算交换顺序比积分与级数运算交换顺序多了一个条件 (3)。

以后我们看到积分算子是有界 (连续) 线性算子；微分算子是无界线性算子，但是它是闭算子。所以他们两者都可以和级数运算交换顺序。而有界算子可以简单的理解为收敛的，无界算子则是非收敛的。所以，有界线性算子与无界线性算子运算性质有很大差别。这也是我们泛函分析的重点。

5 总结

上述数学分析、线性代数、微分方程的一些例子中，在处理问题上有许多相似的方法，包括：问题和元素更一般化 (抽象化)，空间中的元素 (向量) 可以是函数或运算 (矩阵运算、微分运算、积分运算，级数 (极限) 运算。而且我们都建立一种空间的框架，把元素 (可以是函数或运算) 进行坐标分解。我们希望通过类比等方法把它们推广到 (结果可能会有差异) 泛函分析的研究中去，应用几何、代数和分析的综合手段研究解决问题，研究无限维线性空间上的泛函和算子理论。这就是我们研究的主要问题，也是我们下面要研究的泛函分析的主要内容，。

在绪论中，我们希望从问题本身出发，发现问题，解决问题。而不是直接引入公理系统，这样有助于大家理解问题。

而本系列课程的主要内容如下所示：

1. 首先引入**空间，极限**这些概念，讨论它们的性质。包括第一到第三章：距离空间；线性赋范空间；内积空间。这是怎样一个逻辑呢？我们引入内积的概率以后就可以定义长度和距离，自然就有了距离空间和线性赋范空间。

2. 研究线性算子 (线性算子空间) 的性质。包括第四到第五章：有界线性算子，有界线性算子的重要性质；共轭空间；特别是 Hilbert 空间的共轭空间和共轭算子。**重要定理包括：一致有界原则；开映射定理、逆算子定理；闭图像定理；线性泛函的延拓定理 (Hahn Banach 定理)。这是最核心也是最难的部分**。

3. 最后两章是线性算子的谱理论，什么是谱理论呢？实际上就是特征值问题。**谱分解从结构上展示了线性算子的基本运特征，特别是紧的自共轭算子的谱分解和有限维空间对称矩阵的分解十分相似（这里的共轭实际就和对称是非常相似的，紧和有限维空间是非常相似的）**。

而本系列课程的学习目标如下所示：

1. 理解为什么会有泛面分析，明白泛函分析在做什么；
   (a) 最基本的概念 (概念的来源和背景)；
   (b) 数学研究的基本方法: 化归、类比、归纳、联想。
   (c) 一定的抽象思维的能力，**概念清楚，思维清晰**。
2. 感悟数学的美。

老师这里有一句话非常的重要，“泛函分析学习的过程中，一定要牢记类比简单的二维或三维情况，它们都有着类似的地方。如果将这些概念脱离问题背景本身去看是非常复杂，生涩的。而结合其问题背景就会变得很好理解了。”

绪论结束语:

1. 要把数学看成客观世界的简单化。

2. 要从整体上了解数学, 才能培养良好的数学素质。数学知识重要的是，可以用数学抽象出来的观点来看待客观世界，解决实际问题。

3. 要从具体的实例中感悟数学的思想方法.

“数学决不应成为一门十分费解的科学”。

我们借用张恭庆院士等在他的泛函分析序言中的一段话来结。

我们认为要真正理解泛函分析中的一些重要的概念和理论，灵活运用这一强有力的工具，**其唯一的途径就是深入了解它们的来源和背景**，注意研究一些重要的、一般性定理的深刻的、具体的含义。不然的话，如果只是从概念到概念，纯形式地理解抽象定理证明的推演，那么学习泛函分析的结果只能是“如入宝山而空返”，一无所获。

Functional Analysis 1.2 Distance Space 02

Chen Gong

22 August 2020

目录

1 Introduction

在上一小节中我们在集合上引进了距离的概念,有了距离的概念之后随后就定义了极限,极限的定义为: $\lim_{n\to\infty} x_n = x_0 \leftrightarrow d(x_n, x_0) = 0$,那么距离是从两个空间到实数空间上的一个映射: $\mathcal{X}\times\mathcal{X} \to \mathbb{R}$,距离一定是一个数。在空间中定义了距离后，我们就可以在距离空间中引入极限的概率了。这就是我们的重要目的之一。而且，我们还介绍了距离空间点的收敛性，并给出了相应的证明。

2 距离空间函数的连续性

在这一节我们需要证明的是，距离空间可以看成是一个函数，而且是一个二元连续函数。

定理 2.1： $d(x, y)$ 是关于 x 和 y 的二元连续函数。即当 $x_n \to x, y_n \to y(n \to \infty)$ 时，

$$d(x_n, y_n) \to d(x, y)(n \to \infty)$$

如何给出这个证明呢？大家好好想一想，在距离空间 (X, d) 中的任何两点 x, y 都有唯一确定的实数 $d(x, y)$ 与之相对应，这说明 $d(x, y)$ 是一个二元实函数。那么大家想想在之前，我们如何确定一个函数是否连续呢？证明方法是 $f(x, y) \to f(x_0, y_0)(x \to x_0, y \to y_0)$。

而在实数空间中，距离是通常的绝对值距离，定理要证：在条件 $x_n \to x, y_n \to y(n \to \infty)$ 下，有

$$|d(x_n, y_n) - d(x, y)| \to 0 \ (n \to \infty) \tag{1}$$

证明： 由距离的三角不等式有:

$$d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n)$$

即

$$d(x_n, y_n) - d(x, y) \leq d(x_n, x) + d(y, y_n)$$

同理有,

$$d(x, y) \leq d(x, x_n) + d(x_n, y_n) + d(y_n, y)$$

即有,

$$d(x, y) - d(x_n, y_n) \leq d(x, x_n) + d(y_n, y)$$

所以可以得到,

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y, y_n) \to 0 \ (n \to \infty) \tag{2}$$

介绍其是一个函数，便于我们接下来介绍距离空间中收敛的“含义”。在上一小节中，我们介绍了一个集合中的元素都一样，然而定义不同的距离，得到的含义是不一样的。下面我们将在一些距离空间中，研究收敛的“具体含义”。

3 距离空间的具体含义

下面我们将从几个例子中逐步增加难度来分析距离空间的意义。

3.1 $\mathbb{R}^m$ 空间

$\mathbb{R}^m$ 空间，设

$$\begin{gathered}\boldsymbol{x}_n=\left(\boldsymbol{\xi}_1^{(n)},\boldsymbol{\xi}_2^{(n)},\cdots,\boldsymbol{\xi}_m^{(n)}\right)\quad(\boldsymbol{n}=1,2,\cdots),\\ \boldsymbol{x}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_m)\in\mathbb{R}^m\end{gathered}\tag{3}$$

则 $d(x_n,x)\to 0$，等价于，

$$\xi_i^{(n)}\to\xi_i(n\to\infty),\ i=1,2,\cdots,m.\tag{4}$$

在 $\mathbb{R}^m$ 空间中，等价于按坐标收敛。我们可以这样理解，将 $\boldsymbol{x}_n$ 看成是一个向量，那么向量中的每一个维度上的值都将收敛到与 x 上。大概就是这么个意思。

证明： $d(x_n,x)\to 0$，即

$$\sqrt{\left(\xi_1^{(n)}-\xi_1\right)^2+\cdots+\left(\varepsilon_m^{(n)}-\xi_m\right)^2}\to 0.(n\to\infty)\tag{5}$$

首先我们来看第一个不等式，有

$$\begin{aligned}\left|\varepsilon_i^{(n)}-\xi_i\right|\le&\left(\sum_{k=1}^m\left|\xi_k^{(n)}-\xi_k\right|^2\right)^{1/2}\\ &=d\left(x_n,x\right),i=1,2,\cdots,m\end{aligned}\tag{6}$$

其中,i 是 $i=1,2,\cdot m$ 中的任意一个数。这个等式是显然成立的,$\left|\varepsilon_i^{(n)}-\xi_i\right|$ 只是 $\left(\sum_{k=1}^m\left|\xi_k^{(n)}-\xi_k\right|^2\right)^{1/2}$ 中的一项。从第一个式子中，我们可以得到在按坐标收敛中，如果距离收敛，则每一个坐标都将收敛。而同样的思路我们可以得到第二个不等式，要证明每一个坐标都收敛，则可以得到距离收敛。这就是我们证明的整体思路。

$$\begin{aligned}d\left(x_n,x\right)=&\left(\sum_{k=1}^m\left|\xi_k^{(n)}-\xi_k\right|^2\right)^{1/2}\\ \le&\sqrt{\left(\left|\varepsilon_1^{(n)}-\xi_1\right|+\left|\varepsilon_2^{(n)}-\xi_2\right|+\cdots+\left|\varepsilon_m^{(n)}-\xi_m\right|\right)^2}\\ \le&\left|\varepsilon_1^{(n)}-\xi_1\right|+\left|\varepsilon_2^{(n)}-\xi_2\right|+\cdots+\left|\varepsilon_m^{(n)}-\xi_m\right|\end{aligned}\tag{7}$$

而从第二个公式中可以得出，如果每个分量都趋近于 0，而他们的和当然也是趋近于 0 的。那么我们可以得出 $d(x_n,x)\to 0$，等价于 $\xi_i^{(n)}\to\xi_i(n\to\infty)$，$i=1,2,\cdots,m.$。即可得到结论 (空间中点列的收敛，等价于按坐标收敛)。从这个简单的例子中，我们大致导出了距离空间中的收敛的它的含义是什么。下一步将讲述更复杂的例子。

3.2 连续函数空间 $C[a,b]$

3.2.1 在 $C[a,b]$ 中考虑

在连续函数空间 $C[a,b]$ 中的元素都是连续函数。空间中的距离定义为：$d(x,y)=\max_{a\le t\le b}|x(t)-y(t)|$。而且，$C[a,b]$ 中的收敛性是函数列在 $[a,b]$ 上的一致收敛。一致收敛是我们数学分析中一个非常重要的概念。设 $x_n(t)$ $(n=1,2,\cdots)$，$x(t)\in C[a,b]$，且 $d(x_n,x)\to 0$，即为

$$\max_{a\le t\le b}|x_n(t)-x(t)|\to 0(n\to\infty)\tag{8}$$

那么这个公式 (8) 趋近 0 是什么意思呢？实际上就是距离趋近 0。注意，这里的 $x_n(t)$ 代表的是一个数列，关于 n 取不同的值，$x_n(t)$ 会得到不同的结果。所以，这里实际上是一个数列的极限。于是，我们可以用 $\epsilon - N$ 的语言来描述数列的收敛，

于是对 $\forall\epsilon \le 0, \exists N,$ 当$n \le N$时，对$\forall t \in [a, b]$，有

$$|x_n(t) - x(t)| < \max_{a \le t \le b} |x_n(t) - x(t)| < \epsilon \tag{9}$$

即：$x_n(t)$ 一致收敛到 $x(t)$。请注意这里为什么是一致收敛？我们从逻辑上要非常关注 $n \le N$ 和 $\forall t$ 的顺序。在这里，我们是先定义的 N，然后才是 $\forall t$，所以 N 是和 t 没有关系的，而仅仅只和 ϵ 有关。也就对于 $[a, b]$ 上任意一个点，都可以用同一个 ϵ 来约束。这样我们从直观的角度去理解，我们可以保证所有的 $|x_n(t) - x(t)|$ 的值不能无限大，那么这样就很好的控制了收敛的速度，说明在这个区间上函数收敛的速度是一样的。

这就是我对一致收敛的理解，而收敛中的定义为，对 $\forall\epsilon \le 0, \exists N,$ 对$\forall t \in [a, b]$有$n \le N$，

$$|x_n(t) - x(t)| < \max_{a \le t \le b} |x_n(t) - x(t)| < \epsilon \tag{10}$$

那么这里的 N 和 ϵ 和 t 都是相关的。那么，在不同的点，ϵ 实际上是不一样的，所以没有那个一致的概念在里面。一致收敛是指在收敛中和区间上的点没有关系。

因为要证两个关系的等价性，所以又需要反过来证明一遍。$x_n(t)$ 一致收敛到 $x(t)$ 可以推出 $d(x_n, x) \to 0$；事实上，$x_n(t)$ 一致收敛到 $x(t)$，可以得出，

对 $\forall\epsilon \le 0, \exists N,$ 当$n \le N$时，对$\forall t \in [a, b]$，有

$$|x_n(t) - x(t)| < \epsilon \tag{11}$$

上述公式两边同时对 $t \in [a, b]$ 取最大值，则有

$$\max_{a \le t \le b} |x_n(t) - x(t)| \le \max_{a \le t \le b} \epsilon = \epsilon \tag{12}$$

其中，$max_{a \le t \le b}|x_n(t) - x(t)|$ 就是关于距离的定义，那么可以说明 $x_n \to x(n \to \infty)$。

通过以上的证明，我们可以充分的说明，$C[a, b]$ 中的收敛是函数列在 $[a, b]$ 上的一致收敛。

当然我们还可以定义其他的距离，比如

$$d_2(x, y) = \left\{ \int_0^1 |x(t) - y(t)|^2 \mathrm{d}t \right\}^{1/2} \tag{13}$$

这个距离是非常重要的，在这个距离空间下我们可以定义内积，角度，长度等概念。但是这个是距离我们还没有给出证明，这个会在第二章中给出证明。在之前我们分析过了，集合中的元素还是一样的，都是连续函数，而定义的距离不一样，则反映的客观现实是不一样的。

比如，考虑 (X, d_2) 中的点列 $\{x_n\}$，

$$x_n(t) = \begin{cases} 1 - nt, & 0 \le t \le 1/n \\ 0, & 1/n < t \le 1 \end{cases} \tag{14}$$

则 $\{x_n\}$ 收敛到 $x_0=0$，这个函数列实际上是一根直线。事实上，

$$\begin{aligned} d_2\left(x_n, x_0\right) &= \left\{\int_0^1 \left|x_n(t)-x_0(t)\right|^2 \mathrm{d}t\right\}^{1/2} \\ &= \left\{\int_0^{1/n} (1-nt)^2 \mathrm{d}t\right\}^{1/2} = (3n)^{-1/2} \rightarrow 0 \end{aligned} \tag{15}$$

这个东西是会趋近 0 的。和之前公式 (8) 中直接取最大值的做法，公式 (15) 中的距离明显要更平均一些。

收敛性和空间中的距离有很大的关系，就算空间中的元素都一样，定义的距离不一样，那么它的意义会有很大的区别，而且收敛性也有很大的区别。

注 1，上述 $\{x_n\}$ 在下列距离下也可以收敛到 x_0，

$$d(x,y)=\int_0^1 |x(t)-y(t)|dt \tag{16}$$

这个的证明同样也非常的简单，和公式 (15) 中的证明方法是一样的。而在公式 (8) 中给的距离不是一致收敛到 x_0，甚至 $\{x_n\}$ 并不点点收敛（所谓点点收敛就是函数上 $\forall t, x_n(t)\rightarrow x(t)$）到 x_0。因为 $x_n(0)=1$ 是恒成立的。大家把这个函数列画一下就知道，他们之间的距离肯定是等于 1 的。

这说明定义在不同的距离空间下的点列 (函数列) 的收敛与 $C[a,b]$ 中点列的收敛在 “具体意义” 下有很大的不同。

3.2.2 在 $C[0,1]$ 中考虑

如果在 $C[0,1]$ 中我们重新考虑上面的例子。再使用距离为公式 (8) 的情况下，由于对 $\forall n$，都有 $d(x_n,x_0)=1$，于是 $\{x_n\}$ 不收敛于 x_0。但是初学者可能会认为 $\{x_n\}$ 趋近于 y_0，其中，

$$y_0(t)=\begin{cases} 1, & t=0 \\ 0, 0<t<1 \end{cases} \tag{17}$$

但是，这有什么问题呢？问题就是这个函数不是连续函数。都不属于这个空间了，这肯定是不行的。

3.3 空间 s

设 $s=\{\{\xi_n\}\}$，是全体实数列（实数组成的数列）组成的集合，定义，

$$d(x,y)=\sum_{k=1}^{\infty}\frac{1}{2^k}\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|} \tag{18}$$

其中，$x=\{\xi_k\}$，$y=\{\eta_k\}$，则

(1) s 为距离空间。

(1) s 中的收敛按坐标收敛。

下面我们将给出相关的证明。首先可以得知，$d(x,y)$ 中的每一项都比必然是小于 $\frac{1}{2^k}$，所以加起来必然是小于 1 的。那么这样可以保证 $d(x,y)$ 一定是一个数，通常我们在确定距离空间，第一步就是判断其是否是一个数。

而按坐标收敛的意思是，设 $x_n=\left(\xi_1^{(n)},\xi_2^{(n)},\cdots,\xi_k^{(n)},\cdots\right)\in s, x=(\xi_1,\xi_2,\cdots,\xi_k,\cdots)\in s$。则 $d\left(x_n,x\right)\rightarrow 0(n\rightarrow\infty)\Longleftrightarrow \forall k, \xi_k^{(n)}\rightarrow\xi_k(n\rightarrow\infty)$。大致意思就是 $x_n\rightarrow x$ 就需要 x 中的每个维度上的坐标都要趋近于 x_n。

3.3.1 性质 1 的证明

下面我们给出性质 1 的证明，s 为距离空间。

要证 s 为距离空间，只要证明在 s 中所定义的距离 d 满足距离定义的 4 条即可。其中 (1)，(2)，(3）显然成立，只要验证 (4) 三角不等式成立，即

$$d(x,y) \leq d(x,z) + d(z,y) \tag{19}$$

设 $x=\{\xi_k\}$，$y=\{\eta_k\}$，$z=\{\zeta_k\}$，我们要证明的目标是，

$$\sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|\xi_k - \eta_k|}{1+|\xi_k - \eta_k|} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|\xi_k - \zeta_k|}{1+|\xi_k - \zeta_k|} + \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|\zeta_k - \eta_k|}{1+|\zeta_k - \eta_k|} \tag{20}$$

而对此公式进一步化简可得我们的证明目标为。

$$\frac{|\xi_k - \eta_k|}{1+|\xi_k - \eta_k|} \leq \frac{|\xi_k - \zeta_k|}{1+|\xi_k - \zeta_k|} + \frac{|\zeta_k - \eta_k|}{1+|\zeta_k - \eta_k|} \tag{21}$$

注意到 $f(x) = \frac{x}{1+x}$ 上下同加一个正数，这个值会越来越大，我们高中数学老师用了一句话来介绍为：“糖水加糖更甜。” 并且根据三角不等式可得，$|\xi_k - \eta_k| \leq |\xi_k - \zeta_k| + |\zeta_k - \eta_k|$。

$$\begin{aligned}
\frac{|\xi_k - \eta_k|}{1+|\xi_k - \eta_k|} &\leq \frac{|\xi_k - \zeta_k| + |\zeta_k - \eta_k|}{1+|\xi_k - \zeta_k| + |\zeta_k - \eta_k|} \\
&= \frac{|\xi_k - \zeta_k|}{1+|\xi_k - \zeta_k| + |\zeta_k - \eta_k|} + \frac{|\zeta_k - \eta_k|}{1+|\xi_k - \zeta_k| + |\zeta_k - \eta_k|} \\
&\leq \frac{|\xi_k - \zeta_k|}{1+|\xi_k - \zeta_k|} + \frac{|\zeta_k - \eta_k|}{1++|\zeta_k - \eta_k|}
\end{aligned} \tag{22}$$

所以，可以证明 $d(x,y) \leq d(x,z) + d(z,y)$，我们把这个距离空间记为 s。

3.3.2 性质 2 的证明

要证明的是 $d(x_n,x) \to 0 \Longleftrightarrow \forall k, \epsilon_k^{(n)} \to \xi_k$。这个问题我们要从必要性和充分性两个角度来证明。首先是给出必要性证明。必要性要证，对于任意给定的 $k_0 \in N$，要做到：

$$\forall \epsilon > 0, \exists N, \forall n > N, 有 |\xi_{k_0}^{(n)} - \xi_{k_0}| < \epsilon$$

对于任意给定的 k_0，对于 $\forall \epsilon > 0$，令 $\epsilon_0 = \frac{1}{2^{k_0}} \frac{\epsilon}{1+\epsilon} > 0$，由于 $d(x_n,x) \to 0\ (n \to \infty)$，对于这个 $\epsilon_0 > 0$，$\exists N$，当 $n > N$ 时，有 $d(x_n,x) < \epsilon$，即为

$$d(x_n,x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\left|\xi_k^{(n)} - \xi_k\right|}{1+\left|\xi_k^{(n)} - \xi_k\right|} < \epsilon_0 = \frac{1}{2^{k_n}} \frac{\epsilon}{1+\epsilon} \tag{23}$$

由于每一项都是正数，于是我们有，

$$\frac{1}{2^{k_n}} \frac{\left|\xi_k^{(n)} - \xi_k\right|}{1+\left|\xi_k^{(n)} - \xi_k\right|} < \epsilon_0 = \frac{1}{2^{k_n}} \frac{\epsilon}{1+\epsilon} \tag{24}$$

从而可以推出，

$$\frac{\left|\xi_k^{(n)} - \xi_k\right|}{1+\left|\xi_k^{(n)} - \xi_k\right|} < \frac{\epsilon}{1+\epsilon} \tag{25}$$

结合 $f(x)=\frac{x}{1+x}$ 是单调递增函数可得，

$$\left|\xi_k^{(n)}-\xi_k\right|<\epsilon$$

即为

$$\xi_k^{(n)}\rightarrow\xi_k\ (n\rightarrow\infty) \tag{26}$$

在上面我们已经证明了必要性，下一步则是证明充分性，充分性我们要证，

$$\forall k,\xi_k^{(n)}\rightarrow\xi_k(n\rightarrow\infty)\Longrightarrow d\left(x_n,x\right)\rightarrow 0\ (n\rightarrow\infty) \tag{27}$$

即要证明 $d(x_n,x)=\sum_{k=1}^{\infty}\frac{1}{2^k}\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|}\rightarrow 0\ (n\rightarrow\infty)$

注意到收敛的级数，充分靠后面的无穷多项的和可任意小。对于前面的有限项，由条件可以找到共同的 N，当 $n>N$ 时级数中的这些项都一致很小。也就是无穷级数求和的问题非常的麻烦，由于每一项都是无穷小的，而无穷项加起来并不一定是无穷小的。所以，我们的思路是希望得到后面无穷项加起来是无穷小的，然后证明前面有限项是无穷小的，那么就可以将无穷级数求和的问题转换为有限项求和的问题，难度一定会减小很多。

证明：

对于 $\forall\epsilon>0,\exists K$，使得，$\sum_{k=K+1}^{\infty}\frac{1}{2^k}<\frac{1}{2}\epsilon$。这个必然是可以成立的，高中学过等差数列求和的同学都知道。

由于 $\xi_k^{(n)}\rightarrow\xi_k(n\rightarrow\infty)\ (k=1,2,\cdots,K)$。所以存在 N，当 $n>N$ 时，

$$|\xi_k^{(n)}\rightarrow\xi_k|<\frac{\epsilon}{2}\ (k=1,2,\cdots,K) \tag{28}$$

那么这样我们就把一个无穷相加的问题，变成了一个有限的问题，把后面一部分都给甩掉了，问题证明起来就会变得比较简单了。

于是当 $n\geq N$ 时，

$$\begin{aligned}
d\left(x_n,x\right)&=\sum_{k=1}^{\infty}\frac{1}{2^k}\frac{\left|\xi_k^{(n)}-\xi_k\right|}{1+\left|\xi_k^{(n)}-\xi_k\right|}\\
&=\sum_{k=1}^{K}\frac{1}{2^k}\frac{\left|\xi_k^{(n)}-\xi_k\right|}{1+\left|\xi_k^{(n)}-\xi_k\right|}+\sum_{k=K+1}^{\infty}\frac{1}{2^k}\frac{\left|\xi_k^{(n)}-\xi_k\right|}{1+\left|\xi_k^{(n)}-\xi_k\right|}\\
&<\frac{1}{2}\varepsilon\sum_{k=1}^{K}\frac{1}{2^k}+\sum_{k=K+1}^{\infty}\frac{1}{2^k}<\varepsilon
\end{aligned} \tag{29}$$

所以，$\{x_n\}$ 在 s 中可以收敛到 x。在这里老师介绍了一个证明中的通用性思路，由级数的收敛性，由于后面项的和“很小”，所以可以把后面的项加起来，证明其是收敛的。再对前面的有限项进行估计，这是很常用的办法。

3.4 空间 S

S 中的元素为 E 上全体几乎处处有限的可测函数。其中 $E\subset R$ 是一个 Lebesgue 可测集，且 $mE<\infty$。这个 Lebesgue 可测集这一堆复杂的概念是什么意思呢？这里并没有必要做过多的理解，很

多时候数学中掺杂这些很生涩的词汇，会让很多同学直接望而止步。这里理解成一个连续函数就行了，可测的目的是为了可积。

对于 $x=x(t), y=y(t) \in S$，定义

$$d(x,y)=\int_E \frac{|x(t)-y(t)|}{1+|x(t)-y(t)|}\mathrm{d}t \tag{30}$$

这里的定义是不是和我们公式 (18) 直接的定义很类似。一般来说级数和积分的区别就在于，级数是可数的，而积分是不可数的。

则其中，

(1) S 为距离空间；（这个证明和 3.3.1 中的证明几乎一模一样，这里不做过多的说明，大家自己去争，证明一下就行。）

(2) S 中的收敛是依测度收敛。即为 $d(x_n,x)\to 0\ (n\to\infty)$ 等价于 $x_n \underset{m}{\Rightarrow} x(n\to\infty)$(依测度收敛，其中 m 是代表依测度收敛的意思，metric 的首字母)。

首先需要了解什么出测度，非正式的说，测度把每个集合映射到非负实数来规定这个集合的大小。对于一个集合 E 就有一个数与之对应，这个数就是集合的测度，记为 mE。这个集合必须要有良好的性质，不能是随便什么集合都可以定义测度，所以就有了可测集的概念。而可测集具有可加性这样良好的性质，因为我们需要对集合进行测量，可加性是一个很重要的性质，$m\left(\bigcup_{n=1}^{\infty}=\sum_{n=1}^{\infty} mE_n\right)$。概率也是一种测量，并且所有集合的测度和等于 1。

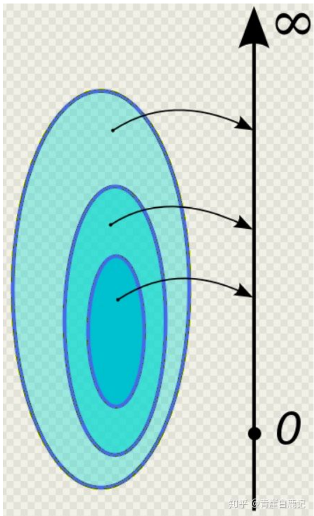

图 1: 依测度收敛的具体意义

下面我们详细的定义一下依测度收敛的定义，

$x_n \overset{m}{\to} x(n\to\infty) \Longleftrightarrow$ 对于 $\forall\sigma>0$，有 $m\{t\in E||x_n(t)-x(t)|\geq\sigma\}\to 0(n\to\infty)$

即对于 $\forall\sigma>0$，$\forall\epsilon>0$，存在 N，当 $n>N$ 时，

$$m\{t\in E||x_n(t)-x(t)|\geq\sigma\}<\varepsilon$$

证明方法是，我们可以将集合 E 分成两个部分：对于任意给定的 $\sigma>0$ 和自然数 n，令

$$\begin{gathered} E_1=\{t\in E||x_n(t)-x(t)|<\sigma\} \\ E_2=\{t\in E||x_n(t)-x(t)|\geq\sigma\} \\ E=E_1\cup E_2 \end{gathered} \tag{31}$$

根据测度的可测性，可得 $mE = mE_1 + mE_2$，并且 $mE > mE_1$，$mE > mE_2$。利用函数 $\frac{t}{1+t}$ 的单调性来证明，

$$m\left\{t \in E \| x_n(t) - x(t) \mid \geq \sigma\right\} \rightarrow 0\ (n \rightarrow \infty) \tag{32}$$

实际上这个证明并不是很难，而其关键是理解什么是依测度收敛。对于任意给定的 $\sigma > 0$，由

$$\begin{aligned} d(x_n, x) &= \int_E \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t \\ &\geq \int_{\{t \in E \| x_n(t) - x(t) | \geq \sigma\}} \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t \end{aligned} \tag{33}$$

因为 $\frac{t}{1+t}$ 单增，而 $|x_n(t) - x(t)| \geq \sigma$，于是有，

$$\begin{aligned} & d\left(x_n, x\right) \geq \textstyle\int_{\{t \in E \| x_n(t) - x(t) | \geq \sigma\}} \frac{\sigma}{1+\sigma} dt \\ & = \tfrac{\sigma}{1+\sigma} m\left\{t \in E \| x_n(t) - x(t) \mid \geq \sigma\right\} \end{aligned} \tag{34}$$

由于 $d(x_n, x) \rightarrow 0\ (n \rightarrow \infty)$，$\sigma$ 给定，由不等式

$$d\left(x_n, x\right) \geq \frac{\sigma}{1+\sigma} m\left\{t \in E \| x_n(t) - x(t) \mid \geq \sigma\right\} \tag{35}$$

推出

$$m\left\{t \in E \| x_n(t) - x(t) \mid \geq \sigma\right\} \rightarrow 0\ (n \rightarrow \infty) \tag{36}$$

反过来呢，我们要从 $x_n \underset{m}{\Rightarrow} x(n \rightarrow \infty)$ 推出 $d(x_n, x) \rightarrow 0\ (n \rightarrow \infty)$。事实上，

$$\begin{aligned} d(x_n, x) &= \int_E \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t \\ &= \int_{E_1} \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t + \int_{E_2} \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t \\ &\leq \int_{E_1} \frac{\sigma}{1+\sigma} dt + \int_{E_2} \frac{|x_n(t) - x(t)|}{1 + |x_n(t) - x(t)|} \mathrm{d}t \\ &= \frac{\sigma}{1+\sigma} mE_1 + mE_2 \\ &\leq \frac{\sigma}{1+\sigma} mE + mE_2 \end{aligned} \tag{37}$$

对于 $\forall \epsilon > 0$，取 $\sigma_0 mE < \frac{\epsilon}{2}$，对于此 $\sigma_0 > 0$，上式成立。存在 N，当 $n > N$ 时，$mE_2 < \frac{\epsilon}{2}$。这里用到了依测度收敛。

于是

$$\begin{aligned} d\left(x_n, x\right) &\leq \frac{\sigma_0}{1+\sigma_0} mE + mE_2 \\ &< \sigma_0 mE + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned} \tag{38}$$

即为 $d(x_n, x) \rightarrow 0\ (n \rightarrow \infty)$。

4 小结

本小节主要通过几个例子介绍了，距离空间中收敛的含义，简单的描述了点点收敛，一致收敛和依测度收敛的意思。1. 说明了收敛性和空间中的距离有很大的关系，就算空间中的元素都一样，定义

的距离不一样，那么它的意义会有很大的区别，而且收敛性也有很大的区别。2. 在有限维实数空间中，证明了距离空间等于 0 等价于点点收敛；在连续函数空间中，证明了证明了距离空间等于 0 等价于一致收敛 (收敛的速度一样)；在全体实数列空间中，证明了距离空间等于 0 等价于点点收敛；在可测函数空间中，证明了距离空间等于 0 等价于依测度收敛。在证明等价性的过程中，需要分别证明其充分性和必要性。

通过本小节的介绍，我们充分的感受到了距离空间的定义和极限之间的关系，从而通过极限定义连续，连续的定义基本离不开距离空间，距离是泛函分析中一个非常重要的概念。小编本身不是数学系出身，在深刻理解点点收敛，一致收敛，依坐标收敛和依测度收敛方面还有所欠缺，之后会好好加强。

Functional Analysis 1.3 Open Set Continuous Mapping

Chen Gong

26 August 2020

目录

1 Introduction

在上一小节中，我们讲述的就是距离空间，我们引入了距离的概念。之后讲述了距离空间上的一些性质，下一节我们要讲的是开集和开映射。

在 $\mathbb{R}^2$,$\mathbb{R}^3$ 空间中引入了距离（欧式距离），就有了

开球

$$B(x_0, r) = \{x \in X | d(x, x_0) < r\}, \tag{1}$$

闭球

$$B(x_0, r) = \{x \in X | d(x, x_0) \leq r\}, \tag{2}$$

于是，就有了开集、闭集这些概念。注意，开球和闭球这两个概率仅仅用到了距离的概念。所以，在一般的距离空间中，我们也同样可以引进开球、闭球，进而引进开集、闭集等一系列概念。

2 开球和闭球

集合 2.1 设 (X, d) 是一个距离空间，$r > 0$，集合

$$B(x_0, r) = \{x \in X | d(x, x_0) < r\}, \tag{3}$$

称为以 x_0 为中心，r 为半径的开球（Open ball）；
集合

$$\bar{B}(x_0, r) = \{x \in X | d(x, x_0) \leq r\}, \tag{4}$$

称为以 x_0 为中心，r 为半径的闭球（Closed ball）；
集合

$$S(x_0, r) = \{x \in X | d(x, x_0) = r\}, \tag{5}$$

称为以 x_0 为中心，r 为半径的球面（Sphere）。

注 1，在 $\mathbb{R}$ 中，开球

$$B(x_0, r) = (x_0 - r, x_0 + r)$$

在 $\mathbb{R}^2$ 中，开球

$$B(x_0, r) = \left\{ x \in \mathbb{R}^2 \middle| \sqrt{(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2} < r \right\}$$

注 2 从 Euclidlean 几何中延伸使用的 “开球”、“闭球、“球面” 这些概念，不一定具有在 $\mathbb{R}^3$ 中球体的几何直观效果。

下面我给出离散空间的定义，X 是一个非空集合，$x, y \in X$，定义

$$d(x, y) = \begin{cases} 1, x \neq y \\ 0, x = y \end{cases} \tag{6}$$

那么在此离散空间中，$B(x_0, \frac{1}{2}) = \{x_0\}$，因为除了他自己，其他所有的元素和他的距离都是 1。同理可得，$B(x_0, \frac{3}{2}) = \{D\}$，$B(x_0, 1) = \{x_0\}$。同理可得，$\bar{B}(x_0, 1) = D$，而 $S(x_0, 1) = D - \{x_0\}$。

设 X 是 $[0,T]$ 上全体连续函数组成的距离空间，距离定义为，

$$d_2(x,y)=\left\{\int_0^T |x(t)-y(t)|^2\right\}^{\frac{1}{2}} \tag{7}$$

对于 $x_0\in X$，开球 $B(x_0,\frac{1}{2})$ 表示定义在 $[0,T]$ 区间上，满足条件

$$\left\{\int_0^T |x(t)-y(t)|^2\right\}^{\frac{1}{2}}<\frac{1}{2} \tag{8}$$

的全体连续函数。

3 内点，开集，邻域

关于这三个概率的介绍，我们可以把 $\mathbb{R}^2$ 空间中的类似概念作为其来源和背景。

定义 3.1 设 X 是一个距离空间，$A\subset X$，若存在开球 $B(x_0,r)$，使得 $A\subset B(x_0,r)$，则称 A 是有界集。

定义 3.2 设 X 是一个距离空间，$G\subset X$，对于 $x_0\in G$，若存在开球 $B(x_0,r)$，使得 $B(x_0,r)\subset G$，则称 x_0 为 G 的内点。**内点是本章中非常重要的概念，之后我们会具体描述，为什么它会这么重要。**

定义 3.3 设 X 是一个距离空间，$G\subset X$，若 G 的每一个点都是内点，则称 G 是一个开集。

对于 $x\in X$，包含 x 的任一开集称为 x 的一个开邻域。

证明，开球 $B(x_0,r)$ 是一个开集。

要证 $B(x_0,r)$ 是开集，即要证明其中的每一个点都是内点。也就是 $B(x_0,r)$ 中的每一个点都存在开球包含的 $B(x_0,r)$ 中。

根据开球的定义可得 $\forall x_1\in B(x_0,r)$，都有 $d(x_0,x_1)<r$。取 $0<r_1<r-d(x_0,x_1)$，我们要证明

$$B(x_1,r_1)\subset B(x_0,r)$$

事实上，根据开球的定义，可以很简单的得出 $\forall x\in B(x_1,r_1)\to d(x_1,x)<r_1$。由距离的三角不等式，有

$$\begin{aligned}
d(x_0,x)&\le d(x_0,x_1)+d(x_1,x)\\
&\le d(x_0,x_1)+r_1\\
&< d(x_0,x_1)+r-d(x_0,x_1)\\
&=r
\end{aligned} \tag{9}$$

所以，$x\in B(x_0,r)$，即为 $B(x_1,r_1)\subset B(x_0,r)$。故 $B(x_0,r)$ 是开集。**注，开球是开集，而开集并不一定是开球。**

3.1 开集的性质

设 X 是距离空间，X 中的开集具有以下性质。

(1) 全空间与空集是开集；

(2) 任意多个开集的并集是开集；

(3) 任意有限多个开集的交集是开集。

第一条性质这里不给出过多的证明，这个结论比较显然。

第二条性质证明，G_a 是开集，其中 $\alpha \in I$。要证明 $G=\bigcup_{\alpha\in I} G_\alpha$ 是开集，$\forall x\in G$，$\exists \alpha_0 \in I$，使得 $x\in G_{\alpha_0}$。

因为 G_{α_0} 是开集，所以存在开球 $B(x,r)$，使得

$$B(x,r)\subset G_{\alpha_0}\subset G \tag{10}$$

所以 G 是开集。

第三条性质证明，设 $G=\bigcap_{k=1}^{n} G_k$，其中 G_k $(k=1,2,\cdots,n)$ 是开集，需要证明 G 是开集。此证明需要采用开集的定义。对于 $\forall x\in G$，则 $x\in G_k$。由于，每一个 G_k 都是开集，则存在 $B(x,r_k)\subset G_k$。由于集合的数量是有限的，并且每一个 r 都必然是大于零的。所以，我们这里取到了 $r=\min\{r_k\}$。注意这里为什么是有限，因为当集合的数量是无限个以后，距离的下确界将不一定大于零了，它有可能会等于零。(所以，在此处我们用到了"有限个开集"这个条件)。

于是，

$$B(x,r)\subset B\left(x,r_k\right)\subset G_k(k=1,2,\cdots,n) \tag{11}$$

即为，

$$B(x,r)\subset \bigcap_{k=1}^{n} G_k = G \tag{12}$$

所以 G 是开集。

如果一个集合 X 中有一个子集族 $\mathcal{J}$ 满足上述三个性质，我们则把它们称为开集族。或者说 $\mathcal{J}$ 是 X 的一个拓扑，$(X,\mathcal{J})$ 成为一个拓扑空间。即全体开集决定了空间的拓扑性质。那么，按这个思路来看，我们一步步从距离定义了开球，从开球定义了开集，从开集定义出了拓扑，定义是一步步逐渐深入的。在这样的拓扑机构下，类似于实数集上定义的连续函数一样，我们可以类似的定义距离空间上的连续映射。

老师从这里引入了一个拓扑的概念，这个概念是来自牛津字典。Definition of topology: Study of geometrical properties and spatial relations unaffected by continuous change of shape or size of figures.

我个人的理解是，图形在连续变换下，比如压缩，拉升等等，几何性质和空间关系没有受到影响的一些性质，被称之为基本的拓扑性质。

3.2 小结

我们从简单的二维空间出发，引出了开集和内点这样一些概念的思维来源。然后将其推广到无穷维空间，和更加一般化的距离空间中，从而得到了更加一般化的概念。有了这样一些概念以后，我们继续研究了开集的性质。抽象出来开集本质的一些特征，从而就可以引出拓扑了，也就是在连续变化下，没有改变的一些性质。下面，我们就要介绍连续映射和等价距离的概念和性质了。

4 等价的距离、连续映射

4.1 等价距离的定义

首先给出等价距离的定义，

定义 4.1 设 (X,d_1)，(X,d_2) 是定义在同一集合 X 上的两个距离，如果存在 $C_1>0$，$C_2>0$，使得对于 $\forall x,y\in X$ 都有

$$C_1d_1(x,y)<d_2(x,y)<C_2d_1(x,y)$$

则称距离 d_1，d_2 是等价的。等价最重要的含义是，收敛性是一样的。为什么说收敛性是一样呢？收敛的定义就是，$x_n\to x_0$ $(n\to\infty)$ 的时候有 $d(x_n,x_0)\to 0$ $(n\to\infty)$。根据这个不等式关系，可以很明显的感受到，如果 $d_1(x_n,x_0)\to 0$ 的话，必然有 $d_2(x_n,x_0)\to 0$。而且反之也是成立的，所以我们可以推出两个距离空间下的收敛性是一样的。即为：

$$x_0\xrightarrow{d}x(n\to\infty)\Leftrightarrow x_n\xrightarrow{d}x(n\to\infty)\tag{13}$$

而且距离 d_1,d_2 是等价的，它们产生的开集族 $\mathcal{J}_1$ 和 $\mathcal{J}_2$ 也是相同。这个也很好理解，因为距离是等价的，而开集的关系都是由距离给出的，所以必然开集族也是相同的。那么，下面我们将给出一个具体的例子。

4.2 等价距离实例

在 $\mathbb{R}^2$ 中，对 $\forall x=(x_1,x_2)$，$y=(y_1,y_2)$，定义距离，

$$\begin{aligned}d_1(x,y)&=|x_1-y_1|+|x_2-y_2|\\ d_2(x,y)&=\left(|x_1-y_1|^2+|x_2-y_2|^2\right)^{1/2}\\ d_\infty(x,y)&=\max\{|x_1-y_1|,|x_2-y_2|\}\end{aligned}\tag{14}$$

则 d_1,d_2,d_∞ 在 $\mathbb{R}^2$ 中是等价的，事实上：$\forall x,y\in\mathbb{R}^2$，我们可以简单的得到如下几个等式，

(1) $d_\infty(x,y)\le d_2(x,y)\le d_1(x,y)$；

(2) $d_2(x,y)\le\sqrt{2}d_\infty(x,y)$；

(3) $d_1(x,y)\le\sqrt{2}d_2(x,y)\le 2d_\infty(x,y)$。

这三个公式的推导非常的简单，这里就不做过多的描述了，显然从以上三个公式中，我们可以得出 $d_\infty(x,y)\le d_2(x,y)\le\sqrt{2}d_\infty(x,y)$。所以有 $d_2(x,y)\Leftrightarrow d_\infty(x,y)$。同理，可以得到，$\frac{\sqrt{2}}{2}d_1(x,y)\le d_2(x,y)\le d_1(x,y)$，所以可以得到 $d_2(x,y)\Leftrightarrow d_1(x,y)$。比较严谨的同学还会想到，这里得到了 $d_2(x,y)\Leftrightarrow d_\infty(x,y)$ 和 $d_2(x,y)\Leftrightarrow d_1(x,y)$，有没有传递性可以得出，$d_2(x,y)\Leftrightarrow d_\infty(x,y)\Leftrightarrow d_1(x,y)$。大家根据距离等价的定义去想想就知道，这个等价性必然是存在的。

而且，这个结果可以推广到 $\mathbb{R}^n$ 中。具体的证明思路是一样的，这里不做过多的描述。

4.3 连续映射的定义

在讲解连续映射之前，我们首先要了解一个连续函数的定义。设 $f(x)$ 是 $\mathbb{R}$ 上定义的实值函数（即为：$\mathbb{R}\to\mathbb{R}$ 的映射），如果满足：对于 $\forall\epsilon>0$，$\exists\delta>0$，当 $|x-x_0|<\delta$ 时，有

$$|f(x)-f(x_0)|<\epsilon$$

则称 $f(x)$ 在 x_0 点连续，如果在函数在其定义域内每个点都连续，则称函数在定义域内是连续函数。实际上把这里的绝对值可以看成是距离，连续和极限的定义都是距离空间为基础的。

而类似于实数空间中的连续函数，可在距离空间上定义连续映射。

下面给出连续映射的定义，

定义 4.2 连续映射：令 (X,d)，(X_1,d_1) 是距离空间，

$$T : X \to X_1$$

是一个映射，$x_0 \in X$，如果对于 $\forall \epsilon > 0$，$\exists \delta > 0$，当 $d(x,x_0) < \delta$ 时，有

$$d(T(x),T(x_0)) < \epsilon$$

则称 T 在 x_0 点连续。

其实这个和前面连续的定义有区别吗？显然都是一样的。可能看到这里还感觉比较简单，接下来我们将用更加一般化的语言来进行抽象。换用一种邻域的观点，从几何的角度来看这个问题，可能会更加的直观一点，注意这里的定义和之前的定义都是等价的。只不过思考的角度不一样，

对 $\forall \epsilon > 0$，存在 $\delta > 0$，使得，

$$TB(x_0,\delta) \subset B(T(x_0),\epsilon) \tag{15}$$

几何映射解释如下图所示，

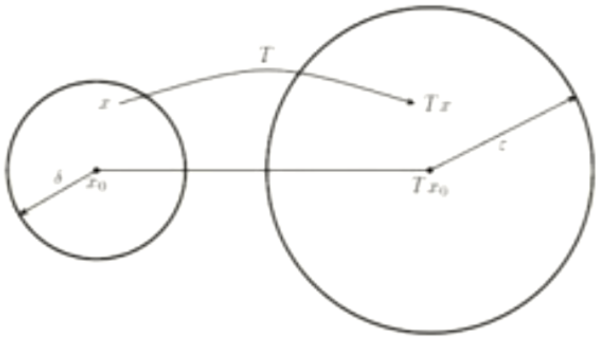

图 1: 二维空间中连续映射的概念图

通过图 1 我们可以看得比较清晰，其中 $B(x_0,\delta)$ 实际上就是 $d(x,x_0) < \delta$ 形成的开球。而 $B(T(x_0),\epsilon)$ 指的是 $d(T(x),T(x_0)) < \epsilon$ 形成的开球。而 $TB(x_0,\delta)$ 是对此开球中的元素做一个映射 T。映射后的元素一定是 $B(T(x_0),\epsilon)$ 中的内点，这就是从邻域的角度进行思考。

定义 4.3 等距映射，若 T 在 X 中的每一个点都连续则称 T 在 X 上连续，特殊的，若

$$d(x,y) = d_1(T(x), T(y)),\ \forall x, y \in X$$

则称 T 是等距映射。注 1 等距映射是连续映射，等距映射是 1-1 的，但不一定是满射的。如果，X, X_1 上存在一个等距在上的映射，

$$T : X \to X_1 (T(X) = X_1) \tag{16}$$

则称 X_1 和 X 是等距的。而且等距的两个空间，在等距的意义下可以认为是同一空间。

定义 4.4 连续映射的一般化定义：在开集的概念，我们可以把连续映射用开集来描述。设 T 是从距离空间 (X, d) 到距离空间 (X_1, d_1) 的映射，T 是连续的当且仅当 (X_1, d_1) 中任何开集的原像仍然是 (X, d) 中的开集。

其实在这个定义中，已经把距离的概念都隐掉了。在我们原始的定义中，本就没有距离的概念的。这才是连续函数的最基本的性质。

证明，

首先要证充分性 "$\Rightarrow$" 假设 T 是连续的，$S_1 \subset X_1$ 是开集。$S \subset X$ 是 S_1 的原像。也就是映射结果是开集，原像也要是开集。我们的条件中最重要的就是这两条标红的。下面的证明中，必须要用到这两个条件，没有用到就一定错了。

如果 $S = \emptyset$，这是显然成立的。

如果 $S \neq \emptyset$，我们需要利用开集的定义来证明，即为要证：对 $\forall x_0 \in S$，存在 x_0 的一个邻域包含在 S 中，而 T 把此邻域都映射到 S_1 中。简单的理解为，S 中关于 X_0 的一个邻域中所有的点，都可以通过 T 映射到 S_1 中 y_0 的一个领域内。

要证 $S \neq \emptyset$，对于 $\forall x_0 \in S$，要证 x_0 是 S 的内点，即证明存在 x_0 的一个邻域映射到 S 中，一个邻域包含在 S 中的含义是：T 把此邻域映射到 S_1 中。

令 $y_0 = Tx_0 \in S_1$，其中 S 是关于 S_1 的原像，关于 T。因为 S_1 是开集，所以必然存在 $B_1(y_0, \epsilon) \subset S_1$。

由 T 是连续的，根据公式 (15) 可得，对于上述 ϵ，存在 $\delta > 0$，使得$TB(x_0, \delta) \subset B_1(y_0, \epsilon) \subset S_1$。这意味着 $B(x_0, \delta) \subset S$，那么 x_0 是一个内点，可知 S 是开集。

根据下面这个图就可以清楚的看到了，

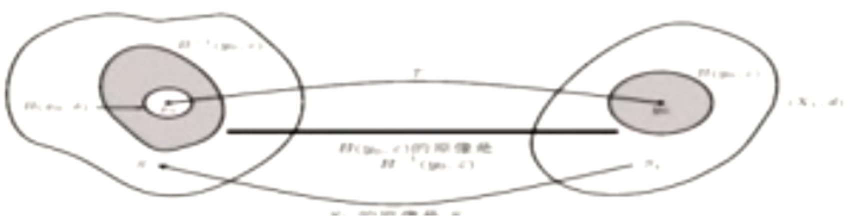

图 2: 开集连续映射的图像

"$\Leftarrow$" 则是关于必要性的证明，假设 X_1 中任何开集的原像仍是 X 中的开集，要证明 T 是连续的。即要证 T 在 X 中的每一点都连续。

对于 $\forall x_0 \in X$ 和 $\forall \epsilon > 0$，$B_1(T(x_0), \epsilon)$ 在 X_1 中是开的。而且他的原像 $N = T^{-1}B_1(T(x_0), \epsilon)$ 也是开集。并且 $x_0 \in N$，因此 N 中包含一个 x_0 的 δ 邻域 $B(x_0, \delta)$。注意左边这个大黑的集合就是映射的原像 N，那么我们必然有 $TB(x_0, \delta) \subset B_1(T(x_0), \epsilon)$。

那么由公式 (15) 式可得，T 在 x_0 处连续，x_0 是任意的，则 T 在 X 上连续。

4.4 小结

我们可以看到连续映射是一个拓扑概念，它最本质的抽象用一句话描述就是，开集的原像也是开集。所以，有了这个概念，推导得出两个连续函数的复合函数也是连续的就变得非常简单了，甚至是推导任意多个连续函数的复合都是连续的。简单想一想就知道，开集的原像也是开集，那么原像的原像也是开集，这样连续映射肯定可以满足拓扑性质。这样可以无限的套下去，就像俄罗斯套娃一样，任意多个连续函数的复合之间开集和其开集的原像也一定是开集。而在数学分析中对这个结论的证明是非常困难的，而在泛函分析中使用开集的概念来证明就变得相对比较简单了，这是当然的，我们使用的数学工具更加抽象了。

5 泛函分析学科思维简介

这是孙炯老师对泛函分析学习过程中的一些思考，以及上升到数学学习的一些想法。

随着课程的进行，大家或许会感觉泛函的学习越来越抽象。泛函分析是分析系列中非常重要，也是非常困难的一门课程。俗话说得好“实变函数学十遍，泛函分析心泛函”并不是没有道理的。

学习泛函分析的困难首先在于它的抽象，抽象是从众多的事物中抽取出共同的、本质性的特征，而舍并其非本质的特征例如苹果、香蕉，葡萄等，它们共同的特性就是都是水果。得出水果概念的过程就是一个抽象的过程，要抽象就必须进行比较，没有比较就无法找到共同的部分。所谓的共同特征是相对的，是指从某一个侧面看是共同的。

在抽象时同与不同，决定于从什么角度上来抽象，抽象的角度取决于分析问题的目的。

其实大家想想我们距离这些概念最原始的都是从之前学习的二维坐标系中的距离中抽象出来的。所以以这个知识点为背景，再去理解整个距离空间会简单很多。

抽象是我们数学研究中，非常重要的一种手段。数学的抽象就是数学建模。而泛函分析不但将古典分析的基本概念和方法抽象化。而且还将这些概念和方法几何论。我们研究泛函的目的是如何学会抽象这种思维方式。

6 可交换的极限顺序思考连续映射

6.1 可交换的顺序思考连续映射

之前我们从开集的角度定义了连续映射，下面我们将从极限运算可交换的角度思考连续映射。

定理 4.1: 连续映射设 T 是从距离空间 (X, d) 到距离空间 (X_1, d_1) 的映射。T 在 x_0点是连续的当且仅当对于每个满足条件 $\lim_{n\to\infty} x_n = X_0$ 的点列都有

$$\lim_{n\to\infty} T(x_n) = T(x_0) = T(\lim_{n\to\infty} x_n) \tag{17}$$

此定理说明，若 T 连续，则极限运算可以和 T 交换顺序。

证明：

首先我们要证的是**充分性**，假设 T 是连续的，且 $\lim x_n = x_0$，要证明

$$\lim_{n\to\infty} T(x_n) = T(x_0) = T(\lim_{n\to\infty} x_n)$$

由 T 是连续的，则对于 $\forall \epsilon > 0$，使得当 $d(x, x_0) < \delta$ 时，有

$$d_1(T(x), T(x_0)) < \epsilon$$

因为 $\lim x_n = x_0$，则对于 $\delta > 0$，存在 N，当 $n > N$ 时，有 $d(x_n, x_0) < \delta$，于是有

$$d_1(T(x_n), T(x_0)) < \epsilon$$

即为

$$\lim_{n\to\infty} T(x_n) = T(x_0)$$

又因为 $\lim x_n = x_0$，所以我们可以得到

$$\lim_{n\to\infty} T(x_n) = T(x_0) = T(\lim_{n\to\infty} x_n) \tag{18}$$

第二步则是要证**必要性**，在已知 $\lim_{n\to\infty} T(x_n) = T(x_0) = T(\lim_{n\to\infty} x_n)$ 的情况下，显然有 $T(x_n) \to T(x_0)$，所以我们要证明 T 在 x_0 点连续。

我们需要采用反证法来进行证明，假若不然，T 在 x_0 点不连续。则存在 $\epsilon_0 > 0$，对于任何的 $\delta > 0$，都存在一个 x_0，使得 $d(x_0, x_n) < \delta$，但是有

$$d_1(T(x_n), T(x_0)) \geq \epsilon_0$$

取 $\delta = 1, \frac{1}{2}, \cdots, \frac{1}{n}, \cdots$。对应的有 $x_1, x_2, \cdots, x_n, \cdots$，满足 $d(x_n, x_0) < \frac{1}{n}$，但是有 $d_1(T(x_n), T(x_0)) \geq \epsilon_0$ $(n = 1, 2, 3, \cdots)$。

那么现在根据上述的推导可以得出 $x_n \to x_0$ $(n \to \infty)$，但是 $T(x_n)$ 不趋近于 $T(x_0)$，这与条件矛盾。

6.2 实例

例 1. 设 $X = C[a, b]$，令 $T(x) = \int_a^b x(t)dt$，则 T 是从 $C[a, b]$ 到 $\mathbb{R}$ 的映射 (这里实际上就是一个线性泛函)，由于

$$\begin{aligned} d(T(x), T(y)) = \left| \int_a^b x(t)\mathrm{d}t - \int_a^b y(t)dt \right| &\leq \int_0^b |x(t) - y(t)|dt \\ &\leq |b - a| d(x, y) \end{aligned} \tag{19}$$

那么，根据公式 (19) 是显然可以符合映射连续的条件的，所以 T 是连续的。

例 2. 若 $\{x_n(t)\}$ 是 $C[a, b]$ 中的收敛序列，即为

$$x_n \underset{d_{C[a,b]}}{\to} x \ (n \to \infty)$$

由定理 4.1 可得，

$$\lim_{n\to\infty} T(x_n) = T(x_0) = T(\lim_{n\to\infty} x_n) \tag{20}$$

这里的 $T(x) = \int_a^b x(t)dt$，我们有：

$$\lim_{n\to\infty} \int_a^b x_n(t)\mathrm{d}t = \int_a^b \lim_{n\to\infty} x_n(t)\mathrm{d}t = \int_a^b x(t)\mathrm{d}t \tag{21}$$

注意到 $\{x_n(t)\}$ 在 $C[a,b]$ 中的收敛是：

$$\max_{a\leq t\leq b} |x_n(t) - x(t)| \to 0 \ (n\to\infty) \tag{22}$$

由于此数列的收敛与具体的点无关，所以可知此函数列是一致收敛的。而我们之前在数学分析中就学习过，如果函数是一致收敛的，那么可以交换积分和极限的顺序，这是数学分析中熟知的结论。

可否交换积分和极限的顺序对于运算来说是非常重要的，比如在重积分中，如果可以交换顺序很多时候是可以计算出来的，如果不可以交换则是算不出来的。而定理 4.1 把结论更一般化了。

只要映射是连续的（在距离空间下），极限运算和这个映射就可以交换运算顺序。而积分和极限运算的交换顺序只是其中的一个特例。

7 总结

本章首先介绍的是开集和连续映射的内容。其中开集的定义非常的重要，开集是集合中的每一个点都是内点组成的集合。然后描述并证明了开集的性质，

(1) 全空间与空集是开集；

(2) 任意多个开集的并集是开集；

(3) 任意有限多个开集的交集是开集。

系统的学习了开集之后，很自然的就可以引出拓扑的概念了，图形在连续变换下，比如压缩，拉升等等，几何性质和空间关系没有受到影响的一些性质，被称之为基本的拓扑性质。

文章的第二部分，描述的是开集引出的重要概念“连续映射”。其中首先介绍的是等价距离，由距离等价可以得出它们产生的开集族是一样的。之后，我们分别从三个角度来定义了连续映射，

(1) 距离的角度，我们从距离的定义，引出了连续映射的定义。

(2) 拓扑的角度，舍弃距离的具体表达形式，采用开集的原像一定是个开集来定义连续映射。

(3) 运算的角度，只要映射是连续的（在距离空间下），极限运算和这个映射就可以交换运算顺序。

同时通过一些例子，我们发现很多分析中很难的问题，在泛函分析中得到了简化和推广，比如 n 个连续函数复合还是连续函数；如果函数是一致收敛的，那么可以交换积分和极限的顺序。

Functional Analysis 1.4 Close Set Separability

Chen Gong

31 August 2020

目录

1 Introduction

上一节中，我们介绍了开集和连续映射的概念。本节讲述的是闭集，可分性和列紧性的内容。闭集的概念比开集的概念就难多了。本节的主要内容是，研究闭集的定义和性质，包含两个内容：

(1) 利用开集研究闭集。

(2) 从点集结构上研究闭集。

2 距离空间的闭集

2.1 闭集的定义

定义 2.1 闭集的定义设 X 是一个距离空间，若它的补集 $A^c = X \setminus A$ 是开的，集合 $A \subset X$ 称为闭的。

定理 2.1 设 X 是一个距离空间,则 $B(x_0, r) = \{x \in X | d(x, x_0) \leq r\}$ 和 $S(x_0, r) = \{x \in X | d(x, x_0) = r\}$ 是闭集。

而之前讲的开集是将上式中的 $\leq$ 换成 $<$。

那么现在要如何证明定义 2.1 呢？根据闭集的定义，只要证明一个集合的补集是开集即可，证明开集使用开集的定义即可。

我们可以将条件改写为，设 $y \in \bar{B}(x_0, r)^c$，要证 y 是 $\bar{B}(x_0, r)^c$ 的内点。根据条件可得，

$$d(y, x_0) = \alpha > r$$

那么令 $\beta = \alpha - r > 0$，对于 $\forall z \in B(y, \beta)$，有

$$\begin{aligned} d(x_0, z) &\geq d(y, x_0) - d(y, z) \\ &= \alpha - d(y, z) \\ &> \alpha - \beta \\ &= r \end{aligned} \tag{1}$$

所以对于 $\forall y \in \bar{B}(x_0, r)^c$，$\exists$ 邻域 $B(y, \beta)$，邻域中的点都是大于 r 的。从而推出 $B(y, \beta) \subset \bar{B}(x_0, r)^c$。所以 $\bar{B}(x_0, r)^c$ 是开的，于是 $\bar{B}(x_0, r)$ 是闭集。

由 $S\left(x_a, r\right)^c = B\left(x_0, r\right) \cup B\left(x_0, r\right)^c$ 是开的。(根据任意多个开集的并集是开集) 因此 $S(x_0, r)$ 是闭集。

记 $\mathcal{F}$ 为距离空间 (X, d) 中全体闭集，利用关于补集的 De Morgan 公式，结合 (开集的性项，决定空间拓扑结构的三条性质)，得

定理 2.2 设 (X, d) 是一个距离空间，则

(1) 全空间与空集是闭集；

(2) 任意多个闭集的交是闭集；

(3) 有限多个闭集的并是闭集。

上面我们是从开集和补集的角度来说明什么是闭集，接下来我们将换个角度来思考。数学概念学习的角度当然不是单一的，概念的引出是多方面的，我们只有从多个角度全面的理解数学概念，才能学得更加的扎实。

2.2 闭集的结构

当时我们也讲了开集的结构，就是都是由内点组成的。刚刚上面给出的闭集的定义是说他的补集是开集的集合，大家可能没有什么感觉。

而闭集的结构相对比较复杂，这从 Cantor 集是闭集就可以反映出来。下面从点集的结构上进一步研究闭集。

定义 2.2 设 X 是一个距离空间，$A \subset X$，$x \in X$。如果对于 $\forall \epsilon > 0$，球 $B(x,\epsilon)$ 中都包含 A 中的点，即为

$$B(X,\epsilon) \cap A \neq \emptyset \ (\forall \epsilon > 0) \tag{2}$$

则称 x 为 A 的接触点。这里和我们之前学的交集的思路很像，我们在这里做一个简单的类比。那么显然会有，A 中的点一定是 A 的接触点，A 的接触点可能属于 A 也可能不属于 A。类比一下交集就知道了。

定义 2.3 聚点： 设 X 是一个距离空间，$A \subset X$，$x \in X$。如果对于 $\forall \epsilon > 0$，球 $B(x,\epsilon)$ 中都包含 A 中不同于 x 的点，即

$$B(x,\epsilon) \cap (A \setminus \{x\}) = \emptyset \tag{3}$$

则称 x 是 A 的聚点。而且聚点一定是接触点，而反过来则不一定。也就是 x 的去心领域在 A 中。下面来举个简单的例子，

例 2.2.1，a 点和 b 点是开区间 (a,b) 上的聚点（也是接触点），但不属于开区间 (a,b)，而闭区间 $[a,b]$ 的聚点全部都在 $[a,b]$ 中。

例 2.2.2，设在连续函数空间 $X = C[0,T]$ 中，且 $A = \{x(t) \mid x(0) = 0, 且|x(t)| < 1(0 \leq t \leq T)\}$。我们可以注意到 $x_0(t) = 1$ 不是 A 的接触点。因为对于所有的 $x \in A$，其中都有 $d(x_0, x) \leq 1$（这里用到的是无穷阶的距离）。所以没有点会落到 A 里面，必然不是接触点。

例 2.2.3，设 X 表示由 $[0,T]$ 上的全体连续函数组成的集合，定义为，

$$d_2(x,y) = \left\{ \int_0^T |x(t) - y(t)|^2 dt \right\}^{1/2}$$

集合还是 2.2.2 中的那个集合，$A = \{x(t) \mid x(0) = 0, 且|x(t)| < 1(0 \leq t \leq T)\}$。则可以证明 $x_0(t) = 1$ 是 A 的接触点（也是 A 的聚点），这个证明就很简单了，此距离求的是两个函数之间的平均距离。所以可以看到，一个点是否是一个集合的接触点和空间的距离有关。而在实际应用中，我们需要根据不同的问题来定义不同的距离空间。

2.2.1 从接触点的定义上定义闭集

下面将从结构上，用接触点来定义闭集的概念。

定义 2.4 设 X 是一个距离空间，$A \subset X$，A 的接触点的全体称为 A 的闭包，记为 $\bar{A}$。又因为 A 中的点一定是 A 的接触点，所以 $A \subset \bar{A}$。

定理 2.3 设 X 是一个距离空间，$A \subset X$，A 是闭集当且仅当 $A = \bar{A}$。

下面我们将给出其证明过程，第一步是证明其必要性：由 $A = \bar{A} \Rightarrow A$ 是闭集。思路比较简单，即为证明 A^c 是开的。

令 $x \in A^c$，由于 $A = \bar{A}$，所以 x 不是 A 的接触点，必然存在 ϵ 使得，

$$B(x,\epsilon) \cap A = \emptyset \tag{4}$$

且，$\bar{A}$ 表示为所有接触点的集合，那么 A^c 表示为所有的非接触点，而根据公式 (4) 可得，x 的一个领域中都是非接触点。则有 $B(x,\epsilon) \subset A^c$，因此 A^c 是开集，即 A 是闭集。

下一步则是证明充分性，即为证明 "A 是闭集，$A \subset X$，$A \Rightarrow A = \bar{A}$"。其中 $\bar{A}$ 是 A 的全体接触点的集合。要证 $A = \bar{A}$，由于有 $A \subset \bar{A}$（因为 A 中的点一定是 A 的接触点），下一步则是得到 $\bar{A} \subset A$ 即可。

这里的思路采用反证法来证明，令 $x \in \bar{A}$，假若有 $x \bar{\in} A$，即有 x 属于开集 A^c（这是因为 A 是闭集）。于是存在 ϵ_0，使得 $B(x,\epsilon_0) \subset A^c$，也就是，

$$B(x,\epsilon_0) \cap A = \emptyset \tag{5}$$

那么，x 不是 A 的接触点，这与 $x \in \bar{A}$ 相互矛盾。所以原证明成立！

2.2.2 从极限的角度来定义闭集

之所以从不同的角度来反复的定义"闭集"，是要从多方面理解闭集的性质。这一部分将从极限的角度来定义闭集。

定理 2.4 设 X 是一个距离空间，$A \subset X$，A 是闭集当且仅当 A 中收敛点列 $\{x_n\} \subset A$ 的极限属于 A。这句话可以理解为**在闭集里极限运算是封闭的**。

证明思路还是一样，我们需要从充分性和必要性两个角度来证明。首先是充分性条件的证明，**设 A 是闭的，且 $\{x_n\} \subset A$ 收敛，$\lim_{n\to\infty} x_n = x_0$，要证明** $x_0 \in A$。那么这个证明思路是怎样的呢?

由于 $\lim_{n\to\infty} x_n = x_0$，根据极限的定义，$\forall \epsilon > 0$，都存在 x_n 满足 $d(x_n, x_0) < \epsilon$。这句话翻译一下就是 $x_n \in B(x_0, \epsilon)$，这表明 x_0 是 A 的接触点（x_n 是 A 中的点），有 $x_0 \in \bar{A}$。由于，A 是闭集，所以 $A = \bar{A}$，$x_0 \in A$。

这一段的逻辑性非常的强，根据极限的定义得出 x_0 是接触点，根据 $\bar{A} = A$ 得出 $x_0 \in A$。

下一步则是证明必要性，假定每个收敛点列的极限都属于 A，要证 A 是闭集。根据定理 2.3 中给出的条件，只要证明 $A = \bar{A}$，即由 $x_0 \in \bar{A}$ 推出 $x_0 \in A$。

令 $x_0 \in A$，根据闭包的定义（闭包是全体接触点），对 $\forall n \in N$，在 $B(x_0, \frac{1}{n}) \cap A$ 中至少存在一点 x_n，显然

$$\lim_{n\to\infty} x_n = x_0 \tag{6}$$

由已知得出，$x_0 \in A$，因此 A 是闭的。

3 可分的距离空间

上述文章中，我们已经比较全面的分析了什么是闭集。本章中将可分的概念从一般实数空间，推广到一般空间中。

实数空间中，有理数是稠密的，有理数是可数的（可列可数的概念相信大家都有较好的认识了）。而任意一个实数都可以用有理数列来逼近。我们希望把这样的性质推广到一般空间中。**重点：理解稠密集的定义，由其定义可分距离空间；判断稠密性，距离空间的可分性**。

定义 3.1（稠密集）：设 A, B 是距离空间 X 中的点集，如果 $A \subset \bar{B}$，则称 B 在 A 中稠密。

而如果用 $\epsilon - \delta$ 语言描述即为，$\forall x \in A, \because x \in \bar{B}, \therefore \forall \epsilon > 0, B(x,\epsilon) \cap B \neq \emptyset$，即为存在 $y \in B$，使得 $d(x,y) < \epsilon$。

也就是说 A 中的每一个点都可以用 B 中的点来逼近：

$$\forall x \in A,\ \forall \epsilon > 0,\ \exists y \in B,\ \text{s.t. } d(x, y) \leq \epsilon$$

注意，定义并没有要求 $B \subset A$。下面举两个例子：

例 3.1 $A = [0, 1]$，B 是 $[0, 1]$ 中全体有理数。$\bar{B} = [0, 1]$，$A \subset \bar{B}$。所以 B 在 A 中稠密，这里的 $B \subset A$ 且 $B \neq A$。

例 3.2 如果 A 是 $[0, 1]$ 中全体无理数，B 是 $[0, 1]$ 中的全体有理数，我们有 $A \subset \bar{B}$，即有 $\bar{B}$ 在 A 中稠密，但是 $\bar{A} \cap A = \emptyset$。

显然有理数是一个稠密的，并且它还是可数的，这个性质非常的不错（比如说实数空间中的一点 π，可以用一个可数的有理数数列来逼近：$\{3, 3.1, 3.14, 3.141, \cdots\}$）。我们可以类似的将可数的特点推广到一般的空间中。

定义 3.2（可分距离空间）：设 X 是距离空间，如果 X 中存在一个可数稠密子集，则称 X 是可分的。即为，对于子集 $A \subset X$，如果 X 中存在可数子集 B，使得 B 在 A 中稠密，则称 A 是可分的。

我们还可以换个角度来以换个角度来思考一下。距离空间 (X, d) 是可分的，当且仅当存在一个具有下列性质的可数集 $\{x_n\}$：

$\forall x \in X$ 和 $\forall \epsilon > 0$，至少存在一个 $x_k \in \{x_n\}$，$s.t.$

$$d(x_k, x) < \epsilon$$

即为，可分空间中的任意一点可通过一个可数集来近似逼近，并且对于 $\forall \epsilon > 0$，$\bigcup_{k=1}^{\infty} B(x_k, \epsilon) = X$。**在数学中，大家可以深刻感受到，同一个问题，我们从不同的角度来思考，得到的感受是完全不一样的。所以，将一个问题放在不同的背景下去琢磨是非常有必要的，可以帮助我们加深对一个概念的理解，更好的掌握这个知识点。**

上面我们介绍了什么是可分空间，接下来则将用几个例子来强化记忆。

例 3.2 $\mathbb{R}^n$ 是可分的。

$\mathbb{R}^n$ 中的有理点（各个坐标都是有理数）是可数集，且在 $\mathbb{R}^n$ 中稠密。这个很显然可以由一维的数轴推广到有限维空间中，还是比较简单的。

例 3.3 $C[a, b]$ 是可分的。我们已经在有限维的实数空间中探究了可分的问题了，下一步则是将其推广到无穷维的空间中。

只要找到 $C[a, b]$ 中的可数稠密子集 $\{x_n(t)\}$ 即可。其中 $\{x_n(t)\}$ 是 $C[a, b]$ 上的连续函数。根据定义 3.2 中判断可分距离空间的方法，我们需要找到 $\{x_n(t)\}$ 应满足：对于 $\forall x(t) \in C[a, b]$ 和 $\forall \epsilon > 0$，至少存在一个 $x_n(t)$，使得

$$d(x_n, x) < \epsilon$$

我们要类比在有限维空间中的证明方法，想想在无穷维的函数空间中如何证明可数稠密子集，其思路为任何一个点，都可以用有理点数列来逼近。借鉴有限维空间中的方法，在函数空间中证明思路为，**连续函数用多项式逼近（类似于泰勒公式），多项式可以用有理多项式来逼近，即为任何一个连续函数，可以用一个系数为有理数的多项式函数列来近似。**全体有理系数多项式是 $C[a, b]$ 中的可数子集，所以 $C[a, b]$ 可分。

证明，由 Weierstrass 逼近定理，对于 $\forall x(t) \in C[a,b]$，存在多项式 $P_n(t)$ 一致收敛到 $x(t)$。(这部分证明，参阅 P.M. Fitzpatriclc 《Advanced Calculus》 p.188)。

即对于 $\forall x(t) \in C[a,b]$，当 $n \to \infty$，$\forall \epsilon > 0$，存在

$$P_n(t) = a_0 + a_1 t + \cdots + a_n t^n$$

其中 $a_i \in \mathbb{R}(i = 1, 2, 3, \cdots, n)$，

$$|P_n(t) - x(t)| < \epsilon \qquad (\forall t \in [a,b])$$

上面得到了每一个连续函数可以用多项式函数列来逼近，而下一步则是得出，每一个多项式函数列可以用系数为有理数的函数列来逼近。则为：$P_n(t)$ 又可以用 $P_n^r(t)$ 一致逼近，其中 $P_n^r(t) = r_0 + r_1 t + \cdots + r_n t^n$，其中 $r_0, \cdots, r_n$ 是有理数。

就可以得出 $\{P_n(t)\}$ 是可数的，$C[a,b]$ 是可分的。

例 3.4 l^∞ 是不可分的。

分析：用反证法来证明。

证明：$l^\infty = \{$全体有界的实数列$\}$。

$$x = \{\xi_k\},\ y = \{\eta_k\} \in l^\infty$$

$$d(x,y) = \sup_k |\xi_k - \eta_k|$$

设 $x = \{\xi_k\}$，其中 $\xi_k = 0/1$。这样的 x 的全体记为 A，而且 A 是连续势（x 可看成无限的由 0/1 组成的数列，而实数空间中 $[0,1]$ 的每一个数的小数部分都可以用二进制数来表示，显然有 $A \subset l^\infty$，并且 A 与实数空间 $[0,1]$ 中的每一个数都对应，所有简单思考可知 A 是不可数的）。而且，$\forall x, y \in A$，$x \neq y$，则

$$d(x,y) = \sup_k |\xi_k - \eta_k| = 1 \tag{7}$$

这里是为什么呢？因为两个二进制数列肯定都是不一样的，那么必然存在两个数位是不同的，而此次采用的距离是取每个数位上的最大差值，所以距离恒等于 1。如果 l^∞ 是可分的，则必然存在可数的稠密子集 E，对于 $\forall x \in E$，作开球 $B(x, \frac{1}{3})$，有 $l^\infty = \bigcup_{x \in E} B(x, \frac{1}{3})$。因此可以得到关系为：

$$A \subset l^\infty = \bigcup_{x \in E} B(x, \frac{1}{3})$$

由于 l^∞ 是可数的，而 A 是不可数的。则必有在某开球内含有 A 中一定可以找到两个不同的点（抽屉原理），$x, y \in A$，使得 $x, y \in B(x_0, \frac{1}{3})$，其中 $x_0 \in E$。所以

$$d(x,y) < \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

这显然与 $d(x,y) = 1$ 矛盾，所以是不可分的。有同学可能会疑问，为什么要取 $\frac{1}{3}$？实际上取任意一个小于 $\frac{1}{2}$ 的数都行，都可以形成与 $d(x,y) = 1$ 的矛盾。**这个例题的证明思路即为找到一个反例来说明 l^∞ 是不可分的。**

下面的例子将说明，距离空间是否可分，与空间上距离的定义密切相关。这个大家应该比较好理解。

例 3.5 空间 s 可分，其中 s 是全体实数列组成的集合，其上距离为，

$$d(x,y)=\sum_{k=1}^{\infty}\frac{1}{2^k}\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|} \tag{8}$$

证明 s 可分，就要找出 s 中的可数稠密子集。那么就来构建一个可数子集，令 $s_0\{(r_1,r_2,\cdots,r_n,0,0,\cdots)\}$（我其实想知道，后面这堆零是怎么想到的，有什么用吗？前面的有理数我还能想到理由），其中 $r_k(k=1,2,\cdots,n)$ 是有理数，有理数必然是可数的，所有 s_0 是可数的。下面的重点则为证明 s_0 是 s 的稠子集。

根据稠密集的定义，要证明一个可数集是稠密性，要找到一个具有下列性质的可数集 $\{x_n\}$：$\forall x\in X$ 和 $\forall\epsilon>0$，至少存在一个 $x_k\in\{x_n\}$，s.t. $d(x_k,x)<\epsilon$。那么在此例题中，对 $\forall x\in s$，$x=(\xi_1,\cdots,\xi_j,\cdots)$，对于每一个 $\xi_j(j=1,2,\cdots)$，存在 $r_j^{(n)}\to\xi_j(n\to\infty)$。这个逻辑是这样的，**对于任何一个 $x_n\in s$ 中的元素，我们都能找到一个 $x\in s_0$，使得在距离空间 d 中 $x_n\to x$。**

其中，$r_j^{(n)}$ 是有理数，令

$$x_n=\{r_1^{(n)},r_2^{(n)},\cdots,r_j^{(n)},\cdots,r_n^{(n)},0,\cdots\}$$

也就是说随着 n 的增大，x_n 中非零元素的个数也在增大，且等于 n，所以说 x_n 收敛到 x。则 $x_n\to x$，且 $x_n\in s_0$，所以 s 是可分的。

注 l^∞ 是全体有界的实数列，从集合的角度来看，l^∞ 是 s 的一个子集。但作为距离空间的 l^∞ 不可分，而 s 可分。那么距离空间的定义，对其是不是一个稠密集非常的重要，大家可以直观的想象一下，要想存在一个可数的稠密集，是不是需要距离空间将其元素“压”得比较紧，而如果距离空间比较松的话就不行了，公式 (7) 和公式 (8) 中的举例，就比较好的验证了我们的想法。

4 列紧的距离空间

在数学分析中，闭区间上的连续函数有着良好的性质。

1. 闭区间满足有限覆盖定理；
2. 进一步说，平面上的有界闭集也有这样的性质；
3. 我们有具有这样的性质的集合，抽象为紧集（紧空间）。

而紧性的刻画，可以从不同的角度给出几种定义，**1. 序紧列（Weierstrass 定理：$\{x_n\}$ 有界，一定存在收敛的子列。）；2. Borel 紧（有限覆盖定理：即为一个闭集可以找到有限个开集来覆盖整个区间。）；3. 完全有界**。

表面上看，序紧列和 Borel 紧似乎没有什么关系。序紧列是从子列的角度来定义紧列的，而 Borel 紧是从有限个开集的角度来定义，小伙伴们似乎有点懵逼。我想了一会儿也没有想出来，但是数学家们发现了序紧列和 Borel 紧之间是紧密的联系的（这或许就是我不是数学家的原因吧）。**将这两者联系起来在数学上是非常具有意义的，这样可以将两个看似没有关联的概念联系起来。**

在数学分析中国，根据 Bolzano-Weierstrass 定理我们可以知道，实数域中每个有界的无穷点集至少有一个聚点，下面将在一般的距离空间中研究这种性质。

定义 4.1 设 A 是距离空间 X 中的一个子集，如果 A 中的每一个无穷点列都有一个收敛的子列，则称为 A 是列紧的集合，闭的列紧集称为是自列紧集 (所以根据定义，一个集合 A 的自列紧的，则要

求收敛子列的极限必须在 A 中)。这是对于集合来说的，如果对于空间来说则是，距离空间 X 称为是列紧的，如果 X 中的每一个无穷点列都有一个收敛的子列。

而我们的目标则是把 Weierstrass 定理推广到一般空间中。

定理 4.1 设 X 是一个距离空间，$A \subset X$ 是列紧集的，则 A 是有界集。

证明：假设 A 是无界的。于是可以从 A 中选取一个点列 $\{y_n\}$，使得 $d(y_n, a) > n$，其中 a 是 X 中的一个点。由于这个点列的任意子列都是无界的，所以这个点列没有收敛的子列（因为收敛的点列是有界的)。这显然与列紧相矛盾，所以 A 有界。

注 1，自列紧集是有界闭集；因为自列集就是闭集，而紧集是有界集。而自列紧集即为有界闭集。

注 2，在一般的距离空间中，有界的闭集不一定是列紧的。这是无穷维空间中，最重要的性质之一。注 1 反过来就不行了。而且，这个和有限维中的情况不一样了。在有限维的情况下，对于闭集，通常可以找到一个子列收敛的。比如：

$$\begin{gathered} 1, 0, 0, \cdots) \\ (0, 1, 0, \cdots) \\ (0, 0, 1, \cdots) \\ \cdots \end{gathered} \tag{9}$$

定理 4.2 设 f 是定义在列紧的距离空间 (X, d) 上的实值连续函数。则 f 是有界的，即为，

$$\begin{aligned} M &= \sup\{f(x)|x \in X\} \\ m &= \inf\{f(x)|x \in X\} \end{aligned} \tag{10}$$

是有限的，进一步，存在点 $x_{\max}$ 和点 $x_{\min}$ 使得，

$$f(x_{\max}) = M, \quad f(x_{min}) = N$$

实际上这里讲的非常抽象，我们可以类比到一维连续函数闭区间上去类比理解，一维连续函数的闭区间上是必然有界的，而且可以取到最大最小值。而实际上，在无穷维空间中，仅仅是有界闭集是不够的，而将有限维空间中有界的条件推广到无穷维空间，还需要列紧的概念。我们为什么要得到列紧的空间呢？因为列紧的性质是最好的，字面上理解就是列紧就是空间被压得很紧，便于研究上下界和收敛性。

所以，下面将研究在具体空间中什么样的集合是列紧的。

例 4.1 $\mathbb{R}^n$ 中有界闭集是自列闭集。例如闭区间 $[a, b]$ 是自列紧集。这个还是比较容易想到的。

定理 4.3 Arzela 定理：$C[a, b]$ 中的子集 A 是列紧的当且仅当 A 中的函数是**一致有界和等度连续**的。

什么是一致有界的定义？及存在 $K > 0$，使得对于每一点 $t \in [a, b]$ 以及一切的 $x \in A$，都有

$$|x(t)| < K \tag{11}$$

可以理解为集合内的所有连续函数的上界都是同一个。什么是等度连续呢？对于任意的 $\epsilon > 0$，存在 $\delta > 0$，当 $|t_1 - t_2| < \delta$ 时，

$$|x(t_1) - x(t_2)| < \epsilon \qquad (\forall x \in A) \tag{12}$$

此连续的定义和一致连续很像，一致连续和区间内具体的点无关，用来控制收敛的速度。而等度连续中将 $(\forall x \in A)$ 去掉就是一致连续。等度连续就是空间中所有的连续函数，对于任何的 $\epsilon > 0$，都可以找到一个共同的 δ，使得公式 (12) 成立。**这是什么意思呢？就是空间中所有的连续函数都是一致连续的，而且所有函数收敛的速度都是一致的**。

大家是不是可以深刻感受到空间列紧的强大之处了，是空间中好的性质。列紧是距离空间中来考虑的，跳出了实数空间。什么是函数集合列紧呢？列紧即为对于任意一个无穷点列都可以找到一个函数子列，这个函数子列是收敛的。并且集合中的还是是有界连续和等度收敛的。

例 4.2 $A = \{x(t) \in C^1[a,b] \mid |x(t)| \leq M, |x'(t)| \leq M_1\}$。集合中的元素是连续的，而且一阶导数也是连续的，而且函数值有界，导数值也有界。则 A 是 $C[a,b]$ 中的列紧集。($C^1[a,b]$ 是在 $[a,b]$ 中全体连续可微的函数)。

首先看是不是一致有界的，集合中的条件 $|x(t)| \leq M$ 已经可以满足了。

等度连续：可以由中值定理得，$\forall x \in A$，$\forall t_1, t_2 \in [a,b]$，存在 $\theta \in (0,1)$，使得

$$\begin{aligned} |x(t_1) - x(t_2)| =& |x'(\theta t_1 + (1-\theta)t_2)(t_1 - t_2)| \\ \leq & M_1 |t_2 - t_1| \end{aligned} \tag{13}$$

所以 A 中的函数等度连续，且一致有界，所以是列紧的。

5 本章小结

本节中，首先介绍了闭集的定义，从三个角度来介绍了闭集的定义，1. 闭集的补集是开集；2. 从结构的角度说，集合中所有的点都是接触点，则是闭集。3. 从极限的角度思考闭集，在闭集里极限运算是封闭的。在数学中，大家可以深刻感受到，同一个问题，我们从不同的角度来思考，得到的感受是完全不一样的。所以，将一个问题放在不同的背景下去琢磨是非常有必要的，可以帮助我们加深对一个概念的理解，更好的掌握这个知识点。

有了闭集的定义，下一步将可分的概念从一般实数空间推广到一般空间中。首先定义了稠密集的概念，稠密集是可以简单理解为集合中的每一个点都可以用另一个集合中的点来逼近。有了稠密集的概念则进一步定义了可分距离空间的定义，即为空间中存在可数稠密子集则为可分空间。

最后，则是介绍了列紧空间，列紧空间定义为，距离空间的子集中的每一个无穷点列都有一个收敛的子集，则为列紧的集合，闭的列紧集为自列紧集。为什么要有列紧的概念呢？在无穷维的空间中，有界的闭集不一定是列紧的，列紧的性质更好，字面上理解就是列紧就是空间被压得很紧，便于研究上下界和收敛性。而且，函数空间是列紧的，其中的函数是等度连续和一致有界的。

完备的距离空间

Chen Gong

03 October 2020

目录

1 Introduction

上一节主要介绍了空间的可分性，列紧性，这是距离空间中非常重要的性质。这节中将主要讲述距离空间的完备性，这将是距离空间中最重要的性质。

2 Cauchy（柯西）列

2.1 实数空间中的 Cauchy 列

之前在数学分析中就已经探讨过柯西列，所谓 Cauchy 列指对于一个数列 $\{x_n\}$，$\forall\epsilon > 0$，$\exists N$，$m, n > N$，有 $|x_n - x_m| < \epsilon$。

而在**全体有理数**组成的距离空间中，Cauchy 列不一定收敛。而在**全体实数**组成的距离空间中，Cauchy 列一定收敛。比如，一个数列收敛到 π，$\{3,\ 3.1,\ 3.14,\ \cdots\}$，他的极限不是有理数，是无理数。

而空间中的 Cauchy 列一定收敛，则反映了实数空间中非常重要的性质，即为**实数空间的完备性**。我们将把这个性质“类比”的推广到一般的距离空间中，从而得到空间的完备性。从上述实数的例子中可以得出，一个点列 $\{x_n\}$ 是否收敛，除了**点列自身构造性质**以外，和**空间的结构**有很大的关系。比如，上述例子中，首先这个点列最后要非常的密集的靠近一个区域，点列不能是发散的。第二空间要是实数空间，有理数就显然不行，因为最终收敛的极限 π 都不在空间中。

例 2.1 设 $X = (0, 1]$ 赋以实数空间通常的距离，$\frac{1}{n}$ 是 $X = (0, 1]$ 中的 Cauchy 列，但是它在 $X = (0, 1]$ 中不收敛，因为 $0 \notin X$。

上述表明，空间中的 Cauchy 列可能不收敛。问题在于这个空间可能存在“缺陷”，或者说距离空间中有一些“缝隙”。在**例 2.1** 中，问题产生于空间 $X = (0, 1]$ 中缺失了 0 点。如果我们加上这个点，则 $\{\frac{1}{n}\}$ 就收敛了。由此可以证明，在新的“更大的”空间中，$X_1 = (0, 1] \cup \{0\}$ 中每个 Cauchy 列都收敛。

2.2 一般空间中的 Cauchy 列

在实数空间中，点列收敛的充要条件是：**这个点列是 Cauchy 列**。类似于实数空间，在距离空间，我们也引入了 Cauchy 列，完备性这些概念。大家注意到，对于每个概念的定义，我们都是可以找到一个背景的，这样可以让抽象的泛函分析没那么抽象。

定义 2.1 设 (X, d) 是一个距离空间，$\{x_n\}_{n=1}^{\infty} \subset (X, d)$。若对于任意的 $\epsilon > 0$，存在正整数 N，当 $m, n \leq N$ 时，有

$$d(x_n, x_m) < \epsilon \tag{1}$$

称 $\{x_n\}$ 是一个 Cauchy 列。实际上和实数空间中的 Cauchy 的定义基本一样。而性质也有很多是相似的。

命题 2.1 设 $\{X, d\}$ 是距离空间中的 Cauchy 列，则集合 $\{x_1, x_2, \cdots\}$ 是有界的。

这里采用 Cauchy 列的定义证明，和数学分析中的证明方法类似。

证明： 由 Cauchy 列的定义，对于 $\epsilon = 1$，存在 N，当 $n, m > N$ 时，有 $d(x_n, x_m) < 1$，令

$$\beta = \max\{d(x_1, x_2), d(x_1, x_3), \cdots, d(x_1, x_{N+1})\}$$

结合三角不等式可得，对于任何的正整数 n，都有

$$d(x_1, x_n) < \beta + 1 \tag{2}$$

即为，集合 $\{x_1, x_2, \cdots\} \subset B(x_1, \beta + 1)$。根据有界集合的定义，命题成立。

命题 2.2 收敛的点列一定是 Cauchy 列。

我们可以用收敛列和 Cauchy 的定义来证明，与数学分析中的证明思路相似。

证明，设 $\lim_{n\to\infty} x_n = x_0$，则对于 $\forall\epsilon > 0$，$\exists N$ 当 $n, m > N$ 时有，

$$d(x_n, x_0) < \frac{\epsilon}{2}, \quad d(x_m, x_0) < \frac{\epsilon}{2}$$

根据三角不等式有，

$$d(x_n, x_m) < d(x_n, x_0) + d(x_m, x_0) < \epsilon \tag{3}$$

所以，$\{x_n\}$ 是一个 Cauchy 列。而实际上在有限维空间中，列紧集合和有界之间是可以等价的，而在无穷维空间中这个就不成立了，列紧集合可以推出有界，而有界集合不能推出列紧。

3 完备的距离空间

3.1 完备的距离空间的定义

其实大部分数学感觉学起来困难，原因就是对定义的背景和概念理解得不够完善。

收敛的点列一定是 Cauchy 列，但是在一般的空间中 Cauchy 不一定收敛。而 **"所有的 Cauchy 都收敛"，** 这样的距离空间是非常重要的距离空间，同时也被称为完备空间。而完备性为什么会这么重要呢？因为有了完备性，极限运算（微积分）才能很好的进行，在一个完备的距离空间判断点列是否收敛，仅仅只需要判断它是否是 Cauchy 列。

例 3.1 设 $\mathbb{Q}$ 为全体有理数组成的集合，赋以通常的距离成为一个距离空间，但是它不完备。例如：以 π 的前 n 个数字组成的数列，

$$\{3, 3.1, 3.14, 3.141, \cdots\}$$

是一个 Cauchy 列，但是它在 $\mathbb{Q}$ 中不收敛，因为 π 不是有理数。

所以，根据以上例子可得：**"一个点列是不是 Cauchy 列由其本身的性质所决定的，但它是否为 Cauchy 列由这个点列自身的结构所决定的。"** 在例 3.1 中，由于无理数的缺失，点列不收敛。

命题 3.1 完备空间的任何一个闭子空间也是完备的。

证明步骤如下所示，设 X 是完备的，子空间 $X_1 \subset X$，且 X_1 是闭集（闭集的意思即为点列收敛的极限在此集合中，也就是极限运算是封闭的）。

要证明 X_1 是完备的，根据完备空间的定义，只需证明 X_1 中的任意 Cauchy 列都在 X_1 中收敛。设 $\{x_n\}_{n=1}^{\infty} \subset X_1$ 是任意一个 Cauchy 列，由于 X 是完备的，因此 $\{x_n\}$ 收敛到 x，$x \in X$。由于

X_1 是闭集，所以 X_1 中收敛点列的极限 x 属于 X_1，所以 X_1 是完备的。

列紧是泛函分析中非常重要的性质，因为有了这个性质，无穷点列就可以找到收敛的点列，然后极限运算就可以较好的考虑了。

命题 3.2 列紧的空间一定是完备的。

分析：设 $\{x_n\}$ 是列紧空间 X 中的任一 Cauchy 列，我们只要证明它是收敛的，则空间是完备的。**由空间的列紧性可得，**首先找到它的一个收敛的子列，再结合它本身是 Cauchy 列，证明这个 Cauchy 列在空间中收敛。证明思路大体上是这样的。

证明： 设 $\{x_n\}$ 是列紧空间 X 中的任意 Cauchy 列。由 Cauchy 的定义，$\forall \epsilon > 0$，$\exists N$，当 $n, m > N$ 时，

$$d(x_n, x_m) < \epsilon \tag{4}$$

由 X 是列紧的，可知存在 $\{x_k\}$ 的收敛子列 $\{x_{n_k}\}$ 及 $x_0 \in X$，

$$x_{n_k} \to x_0 \ (k \to \infty) \tag{5}$$

既然极限已经存在了，下一步则是证明 x_0 也是 x_n 在空间 X 中的极限。

令 $K = N$，当 $k > N$ 时，有 $n_k \leq k > K = N$，于是有

$$d(x_n, x_{n_k}) < \epsilon \ (n > N) \tag{6}$$

令 $k \to \infty$，由于 d 连续，我们有，

$$d(x_n, x_0) \leq \epsilon \ (n > N) \tag{7}$$

即为 $x_n \to x_0$，$x_0 \in X$，这证明 X 是完备的。从证明中可以看出，对于一个 Cauchy 列 $\{x_n\}$，只要他有一个子列收敛到 x_0，则有 $x_n \to x_0$ $(n \to \infty)$。

命题 3.3 设 X 是一个距离空间，$\{x_n\}$ 是 X 中的 Cauchy 列，如果 $\{x_n\}$ 有一个子列 $\{x_{n_k}\}$ 收敛到 x_0，则 $\{x_n\}$ 也收敛到 x_0 $(n \to \infty)$。

这个命题可以从直观上去理解，Cauchy 的可以从直观上去理解，Cauchy 的尾部会收缩的非常的紧，那么如果有一个子列可以收敛到 x_0 的话，子列收敛点一定会在 Cauchy 列的尾部区间里，而这个区间内的所有点的距离都是小于 ϵ 的，所以 Cauchy 也会收敛到同样的 x_0 点。

3.2 完备与不完备距离空间的例子

3.2.1 $C[a,b]$ 是完备的

$C[a,b]$ 这个空间前面已经讨论了很多次了，第一次给出了什么是 $C[a,b]$ 空间；然后描述了它的性质是怎样的，它的收敛性怎么样；讲了它是不是列紧的等等。为什么要反复分析了，因为我们现在数学研究最重要的还是连续函数。

分析：设 $\{x_n\}_{n=1}^{\infty}$ 是 $C[a,b]$ 中的任一 Cauchy 列。要证明 $C[a,b]$ 是完备的，需要做到以下三点：

1. 找出 $x(t)$ (即 $\{x_n\}$ 的极限)；
2. 证明 $x(t) \in C[a,b]$；
3. 证明 $x_n(t) \to x(t)$ $(n \to \infty)$ (按 $C[a,b]$ 空间中的距离收敛)。

这三点的逻辑是这样的，首先我们需要构造出一个极限，然后要得到这个极限在 $C[a,b]$ 中，最后一步则是证明任一 Cauchy 的极限会收敛到构造出的这个极限。这样就可以满足完备空间中，所有的 Cauchy 都收敛的条件。

证明：

1. 由 $\{x_n(t)\}$ 是 $C[a,b]$ 中的 Cauchy 列。那么对于 $\forall\epsilon>0$，$\exists n,m>N$ 时，

$$d(x_n,x_m)<\epsilon$$

即：

$$\max_{a\le t\le b}|x_n(t)-x_m(t)|<\epsilon \tag{8}$$

注意这里 x_n 代表的是函数列，而 $x_n(t)$ 表示函数 x_n 在 t 点的取值，所以是一个数。所以，$\forall t\in[a,b]$，$|x_n(t)-x_m(t)|<\epsilon$ $(n,m>N)$。而 $\{x_n(t)\}$ 是 $\mathbb{R}$ 空间中的数列，由于 $\mathbb{R}$ 是完备空间，存在 $x(t)$，使得，

$$x_n(t)\to x(t)\ (n\to\infty) \tag{9}$$

2. 下一步则是证明，$x(t)\in C[a,b]$。当 $n,m\ge N$ 时，有

$$|x_n(t)-x_m(t)|<\epsilon,\ \forall t\in[a,b] \tag{10}$$

对于固定的 t，令 $m\to\infty$，有

$$|x_n(t)-x(t)|\le\epsilon\ (n\ge N),\ \forall t\in[a,b] \tag{11}$$

这里的逻辑是，距离函数 $|x_n(t)-x_m(t)|$ 是连续函数，所以可以令 $\lim_{m\to\infty}|x_n(t)-x_m(t)|=|x_n(t)-\lim_{m\to\infty}x_m(t)|$。同样，也可以这样来证明：

$$|x_n(t)-x_m(t)|\le\epsilon\to-\epsilon\le x_n(t)-x_m(t)\le\epsilon\to\lim_{m\to\infty}-\epsilon\le\lim_{m\to\infty}x_n(t)-x_m(t)\le\lim_{m\to\infty}\epsilon$$
$$\le\epsilon x_n(t)-x(t)\le\epsilon\to|x_n(t)-x(t)|\le\epsilon$$

根据上述公式，可以推出 $x_n(t)$ 一致收敛到 $x(t)$，即为 $x(t)$ 是连续的，即为：$x(t)\in C[a,b]$。

3. 最后一步则是证明，按在 $C[a,b]$ 空间中的距离收敛。已证得：

$$|x_n(t)-x(t)|\le\epsilon,\ \forall t\in[a,b] \tag{12}$$

所以，有

$$\max_{a\le t\le b}|x_n(t)-x(t)|\le\epsilon \tag{13}$$

即为 $d(x_n,x)<\epsilon$，$(n\ge N)$，所以 $\lim_{n\to\infty}x_n=x$。

3.3 l^∞ 是完备的

l^∞ 表示全体有界的数列。设 $\{x_n\}$ 是 l^∞ 中的任意一个 Cauchy 列，其中 $x_n=\left\{\xi_k^{(n)}\right\}_{k=1}^\infty$，同样的思路，只需要证明以下三点：

(1) 找出 x (即 $\{x_n\}$ 的极限)；

(2) $x\in l^\infty$；

(3) $x_n\to x$ $(n\to\infty)$ (按照 l^∞ 空间中的距离收敛)。

证明：

1. 设 $\{x_n\}$ 是 l^∞ 中的任意一个 Cauchy 列，其中 $x_n = \left\{\xi_k^{(n)}\right\}_{k=1}^{\infty}$。由 Cauchy 的定义，对于 $\forall\epsilon > 0$，$\exists n, m > N$ 时，

$$d(x_n, x_m) < \epsilon$$

即为：$\sup_k \left|\xi_k^{(n)} - \xi_k^{(m)}\right| < \epsilon$。对任意的 k，$|\xi_k^{(n)} - \xi_k^{(m)}| < \epsilon \ (n, m \geq N)$。注意，这里需要清晰一下符号的意义，$\xi$ 的上标 n 表示第 n 条有界实数列，而下标 k 表示有界实数列中的第 k 个元素。即 $\left\{\xi_k^{(n)}\right\}_{n=1}^{\infty}$ 是 $\mathbb{R}$ 中的 Cauchy 列。由于 $\mathbb{R}$ 的完备性，存在 ξ_k，使得 $\lim_{n\to\infty} \xi_k^{(n)} = \xi_k$。(**这里谈一点个人的理解，这个有界数列极限就是对于一个 Cauchy 列中所有的有界数列，都去取他们的极限 $\xi_1^\infty, \xi_2^\infty, \xi_3^\infty, \cdots$ 得来的。**)

2. 令 $x = \{\xi_k\}$，第二步则是验证 $x \in l^\infty$。方法是一样的，由于，$|\xi_k^{(n)} - \xi_k^{(m)}| < \epsilon \ (n, m \geq N)$，令 $m \to \infty$，有

$$\left|\xi_k^{(n)} - \xi_k\right| \leq \epsilon \ (n \geq N)$$

我们的目标是要证 $\forall \xi_k$ 都有 $|\xi_k| < M$，其中 M 是一个定值。运用一下三角不等式，对 $\forall k$，

$$|\xi_k| = |\xi_k - \xi_k^{(N)} + \xi_k^{(N)}| \leq |\xi_k^{(N)} - \xi_k| + |\xi_k^{(N)}| \leq \epsilon + |\xi_k^{(N)}| \tag{14}$$

由于 $\xi_k^{(N)}$ 是 l^∞ 中的元素，他们是必然有界的，所以 $\{\xi_k\}$ 是有界数列，$\{\xi_k\} \in l^\infty$。而这里我们为什么会想到使用三角不等式呢？因为 ξ_k 有没有界，不知道，而 $\xi_k^{(N)}$ 是必然有界的。通过三角不等式的方式，可以成功的将两者结合起来来，证明 l^∞ 有界。

3. 且 $n \geq N$ 时，对于 $\forall k$，有

$$|\xi_k^{(n)} - \xi_k| \leq \epsilon$$

即为：

$$d(x_n, x) = \sup_k \left|\xi_k^{(n)} - \xi_k\right| \leq \epsilon \ (n \leq N)$$

这就是极限的定义，所以 $x_n \to x = \{\xi_k\}$。所以，l^∞ 是完备的。

3.4 不完备的例子

在距离空间 $C[0,T]$ 中，$P[0,T]$ 记为定义在 $[0,T]$ 上的全体多项式，显然 $P[0,T] \subsetneq C[0,T]$，且 $P[0,T]$ 是 $C[0,T]$ 一个子空间。**距离空间 $C[0,T]$ 是完备的，但是 $P[0,T]$ 在 $C[0,T]$ 中并不完备**。事实上，

$$\left\{1, 1+t, 1+t+\frac{1}{2!}t^2, 1+t+\frac{1}{2!}t^2+\frac{1}{3!}t^3, \cdots\right\}$$

在 $C[0,T]$ 中收敛到 e^t (一致收敛，因为区间是有限的)，但 $e^t \notin P[0,T]$，即 $P[0,T]$ 不是 $C[0,T]$ 中的闭集。而

$$\left\{1, 1+t, 1+t+\frac{1}{2!}t^2, 1+t+\frac{1}{2!}t^2+\frac{1}{3!}t^3, \cdots\right\}$$

一定是 Cauchy 列 (收敛的点列一定都是 Cauchy 列)，但是它极限不在子空间 $P[0,T]$ 中，所以 $P[0,T]$ 在距离：

$$d(p_1, p_2) = \max_{0 \leq t \leq T} |p_1(t) - p_2(t)| \tag{15}$$

下不完备。说明这个空间中有缝隙，收敛出去了。

下面再来看一个例子。连续函数换一个距离也不是完备的。

设 X 是全体在 $[0,1]$ 上定义的连续函数，在 X 上定义距离

$$d(x,y)=\int_0^1 |x(t)-y(t)|dt \tag{16}$$

X 是一个距离空间，但不完备。下面我们构造一个这个空间中的 Cauchy 列，但它在这个空间中不收敛的例子。考虑连续函数列 $\{x_n(t)\}$ $(n>2)$:

$$x_n(t)=\begin{cases}0, & 0\le t\le \frac{1}{2}-\frac{1}{n}\\ 1, & \frac{1}{2}\le t\le 1\\ \text{直线连接}, & \text{else}\end{cases} \tag{17}$$

$\{x_n\}$ 是 X 中的 Cauchy 列，事实上

$$\begin{aligned}d(x_n,x_m)=&\int_0^1 |x_m(t)-x_n(t)|dt\\ =&\frac{1}{2}\left|\frac{1}{n}-\frac{1}{m}\right|\to 0\ (m,n\to\infty)\end{aligned} \tag{18}$$

若空间完备，则存在 X 中的连续函数 $y(t)$，使得 $d(x_n,y)\to 0$，$(n\to\infty)$。考虑

$$x(t)=\begin{cases}0, & 0\le t\le \frac{1}{2}\\ 1, & \frac{1}{2}\le t\le 1\\ \frac{1}{2}, & t=\frac{1}{2}\end{cases} \tag{19}$$

由三角不不等式得：

$$\begin{aligned}&\int_0^1 |x(t)-y(t)|\,dt\\ &\le \int_0^1 |x(t)-x_n(t)|\,dt+\int_0^1 |x_n(t)-y(t)|\,dt\\ &=\frac{1}{2n}+d(x_n,y)\to 0\end{aligned} \tag{20}$$

因此，

$$\int_0^1 |x(t)-y(t)|dt=0 \tag{21}$$

即 $x(t)$ 几乎处处等于 $y(t)$，因为 $x(t)$ 和 $y(t)$ 在 $[0,\frac{1}{2})\cup(\frac{1}{2},1]$ 上连续，两个连续函数几乎处处相等，即为点点相等。于是在 $[0,\frac{1}{2})$ 上 $y(t)=x(t)=0$，在 $(\frac{1}{2},1]$ 上 $y(t)=x(t)=1$，这显然与 $y(t)$ 连续的假设是相违背的。所以此空间是不完备的。这里不是 Cauchy 没有极限，而是空间有缝隙，极限跑出去了。上述的例子中，很好的描述了**组成空间的元素一样，但是空间中定义的距离不一样，空间 X 的完备性不一样**。

4 距离空间的完备化

4.1 距离空间的完备化的定义

从前面的例子中我们看到了有的空间是完备的，有的空间并不完备。但是我们可以将不完备的空间完备化。**事实上任意一个空间都可以完备化，**这是本节要证明的重要结论。

为了方便理解，下面做出一些直观性的解释。设 (X,d) 是完备距离空间，$(X_0,d)\subset(X,d)$ 是一子空间。

(1). 如果 X_0 在 (X,d) 中是闭的 (对极限运算是封闭的)，则 (X,d) 是完备的。

(2). 如果 X_0 不是闭集，我们知道 X_0 在 (X,d) 中是闭的，且 $\widetilde{X}_0$ 在 (X,d) 中是闭的，且是包含 X_0 的最小闭集。因此 $(\widetilde{X}_0,d)$ 是完备的，且 X_0 在 $(\widetilde{X}_0,d)$ 中稠密。

大家可能看起来会觉得,非常难以理解。而通俗的来说,之前在空间中不完备的原因是有的 Cauchy 列不在空间中收敛，就是因为空间中有缝隙。从 (X_0,d) 到 $(\widetilde{X}_0,d)$ 的过程，就是**填满**了原来 (X_0,d) 中存在的“缝隙”，使之成为一个完备空间。

举个例子：有理数全体组成的空间 $\mathbb{Q}$ 是不完备的，即：存在有理数组成的 Cauchy 列，收敛的极限不是有理数。而实数空间 $\mathbb{R}$ 是一个完备的距离空间。而我们可以通过“做闭包”的方法，把 $\mathbb{Q}$ 扩展为完备的实数空间 $\mathbb{R}$，或者是说把 $\mathbb{Q}$ 嵌入到另一个完备空间 $\mathbb{R}$ 中，也就是将 $\mathbb{Q}$ 作为 $\widetilde{\mathbb{Q}}=\mathbb{R}$ 这时就空间就完备了。比如有理数列 $\{3,3.1,3.14,3.141,\cdots\}$，最终收敛到 π，并不是一个有理数，那我们就把这个数补进去。

这意味着：

1) $\mathbb{Q}$ 中的元素的距离不变（等距离嵌入）;

2) $\mathbb{Q}\subset\mathbb{R}$，$\mathbb{Q}$ 在 $\mathbb{R}$ 中稠密;

3) $\mathbb{R}$ 是 $\mathbb{Q}$ 的完备化空间。

一般的距离空间 (X,d)，如果 X 不完备，我们可以使用类似的方法，把 X 嵌入到一个完备的距离空间中，或者说把 X 扩充进一些元素，使之完备。

加进去的“点”就是 X 闭包中不属于 X 的那些点。这是我们下面要证明的距离空间完备化定理的基本含义,。

最大的问题在于如何用 X 的元素来刻画新加进去的点，这些点原来是没有的。比如，无理数的粗略说法是无限不循环小数，说实话这是无法验证的，因为算都算不完，怎么知道从当前计算的下一位开始会不会循环起来。而我们现在只有有理数，没有无理数。那么想要将空间完备化的话，就需要用已有的有理数来刻画无理数，问题就是怎么用有理数来定义出无理数。

而空间完备化的定义直观的理解还比较简单，而在数学书上用规范的语言来描述就变得很复杂了。这也是数学学习的难点,所以大家在数学学习的过程中,一定要注意定义的背景和它所解决的问题,而不是死记硬背定义本身。

所谓距离空间的完备化，是将不完备的空间 $\widetilde{X}$，定义 $\widetilde{\widetilde{X}}$ 是完备的，其 $\widetilde{X}\subsetneq\widetilde{\widetilde{X}}$，并且用 $\widetilde{X}$ 定义出 $\widetilde{\widetilde{X}}$ 中的元素。

那么，我们怎么用当前的元素来刻画新的元素呢？从 Cauchy 列入手，**这个逻辑是这样的。因为距离空间不完备本就是由其中的 Cauchy 列不收敛造成了，那么将这些 Cauchy 列的极限补进去就好了**。

4.2 距离空间的完备化

定义 4.1 任何距离空间 (X,d)，都存在一个完备的距离空间 $(\widetilde{X}.\widetilde{d})$，使得 (X,d) 和 $(\widetilde{X}.\widetilde{d})$ 的一个子空间等距，且在等距的意义下，这样的空间 $(\widetilde{X}.\widetilde{d})$ 是唯一的。称 $(\widetilde{X}.\widetilde{d})$ 为 (X,d) 的完备化空间。

这个定义，现在看起来非常的抽象，我们同样可以用有理数和实数的例子来理解它。大家一定要记住实数这个例子，有了这个例子可以让大家头脑中有个清晰的认识，对概念有更好的理解。有理数空间并不完备，通过将缝隙填满以后，可以完备成一维实数空间。那么可不可以完备成二维的平面呢？不行的，太大了导致有理数在其中不是稠密的，所以稠密等距就是来限制完备化空间的大小的。

分析：证明可以分为四步：

(1) 先构造空间 $\widetilde{X}$ 和距离 $\widetilde{d}$；

(2) 证明 (X,d) 与 $(\widetilde{X},\widetilde{d})$ 中的一个稠子集等距；

(3) $(\widetilde{X},\widetilde{d})$ 完备，这个空间 $\widetilde{X}$ 就是我们需要的空间；

(4) 在等距的意义下，完备化空间是唯一的。

第一步是非常困难的，我们之前就分析了如何用当前的元素来构造未知的元素是一个难点，之前想到了用 Cauchy 列作为媒介来解决。下面来看有理数中具体的例子，现在有两个数列：

$$\begin{aligned}&(3, 3.1, 3.14, \cdots, \pi)\\&(1, 1.4, 1.41, \cdots, \sqrt{2})\end{aligned} \tag{22}$$

显然这两个数列收敛的极限都不属于有理数空间，那么我们令 $x_n = (3, 3.1, 3.14, \cdots) = \pi$，$y_n = (1, 1.4, 1.41, \cdots) = \sqrt{2}$，记为：

$$\lim_{n\to\infty} x_n = \pi,\ \lim_{n\to\infty} y_n = \sqrt{2} \tag{23}$$

在等距的要求下，显然有：

$$\lim_{n\to\infty} d(x_n, y_n) = d(\pi, \sqrt{2}) = \left|\pi - \sqrt{2}\right| \tag{24}$$

显然这样我们就定义出了完备化空间 $\widetilde{X}$，用规范化的语言描述即为：对于当前空间 X 中任意一个 Cauchy 列，$\{x_n\} \to \widetilde{x}$，$\{y_n\} \to \widetilde{y}$，距离为 $\widetilde{d}(\widetilde{x}, \widetilde{y})$，(注意，新空间 $\widetilde{X}$ 中的元素都是 Cauchy 列。)。这样就构造空间 $\widetilde{X}$ 和距离 $\widetilde{d}$，下一步就理所当然的要证明 (X,d) 与 $(\widetilde{X},\widetilde{d})$ 中的一个稠子集等距了。有了这样一个具体的例子之后，可以让抽象的概念变得没那么复杂，所以数学学习中一定要注意问题提出的背景。

证明：

1. 构造距离空间 $(\widetilde{X},\widetilde{d})$；

   (a) 构造集合 $\widetilde{X}$；

   把 (X,d) 中的 Cauchy 列 $\widetilde{x} = \{x_n\}$，$\widetilde{y} = \{y_n\}$，满足：

   $$\lim_{n\to\infty} d(x_0, y_0) = 0 \tag{25}$$

   则称他们为 $\widetilde{X}$ 中的同一元素，即为 $\widetilde{x} = \widetilde{y}$。为什么要这样定义呢？比如趋近 π 的数列有很多个比如：$\{4, 3.5, 3.2, 3.15, \cdots\}$，如果没有这个等价的定义，数列太多了，会乱套。

(b) 定义距离

对任何 $\widetilde{x} = \{x_n\}$, $\widetilde{y} = \{y_n\} \in \widetilde{X}$ 定义

$$\widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} d(x_n, y_n) \tag{26}$$

这样距离就定义完了，下一步则是考虑距离的合理性，它是距离吗？

(c) 验证距离的合理性。$\{d(x_n, y_n)\}$ 要是一个数。

由于 $\{x_n\}, \{y_n\}$ 是 X 中的 Cauchy 列，

对于 $\forall \epsilon > 0$, $\exists N$, 当 $m, n > N$ 时，

$$d(x_n, x_m) < \frac{\epsilon}{2}, \text{ 且 } d(y_n, y_m) < \frac{\epsilon}{2} \tag{27}$$

由三角不等式可得：

$$|d(x_n, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_n, y_m) < \epsilon \tag{28}$$

由此可得 $\{d(x_n, y_n)\}$ 是实数空间 $\mathbb{R}$ 中的一个 Cauchy 数列。

(d) 证明 $\widetilde{X}$ 中定义的 $\widetilde{d}$ 与 $\widetilde{x}, \widetilde{y}$ 想对应的 Cauchy 列的选择无关。并且容易验证 $\widetilde{d}$ 的非负，恒正，对称等性质。

2. 验证 (X, d) 与 $\widetilde{X}, \widetilde{d}$ 中的一个稠子空间等距。

(a) 构造稠子空间，等距映射

设 $\widetilde{X}_0$ 是全体由 X 中元素作成的常驻列 $\{x\}$。什么意思呢？回到之前有理数的例子，新的空间 $\widetilde{X}$ 中的所有元素都是 Cauchy 列，那么就将原来的有理数元素都变成常数 Cauchy 列，比如：$3 = \{3, 3, 3, 3, \cdots\}$, $4 = \{4, 4, 4, 4, \cdots\}$。这样在新的距离 $\widetilde{d}$ 下，两个元素间的距离也是一样的。

即为 $\widetilde{X}_0 \subset \widetilde{X}$, $\widetilde{X}_0$ 是 $\widetilde{X}$ 的一个子空间。令 $T : (X, d) \to (\widetilde{X}, \widetilde{d})$, $x \in X$, $T(x) = (x, x, x, \cdots)$。

(b) 验证等距映射。

显然 $Tx \in \widetilde{X}$，任何 $x, y \in X$，$\widetilde{x} = (x, x, x, \cdots)$, $\widetilde{y} = (y, y, y, \cdots)$,

$$\widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} = d(x, y) \tag{29}$$

即为：$T : X \to \widetilde{X}$ 是等距映射。

(c) 证明 $TX = \widetilde{X}_0$ 在 $\widetilde{X}$ 中稠密。

对于 $\forall \widetilde{x} = \{x_n\} \in \widetilde{X}$，令 $\widetilde{x}_k = (x_k, x_k, \cdots, x_k, \cdots)$。显然有 $\widetilde{x}_k \in \widetilde{X}_0$ 是 X 中的 Cauchy 列，存在 N，当 $k, n > N$ 时，$d(x_n, x_k) < \epsilon$。即为：

$$\widetilde{d}(\widetilde{x}, \widetilde{x}_k) = \lim_{n\to\infty} d(x_n, x_k) \leq \epsilon \ (k \geq N) \tag{30}$$

所以 $(\widetilde{X}_0, \widetilde{d})$ 在 $\widetilde{X}, \widetilde{d}$ 中稠密，即为 TX 在 $\widetilde{X}$ 中稠密。核心就是在 $\widetilde{X}$ 中找到一个数列来逼近 $\widetilde{X}_0$ 中的数列。

3. 第三步则是证明，新得到的空间 $\widetilde{X}$ 是完备的。

4. 第四步证明，在等距意义下完备化空间是唯一的。

数学有的时候定义的东西很简单，但是由于符号的表达，概念的证明等等一套很复杂的东西交织在一起会让大家觉得很难。只有知道了问题的背景，才知道问题是怎么来的，才能很好的理解数学抽象的概念。上述过程结合有理数的完备化这一具体的背景，可以帮助大家更好的理解空间完备化的知识点。

4.3 有理数空间完备化详细流程

$\pi, \sqrt{2}$ 对应的 Cauchy 数列“代表元”可以是，

$$\begin{aligned} &(3, 3.1, 3.14, \cdots) \\ &(1, 1.4, 1.41, \cdots) \end{aligned} \tag{31}$$

$(\widetilde{X}, \widetilde{d})$ 就是由这样的一些 Cauchy 列组成的 ($\widetilde{X}$ 实际上就是全体实数) 距离的定义方式：

$$d(\{3, 3, 1, 3.14, 3, 141, \cdots\}, \{1, 1.4, 1.41, 1.414, \cdots\})$$

是数列

$$\{3-1, 3.1-1.4, 3.14-1.41, 3.141-1.414, \cdots\}$$

的极限，等于 $\pi - \sqrt{2}$。

$$\begin{aligned} &\{3, 3, \cdots\}, \{3.1, 3.1, \cdots\}, \{3.14, 3.14, \cdots\}, \cdots \\ &\{1, 1, \cdots\}, [1, 4, 1, 4, \cdots\}, \{1.41, 1.41, \cdots\}, \cdots \end{aligned}$$

这个是 $(\widetilde{X}_0, \widetilde{d})$ 中与其相关的常值 Cauchy 列 ($\widetilde{X}_0$ 实际上是全体有理数)。

从形式上看，完备化的距离空间 $\widetilde{X}$ 是个全新的空间，但应注意完备化的距离空间 $\widetilde{X}$ 包含了一个稠子空间 $\widetilde{X}_0$ 与原来的空间 (X, d) 等距同构，也就是说 X 嵌入到 $\widetilde{X}$ 中，作为它的一个稠子集。

注：完备化以后的空间是原来空间的一种“闭扩张”，不能仅仅理解为是一个由 Cauchy 列组成的一个十分抽象的的新空间。就像有理数的例子一样，我们之所以把其中的元素都变成 Cauchy 列，是我们用有理数不好去描述无理数。

设 X 是一个不完备的距离空间，X_i 是一个完备的距离空间，如果满足:

1. $X \subset X_1$，

2. X 在 X_1 中稠密，

那么，X_1 是 X 的完备化空间。例如实数空间就是有理数空间的完备化空间，L^2 空间 (平方可积函数空间) 就是全体连续函数在距离

$$d(x, y) = \left\{ \int_a^b |x(t) - y(t)|^2 dt \right\}^{1/2} \tag{32}$$

下的完备化空间。

4.4 空间完备化的意义

距离空间完备化后，空间中的 Cauchy 列都收敛。从另一个角度说，空间 X 被适度地扩大为 $\widetilde{X}$。原来的“缝隙”已经被全部填满。这点是十分重要的。以后我们会看到，这使得一些在原空间 X 中无

解的问题 (例如微分方程)，在新的扩大了的空间 $\widetilde{X}$ 中就可以有“较弱”意义下的解，运算就可以进行了，这也是完备化的意义所在。比如我们高中学习的 $x^2=-1$ 在实数域上无解，但是我们扩充到复数域上，就有解了。

5 小结

本节主要介绍的是完备的距离空间等相关的知识。首先介绍了最重要的 Cauchy 列，因为空间完备的定义就是空间中所有的 Cauchy 列都收敛。其中讲到了点列是否收敛和点列自身的构造性质和空间的结构都有关。

然后介绍了完备空间的定义，举了连续函数空间是完备的，全体有界数列空间是完备的，和一个不完备的例子来帮助大家理解空间完备的概念。用直观的感觉去理解，即为完备的空间是没有缝隙的，所有的 Cauchy 列都在此空间中收敛。

最后一部分介绍的是空间的完备化，这使得一些在原空间 X 中无解的问题 (例如微分方程)，在新的扩大了的空间 $\widetilde{X}$ 中就可以有“较弱”意义下的解，运算就可以进行了，这也是完备化的意义所在。全程我们都引入了有理数空间完备成实数空间的例子，通过这个例子来帮助我们理解空间完备化的过程。这个例子非常的重要，数学很多时候太难了，实际上就是太抽象了，而直观的理解并不困难。所谓完备化，即为将其中 Cauchy 列收敛的极限给补上，即为构造一个新的空间，使之与原空间形成等距映射，而且原空间是新空间的等距稠子集。其中较难解决的问题是如何用当前的元素来构造未知的元素是一个难点，可以用 Cauchy 列作为媒介来解决。将原空间中的元素映射为常驻列，未知的元素用原空间的元素组成的 Cauchy 表示。逻辑比较清晰!

完备的距离空间的性质和应用

Chen Gong

15 January 2021

目录

1 Introduction

上一节讲到了，对于不完备的空间，我们可以想办法让它变得完备。那么为什么要在完备的空间下研究问题呢？有什么样的好处吗？这一节，将介绍完备空间的性质。让大家深刻理解，为什么要在完备的空间下研究问题呢？

2 闭球套定理

在距离空间中引入了开集，闭集的概率后，我们研究了空间中序列的收敛性，Cauchy 列，讨论了空间的列紧性，可分性，完备性。

在完备空间中，类似于数学分析中的区间套定理，也有闭球套定理。所谓，区间套定理，直观的来说就是一个区间套一个区间，被套住的区间大小比之前要小，这样不停地套下去，最终会收敛到一个点，而且这个点是唯一的。

~

定理 2．1 X 是完备的距离空间，$\bar{S}_n = \bar{S}(x_n, r_n)(n = 1, 2, \cdots)$ 是 X 中的一系列闭球集：

$$\bar{S}_1 \supset \bar{S}_2 \supset \cdots \supset \bar{S}_n \supset \cdots$$

且 $r_n \to 0(n \to \infty)$，则存在 X 中唯一的一点，

$$x \in \bigcap_{n=1}^{\infty} \bar{S}_n$$

此定理的证明过程，主要分三个步骤，1．找到这样的 x；2. 证明 $x \in \bigcap_{n=1}^{\infty} \bar{S}_n$； 3．证明其唯一性。

证明：

1. 设 $\{x_n\}$ 是球心组成的点列，所以，有

$$d(x_n, x_m) < r_n \ (m > n)$$

这是由于第 m 个球，一定在第 n 个球里面。且，当 $r_n \to 0 \ (n \to \infty)$。所以，当 $n, m > N$ 时

$$d(x_n, x_m) < r_n < \epsilon \ (m > n > N)$$

根据 Cauchy 的定义，可得 $\{x_n\}$ 是 Cauchy 列。又因为X 是完备的，所以，$\{x_n\}$ 可以收敛到一点，使得 $\lim_{n\to\infty} x_n = x$。

2．由于 $d(x_n, x_m) < r_n \ (m > n)$，根据距离的连续性，令 $m \to \infty$，根据第一步推出的结果，有 $d(x_n, x) \le r_n$，所以有 $x \in \bar{S}(x_n, r_n)$。且，$\bar{S}_n$ 是闭球集，那么点 x 肯定包含在所有的球中。于是我们有 $x \in \bigcap_{n=1}^{\infty} \bar{S}_n$。

3．第三步则是证明此点唯一。如果，存在 $y \in \bigcap_{n=1}^{\infty} \bar{S}_n$。那么，对于任意的 x_n，都有

$$d(x_n, y) \le r_n,$$

令 $n \to \infty$，有 $d(x, y) \le 0$，由于距离的连续性，所以有 $x = y$。

3 压缩映射原理

终于系统的看到了压缩映射原理了，小编研究的是强化学习，压缩映射原理是非常重要的。因为，我们设计一种算法，往往需要迭代的方法来进行求解。并且，需要证明每一次迭代，数值解和真实解之间的距离都会更近，而且最终会收敛到一个不动点，这是算法的理论基础，证明不动点的存在非常重要，因为只有证明了这个点的存在，我们才能采用一些办法来逼近它。而不动点问题，是数学研究中的重要问题之一。

3.1 压缩映射定理的基本背景

本节主要介绍压缩映射原理，所谓一个映射 T 的不动点是指，T把这个点映射为自身，即有 $\mathrm{T}x = x$。实际上，机器学习中非常重要的一类方法即为迭代法解方程。而任何解方程的问题都可以转化为求不动点的问题：

$$F(x) = 0 \Longleftrightarrow F(x) + x = x$$

那么，可以令 $F_1(x) = F(x) + x = x$，令 $\mathrm{T}(x) = F_1(x)$。则解方程的问题转化为求不动点 $x : \mathrm{T}x = x$。因而研究不动点理论及其应用具有重要的理论意义及应用价值。我们数学上最重要的一类问题就是解方程的问题，不论是微分方程还是积分方程等。

~

例 3 .1 在实数范围内求解方程 $y = x^2 - 2x + 1 = 0$，令

$$\mathrm{T}x = x^2 - x + 1,$$

则求解一元二次方程的问题转化为了：什么时候 $\mathrm{T}x = x, x \in \mathbb{R}$，即为：映射 T 有没有不动点。

在代数方程、微分方程、积分方程及其它各类方程理论中解的存在性，唯一性以及近似解的收敛性都是很重要的课题，在许多关于存在唯一性的定理的证明中，"不动点" 是一个有力的工具。可以说，不动点定理是泛函分析中最基本的一个存在性定理。分析中的许多存在性定理都是不动点定理的特例。不动点理论已发展成非线性泛函分析的重要内容之一。

3.2 压缩映射定理起源

多项式根的近似计算最显著的技巧是逐次迭代法，比如 Newton 迭代求根法，

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},\ k = 0, 1, 2, \cdots$$

如果 $\lim_{n\to\infty} x_n = x^*$，我们有 $f(x^*) = 0$。这实际就是压缩映射的一种简单表现形式。而 $\mathrm{T}x = x_k - \frac{f(x_k)}{f'(x_k)}$。

首先将这个技巧应用于无穷维情形的是一个法国数学家 Liourville，他成功地利用这个技巧求解常微分方程初值问题，1922 年，Banach 把这个结果抽象化、用距离空间及压缩映射 (压缩映射是一种特殊的非线性映射) 等概念更一般地描述这个方法，这就是著名的 Banach 不动点定理，或压缩映像原理。

考虑微分方程的初值问题：

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x(t), t) \\ x|_{t=0} = x_0 \end{cases} \tag{1}$$

对公式左右两边 $[0,\tau]$ 进行积分有，

$$x(t)-x(0)=\int_0^t f(x(\tau),\tau)d\tau$$
$$x(t)=x_0+\int_0^t f(x(\tau),\tau)d\tau$$

其中，$\int_0^t f(x(\tau),\tau)d\tau$ 是关于 $x(t)$ 的函数，记为 $\bar{x}(t)$。那么从泛函分析的角度来看，可以看成一个映射的问题，令

$$\mathrm{T}x=x_0+\int_0^t f(x(\tau),\tau)d\tau \tag{2}$$

其中，x 是一个函数 $x(t)$，经过映射 T 之后，得到一个新的函数 $\bar{x}(t)$，T 是函数到函数的映射。那么微分方程就转化成立这个积分算子 $\mathrm{T}x$ 是否有不动点的问题，即在空间 X 是否存在元素 x，满足 $\mathrm{T}x=x$。如果满足压缩映射定理，就有不动点。

3.3 压缩映射原理，Banach 不动点定理

设 (X,d) 是完备的距离空间，$\mathrm{T}:X\to X$。如果对于任意的 $x,y\in X$，不等式

$$d(\mathrm{T}x,\mathrm{T}y)\le\theta d(x,y) \tag{3}$$

成立，其中 $0<\theta<1$，则存在唯一的 $\bar{x}\in X$，使得

$$\mathrm{T}\bar{x}=\bar{x} \tag{4}$$

分析，首先找到 T 的不动点，然后证明其唯一性。由公式 (3) 可以看到，T 作用后两点间的距离成比例地压缩，是一压缩映射，希望用迭代法找到不动点。

任取 $x_0\in X$，令

$$x_1=\mathrm{T}x_0,\ x_2=\mathrm{T}x_1,\cdots,\ x_{n+1}=\mathrm{T}x_n,\ \cdots$$

需要证明

1. $x_n\to\bar{x}\ (n\to\infty)$,
2. T 连续,

则可以推出 $\bar{x}=\mathrm{T}\bar{x}$。而为什么需要这两个条件呢?我们最终希望得到 $x_{n+1}=\mathrm{T}x_n$，即有 $\lim_{n\to\infty}x_{n+1}=\lim_{n\to\infty}\mathrm{T}x_n$，推出 $\bar{x}=\lim_{n\to\infty}\mathrm{T}x_n$。右边，我们想得到，$\lim_{n\to\infty}\mathrm{T}x_n=\mathrm{T}\lim_{n\to\infty}x_n=\bar{x}$。而这里要想交换 T 和 lim 运算，则需要保证 T 是连续的。

要想证明 1，由于空间是完备的，要证收敛，只需要证明 $\{x_n\}$ 是 Cauchy 列即可。对于证明 2，因为 T 是压缩映射，由连续映射的定义可知 T 是连续的。

1 . T 是连续的 (一致连续的)。

事实上，$\forall\epsilon>0$，取 $\delta=\epsilon>0$，当 $d(x,y)<\delta$ 时，

$$d(\mathrm{T}x,\mathrm{T}y)\le\theta d(x,y)<\delta=\epsilon$$

2. 用迭代法求 x，

任取 $x_0\in X$，令

$$x_1=\mathrm{T}x_0,\ x_2=\mathrm{T}x_1,\cdots,\ x_{n+1}=\mathrm{T}x_n,\ \cdots$$

下面的目标则是证明 $\{x_n\}$ 为 Cauchy 列，由于对任意的自然数 p，

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+p-1}, x_{n+p}) \\ &\leq \theta^n d(x_0, \mathrm{T}x_0) + \theta^{n+1} d(x_0, \mathrm{T}x_0) + \cdots + \theta^{n+p-1} d(x_0, \mathrm{T}x_0) \\ &< \frac{\theta^n}{1-\theta} d(x_0, \mathrm{T}x_0) \end{aligned} \tag{5}$$

因为 $0 < \theta < 1$，所有当 $n \to \infty$ 时，必然有 $\frac{\theta^n}{1-\theta} d(x_0, \mathrm{T}x_0) \to 0$。所以，$\{x_n\}$ 是 Cauchy 列。而由于 (X, d) 是完备的，所有存在 $\bar{x}$，使得 $x_n \to \bar{x}(n \to \infty)$，由于 T 是连续的，所以有，

$$\lim_{n\to\infty} \mathrm{T}x_n = \mathrm{T} \lim_{n\to\infty} x_n = \bar{x},$$

由迭代公式，$x_{n+1} = \mathrm{T}x_n$，我们得到：$\bar{x} = \mathrm{T}\bar{x}$。

3 . 唯一性：若存在 $\bar{y}$ 使得，$\mathrm{T}\bar{y} = \bar{y}$，则

$$d(\bar{x}, \bar{y}) = d(\mathrm{T}\bar{x}, \mathrm{T}\bar{y}) \leq \theta d(\bar{x}, \bar{y})$$

由于 $\theta \in (0, 1)$，所以只有可能 $d(\bar{x}, \bar{y}) = 0$，故 $\bar{x} = \bar{y}$。

~

存在这样一种情况，开始不是压缩映射，后来就变成压缩映射了。也就是映射做一次并不是压缩映射，而做了多次以后，变成了压缩映射。

设 (X, d) 是完备的距离空间，T 是 X 到 X 的映射，如果存在正整数 n_0，使得对所有的 $x, y \in X$，

$$d(\mathrm{T}^{n_0}x, \mathrm{T}^{n_0}x) \leq \theta d(x, y) \tag{6}$$

其中，$\theta \in (0, 1)$，则 T 有唯一的不动点。

分折：由 (6) 式，我们看到 T^{n_0} 满足不动点定理的条件，存在唯一不动点 $\bar{x}$。要证 T 有唯一不动点，考虑 T^{n_0} 的不动点是否就是 T 的不动点?

证明：

因为 T^{n_0} 满足不动点定理，顾存在唯一的 $\bar{x}$，使得 $\mathrm{T}^{n_0}\bar{x} = \bar{x}$。因为，

$$\mathrm{T}^{n_0}(\mathrm{T}\bar{x}) = \mathrm{T}(\mathrm{T}^{n_0}\bar{x}) = \mathrm{T}\bar{x}$$

所以，观察第一项和第三项可得，$\mathrm{T}\bar{x}$ 也是 T^{n_0} 的不动点。又因为 T^{n_0} 的不动点是唯一的，所以，

$$\mathrm{T}\bar{x} = \bar{x}.$$

所以，$\bar{x}$ 是 T 的不动点。

唯一性：反证法（小编发现，证明唯一性，非常喜欢用反证法！）

假设，x_1 也是 T 的不动点，则有，

$$\mathrm{T}^{n_0}\bar{x}_1 = \mathrm{T}^{n_0-1}(\mathrm{T}\bar{x}_1) = \mathrm{T}^{n_0-1}(\bar{x}_1) = \cdots = \bar{x}_1$$

所以，$\bar{x}_1$ 也是 T^{n_0-1} 的不动点，由于 T^{n_0-1} 不动点的唯一性，我们有 $\bar{x} = \bar{x}_1$。

4 不动点定理

满足压缩映射的是不动点，而满足压缩映射的情况有很多，从而就衍生出了一系列的不动点定理。比较出名的有以下几个，

定理 4 .1(Brouwer) 设 B 是 $\mathbb{R}^n$ 中的闭单位球。设 $\mathrm{T}: B \to B$ 是一个连续映射，则 T 必有一个不动点 $x \in B$。

例 4 .1 设 $f(x)$ 是定义在 $[-1,1]$ 上的连续函数，且其值域包含 $[-1,1]$ 中，则存在 $\bar{x} \in [-1,1]$，使得：$f(\bar{x}) = \bar{x}$。

定理 4 .2(Schauder) 设 C 是线性赋范空间 X 中的一个闭凸子集，$\mathrm{T}: C \to C$，连续且 $T(C)$ 列紧，则 T 在 C 上必有一个不动点。

5 压缩映射原理的应用

例 5.1

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x(t), t) \\ x|_{t=0} = x_0 \end{cases} \tag{7}$$

其中 $f(x,t)$ 在平面上连续，且对于变量 x 满足 Lipschitz 条件：

$$|f(x_1,t) - f(x_2,t)| \leq K|x_1 - x_2| \tag{8}$$

这个条件非常的好，把右边的 $|x_1 - x_2|$ 除过来有，

$$\frac{f(x_1,t) - f(x_2,t)|}{|x_1 - x_2|} \leq K \tag{9}$$

对于二元函数 $f(x,t)$，可以保证 x 是连续的，这个很好判断，但是并不能保证可导。所以，Lipschitz 是介于偏导数存在和连续之间的一个条件。如果满足此条件，则有方程 (9) 在 $t=0$ 的某个邻域中有唯一解。

分析，和 3.2 中的做法一样，把微分运算换成积分运算，大家都知道，微分运算和积分运算是互为逆运算的。

$$\mathrm{T}x = x_0 + \int_0^t f(x(\tau), \tau) d\tau \tag{10}$$

我们先证明这个积分运算是压缩映射，然后利用不动点存在定理得出方程有唯一解。那么如何证明积分运算是压缩映射呢？压缩映射有以下几个条件，1. X 要是个完备的距离空间；2. 满足压缩映射的条件； 3 . 得到条件，一定存在不动点。所以，我们首先要意识到，我们在哪个空间上研究这个问题。

证明：

1. 确立距离空间，建立映射；

考虑连续距离空间 $C \in [-\delta, \delta]$ 上的如下映射 (积分算子)：

$$\mathrm{T}x = x_0 + \int_0^t f(x(\tau), \tau) d\tau$$

则 T 是从 $C \in [-\delta, \delta]$ 到 $C \in [-\delta, \delta]$ 自身的映射。

2. 验证映射满足不动点定理的条件。

$$
\begin{aligned}
d(\mathrm{T}x,\mathrm{T}y)=&\max_{-\delta\le t\le\delta}\left|\int_0^t|f(x(\tau),\tau)-f(y(\tau),\tau)|d\tau\right| \\
\le&K\max_{-\delta\le t\le\delta}\left|\int_0^t|x(\tau)-y(\tau)d\tau\right| \\
\le&K\delta\max_{-\delta\le t\le\delta}|x(t)-y(t)| \\
=&K\delta d(x,y)
\end{aligned}
\tag{11}
$$

我们肯定可以找到一个 δ 令 $0<K\delta<1$ 的，这个可以放心，且 $C[-\delta,\delta]$ 是完备的。由压缩映射定理可得，方程必有唯一解。而为什么我们要将微分运算变成积分运算呢?

这里老师提出了一种直观的理解方法，我们知道积分运算和微分运算是互逆的。比如，a 和 $\frac{1}{a}$ 是互逆的，那么当 a 越来越大，则 $\frac{1}{a}$ 则会越来越小。所以类比过来，积分运算每一次运算会导致距离会越来越小，而微分运算每一次运算会导致距离会越来越大。那么，在微分运算中，其实达不到压缩映射的要求，而在积分运算中可以证明是压缩映射。那么，可以证明积分方程可以收敛到唯一的解，所以反向得出，微分方程也可以收敛到唯一的解。而我们已经证明了，积分运算是压缩映射，而微分运算是积分运算的逆运算，所以理论上，微分方程是肯定做不到这一点的。所以，数学中经常将微分方程转化为积分方程。

~

例 5.2 考虑线性方程组

$$
\xi_i-\sum_{j=1}^n a_{ij}\xi_j=b_i\quad(i=1,2,\cdots,n)
\tag{12}
$$

其中，

$$
\sum_{i=1}^n\sum_{j=1}^n|a_{ij}|^2<1
\tag{13}
$$

证明，此方程组有唯一解。

分析：考虑将方程组转化为映射的不动点问题。假定 $x=\xi_1,\xi_2,\cdots,\xi_n$，并定义 $\mathrm{T}x$，使得其第 i 个分量的作用形式为 $\sum_{j=1}^n a_{ij}\xi_j+b_i$，则方程组可转化为 $\mathbb{R}^n$ 空间上 $\mathrm{T}x=x$ 的不动点问题。

证明：

1．建立映射：设，

$$
(\mathrm{T}x)_i=\sum_{j=1}^n a_{ij}\xi_j+b_i,\quad(i=1,2,3,\cdots).
\tag{14}
$$

$x=\{\xi_1,\xi_2,\cdots,\xi_n\}$，则 T 是 $\mathbb{R}^n\to\mathbb{R}^n$ 的一个映射。

2. 验证映射满足不动点定理条件。由于，

$$
\begin{aligned}
d\left(Tx_1,Tx_2\right)&=\left\{\sum_{i=1}^n\left[\left(\sum_{j=1}^n a_{ij}\xi_j^{(1)}+b_i\right)-\left(\sum_{j=1}^n a_{ij}\xi_j^{(2)}+b_i\right)\right]^2\right\}^{\frac{1}{2}} \\
&=\left\{\sum_{i=1}^n\left[\sum_{j=1}^n a_{ij}\left(\xi_j^{(1)}-\xi_j^{(2)}\right)\right]^2\right\}^{\frac{1}{2}}
\end{aligned}
\tag{15}
$$

而，

$$\left\{\sum_{i=1}^{n}\left[\sum_{j=1}^{n}a_{ij}\left(\xi_j^{(1)}\quad-\xi_j^{(2)}\right)\right]^2\right\}^{\frac{1}{2}} \leq \left\{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2\sum_{j=1}^{n}\left|\xi_j^{(1)}-\xi_j^{(2)}\right|^2\right\}^{\frac{1}{2}}$$

$$=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}\left(\sum_{j=1}^{n}\left|\xi_j^{(1)}-\xi_j^{(2)}\right|^2\right)^{\frac{1}{2}} \tag{16}$$

$$=\theta d\left(x_1,x_2\right)$$

其中，

$$\theta=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}<1 \tag{17}$$

根据压缩映射原理，方程组有唯一解。

~

例 5.3 在上例中，若将条件改为：

$$\alpha=\max_{1\leq i\leq n}\sum_{j=1}^{n}\left|a_{ij}\right|<1 \tag{18}$$

通过这一系列例子，我们希望看到不同的空间中，对于压缩映射定理不同的处理方法。其解决问题的思路和方法同上。事实上，如果 n 维向量空间的距离定义为：

$$d(x,y)=\max_{1\leq i\leq n}\left|\xi_i-\eta_i\right| \tag{19}$$

其中，$x=(\xi_1,\xi_2,\cdots,\xi_n)$，$y=(\eta_1,\eta_2,\cdots,\eta_n)$。

1．建立映射：设，

$$(\mathrm{T}x)_i=\sum_{j=1}^{n}a_{ij}\xi_j+b_i,\quad(i=1,2,3,\cdots). \tag{20}$$

$x=\{\xi_1,\xi_2,\cdots,\xi_n\}$，则 T 是 $\mathbb{R}^n\to\mathbb{R}^n$ 的一个映射。

2.

$$d\left(\mathrm{T}x_1,\mathrm{T}x_2\right)=\max_{1\leq i\leq n}\left|\left(\sum_{j=1}a_{ij}\xi_j^{(1)}+b_i\right)-\left(\sum_{j=1}a_{ij}\xi_j^{(2)}+b_i\right)\right|$$

$$=\max\nolimits_{1\leq i\leq n}\left|\sum_{j=1}^{n}a_{ij}\left(\xi_j^{(1)}-\varepsilon_j^{(2)}\right)\right| \tag{21}$$

$$=\max\nolimits_{1\leq j\leq n}\left|\xi_j^{(1)}-\xi_j^{(2)}\right|\cdot\max_{1\leq i\leq n}\sum_{j=1}^{n}\left|a_{ij}\right|$$

$$=\alpha d\left(x_1,x_2\right)$$

同样是完备的距离空间，满足压缩映射的条件，所有有唯一的解。这两个例题充分描述了，根据研究问题的不同，我们可以选择不同的距离空间，只要是完备的就行。

例 5.4 Fredhom 积分方程

$$x(t)=\varphi(t)+\mu\int_a^b k(t,s)x(s)ds \tag{22}$$

其中，$k(t,s),\varphi(t)$ 是 $a\le t\le b, a\le s\le b$ 上的连续函数，则当 $|\mu||b-a|M<1$ 时，方程存在唯一解，其中，

$$M=\max_{a\le t\le b, a\le s\le b}|k(t,s)| \tag{23}$$

证明，1.

$$\mathrm{T}t=\varphi(t)+\mu\int_a^b k(t,s)x(s)ds \tag{24}$$

T 是从 $C[a,b]$ 到 $C[a,b]$ 的映射。

2. 对任意的 $x,y\in C[a,b]$，有

$$\begin{aligned} d\left(Tx_1,Tx_2\right) &= \max_{a\le t\le b}|\mu|\left|\int_a^b\left[k(t,s)\left(x_1(s)-x_2(s)\right)\right]\mathrm{d}s\right| \\ &\le |\mu||b-a|M\max_{a\le t\le b}\left|x_1(t)-x_2(t)\right| \\ &= |\mu||b-a|\,\mathrm{Md}\left(x_1,x_2\right) \end{aligned} \tag{25}$$

其中，

$$M=\max_{a\le t\le b, a\le s\le b}|k(t,s)| \tag{26}$$

且因为，$|\mu||b-a|\,\mathrm{M}\in(0,1)$，所以，满足压缩映射条件，方程有唯一解。

例 5.5 Volterra 积分方程

$$x(t)=\varphi(t)+\mu\int_a^t k(t,s)x(s)ds \tag{27}$$

其中，$k(t,s),\varphi(t)$ 是 $a\le t\le b, a\le s\le b$ 上的连续函数，则方程存在唯一解。这个 Volterra 方程和前面的 Fredhom 方程最大的区别即为，不再是定积分，而是变限积分函数。和前面一样，首先要确定完备的距离空间，并确立映射关系。

1.

$$\mathrm{T}x=\varphi(t)+\mu\int_a^t k(t,s)x(s)ds \tag{28}$$

T 是从 $C[a,b]$ 到 $C[a,b]$ 的映射。

2. 对任意的 $x,y\in C[a,b]$，有

$$\begin{aligned} \left|Tx_1-Tx_2\right| &= |\mu|\left|\int_a^t k(t,s)\left[x_1(s)-x_2(s)\right]\mathrm{d}s\right| \\ &\le |\mu|M(t-a)\max_{a<t\le b}\left|x_1(t)-x_2(t)\right| \end{aligned} \tag{29}$$

其中，$M=\max_{a\le t\le b, a\le s\le b}|k(t,s)|$。

3. 进一步的有

$$\begin{aligned}\left|T^2x_1 - T^2x_2\right| =& \left|T\left(Tx_1\right) - T\left(Tx_2\right)\right| \\ \leq& |\mu|^2M^2\int_a^t(\tau-a)\max_{a\leq t\leq b}|x_1(\tau)-x_2(\tau)|\,\mathrm{d}\tau \\ =& |\mu|^2M^2\frac{(t-a)^2}{2}\max_{a\leq t\leq b}|x_1(t)-x_2(t)|\end{aligned} \tag{30}$$

4. 一般来说，

$$|T^nx_1 - T^nx_2| \leq |\mu|^nM''\frac{(t-a)^n}{n!}\max_{a\leq t\leq b}|x_1(t)-x_2(t)| \tag{31}$$

5. 所以，

$$d(\mathrm{T}^nx_1, \mathrm{T}^nx_2) \leq |\mu|^nM''\frac{(t-a)^n}{n!}d(x_1, x_2) \tag{32}$$

由于，

$$|\mu|^nM''\frac{(t-a)^n}{n!} \to 0 \quad (n\to\infty)$$

所以，

$$|\mu|^nM''\frac{(t-a)^n}{n!} \in (0,1)$$

由 3.3 节的描述，可以得出有此积分方程有唯一解。而此证明过程，放大的比较细致，所以对条件的要求比较少；放大的越粗糙，当然对条件的要求会比较高。

上述这些对于求解方程的例子，都是通过把原来的问题转化为不动点问题解决的。压缩映射原理在隐函数存在性定理。研究各类方程解的存在性、唯一性以及近似解的收敛性等理论中起到了十分重要的作用。

6 本章小结

本章主要介绍的是完备空间中的应用，压缩映射定理。压缩映射原理在隐函数存在性定理，研究各类方程解的存在性、唯一性以及近似解的收敛性等理论中起到了十分重要的作用。也是小编研究的强化学习中非常重要的定理。本章首先对压缩定理进行了证明，然后举了一些例子来讲述压缩映射定理在其中的应用。压缩映射的使用步骤整体为两步，第一步证明是在完备的举例空间中，第二步验证是否满足压缩映射定理的条件。其中，学习到了放缩是数学中最重要的技巧之一，放缩的方法有很多，而往往越精细的方法得到的结果会越好。

第一章习题课

Chen Gong

18 January 2021

目录

1 Introduction

俗话说，学习不刷题等于白学。所以，在学习完一章后，通过刷题来提高对知识点的理解是非常必要的。本次习题课老师主要讲解了七道习题，来帮助我们进一步提高对本章知识的理解。

2 第一题

在 $\mathbb{R}$ 上定义 $d(x,y)=\arctan|x-y|$，问 $(\mathbb{R},d)$ 是不是距离空间？

分析：我们只要一一验证，$d(x,y)$ 是否可以满足距离的四个条件就行。包括，非负性；严格正 $(d(x,y)=0\to x=y)$；对称性；满足三角不等式。一般来说前三条是比较好验证的，主要在于第四条的验证。

其中，前三个条件是显然成立的，下面验证第四个条件是否可以成立。

$$\arctan|x-y|\le\arctan|x-z|+\arctan|z-y| \tag{1}$$

由于，$\arctan(\cdot)\in[0,\frac{\pi}{2})$，所以易得，

$$\begin{aligned}&\arctan|x-y|\in[0,\frac{\pi}{2})\\&\arctan|x-z|+\arctan|z-y|\in[0,\pi)\end{aligned} \tag{2}$$

显然，当 $\arctan|x-z|+\arctan|z-y|\in[\frac{\pi}{2},\pi)$ 的时候，是显然满足公式 (1) 的。那么，我们只需考虑另外一半情况，$\arctan|x-z|+\arctan|z-y|\in[0,\frac{\pi}{2}]$。

$$\begin{aligned}\tan\{\arctan|x-z|+\arctan|z-y|\}=&\frac{|x-z|+|z-y|}{1-|x-z||z-y|}\\\ge&|x-z|+|z-y|\\\ge&|x-y|=\tan\{\arctan|x-y|\ |\end{aligned} \tag{3}$$

由于，tan 还是是单调递增的函数，所以可以得出，$\arctan|x-z|+\arctan|z-y|\}>\arctan|x-y|$。有的小伙伴，可能会疑惑，为什么可以确定 $1-|x-z||z-y|\in(0,1]$。因为，可以确定 $\arctan|x-z|+\arctan|z-y|\in[0,\frac{\pi}{2}]$，这个区间的 tanh 值一定是正数。所以，$1-|x-z||z-y|$ 一定是大于零的，小于等于 1 是一件很显然的事情。那么，我们可以很简单的得出满足三角不等式。所以，$(\mathbb{R},d)$ 是距离空间。

一般的，对于 $\forall x,y\in\mathbb{R}$，他们之间的距离为 $|x-y|$，实数空间在这个距离下是无界的。本题在实数空间中定义了一个新的距离：$d(x,y)=\arctan|x-y|$。这样把实数空间 $\mathbb{R}$ 中一个无界的集合，转变成了有界的集合。所以，有界和无界是相对的概念，和空间定义的距离有关。

3 第二题

在 n 维欧几里得空间 $\mathbb{R}^n$ 中，对于

$$x=(x_1,x_2,\cdots,x_n),\quad y=(y_1,y_2,\cdots,y_n)$$

定义

$$d(x,y)=\sum_{i=1}^{n}\lambda_i|x_i-y_i| \tag{4}$$

其中，$\lambda_1, \cdots, \lambda_n$ 是 n 个正数。证明，d 是 $\mathbb{R}^n$ 中的距离，并且按距离收敛等价于按坐标收敛。

按距离收敛等价于按坐标收敛，即有，

$$d(x_k, x) \to 0 \ (k \to \infty) \tag{5}$$

其中，$x_k = \{x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}\} \in \mathbb{R}^n$，$x = \{x_1, x_2, \cdots, x_n\} \in \mathbb{R}^n$，

$$\Longleftrightarrow x_i^{(k)} \to x_i \ (k \to \infty),\ i = 1, 2, \cdots, n.$$

证明：

首先，需要证明 d 是 $\mathbb{R}^n$ 中的距离，很显然前三条条件是满足的，重点则是判断是否满足三角不等式，对于 $\forall x, y, z \in \mathbb{R}^n$，

$$\begin{aligned} d(x,y) &= \sum_{i=1}^{n} \lambda_i \left| x_i - y_i \right| = \sum_{i=1}^{n} \lambda_i \left| x_i - z_i + z_i - y_i \right| \\ &\leq \sum_{i=1}^{n} \lambda_i \left(\left| x_i - z_i \right| + \left| z_i - y_i \right| \right) \\ &= d(x,z) + d(z,y) \end{aligned} \tag{6}$$

所以，d 是 $\mathbb{R}^n$ 上的距离。

2. 首先需要证明的是，按距离收敛可以推出按坐标收敛。回顾一下，什么是按坐标收敛，按坐标收敛即为向量中的每一个值都收敛。则要证明，$\{x_k\}$ 是 $\mathbb{R}^n$ 中按距离收敛于 x，即

$$d(x_k, x) \to 0 \ (k \to \infty)$$

对 $\forall \epsilon > 0$，$\exists K(\epsilon) \in Z$，当且仅当 $k > K(\epsilon)$ 时，有

$$d(x_k, x) = \sum_{i=1}^{n} \lambda_i \left| x_i^{(k)} - x_i \right| < \lambda \epsilon \tag{7}$$

其中，$\lambda = \min\{\lambda_1, \lambda_2, \cdots, \lambda_n\} > 0$。而对于每一个 i，有

$$\lambda_i \left| x_i^{(k)} - x_i \right| \leq \sum_{i=1}^{n} \lambda_i \left| x_i^{(k)} - x_i \right| < \lambda \epsilon \tag{8}$$

所以，有

$$\left| x_i^{(k)} - x_i \right| < \frac{\lambda}{\lambda_i} \epsilon \Longleftrightarrow \epsilon,\ i = 1, 2, \cdots, n$$

这表明 $\{x_k\}$ 在 $\mathbb{R}^n$ 中按坐标收敛于 x。

3. 第三步则是证明充分性，反之，设 $\{x_k\}$ 在 $\mathbb{R}^n$ 中按坐标收敛于 x，则对于每一个 i，对 $\forall \epsilon > 0$，$\exists K_i \in Z$，使得当 $k > K_i$，时

$$|x_i^{(k)} - x_i| < \frac{\epsilon}{n\lambda} \tag{9}$$

其中，$\lambda = \max\{\lambda_1, \lambda_2, \cdots, \lambda_n\} > 0$。令 $K = \max\{K_1, K_2, \cdots, K_n\}$，当 $k > K$ 时，则对于每个 i 有，

$$|x_i^{(k)} - x_i| < \frac{\epsilon}{n\lambda}$$

于是，

$$\sum_{i=1}^{n} \lambda_i \left| x_i^{(k)} - x_i \right| < \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{n\lambda} \epsilon < \epsilon.$$

故 $d(x_k, x) \to 0 \ (k \to \infty)$，即 $\{x_k\}$ 在 $\mathbb{R}^n$ 中按距离收敛于 x。

4 第三题

设 X 是一个距离空间，$A \subset X, x \in X$，称，

$$d(x, A) = \inf\{d(x, \omega) | \omega \in A\} \tag{10}$$

为点 x 到集合 A 的距离，证明

$$\bar{A} = \{x | d(x, A) = 0\}. \tag{11}$$

且 $\bar{A}$ 是包含 A 的最小闭集。

实际上，大家自己画画就会觉得这是显然的。那么，我们证明，$A \subset \bar{A} \wedge \bar{A} \subset A$，对于任意一个 A 中的点都属于 $\bar{A}$，且任意一个 $\bar{A}$ 中的点都属于 A 即可。

本题中需要使用到几个简单的数学概念，1. 下确界的定义，(下确界是最大下界，1. 集合中的任何点都比下确界 X 大；而且对于任意的 $\epsilon > 0$，都存在点 x_0，有 $X + \epsilon > x_0$)；2. 闭包的定义，(A 的接触点全体称为 A 的闭包，记为 $\bar{A}$)；3. 接触点的定义：$x \in X$，如果对于 $\forall \epsilon > 0$，

$$B(x, \epsilon) \cap A \neq \emptyset$$

则称 x 为 A 的接触点。那么，如何来证明呢？首先想到的即为，$\bar{A} \subset$ 任意的接触点的集合。用数学的符号表达即为：

令 $B = \{x | d(x, A) = 0\}$，先证明 $B \in \bar{A}$。若 $x_0 \in B$，即，

$$d(x_0, A) = \inf\{d(x_0, \omega) | \omega \in A\} \tag{12}$$

由下确界的定义，对于 $\forall \epsilon > 0$，$\exists \omega \in A$，使得：

$$d(x_0, \omega) < \epsilon,$$

即 $B(x_0, \epsilon) \cap A \neq \emptyset \rightarrow x_0 \in A$。

反之，若 $x_0 \in \bar{A}$，即 x_0 是 A 的接触点。那么，对于 $\forall \epsilon > 0$，

$$B(x, \epsilon) \cap A \neq \emptyset$$

即存在，$\omega \in A$，使得，

$$d(x_0, \omega) < \epsilon \rightarrow \inf\{d(x_0, \omega) | \omega \in A\} \leq \epsilon \tag{13}$$

由于 ϵ 是任意的，我们有，

$$\inf\{d(x_0, \omega) | \omega \in A\} = 0$$

即 $x_0 \in B = \{x | d(x, A) = 0\}$。综上可得，

$$\bar{A} = \{x | d(x, A) = 0\} \tag{14}$$

下面则是证明 $\bar{A}$ 是包含 A 的最小闭集，设 M 为包含 A 的任意闭集，$A \subseteq M$。对 $\forall x \in \bar{A}$，存在 $\{x_0\} \subseteq A \subseteq M$ 使得，

$$\lim_{n \to \infty} d(x, x_0) = 0. \tag{15}$$

由 M 闭可知 $x \in M$。因此，$\bar{A} \subset M$。这也证明了 $\bar{A}$ 是包含 A 的最小闭集。通过这个题，非常详细的刻画了 A 的闭包的几何特征，其描述的是到 A 距离等于 0 的所有的点构成的集合。

5 第四题

设 X 按照距离 d 为距离空间，$A \subset X$ 非空，令，

$$f(x) = \inf_{y \in A} d(x, y) \qquad (\forall x \in X). \tag{16}$$

证明，$f(x)$ 是 X 上的连续函数。分析：这里的 $f(x)$ 是定义在 X 上的非负的实值函数。由连续函数的定义可知，需证明对 $\forall x_0 \in X$，对 $\forall \epsilon > 0$，存在 $\delta > 0$，当 $d(x, x_0) < \delta$ 时有，

$$d_1(f(x), f(x_0)) = |f(x) - f(x_0)| < \epsilon \tag{17}$$

此证明比较容易，需要用到距离定义中的三角不等式。实际上，此证明过程相对简单。我们已知的有 $d(x, x_0) < \delta$，是不是想办法构建 $d(x, x_0)$ 和 $d_1(f(x), f(x_0))$ 之间的关系就可以了。首先有，

$$\begin{aligned} d(x, y) &\leq d(x, x_0) + d(x_0, y) \\ d(x_0, y) &\leq d(x, x_0) + d(x, y) \end{aligned} \tag{18}$$

左右两边都对 $y \in A$ 取下界有，

$$\begin{aligned} \inf_{y \in A} d(x, y) &\leq d(x, x_0) + \inf_{y \in A} d(x_0, y) \\ \inf_{y \in A} d(x_0, y) &\leq d(x, x_0) + \inf_{y \in A} d(x, y) \end{aligned} \tag{19}$$

将公式 (16) 代入即可得，

$$|f(x) - f(x_0)| \leq d(x, x_0) < \delta = \epsilon \tag{20}$$

所以，可以证明，$f(x)$ 为连续函数。

6 第五题

证明集合 $M = \{\sin nx | n = 1, 2, \cdots\}$ 在空间 $C[0, \pi]$ 中是有界集，但不是列紧集。

什么是列紧集呢？$C[a, b]$ 中的子集 A 是列紧的当且仅当 A 中的函数是一致有界和等度连续的。

什么是一致有界呢？即存在 $K > 0$，使得对于每一点 $t \in [a, b]$ 及一切的 $x \in A$，有

$$|x(t)| \leq K$$

通俗的讲，是所有的函数都有共同的界。而什么又是等度连续呢？对于任意的 $\epsilon > 0$，存在 $\delta > 0$，当 $|t_1 - t_2| < \epsilon$ 时，$|x(t_1) - x(t_2)| < \epsilon \quad (\forall x \in A)$。

分析：$C[0, \pi]$ 中的子集 M 是列紧的当且仅当 M 中的函数是一致有界和等度连续的。显然集合 M 是一致有界的 (因为 $\sin(\cdot)$ 是一定 $\in [0, 1]$ 的。)，要证明 M 不是列紧集，则证明 M 不是等度连续的。在此题中，即为，

对于任意的 $\epsilon > 0$，存在 $\delta > 0$，当 $|t_1 - t_2| < \delta$ 时，

$$|\sin n \cdot t_2 - \sin n \cdot t_1| < \epsilon \qquad (\forall n \in Z^+)$$

我们的目标是要证伪这个命题，即为得到命题的否命题。否命题即为，$\exists \epsilon > 0$，$\forall \delta > 0$，$\exists n_0 \in Z^+$，虽然 $\exists t_1, t_2 \in [0, \pi]$，且 $|t_1 - t_2| < \delta$，但是，$|\sin n \cdot t_2 - \sin n \cdot t_1| \geq \epsilon$。实际上，我们的主要目标是找 t_1, t_2 和 n_0。实际上证伪，只需要找到一个反例就行。

为了简单起见，我们令 $t_1 = 0$，令 $t_2 = \frac{\pi}{2n_0} \in [0, \pi]$。想要，

$$|t_1 - t_2| < \delta \rightarrow n_0 > \frac{\pi}{\delta} > \frac{\pi}{2\delta}$$

且，n_0 要求为整数，所以，令 $n_0 = [\frac{\pi}{2\delta}] + 1$，于是就可以顺理成章的得到，

$$|t_1 - t_2| < \frac{\pi}{2n_0} \leq \frac{\pi\delta}{4} < \delta$$

而，

$$|\sin n \cdot t_2 - \sin n \cdot t_1| = \sin \frac{\pi}{2} = 1 = \epsilon_0$$

由此可得，$M = \{\sin nx | n = 1, 2, \cdots\}$ 是非等度连续。通过这个题，非常形象的描述了，有限维空间和无穷维空间的区别。在有限维空间中，有界就一定列紧，而无限维空间中则是不一定。非常有可能是非等度连续的。

7 第六题

设 D 是 $[0, 1]$ 区间上具有连续导数的实函数全体，在 D 上定义，

$$d(x, y) = \max_{0 \leq t \leq 1} |x(t) - y(t)| + \max_{0 \leq t \leq 1} |x'(t) - y'(t)| \tag{21}$$

(1) 证明 D 是距离空间;

(2) 指出 D 中点列按距离收敛的意义;

(3) 证明 D 是完备的.

要证明 D 是完备的，即为证明 D 中的每一个 Cauchy 列是收敛的。这个证明流程，大致分为三步，1. 需要找一个 $x_0(t)$；2. 确定 $x_0(t) \in C[a, b]$；3. 确定 $x_n(t) \rightarrow x_0(t)$ $(n \rightarrow \infty)$。但是，空间 D 是 $C[0, 1]$ 空间的一个真子空间，但是在其上定义的距离不同。

(1) 要证明 D 是距离空间，只要证明在 D 中所定义的距离 d 满足距离的定义的四条即可。其主要为验证三角不等式成立。

设 $x, y \in D$，则 $\forall t \in [0, 1]$，有，

$$\begin{aligned} |x(t) - y(t)| \leq |x(t) - z(t)| + |z(t) - y(t)| \\ |x'(t) - y'(t)| \leq |x'(t) - z'(t)| + |z'(t) - y'(t)| \end{aligned} \tag{22}$$

二式对 $t \in [0, 1]$ 取最大值并相加得:

$$d(x, y) < d(x, z) + d(z, y)$$

(2) D 中的点列 $\{x_n\}$ 收敛于 x 的充要条件的 $\{x_n(t)\}$ 在 [0,1] 上一致收敛于 $x(t)$ 且 $\{x'_n(t)\}$ 在 [0,1] 上一致收敛于 $x'(t)$。

证明：必要性：设 $x_n(t)$ $(n = 1, 2, \cdots)$，$x(t) \in D$。且 $d(x_n, x) \rightarrow 0$。于是对，$\forall \epsilon > 0, \exists N$，当 $n > N$ 时，$d(x_n, x) < \epsilon$，即:

$$\max_{0 \leq t \leq 1} |x_n(t) - x(t)| + \max_{0 \leq t \leq 1} |x'_n(t) - x'(t)| < \epsilon \tag{23}$$

于是对 $\forall t\in[0,1]$，有

$$\begin{aligned}|x_n(t)-x(t)| &\le \max_{0\le t\le 1}|x_n(t)-x(t)|<\epsilon\\ |x_n'(t)-x'(t)| &\le \max_{0\le t\le 1}|x_n'(t)-x'(t)|<\epsilon\end{aligned} \tag{24}$$

显然可以得出结论！

充分性：即 $\{x_n(t)\}$，$\{x_n'(t)\}$ 在 [0,1] 上分别一致收敛到 $x(t)$，$x'(t)$。反之，$\{x_n(t)\}$，$\{x_n'(t)\}$ 在 [0,1] 上分别一致收敛到 $x(t)$，$x'(t)$。对 $\forall\epsilon>0,\exists N$，当 $n>N$ 时，对 $\forall t\in[0,1]$，有

$$|x_n(t)-x(t)|<\frac{\epsilon}{2},\quad |x_n'(t)-x'(t)|<\frac{\epsilon}{2}, \tag{25}$$

上两式分别对 $t\in[0,1]$ 取最大值，并相加得：

$$\max_{0\le t\le 1}|x_n(t)-x(t)|+\max_{0\le t\le 1}|x_n'(t)-x'(t)|<\epsilon \tag{26}$$

说明 $x_n\to x$ $(n\to\infty)$。即 D 中的收敛是函数列和函数的导数列在 [0,1] 上的一致收敛。

(3) 证明 D 是完备空间，需要用到 Cauchy 列的定义，已经完备空间的性质，之前都分析过了。设 $\{x_n(t)\}$ 是 D 中的任意一个 Cauchy 列，要证明存在 $x_0(t)\in D$，使得 $\{x_n(t)\}$ 按 D 中的距离收敛到 $x_0(t)$，即：

$$\max_{0\le t\le 1}|x_n(t)-x_0(t)|+\max_{0\le t\le 1}|x_n'(t)-x_0'(t)|\to 0\ (n\to\infty) \tag{27}$$

证明中应当注意以下三点：

1. 首先证明存在 $x_0(t)$，$\{x_n(t)\}$ 一致收敛到 $x_0(t)$；
2. 首先证明存在 $y_0(t)$，$\{x_n'(t)\}$ 一致收敛到 $y_0(t)$；
3. 关键说明：$y_0(t)=x_0'(t)$，这说明 $\{x_n(t)\}$ 按 D 中的距离收敛到 $x_0(t)$。

8 第七题

证明存在闭区间 [0,1] 上的连续函数 $x(t)$，使得

$$x(t)=\frac{1}{2}\sin x(t)-a(t) \tag{28}$$

其中，$a(t)$ 是给定的 $[0,1]$ 上的连续函数。完备空间中的重要应用就是，压缩映射定理。

证明，在空间 $C[0,1]$ 上考虑如下映射：

$$\mathrm{T}x(t)=\frac{1}{2}\sin x(t)-a(t)$$

显然，T 是 $C[0,1]$ 到 $C[0,1]$ 自身的映射，证明目标是得到：$d(\mathrm{T}x,\mathrm{T}y)<\theta d(x,y)$。对于 $\forall x,y\in C[0,1]$，$\forall t\in[0,1]$，有，

$$\begin{aligned}|Tx(t)-Ty(t)| &= \left|\left(\frac{1}{2}\sin x(t)-a(t)\right)-\left(\frac{1}{2}\sin y(t)-a(t)\right)\right|\\ &= \left|\frac{1}{2}\sin x(t)-\frac{1}{2}\sin y(t)\right|\\ &= \frac{1}{2}\left|2\cos\frac{x(t)+y(t)}{2}\sin\frac{x(t)-y(t)}{2}\right|\\ &\le \left|\sin\frac{x(t)-y(t)}{2}\right|\\ &\le \frac{1}{2}|x(t)-y(t)|\end{aligned} \tag{29}$$

所以，

$$\begin{aligned} d(Tx, Ty) &= \max_{\ell \in 0,1} |Tx(t) - Ty(t)| \\ &\leq \max_{t \in 0,1} \frac{1}{2} |x(t) - y(t)| = \frac{1}{2} d(x, y). \end{aligned} \tag{30}$$

由压缩映射原理知存在唯一的 $x_0 \in C[0,1]$ 使得：

$$Tx_0 = x_0 \tag{31}$$

即为：

$$x(t) = \frac{1}{2} \sin x(t) - a(t) \tag{32}$$