分析学：数学中的“空间感”（第二版）

邓月, 电子科技大学，基础与前沿研究院，941951631@qq.com

2022 年 10 月 29 日

分析学是数学家对局部微小演化进行精确量化的显微镜，它扮演着“鸟与青蛙”中的青蛙。本文覆盖了我硕士期间所学的一系列数学课程中与分析学相关的知识，核心内容总结自《泛函分析》、《非线性分析》、《实分析》、《高等概率论》、《随机过程》、《矩阵理论》等课程。其中，《非线性分析》将空间的概念下潜到了更抽象的拓扑中，将空间的维数拓展到了更一般的无穷维。应用内容联系到《数值分析》、《最优化理论》、《量子力学》、《量子场论》等课程。注意不断积累和反复品味本文所给出的例题中巧妙的证明思维和套路，这才是真正能够受益终身的能力。

特别说明，本文并没有特意地将向量或矩阵符号用粗体标识出来以与数值进行区分，而是默认读者能够基于“场景”自行识别开来。此外，本文也默认读者能够自行识别出文中一些地方对同一个概念或含义采用的不同记号形式。

主要参考资料：

【1】陈天权. (2010). 数学分析讲义（第二册）. 北京大学出版社.

【2】A.H. 柯尔莫戈罗夫, C.B. 佛明著. (1963). 函数论与泛函分析初步（第 7 版）. 高等教育出版社.

【3】姚泽清, 苏晓冰, 郑琴, 王在华. (2007). 应用泛函分析. 科学出版社.

【4】张世清. (2017). 泛函分析及其应用. 科学出版社.

【5】缪柏其. (2009). 概率论教程（第二版）. 中国科学技术大学出版社.

【6】徐全智. (2013). 随机过程及应用. 高等教育出版社.

【7】Munkres, J. R. (2000). Topology.

【8】张维[illegible]. (2013). Sobolev 空间与变分原理. 中国科学技术大学出版社.

【9】李明奇, 田太心. (2014). 数学物理方程 (第二版). 电子科技大学出版社.

目录

1 基本概念

1.1 常用的术语

1.1.1 度量空间 X 中以 x_0 为中心，r 为半径的“开球（或邻域）$Br(x_0)$”、“闭球 $\overline{Br}(x_0)$”、“球面 $Sr(x_0)$”

(1) $Br(x_0) = \{x \in X : d(x, x_0) < r\}$；

(2) $\overline{Br}(x_0) = \{x \in X : d(x, x_0) \leq r\}$；

(3) $Sr(x_0) = \{x \in X : d(x, x_0) = r\}$。

理解：

(1) 距离空间又称度量空间，而赋范空间、Banach 空间、内积空间、Hilbert 空间它们实则也都是度量空间，但反过来不一定成立；

(2) 球的形状随距离定义的不同而不同，即便在相同的距离定义下，不同的空间中球的形状也会不同，特别注意：这个“球”并不意味着形状一定要是圆的。

1.1.2 x_0 为 $A(A \subset X)$ 的“内点”、“边界点”、“触点”、“聚点（或极限点）”

(1) $\exists r > 0, B_r(x_0) \subset A$；

(2) $\forall r > 0, B_r(x_0) \cap A \neq \emptyset$ 且 $\forall r > 0, B_r(x_0) \cap A^c \neq \emptyset$；

(3) $\forall r > 0, B_r(x_0) \cap A \neq \emptyset$；

(4) $\forall r > 0, B_r(x_0) \cap (A \backslash \{x_0\}) \neq \emptyset$。

理解：

内点不一定是聚点，比如 $n \in V \subset Z, \exists r = \frac{1}{2} > 0$，是内点，但不满足聚点的定义。

1.1.3 A 的“内部 A° ”、“边界 ∂A”、“闭包 $\overline{A}$”、“导集 A' ”

(1) A 的内点的全体；

(2) A 的边界点的全体；

(3) A 的触点的全体；

(4) A 的聚点（或极限点）的全体。

理解：

(1) 它们四者之间的联系有：$\overline{A}=A^{\circ}\cup\partial A=A\cup A'$，$A^{\circ}=\overline{A}\backslash\partial A=A\backslash\partial A,\partial A=\overline{A}\backslash A^{\circ}=\overline{A}\cap\overline{A^{c}}$。

1.1.4 A 是"开集"、"闭集"、"有界集"

(1) $A=A^{\circ}$；

(2) $A=\overline{A}$；

(3) $\operatorname{diam}(A)<\infty$，其中 $\operatorname{diam}(A)=\sup\{d(x,y):x,y\in A\}$。

理解：

(1) 一个集合是开集还是闭集，依赖于它所依附的空间 X，例如集合 $A=\{1,\frac{1}{2},\frac{1}{3},...\}$，在 R 中非开非闭，有聚点 0；在 $X=A\cup\{0\}$ 中开，非闭，有聚点 0，在 $X=A$ 中既开又闭，无聚点（上述分析可看做集合 A 为 X 的什么集，因为集合 A 中每个点是 X 中的什么点，直观上理解，将集合 A 看做是大球 X 中的小球）；

(2) 空集既是开集又是闭集。

1.1.5 稠集、稠子集、稠密、可分

(1) A 为 X 的稠集，若 $\overline{A}=X$；

(2) A 为 B 的稠子集，若 $A\subset B\subset\overline{A}$；

(3) A 在 B 中稠密，若 $B\subset\overline{A}$；

(4) 空间 X 是可分的，若 X 具有可数的稠集。X 的子集 A 称为可分的，若 A 具有可数的稠子集。【1.1-1 的第 2 题例子，体会"可数"的重要性】

理解：

(1)A 在 B 中稠密，表示 B 中的点可以用 A 中的一串点来逼近，即：$\forall x\in B,\exists\{x_n\}\subset A,\mathrm{s.t.}\,x_n\to x(n\to\infty)$。注意：并不要求 A 一定要是 B 的子集。

(2) **可数**和可列是一个概念。我们说一个集合 X 可数，如果 X 中的元素可以与全体正整数 $\mathcal{N}$ 建立一一对应，即存在一个双射 $f:\mathcal{N}\to X$。

可数集合是"最小"的无穷集合。如果一个集合可数，说明它和 $\mathcal{N}$ 一样大。

1.1.6 列紧集、紧集、紧空间

(1) 设 A 是空间 X 的子集，若 A 中的任一点列都有收敛子列，则称 A 是列紧集；

(2) 若 A 中的任一点列都有收敛于 A 的子列，则称 A 为紧集；

(3) 当 X 为列紧集时，X 必为紧集，此时称 X 为紧空间。

理解：

(1) **子列**的定义：在点列中任取无穷多项，不改变它们在原来点列中的先后次序，所得到的点列称为原来点列的一个子列；

(2) 列紧集与紧集的区别在于：前者只要求极限属于大空间 X，后者则要求极限属于集合本身。

1.1.7 满射、单射、双射

首先要注意，映射是不可以一对多的！只能一对一或多对一！

(1) 对映射 $f:X\to Y$，若 f 的值域 $f(X)=Y$，则称 f 是 X 到 Y 的满射（满射可以从直观上理解为二部图中没有孤立节点~）；

(2) 若 $\forall x_1,x_2\in X,x_1\neq x_2$，有 $f(x_1)\neq f(x_2)$，则称 f 为 X 到 Y 上的单射；

(3) 若 f 既是单射又是满射，则称 f 为 X 到 Y 上的双射（又叫"一一对应"或"一一映射"，直观理解）。

1.1.8 点列的“收敛性”、映射的“连续性”

(1) 设 $\{x_n\}$ 是空间 X 中的一个点列，$x_0\in X$，若 $\lim_{n\to\infty}d(x_n,x_0)=0$，则称点列 $\{x_n\}$ 收敛于 x_0，记作 $\lim_{n\to\infty}x_n=x_0$ 或 $x_n\to x_0(n\to\infty)$；

(2) 若映射 $f:X\to Y$ 满足$\forall\epsilon>0,\exists\delta>0,\mathrm{s.t.}\,f(B_\delta(x_0))\subset B_\epsilon(f(x_0))$，则称映射 f 是连续的。

理解：

(1) 用点列的收敛性刻画“闭集”：A 为闭集的充分必要条件是 $\forall\{x_n\}\subset A$，若 $x_n\to x_0$，则 $x_0\in A$；

(2) 映射的连续性也可表述为：$\forall\epsilon>0,\exists\delta>0$，只要 $d_1(x,x_0)<\delta$，就有 $d_2(f(x),f(x_0))<\epsilon$。

1.1.9 【在拓扑空间中对上述一些基本概念进行更抽象的定义】

拓扑和拓扑空间的定义见1.3.7。时刻谨记，拓扑空间中考虑的是集合（可以从简单的“点”开始理解，同时集合中的元素还可以是函数等等），与上述通过度量来“限定范围”的方式比起来更抽象。理解时将两种定义结合起来相辅相成。设 X 是拓扑空间（这里是 $(X,\mathcal{T})$ 的简便形式，即本质上就是个子集族），则：

(1) 拓扑空间中的元素称为 (关于拓扑 $\mathcal{T}$ 的) **开集**。对应于拓扑空间的 3 个条件，我们同样可以对开集进行公理化定义，即：“任意多个开集的并是开集”、“有限多个开集的交是开集”、“空集和全空间是开集”。

(2) $x\in X,N\subset X$，N 称为 x 的一个**邻域**，若有开集 (拓扑空间中的元素) G 使得 $x\in G\subset N$。这时，我们称 x 为 N 的一个**内点**；（这里的 G 可以先理解为度量形式下的 r，本质是一个意思，只是这里更抽象，邻域是针对“半径”而言的）

(3) $F\subset X$，F 称为**闭集**，假若 $F^C=X\backslash F$ 是开集。

(4) $E\subset X$，E 的**闭包**，记作 E，定义为所有包含 E 的闭集之交。E 的闭包是包含 E 的最小闭集，即任何包含 E 的闭集必包含 E 的闭包 $\overline{E}$ (正所谓“必 (然) 包 (含)” 哈哈)；

(5) $E\subset X$，点 $x\in X$ 称为集 E 的一个**聚点 (或极限点)**，假若对 x 的任何邻域 U，有 $E\cap(U\backslash\{x\})\neq\emptyset$。$E$ 的聚点的全体称为 E 的**导集**，记作 E'。$x\in E$ 称为 E 的一个**孤立点**，若 x 非 E 的极限点。$\overline{E}=E'\cup E$；

(6) $E\subset X$，E 的**内核**，记作 E°，定义为 E 的全体开子集之并。E 的内核是 E 的最大开子集，也是 E 的内点的全体；

(7) $E\subset X,\overline{E}\backslash E^\circ$ 称为 E 的**边界**；

(8) $D\subset E\subset X$，D 称为 E **中的稠密集**，若 $\overline{D}\supset E$。特别地，$D\subset X$ 称为是拓扑空间 X **中的稠密集**，若 $\overline{D}=X$。$D\subset X$ 在 X 中**稠密**，当且仅当对于 X 的任何非空开集 G，有 $G\cap D\neq\emptyset$；

(9) 若 X 有至多可数个点组成的稠密子集，则 X 称为**可分的**；

(10) 设 X 和 Y 是两个拓扑空间，映射 $f:X\to Y$ 称为**在点** $x\in X$ **处连续**，假若对于点 $f(x)$ 在 Y 中的任何邻域 N，$f^{-1}(N)$ 是点 x 的邻域。X 到 Y 的映射 f 称为**在集合** $A\subset X$ **上连续**，若对于任何点 $x\in A$，它在点 x 处连续。若 f 在 X 上连续，便简称 f **连续**【《数学分析讲义 (第二册)》P10 例子】；

(11) 设 X 和 Y 是两个拓扑空间，双射 $f:X\to Y$ 称为 X 和 Y 之间的一个**同胚 (映射)**，若 f 和 f^{-1} 分别是 $X\to Y$ 和 $Y\to X$ 的连续映射。两个拓扑空间称为**同胚的**，若这两个拓扑空间之间存在同胚 (映射)【《数学分析讲义 (第二册)》P11 例子，判断依据就是看在两个拓扑空间之间能不能找到满足或不满足条件的映射，比如 $(-\pi/2,\pi/2)$ 和 R 之间能找到映射 $\tan$；$(-\pi,\pi)$ 和 (a,b) 之间能找到平移和相似变换映射】【1.1-4 的 14 题体会】；

(12) 设 $\{x_n\}$ 是 X 中的一个点列，$x\in X$。$\{x_n\}$ 被称为**收敛于** x 的，记作 $n\to\infty$ 时 $x_n\to x$，或记作 $\lim_{n\to\infty}x_n=x$，若对于 x 的任何邻域 U，有一个 n_0，使得 $n\geq n_0\Rightarrow x_n\in U$（从集族与元素的角度看哦）。

(13) 对“**覆盖**”的理解参见1.1.58节的（2）芝诺悖论。设 X 是拓扑空间，$E\subset X$。设 A 是个指标集，$\mathcal{G}=\{G_\alpha:\alpha\in A\}$ 称为 E 的一组**开覆盖**，假若对于每个 $\alpha\in A$，G_α 是开集，且 $E\subset\bigcup\limits_{\alpha\in A}G_\alpha$。若 A 是有限集，则 $\mathcal{G}=\{G_\alpha:\alpha\in A\}$ 称为 E 的一组**有限开覆盖**。若 $B\subset A$，且 $E\subset\bigcup\limits_{\alpha\in B}G_\alpha$，则称 $\mathcal{G}'=\{G_\alpha:\alpha\in B\}$ 为 $\mathcal{G}=\{G_\alpha:\alpha\in A\}$ 的（关于 E 的）**子覆盖**。E 的一组开覆盖 $\mathcal{G}=\{G_\alpha:\alpha\in A\}$ 称为**能有限覆盖** E，假若它有（关于 E）的**有限子覆盖**。

(14) 设 X 是拓扑空间，$K \subset X$。K 称为**紧集**，若 K 的任何开覆盖有有限子覆盖。换一个说法：给定 $\mathbb{R}$ 的子集 X，如果对于任意一个并集包含 X 的开区间族 $\mathcal{U} = \{U_i\}_{i \in I}$，都存在 $\mathcal{U}$ 中有限个元素（开区间）$U_1, U_2, ..., U_n$ 使得 $X \subset \bigcup\limits_{i=1}^{n} U_i$，则说 X 是紧的。当 $K = X$ 时，X 称为**紧空间**。【1.1-4 的 15 题例子】

在【7】中还对 compact 有如下定义：A collection $\mathcal{A}$ of subsets of a space X is said to cover X, or to be a **covering** of X, if the union of the elements of A is equal to X. It is called an **open covering** of X if its elements are open subsets of X. A space X is said to be **compact** if every open covering $\mathcal{A}$ of X contains a finite subcollection that also covers X.

对于紧性的直观理解，可以参见知乎链接 https://www.zhihu.com/question/31734712。简而言之就是：紧致性是把性质从局部往整体推的法宝。比如证明“有界闭区间上的连续函数一定有界”，即从无穷到有限，类似于上面测度论中所揭示的“长度”和覆盖。再比如“半开半闭区间 $[0,1)$ 不是紧致的”。又比如“$\mathbb{Q} \cap [0,1]$ 不是紧致的（因为这里的子集没有上确界）”

(15) 设 E 是度量空间 X 的子集，E 的**直径**定义为 $\operatorname{diam} E = \sup\limits_{x,y \in E} \rho(x,y)$，集合 E 称为**有界的**，若 $\operatorname{diam} E < \infty$。

(16) 设 E 是度量空间 X 的子集，$\varepsilon > 0$。集合 $F \subset X$ 称为 E 的一个 **ε-网**，若 $E \subset \bigcup\limits_{x \in F} B(x, \varepsilon)$（这里的 ε 是固定的）。集合 E 称为**全有界的**，若对于任何 $\varepsilon > 0$，E 有一个由有限点集组成的 $\varepsilon-$ 网（简称有限 ε-网）。换言之，有 X 的有限子集 $\{x_1, ..., x_n\}$，使得 $E \subset \bigcup\limits_{j=1}^{n} B(x_j, \varepsilon)$。

(17) 拓扑空间 X 称为**列紧的（或称具有 Bolzano-Weierstrass 性质的）**，若 X 的每个点列都有收敛子列。

(18) 先引入一个记号：设 X 和 Y 是两个拓扑空间，$C(X,Y)$ 表示 $X \to Y$ 的连续映射的全体，若 $Y = R$，则简记做 $C(X) = C(X,R)$。

连续函数列 $\{f_n\}$ 在度量空间 $(C(X,Y),d)$ 中收敛于 f，当且仅当连续函数列 $\{f_n\}$ 在 X 上一致收敛于 f。这就是**一致度量**称谓的来源。由一致度量诱导出的 $C(X,Y)$ 上的拓扑称为**一致收敛拓扑（或一致拓扑）**。

(19) 设 X 是拓扑空间，(Y,ρ) 是度量空间，$\mathcal{F} \subset C(X,Y), x \in X$。我们称 **$\mathcal{F}$ 在 x 处等度连续**，若 $\forall \varepsilon > 0, \exists x$ 的邻域 U，$\forall y \in U, \forall f \in \mathcal{F}$，$\rho(f(y), f(x)) < \varepsilon$。若对于任何 $x \in X$，$\mathcal{F}$ 在 x 处等度连续，便称 **$\mathcal{F}$ 在 X 上等度连续**的。

(20) 设 X 是拓扑空间，$A \subset X$ 称为 X 中的**处处不稠密集**，假若 $\forall$ 非空开集 $G \subset X$，$\exists$ 非空开集 $O \subset G \cap A^C$；

(21) 设 X 是拓扑空间，$A \subset X$ 称为 X 中的**第一纲集**，若有可数个 X 中的处处不稠密集 $\{M_n : n \in N\}$，使得 $A = \bigcup_{n \in N} M_n$。$X$ 中的非第一纲集称为 X 中的**第二纲集**。例如：$\mathbb{Q}$ 是 $\mathbb{R}$ 中的第一纲集，$\mathbb{R} \backslash \mathbb{Q}$ 是 $\mathbb{R}$ 中的第二纲集。

性质：

我们可以得到以下几个命题：

(1) X 的子集 G 是开集，当且仅当 G 的任何点都是 G 的内点【1.1-4 的 13 题证明】；

(2) 设 X 是拓扑空间，$E \subset X$，则对于任何 $x \in X$，有 $x \in \overline{E} \Leftrightarrow \forall x$ 的开邻域 $U, U \cap E \neq \emptyset \Leftrightarrow \forall x$ 的邻域 $V, V \cap E \neq \emptyset$【《数学分析讲义 (第二册)》P5-6 证明】；

(3) 设 X 和 Y 是两个拓扑空间，映射 $f : X \to Y$，则映射 f 连续的充分必要条件是：对于任何 Y 中的开集 G，$f^{-1}(G)$ 在 X 中是开集【《数学分析讲义 (第二册)》P10 证明】；

(4) 相比于“紧性”的定义，“全有界”还要更强一些，因为紧集中开覆盖的 ε_i 与 x_j 对应，不固定，而“全有界”中的 ε 是固定的。

(5) 非空完备度量空间 X 是 X 的第二纲集。

(6) 设 X 是非空完备度量空间，则 X 的非空开集是第二纲集；(3) 设 X 是非空完备度量空间，则 X 的第一纲集的余集是 X 中的稠密集。【《数学分析讲义 (第二册)》P44-45 习题 7.5.3，证明过程中熟练并灵活地运用 4.1 中的 3 和 4】

理解：

（1）不论是用拓扑空间重新定义的上述概念，还是关于命题的证明，我们要摆脱度量空间中“长度”的想象，把这里进一步抽象成仅有“集合和元素 (不一定是点，还可以是函数等等)”的情形；

（2）两个拓扑空间之间的同胚关系是等价关系，拓扑空间按同胚关系可以分成许多等价类，拓扑学就是研究刻画按同胚关系分成的同一等价类中的拓扑空间的特征的数学（与抽象代数中的等价关系和等价类有着异曲同工之妙）；

（3）对于 (12) 中的关于点列收敛的定义，把 $\forall$ 理解为“可任意小”，$\exists n_0$ 相当于一个“截断操作”。

（4）X 中的第一纲集被认为是 X 中极稀疏之集，或在 X 中所占“份量”很小之集。X 中的第二纲集被认为是 X 中极不稀疏之集，或在 X 中所占“份量”很大之集。在数学研究 (例如在动力系统的研究) 中，我们无法证明一个依赖于 $x \in X$ 的命题对于一切 $x \in X$ 成立，但却能证明它对于某第二纲集中的一切 x 均成立，这时便认为该命题“相当普遍地”成立了。不难看出，X 中的第一纲集的子集仍是 X 中的第一纲集。可数个第一纲集之并仍为第一纲集。【《数学分析讲义 (第二册)》P45 习题】

1.1.10 压缩映射、不动点

(1)$T: X \to X$ 是 X 到自身的一个自映射，若存在常数 $0 < \theta < 1$，使得 $\forall x, y, \in X$ 有 $d(Tx, Ty) \leq \theta d(x, y)$，则称 T 是 X 上的压缩映射；

(2) 对 X 上的自映射 T，若 $\exists x^* \in X$，使得 $Tx^* = x^*$，则称 x^* 为 T 的一个不动点。

理解：

(1) 压缩映射是 X 上的连续映射【《应用泛函分析》P64 证明，取 $\delta = \epsilon/\theta$；

(2) 将一般方程 $f(x) = 0$ 求解的问题转换为求 T 的不动点问题的方法即令 $T = f + I$ 即可；

(3) 将压缩映射和有界线性算子及其范数结合起来理解。

1.1.11 （自）相容的矩阵范数

设 $\|\cdot\|_a : K^{m\times l} \to R, \|\cdot\|_b : K^{l\times n} \to R, \|\cdot\|_c : K^{m\times n} \to R$ 是矩阵范数, 如果 $\|AB\|_c \leq \|A\|_a \|B\|_b, \forall A \in K^{m\times l}, B \in K^{l\times n}$，则称 $\|\cdot\|_a, \|\cdot\|_b$ 和 $\|\cdot\|_c$ 是相容的。特别地, 如果 $\|AB\| \leq \|A\|\|B\|$，则称 $\|\cdot\|$ 是自相容的矩阵范数。【3.2-2 的 5 题、6 题和 7 题】

理解：

(1) 关于矩阵范数相容的一个不太恰当的解释：比如我有一瓶 500ml 的沙子和一瓶 500ml 水，如果分开装，总共需要 1000ml 的容器，但如果把他们混合到一起变成泥，这时候可能 990ml 的容器就够了，因为水可以进入到沙子的空隙中（把沙子换成石子更直观），沙子也够大方的，可以容忍水进入到自己的空隙，我们就说他俩是相容的。但如果把沙子换成钠（可以有其他更好的选择），两者就会发生化学反应放出氧气，可能还要 2000ml 的瓶子来装那生成的氧气。总之，这种相容性保证了我们不需要更多的容器，要么还是 1000ml，要么比 1000ml 要小，比如上面的 990ml（摘抄自知乎-做个人吧）；

(2) 要证明矩阵范数不是自相容的，则举反例推出矛盾即可，一般用 0 和 1 来构造反例比较方便。

1.1.12 与向量范数相容的矩阵范数、从属于向量范数的算子范数、列和范数/行和范数/谱范数

(1) 设 $\|\cdot\|_\alpha$ 是 K^n 上的向量范数，$\|\cdot\|_m$ 是 $K^{n\times n}$ 上的矩阵范数，且 $\|Ax\|_\alpha \leq \|A\|_m \|x\|_\alpha, A \in K^{n\times n}, x \in K^n$，则称 $\|\cdot\|_m$ 为与向量范数 $\|\cdot\|_\alpha$ 相容的矩阵范数；

(2) 设 $\|x\|_\alpha$ 是 K^n 上的向量范数，$A \in K^{n\times n}$，则 $\|A\|_a = \max\limits_{x\neq 0} \dfrac{\|Ax\|_\alpha}{\|x\|_\alpha} (= \max\limits_{\|u\|_\alpha = 1} \|Au\|_\alpha)$ 是与向量范数 $\|x\|_\alpha$ 相容的矩阵范数，称此矩阵范数为从属于向量范数 $\|x\|$ 的算子范数；

(3) 从属于 $\|x\|_1$ 、$\|x\|_\infty$ 、$\|x\|_2$ 的算子范数分别称为列和范数/行和范数/谱范数, 记为
$\|A\|_1 = \max\limits_j \sum\limits_{i=1}^n |a_{ij}|$ 、$\|A\|_\infty = \max\limits_i \sum\limits_{j=1}^n |a_{ij}|$ 、$\|A\|_2 = \sqrt{r(A^H A)}$。【3.6-1 的 3 题和 4 题，3.6-2 的 5 题证明】

理解：

(1) 为什么需要这个概念呢？比如在信号估计中，理想情况下应该是 $y = Ax$，但真实环境下会有扰动 Δ, 那么上面的情形就会成为 $y + \Delta y = A(x + \Delta x)$, 其中的扰动为 $\Delta y = A\Delta x$。要想估计扰动大小, 则两边

取范数为 $\|\Delta y\|_\alpha = \|A\Delta x\|_\alpha$，倘若关系 $\|A\Delta x\|_\alpha \le \|A\|_m \|\Delta x\|_\alpha$ 成立，那么就可以通过 $\frac{\|\Delta y\|_\alpha}{\|\Delta x\|_\alpha} \le \|A\|_m$ 来得到估计的相对误差，算子范数则研究怎样才能使该假设的关系成立；

(2) 对于 $\|A\|_m$，它可能有多个不同的选择，也就是说相对误差 $\frac{\|\Delta y\|_\alpha}{\|\Delta x\|_\alpha}$ 的上界是不定的，而在误差估计中无门往往希望上界最小，因此引入了算子范数来来找出所有与向量范数相容的矩阵范数中最小的那一个；

(3) 从代数上看，如果不是误差估计情形，而是压缩映射情形，那么算子范数是满足 $\|Tx\| \le M\|x\|$ 的所有 M 的下确界，也就是说，通过这个方法可以把 θ 找出来，特别注意：上面的所有，前提得是有界线性算子才行。

1.1.13 范数等价

设在 $V_n(P)$ 上定义了 $\|\cdot\|_a, \|\cdot\|_b$ 两种向量范数，若存在常数 $C_1 > 0, C_2 > 0$，使得 $\|x\|_a \le C_1\|x\|_b, \|x\|_b \le C_2\|x\|_a, \forall x \in V_n(P)$，则称 $V_n(P)$ 上的两个向量范数 $\|\cdot\|_a, \|\cdot\|_b$ 等价。

理解：

或定义为：设 $\|\cdot\|_a, \|\cdot\|_b$ 是线性空间 X 上的两个范数，若 $\forall \{x_n\} \subset X$，$\lim_{n\to\infty} \|x_n\|_a = 0 \Leftrightarrow \lim_{n\to\infty} \|x_n\|_b = 0$，则称 $\|\cdot\|_a, \|\cdot\|_b$ 是等价的。

1.1.14 半范

P 是线性空间 X 上的泛函，若 P 满足：(1)$P(\alpha x) = |\alpha| P(x)$；(2)$P(x+y) \le P(x) + P(y)$，则称 P 为空间 X 上的一个半范。

1.1.15 范数诱导范数

设 $\|\cdot\|$ 是 C^m 上的范数，$A \in C_n^{m\times n}$ 为列满秩，则 $\|A\cdot\|$ 是 C^n 上的范数。

1.1.16 范数诱导距离

对赋范空间 $(X, \|\cdot\|)$，若定义 $d(x.y) = \|x-y\|$，则 d 满足距离的三个条件，故在距离 d 下，(X,d) 成为一个距离空间。

理解：

(1) 反过来不一定成立，例如数列空间 R^∞ 按距离 $d(x,y) = \sum_{k=1}^\infty \frac{|x_k - y_k|}{2^k(1+|x_k-y_k|)}$ 不能扩展为赋范空间，因为不满足绝对齐次性【《应用泛函分析》P76 证明，取特殊值举反例】；

(2) 在由范数诱导的距离下，赋范空间中的点列收敛 $x_n \to x$ 等价于 $d(x_n, x) = \|x_n - x\| \to 0 (n\to\infty)$，故也称 $\{x_n\}$ **依范数收敛**于 x（这也是为什么赋范空间等同样能够应用压缩映射原理能的原因）。

1.1.17 内积诱导范数

设 $(X, \langle\cdot,\cdot\rangle)$ 是内积空间，$\forall x \in X$，定义 $\|x\| = \sqrt{\langle x, x\rangle}$，则 $\|\cdot\|$ 是 X 上的范数，称之为由内积诱导出的范数（《应用泛函分析》P93 证明，证明范数的 3 个条件即可）。

理解：

(1) 也就是说，每一个内积空间都是线性赋范空间，因为每一个内积空间都定义了一个范数，因此线性赋范空间中存在的一些结论，在内积空间中也成立。

1.1.18 线性算子/线性泛函、零算子/单位算子 (恒等算子)/纯量算子、有界线性算子/无界线性算子/下有界算子、对偶空间 (共轭空间)、二次共轭空间

(1) 设 X, Y 是数域 K 上的线性空间，$T: X \to Y$ 称为是从 X 到 Y 的**线性算子**，若 $\forall \alpha, \beta \in K, x, y \in X$，有 $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$。当 $Y \in K$ 时，称 T 为**线性泛函**；

(2) 若映射 $I: X \to X$ 满足 $Ix = x, \forall x \in X$，则称 I 为 X 上的**单位算子 (恒等算子)**；若映射 $O: X \to Y$ 满足 $Ox = 0, \forall x \in X$，则称 O 为 X 上的**零算子**；若映射 $T: X \to X$ 满足 $Tx = \alpha x, \forall x \in X$，其中 $\alpha \in K$，则称 T 为 X 上的**纯量算子**。它们都是线性算子；

(3) 设 X, Y 是赋范空间，$T: X \to Y$ 是线性算子，若 $\exists k > 0$，使得 $\forall x \in X$，有 $\|Tx\| \le k\|x\|$，则称 T 为**有界线性算子**，否则称为**无界线性算子**。若 $\exists k > 0$，使得 $\forall x \in X$，有 $\|Tx\| \ge k\|x\|$，则称 T 为**下有界线性算子**；

(4) 赋范空间 X 上的有界线性泛函的全体记为 X^*，X^* 称为 X 的**对偶空间 (共轭空间)**；

(5) 设 X 是 Banach 空间，X^* 是 X 的共轭空间，称 $(X^*)^*$ 为**二次共轭空间**，记为 X^{**}。

理解：

(1) 区分“数函”、“泛函”和“算子”的概念，它们分别是“数域 $\to$ 数域”、“函数空间 $\to$ 数域”和“函数空间 $\to$ 函数空间”；

(2) 对于“有界性”而言，从定义中可以看出它是对于赋范空间（同样有 Banach 空间、内积空间、Hilbert 空间等，因为它们都是“进一步”的赋范空间）而言的；

(3) 要会证明一个算子是否为线性有界算子。【3.2-1 的 3 题】【参见《应用泛函分析》P121-122 例题、P124 习题 1-4】【3.6-1 的 1 题和 2 题】【3.6-2 的 6 题和 7 题】；

(4) 对偶空间 (共轭空间) 的例子，参【见《泛函分析》课程笔记 P2】。

1.1.19 和、直和

(1) 设 V_1, V_2 是线性空间 V 的子空间，所谓 V_1 与 V_2 的和，是指由所有能表示成 $a_1 + a_2, a_1 \in V_1, a_2 \in V_2$ 的向量组成的子集合，记为 $V_1 + V_2$；

(2) 设 V_1, V_2 是线性空间 V 的子空间，若对 $\forall \alpha \in V_1 + V_2$，有 $\alpha = \alpha_1 + \alpha_2 \quad (\alpha_1 \in V_1, \alpha_2 \in V_2)$，且这种表示是唯一的，这个和 $V_1 + V_2$ 就称为直和，记为 $V_1 \oplus V_2$。

(3) 进一步推广，设 $V_1, V_2, ..., V_s$ 是线性空间 V 的子空间，如果和 $V_1 + V_2 + ... + V_s$ 中的每个向量 α 的分解式 $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_s \ \alpha_i \in V_i$ 是唯一的，这个和就称为直和，记为 $V_1 \oplus V_2 \oplus ... \oplus V_s$。

理解：

(1) 直和的判定定理：设 V_1, V_2 是线性空间 V 的子空间，则下列几条命题相互等价：(a)$V_1 + V_2$ 是直和；(b) 零向量表示法唯一，即由 $0 = \alpha_1 + \alpha_2 (\alpha_i \in V_i)$，必有 $\alpha_1 = 0, \alpha_2 = 0$；(c)$V_1 \cap V_2 = \{0\}$；(d)$W = V_1 \oplus V_2 \Leftrightarrow \dim(W) = \dim(V_1) + \dim(V_2)$；

(2) 设 $V_1, V_2, ..., V_s$ 是 V 的子空间，下面几个命题互相等价：(a)$W = \sum V_i$ 是直和；(b) 零向量的表示法唯一；(c)$V_i \cap (\sum_{j \neq i} V_j) = \{0\}, 1 \le i \le s$；(d)$\dim(W) = \sum \dim(V_i)$；

(3) 对于子空间和直和的理解，考虑二维平面 $V_2(C)$，它有两个一维子空间 V_1 和 V_2，且 $V_2(C) = V_1 \oplus V_2$。

1.1.20 零空间 $N(T)$ 或核 $\mathrm{ket}(T)$、值空间 $R(T)$

(1) 设 $T: X \to Y$ 是线性算子，称 $T^{-1}(\{0\}) = \{x \in X : Tx = 0\}$ 为 T 的零空间或核，记作 $N(T)$ 或 $\ker(T)$；

(2) 称 $T(x) = \{y \in Y : y = Tx, x \in X\}$ 为 T 的值空间，记作 $R(T)$。

性质：

(1) 在这里，将算子 T 形象成三个矩阵 $A \in C^{m \times n}, B \in C^{n \times p}$ 和 $C = AB \in C^{m \times p}$，关于值空间常用的性质有：(a) $R(AB) \subset R(A)$，(b)$R(AB) \subset R(A)$ 且 $\mathrm{rank}(AB) = \mathrm{rank}(A)$，则 A 和 AB 是同一个空间，(c) 若 $R(A) = R(B)$，则一定存在 $X \in C^{n \times p}$，使得 $B = AX$ 成立【《矩阵理论》P185 证明】。

(2) 关于零空间 (核空间) 常用的性质有：(a)$N(B) \subset N(AB)$；(b) 若 $\mathrm{rank}(AB) = \mathrm{rank}(B)$，则有 $\dim N(B) = p - \mathrm{rank}(B)$ 和 $\dim N(AB) = p - \mathrm{rank}(AB)$，以及 $N(B) = N(AB)$【《矩阵理论》P185 证明】。

(3) 设 $A \in C^{m \times n}$，则有：(a) $\dim R(A) + \dim N(A^H) = m$；(b)$\dim R(A^H) + \dim N(A) = n$；(c) $C^m = R(A) \oplus N(A^H)$；(d)$C^n = R(A^H) \oplus N(A)$。

1.1.21 图像、闭算子

(1) 设 X, Y 为集合，有映射 $f: D \to Y, D \subset X$，f 的图像为集合 $\text{graph}(f) = \{(x,y) \in X \times Y | x \in D, y = f(x)\}$；

(2) 设 X, Y 是线性赋范空间，$D \subset X$ 是 X 的线性子空间，线性算子 $T: D \to Y$ 称为是闭算子，如果 $\text{graph}(T)$ 是 $X \times Y$ 中的闭集。**或者定义为：** 设 X, Y 均为 Banach 空间，T 是 $D(T) \subset X \to Y$ 的线性算子。对于任意的 $\{x_n\} \subset D(T)$，若由 $x_n \to x, Tx_n \to y$ 可得 $x \in D(T)$，且 $y = Tx$，则称 T 为闭线性算子，简称闭算子。

理解：

每个连续线性算子必是闭算子；但一般的闭线性算子不一定是连续算子。

1.1.22 紧算子

设 X, Y 是赋范空间，若算子 $T: X \to Y$ 把 X 中的有界集都映射成 Y 中的列紧集，则称 T 为紧算子。从 X 到 Y 的紧线性算子的全体记作 $C(X, Y)$，并记 $C(X) = C(X, X)$.

理解：

赋范空间上的有界线性泛函是紧算子。(《应用泛函分析》P177 证明)

1.1.23 Hilbert 空间有界线性算子理论中的共轭算子、自共轭算子、正算子、正交算子、酉算子

(1) 设 X 为 Hilbert 空间，$T: X \to X$ 为有界线性算子，其共轭算子 T^* 定义为 $\langle Tx, y\rangle = \langle x, T^*y\rangle, \forall x, y \in X$；【3.6-2 的 8 题，3.6-3 的 9-11 题】

(2) 若满足 $T = T^*$，则称 T 是自共轭算子 (Hermitian Operator)；

(3) 称自共轭算子 T 是正算子，若满足 $\langle Tx, x\rangle \geq 0, \forall x \in X$（若 $\langle Tx, x\rangle > 0, \forall x \in X$，则称为是严格正)；

(4) 若算子 T 满足 $T^*T = TT^*$，则称 T 是正交算子；

(5) 若算子 T 满足 $T^*T = TT^* = I$，则称 T 是酉算子。

理解：

对共轭的理解，可以参见【3.6-2 的 8 题】的图，求 T 的 T^* 要用 Riesy 表示定理，且建议画图。具体来说，对于 $f(Tx) = (T^*f)(x)$，相当于用 Y^* 中的 f 作用在从 X 映射到 Y 后的元素，等价于将 f 先映射到 X^*，再用这个对应于 X^* 的 f^* 去作用到 x 这个本来的元素，结果相等。有种“对称”和“纠缠”的味道在里面。可联想在知乎上看到的 X 光片的那个例子。在表述上，$\langle Tx, y\rangle = \langle x, T^*y\rangle$ 也是这个道理，只不过是在 Hilbert 空间中更直观的形式。

1.1.24 Hilbert 空间无界线性算子理论中的共轭算子、算子的逆、对称算子

(1) 设 T 是 Hilbert 空间 X 上的无界算子，设 $D(T)$ 为其定义域，$D(T) \subset X$ 稠密，T 的共轭 T^* 定义为 $\langle Tx, y\rangle = \langle x, T^*y\rangle, \forall x \in D(T), y \in D(T^*), D(T^*) = \{y \in X | \langle Tx, y\rangle = \langle x, z\rangle\}, z = T^*y$；

(2)T 是 Hilbert 空间 X 上的线性算子，在 $D(T)$ 上是一对一的，则 T^{-1} 定义在 $D(T) = R(T)$ 上，且满足 $T^{-1}(Tx) = x, \forall x \in D(T)$ 称为 T 的逆；

(3)T 是 Hilbert 空间 X 上的无界算子，称 T 是对称的，如果 T^* 是 T 的延拓，即 $T^* = T$ in $D(T), D(T) \subseteq D(T^*)$，在此情形下，有 $\langle T^*x, y\rangle = \langle x, Ty\rangle, \forall x, y \in D(T^*)$。

1.1.25 半双线性型、双线性型、双线性型是对称的、正的、正定的、二次型、有界线性的、保制的 (X-椭圆的)、实的

(1) Hilbert 空间 X 中的映射 $a(\cdot,\cdot): X * X \to C$ 称为半双线性型，如果满足以下 4 个条件：$a(x_1 + x_2, y) = a(x_1, y) + a(x_2, y)$、$a(\alpha x, y) = \alpha a(x, y)$、$a(x, y_1 + y_2) = a(x, y_1) + \alpha(x, y_2)$、$a(x, \alpha y) = \alpha a(x, y)$，其中 $\alpha \in C$；

(2) 如果还满足条件 $a(x, \alpha y) = \alpha a(x.y)$，则称该映射为双线性型；

(3) 称 $a(\cdot,\cdot)$ 是对称的，如果 $a(x,y)=\overline{a(y,x)}$；

(4) 称 $a(\cdot,\cdot)$ 是正的，如果 $a(x,x)\geq 0$；

(5) 称 $a(\cdot,\cdot)$ 是正定的，如果 $a(x,x)\geq 0$ 且 $a(x,x)=0$ 时 $x=0$；

(6) $F(x)=a(x,x)$ 称为二次型；

(7) 称 $a(\cdot,\cdot)$ 是有界线性的，如果 $\exists M>0$ 使得 $|a(x,y)|\leq M\|x\|\|y\|$；

(8) 称 $a(\cdot,\cdot)$ 是保制的，如果 $\exists\alpha>0$，使得 $a(x,x)\geq\alpha^2\|x\|^2$；(9)$F$ 是实的，如果对 $\forall x\in X$, $F(x)$ 是实的。

1.1.26 一组基底、维数、子空间、生成子空间

(1) 在线性空间 V 中，以向量空间为例，如果有 n 个向量 $\varepsilon_1,\varepsilon_2,...,\varepsilon_n$ 线性无关，而 V 中任意 $n+1$ 个向量线性相关，则称 $\varepsilon_1,\varepsilon_2,...,\varepsilon_n$ 为 V 的一组基底；

(2) 由于线性空间的所有基底总含有相同数目的向量，则 n 称为线性空间 V 的维数；

(3) 如果数域 K 上线性空间 V 的一非空子集 W 对于 V 的两种运算也构成线性空间，则称 W 为 V 的一个线性子空间 (简称子空间)；

(4) 称 $W=\{\beta|\beta=\sum\limits_{i=1}^{r}k_i\alpha_i\}$ 为 V 的生成子空间，其中 $\alpha_1,\alpha_2,...,\alpha_r\in V$（不一定线性无关）。

理解：

(1) **线性相关和线性无关**：如果一组向量中的某一个或多个向量可以由这组向量中的其余向量通过加法或数乘表达（即使 $c_1v_1+c_2v_2+...+c_nv_n=0$ 成立的 c_i 只有平凡解），则该向量组线性相关，否则线性无关；

(2) 上面的概念对于矩阵或函数等空间同样成立。

1.1.27 投影、正交投影、凸集上的投影

(1) 设 $V_n(C)$ 是线性空间，如果线性变换 $T:V_n(C)\to V_n(C)$ 具有 $T^2=T$ 的性质，则称 T 是 $V_n(C)$ 上的投影（也称投影算子或幂等算子）；

(2) 设 T 是 $V_n(C)$ 上的投影，$V_n(C)=R(T)\oplus N(T)$，如果 $R^{\perp}(T)=N(T)$，则称 T 是正交投影；

(3) 设 X 是 Hilbert 空间，$K\subset X$ 是非空闭凸子集，对 $\forall x\in X,z\in K$，若投影 $P_K(x)$ 满足：$\|x-P_k(x)\|_X\leq\|x-y\|_X,\forall y\in K\Leftrightarrow\|x-z\|_X=\inf\|x-y\|_X\Leftrightarrow\langle x-z,x-z\rangle_X\leq\langle x-y,x-y\rangle,\forall y\in K$，则称 $P_K(x)$ 为 x 到 K 的投影。

性质：

(1) 投影的值域和核互为直和补，即 $V_n(C)=R(T)\oplus N(T)$【《矩阵理论》P35 证明】；

(2) 设 $V_n(C)=V_1\oplus V_2$，则存在投影 T，使 $R(T)=V_1,N(T)=V_2$【《矩阵理论》P35 证明】；

(3) 设 V_1 是 $V_n(C)$ 的任一子空间，$V_n(C)=V_1\oplus V_1{}^{\perp}$，若定义 $T(\alpha)=T(\alpha_1+\alpha_2)=\alpha_1$, 其中 $\alpha\in V_n(C),\alpha_1\in V_1,\alpha_2\in V_1{}^{\perp}$, 则 T 是正交投影。

理解：

考虑用二维平面来理解投影，它有两个一维子空间 V_1,V_2，即该二维平面的两个坐标轴，那么对于每一个向量 α 可唯一地表示为 $\alpha=\alpha_1+\alpha_2,\alpha_1\in V_1,\alpha_2\in V_2$。投影算子 T 则定义为 $T(\alpha)=T(\alpha_1+\alpha_2)=\alpha_1$。如此一来，$T$ 把平面上的点沿着坐标轴 V_2 的方向投影到坐标轴 V_1 上，并把 V_1 上的点变为 V_1 上的同一个点，因此 T 是限制在 V_1 上的一个恒等变换，且 $T^2=T$（上述例子实际上还是正交投影的直观理解）。

1.1.28 正交的、正交补 (子空间)

(1) $x,y\in H$ 称为**向量与向量正交**，若 $\langle x,y\rangle=0$，记为 $x\perp y$；$x\in H$，非空子集 $A\subset H$ 称为**向量与集合正交**，若 $\langle x,y\rangle=0,\forall y\in A$，记为 $x\perp A$；$A,B\subset H$ 称为**集合与集合正交**，若 $\langle x,y\rangle=0,\forall x\in A,y\in B$，记为 $A\perp B$；

(2) $A\subset H$，所有正交于 A 的向量之集记为 $A^{\perp}$，称为 A 的正交补 (子空间)，即 $A^{\perp}=\{x\in H|\langle x,y\rangle=0,\forall y\in A\}$，且满足 $H=A\oplus A^{\perp}$。

理解：

(1) 对于正交的概念，前提是定义在内积空间中，因为从内积空间开始才引入了角度等概念；(2)$A^{\perp\perp} = (A^{\perp})^{\perp}$。

1.1.29 向量集是正交的、正交基、标准正交基、标准正交系是极大的 (完备的)

(1) 内积空间 X 中的向量集 $\{u_i\}$ 称为是正交的，如果 $\langle u_i.u_j\rangle = \delta_{ij}$；

(2) 内积空间 X 中的正交向量集 $\{\psi_i\}$ 称为是一组正交基，如果对 $\forall u \in X, \exists \alpha_i$，使得 $u = \sum_{i=1}^{\infty} \alpha_i \psi_i$；

(3) 如果 ψ_i 是标准正交的，称 $\{\psi_i\}$ 为标准正交基；

(4) Hilbert 空间中的一组标准正交基 $\{\psi_i\}$ 称为是极大的 (完备的)，如果不存在单位向量 ψ_0，使得 $\{\psi_0, \psi_1, ..., \psi_i, ...\}$ 是一个正交基，也就是说 $\{\psi_i\}$ 是完备的 $\Leftrightarrow \langle\psi_0, \psi_i\rangle = 0 \to \psi_0 = 0$。

理解：

(1) **标准正交基**指一组正交基中的基向量的模长都是单位长度 1。

1.1.30 预解算子或预解式、正则值、正则集 $\rho(T)$、谱点、谱 $\sigma(T)$

设 X 是复 Banach 空间，T 是 X 到其自身中的有界线性算子，即 $T \in B(X)$,I 是 X 中的单位算子，$\lambda \in C$:

(1) 若 λ 使得 $T - \lambda I$ 具有定义在整个空间 X 上的逆算子 $R_\lambda = (T - \lambda I)^{-1}$（称作算子 T 的预解算子或预解式)，则称 λ 为算子 T 的正则值。正则值的全体记作 $\rho(T)$，称为 T 的正则集；

(2) 若 λ 不是 T 的正则值，则称为 T 的谱点。谱点的全体记作 $\sigma(T)$，称为 T 的谱。

理解：

上述概念其实是从解各种方程的过程中提炼出来的，例如线性方程组 $Ax - \lambda x = y$，其中 A 是 K 中的 n 阶方阵，$\lambda \in K, x, y$ 是 K^n 中的 n 维列向量；线性微分方程 $\frac{d}{dt}x(t) - \lambda x(t) = f(t)$；Fredholm 积分方程 $\int_a^b K(s,t)x(t)dt - \lambda x(s) = \varphi(s)$。这些方程都可以抽象成算子方程 $Tx - \lambda x = y$ 或 $(T - \lambda I)x = y$，T 是 K 上某个赋范空间 X 到自身的线性算子。于是，研究上述各种方程解的存在性、唯一性和稳定性问题均可归结于研究对任意给定的 $y \in X$，算子方程的解 x 的存在性、唯一性和稳定性。进一步地，可以继续将问题转换为研究算子 $T - \lambda I$ 是否有定义在整个空间 X 上的、有界的逆算子，因为算子的逆有意义 (但不一定在全空间 X 上有定义) 等价于算子方程若有解则解唯一、逆算子定义在 X 上等价于 $\forall y \in X$，算子方程有解、逆算子有界等价于方程的解稳定。

1.1.31 算子的特征值、算子的特征向量

设 X 是复 Banach 空间，$T \in B(X)$，若复数 λ 使得方程 $(T - \lambda I)x = 0$ 有非零解 $x \neq 0$，则称 λ 为 T 的一个特征值，x 为 T 的对应于特征值 λ 的特征向量。

理解：

类比矩阵的 $Ax = \lambda x \Rightarrow (A - \lambda I)x = 0$。

1.1.32 点谱 $\sigma_p(T)$、连续谱 $\sigma_c(T)$、剩余谱 $\sigma_r(T)$

谱点可以根据它破坏正则性的方式分为三类:

(1) 若 $(T - \lambda I)^{-1}$ 不存在，即 $\ker(T - \lambda I) \neq \{0\}$，则 λ 为 T 的特征值，其全体构成算子 T 的点谱，记作 $\sigma_p(T)$；

(2) 若 $(T - \lambda I)^{-1}$ 存在，其定义域不是 X 但在 X 中稠密，即 $(T - \lambda I)(X) \neq X, \overline{(T - \lambda I)(X)} = X$，这样的复数 λ 的全体称为 T 的连续谱，记作 $\sigma_c(T)$；

(3) 若 $(T - \lambda I)^{-1}$ 存在，且其定义域在 X 中不稠密，即 $\overline{(T - \lambda I)(X)} \neq X$，这样的复数 λ 的全体称为 T 的剩余谱，记作 $\sigma_r(T)$。

理解：

(1) 对于 X 中的零元素 0，必有 $(T-\lambda I)0=0$，故当 λ 为 T 的特征值时，$T-\lambda I$ 不是单射，从而 $(T-\lambda I)^{-1}$ 不存在，故算子 T 的任一特征值都是谱点，即 $\sigma_p(T)\subset\sigma(T)$。而对有限维赋范空间上的线性算子 $A,\forall\lambda\in C$，只有两种可能，行列式 $|A-\lambda I|=0$，此时 λ 为算子 A 的特征值，或者行列式 $|A-\lambda I|\neq 0$，此时 λ 为算子 A 的正则值。因此在有限维空间中 $\sigma_p(A)=\sigma(A)$，但是在无限维空间中算子的谱并不都是由它的特征值组成；

(2) 有关系 $C=\sigma(T)\cup\rho(T)$ 和 $\sigma(T)=\sigma_p(T)\cup\sigma_c(T)\cup\sigma_r(T)$ 成立；

（3）对于 $(T-\lambda I)x=0$ 而言，$\ker(T-\lambda I)\neq\{0\}$，即说明有非零解。

1.1.33 奇异系统

考虑 $T:X\to Y$ 是紧算子，X,Y 是 Hilbert 空间，有 $TT^*:Y\to Y$。$\{v_j,u_j,\mu_j\}$ 称为 T 的奇异系统，其中 $\{u_j\}$ 是 TT^* 的正交特征向量集，μ_j 称为 T 的奇异值，$Tv_j=\mu_j u_j,T^*u_j=\mu_j v_j$。

理解：

在一些情形，如 $T:X\to Y$ 是紧算子，X,Y 是 Hilbert 空间，则不一定都能像之前一样找到特征值和特征向量，因此在这里考虑奇异值，进而能够去分析问题。

1.1.34 广义函数、广义微商运算、（弱）广义导数

(1) 定义在实验函数空间 $D(\Omega)$ 上的有界线性泛函 (有书上也把这里说成是线性连续泛函，因为定义已经证明了算子的连续性与有界性是等价的) 的全体 $D'(\Omega)$，即若 $f:D(\Omega)\to R^1$ 是广义函数，即 $f\in D'(\Omega)$，则它满足 (a) 线性性：$\forall\varphi_1,\varphi_2\in D(\Omega),\forall\lambda_1,\lambda_2\in R^1$，有 $\langle f,\lambda_1\varphi_1+\lambda_2\varphi_2\rangle=\lambda_1\langle f,\varphi_1\rangle+\lambda_2\langle f,\varphi_2\rangle$（与内积空间“对第二变元共轭线性” 区分）；(b) 对于任意的 $\{\varphi_j\}\subset D(\Omega)$，若按上面定义的拓扑只要 $\varphi_j\to\varphi_0\in D(\Omega)$，都有 $\langle f,\varphi_j\rangle\to\langle f,\varphi_0\rangle(j\to\infty)$。

其中 $\langle f,\varphi\rangle$ 表示 f 在 φ 上的取值，即做运算 $\langle f,\varphi\rangle=\displaystyle\int_\Omega f\varphi dx$。

(2) 称 $D^\alpha=(-1)^{|\alpha|}(\partial^\alpha)^*$ 为 α 阶广义微商运算，即 $\forall f\in D^*(\Omega)$，定义 $\langle D^\alpha f,\varphi\rangle=(-1)^{|\alpha|}\langle f,\partial^\alpha\varphi\rangle,\forall\varphi\in D(\Omega)$；

（3）（弱）广义导数：$\langle Df,\varphi\rangle=-\langle f,\partial\varphi\rangle$。

【3.6-4 的 14-16 题】

理解：

（1）广义函数对微商运算是封闭的，即任意广义函数都可以作任意次广义微商；

（2）广义函数与普通函数相比：（a）对应法则变了。普通函数在定义域和值域之间由映射对应，而广义函数则是由线性性和连续性对应。（b）定义域变了。广义函数的定义域是特殊的函数空间，即 $D(\Omega)$。

1.1.35 Delta 函数

$$\delta(x)=\begin{cases}\infty, & x=0,\\ 0, & x\neq 0.\end{cases}\text{，且有}\int_{-\infty}^{+\infty}\delta(x)dx=1\text{。}$$

1.1.36 连续嵌入

若 X,Y 是两个 Banach 空间且作为集合有包含关系 $X\subset Y$ 并且成立不等式 $\|u\|_Y\le C\|u\|_X,\forall u\in X$，则称 X（连续）嵌入 Y，记为 $X\hookrightarrow Y$。嵌入称为是紧的，当且仅当算子是紧算子。

1.1.37 代数、可交换代数、Banach 代数

(1) 线性空间 A 称为代数，如果乘法有如下 4 个性质：$x(yz)=(xy)z,\forall x,y,z\in A$、$x(y+z)=xy+xz$、$(x+y)z=xz+yz$、$\alpha(xy)=\alpha xy=x(\alpha y),\forall\alpha\in A$；

(2) 若代数 A 满足 $xy=yx$，则称 A 是可交换代数，特别地，若 $e\in A$ 满足 $ex=xr=x$，则称 e 为**幺元**；

(3) 若代数 A 满足 $\|xy\|\le\|x\|\cdot\|y\|$，则称 A 是 Banach 代数。

1.1.38 强收敛 s、弱收敛 ω、弱 * 收敛 ω^*

(1) 设 X 是赋范空间，$\{x_n\}\subset X$，若 $\exists x\in X$，使得 $\{x_n\}$ 依范数收敛于 x, 则称点列 $\{x_n\}$ **强收敛**于 x；

(2) 若 $\exists x\in X$，使对 $\forall f\in X^*$，有 $f(x_n)\to f(x),(n\to\infty)$，则称点列 $\{x_n\}$ 若收敛于 x，x 称为 $\{x_n\}$ 的**弱极限**；

(3) 设 X 是赋范空间，$\{f_n\}\subset X^*$，若 $\exists f\in X^*$，使对 $\forall x\in X, f_n(x)\to f(x)(n\to\infty)$，则称泛函列 $\{f_n\}$ **弱 * 收敛**于 f。

理解：

(1) 在点列的依范数收敛下，无限维赋范空间中的单位闭球不是紧集【《应用泛函分析》P88 定理 3.4.4，针对有限维的赋范空间】，故一些经典的数学分析技巧（即本科数学分析中完备性的 7 个等价定理在无穷维空间中失效历程）派不上用场，因此需要定义较弱的收敛性，以得到更多的结论；

(2) 弱收敛的例子参见《应用泛函分析》P158 例 5.9.1；

(3) 弱收敛与强收敛的关系：(9)$x_n\to x(n\to\infty)\Rightarrow x_n\xrightarrow{\omega}x(n\to\infty)$；(ii) 若 $\dim X=m<\infty$, 则弱收敛与强收敛等价【《应用泛函分析》P160 证明】；

(4) 若将 $\{f_n\}$ 看作是一个算子列，则弱 * 收敛就是算子列的强收敛，只是在不同的场合下习惯的叫法不同罢了。

1.1.39 可微、微分、导数

（1）设 $U\subset R^n$，p 是 U 的一个内点，映射 $f:U\to R^m$。f 称为**在点** $p\in U$ **处可微的**，假若有一个线性映射 $L:R^n\to R^m$（前提是线性映射哦！），使得 $\lim\limits_{h\to 0}\dfrac{1}{|h|}[f(p+h)-f(p)-Lh]=0$。若 f 在 U 的每一点处可微，则 f 称为**在** U **上是可微的**。

（2）假若映射 $f:U\to R^m$ 在点 $p\in U$ 处可微，则上述定义中惟一确定的线性映射 L 称为**映射** f **在点** $p\in U$ **处的微分**，记作 df_p。线性映射 df_p 相对于 R^n 和 R^m 的标准基的矩阵称为**映射** f **在点** $p\in U$ **处的导数**，记作 $f'(p)$。

【《数学分析讲义（第二册）》P99 例 8.1.2，R^n 上（实）二次型 $Q(x)=xBx^T$ 在 x 处的微分是 $dQ_x=2xB$】

拓展：

有时映射 $f:U\to R^m$ 的定义域 U 并非 R^n 中的开集。比如 “U 是 R^n 的某个线性子空间 E 的开集” 或 “U 是 R^n 的某仿射子空间 $E+a+F$ 的开集，其中 F 是 R^n 的某个线性子空间”，可微的定义会稍有变化，参见【《数学分析讲义》（第二册）P98】

理解：

（1）上述定义（1）中的线性映射 L 假若存在，则它是惟一确定的。

（2）直观地理解 $f'(p)$ 定义中相对于标准基的矩阵的含义：设 R^n 的一组基是由向量组 $\{e_1,e_2,...,e_n\}$ 组成的，又设 R^m 的一组基是由向量组 $\{f_1,f_2,...,f_m\}$ 组成的，而线性映射 L 由下式确定：$Le_j=\sum\limits_{i=1}^{m}a_j^i f_i, j=1,2,...,n$（即先得出标准基之间的变换，然后再用它们表示出任意两向量之间的变换）。若 $u=\sum\limits_{j=1}^{n}u^je_j$ 和 $v=\sum\limits_{i=1}^{m}v^if_i$，则有 $Lu=\sum\limits_{i=1}^{m}(\sum\limits_{j=1}^{n}a_j^iu^j)f_i$。则线性映射 L 相对于 R^n 的基和 R^m 的基的矩阵是 $[a_j^i]$=

$$\begin{pmatrix} a_1^1 & a_2^1 & ... & a_n^1 \\ a_1^2 & a_2^2 & ... & a_n^2 \\ ... & ... & ... & ... \\ a_1^m & a_2^m & ... & a_n^m \end{pmatrix}。$$

1.1.40 方向导数、Gateaux 导数、偏导数、泛函梯度

（1）设 $U \subset R^n$ 是开集，$f: U \rightarrow R^m, p \in U, q \in R^n$，且有 $\varepsilon > 0$ 使得 $[p-\varepsilon q, p+\varepsilon q] \subset U$，函数 $g(t) = f(p+tq)$（t 是个极小增量）在区间 $[-\varepsilon, \varepsilon]$ 上有定义。假若 g 在 $t=0$ 处可微，即 $g(t) = g(0) + dg_0 t + r(t), \lim_{t \to 0} \frac{|r(t)|}{t} = 0$（即对于 0 点处而言的，这样更方便）。则称 f 在点 p 处沿 q 方向是可微的。dg_0 称为 f 在 p 处沿 q 方向的方向微分，或称 f **在点 p 处沿 q 方向的方向导数**，记作 $D_q f(p) = \lim_{t \to 0} \frac{f(p+tq)-f(p)}{t} = dg_0(1)$，其中 $dg_0(1)$ 表示线性映射 dg_0 作用在一维单位向量 1 上的值。

（2）假若 f 在点 p 处沿任何方向 $q \in R^n$ 都是可微的，称 f 在点 p 处是 Gateaux 可微的，f 在点 p 处的方向导数也常称为 **Gateaux 导数**。

（3）设 $U \subset R^n$ 是开集，映射 $f: U \rightarrow R^m, p \in U$，则 f 在点 p 处沿 e_k 方向的方向导数（假若存在）称为 f **在点 p 处的第 k 个偏导数**，记作 $D_k f(p) = \frac{\partial f}{\partial x_k}(p) = \lim_{t \to 0} \frac{f(p+te_k)-f(p)}{t}$，其中 e_k 表示第 k 个标准基向量：$e_k = (\delta_{1k}, ..., \delta_{nk})$。

（4）设 F 是 U 上的泛函，对于 $x \in U$, 映射 $x \rightarrow DF(x)$ 称为是 F 的泛函梯度，记为 DF。

性质：

Gateaux 导数具有以下 4 条性质：

（1）$D_{\lambda h} f(x) = \lambda D_h f(x)$；

（2）若映射 f 在点 x 处可微，它将在点 x 处相对于任何向量 h 有方向导数，且 $D_h f(x) = f'(x)h$。由此，在上述条件下，$D_{\lambda_1 h_1 + \lambda_2 h_2} f(x) = \lambda_1 D_{h_1} f(x) + \lambda_2 D_{h_2} f(x)$；

(3) 假设 L 是一个赋范线性空间 F 到赋范线性空间 G 的连续线性映射，则 $D_h(L \circ f)(x) = L \circ D_h f(x)$；

(4) 假若映射 f 在 Banach 空间 E 的开集 U 上 n 次可微，则 $f^{(n)}(x)(h_1, ..., h_n) = D_{h_1} D_{h_2} ... D_{h_n} f(x)$。

理解：

(1)Gateaux 导数是数学分析中方向导数概念的推广，一般要求 X 和 Y 都满足 Banach 空间，它比下面即将介绍的 Frechet 微分更弱；

(2) 对 $x_0 \in \Omega \subset X, h \in X$，定义 $f: \Omega \subset X \rightarrow Y$ 在 x_0 处沿方向 h 的 **n 阶 Gateaux 导数**为 $D^n f(x_0; h) = \frac{d^n}{d\varepsilon^n} f(x_0 + \varepsilon h)|_{\varepsilon=0} = f^{(n)}(x_0) h^n$，它是关于 h 的 n 次齐次函数。更一般地，设 $x_0 \in \Omega \subset X, h_1, h_2, ..., h_n \in X$，定义：$D^n f(x_0; h_1, h_2, ..., h_n) = D_n D_{n-1} ... D_1 f(x_0 + \sum_{i=1}^{n} t_i h_i)|_{t_1 = t_2 = ... = t_n = 0}$，其中算子 $D_i = \frac{\partial}{\partial t_i}$ 之间可交换，可以看出 $D^n f(x_0; h_1, h_2, ..., h_n)$ 是关于 $(h_1, h_2, ..., h_n)$ 是对称的 n 重线性型。

1.1.41 Jacobi 矩阵、Hesse 矩阵

在有限维空间中，因为能找到基，所以以偏导数为元的 Jaccobi 矩阵在有限维多元函数微分中扮演着重要的角色。

(1) 设 $U \subset R^n$ 是开集, 映射 $f: U \rightarrow R^m, p \in U$, 而 $f = \begin{pmatrix} f_1 \\ ... \\ f_m \end{pmatrix} \in R^m$。假若 f 在点 p 处是可微的, 则

$$f'(p) = [D_1 f(p), D_2 f(p), ..., D_n f(p)] = \begin{pmatrix} D_1 f_1(p) & D_2 f_1(p) & ... & D_n f_1(p) \\ D_1 f_2(p) & D_2 f_2(p) & ... & D_n f_2(p) \\ ... & ... & ... & ... \\ D_1 f_m(p) & D_2 f_m(p) & ... & D_n f_m(p) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{x_1}(p) & \frac{\partial f_1}{x_2}(p) & ... & \frac{\partial f_1}{x_n}(p) \\ \frac{\partial f_2}{x_1}(p) & \frac{\partial f_2}{x_2}(p) & ... & \frac{\partial f_2}{x_n}(p) \\ ... & ... & ... & ... \\ \frac{\partial f_m}{x_1}(p) & \frac{\partial f_m}{x_2}(p) & ... & \frac{\partial f_m}{x_n}(p) \end{pmatrix}。$$

上式的矩阵便是用 f 的分量的偏导数表示的 f 的导数，或称为 **Jacobi 矩阵**。

（2）若 f 具有二阶连续偏导，且当 $m=n$ 时，则 $\begin{pmatrix} \frac{\partial^2 f_1}{\partial^2 x_1} & \frac{\partial^2 f_1}{\partial x_2 x_1} & ... & \frac{\partial^2 f_1}{\partial x_n x_1} \\ \frac{\partial^2 f_2}{\partial x_1 x_2} & \frac{\partial^2 f_2}{\partial^2 x_2} & ... & \frac{\partial^2 f_2}{\partial x_n x_2} \\ ... & ... & ... & ... \\ \frac{\partial^2 f_n}{\partial x_1 x_n} & \frac{\partial^2 f_n}{\partial x_2 x_n} & ... & \frac{\partial^2 f_n}{\partial^2 x_n} \end{pmatrix}$ 被称为 **Hesse 矩阵**。显然其对称。

理解：

f 很直观地展示了如何从高维映射到低维。

1.1.42 直线段、凸集合

设 $a, b \in R^n$，记 $[a,b]=\{x\in R^n: \exists\lambda\in[0,1], x=\lambda a+(1-\lambda)b\}$，它被称为**连接点 a 和点 b 的直线段**。集合 $E\subset R^n$ 称为**凸集合**，假若 $\forall a,b\in E,[a,b]\subset E$。

1.1.43 高阶偏导数

设 $f:\Omega\to R$。假若对于某开集 $U\subset\Omega$ 中的一切点 p 偏导数 $D_jf(p)$ 都存在。D_jf 可以看成 Ω 到全体一行一列矩阵组成的线性空间的映射，也就是说，D_jf 可以看成 Ω 到 R 的映射。这样，就有可能讨论在某点 $p\in U$ 处的偏导数 $D_kD_jf=D_k(D_jf)$，其中 $1\le k\le n$。D_kD_jf 称为 f 的**二阶偏导数**。归纳地，我们可以定义 f 的 r **阶偏导数** $D_{j_1}D_{j_2}...D_{j_r}f, j_i\in\{1,...,n\}$。常用 D_j^2f 表示 D_jD_jf。有时也用以下记法表示 r 阶偏导数：$D_{j_1}D_{j_2}...D_{j_r}f=\dfrac{\partial^r f}{\partial x_{j_1}\partial x_{j_2}...\partial x_{j_r}}$。

1.1.44 $C^r(\Omega)$ 类、支集

设 Ω 是开集，映射 $f:\Omega\to R$。f 称为**（属于）$C^r(\Omega)$ 类的**，记作 $f\in C^r(\Omega)$，假若 f 的所有 r 阶偏导数在 Ω 上都存在且连续（当然，这蕴涵了 f 的阶数低于 r 的偏导数在 Ω 上存在且连续）。若对于任何 $r\in N, f\in C^r(\Omega)$，则称 $f\in C^\infty(\Omega)$。$C^0(\Omega)$ 表示 Ω 上连续函数的全体。又记 $C_0^r(\Omega)=\{f\in C^r(\Omega):\operatorname{supp}f\text{紧}\}$，其中 $r\in N\cup\{0\}\cup\{\infty\}$，而 $\operatorname{supp}f$ 表示使（连续）函数 f 不等于零的点全体构成的集合的闭包：$\operatorname{supp}f=\overline{\{x\in\Omega:f(x)\neq 0\}}$。$\operatorname{supp}f$ 称为（连续）函数 f 的**支集**。

向量值函数 $f=(f_1,f_2,...,f_m)$ 称为 C^r 类的，若它的每个分量函数 f_j 是 C^r 类的。

性质：

显然，我们有 $C^0(\Omega)\supset C^1(\Omega)\supset...\supset C^{r-1}(\Omega)\supset C^r(\Omega)\supset...$，且 $C^\infty(\Omega)=\bigcap\limits_{r=0}^{\infty}C^r(\Omega)$。

1.1.45 有限维多元函数的 Taylor 展开公式

设 U 是 R^n 中的开凸集，$f:U\to R$ 是 C^{r+1} 类的函数，$r\ge 0$。若 $p\in U$ 和 $q=p+h\in U$，则有 $\theta\in(0,1)$，使得以下多元函数的公式成立：$f(q)=\sum\limits_{|\alpha|\le r}\dfrac{D^\alpha f(p)}{\alpha!}h^\alpha+\sum\limits_{|\alpha|=r+1}\dfrac{D^\alpha f(p+\theta h)}{\alpha!}h^\alpha$。上式第一项称为**多元函数的 Taylor 多项式**，第二项称为**多元函数的 Taylor 公式的 Lagrange 余项**。**【**《数学分析讲义（第二册)》P122-123 证明**】**

1.1.46 重线性映射（n-线性映射）

设 $E_j(j=1,2,...n)$ 和 F 是实数域 R(或复数域 C)上的 $(n+1)$ 个线性空间,映射 $T:E_1\times...\times E_n\to F$ 称为**重线性（或 n-线性）映射**。若对于任何 $j\in\{1,...,n\}$，当 $x_1,...,x_{j-1},x_{j+1},...,x_n$ 取任何固定的值时，映射 $x_j\mapsto T(x_1,...,x_j,...,x_n)$ 是个线性映射。当 $n=2$ 时, n-线性映射称为**双线性映射**。$E_1\times...\times E_n\mapsto F$ 的 n-线性映射全体记作 $\mathcal{L}(E_1,...,E_n;F)$。

理解：

（1）这个定义将重线性映射的概念推广到无限维线性空间，因为这里的 n 允许是无限的。

（2）在有限维线性空间上有有限个向量组成的标准基，因而映射的微分在标准基下可以用映射的分量的偏导数组成的 Jacobi 矩阵表示，这就是偏导数在传统的微积分中为什么扮演十分重要且方便的角色的缘由。无限维赋范线性空间上的微分学一般不假定空间有基，映射的微分将不可能用映射的分量的偏导数组成的某个无限矩阵表示。以偏导数为元的 Jacobi 矩阵将不再扮演有限维线性空间的微分学中那样重要的角色。

（3）$x_1,...,x_{j-1},x_{j+1},...,x_n$ 中默认不能有 0 哦，因为只要一个为 0，则映射为 0。

1.1.47 重线性映射的范数、有界重线性映射

设 $E_1,...,E_n$ 和 F 是 $(n+1)$ 个线性赋范空间，我们定义**重线性映射** $A\in\mathcal{L}(E_1,...,E_n;F)$ **的范数**为 $\|A\| = \sup\limits_{x_j\in E_j,|x_j|_{E_j}\neq 0,j=1,...,n}\dfrac{|A(x_1,...,x_n)|_F}{|x_1|_{E_1}...|x_n|_{E_n}}$。若 $\|A\|<\infty$，A 称为**有界重线性映射**。$\mathcal{L}(E_1,...,E_n;F)$ 中全体有界 n-线性映射组成的集合记作 $L(E_1,...,E_n;F)$。

从上述定义中不难看出：$|A(x_1,...,x_n)|_F\le\|A\||x_1|_{E_1}...|x_n|_{E_n}$ 和 $\|A\| = \sup\limits_{x_j\in E_j,|x_j|_{E_j}=1,j=1,...,n}|A(x_1,...,x_n)|_F$。

【1.1-4 的 16 题例子】

理解：

定义范数的目的就是判断是否有界，与泛函中学的有限维情形一样。

1.1.48 Frechet 导数、广义梯度 (次微分)、临界点

在无限维多元函数微分中，导数不再像有限维多元函数微分那样是个数，而是用算子来表示的。

(1) 设 E 和 F 为两个赋范线性空间，$G\subset E$。映射 $f:G\to F$ 称为在 G 的内点 x 是可微的，若有连续线性映射（或有界线性映射）$L(x)\in L(E;F)$，使得 $f(x+h)-f(x)=L(x)h+\alpha(x;h)$，其中 $\alpha(x;h)=o(h)$。确切些说，当 $h\to 0, x+h\in G$ 时，$|\alpha(x;h)|=o(|h|)$。若 G 是 E 的开集，映射 $f:G\to F$ 在 G 的任何点都可微，则称映射 f 在 G 上可微。【3.6-3 的 12 题，3.6-4 的 13 题】

上述定义中的 $L(x)\in L(E;F)$（假若存在）称为映射 $f:G\to F$ 在 G 的内点 x 处的微分，或切映射，或导数。$L(x)$ 常被记作 $df_X, Df(x)$ 或 $f'(x)$，称为 **Frechet 导数（强导数）**。（因为无穷维的线性赋范空间的研究通常不借助于它的一组基，所以不存在线性映射的矩阵，我们不再区分导数和微分）。【1.1-5 例子，求 G 导数和 F 导数】

(2)$F:X\to R$ 在 $x\in X$ 的广义梯度记为 $\partial F(x)$, $\partial F(x)=\{F'\in X^*|F'(h)\le F(x+h)-F(x), h\in X\}$, $\partial F(x)$ 中的元素 F' 称为是 X 处支撑泛函的**次梯度**；

(3) 设 $f:E\to F$，若存在点 $x_0\in E$，使得 f 的 Frechet 微分 $f'(x_0)=0$，则称 x_0 是 f 的一个**临界点**。

拓展：

有时映射 $f:G\to F$ 的定义域 G 并非 E 中的开集，比如“U 是 G 的某线性子空间 E 的开集”或“U 是 G 的某仿射子空间 $E=a+H$ 的开集，其中 H 是 G 的某线性子空间”，可微的定义会稍有变化，参见【《数学分析讲义（第二册)》P188-189】

性质：

和有限维的情形一样，无穷维赋范线性空间上的微分（求导）运算也有以下规则：

（1）微分（求导）运算是线性的；

（2）复合映射的求导规则（锁链法则）成立：设 E,F 和 K 是三个线性赋范空间，G 和 H 分别是 E 和 F 的开集。若映射 $f:G\to H$ 在点 $x\in E$ 处可微，而映射 $g:H\to K$ 在点 $f(x)=y\in H\subset F$ 处可微，则复合映射 $g\circ f:G\to K$ 在点 x 处可微，且 $(g\circ f)'(x)=g'(f(x))\circ f'(x)$;

（3）逆映射的求导规则：设 $f:G\to F$ 是一个在点 $x\in G$ 处连续的映射，而且有一个在点 $y=f(x)$ 的一个邻域 H 内有定义的 f 的逆映射 $f^{-1}:H\to E$，它在点 $y=f(x)$ 处连续。若映射 f 在点 x 处可微，且它的切映射 $f'(x)\in L(E;F)$ 有连续的（或称为有界的）逆映射 $[f'(x)]^{-1}\in L(F;E)$，则映射 f^{-1} 在点 $y=f(x)$ 处可微，且 $[f^{-1}]'(f(x))=[f'(x)]^{-1}$。（该规则也可理解为：(a)$f:G\to H$ 为同胚映射；(b)$f'(x)$ 为 $E\to F$ 的线性同胚）

【《数学分析讲义（第二册)》P190-192 例子】

理解：

(1) 对于定义（1），我们一般讨论开集，因为边界条件有时很麻烦。$L(x)$ 是连续线性（或有界线性）的，这是两个很关键的前提条件。

(2) 上述内容的价值在于线性近似，即 $S(x+t)=S(x)+T(t)+E$, 其中 E 是一个误差项，它用于非线性优化中，可以用这个线性近似去逼近。也就是说，其目的是用线性映射 L 去代替非线性映射 f，是个

线性化的过程。对 T 的模长控制越小越好，但在实际应用中不能太小，否则收敛速度很慢，一个例子就是 $\sin(0.1)=\sin(0)+\sin' 0\cdot(0.1)+\text{error}$；

(3)Frechet 导算子具有“唯一性”、“线性性”和“链式法则”等简单的性质。

1.1.49 偏映射

因为一般的无限维线性空间未必有合适的基，以偏导数为元的 Jacobi 矩阵在无限维线性空间上的微分学中不再扮演有限维线性空间上的微分学中那样重要的角色。但是当无限维线性空间是给定的 m 个赋范线性空间的笛卡尔积时，偏导数概念仍然有用。

设 $E_j(j=1,...,m)$ 是 m 个线性赋范空间，$E=E_1\times...\times E_m$ 是它们的笛卡尔积。设 $U=U(a)$ 是点 $a=(a_1,...,a_m)\in E$ 在 E 中的一个邻域，$f:U\to F$ 是 U 到赋范线性空间 F 的一个映射。当 $a_1,...,a_{j-1},a_{j+1},...,a_m$ 固定时，映射 $\varphi_j:E_j\to F$ 定义为 $\varphi_j(x_j)=f(a_1,...,a_{j-1},x_j,a_{j+1},...,a_m)$，称为映射 $f:U\to F$ 在 $a=(a_1,...,a_m)$ 处关于变量 x_j 的**偏映射**。假若偏映射 $\varphi_j:E_j\to F$ 在点 $x_j=a_j$ 处可微，它在该点的导数称为**映射 $f:U\to F$ 在点 $a=(a_1,...,a_m)\in E$ 处关于变量 x_j 的偏导数（或偏微分）**，记作 $\partial_j f_a=D_jf(a)=\dfrac{\partial f(a)}{\partial x_j}=f'_{x_j}(a)=\varphi'_j(a_j)$。

1.1.50 高阶导数

设 U 是赋范线性空间 E 中的开集，又设映射 $f:U\to F$ 是从 U 到 Banach 空间 F 的一个可微映射，则 f 的导数 f' 是 U 到 $L(E;F)$ 的映射：$x\in U\mapsto f'(x)\in L(E;F)$。已知 $L(E;F)$ 也是个 Banach 空间，若 $f':U\to L(E;F)$ 在 $x\in U$ 处是可微映射，则它在 x 处的导数，称为 **f 在 x 处的二阶导数（或二阶微分）**，记作 $d^2f_x=f^{(2)}(x)=f''(x)=(f')'(x)\in L(E;L(E;F))$。若 $f':U\to L(E;F)$ 在 U 上是可微映射，则记 $d^2f=f^{(2)}=f''=(f')':U\to L(E;L(E;F))$。

映射 $f:U\to F$ 在点 x 处的 n 阶导数（或 n 阶微分）定义为映射 $f:U\to F$ 的 (n-1) 阶导数在 x 点处的导数或切映射（假若后者存在的话）：$f^{(n)}(x)=(f^{(n-1)})'(x)$。显然，若 $f^{(n)}(x)$ 有定义，则 $f^{(n)}(x)\in L(E;L(E;...;L(E;F)))$，其中右端的 L 出现了 n 次，因为 E 也出现 n 次。由 4.9 的定理 4，$f^{(n)}(x)$ 是 E 到 F 的 n-线性映射：$f^{(n)}(x)\in L(E,E,...,E;F)$（n 个 E），而 $f^{(n)}$ 是如下的映射：$f^{(n)}:U\to L(E,E,...,E;F)$，$f^{(n)}:x\mapsto f^{(n)}(x)$。

1.1.51 无限维多元函数的 Taylor 展开公式

记号约定见【《数学分析讲义》（第二册）P198-199】。

假设 E 和 F 是两个线性赋范空间，U 是 E 的一个开集，映射 $f:U\to F$ 在 U 上有直到 $(n-1)$ 阶的导数，而在点 $x\in U$ 处有 n 阶导数 $f^{(n)}(x)\in L(E,...,E;F)$，则当 $h\to 0$ 时，有 $f(x+h)=f(x)+f'(x)h+...+\dfrac{1}{n!}f^{(n)}(x)h^n+\circ(|h|^n)$。称为**无限维空间中的带 Peano 余项的 Taylor 公式**。【《数学分析讲义（第二册）》P199 证明】

1.1.52 Fourier 展开

设 $\{\psi_i\}$ 是内积空间 X 中的一组标准正交基（即 $\{\psi_i\}$ 是一个标准正交系），$\forall u\in X$，称级数 $u=\sum\limits_{i=1}^{\infty}\alpha_i\psi_i$ 为 u 关于标准正交系 $\{\psi_i\}$ 的 **Fourier 级数**，其中 $\alpha_i=\langle u,\psi_i\rangle$ 为 u 关于 ψ_i 的 **Fourier 系数**。

理解：

前提是按一组标准正交系进行展开，因此对于一线性无关的向量组，首先应当对其进行 **Schmit 正交化**。【《应用泛函分析》P101 和《矩阵分析》P86 有形象的理解】

1.1.53 三角函数系、傅里叶级数

(1) 在闭区间 $[-\pi,\pi]$ 上的平方可和函数空间 $L_2[-\pi,\pi]$ 是完备的无穷维欧几里得空间，即希尔伯特空间。诸函数 $1,\cos x,\sin x,...,\cos nx,\sin nx$ 称为**三角函数系**。

（2）对于每一个函数 $f \in L_2[-\pi,\pi]$，**傅里叶级数** $\dfrac{a_0}{2}+\displaystyle\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$ 均方收敛于 f，也就是在空间 $L_2[-\pi,\pi]$ 的度量下收敛于 f。其中 $a_n=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(x)\cos nxdx, b_n=\dfrac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nxdx, a_0=\dfrac{1}{\pi}\int_{-\pi}^{\pi}f(x)dx$。

性质：

三角函数系具有正交性，即 $\displaystyle\int_{-\pi}^{\pi}\cos nxdx=0,\int_{-\pi}^{\pi}\sin nxdx=0,\int_{-\pi}^{\pi}\cos mx\cos nxdx=\begin{cases}\pi, & m=n,\\ 0, & m\neq n.\end{cases}$，$\displaystyle\int_{-\pi}^{\pi}\sin mx\sin nxdx=\begin{cases}\pi, & m=n,\\ 0, & m\neq n.\end{cases}$，$\displaystyle\int_{-\pi}^{\pi}\sin mx\cos nxdx=0$。（由积化和差公式可得）

理解：

在实际应用中，我们需要确定保证傅里叶级数不仅仅是平均收敛于 f，而且在已知点、处处收敛或者甚至是一致收敛于 f 的条件是重要的。在 4.10 中的定理 1 建立了三角级数在已知点收敛的充分条件，且需要用到下面的 1.1.54 的概念。

1.1.54 狄利克雷积分、狄利克雷核

（1）表达式 $S_n(x)=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}f(t)\frac{\sin\frac{2n+1}{2}(t-x)}{2\sin\frac{t-x}{2}}dt$ 及它的各种变形称为**狄利克雷积分**；【4.10-1 的 1 题推导】

（2）对迪狄利克雷积分继续进行化简，令 $t-x=z$，有 $S_n(x)=\dfrac{1}{\pi}\displaystyle\int_{-\pi-x}^{\pi-x}f(x+z)\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}dz=\dfrac{1}{\pi}\int_{-\pi}^{\pi}f(x+z)\frac{\sin\frac{2n+1}{2}z}{2\sin\frac{z}{2}}dz$。函数 $D_n(z)=\dfrac{1}{2\pi}\dfrac{\sin\frac{2n+1}{2}z}{\sin\frac{z}{2}}$ 称为**狄利克雷核**。

性质：

显然，有 $\displaystyle\int_{-\pi}^{\pi}D_n(z)dz=1,\forall n\in N^+$。

1.1.55 费耶和、费耶积分、费耶核

由于连续函数的傅里叶级数一般来说不一定收敛，这样的函数 f 不能由直接求它的傅里叶级数和来得到，因此费耶给出了由一个连续函数的傅里叶级数得到这个连续函数的方法，即 4.10 中的定理 3。在此之前需要先熟知以下几个概念：

（1）设 $S_k(x)=\dfrac{a_0}{2}+\displaystyle\sum_{j=1}^{k}a_j\cos jx+b_j\sin jx$ 是函数 f 的傅里叶级数的部分和。

令 $\sigma_n(x)=\dfrac{S_0(x)+S_1(x)+...+S_{n-1}(x)}{n}$。表达式 σ_n 称为函数 f 的**费耶和**；

（2）可以将上述表达式表示成**费耶积分** $\sigma_n(x)=\dfrac{1}{2n\pi}\displaystyle\int_{-\pi}^{\pi}(\frac{\sin n\frac{z}{2}}{\sin\frac{z}{2}})^2f(x+z)dz$；【4.10-2 的 5 题推导】

（3）表达式 $\Phi_n(z)=\dfrac{1}{2n\pi}(\dfrac{\sin n\frac{z}{2}}{\sin\frac{z}{2}})^2$ 称为**费耶核**。

性质：

（1）显然，$\Phi_n(z)\geq 0$；

（2）$\displaystyle\int_{-\pi}^{\pi}\Phi_n(z)dz=1$；【4.10-2 的 7 题证明】

（3）对于任意固定的 $\delta>0$，当 $n\to\infty$ 时，有 $\displaystyle\int_{-\pi}^{-\delta}\Phi_n(z)dz=\int_{\delta}^{\pi}\Phi_n(z)dz=\eta_n(\delta)\to 0$。【4.10-3 的 8 题证明】

1.1.56 傅里叶积分、复数形式的傅里叶积分

（1）傅里叶积分公式：$f(x)=\dfrac{1}{\pi}\displaystyle\int_0^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)\cos\lambda(t-x)dt$。引入记号 $a_\lambda=\dfrac{1}{\pi}\displaystyle\int_{-\infty}^{\infty}f(t)\cos\lambda tdt, b_\lambda=\dfrac{1}{\pi}\int_{-\infty}^{\infty}f(t)\sin\lambda tdt$，可以将傅里叶积分写成傅里叶级数的形式：$f(x)=\displaystyle\int_0^{\infty}(a_\lambda\cos\lambda x+b_\lambda\sin\lambda x)d\lambda$。

【4.10-3 的 10 题推导】

（2）复数形式的傅里叶积分公式：$f(x)=\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}d\lambda\int_{-\infty}^{\infty}f(t)e^{-i\lambda(t-x)}dt$。【4.10-4 的 11 题推导】

1.1.57 傅里叶变换

设 $f(x)\in L_1(-\infty,+\infty)$，则：

$g(\lambda)=\displaystyle\int_{-\infty}^{\infty}f(t)e^{-i\lambda t}dt$ 称为**傅里叶变换公式**；

$f(x)=\dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}g(\lambda)e^{i\lambda x}d\lambda$ 称为**反演公式**。

如果用公式 $g(\lambda)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty}f(t)e^{-i\lambda t}dt$ 确定了 g，就可以使公式变得更对称，那么反演公式取形式为：$f(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty}g(\lambda)e^{i\lambda x}d\lambda$。

性质：

（1）卷积性质：$f,g\in L(-\infty,+\infty)$, 定义卷积运算为 $(f*g)(t)=\int_{-\infty}^{\infty}f(t-u)g(u)du$。则有 $F[f*g]=F[f]\cdot F[g]$；

（2）$\{f_n\}\in L_1(-\infty,+\infty)$，$\{f_n\}$ 在 $L_1(-\infty,+\infty)$ 的度量收敛于 $f(x)\in L_1$，则傅里叶变换序列 $g_n=F[f_n]$ 在整个直线上 (R 上) 是一致收敛的；

（3）线性性质：$\forall\alpha,\beta\in R,f,g\in L_1$，有 $F[\alpha f+\beta g]=\alpha F[f]+\beta F[g]$；

（4）连续性与有界性：若 $f\in L_1$，则 $F[f]$ 是有界连续函数，且 $\lim\limits_{|\lambda|\to\infty}F[f]=0$；

（5）时移性质：$f\in L_1$，则 $F[f(x-x_0)]=e^{-i\lambda x_0}F[f]$；

（6）平移性质：$F[e^{i\omega_0 x}f(x)](\lambda)=F[f](\lambda-\omega_0)$；

（7）尺度性质：$a\neq 0,F[f(ax)](\lambda)=\dfrac{a}{|a|}F[f](a\lambda)$；

（8）微商性质：若 $f\in L_1$ 且 $f'\in L_1$，f 在每一个有限区间上绝对连续（比一致连续还强），则 $F[f']=i\lambda F[f]$。类似地，$F[f^{(k)}]=(i\lambda)^kF[f]$；

（9）$\dfrac{|F[f^{(k)}]|}{|\lambda|^k}=F[f](\lambda)$。反映了函数光滑程度与傅里叶变换在无穷远处减小的速度之间的联系，即 $F[f]$ 在无穷远处比 $\dfrac{1}{|\lambda|^k}$ 下降得更快；

（10）若 f' 存在，且 $f''\in L_1$ 及 $\lim\limits_{x\in\pm\infty}f'(x)=0$，则 $F[f]$ 绝对可积；

（11）若 $f(x),xf(x)\in L_1$，则 $F[f](\lambda)=g(\lambda)$ 可微，且 $g'(\lambda)=F[-ixf(x)]$。更一般地，若 $f(x),xf(x),...,x^nf(x)\in L_1$，则 $g^{(k)}(\lambda)=F[(-ix)^kf(x)]$。

【证明参见任何一本数学物理方程书籍，以及参考资料【2】，这里的内容有扩展】

理解：

（1）傅里叶变换只是将时域与频域相互变换的众多 transforms 中的一种。这类 transforms 变换后得到的频域上的谱。傅里叶反演公式的意义在于可以将这些分析学的运算转换为代数运算。

（2）傅里叶变换得到的特征值表示 frequencies，特征向量 $\{e^{-i\lambda t}\}$ 表示 Fourier basis。类比图傅里叶变换。

（3）至于为什么我们需要傅里叶变换，或更进一步说为什么我们需要积分变换？可以参见【9】的 P116 起的傅里叶变换的应用和 P133 起的拉普拉斯变换的应用。简而言之就是简化积分计算。

1.1.58 【测度和勒贝格 (Lebesgue) 积分的相关概念】

以下内容部分来自知乎链接：https://zhuanlan.zhihu.com/p/343599708。

（1）**扩展的实数集** $\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\}=[-\infty,+\infty]$ 是为了更方便地处理无穷而定义的，由此可以规定一些无穷的运算。以前总是强调不要把无穷大当成一个数来运算和比较大小，但是在这里我们有必要做出上面的规定，因为这样可以简化很多符号和陈述，帮助我们抓住很多问题的重点（例如定义非负级数时，如果级数发散，则和为 $+\infty$，因此，不管两个非负级数是否收敛，不等式 $\sum_n a_n\le\sum_n b_n$ 都有意义，其含义是如果它俩均收敛，则 a_n 的和小于或等于 b_n，并且如果前者发散，则后者发散。所以当后者发散

时，无论 a_n 取何值，这个不等式都成立）【对**级数**的理解参见阿基里斯追乌龟的悖论，即：可数多个数相加的和，详情见最上面的知乎链接】。

（2）**芝诺悖论**：

先从直观上尝试定义子集的**长度**。直觉上，$[0,2]$ 这个区间的长度为 2，集合 $(0,1)\cup(5,8)$ 的长度为 $1+3=4$。如果我们有可数多个互不相交的区间 $(I_j)_{j=1}^{\infty}$，可能是开区间、闭区间或半开半闭区间，接着上述的直观上尝试，$l(\bigcup_{j=1}^{\infty} I_j)=\sum_{j=1}^{\infty}(b_j-a_j)=\sum_{j=1}^{\infty} l(I_j)$。顺便提一下，空集 $\emptyset=(0,0)=(a,a)$ 也是一个开区间，$l(\emptyset)=0$。那么，对于单个点，比如 $\{0\}$，或若干个点，比如 $\{1,2,3\}$，它们的长度又是多少呢？因为至今都还没严格定义（即后面要讲的勒贝格测度）"长度"的概念，所以现在我们只是在寻找合理的定义方式，尽可能地捕捉对长度的直观感受，只是要注意做了什么样的假设，结论在怎样的前提下成立。稍微想一下会发现，$\{0\}=[0,0]$ 的长度为 0，根据平移不会改变长度（假设之一），所以若干点集合的长度也是 0。但是，问题来了，对于区间 $[0,1]$，它不也是无穷多个点构成的吗？每个点的长度都是 0，为什么"加起来"之后 $[0,1]$ 的长度是 1?

关于该悖论详情见最上面的知乎链接系列之二，简而言之就是：（1）**阿基里斯追乌龟**。位置是时间的函数，"追不上"仅限 $0<t<1$，这个版本的悖论在允许可数多个长度/时间相加后就解决了，前面我们就是用可数多个数相加得出 $\mathcal{N}$ 的"长度"的，用这种方法可以计算很多集合的长度。这个例子说明一个实数的小数表示可能是不唯一的。小数是表示实数的一种方法，而不是一个实数的"名字"，实数 $1/2=0.5=0.49999...$ 就有若干个不同的名字。（2）让无穷多个数相加解决了阿基里斯追乌龟悖论，却引发了**飞矢不动悖论**，也就是上一段落的最后一个问题。但是，前面我们已经强调了结论**前提**的重要性！"飞矢不动"的结论基于一个假设，就是无穷多个 0 加起来等于 0，但是之前我们只定义了可数多个实数的和（即级数的和）。所以问题是，$[0,1]$ 是不可数的，所以这个上述逻辑自然也不成立了。

良好定义的概念参见上述知乎链接。所以说，我们上述从直觉上定义的"长度"（既允许使用不可数个实数的和，又允许使用不可数个区间的并）就不是个良好定义了。为了避免这个问题，我们不引入不可数个实数的求和，并且只用可数多个区间**覆盖**（在拓扑中的体现参见1.1.9节的 (13)）。即 $[0,1]$ 不可以写成可数多个点的并集，因此用可数多个区间覆盖，给可数多个数求和得到的"长度"不会以同样的方式导致 $[0,1]$ 的长度为 0 的悖论（这只是解决方式中的一种，但这里仿佛明白了拓扑中讲覆盖的意义了）。

那么，芝诺的飞矢不动悖论到底该怎么解释呢？其实该悖论有一些假设（一个区间是由很多个点构成的；一个集合的长度是实数；每个集合都有长度；一个集合的长度在经过平移之后不变；长度无限可加等等）。在严格定义"长度"之后，这些性质中的一部分会得到证明。简而言之的**思想**就是：放弃或弱化上面的一条或几条假设，就可以很大程度上避免矛盾或者反直觉的情况发生，即如果要更好地满足其中一条，就要更多地放弃另一条。

（3）如前面所述，要测量 A 的"长度"。如果一些开区间 $(a_1,b_1),(a_2,b_2),...$ 覆盖了 A，即 $A\subseteq\bigcup\limits_{j=1}^{\infty}(a_j,b_j)$，则 A 的"长度"不能超过这些区间的总长度。我们可以想办法让这些区间的总长度尽可能地小，从而得到比较好的**近似**。理想情况下这个覆盖的总长度所能取得的最小值应该就是我们想要的长度，但是这个最小值不一定存在，因此我们用**下确界**给出如下对"长度"的定义（仿佛理解到了"测度"的含义，就是如何去测准上述直观上定义的"长度"）。

直观上来看，定义任何一个集合 $A\subseteq\mathbb{R}$ 的"Lebesgue 外测度"为 $\mu^*(A)=\inf\{\sum\limits_{j=1}^{\infty}(b_j-a_j):X\subseteq\bigcup\limits_{j=1}^{\infty}(a_j,b_j)\}$，其中 $\mu^*:2^X\to\overline{\mathbb{R}}_+$。即一组开区间覆盖的"极小"长度。我们允许用无穷可数多个区间覆盖。

（4）至此，上述引入给出了对"测度"的直观理解。然而，构造一个测度，或证明具有某种性质的测度的存在，远非如想象的那样简单。为了构造一个 X 上的 σ 代数上的具有可数可加性质的非负集合函数，也即测度，我们先退一步，引进一个定义在 2^X 上满足比可数可加性要若的条件（称为次可数可加性）的非负集合函数，我们把这样的集合函数叫做外测度，定义如下：

设 X 是个非空集合，具有以下性质的映射 $\mu^*:2^X\to\overline{\mathbb{R}_+}$ 称为 X 上的一个**外测度**：（1）$\mu^*(\emptyset)=0$；（2）$A\subset B$，则 $\mu^*(A)\le\mu^*(B)$；（3）$\forall n\in\mathbb{N},A_n\subset X$，则 $\mu^*(\bigcup\limits_{n=1}^{\infty}A_n)\le\sum\limits_{n=1}^{\infty}\mu^*(A_n)$。其中条件（3）称为**次可数可加性**，由条件（1）和（2）知外测度的取值必为非负。【**一个构造外测度的常用方法**参见【1】的 P235

的引理 9.3.1。上述（3）中最后的那个式子称为 **Lebesgue 外测度**，详情参见【1】的 P236】

但所构造的外测度 μ^* 与用于构造 μ^* 的 λ 之间的关系仍很不清楚。确切些说，我们不知道有没有以下关系：$\forall A\in\mathcal{C},\mu^*(A)=\lambda(A)$。假若有这个关系，则 μ^* 是定义在 $\mathcal{C}$ 上的 λ 到 2^X 上的一个延拓。我们不知道什么样的条件可以保证 μ^* 是定义在 $\mathcal{C}$ 上的 λ 到 2^X 上的一个延拓。

(5) 但是上述外测度的概念有一个 BUG，就是当 $A\cap B=\emptyset$ 时，我们希望 $\mu^*(A\cup B)=\mu^*(A)+\mu^*(B)$。然而外测度不一定能很好满足这个性质，一个重要的原因是开集的覆盖不一定能很好地对 A 形状进行近似。在外测度的基础上，我们构造测度的思路是：设法限制外测度的定义域，以确保可数可加性在限制了的定义域上得以满足。

首先明确一个概念：设 X 是个非空集合，μ^* 是 X 上的一个外测度。X 的子集 E 称为 μ^* **可测的**，若以下条件得以满足：$\forall A\subset X,\mu^*(A)=\mu^*(A\cap E)+\mu^*(A\cap E^C)$【还可以证明命题：零集总是可测的，证明见【1】的 P237】。

基于这个概念，我们可以通过 **Caratheodory 定理**来构造测度：设 X 是非空集合，μ^* 是 X 上的一个外测度，则 X 的全体 μ^* 可测集全体 $\mathcal{M}$ 是个 σ 代数，且限制在 $\mathcal{M}$ 上的 μ^* 是可数可加的【其证明了 Lebesgue 测度的可数可加性，证明参见知乎链接之四或参见【1】的 P237 起】。对于 μ^* 可测的 E，E 的外测度常称为 E 的**测度**，并把可测集 E 的外测度（即测度）$\mu^*(E)$ 简记为 $\mu(E)$。为了方便，以后 μ^* 可测集也简称为 μ 可测集。

最后，这样得到的测度在什么样的条件下才是 λ 的延拓呢？参见【1】的 P240 的定理 9.4.2。在 R^n 的特殊情况下，通过该定理延拓成的一个 σ 代数上的测度称为 **Lebesgue 测度**，这个 σ 代数中的元素称为 Lebesgue 可测集，定义参见【1】的 P242 或参见知乎链接之三。同时，Lebesgue 可测集全体构成的 σ 代数包含 R^n 上的 Borel 代数，换言之，Borel 可测集比 Lebesgue 可测。

(6) 有时，测度并非由某个代数上的非负可数可加集函数延拓成的。我们将来要遇到的许多测度是在度量空间上的，这种测度常常不是由某个代数上的非负可数可加集函数延拓成的因为度两款空间上不存在一个天然的代数。

设 (X,ρ) 是个度量空间，$A\subset X,B\subset X$。我们称 A 和 B 是**完全分离的**，若 $\inf\{\rho(x,y):x\in A,y\in B\}>0$。$X$ 上的外测度 μ^* 称为**度量外测度**，若对于 X 的任何两个完全分离的子集 A 和 B 都有 $\mu^*(A\cup B)=\mu^*(A)+\mu^*(B)$。

(7) 有了上述的“长度”的概念后，就可以定义 Lebesgue 积分了。为什么需要它呢？Cauchy 和 Riemann 的积分概念虽然是严格的，但它包含的可积函数类的局限性太大，仍有改进的必要。其中之一就是：定义的可积函数列（在某种意义下）的极限，即使是在有界闭区间上的有界函数，仍有可能在黎曼积分的定义下是不可积的。所以 Lebesgue 积分就是为了解决“可积函数的极限有可能不是可积函数的这个数学缺陷”。在测度理论的基础上现代积分理论建立起来了。积分理论中的很大一部分结果的成立并不要求 μ 是完备的。但有些结果是只对完备测度才成立。

映射 $f:X\to\overline{\mathbb{R}}$ 称为 $\mathcal{A}$ **可测的**，若 (a) $f^{-1}(\infty)\in\mathcal{A}$ 且 $f^{-1}(-\infty)=\mathcal{A}$；(b) $\forall$ 开集 $U\subset\mathbb{R},f^{-1}(U)\in\mathcal{A}$【判定方法参见【1】的 P255 的引理 10.1.1 和【1】的 P256 的引理 10.1.2】。若 X 是拓扑空间，$\mathcal{A}$ 是 X 上的 Borel 代数，则 $\mathcal{A}$ 可测函数称为 **Borel 可测函数**。若 $X=\mathbb{R}^n$，$\mathcal{A}$ 是 $\mathbb{R}^n$ 上的 Lebesgue 可测集全体构成的代数，则 $\mathcal{A}$ 可测函数称为 **Lebesgue 可测函数**。

集合 X 上定义的函数 f 称为**简单函数**，假若它只取有限个值。任何集合的指示函数都是简单函数。每个简单函数都可表示成有限个指示函数的线性组合，但表示方法并不唯一。任何简单函数有唯一的标准表示式。f 是 $\mathcal{A}$ 可测的简单函数，当且仅当每个 $\mathcal{A}_j\in\mathcal{A}$【参见【1】的 P257】。

设 $P(x)$ 是一个依赖于 $x\in X$ 的命题。$P(x)$ 被称为在 X 上**几乎处处**（常缩写为 a.e.，即 almost everywhere 的缩写）成立的，假若集合 $\{x\in X:P(x)\text{不成立}\}$ 是零集。

(8) 首先定义下面三种积分:

(a) 设 f 是**非负简单可测函数**，它写成指示函数的线性组合的标准表示是 $f=\sum_{j=1}^{n}c_j1_{A_j}$（看成是展开式哦！即在或者不在里面?），则它的**积分**定义为 $\int fd\mu=\sum_{j=1}^{n}c_j\mu(A_j)$（从这里就能体会到“从值域出发，测度为基础”了。因为真实情况下多个 A_j 可能对应一个 f，所以我们把对应于一个“输出”的 A_j 全找

出来，再用 $\mu(A_j)$ 测量它们的“长度”，所以说是一层一层横着来的！）。对于任何 $A\in\mathcal{A}$，f 在 A 上的积分定义为 $\int_A fd\mu=\int f1_A d\mu$。【非负简单可测函数的积分性质参见【1】的 P261】。注意，积分 $\int fd\mu$ 也常写成 $\int f(x)\mu(dx)$ 或 $\int f$。

（b）设 f 是**非负可测函数**，f 的**积分**定义为 $\int fd\mu=\sup\{\int gd\mu : g\text{简单可测且}0\le g\le f\}$（逼近的方法，因为学黎曼积分的时候也会加个 C，即逼近嘛！）。对于任何 $A\in\mathcal{A}$，定义 $\int_A fd\mu=\int f1_A d\mu$。【非负可测函数的积分性质参见【1】的 P263-264】

（c）设 f 是**可测函数**，则 $f=f^+-f^-$，其中 $f^+(x)=\begin{cases} f(x), & f(x)\ge 0,\\ 0, & f(x)<0.\end{cases}$，$f^-(x)=\begin{cases} -f(x), & f(x)<0,\\ 0, & f(x)\ge 0.\end{cases}$。$f^+$ 与 f^- 分别称为 f 的正部与负部，显然 f^+ 与 f^- 均非负可测，若以下两个积分 $\int f^+d\mu$ 和 $\int f^-d\mu$ 中至少有一个取有限制，则 f 的**积分**定义为 $\int fd\mu=\int f^+d\mu-\int f^-d\mu$。若上述两个积分均有限，则 f 称为**可积函数**（或称可和函数）。全体可积函数组成的集合记作 $L^1(\mu)$，当测度由上下文不言自明时，简记做 L^1。若 μ 是 R^n 上的 Lebesgue 测度，则可积函数称为 **Lebesgue 可积的**，对应的积分称为 **Lebesgue 积分**。【可测函数的积分性质参见【1】的 P265-266，注意**绝对连续性**的概念】

针对上述 Lebesgue 积分，其体现于 Riemann 积分的优越性之一在于积分号与极限号的交换的一些结论，这些结论参见4.10节。在此之前需要先定义下述一些基本的概念：

（a）记 N 为所有几乎处处等于零的函数组成的集合，则 $L^1(\mu)/N$ 为**商空间**。

（b）可测函数列 $\{f_n\}$ 称为**按测度 μ 收敛**于可测函数 f 的，假若对于任何 $\varepsilon>0$ 有 $\lim\limits_{n\to\infty}\mu(\{x\in X : |f(x)-f_n(x)|>\varepsilon\})=0$。【例子参见【1】的 P274】

（c）设 $\{\mu_n\}$ 是定义在 σ 代数 $\mathcal{A}\subset 2^X$ 上的一个测度序列和 v 是定义在 σ 代数 $\mathcal{A}\subset 2^X$ 上的一个测度。测度序列 $\{\mu_n\}$ 称为关于测度 v 是**等度绝对连续**的，假若对于任何 $\varepsilon>0$，有一个 $\delta>0$，使得 $\forall n\in\mathbb{N},\forall E\in\mathcal{A}$，若 $v(E)<\delta$，则 $\mu_n(E)\le\varepsilon$。

理解：

（1）将可测和测度配套起来理解吧。可测就是说能在上面定义测度，因为测度需要满足将其作用在这个空间后进行运算的一些条件。

（2）Lebesgue 积分的优越性和它与黎曼积分的比较，参见【1】的 P267 起和 P285 起。

1.1.59 【Sobolev 空间的相关概念】

用含检验函数的积分形式定义导数要比传统导数定义更先进、更合理。当一个函数在某个区间上不是一次可导，如有奇点或间断点，则它的广义导数与传统导数相比就有很大的不同，并且这些不同还依赖于空间维数。对此，关键的研究问题包括：检验函数的集合怎样定义，广义函数怎样定义，广义函数的导数怎样定义，广义导数比传统导数广在何处。

（1）设 $\Omega=(a,b),u$ 在 (a,b) 上一次可导，设 $C_0^1(a,b)=\{\phi|\phi\text{在}(a,b)\text{上一次可导，且}\phi(a)=\phi(b)=0\}$，则可以得到 $\int_a^b u'\phi dx=-\int_a^b u\phi' dx,\forall\phi\in C_0^1(a,b)$【推导参见【8】的 P1，其中 $\phi(a)=\phi(b)=0$ 是关键】。设存在一个函数 w 满足 $\int_a^b w\phi dx=-\int_a^b u\phi' dx,\forall\phi\in C_0^1(a,b)$，我们暂称 w 为 u 的**广义导数**并记为 Du，称 ϕ 为**检验函数**。引入记号 $\langle Du,\phi\rangle=-\langle u,D\phi\rangle$，就回到了之前学过的（弱）广义导数，见1.1.34节的 (3)。

（2）设 $T\in D'(\Omega)$，我们定义 T 的导数 $D^\alpha T$ 满足 $\langle D^\alpha T,\phi\rangle=(-1)^\alpha\langle T,D^\alpha\phi\rangle,\forall\phi\in D(\Omega)$，其中 $|\alpha|\ge 1$，称为**广义导数**【计算实例参见【8】的 P4，Dirac 函数的引入，以及【8】的 P6-10 的那些计算例子，做法都是把形式写出来，然后按照上述定义式把导数部分提取出来即可】。

广义导数与传统导数的对比参见【8】的 P6-10，包括：有奇点的函数、关于奇点的注记、函数有间断、收敛于 Dirac 函数 δ 的序列、多维分片光滑、Cantor 函数的广义导数与一些不妥提法。

（3）设 $T\in D'(\Omega),\{T_i\}\subset D'(\Omega)$，我们定义 $T_i\to T$ 为 $\lim\limits_{i\to\infty}\langle T_i,\phi\rangle=\langle T,\phi\rangle,\forall\phi\in D(\Omega)$。

1.1.60 事件、随机试验、试验（样本点）、试验空间（样本空间）、联合试验、笛卡尔积集、n 维乘积空间

（1）**事件**：人类在对自然界的认识过程中最基本的概念。比如：不可能事件 $\emptyset$、必然事件 Ω、事件 “A 不发生” A^c（所以不能说某某事件不发生，应当把某某事件不发生也当做一个事件）、至少发生一个的事件 $\bigcup$、同时发生的事件 $\bigcap$。此外，还包括事件的运算：不相容 $A+B$、事件差 $A-B$、对称差 $A\Delta B$（事件 A 和 B 中有且仅有一个发生）、包含、相等【具体参见【5】的 P2-3】。

上述这些将自然语言与数学语言对应起来。比如：当指标集 I 是空集时，有 $\bigcup\limits_{i\in I} A_i=\emptyset, \bigcap\limits_{i\in I} A_i=\Omega$，将它们“翻译”成自然语言都很好理解。而且同一个自然语言可以有多重不同的表达形式，如 $A\Delta B = (A-B)+(B-A)=(A\bigcup B)-(A\bigcap B)$。

（2）**随机试验**的一个结果称为一次（一个）**试验（样本点）**ω。比如：在考虑或不考虑顺序的情形下，从 n 件产品中随机抽取 m 件进行检测；观察日光灯管的寿命，这时每一时刻都是试验的一个可能的结果。试验结果的全体称为**试验空间（样本空间）**Ω

（3）将两个随机试验合并为一个试验，可记为 $E=E_1\times E_2$，称为**联合试验**。则其样本空间是 Ω_1 和 Ω_2 的**笛卡尔积集** $\Omega=\Omega_1\times\Omega_2=\{(\omega_1,\omega_2)|\omega_1\in\Omega_1,\omega_2\in\Omega_2\}$。假设随机试验 E 的样本空间为 Ω，若将 E 重复进行 n 次的试验记为 $E=E\times E\times ...\times E=E^n$，则其样本空间为 $\Omega_n=\{(\omega_1,\omega_2,...,\omega_n)|\omega_i\in\Omega, i=1,2,...,n\}=\Omega\times\Omega\times...\times\Omega=\Omega^n$，称为 Ω 的 **n 维乘积空间**。

理解：

（1）对样本空间的理解要把握其：（a）充分大，使得能够找到一个试验把两个不同的事件区分开来；(b)“不可再分”、“最基本的”。比如硬币投 3 次，$\Omega=\{$ 全正，全反，两正一反，两反一正 $\}$ 就不是，因为“两正一反”还可以再分为“正正反”、“正反正”、“反正正”；“两反一正”同理)。

（2）对于试验与事件的联系：我们可以把每个事件 A 与试验空间中实现事件 A 的试验所组成的实验空间的那个子集 A' 联系起来，而且很自然地试图把 A 和 A' 等同起来【参见【5】的 P4-5 例子。简而言之就是用各试验结果来为事件赋予具体的含义或名称】。从而将事件运算转化成了集合论中集合的运算。可以将其理解为拓扑空间中的开集。

1.1.61 逆映射、可测映射、可测函数

（1）设 X 是从 Ω 到 Ω' 中的一个映射，则存在一个**逆映射** $X^{-1}:\mathcal{F}(\Omega')\to\mathcal{F}(\Omega)$ 满足 $X^{-1}(A')=\{\omega: X(\omega)\in A'\}, A'\in\mathcal{F}(\Omega')$。逆映射对余运算、交运算和并运算有 5 个性质，参见【5】的 P42。

（2）给定两个可测空间 $(\Omega_1,\mathcal{A}_1)$ 和 $(\Omega_2,\mathcal{A}_2)$,X 为 Ω_1 到 Ω_2 中的映射，若 $X^{-1}(\mathcal{A}_2)\subseteq\mathcal{A}_1$，则称 X 为**可测映射**，记 $X\in\mathcal{A}_1|\mathcal{A}_2$【判定定理（充要条件）见【5】的 P44】。若 $(\Omega_2,\mathcal{A}_2)=(\overline{\mathbb{R}}^n,\overline{\mathcal{B}}^{(n)})$，则简记为 $X\in\mathcal{A}_1$，并称 X 为 $\overline{\mathbb{R}}^n$ 上的 Borel 可测函数，或简称为**可测函数**【相关的符号含义见【5】的 P11，如 Borel 域、扩张直线（扩张这个表示方法在刘晓老师课上也用过)】。

结论：

（1）若 $\mathcal{A}'$ 为 Ω' 上的一个代数（或 σ 代数），则 $X^{-1}(\mathcal{A}')$ 也是 Ω 上的一个代数（或 σ 代数）【证明见【5】的 P42】。

（2）设 $\mathcal{F}'$ 为 Ω' 的一个子集类，则由 $\mathcal{F}'$ 生成的 σ 代数 $\mathcal{A}'$ 的逆像 $X^{-1}(\mathcal{A}')$ 等于由 $X^{-1}(\mathcal{F}')$ 生成的 σ 代数【证明见【5】的 P43】。

理解：

（1）X 可测是指 $\mathcal{A}_1$ 经 X 导出的 Ω_2 上的子集 σ 代数比 $\mathcal{A}_2$ 更精细。

1.1.62 阶梯随机变量、随机变量、分布函数

（1）设 $\{A_i, i\in I\}$ 为 Ω 的一个有限分割【概念参见【5】的 P6】，且 $A_i\in\mathcal{A}, i\in I$，$X$ 为 Ω 到实直线 $\mathbb{R}$ 的一个映射。若在每个 A_i 上 X 取常值，即 $X(\omega)=x_i,\omega\in A_i, i\in I$，则称 X 是**阶梯随机变量**【判定定理（充要条件）参见【5】的 P46】【性质参见【5】的 P46】。

（2）设 X 是 $(\Omega,\mathcal{A})$ 到 $(\mathbb{R},\mathcal{B})$ 中的映射，若存在阶梯随机变量序列 X_n，使对每个 $\omega\in\Omega$，都有 $X_n(\omega)\to X(\omega), n\to\infty$，即 X 是一系列阶梯随机变量的逐点极限，则称 X 是 $(\Omega,\mathcal{A},P)$ 上的一个**随机变**

量。若 $X(\Omega)\subseteq\mathbb{R}_+$，则称 X 为**非负随机变量**【判定定理（充要条件）参见【5】的 P47】【性质参见【5】的 P48 起，把随机变量看做是可测函数的一个特例】。

关于随机变量，在【6】的 P5 定义为：设 $(\Omega,\mathcal{A},P)$ 是概率空间，$X(\omega)$ 是定义在 Ω 上的单值实函数（$f(x)$ 唯一，单射），若对于任意实数 $x\in\mathbb{R}$，有 $\{\omega|X(\omega)\le x\}\in\mathcal{A}$（保证了 $P\{X\le\omega\}$ 总有意义），则称 $X(\omega)$ 是定义在 $(\Omega,\mathcal{A},P)$ 上的随机变量。

(3) 设 $X(\omega)$ 是定义在概率空间 $(\Omega,\mathcal{A},P)$ 上的随机变量，x 是任意实数，称函数 $F(x)=P\{X\le x\}=P\{\omega|X(\omega)\le x\}$ 为随机变量 X 的**分布函数**，也记为 $F_X(x)$。

例子：

"试验空间（样本空间）、阶梯随机变量、随机变量、事件"之间的关系参见【1】的 P215-217 的例 9.1.3（注意，前提是**有限可加的集函数（可加集函数）**【参见【1】的 P216，而这个概念的前提又是建立在代数的基础上的哦，所以要理清拓扑和概率的联系】。独立重复试验，Bernoulli 模型和二项概率模型），明白它们在这个例子中各自指什么，联系起拓扑和概率论，很清晰!

推广：

对于二**维随机变量、联合分布函数**、**n 维随机变量**、**n 维联合分布函数**的定义也是如此，以及**边缘分布函数**的定义，详细参见【6】的 P6-7。

理解：

(1) 随机变量许多性质的研究都是通过用阶梯随机变量逼近的方法来研究的（**逼近研究法**）【具体见【5】的 P45】。

(2) 随机变量是一个函数（到实值空间的映射）。

(3) 随机变量与可测函数的差别仅在于是否可取值 ∞。

1.1.63 随机过程（随机函数）、n 维随机向量、时间序列、随机场（多维指标集随机过程）、样本函数

(1) 设给定概率空间 $(\Omega,\mathcal{A},P)$ 和指标集 T，若对每个 $t\in T$，有定义在 $(\Omega,\mathcal{A},P)$ 上的随机变量 $X_t(\omega),\omega\in\Omega$ 与之对应。称依赖于 t 的随机变量族 X_t 为**随机过程（随机函数）**，记为 $\{X_t(\omega),t\in T\}$ 或 $\{X_t,t\in T\}$ 及 $\{X(t),t\in T\}$。当参数集 $T=\{1,2,...,n\}$，则 $\{X_t,t\in T\}=(X_1,X_2,...,X_n)$ 为 **n 维随机向量**。当集 $T=\{1,2,...,n,...\}$，则为**时间序列**。当参数集是多维指标集的堆积过程称为**随机场（多维指标集随机过程）**。

关于它们的更"数学"一些的定义参见【5】的 P117-118，包括：一族可测空间、空间 Ω_T、投影映射、柱集等等

(2) 对每一固定 $\omega\in\Omega$，称 $x_t(\omega)$ 是随机过程 $\{X_t(\omega),t\in T\}$ 相应于 ω 的**样本函数**（或轨道，路径，现实）。

理解：

(1) 随机过程的参数 t 不一定是时间变量或一维实数。比如海平面上波浪的浪高随地点和时间的改变而变化，因此为描述一个区域内浪高 H 的变化，需要用到三维参数 (x,y,t) 构成一个参数集。

(2) 显然，一条样本函数可看成是在特定条件下对随机过程的一次观察结果。

1.2 常用的离散域和连续域

1.2.1 离散域

1. K^n：n 维向量空间。
2. $K^{m\times n}$：$m\times n$ 维矩阵空间。
3. R^∞：一般数列空间，$R^\infty=\{(x_1,x_2,...,x_k,...):x_k\in R,k\in N\}$。
4. l^∞：有界数列空间, $l^\infty=\{(x_1,x_2,...,x_k,...):\sup_{k\in N}|x_k|<\infty\}$。
5. l^p： p 次幂可和数列空间, $l^p=\{(x_1,x_2,...,x_k,...):\sum_{k=1}^\infty|x_k|^p<\infty\}$。
6. (1) Ω 的子集类（就是子集组成的集族）$\mathcal{F}$ 称为 Ω 中的**集合代数或集合域，简称为代数或域**，若它满足：(a) $\Omega\in\mathcal{F}$；(b) 若 $A\in\mathcal{F}$，则 $A^C\in\mathcal{F}$（前面说过，事件 A 不发生也应该当做是一个事件）；(c)

若 $A, B \in \mathcal{F}$，则 $A \cup B \in \mathcal{F}$（即"有限并封闭"）。其任意元素称为随机事件，从这三个条件或性质出发可以推导出更多的性质。【参见【5】的 P6 等价叙述，即充要条件，用 $\emptyset$ 和 $\cap$】

（2）在 $\mathcal{F}$ 的基础上构造代数的方法（要合理构造，否则太大会加重负担）:

给定 Ω 的一个子集类 $\mathcal{F}$，必然存在包含 $\mathcal{F}$ 且由 Ω 的子集构成的**最小代数** $\mathcal{A}$【证明见【5】的 P6】。给定 Ω 的任意一个子集类 $\mathcal{F}$，由此可以得到 $\mathcal{F}$ 的**生成代数**，即最小代数【定理和证明见【5】的 P7-8】（这实际上就是求最小代数的方法）。

（3）然而，代数 $\mathcal{A}$ 仅对有限交、有限并运算封闭。我们需要进一步研究对极限运算封闭的类，现在对 $\mathcal{A}$ 进行推广:

若 Ω 的子集类 $\mathcal{A}$ 包含 $\emptyset, \Omega$ 以及在余运算、可列交和可列并运算下封闭，则称 $\mathcal{A}$ 是 Ω 的一个子集 σ **代数**。(注意区别：这里是"可列"，而非"有限"，也就是说这里可以是"无限"的，这也贴合了【6】P4 对 σ 代数的定义）

那我们为什么还要在（2）中用"有限"来定义"代数"的概念，而不直接定义 σ 代数呢？因为首先定义代数可以帮助我们验证或构造 σ 代数：为验证 $\mathcal{A}$ 是一个 σ 代数，只要验证 $\mathcal{A}$ 是一个代数以及在可列交或可列并之一运算下封闭即可【例子见【5】的 P11】【σ 代数的 2 个判定定理见【5】的 P12-13】。

（4）设 X 是个拓扑空间，则由 X 上全体闭集组成的集族生成的 σ 代数称为 X 上的 **Borel 代数**。Borel 代数中的元素（它本身是 X 的子集）称为 X 中的 **Borel（可测）集**。

1.2.2 连续域

1. $C([a,b])$：连续函数空间, 即 $[a,b]$ 上连续函数的全体。

2. $C^k([a,b])$：k 阶连续可导函数空间，即 $[a,b]$ 上具有直到 k 阶连续导函数的函数的全体。

3. $C^\infty([a,b])$：无穷阶连续可导函数空间，即 $[a,b]$ 上具有无穷阶连续导函数的函数的全体。

4. $B(\Omega)$：有界函数空间，即 Ω 上有界函数的全体，$B(\Omega) = \{u : \sup_{t\in\Omega} |u(t)| < \infty\}$。

5. $L^p(E)$：p 次幂可积函数空间，即 $L^p(E) = \{f(x) | \int_E |f(x)|^p dx < +\infty\}, 1 \le p < \infty$，$E$ 是 R^n 中的可测集，$f(x)$ 是 E 上的可测函数，且 $|f(x)|^p$ 在 E 上 Lebesgue 可积，则称 $f(x)$ 在 E 上是 p 次幂可积的，所有这样的函数的全体称为 E 上的 L^p 空间，记为 $L^p(E)$ 或 L^p。

理解：

1. 集合中最大/最小的元素可能不存在，比如不存在最小的正实数（没有最小、只有更小），但是正数集合的"底部"是 0。所以我们使用**最小上界（上确界）和最小下界（下确界）**的概念来替代最大值和最小值，因为上确界和下确界总是存在的。定义：

设 $A \subseteq \mathbb{R}$ 是一个上有界的集合（即存在 $M \in \mathbb{R}$ 使得 A 中的每一个数都小于或等于 M）。如果一个实数 s 满足以下两个条件：（1）对于任意的 $x \in A, x \le s$；（2）无论 $\epsilon > 0$ 有多小，总存在 $x \in A$，使得 $|x - s| < \epsilon$，即：A 中的数可以无限接近于 s。则称 s 为 A 的最小上界，又称为上确界，记为 $s = supA$。如果 A 上无界，则可以说 $supA = +\infty$。

用类似的方法可以定义集合的最大下确界（下确界）。

1.3 常用的空间

1.3.1 度量 (距离) 空间 (X, d)

在非空集合 X 上定义映射 $d : X \times X \to R$，使得 $d(x, y)$ 满足："非负唯一性"、"对称性"和"三角不等式" 3 个条件。这个定义的映射则被称为是 X 上的一个度量 (或距离)【《应用泛函分析》P36】。

1.3.2 赋范空间 $(X, \|\cdot\|)$

在数域 K 上的线性空间 X（即为了满足对线性运算封闭的前提条件）中定义映射 $\|\cdot\| : X \to R$，使得 $\|x\|$ 满足："正定性"、"绝对齐次性"和"三角不等式" 3 个条件【《应用泛函分析》P74】。则 $(X, \|\cdot\|)$ 称为赋范线性空间或线性赋范空间。

赋范线性空间称为**无限维的**，假若在任何有限个线性无关的向量 $x_1, x_2, ..., x_n \in X$ 之外还可以找到一个向量 $x_{n+1} \in X$，使得 $x_1, x_2, ..., x_n, x_{n+1}$ 是线性无关的。

理解：

(1) 如果数集 K 中任意两个数作某一运算后仍在 K 中，则称**数集 K 对这个运算是封闭的**。对加、减、乘、除（除数不为零）四则运算封闭的数集 K 称为**数域**，常见的数域有**实数域** R、**复数域** C **和有理数域** Q 等；

(2) 若非空集合 X 满足加法封闭性和在数域 K 上的数乘封闭性，且唯一，则称 X 为**数域 K 上的线性空间**，记为 $X(K)$ 或 $X(K)(+,\cdot)$。其中，非空集合 X 如果是有限个元素，除了零空间外，不能构成线性空间；加法和数乘运的算定义不唯一，只要它们各自满足 4 条规则即可【《矩阵理论》P1】；

(3) 赋范空间比度量空间要求更宽，换言之，范数可以诱导距离，但距离不一定能诱导范数，这也是为什么在损失函数收敛分析时可以将压缩映像原理应用于范数的原因吧；

(4) 由三角不等式条件还可以推导出赋范空间满足 $|\|x\|-\|y\||\leq\|x-y\|$【《应用泛函分析》P75】；

(5) 从范数的定义中，可以立刻推出范数性质：(i) 零向量的范数是 0；(ii)$x\neq 0$ 时，$\|\frac{1}{\|x\|}x\|=1$；(iii) 对任意 $x\in C^n,\|-x\|=\|x\|$；(iiii) 对任意 $x,y\in C^n$，有 $|\|x\|-\|y\||\leq\|x-y\|$【《矩阵理论》P50 证明】；

(6) 范数定义中的“三角不等式”性质主要用于证明中的验证，但在实际应用当中用得更多的还是 $|\|x\|-\|y\||\leq\|x-y\|$，从中可以看出范数具有 lipschiz 连续性（则一定连续）。

1.3.3 Banach 空间

完备的赋范空间。

理解：

(1) 若 $\forall\epsilon>0,\exists N\in\mathbf{N}$, 使得当 $m,n>N$ 时，有 $|x_m-x_n|<\epsilon$，则称数列 $\{x_n\}$ 为 **Cauchy 列（或基本列）**；

(2) 收敛列必是 Cauchy 列，但 Cauchy 列不一定是收敛列。例如：在有理数集 Q 中，令 a_n 为 π 的精确到 10^{-n} 的不足近似值，即 $a_1=3.1,a_2=3.14,a_3=3.141,...$，$\forall\epsilon>0$，取 $N>-lg\epsilon$，则当 $m>n>N$ 时，就有 $|x_m-x_n|<\frac{1}{10^n}<\epsilon$，故 $\{a_n\}$ 是 Cauchy 列，但其极限 $\pi\notin Q$，所以它在 Q 内不收敛。也就是说，它的收敛极限在集合 Q 中不存在，这是一个今后可以用来做验证的 Cauchy 列，另外还可以构造很多其他的 Cauchy 列作为反例；

(3) 若集合中的任一 Cauchy 列都收敛到该集合中的对应元素，则称这个集合是**完备的**。例如：R 中的任一 Cauchy 列都收敛到一实数，因此称 R 是完备的；K 也是完备的；

(4) 对于自反的 Banach 空间 X，$\forall x_0\in X$，定义 X^* 上的有界线性泛函 x_0^{**}，对于 $\forall f\in X^*$，有 $x_0^{**}(f)=f(x_0)$ 成立。对于范数，有 $\|x_0^{**}\|_{X^{**}}=\|x_0\|_X$ 成立【《泛函分析》课程笔记 P4】。

1.3.4 内积空间 $(X,\langle\cdot,\cdot\rangle)$、欧几里得空间、酉空间

(1) 在数域 K 上的线性空间 X 中定义映射 $\langle\cdot,\cdot\rangle:X\times X\to K$，使得 $\langle x,y\rangle$ 满足：“正定性”、“共轭对称性”、“对第一变元的线性” 3 个条件【《应用泛函分析》P91】；

(2) 有限维的实内积空间称为欧几里得空间（欧氏空间）；

(3) 有限维的复内积空间称为酉空间。

理解：

(1)Banach 空间由于太过于抽象，因此很多东西在实际中不易应用，因此研究是否能够将 Banach 空间再缩小一些，使得能够得到更好的空间性质，进而应用到优化、解的存在性等实际问题中去；

(2) 由“对第一变元的线性”和“共轭对称性”还可以推导出内积具有“对第二变元的共轭线性”【《应用泛函分析》P91-92】。

例子：

在欧几里得空间中定义运算 $\langle\alpha,\beta\rangle=\alpha^TA\beta$，易得该运算为内积的前提条件为 A 为对称正定矩阵。同理，在酉空间中 A 为 Hermite 正定矩阵。这个例子具有现实应用，因为线性无关的向量不一定正交 (其中 $\langle\alpha,\beta\rangle=\arccos\dfrac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}$)，而正交的向量一定线性无关，所以我们在定义内积时往往希望同时能够把不正交的空间转换为正交的空间，因此这里的 A 就起到了这个作用。正交空间中，任意两个向量 $x\perp y$，由勾股定理有 $\|x+y\|=\|x\|+\|y\|$。此外，在大规模科学计算中，我们可以将内积计算转换为矩

阵计算，即假设 $\varepsilon_1,\varepsilon_2,...,\varepsilon_n$ 为 V 的一组基底，则 $\forall\alpha,\beta\in V_n(R)$，有 $\alpha=\sum_{i=1}^{n}x_i\varepsilon_i,\beta=\sum_{j=1}^{n}y_j\varepsilon_j$，此时 $\langle\alpha,\beta\rangle=\sum_{i,j=1}^{n}x_iy_j\langle\varepsilon_i,\varepsilon_j\rangle=\sum_{i,j=1}^{n}x_iy_ja_{ij}$，这里的 $A=(a_{ij})$ 是度量矩阵。

同时也要会快速判断某空间不是内积空间。【3.4-1 的 2 题、3 题和 4 题】

1.3.5 Hilbert 空间

完备的内积空间。

理解：

(1) 内积空间的完备性也可描述为：$\forall x_n,x_m\in X$，有 $\langle x_n-x_m,x_n-x_m\rangle\to 0$，当 $n,m\to\infty$（Cauchy 列），可以找到 $\mu\in X$，使得 $\langle x_n-\mu,x_n-\mu\rangle\to 0$，当 $n\to\infty$（$\{x_n\}$ 收敛点在 X 中存在，由此也可得到 $\{x_m\}$ 亦如此）；

(2) 借助 Banach 空间视角也可以判断内积空间是否为 Hilbert 空间，即：内积空间是 Hilbert 空间，若它是 Banach 空间，且范数由内积诱导。因此 Banach 空间的一些结论可以移植到 Hilbert 空间。

1.3.6 有界线性算子空间、对偶空间 (共轭空间)、二次对偶空间、自反空间

(1) 设 X,Y 是数域 K 上的赋范空间，记 $B(X,Y)$ 为从 X 到 Y 的有界线性算子的集合。$\forall x\in X,\alpha\in K$，在 $B(X,Y)$ 中定义线性运算 $(T_1+T_2)x=T_1x+T_2x,(\alpha T)x=\alpha(Tx)$，则 $B(X,Y)$ 是数域 K 上的线性空间，称为有界线性算子空间；

(2) 称赋范空间 X 上的全体有界线性泛函所成的赋范空间 $B(X,K)$ 为 X 的对偶空间 (或共轭空间)，记作 X^*；

(3)X^* 的对偶空间 $(X^*)^*$ 称为 X 的二次对偶空间 (或二次共轭空间)，记作 X^{**}；

(4) 赋范空间 X 称为自反空间，若映射 $J:X\to X^{**}:J(x)=J_x,\forall x\in X$ 是满射，即 $J(X)=X^{**}$。

理解：

(1) 注意有界线性算子空间与对偶空间 (共轭空间)、二次共轭空间的区别；

(2)$B(X,Y)$ 中的算子列 $\{T_n\}$ 称为**依算子范数收敛**，若 $\exists T\in B(X,Y)$，使得 $\|T_n-T\|\to 0(n\to\infty)$，记作 $T_n\to T(n\to\infty)$；

(3) 设 $B(A)$ 是赋范空间 X 到其自身上所有的有界线性算子之集，且 $S,T\in B(A)$，定义其乘积 $ST(x)=S(T(x)),\forall x\in A,T\in B(A)$ 称为**可逆的**，如果 $\exists T^{-1}\in B(A)$ 称为T 的逆，使得 $TT^{-1}=T^{-1}T=I$；

(4) 对于上述的概念 (2)，结合【3.6-5 的 17 题】再理解下，用同构去理解。这些线性泛函在 X 的对偶空间 X^* 中。

1.3.7 拓扑、拓扑空间

设 X 是个集合，X 上的一个**拓扑** $\mathcal{T}$ 是由 X 的一些子集组成的满足以下条件的子集族：

(1) 设 A 是一个指标集，则 $\forall\alpha\in A,U_\alpha\in\mathcal{T}\Rightarrow\bigcup_{\alpha\in A}U_\alpha\in\mathcal{T}$ ($\alpha\in A$ 表示任意多个)（任意并封闭）；

(2) $F=\{1,...,n\}$ 是有限个指标组成的指标集，则 $\forall j\in F,U_j\in\mathcal{T}\Rightarrow\bigcap_{j\in F}U_j\in\mathcal{T}$ （有限交封闭）；

(3) $\emptyset\in\mathcal{T},X\in\mathcal{T}$。

集合 X 和 X 上的一个拓扑 $\mathcal{T}\subset 2^X$ 组成的二元组 $(X,\mathcal{T})$ 称为一个**拓扑空间**，其中 2^X 表示 X 的子集的全体构成的集合。当 $\mathcal{T}$ 已从上下文可以无误地确定时，也简称 X 是个拓扑空间。

理解：

(1) 和之前泛函分析中所学的空间的关系：拓扑空间进入了一个更底层、更抽象的数学层次，同样与之前对空间的理解 (集合 + 操作) 类似，这里的拓扑空间可以理解为定义一种操作 $\mathcal{T}$ 来从集合 X 的 2^X 中提取满足 3 个公理化条件的元素；

(2) 上述公理化定义中的前两条可以概括为“任意并封闭”和“有限交封闭”。

1.3.8 Hausdorff 空间

拓扑空间 X 称为 Hausdorff 空间，若对于 X 中的任何两个不同的点 x 和 y，必有 x 的邻域 U 和 y 的邻域 V，使得 $U \cap V = \emptyset$。

理解：

由度量空间 (X,ρ) 诱导出的拓扑空间 X 必是 Hausdorff 空间。【《数学分析讲义（第二册)》P35 证明】

1.3.9 可测空间、测度空间、概率空间、导出、乘积可测空间

（1）试验空间（样本空间）Ω 和 σ 代数 $\mathcal{F}(\mathcal{F} \subset 2^{\Omega})$ 的二元体 $(\Omega,\mathcal{F})$ 称为**可测空间**。$\mathcal{F}$ 中的元素称为 $\mathcal{F}$-可测集，有时简称可测集。

（2）设 X 是个非空集合，$\mathcal{A} \subset 2^X$。集函数 $\mu : \mathcal{A} \to \overline{\mathbb{R}}_+$【定义参见【1】的 P214】称为**可数可加**的，假若对于任何可数个 $A_n \in \mathcal{A}, n = 1, 2, ...$，只要这些 A_n 是两两不相交的，且 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$，我们便有 $\mu(\bigcup\limits_{n=1}^{\infty} A_n) = \sum\limits_{n=1}^{\infty}\mu(A_n)$。又若 $\mathcal{A}$ 是个 σ 代数，则可数可加的映射 μ 称为 $(X,\mathcal{A})$ 上的一个**测度**。三元组 $(X,\mathcal{A},\mu)$ 称为**测度空间**。

测度空间 $(X,\mathcal{A},\mu)$ 称为**完备的**，假若它的任何零集的子集都是可测的。**测度空间的完备化**方法参见【1】的 P232。

（3）首先定义**概率**：Ω 的某个子集代数 $\mathcal{A}$（最小代数哟）上的一个概率 P 是 $\mathcal{A}$ 到 $[0,1]$ 的一个映射，满足如下的三条公理：(a) $P(\Omega) = 1$（规范性）；(b) 对于任意两两不交的有限子集族 $\{A_i, 1 \le i \le n\}, P(\sum\limits_{i=1}^{n} A_i) = \sum\limits_{i=1}^{n} P(A_i)$（有限加性）；(c) 设 $A_n \in \mathcal{A}, n \in \mathbb{N}$，满足 $A_n \downarrow \emptyset$，则 $\lim\limits_{n\to\infty} P(A_n) = 0$（在 $\emptyset$ 处的连续性，反之亦然）(这一点区别于本科学习的初等概率中的“可列可加性”，其在如今的代数定义下可能都不在定义域内了！)。【概率的 7 个性质见【5】的 P9-10】

那么，我们已经在代数 $\mathcal{A}$（最小代数）上定义了概率，那如何把 $\mathcal{A}$ 上的代数唯一扩张到 $\sigma(\mathcal{A})$ 上去呢？那就要用到**概率的扩张**【参见【5】的 1.4 节，P16-26】。

至此，设 $\mathcal{A}$ 是 Ω 的一个 σ 代数，P 是定义在 $\mathcal{A}$ 上的概率，则三重偶 $(\Omega,\mathcal{F},P)$ 称为一个**概率空间**【例题见【6】的 P4-5】。

（4）给定一个 Ω 到 Ω' 的映射 X 以及 Ω 的 σ 代数 $\mathcal{A}$，有 $\mathcal{A}' = \{A' | X^{-1}(A') \in \mathcal{A}\}$ 是 Ω 上的一个 σ 代数。若 P 是可测空间 $(\Omega,\mathcal{A})$ 上的一个概率，则由公式 $P'(A') = P(X^{-1}(A'))$ 在可测空间 $(\Omega',\mathcal{A}')$ 上定义一个概率 P'，我们称概率空间 $(\Omega',\mathcal{A}',P')$ 是通过 X 由 $(\Omega,\mathcal{A},P)$ **导出**的（相关概念回归1.1.61）【例子参见【5】的 P43-44】。

(5) 可测空间 $(\Omega_1\times\Omega_2, \mathcal{A}_1\times\mathcal{A}_2)$ 称为可测空间 $(\Omega_1,\mathcal{A}_1)$ 和 $(\Omega_2,\mathcal{A}_2)$ 的**乘积可测空间**【“投影映射”、“截口”的概念见【5】的 P108-109】【性质定理见【5】的 P109-110】。

1.3.10 状态空间、无穷维乘积可测空间

（1）将事件“$X_t = x$”称为在时刻 t 时随机过程 X_t 处于状态 x【理解参见【6】的 P8】。将随机过程 $\{X_t(\omega), t \in T\}$ 可能取值的全体构成的集合（所有状态构成的集合）称为**状态空间**，记为 E，即 $E = \{x | X_t(\omega) = x, t \in T\}$。

随机过程在时刻 t 的状态不完全由 t 来确定，它还会受到随机因素 ω 的影响。

状态空间看似也就是【5】P119 的**无穷维乘积可测空间**：将可测空间 $(\Omega_T,\mathcal{A}_T), \Omega_T = \prod_{t\in T}\Omega_t = \{\omega_T = (\omega_t, t\in T) | \omega_t \in \Omega_t, t \in T\}$（当 $\Omega_t = \Omega, \forall t \in T$ 时，常用 Ω^T 表示 Ω_T）称为 $(\Omega_t,\mathcal{A}_t), t \in T$ 的乘积可测空间。

其目的大致有两个，一个是将有限乘积空间上的一些定理推广到无穷维乘积空间可测空间中来【具体参见【5】的 P119 起】。二是空间 Ω_t 常取为离散空间或欧式空间，而且往往不依赖于 t。这个不依赖于 t 记得当时在随机过程中学过，而且也是这门课的精华：将问题转换进行分析，“摆脱”对 t 的一个依赖。

1.3.11 Lebesgue 函数空间

（1）设 X 是个非空集合，$\mathcal{A}$ 是 X 上的一个 σ 代数，μ 是 $\mathcal{A}$ 上的一个测度，$1 \leq p < \infty$。$L^p(X,\mu)$（当 X 和 μ 已从上下文不言自明时，也简记做 L^p）定义为 $L^p = \{f : X \to C, f\text{是}\mu\text{可测的，且}|f|^p\text{可积}\}$。在 $L^p(X,\mu) = L^p$ 中引进等价关系 $\sim$ 如下：$f \sim g \Leftrightarrow f = g, a.e.(\mu)$。$L^p(X,\mu) = L^p$ 相当于如上定义的等价关系 $\sim$ 被分成许多等价类，这样的等价类的全体称为 L^p **空间**。记作 $L^p(X,\mu)$，有时简记做 L^p。可以证明，L^p 是线性空间【证明参见【1】的 P323-324】。

(2)对于 $f \in L^p$,它的**范数**定义为 $|f|_p = (\int |f|^p d\mu)^{1/p}$。它具有以下三条重要性质:(a)$\forall \lambda \in \mathbb{C}, \forall f \in L^p$，有 $|\lambda f|_p = |\lambda||f|_p$；(b) $\forall f, g \in L^p$，有 $|f+g|_p \leq |f|_p + |g|_p$；(c) $|f|_p = 0 \Leftrightarrow f = 0$【性质 (b) 的证明参见【1】的 P324-330, 要用到本质上界 $|f|_\infty$ 和测度形式的 Jensen 不等式 (P324)、Holder 不等式 (P326) 和 Minkowski 不等式 (P328)】。

L^p 中有了范数 $|\cdot|_p$，L^p 中可以通过范数 $|\cdot|_p$ 引进以下形式的距离：$\forall f, g \in L^p, \rho(f,g) = |f-g|_p$，不难检验 L^p 相对于 ρ 构成度量空间。关于度量空间的全部理论都可用到 L^p 上, 但是关于度量空间的较深刻的理论主要集中在完备度量空间上，因此我们同时也有必要研究 L^p 空间的完备性，参见4.11节。

（3）映射 $L : L^p \to \mathbb{C}$ 称为**线性泛函**，假若 $\forall a, b \in \mathbb{C}, \forall f, g \in L^p$，有 $L(af+bg) = aL(f) + bL(g)$。假若 L 又是连续的（相对于 L^p 的由范数确定的拓扑和 $\mathbb{C}$ 上的二维欧氏空间的拓扑），则 L 称为**连续线性泛函**。假若有一个常数 K，使得 $\forall f \in L^p$，有 $|L(f)| \leq K|f|_p$，则称 L 为**有界线性泛函**。不难证明以下两点：(a) 线性泛函在 L^p 上是连续的，当且仅当它在 0 点是连续的；(b) 线性泛函是连续的，当且仅当它是有界的。

L^p 上全体连续线性泛函也构成一个线性空间,称为 L^p 的**对偶空间**(老一点的文献也常称为共轭空间),并记做 $(L^p)'$。若在这个线性空间上引进范数：$\forall L \in (L^p)'$，有 $|L|_{(L^p)'} = \sup\limits_{|f|_p=1} |L(f)| = \sup\limits_{|f|_p \neq 0} |L(f)|/|f|_p$。易证 $(L^p)'$ 也是个度量空间，因而，也是拓扑空间。

（4）空间 $L^2(\mu)$ 具有一般的 $L^p(\mu)$ 空间所不具有的特殊概念：内积。利用这个概念，空间 $L^2(\mu)$ 有着许多漂亮的结果。

理解：

（1）事实上，L^p 空间可以理解成把 L^p 空间中的几乎处处相等的函数看成同一元素后而得到的空间。

1.3.12 试验函数空间

试验函数空间 $D(\Omega)$ 感觉就是1.1.59节开头提到的“检验函数”构成的空间，即解决“检验函数的集合怎样定义”这个问题。

设 $\Omega \subset R^n$ 是开集，$C_0^\infty(\Omega)$ 是定义在 Ω 上具有紧支集的 C^∞ 函数集合。在集合 $C_0^\infty(\Omega)$ 上定义收敛性如下：我们说序列 $\{\varphi_j\}$ 收敛于 φ_0，如果（1）存在一个相对于 Ω 的紧集 $K \subset \Omega$，使得 $\text{supp}(\varphi_j) \subset K(j = 0,1,2,...)$；（2）对于任意指标 $\alpha = (\alpha_1, ..., \alpha_n)$，都有 $\max\limits_{x \in K} |\partial^\alpha \varphi_j(x) - \partial^\alpha \varphi_0(x)| \to 0 (j \to +\infty)$（也就是说 $\varphi_j(x)$ 及所有导数一致收敛到 $\varphi_0(x)$）。赋予上述拓扑收敛性的线性空间 $C_0^\infty(\Omega)$ 称为试验函数空间，记为 $D(\Omega)$。

其中，**支集**的概念为：一个定义在集合 X 上的实值函数 f 的支撑集，或简称支集，是指 X 的一个子集，满足 f 恰好在这个子集上非 0。$D(\Omega)$ 与一般函数空间的不同之处就体现在紧支集与 2 个条件。而**紧支集**意思就是说这个支集是紧的。【关于紧性、紧空间等的概念和理解，参见1.1.9节的 (14)】

关于试验函数空间的定义还可参见【8】的 P2，其将 Ω 定义成的是有界开集，其边界是无限可微的 $(n-1)$ 维 Riemann 流形【其相关定义可参见知乎链接：https://zhuanlan.zhihu.com/p/81945115，包括：流形、黎曼空间、黎曼流形等】。并且在【8】中将 φ_0 视为了 0，$\lim\limits_{i \to \infty} \sup\limits_{x \in K} |D^\alpha \varphi_i(x)| = 0$，其中 $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n), |\alpha| = \alpha_1 + ... + \alpha_n \geq 0, D^\alpha = \dfrac{\partial^{\alpha_1+\alpha_2+...+\alpha_n}}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}}$。

1.3.13 Sobolev 空间

其实下面的 μ 就是在试验函数空间中选的吧。

（1）设 $m(m \geq 1)$ 为整数，我们定义 $H^m(\Omega)=\{u|D^\alpha u \in L^2(\Omega), \forall \alpha, |\alpha| \leq m\}$，其中 $D^\alpha u$ 是广义导数。我们称 $H^m(\Omega)$ 为**整数阶 Sobolev 空间**。记 $H^0(\Omega)=L^2(\Omega)$。并且它是完备的【证明参见【8】的 P10】。$H^m(\Omega)$ 的范数定义为 $\|u\|_{H^m(\Omega)}=(\sum\limits_{|\alpha|\leq m}\|D^\alpha u\|^2_{L^2(\Omega)})^{\frac{1}{2}}$。我们可以证明 $H^m(\Omega)$ 是 Hilbert 空间。

（2）设 $\Omega \subset R^n$ 为开区域，m 为非负整数，$1 \leq p < +\infty$。称集合 $W^{m,p}(\Omega)=\{u|u \in L^p(\Omega), D^\alpha u \in L^p(\Omega), |\alpha| \leq m\}$ 按模 $\|u\|_{W^{m,p}(\Omega)}=\|u\|_{m,p}=\begin{cases}\sum_{|\alpha|\leq m}(\int_\Omega |D^\alpha u|^p dx)^{\frac{1}{p}}, & 1 \leq p < +\infty, \\ \sum_{|\alpha|\leq m} \operatorname{esssup}|D^\alpha u|, & p=+\infty.\end{cases}$ 构成的空间为 Sobolev 空间，记作 $W^{m,p}(\Omega)$ 或 $W^m_p(\Omega)$。

除此之外，在【8】的 P10 是这样定义的：

$W^{m,p}(\Omega)=\{u|D^\alpha u \in L^p(\Omega), \forall \alpha, |\alpha| \leq m, p>1, p \neq 2\}$。范数定义为 $\|u\|_{W^{m,p}(\Omega)}=(\sum\limits_{|\alpha|\leq m}\|D^\alpha u\|^p_{L^p(\Omega)})^{1/p}$。类似地，我们可以证明 $W^{m,p}(\Omega)$ 是完备的 Banach 空间。

（3）$H^m_0(\Omega)$ 是 $D(\Omega)$ 在 $\|\cdot\|_{H^m(\Omega)}$ 意义下的闭包。$W^{m,p}_0(\Omega)$ 是 $D(\Omega)$ 在 $\|\cdot\|_{W^{m,p}(\Omega)}$ 意义下的闭包。

（4）非整数阶 Spbolev 空间 $H^s(\mathbb{R}^n)$，其中 s 为实数【定义参见【8】的 P13，要用到【8】的 P11-12 的极降函数空间和缓增函数空间】。

理解：

(1) **弱导数**的定义为：$f \in L^p(\Omega), \alpha \in N, \int_\Omega D^\alpha f \varphi dx=(-1)^{|\alpha|}\int_\Omega f D^\alpha \varphi dx$;

(2) 当 $m=0$ 时，$W^{0,p}(\Omega)=L^p(\Omega)$;

(3) Soblev 空间中有许多重要性质, 例如: (i)$W^{m,p}(\Omega)$ 是 Banach 空间; (ii) 如果 $p<+\infty$, 则 $W^{m,p}(\Omega)$ 是可分的; (iii)$W^{m,p}(\Omega)$ 空间也可以看作是 $C^\infty(\Omega)$ 在范数 $\|\cdot\|_{m,p}$ 下的完备化空间。

1.3.14 广义函数空间

$D(\Omega)$ 上的一切广义函数【定义参见1.1.34节】所组成的集合称为广义函数空间（或分布空间），记作 $D^*(\Omega)$ 或者 $D'(\Omega)$，它是 $D(\Omega)$ 的共轭 (或对偶) 空间。

一个重要的判定定理(充要条件):设紧集 K 包含于 Ω,引入 $D_K(\Omega)=\{\varphi|\varphi\text{在}\Omega\text{上无限可微且支集包含于}K\}$。定义在 $D(\Omega)$ 上的线性形式（即有界线性泛函）$f \in D'(\Omega)$ 的充要条件是对任意紧集 $K \subset \Omega$，存在 k, c 成立 $|\langle f, \varphi\rangle| \leq c\sum\limits_{x\in K}\sup\limits_{|\alpha|\leq k}|D^\alpha\varphi(x)|, \forall \varphi \in D_k(\Omega)$【证明参见【8】的 P3 的定理 1.1.1】。

不成分布的例子参见【1】的 P3。

1.3.15 Holder 空间 $C^{k,\alpha}(\Omega)$

$C^{0,\alpha}(\Omega)=\{f \in C^0(\Omega)|\sup\limits_{x,y\in\Omega,x\neq y}\frac{|f(x)-f(y)|}{|x-y|^\alpha}<\infty\}, \|f\|_{C^{0,\alpha}(\Omega)}=\|f\|_{C^0(\Omega)}+\sup\limits_{x,y\in\Omega,x\neq y}\frac{|f(x)-f(y)|}{|x-y|^\alpha}, 0<\alpha<1$，$|f(x)-f(y)| \leq C|x-y|^\alpha$。

理解：

(1)$\forall \alpha \in [0,1], C^{0,\alpha}(\Omega)$ 为 Banach 空间；

(2)$\forall 0 \leq \alpha \leq \beta \leq 1, C^{0,\beta}(\Omega) \subset C^{0,\beta}(\Omega)$;

(3)$\alpha>1, f \in C^{0,\alpha}(\Omega) \Leftrightarrow f$ 为常函数；

(4)$0 \leq \alpha \leq \beta \leq 1$，对任意有界函数序列 $\{f_n\} \subset C^{0,\beta}(\Omega), \operatorname{supp} f_n \subset K \subset \Omega$，则 $\exists\{f_n\}$ 的一个收敛于 $C^{0,\alpha}(\Omega)$ 的子列。

2 重要不等式

不等式的巧妙使用在偏微分方程中是基本的。在学了测度后，对下面一些不等式可以放在测度论的框架下进一步升华理解，详情参见【1】的 Jensen 不等式 (P324)、Holder 不等式 (P326) 和 Minkowski 不等式 (P328)。

2.1 Cauchy-Schwarz 不等式

(级数形式) $(\sum_{i=1}^{n} x_i y_i)^2 \leq \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2$;

(积分形式) $|\int_E x(t)y(t)dt|^2 \leq \int_E |x(t)|^2 dt \int_E |y(t)|^2 dt$。**【2-1 的 1 题证明】**

2.2 Young 不等式

设 $p, q > 1, \frac{1}{p} + \frac{1}{q} = 1, a, b \geq 0$，则有 $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ **【2-1 的 2 题证明】**。

2.3 Holder 不等式

(级数形式) 若 $p, q > 1$，且 $\frac{1}{p}+\frac{1}{q} = 1$，则对 C^n 中任意向量 x, y 都有 $\sum_{i=1}^{n} |x_i||y_i| \leq (\sum_{i=1}^{n} |x_i|^p)^{1/p} (\sum_{i=1}^{n} |y_i|^q)^{1/q}$;

(积分形式) 若 $p, q > 1$，且 $\frac{1}{p} + \frac{1}{q} = 1, x \in L^p(E), y \in L^q(E)$，则 $xy \in L(E)$，且 $\int_E |x(t)y(t)|dt \leq (\int_E |x(t)|^p dt)^{1/p} (\int_E |y(t)|^q dt)^{1/q}$。**【2-2 的 3 题证明】**

2.4 Minkowski 不等式

(级数形式) 设 $p \geq 1, x, y \in C$，则有 $(\sum_{i=1}^{n} |x_i + y_i|^p)^{1/p} \leq (\sum_{i=1}^{n} |x_i|^p)^{1/p} + (\sum_{i=1}^{n} |y_i|^p)^{1/p}$;

(积分形式) 设 $p \geq 1, x, y \in L^p(E)$，则 $x + y \in L^p(E)$，且 $(\int_E |x(t) + y(t)|^p dt)^{1/p} \leq (\int_E |x(t)|^p dt)^{1/p} + (\int_E |y(t)|^p dt)^{1/p}$。**【2-2 的 4 题证明】**

理解：

(1) 该不等式左端实际就是 L_p 空间的操作，因此该不等式在判定两 L_p 空间中的函数叠加后，仍属于 L_p 空间。联系量子力学的操作。

(2) 若级数收敛，那么级数形式对于无限维情形也同样适用，积分形式要给定区间 ($x, y \in C[a, b]$)。上述所有不等式，在复空间中，取模后同样适用。

2.5 Schwarz 不等式

设 $(X, \langle\cdot,\cdot\rangle)$ 是内积空间，则 $\forall x, y \in X$，有 $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$。**【2-3 的 5 题证明】**

2.6 Bessel 不等式

设 $\{e_k\}$ 是内积空间 X 中的一个标准正交系，则 $\forall x \in X, n \in N$，有:

(1) $\|x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2$;

(2) $\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leq \|x\|^2$。

【2-3 的 6 题证明，及最后的体会】

理解:

当 Bessek 不等式中的等号成立时，则称为该等式为 **Parseval 等式**，再由 $\|x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k\|^2 = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2$，可以说明为什么内积空间中的任意元素可以按一个标准正交系进行展开。

2.7 Jensen 不等式

【参见【1】的 P324 的定理 10.7.1】

2.8 Hanner 不等式

【参见【1】的 P332 的定理 10.7.5】

3 在离散域和连续域上构造常用的空间

3.1 度量 (距离) 空间

1. $(K^n, d_p(x,y) = (\sum\limits_{k=1}^{n} |x_k - y_k|^p)^{\frac{1}{p}}, 1 \le p < \infty)$、$(K^n, d_\infty(x,y) = \max\limits_{1\le k\le n} |x_k - y_k|)$；【2.1-1 的 1 题和 2 题证明】

2. $(R^\infty, d(x,y) = \sum\limits_{k=1}^{\infty} \dfrac{|x_k - y_k|}{2^k(1+|x_k - y_k|)})$；【3.1-1 的 3 题证明，对于 $R^\infty, d_p(x,y)$ 和 $d_\infty(x,y)$ 全部无意义】

3. $(l^\infty, d_\infty(x,y) = \sup\limits_{k\in N} |x_k - y_k|)$；【3.1-1 的 4 题证明】

4. $(l^p, d_p(x,y) = (\sum\limits_{k=1}^{\infty} |x_k - y_k|^p)^{\frac{1}{p}})$；

5. $(C([a,b]), d_p(x,y) = (\int_a^b |x(t) - y(t)|^p dt)^{\frac{1}{p}}, 1 \le p < \infty)$、$(C([a,b]), d_\infty(x,y) = \max\limits_{t\in[a,b]} |x(t) - y(t)|)$；【3.1-2 的 6 题证明】

6. $(C^k([a,b]), d(x,y) = \sum\limits_{i=0}^{k} \max\limits_{t\in[a,b]} |x^{(i)}(t) - y^{(i)}(t)|)$；

7. $(L^p([a,b]), d_p(x,y) = (\int_a^b |x(t) - y(t)|^p dt)^{\frac{1}{p}}, 1 \le p < \infty)$。

【3.1-2 的 7 题】

空间的常见完备性：

（1）完备的距离空间：(K^n, d_2)、(l^p, d_p)、(l^∞, d_∞)、$(C([a,b]), d_\infty)$、$(L^p([a,b]), d_p)$；

（2）不完备的距离空间：$(C([a,b]), d_p)$。【3.1-3 的 8 题证明】

空间的常见可分性：

（1）可分的距离空间：(K^n, d_2)、(l^p, d_p)、$(C([a,b]), d_\infty)$、$(L^p([a,b]), d_p)$；

（2）不可分的距离空间：(l^∞, d_∞)。

理解：

(1) 若 $1 \le p < q < \infty, k > 1$，则上述空间具有下列包含关系：$l^p \subset l^q \subset l^\infty$,$C^k([a,b]) \subset C([a,b]) \subset L^q([a,b]) \subset L^p([a,b]) \subset L([a,b])$，故小的空间也可以作为大空间的子空间而成为距离空间【3.1-2 的 5 题证明】；

(2) 可以看出，为了使得定义后的距离能够满足“非负唯一性”，因此在定义距离时都会加上绝对值符号来保障；

(3)$d_\infty(x,y)$ 其实可以看作是 $d_p(x,y)$ 在 $p \to \infty$ 时的情形，因为 $\max\limits_{1\le k\le n} |x_k - y_k| \le (\sum\limits_{k=1}^{n} |x_k - y_k|^p)^{\frac{1}{p}} \le n^{\frac{1}{p}} \max\limits_{1\le k\le n} |x_k - y_k|$，由夹逼定理有 $\lim\limits_{p\to\infty} d_p(x,y) = \max\limits_{1\le k\le n} |x_k - y_k|$；

(4) 证明非负唯一性时，因为是当且仅当，所以需要对两个方向都证明，特别地，对于函数空间，需要证明 $d(x,y) = 0$ 时，不仅需要证明 $x(t) = y(t)$，还需要近一步证明 $x(t) \equiv y(t)$，因为“等于”是指一般情况下需要满足一定条件时才成立的，而“恒等于”指对区间中任意都成立，即无条件的；

(5) 记住上述空间的常见完备性和可分性，实际上也就对应了后面将介绍的部分一步步继续缩小后的空间。讨论空间的完备性简而言之就是讨论空间（范围）内对极限运算的封闭性。

3.2 赋范空间

1. $(K^n, \|x\|_p = (\sum\limits_{k=1}^{n} |x_k|^p)^{\frac{1}{p}}, 1 \le p < \infty)$、$(K^n, \|x\|_\infty = \max\limits_{1\le k\le n} |x_k|)$；【《矩阵理论》P51 证明】

2. $(K^{m\times n}, \|A\|_{m_p}=(\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|^p)^{\frac{1}{p}}, 1\le p<\infty)$、$(K^{m\times n}, \|A\|_{m_\infty}=\max_{i,j}\{|a_{ij}|\}, 1\le i\le m, 1\le j\le n)$；

3. $(l^\infty, \|x\|_\infty=\sup_{i\in N}|x_i|)$；

4. $(l^p, \|x\|_p=(\sum_{k=1}^{\infty}|x_k|^p)^{\frac{1}{p}}, 1\le p<\infty)$；【3.2-1 的 1 题证明】

5. $(C([a,b]), \|x\|_\infty=\max|x(t)|)$；

6. $(C^k([a,b]), \|x\|=\sum_{i=0}^{k}\max_{t\in[a,b]}|x^{(i)}(t)|)$、$(C^k([a,b]), \|x\|_\infty=\max_{0\le i\le k}\{\max_{t\in[a,b]}|x^{(i)}(t)|\})$；

7. $(C^\infty([a,b]), \|x\|=\max_{t\in[a,b]}\sum_{i=1}^{\infty}|x^{(i)}(t)|)$；

8. $(B(\Omega), \|x\|_0=\sup_{t\in\Omega}|x(t)|)$；

9. $(L^p([a,b]), \|x\|=(\int_a^b|x(t)|^p dt)^{\frac{1}{p}})$；

10. $(C^{n,p}([a,b]), \|x\|_{n,p}=(\int_a^b\max\sum_{i=1}^{n}|x^{(i)}(t)|^p)^{\frac{1}{p}})$。

【3.2-1 的 2 题】

理解：

(1) 证明定义的范数不能构成赋范空间，举例反证即可，比如 $K^n, \|x\|_p=(\sum_{k=1}^{n}|x_k|^p)^{\frac{1}{p}}, 0<p<1$ 就不能构成赋范空间，在 R^2 上取 $x=(1,0)^T, y=(0,1)^T, \|x+y\|_{\frac{1}{2}}\ne\|x\|_{\frac{1}{2}}+\ne\|y\|_{\frac{1}{2}}$ 即可反证。从这个例子也提醒，上述所有定义的 p 没说必须为整数奥；

(2)$\|x\|_p$ 又称为是 p 范数或 Holder 范数，其中的特殊形式 $\|\cdot\|_2$ 称为 Euclid 范数。

3.3 Banach 空间

上述 3.2 中，1、2、3、4、5、6、8、9 在其各自的范数定义下是 Banach 空间。【3.3-1 的 1 题和 2 题】

3.4 内积空间

1. $(K^n, \langle x,y\rangle=\sum_{k=1}^{n}x_k\overline{y_k})$；

2. $(l^2, \langle x,y\rangle=\sum_{k=1}^{\infty}x_k,\overline{y_k})$；

3. $(L^2([a,b]), \langle x,y\rangle=\int_a^b x(t)\overline{y(t)}dt)$；【3.4-1 的 1 题】

3.5 Hilbert 空间

上述 3.4 中的 1、2、3 在其各自的内积定义下是 Hilbert 空间。

理解：

(1)$C([a,b])$ 作为 $L^2([a,b])$ 的子空间在 3.4-3 的内积下是内积空间，但不是 Hilbert 空间

3.6 算子空间

1. R^n 的对偶空间是 R^n；l^2 的对偶空间是 l^2；$L^2([a,b])$ 的对偶空间是 $L^2([a,b])$。
2. K^n 的对偶空间是 K^n；
3. l^1 的对偶空间是 l^∞；【3.6-5 的 17 题证明】
4. l^p 的对偶空间是 l^q，其中 $\frac{1}{p}+\frac{1}{q}=1$。
5. $L^p([a,b])$ 的对偶空间是 $L^q([a,b])$，其中 $\frac{1}{p}+\frac{1}{q}=1$。

【《应用泛函分析》P150-151 证明，及更多其他的例子】

常见的自反空间包括：Hilbert 空间,$K^n, l^p, L^p([a,b])(1<p<\infty)$

【《应用泛函分析》P154 证明。同时也要会识别哪些空间不是自反空间，如 $l^1, L^1(\Omega)$ 空间】。

3.7 拓扑空间

1. 离散拓扑空间 $(X, 2^X)$；

2. 平凡拓扑空间 $(X, \emptyset)$；

3. R 上的普通拓扑是指 R 的所有可以表示成开区间的并的集合组成的集族；

4. $\mathbb{R}^n$ 上的普通拓扑 $\mathcal{U}$ 是由这样的子集组成的子集族 $\mathcal{U}$: $G \in \mathcal{U} \Leftrightarrow \forall x \in G, \exists \epsilon > 0, B(x,\epsilon) \subset G$，其中 $B(x,\epsilon) = \{y \in \mathcal{R}^n : \|y-x\|_2 < \epsilon\}$。（这里本质上就是对 $\mathcal{U}$ 中的 G 做了个限定）【3.8-1 的第 1 题证明】

5. 由度量空间诱导的拓扑空间。【4.8-1 的 1 题】

更多的拓扑空间（如：$\mathcal{R}$ 上的上拓扑、$\mathcal{R}$ 上的下拓扑、$C[0,1]$ 上的最大绝对值拓扑 (或一致收敛拓扑)、$C[0,1]$ 上的绝对值积分拓扑等等)【参见《数学分析讲义 (第二册)》P2-4】。

理解：

离散拓扑空间和平凡拓扑空间是拓扑空间概念的两个极端：最强和最弱的拓扑。

3.8 Sobolev 空间

1. $W^{m,2}(\Omega) = H^m(\Omega) = \{f \in L^2(\Omega) | D^\alpha f \in L^2(\Omega), |\alpha| \le m\}$；

2. $H_0^m(\Omega) = \{f \in H^m(\Omega) | f/\partial\Omega = 0\} = \{f \in C_c^\infty(\Omega)\}$。

【3.7-1 的 1 题】

理解：

(1)$H^m(\Omega)$ 是 Hilbert 空间，$\langle f,g\rangle_{H^m(\Omega)} = \sum\limits_{|\alpha|\le m} \langle D^\alpha f, D^\alpha g\rangle_{L^2(\Omega)}$；【3.7-1 的 2 题证明】

(2)$f \in H^{-m}(\Omega)$ 为 $H_0^m(\Omega)$ 的对偶空间，$f(g) = \int_\Omega fg dx, g \in H_0^m$。$f \in H^{-m}(\Omega)$ 当且仅当 $f = \sum\limits_{|\alpha|\le m} D^\alpha g, g \in H_0^m(\Omega)$。

4 定理

4.1 度量 (距离) 空间定理

1. 在距离空间中, (1) 同一个收敛点列的极限存在且唯一；(2) 两个不同收敛点列 $x_n \to x, y_n \to y$ 有 $d(x_n,y_n) \to d(x,y)$ 成立。【《应用泛函分析》P44 证明，反证法 + 三角不等式 + 夹逼定理，《应用泛函分析 P44-45》证明，扩展的三角不等式 + 夹逼定理】【4.1-1 的 1 题证明】

理解：

从这两定理可以看出为什么在实际的机器学习中要定义这些特殊的空间（即距离空间基础上进一步缩小的那些空间）了，否则连收敛的唯一性都不一定能够满足，因为正定性和三角不等式，以及对称性，是证明过程中所必须的。

2. 压缩映射原理: 设 X 是完备的距离空间，$T: X \to X$ 是一压缩映射，则 T 在 X 中有唯一的不动点。【4.1-1 的 2 题证明，还有《数学分析讲义（第二册)》P39-40 证明】【4.1-2 的 3 题】

理解：

(1) 该定理中的完备性条件是必不可少的；

(2) 从该定理的证明过程中还可以得到迭代法求不动点时的收敛速度的估计式 $d(x_n,x^*) \le \dfrac{\theta^n}{1-\theta} d(x_0, Tx_0)$；

(3) 该定理其实是 Banach 压缩映射原理，是最原始和经典的版本。后来又有许多著名的不动点定理，可参见【张世清的《应用泛函分析》P188-213 的第 10 章】。

3. 若 T^n 是压缩映射，则 T 在其空间中有唯一的不动点【4.1-2 的 4 题证明】。

4. 设 X 是度量空间，$x \in X, E \subset X$，则 $x \in \overline{E}$ 的充分必要条件是：有一个收敛于 x 的 E 的点列 $\{x_n\}$。特别地，E 是闭集的充分必要条件是：E 的任何收敛点列的极限都在 E 中。【《数学分析讲义 (第二册)》P35-36 证明】

5. 设 X 是完备度量空间，$E \subset X$，则 E 作为度量空间 X 的度量子空间是完备的，当且仅当 E 在 X 中是闭集。【《数学分析讲义 (第二册)》P36-37 证明】

6. Baire 纲定理： 设 X 是完备度量空间，$\{G_n \subset X : n \in N\}$ 是一串 X 的稠密开集，则 $\bigcap_{n=1}^{\infty} G_n$ 在 X 中稠密【《数学分析讲义 (第二册)》P45 习题 7.5.3(i)】。【4.1-2 的 5 题证明，重点在于自己构造出 Cauchy 列】

理解：

（1）Baire 纲定理的其它刻画：（a）非空完备度量空间一定是第二纲的;（b）若 $\{M_n\}$ 是 X 中可数个处处不稠密集，则 $X \neq \bigcup\limits_{n=1}^{\infty} M_n$；(c) G_n^C 处处不稠密。

（2）这个定理在泛函分析等整个分析学中是最基本的定理，十分重要！因为它可以进一步证明许多其他的定理。比如【《数学分析讲义（第二册）》P44-45 的 7.5.3 题】。又比如泛函分析的“共鸣定理”（即“一致有界”定理），或者是下面的 Osgood 定理。

7. Osgood 定理： 设 X 是个完备度量空间，又设对于每个点 $x \in X$，连续实值函数列 $\{f_n : n \in N\}$ 在 x 点的值集 $\{f_n(x) \in \mathbb{R} : n \in N\}$ 在 $\mathbb{R}$ 中有界，则有个非空开集 $V \subset X$ 和 $M \in \mathbb{R}$，使得 $\forall x \in V, \forall n \in N, |f_n(x)| \leq M$。【《数学分析讲义 (第二册)P39》证明】

4.2 赋范空间定理

1. n 维线性空间 X 上的任意两个范数都是等价的。【4.2-1 的 1 题证明，《应用泛函分析》P85 证明、《矩阵理论》P55 向量情形证明、《矩阵理论》P57 矩阵情形证明】【4.2-1 的 2 题】

理解：

(1) 该定理支撑了科研中的一个套路，那就是在具体的问题中，我们常常挑选一个最便于对该问题进行计算操作的范数；

(2) 该定理强调 n 维有限空间，对于无穷维空间而言，该定理不一定成立；

(3) 在该定理的证明过程中还涉及到另外一个定理：范数 $\|x\|$ 是 x 的连续泛函【《应用泛函分析》P77 证明】。

4.3 Banach 空间定理

1. 设 X 是赋范空间，则 X 为 Banach 空间的充分必要条件是：$\forall \{x_n\} \subset X, \sum\limits_{k=1}^{\infty} \|x_k\|$ 收敛 $\Rightarrow \sum\limits_{k=1}^{\infty} x_k$ 收敛。【《应用泛函分析》P82 证明】

2. 有限维线性赋范空间是完备、可分的。

3. Riesz 引理： 设 Y 是赋范空间 X 的闭子空间，$Y \neq X$，则 $\forall r \in (0,1), \exists x_0 \in X$，使得 $\|x_0\| = 1$，且 $d(x_0, Y) \geq r$。【《应用泛函分析》P88 证明】

4. 开映像原理： 设 X, Y 是 Banach 空间，$T \in B(X,Y)$。若 T 是满射，则 T 是开映射。

5. 闭图像定理： X, Y 是 Banach 空间，$T : X \to Y$ 是线性算子，若 T 是闭算子，则 T 为有界算子（即连续）。【《应用泛函分析》P136 证明】

6. 一致有界原理： 设 X 是 Banach 空间，Y 是赋范空间，算子族 $F \subset B(X,Y)$，若 $\forall x \in X$，有 $\sup\limits_{T \in F} \|Tx\| = M(x) < \infty$，则算子族 F 一致有界。【《应用泛函分析》P138 证明】

7. 共鸣定理： 设 X 是 Banach 空间，Y 是赋范空间，若算子族 $F \subset B(X,Y)$ 不是一致有界的额，即 $\sup\limits_{T \in F} \|T\| = \infty$，则 $\exists x_0 \in X$，使得 $\sup\limits_{T \in F} \|Tx_0\| = \infty$。此时 x_0 为算子族 F 的共鸣点。

4.4 内积空间定理

1. 内积与由内积导出的范数具有下列等式关系：对于实内积空间有 $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$，对于复内积空间有 $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2)$。

2. 平行四边形法则： 赋范空间 X 是内积空间的充分必要条件是：其中的范数满足平行四边形法则，即 $\forall x, y \in X$, 有 $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$。【4.4-1 的 1 题证明】

理解：

线性赋范空间 + 范数满足平行四边形法则 = 内积空间、Banach 空间 + 范数满足平行四边形法则 =Hilbert 空间，很好的一个判别定理；

3. 若在内积空间 $(X, \langle\cdot,\cdot\rangle)$ 中有 $x_n \to x, y_n \to y$，则 $\lim\limits_{m,n\to\infty}\langle x_n, y_n\rangle = \langle x, y\rangle$。【《应用泛函分析》P93-94 证明】

理解：

也就是说，内积是连续泛函。

4.5 Hilbert 空间定理

1. 设 $\{e_n\}$ 是 Hilbert 空间 H 中的一个标准正交系，则 $\forall x \in H$,x 的 Fourier 级数 $\sum\limits_{k=1}^{\infty}\langle x, e_k\rangle e_k$ 在 H 中收敛。【《应用泛函分析》P104】

2. 每个 Hilbert 空间都是自反的。【4.5-1 的 1 题证明】

3. 设 H 是数域 K 上的 Hilbert 空间，(1) 若 $K = R$，则 H^* 与 H 等距线性同构；(2) 若 $K = C$，则 H^* 与 H 等距共轭线性同构。【《应用泛函分析》P149 证明】

4.6 算子空间定理

1. 设 T 是线性赋范空间 X 到线性赋范空间 Y 的线性算子, 则下列条件等价: (1)T 一致连续; (2)T 连续; (3)T 在 $x = 0$ 处连续; (4)$\exists k > 0, \forall x \in X, \|x\| \le 1, \|Tx\| \le k$; (5)$\exists k > 0, \forall x \in X, \|Tx\| \le k\|x\|$。【《应用泛函分析》P123 证明】

2. 设 X, Y 是数域 K 上的赋范空间，记 $B(X,Y)$ 为从 X 到 Y 的有界算子的集合，则 $B(X,Y)$ 按算子范数称为一个赋范空间。此外，若 Y 是 Banach 空间，则 $B(X,Y)$ 也是 Banach 空间。【《应用泛函分析》P127 证明】

3. Riesy 表示定理: 设 f 是 Hilbert 空间 H 上的有界线性泛函，则存在唯一的 $y \in H$，使得对 $\forall x \in H$，$f(x) = \langle x, y\rangle$，且有 $\|f\| = \|y\|$。【《应用泛函分析》P143 证明】【3.6-2 的 8 题，若不再同一个空间，则“存在唯一”会被宽泛到“任意”】

4. 假设 T 是 Hilbert 空间 H 上的有界线性算子, 则 T 的共轭 T^* 有如下性质: (1)$I^* = I$; (2)$(T+S)^* = T^* + S^*$；(3)$(\alpha T)^* = \alpha T^*$；(4)$(TS)^* = S^*T^*$；(5)$T^{**} = T$；(6)$\|T^*\| = \|T\|$；(7) 若 T 可逆，则 T^* 和 $(T^*)^{-1} = (T^{-1})^*$ 互逆。

5. Hilbert 空间中自共轭算子的特征值为实数，对自共轭算子相应于不同特征值的特征向量是正交的。特别地，酉算子的特征值满足 $|\lambda| = 1$。

6. 算子的 Gateaux 导数如果存在，则必唯一。

7. 若在 x_0 的邻域内，f 是 n 次 Frechet 可微的，则 f 也是 n 次 Gateaux 可微的，且它们相应的微分相等。反之不一定成立 (或者说一般都不成立)。

8. 设 X, Y 是赋范空间，$T: X \to Y$ 是紧算子，则 T 的对偶算子 $T^*: Y^* \to X^*$ 是紧算子。【《应用泛函分析》P176 证明】

9. 设 X, Y 是赋范空间，$T: X \to Y$ 是紧算子，$\{x_n\} \subset X, x \in X$，当且仅当 $x_n \xrightarrow{\omega} x \Rightarrow Tx_n \to Tx$。【《应用泛函分析》P176 证明】

理解：

即 T 能够把 X 中的弱收敛点列映成 Y 中的强收敛点列。

10. 设 X 是赋范空间，Y 是 Banach 空间，$\{T_n\} \subset C(X,Y)$。若 $\{T_n\}$ 依范数收敛于 T，则 T 也是紧算子。

11. 自共轭紧算子的相互不同的非零特征值的集合 $\{\lambda_n\}$ 或是有限的，或满足 $\lim\limits_{n\to\infty}\lambda_n = 0$。

12. Hilbert-Schmidt 定理: 每一个无穷维 Hilbert 空间 X 的紧自共轭算子 T 都有相应于非零特征值 $\{\lambda_n\}$ 的正交特征向量系统 $\{u_n\}$，使得 $\forall x \in X$，都有唯一的表达式 $x = \sum\limits_{n=1}^{\infty}\lambda_n u_n + v$ 成立，其中 $Tv = 0$。如果 T 有无穷多互补相同的特征值 $\{\lambda_n\}$，则 $\lambda_n \to 0\,\mathrm{a.s.}\,n \to \infty$。

13. 自共轭紧算子的谱定理： 设 T 是无穷维 Hilbert 空间 X 上的紧自共轭算子，则 X 中存在一个完备正交系（标准正交基）$\{u_n\}$ 组成的 T 的正交向量，近一步，$\forall x \in X, Tx = \sum\limits_{n=1}^{\infty} \lambda_n \langle x, u_n \rangle u_n$，其中 $\{\lambda_n\}$ 是特征值。

4.7 拓扑空间定理

1. 度量空间诱导拓扑空间： 设 (X, ρ) 为度量空间，$A \in 2^X$，定义在 A 上取正实数值或 ∞ 的函数之全体记做 $\mathcal{P}_A$。给定了 $A \in 2^X$ 和 $r \in \mathcal{P}_A$，令 $G_{A,r} = \bigcup_{a \in A} B(a, r(a))$，则 X 的子集族 $\mathcal{T} = \{G_{A,r} : A \in 2^X, r \in \mathcal{P}_A\}$ 满足拓扑空间的 3 个条件。我们将 $(X, \mathcal{T})$ 简称为度量空间 (X, ρ) 的 (诱导) 拓扑空间。【4.8-1 的 1 题证明】

推论：

（1）设 (X, ρ) 是度量空间，$G \subset X$ 是开集的充分必要条件是: $\forall x \in G, \exists r > 0, B(x, r) \subset G$;

（2）设 (X, ρ) 为度量空间，$x \in X$，则单点集 $\{x\}$ 是闭集；

（3）开集是开球的并；

（4）开球是开集。

2. 设 X 是 Hausdorff 空间，$\{x_n\}$ 是 X 中的一个点列。假若两个等式 $\lim_{n\to\infty} x_n = x$ 和 $\lim_{n\to\infty} x_n = y$ 同时成立，则 $x = y$。【《数学分析讲义 (第二册)》 P35 证明，反证法】

3. 设 X 是度量空间，$x \in X, E \subset X$，则 $x \in \overline{E}$ 的充分必要条件是：有一个收敛于 x 的 E 的点列 $\{x_n\}$。特别，E 是闭集的充分必要条件是：E 的任何收敛点列的极限都在 E 中。（简而言之就是 $x \in \overline{E} \Leftrightarrow \exists \{x_n\} \subset E$，有 $\lim_{n\to\infty} x_n = x$）【《数学分析讲义（第二册)》 P35-36 证明】

4. 设 X 是度量空间，Y 是拓扑空间，映射 $f : X \to Y$ 在点 $x \in X$ 连续的充分必要条件是：对于任何极限为 x 的 X 中的收敛点列 $\{x_n\}$，必有 $\lim\limits_{n\to\infty} f(x_n) = f(\lim\limits_{n\to\infty} x_n)$。【《数学分析讲义 (第二册)》 P36 证明】

5. 设 X 是拓扑空间，$K \subset X$，则 K 是紧集，当且仅当 K 看做 X 的拓扑子空间时是紧空间。【《数学分析讲义（第二册)》 P80-81 证明】

6. （1）设 X 是拓扑空间，$K \subset X$ 是紧集，$F \subset K$ 是闭集，则 F 也是紧集。【4.8-2 的 2 题证明】

（2）设 X 是 Hausdorff 空间，$K \subset X$ 是紧集，则 K 是闭集。【4.8-2 的 3 题证明】

推论：

R 的子集是紧的，当且仅当它是有界闭集。

理解：

上述两个定理说明了紧集和闭集之间的关系。对于（1），体会这里的前提条件 F 要是闭集。为什么呢? 既然 $F \subset K \subset \bigcup_{j=1}^{n} G_{\alpha_j}$，为何还要求 F 是闭集呢? 因为 K 的任意开覆盖不一定是 F 的任意开覆盖!

7. 设 X 是紧空间，Y 是拓扑空间，$f : X \to Y$ 是连续映射，则 $f(X)$ 是紧集。【《数学分析讲义（第二册)》 P53-54 证明】

推论：

(1) 设 X 是紧空间，$f : X \to R$ 是连续映射，则 $f(X) \subset R$ 是有界集，且 $\sup f(X), \inf f(X) \in f(X)$，换言之，$f$ 在 X 上是有界函数，且达到最大与最小值。【《数学分析讲义（第二册)》 P54 证明】

（2）设 X 是紧空间，Y 是 Hausdorff 空间，映射 $f : X \to Y$ 是连续双射，则 f^{-1} 连续，换言之，f 是同胚的。

理解：

连续映射可以将紧集映射称紧集，将闭区间上的连续函数有界的性质推广到了高维的拓扑空间中。

8. 设 (X, ρ) 是紧度量空间，(Y, d) 是度量空间，$f : X \to Y$ 是连续映射，则 f 在 X 上是一致连续的。换言之，$\forall \varepsilon > 0, \exists \delta > 0\ \forall x, y \in X, \rho(x, y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$。【《数学分析讲义（第二册)》 P54-55 证明】

理解：

（1）上述定理为数学分析中“康托定理”的推广；

（2）紧度量空间的定义：X 是紧的，则有 X 的有限个点 $x_1,x_2,...,x_n$，使得 $X\subset\bigcup\limits_{j=1}^{n}B(x_j,\delta(x_j))$。

9. 设 E 是度量空间 X 的子集。若集合 E 是全有界的，则对于任何 $\varepsilon>0$，E 有一个子集，它是 E 的有限 ε-网，换言之，有有限集 $\{x_1,...,x_n\}\subset E$，使得 $E\subset\bigcup\limits_{j=1}^{n}B(x_j,\varepsilon)$。【《数学分析讲义（第二册)》P56 证明】

10. X 是度量空间，则以下三个关于 X 的条件中有一条成立时，另两条也成立：（1）X 是紧的；（2）X 是列紧的；（3）X 是全有界且完备的。（该定理是对度量空间的紧性的三种等价的刻画）【《数学分析讲义（第二册)》P58-60 证明】

引理：

在证明该定理时，需要用以下两条引理做铺垫：

（1）有收敛子列的 Cauchy 列必收敛；（这个可以机器学习损失函数非凸优化中所得收敛联系起来）【《数学分析讲义（第二册)》P56-57 证明】

（2）设 $\{x_n\}$ 是度量空间 X 的一个点列。$y\in X$ 是 $\{x_n\}$ 的某子列的极限的充分必要条件是 $\forall\varepsilon>0$，$\{m\in N:x_m\in B(y,\varepsilon)\}$ 是无限集。（对收敛点列的一种新的刻画）【《数学分析讲义（第二册)》P57-58 证明】

理解：

假若度量空间是某个具体的函数空间，它的紧子集的刻画在函数论或微分方程理论中常常十分有用。

11. 设 X 和 (Y,ρ) 分别是紧拓扑空间和度量空间。引进映射 $d:C(X,Y)\times C(X,Y)\to R$ 如下：$\forall f,g\in C(X,Y),d(f,g)=\sup\limits_{x\in X}\rho(f(x),g(x))$，则 $(C(X,Y),d)$ 是个度量空间。如上定义的 d 常称为 $C(X,Y)$ 上的**一致度量**，它是最大绝对值范数的推广。若 Y 完备，则 $(C(X,Y),d)$ 也完备。【《数学分析讲义（第二册)》P61-62 证明】

12. Arzela-Ascoli 定理：

设 X 是紧空间，(Y,ρ) 是紧度量空间，则度量空间 $(C(X,Y),d)$ 的子集 $\mathcal{F}$ 是全有界的，当且仅当 $\mathcal{F}$ 是等度连续的。【《数学分析讲义（第二册)》P62-63 证明】

理解：

该定理讨论的是函数空间，定理中的条件都很强，并且由该定理还可以直接推出 $(C(X,Y),d)$ 是完备的。

推论：

（1）设 X 是紧空间，(Y,ρ) 是紧度量空间，$\mathcal{F}\subset C(X,Y)$。若 $\mathcal{F}$ 是等度连续的，则 $\mathcal{F}$ 的任何点列都有 $C(X,Y)$ 中的收敛子列；【《数学分析讲义（第二册)》P64 证明】

（2）设 X 是紧空间，$\mathcal{F}\subset C(X)$ 是等度连续的，且对于任何 $x\in X$，有 $M_x\in R$ 使得 $\forall f\in\mathcal{F}(|f(x)|\le M_x)$，则 $\mathcal{F}$ 中的任何函数列必有一致收敛的子列。【《数学分析讲义（第二册)》P64 证明】

13. Stone-Weierstrass 逼近定理：

设 X 是紧空间，$A\subset C(X)$ 具有以下性质：

（a）$\forall f,g\in A,\forall a,b\in R$，$af+bg\in A$；

（b）$\forall f,g\in A$，$fg\in A$；

（c）$\forall x,y\in X$，$x\neq y\Rightarrow\exists f\in A,f(x)\neq f(y)$；

（d）$\forall c\in R$，常函数 $c\in A$，

则 $\overline{A}=C(X)$。

【《数学分析讲义（第二册)》P71-72 证明】

引理：

在证明该定理时，需要用以下两条引理做铺垫：

（1）对于任何正数 M，在闭区间 $[-M,M]$ 上，有一串多项式 $\{q_n\}$ 一致地收敛于 $|x|$；（这是 Weierstrass 逼近定理的特殊情形）【《数学分析讲义》（第二册）P69】

（2）首先引入记号：对于 $a,b\in R$，记 $a\vee b=\max(a,b)$ 和 $a\wedge b=\min(a,b)$。若 $f,g:X\to R$，定义 $f\vee g:X\to R$ 为 $\forall x\in X,(f\vee g)(x)=f(x)\vee g(x)$，和 $f\wedge g:X\to R$ 为 $\forall x\in X,(f\wedge g)(x)=f(x)\wedge g(x)$。

（Stone）设 X 是紧空间，$\mathcal{L} \subset C(X)$ 具有以下性质：

（a）$\forall f,g \in \mathcal{L}, \forall a,b \in R, af+bg \in \mathcal{L}$；（即要求 $\mathcal{L}$ 是个向量（线性）空间）

（b）$\forall f,g \in \mathcal{L}, f \vee g, f \wedge g \in \mathcal{L}$；（即要求 $\mathcal{L}$ 是个格，满足条件 (a) 和 (b) 的集合称为向量格）

（c）$\forall x,y \in X, x \neq y \Rightarrow \exists f \in \mathcal{L}, f(x) \neq f(y)$；（即要求 $\mathcal{L}$ 能分辨 X 的点）

（d）$\forall c \in R$，常函数 $c \in \mathcal{L}$。

则 $\overline{\mathcal{L}} = C(X)$，其中 $\overline{\mathcal{L}}$ 表示 $\mathcal{L}$ 在 $C(X)$ 的一致拓扑中的闭包。

（综上，这个定理是说，含有常函数的能分辨 X 的点的 $C(X)$ 的子向量格在 $C(X)$ 中稠密）

【《数学分析讲义（第二册)》P70-71 证明】

推论：

（1）设 K 是 R 的有界闭集，A 表示实系数多项式（看做 K 上的函数）的全体，则 $\overline{A} = C(K)$；

(2)(Weierstrass 多项式逼近定理)有界闭区间 $[a,b]$ 上的实系数多项式全体记作 $P([a,b])$, 则 $\overline{P([a,b])} = C([a,b])$。换言之，对于 $[a,b]$ 上的任何（实值）连续函数 f，总有一串实系数多项式 $\{p_n\}$ 在闭区间 $[a,b]$ 上一致收敛于 f；

（3）（Weierstrass 三角多项式逼近定理）R 上以 2π 为周期的实值连续函数可以被以下形式的“三角多项式”一致逼近：$T_n(x) = \dfrac{a_0}{2} + \sum\limits_{k=1}^{n}(a_k \cos kx + b_k \sin kx)$。其中 $a_n, b_n \in R$。

14. Urysohn lemma:

Let X be a normal space, let A and B be disjoint closed subsets of X. Let $[a,b]$ be a closed interval in the real line. Then there exists a continuous map $f: X \to [a,b]$ such that $f(x) = a$ for every x in A, and $f(x) = b$ for every x in B. 【证明参见【7】的 P207 起】

理解：

（1）刘老师说这个定理很重要，不亚于库伦定理在物理学中的地位。为了更好地理解它，需要搞清楚：(a) regular space 和 normal space 的定义见【7】的 P195。很明显地，一个 regular space 是 Hausdorff 的（将定义联系起来理解），一个 normal space 是 regular 的；（b）映射到 $[a,b]$ 是连续的，则在 X 中的原像也是连续的，有个定理忘记在哪儿了。同时再联系【7】的 P155 的 path 的定义。

（2）我的理解：将拓扑空间与 σ 代数对应后，这个定理告诉我们在此基础上通过构造概率来构造概率空间这个做法是可行的，因为其存在性在这里得到了证明。

4.8 多元微分学定理

1. 锁链法则：

设 $U \subset R^n$ 和 $V \subset R^m$，U 和 V 分别是 R^n 和 R^m 中的开集。映射 $f: U \to R^m$ 和 $g: V \to R^k$ 且 $f(U) \subset V$。又设 f 在点 p 处可微，$q = f(p)$，g 在点 q 处可微，则 $g \circ f$ 在点 p 处可微，且 $d(g \circ f)_p = dg_q \circ df_p$。用导数的语言表示，有 $(g \circ f)'(p) = g'(q)f'(p)$。【《数学分析讲义》（第二册）P105 证明】

2. 设 $A \in \mathcal{L}(E_1,...,E_n;F)$，则以下三个论述等价：

（1）$A \in L(E_1,...,E_n;F)$；（L 即表示有界的意思）

（2）A 是 $E_1 \times ... \times E_n \to F$ 的连续映射；（有界即连续，和泛函分析中得出的结论是一样的）

（3）A 在点 $(0,...,0) \in E_1 \times ... \times E_n$ 处连续。

【4-9.1 的 1 题证明。对于高维情形的证明，可以先从 $n=2$ 等特殊情形开始】

3. 设 $E_j(j=1,...,n)$ 和 F 是实数域 R(或复数域 C)上的 $(n+1)$ 个赋范线性空间，则 $(L(E_1,...,E_n;F), \|\cdot\|)$ 是个赋范线性空间。若 F 是 Banach 空间，则 $(L(E_1,...,E_n;F), \|\cdot\|)$ 也是个 Banach 空间。

【《数学分析讲义（第二册)》P186-187 证明，记得一定要先证线性，再验证是赋范】

理解：

与“$C(X,Y)$，X 紧，Y 完备 $\Rightarrow C(X,Y)$ 也完备”类似。

4. 设 $E_j(j=1,...,n)$ 和 F 是实数域 R(或复数域 C)上的 $(n+1)$ 个赋范线性空间，$(L(E_1,...,E_n;F), \|\cdot\|)$ 是个赋范线性空间。设 $m<n$，则有一个双射:$\Psi: L(E_1,...,E_m; L(E_{m+1},...,E_n;F)) \to L(E_1,...,E_n;F)$，且 Ψ 保持线性空间结构和范数不变。(即这两个重线性空间是等距同构的，所以讨论起来是一样的)

【《数学分析讲义（第二册)》P187-188 证明，体会等距的含义，即要终极证明 $\|B\| = \|A\| = \|\Psi(B)\|$】

理解：

这个定理很有用，解决了高阶微分、导数的部分。具体来说，比如在 $n=2$ 的情况下，该定理可以将乘积空间中的二次问题 $L(E_1;L(E_2;F))$（即 $L(E_2;F)$ 为第一次，外面的为第二次）转换为二重问题 $L(E_1,E_2;F)$。对于高维情形也亦如此。

5. 若映射 $f:G\to F$ 在 G 的内点 x 处是可微的，则它在点 x 处的微分是惟一确定的。【《数学分析讲义（第二册)》P189 证明】

6. 若映射 $f:G\to F$ 在 G 的内点 x 处是可微的，则它在点 x 处是连续的。【《数学分析讲义（第二册)》P189 证明】

7. 设 $U\subset R^n$ 是开集，映射 $f:U\to R^m, p\in U$。假若 f 在点 p 处是可微的，则 f 在点 p 处沿任何方向 q 都是可微的，且 $D_qf(p)=df_p(q)$（Gateaux 导数与一般导数（即函数可微概念中的）联系起来了）。特别，$D_kf(p)=\dfrac{\partial f}{\partial x_k}(p)=df_p(e_k)$。【《数学分析讲义（第二册)》P111-112 证明】

推论：

设 $U\subset R^n$ 是开集，$p\in U, q=(q_1,...,q_n)\in R^n$，假若 f 在点 p 处是可微的，则 $df_p(q)=\sum_{k=1}^{n}q_kD_kf(p)=\sum_{k=1}^{n}q_k\dfrac{\partial f}{\partial x_k}(p)$；【《数学分析讲义（第二册)》P112 证明】

理解：

函数在某点各个方向的方向导数存在（更别说只有各个偏导数了）未必能保证它一定可微。【《数学分析讲义（第二册)》P114 例子】

8. 设 $U\subset R^n$ 是开集，映射 $f:U\to R, p\in U$。假若 $\exists p$ 的邻域 V，$\forall j\in\{1,...,n\},\forall q\in V$，$D_jf(q)$ 存在且在 V 上连续，则 f 在 p 处可微。【《数学分析讲义（第二册)》P114-115 证明】

9. 中值定理：

设 $E\subset R^n$ 是 R^n 的开集，若映射 $f:E\to R$ 在 E 的每一点处都可微，则对于任何 $[a,b]\subset E$（$[a,b]$ 是个集合哦），有 $c\in[a,b]$，使得 $f(b)-f(a)=df_c(b-a)=f'(x)(b-a)$。【《数学分析讲义（第二册)》P108-109 证明】

推论：

（1）（数值多元函数的有限增量定理）设 $U\subset R^n$ 是凸的开集，映射 $f:U\to R$ 在 U 上可微，且 $\|df_p\|\le M$ 对于一切点 $p\in U$ 成立，则 $\forall a,b\in U, |f(b)-f(a)|\le M|b-a|$；

（2）（向量值多元函数的有限增量定理）设 $U\subset R^n$ 是凸的开集，映射 $f:U\to R^m$ 在 U 上可微，且 $\|df_p\|\le M$ 对于一切点 $p\in U$ 成立，则 $\forall a,b\in U, |f(b)-f(a)|\le M|b-a|$。【《数学分析讲义（第二册)》P109-110 证明】

10. 有限增量定理：

设 U 是赋范线性空间 E 的开集，$f:U\to F$ 是 U 到赋范线性空间 F 的连续映射。记 $[x,x+h]=\{p=x+\theta h:0\le\theta\le1\}$，称它是端点为 x 和 $x+h$ 的闭区间；记 $(x,x+h)=\{p=x+\theta h:0<\theta<1\}$，称它是以 x 和 $x+h$ 为端点的开区间。假设 $[x,x+h]\subset U$，且映射 f 在开区间 $(x,x+h)$ 的每一点处都可微，则以下不等式成立：$|f(x+h)-f(x)|_F\le\sup\limits_{z\in[x,x+h]}\|f'(z)\|_{L(E;F)}|h|_E$。【《数学分析讲义（第二册)》P192-194 证明】

老师说上面这个定理条件太强，且书上的证明麻烦，所以就补充了下面这个类似的，但条件更弱一些的定理：

$M\triangleq\sup_{z\in(a,b)}|f'(z)|, X=R, Y$ 为 Banach 空间，$I\triangleq[a,b]\subset R$。有定理：设 $f:I\triangleq[a,b]\to Y$ 连续，$\varphi:I\to R$ 单调递增的连续映射。若在 I 的子集 D，D 至多是可数的，使 $\forall\varepsilon\in I\backslash D$，$f$ 与 φ 在 ε 处是可导的，且 $\|f'(\varepsilon)\|\le\|\varphi'(\varepsilon)\|$，则有 $\|f(b)-f(a)\|\le\varphi(b)-\varphi(a)$。令 $\varphi(x)=M(x-a), x\in I$，则 $\|f(b)-f(a)\|\le M(b-a)$。【4.9-2 的 3 题证明】

同样还可以得到如下版本的中值定理：

设 X,Y 是赋范线性空间，$a,b\in X$，称 $S_{ab}=\{x\in X:x=a+\varepsilon(b-a),0\le\varepsilon\le1\}$ 为连结 a,b 的线段，S 为连结 X 的两点 x_0 与 x_0+t 的线段，$f:U\to Y, U\subset S$ 开邻域，连续，f 在 S 上可导，则 $\|f(x_0+h)-f(x_0)\|\le|h|\sup_{0\le\varepsilon\le1}\|f'(x_0+\varepsilon h)\|$。

推论：

设 E 和 F 是两个线性赋范空间，U 是 E 的开集。若 $A\in L(E;F)$，而映射 $f:U\to F$ 满足上述定理的条件，则 $|f(x+h)-f(x)-Ah|_F\le \sup_{z\in(x,x+h)}\|f'(z)-A\|_{L(E;F)}|h|_E$。特别，$|f(x+h)-f(x)-f'(x)(h)|_F\le \sup_{z\in(x,x+h)}\|f'(z)-f'(x)\|_{L(E;F)}|h|_E$。

理解：

这个定理及其推论用得比较多。推论即把闭区间推广到线性赋范空间上。

11. Gateaux 可导与 Frechet 可导的联系：

（1）Frechet 可导 $\Rightarrow$ Gateaux 可导。

（2）Gateaux 可导 $\Rightarrow$ Frechet 可导的条件（即反过来不一定能推出，必须得满足下列条件）：

（a）$f:U\to F, U\subset E$ 开集，点 $x\in U$，f 在 x 的邻域 V 内 G 可导，且 $f'(x)$ 关于 x 在 V 上连续，则 f 在 x 处 F 可导，且 $f'(x)=A$；

（b）f 在 x 处 G 可导，且 $\lim\limits_{t\to 0}\dfrac{f(x+th)-f(x)}{t}=f'(x)h$ 关于 $|h|=1$ 一致成立，则 f 在 x 点 F 可导。

12. Schwarz 定理：

设 $\Omega\subset R^n$ 是开集，点 $p\in\Omega$，映射 $f:\Omega\to R$。又设 D_if, D_jf 和 D_jD_if 在 p 的一个邻域中存在且它们在 p 处都是连续的，则二阶偏导数 $D_iD_jf(p)$ 存在，且 $D_iD_jf(p)=D_jD_if(p)$。【4.9-2 的 4 题证明，用连续性条件来引出极限】

引理：

证明上述定理需要用到该引理：

设 $(0,0)\in U\subset R^2$，U 是开集，映射 $f:U\backslash\{(0,0)\}\to R$。若二重极限（即把 f 看成由度量空间 $U\backslash\{(0,0)\}$）到度量空间 R 的映射后取的极限）$\lim\limits_{(x,y)\to(0,0)}f(x,y)=A$ 存在，且对于绝对值充分小的 y，极限 $\lim\limits_{x\to 0}f(x,y)$ 存在，则累次极限 $\lim\limits_{y\to 0}\lim\limits_{x\to 0}f(x,y)$ 存在，而且 $\lim\limits_{y\to 0}\lim\limits_{x\to 0}f(x,y)=\lim\limits_{(x,y)\to(0,0)}f(x,y)$。【《数学分析讲义（第二册)》P119 证明，回归极限定义的 $\varepsilon-\delta$ 语言】

理解：

该定理说明了在什么条件下，r 阶偏导数与求导的顺序无关。

13. Schwarz 定理在无限维情形的推广：

假设 E 和 F 是两个线性赋范空间，U 是 E 的一个开集，映射 $f:U\to F$ 在点 $x\in U$ 处有 n 阶导数 $f^n(x)\in L(E,...,E;F)$，则 n-线性映射 $f^{(n)}(x)$ 是对称的：对于任何集合 $\{1,...,n\}$ 到自身的双射（n 元置换）σ，有 $f^{(n)}(x)(h_1,...,h_n)=f^{(n)}(x)(h_{\sigma(1)},...,h_{\sigma(n)})$。【《数学分析讲义》（第二册）P197-198 证明】

理解：

事实上，该定理不光是 Schwarz 定理在无限维情形的推广，而且 Schwarz 定理也是该定理的推论。这两个定理之间连条件都不同，为什么说它俩有很强的联系呢？注意思考，这也是在一次作业题中做了的。

14. 设 $U=U(a)$ 是点 $a=(a_1,...,a_m)\in E=E_1\times...\times E_m$ 在 E 中的一个邻域，$f:U\to F$ 是 U 到线性赋范空间 F 的一个可微映射，则它在点 a 处关于每个变量都有偏导数，偏导数和微分（有时也称全微分）之间的关系是 $df_ah=\sum\limits_{j=1}^{m}\partial_jf_ah_j$，其中 $h=(h_1,...,h_m)\in E_1\times...\times E_m=E$。

15. 无限维空间的隐函数定理：

设 X,Y,Z 是 Banach 空间，U 是点 $(x_0,y_0)\in X\times Y$ 的一个邻域。$F:U\to Z$ 具有以下性质：（1）F 在点 (x_0,y_0) 连续；（2）$F(x_0,y_0)=0$；（3）$F'_y(x,y)$ 在 U 内存在且在点 (x_0,y_0) 连续，且 $F'_y(x_0,y_0)$ 有有界逆算子（即 $\exists M>0$，有 $|(F'_y(x_0,y_0))^{-1}|\le M$）。则存在 (x_0,y_0) 的某一个邻域，方程 $F(x,y)=0$ 有解，即由 $F(x,y)=0$ 确定 $y=f(x)$。有：

（a）若 $F(x,y)$ 在 U 上连续，$(x_0,y_0)\in U$，则 $y=f(x)$ 在 x_0 点的某一邻域内连续；

(b)若 F'_x 在 U 内存在且点 (x_0,y_0) 连续，则 $y=f(x)$ 在 x_0 点可微，且 $f'(x_0)=-[F'_y(x_0,y_0)]^{-1}F'_x(x_0,y_0)$。

（记号约定：$F'(x,y)=dF_{(x,y)}$，$F'_y(x_0,y_0)=dF_{(x_0,y_0)}(0,\cdot)$。分析学中对特殊情况的讨论一般都围绕零点）

【4.9-3 证明，证明思路包括：证 A 是一个压缩映射、证明范数、证明 A 是自己到自己的映射】

16. 反函数定理：

设 $U \subset X$ 开集，$x_0 \in U, y_0 = f(x_0)$，$f \in C^m(U;Y), m \geq 1$，若 $f'(x_0): X \to Y$ 的逆，$[f'(x_0)]^{-1}$ 存在且有界（正则），则 $\exists x_0$ 的邻域 V 与 y_0 的邻域 W，使 $f: V \to W$ 是 C^m 微分同胚（f, f^{-1} 都是 C^m 类的映射）

【证明思路：用两次隐函数定理，$F(x,y)=0$，令 $F(x,y)=f(x)-y, F'_y(x_0,y_0)=-I \Rightarrow y(x)=f$ 且 $y'(x)$ 连续，$F'_x(x_0,y_0)=f'(x_0) \Rightarrow x(y)=f^{-1}$ 且 $x'(y)$ 连续（老师的课堂板书中还显示同样可以 $\Rightarrow$ 出上半部分的结论）】

4.9 傅里叶分析定理

1. 如果 $f \in L_1[-\pi,\pi]$，且对固定的 x 与某个 $\delta > 0$（即充分小），积分 $\int_{-\delta}^{\delta} |\frac{f(x+t)-f(x)}{t}| dt$ 存在（称为 **Dini 条件**），那么函数 f 的傅里叶级数的部分和 $S_n(x)$ 在 x 处收敛于 $f(x)$。【4.10-1 的 3 题证明】

引理：

要证明上述定理，需要借助 **Riemann-Lebesgue 引理**：若函数 $\varphi \in L_1[a,b]$，则 $\lim\limits_{p\to\infty} \int_b^a \varphi(x)\sin px dx = 0$。【4.10-1 的 2 题证明】

推论：

（1）如果积分 $\int_{-\delta}^{0} |\frac{f(x+z)-f(x-0)}{2}| dz$ 与 $\int_0^{\delta} |\frac{f(x+z)-f(x+0)}{2}| dz$ 收敛，其中 x 是 f 的第一类间断点。则可以用来代替上述定理中的 Dini 条件，使得证明仍然有效。即不要求连续，有第一类间断点即可，也就是说左极限与右极限存在即可。

（2）设 f 是以 2π 为周期的有界函数，仅有第一类间断点，且 f 在每一点都有左、右导数（即 $\lim\limits_{h\to 0+} \frac{f(x-h)-f(x-0)}{-h}, \lim\limits_{h\to 0+} \frac{f(x+h)-f(x+0)}{h}$）。则 f 的傅里叶级数处处收敛，且在连续点处级数和等于 $f(x)$，在间断点处级数和等于 $\frac{f(x+0)+f(x-0)}{2}$。【4.10-2 的 4 题证明】

2. 设 $f \in L_1[-\pi,\pi]$，f 在 x 处附近满足 2 阶 Lipschiz 条件 $|f(x+t)-f(x+0)| \leq Lt^{\alpha}$（或 $|f(x-t)-f(x-0)| \leq Lt^{\alpha}$）（$\exists \delta > 0, L > 0, \alpha \in (0,1]$，当 $t \in (0,\delta]$）。则 f 的傅里叶级数在 x 处收敛于 $\frac{f(x+0)+f(x-0)}{2}$。

3. 费耶定理：

如果 f 是以 2π 为周期的连续函数，那么 f 的费耶和的序列 $\{\sigma_n\}$ 在整个数轴上一致收敛于 f。

4.W-三角多项式逼近定理：

R 上以 2π 为周期的连续（实值）函数 f 可以被三角多项式一致逼近。【4.10-3 的 9 题证明】

5. Dini 定理：

如果函数 $f \in L_1(-\infty,+\infty)$ 且 $x \in R$，在 x 处满足 Dini 条件，则 $f(x) = \frac{1}{\pi}\int_0^{\infty} d\lambda \int_{-\infty}^{\infty} f(t)\cos\lambda(t-x)dt$ 成立（$=\int_0^{\infty}(a_\lambda \cos\lambda x + b_\lambda \sin\lambda x)dx$）。【4.10-4 的 12 题证明】

6. 如果对函数 $f \in L_1(-\infty,\infty)$，有 $\int_{-\infty}^{\infty} f(x)e^{-i\lambda x}dx \equiv 0$，那么 $f(x)=0$ 几乎处处成立。【4.10-5 的 13 题证明】

4.10 测度和 Lebesgue 积分定理

1. 测度常用的简单性质：【1】的 P224 的命题 9.2.1、【1】的 P226 的推论 9.2.1、【1】的 P226 的命题 9.2.2、【1】的 P227 的定理 9.2.1、【1】的 P229 的推论 9.2.2。

2. 度量外测度的主要结果：【1】的 P244 的定理 9.5.1、【1】的 P244 的引理 9.5.1。

3. Lebesgue 可测集组成的 σ 代数与 Borel 集组成的 σ 代数之间的关系：【1】的 P246 的定理 9.5.2、【1】的 P246 的引理 9.5.2。

4. $\mathbb{R}$ 上的 Lebesgue 可测集以及 Lebesgue 外测度 m^* 是平移不变的【证明参见【1】的 P248】。

5. $\mathbb{R}$ 上的 Lebesgue 不可测集是存在的【证明参见【1】的 P248】。

6. 设 f 是一个 X 上的非负可测函数，则有一串非负可测简单函数 $\{f_n\}$，使得 $\forall n \in \mathbb{N}, \forall x \in X, f_n(x) \leq f_{n+1}(x)$，且 $\forall x \in X, f(x) = \lim\limits_{n\to\infty} f_n(x)$。若 f 是一个 X 上的非负有界可测函数，则有一串单调非负可测

简单函数 $\{f_n\}$，使得函数极限等式 $f(x)=\lim\limits_{n\to} f_n(x)$ 在 X 上是一致收敛【证明参见【1】的 P258】。

推论：

设 f 是一个 X 上的可测函数，则有一串可测简单函数 $\{f_n\}$，使得 $\forall x\in X, f(x)=\lim\limits_{n\to\infty} f_n(x)$。若 f 有界，则可选择一串可测简单函数 (f_n)，使得上述收敛是一致的【证明参见【1】的 P259】。

7. 若 f 是 R^n 上的 Lebesgue 可测函数，则 R^n 上有 Borel 可测函数 g，使得 $g=f$,a.e.【证明参见【1】的 P259-260】。

8. Lebesgue 积分的积分号与极限号的交换相关结论：围绕 $\{f_n\}$ 为一串（非负）可测（可积）函数而言的。包括：【1】的 P267 的定理 10.3.1（Beppo Levi 单调收敛定理）、【1】的 P269 的推论 10.3.1、【1】的 P269 的定理 10.3.2（Fatou 定理）、【1】的 P270 的命题 10.3.1、以及很重要的 **3 个 Lebesgue 控制收敛定理**在【1】的 P271、P272 和 P278、【1】的 P280 的定理 10.3.8（Vitali 定理，可以推论出 Lebesgue 控制收敛定理）。

其他结论还包括【1】的 P273 的定理 10.3.4（Fatou 引理的 Lieb 形式）、【1】的 P274 的引理 10.3.1、【1】的 P275 的引理 10.3.2、【1】的 P276 的定理 10.3.5 和定理 10.3.6（F.Riesz）、【1】的 P278 的定理 10.3.7（Egorov）。

9. Lebesgue 积分和 Riemann 积分的比较的相关结论，包括：【1】的 P285 的命题 10.4.1、【1】的 P287 的引理 10.4.1、最主要的【1】的 P289 的定理 10.4.1。以及【1】的 P291 的与反常积分的联系。

10. 对于二元函数或多元函数的重积分的计算，参见【1】的 P291 起的 10.5 节（**Fubini-Tonelli 定理**，有 2 个，在【1】的 P300 的定理 10.5.5 和【1】的 P302 的定理 10.5.5'），主要讨论重积分与累次积分之间的关系。

11. 对于将一维情形的 Riemann 积分换元公式推广到任意维的 Lebesgue 积分上，参见【1】的 P307 起（先讨论线性变换下测度的变换公式，进而讨论线性变换下的积分换元公式，最后才讨论一般变换下的积分换元公式）。

4.11 Lebesgue 空间定理

1.（L^p 空间的完备性） 设 X 是个非空集合，$\mathcal{A}$ 是 X 上的 σ 代数，μ 是定义在 $\mathcal{A}$ 上的测度，$f_i\in L^p(i=1,2,...)$ 是 L^p 中的 Cauchy 列，其中 $1\le p\le\infty$。换言之，$\forall\varepsilon>0,\exists N\in\mathbb{N}$，$i>N,j>N\Rightarrow|f_i-f_j|_p<\varepsilon$，则有一个 $f\in L^p$，当 $i\to\infty$ 时，有 $|f-f_i|_p\to 0$。这时，我们常称 f_i（在 L^p 中）强收敛于 f（简称收敛于 f），记做，当 $i\to\infty$ 时，$f_i\to f$（在 L^p 中）。而且，L^p 中的 Cauchy 列 $f_i\in L^p(i=1,2,...)$ 有子列 $f_{i_1},f_{i_2},...(i_1<i_2<...)$ 和一个非负函数 $F\in L^p$，使得：

（1）$|f_{i_k}(x)|\le F(x), a.e.(\mu)$；

（2）$\lim\limits_{k\to\infty} f_{i_k}(x)=f(x), a.e.(\mu)$。

【证明参见【1】的 P330-332】

理解：

该定理告诉我们 L^p 是个完备度量空间。若把 Lebesgue 积分换成 Riemann 积分，这个重要的结论是不成立的（可参见【1】的 P364 的练习 10.7.10）。这是 Riemann 积分不得不让位于 Lebesgue 积分的一个重要原因。

2.（L^p 范数的可微性定理） 设 $(X,\mathcal{A},\mu)$ 是个测度空间，$1<p<\infty, f,g\in L^p$，则定义在 $\mathbb{R}$ 上的函数 $N(t)=\int_X|f(x)+tg(x)|^p d\mu$ 是可微的，且 $N'(0)=\dfrac{p}{2}\int_X|f(x)|^{p-2}[\overline{f}(x)g(x)+f(x)\overline{g}(x)]d\mu$。【证明参见【1】的 P333】

3.（到闭凸集的投影定理）【参见【1】的 P334 的定理 10.7.7 及证明】

4.（F.Riesz 表示定理）【参见【1】的 P336 的定理 10.7.8 及证明】

5.（Radon-Nikodym 定理） 它是 F.Riesz 表示定理的一个应用，概率论中的条件期望的概念就是建立在这个定理基础上的。【参见【1】的 P340 的定理 10.7.9】

6.（Lebesgue 分解定理）【参见【1】的 P341 的定理】

7.（一些关于微积分学基本定理）【参见【1】的 P351 起】

4.12 Sobolev 空间定理

1. 设 $T\in D'(\Omega),\{T_i\}\subset D'(\Omega)$ 且 $T_i\to T$，则成立 $D^\alpha T_i\to D^\alpha T$【证明参见【8】的 P5】。

这一节以后学到或用到时在自行补充，比如“迹定理”等重要的结论。

1. 设 $\Omega\subset R^n$ 为具有 C^1 边界的有界开集，则 $W_0^{1,p}\hookrightarrow\begin{cases}L^q(\Omega), & 1\le q\le\frac{np}{n-p}, 1\le p<n,\\ L^\infty(\overline{\Omega}), & p>n.\end{cases}$，对于该定理中的第二部分，可以进一步地表示为：若 $p>n$ 且 $\Omega\subset R^n$ 是具有 C^1 边界的有界开集，则 $W_0^{1,0}(\Omega)\hookrightarrow C^\alpha(\overline{\Omega}),\alpha=1-\dfrac{n}{p}$。【《泛函分析及其应用》张世清 P32 证明】

理解：

例如 $W^{1,1}(R)$ 连续嵌入到 $W^{0,\infty}(R)=L^\infty(R)$。

2. 设 $\Omega\subset R^n$ 为具有 C^1 边界的有界开集，则对 $k\ge 2$，$W_0^{k,p}\hookrightarrow\begin{cases}L^{\frac{np}{n-kp}}(\Omega), & kp<n,\\ C^m(\Omega), & 0\le m\le k-\frac{n}{p}.\end{cases}$

5 应用

5.1 用范数来研究序列的收敛性

(1) 设 $\|\cdot\|$ 是 C^n 上的任一向量范数，$a,x^{(k)}\in C^n$，则 $\lim\limits_{k\to\infty}x^{(k)}=a\Leftrightarrow\lim\limits_{k\to\infty}\|x^{(k)}-a\|=0$。

(2) 设 $\|\cdot\|$ 是 $C^{m\times n}$ 上任一矩阵范数，$C^{m\times n}$ 中矩阵序列 $\{A^{(k)}\}$ 收敛于矩阵 A 的充分必要条件是 $\lim\limits_{k\to+\infty}\|A^{(k)}-A\|=0$。【《矩阵理论》课件 5-1 证明，以 ∞ 范数为切入，再利用范数的等价定理】

理解：

在更抽象的领域中，用范数定义序列的极限往往更方便，至少它以研究一个数列 $\{\|x^{(k)}-a\|\}$ 的极限代替了研究 n 个数列 $\{x_i^{(k)}\}$ 的极限。

5.2 变分法初步

微分学的一个最早的应用就是求极值，先是一元实值函数的极值，然后是多元实值函数的极值。建立了无穷维赋范线性空间上的微分学，当然会利用它去探讨无穷维赋范空间上的实值函数的极值问题。

1. 假设 E 是个赋范线性空间，U 是 E 的一个开集，映射 $f:U\to R$ 在 U 上有直到 $(k-1)$ 阶的导数，而在点 $x\in U$ 处有 k 阶导数 $f^{(k)}(x)\in L(E,...,E;R)$，其中 $k\ge 2$。又设 $f'(x)=0,...,f^{(k-1)}(x)=0$，而 $f^{(k)}(x)\ne 0$。则：

（a）x 是函数 f 的极值点的**必要条件**是：k 是偶数，且 $f^{(k)}(x)h^k$ 不取相异的符号；

（b）x 是函数 f 的极值点的**充分条件**是：$f^{(k)}(x)h^k$ 在单位球面 $|h|=1$ 上与零保持一个正的距离。若在单位球面上有不等式：$|h|=1\Rightarrow f^{(k)}(x)h^k\ge\delta>0$，其中 δ 是个不依赖于 h 的正数，则 x 是函数 f 的局部极小值点；若在单位球面上有不等式 $|h|=1\Rightarrow f^{(k)}(x)h^k\le\delta<0$，其中 δ 是个不依赖于 h 的负数，则 x 是函数 f 的局部极大值点。

【《非线性分析（第二册)》P200-201 证明，从 Tylor 公式切入】

2. 构造一个 Banach 空间

（1）域：$C^1(K,R)$ 表示定义在 K 上的在 K° 上一次连续可微，且一阶偏导数可连续延拓至 K 上的实值函数全体。其中 K 是 R^n 中满足条件 $K=\overline{K^\circ}$ 的紧子集；

（2）范数：$|f|_{C^1(K)}=\max\{|f|_{C(K)},|\partial_j f|_{C(K)},j=1,...,n\}$。该范数与以下定义的范数等价：$|f|'_{C^1(K)}=|f|_{C(K)}+\sum\limits_{j=1}^n|\partial_j f|_{C(K)}$。（为了讨论极值是否存在，所以要先把范数定义出来）

可以证明它们构成一个 Banach 空间。

3. Du Bois Reymond 引理：

设 $\varphi\in C([a,b],R)$，它满足条件：$\forall h\in C^\infty([a,b],R)$，$h(a)=h(b)=0\Rightarrow\int_a^b\varphi(x)h(x)dx=0$，则 $\forall x\in[a,b],\varphi(x)=0$。【《数学分析讲义（第二册)》P203 证明】

4. 引理： 设 $V=\{u:|u-p|<\delta\}$ 是 R^n 中的一个开球，则在 R^n 上有一个 C_0^∞ 类的函数 ψ，它具有如下性质：当 $u\in V$ 时，$\psi(u)>0$，而当 $u\notin V$ 时，$\psi(u)=0$。【《数学分析讲义（第二册)》P159 证明】

5. Euler-Lagrange 方程：

设 $L \in C^1(R^3, R)$（R^3 表示 L 是个三元函数）和 $f \in C^1([a,b], R)$。映射 $F : C^1([a,b], R) \to R$ 如下：$F(f) = \int_a^b L(x, f(x), f'(x))dx$。

首先我们可以得到 F 可微，且 $F'(f)h = \int_a^b [\partial_2 L(x, f(x), f'(x))h(x) + \partial_3 L(x, f(x), f'(x))h'(x)]dx$ 其中 ∂_i 表示对 L 的第 i 个自变量求偏导数运算。【《数学分析讲义（第二册)》P201-202 推导】

常常遇到这样的极值问题，我们要求 f 限制在 C^1 这样的仿射子空间上：$\{f \in C^1([a,b], R) : f(a) = A, f(b) = B\}$，即 h 应满足 $f(a) = f(a) + h(a), f(b) = f(b) + h(b)$，换言之，$h$ 应满足条件：$h(a) = h(b) = 0$。则在此条件下 F 的极值问题的解应满足条件，即为 f 应满足以下的方程：$\partial_2 L(x, f(x), f'(x)) - \frac{d}{dx}\partial_3 L(x, f(x), f'(x)) = 0$，称为 **Euler-Lagrange 方程**。【《数学分析讲义（第二册)》P202-203 推导】

推广：

Euler-Lagrange 方程很容易被推广到多维情形：设 $L \in C^1(R^{2n+1}, R)$ 和 $f_j \in C^1([a,b], R)(j = 1, ..., n)$。用下式定义了映射 $F : (C^1([a,b], R))^n \to R$：$F(f_1, ..., f_n) = \int_a^b L(x, f_1(x).f_1'(x), ..., f_n(x), f_n'(x))dx$。上式中的 $f_1, ..., f_n$ 被要求限制在 $(C^1)^n$ 的这样的仿射子空间上：$\{(f_1, ..., f_n) \in (C^1([a,b], R))^n : f_j(a) = A_j, f_j(b) = B_j, j = 1, ..., n\}$。当考虑 h 满足同样条件的仿射子空间上的极值问题时，$(f_1, ..., f_n)$ 应满足以下方程组：$\frac{\partial L}{\partial f_j} - \frac{d}{dx}(\frac{\partial L}{\partial f_j'}) = 0, j = 1, ..., n$，称为 **Euler-Lagrange 方程组**。

5.3 修正牛顿法和牛顿法

接着无限维空间的隐函数定理（4.9 的定理 15），令 $A_x(y) = y - [F_y'(x_0, y_0)]^{-1}F(x, y)$, 则有：

（1）修正后牛顿法：$y_{n+1} = A_x(y_n) = y_n - [F_y'(x_n, y_n)]^{-1}F(x, y_n)$。且 $|y_n - y^*| \leq \frac{q^n}{1-q}|y_1 - y_0|$，$q$ 是 A_x 的压缩系数，$|A_x'| \leq q$；

（2）牛顿法：$y_{n+1} = A_x(y_n) = y_n - [F_y'(x, y_n)]^{-1}F[x, y_n]$。且 $|y_n - y^*| \leq \frac{4}{M^2}(\frac{1}{2})^{2^n}$ 和 $|F_y'(x, y_1) - F_y'(x, y_2)| \leq M|y_1 - y_2|$。

【5-1 的 2 题和 3 题证明】

引理：

要证明上述第二个结论时（即牛顿法），需要用到以下定理：

记 $B_x(x_0, \delta) = \{x \in X||x - x_0| \leq \delta\}$ 和 $B_y(y_0, r) = \{y \in Y||y - y_0| \leq r\}$，设 $F : B_x(x_0, \delta) \times B_y(y_0, r) \to Z$ 连续，且满足存在 $M \geq \sqrt{\frac{2}{r}}$ 使：

（1）当 $x \in B_x(x_0, \delta), y \in B_y(y_0, r)$ 时，$F_y'(x, y) : Y \to Z$ 是正则算子，$[F_y'(x, y)]^{-1}$ 存在且 $|[F_y'(x, y)]^{-1}| \leq M$；

（2）当 $x \in B_x(x_0, \delta), y_1, y_2 \in B_y(y_0, r)$ 时，$|F_y'(x, y_1) - F_y'(x, y_2)| \leq M|y_1 - y_2|$；

（3）$|F(x, y_0)| \leq \frac{1}{M^3}$，则 $\exists y(x) \in B_y(y_0, r)$ 使 $F(x, y(x)) = 0$ 且 Newton 迭代序列收敛到 $y(x)$。

理解：

在数值分析中学过，一元函数下的牛顿法迭代公式为：$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$。类似地，当推广到上述高维情形后，导数即变成了逆算子。

5.4 经典力学中的 Hamilton 原理

参见【《数学分析讲义（第二册)》P205-212】。