

Exercise Sheet – Advanced Calculus III

Taiyang Xu*

25/09/2025, Week 3

下面的练习帮助我们证明隐函数存在定理以及反函数定理。在直击最后目标之前, 我们先有目的地回顾 (也可能是学习) 一些知识.

定义 1 (度量与度量空间). 设 X 为非空集合, $\forall x, y \in X$, 存在唯一实数 d , 记为 $d(x, y)$ 与之对应, 且满足下述三个条件

- (1) $d(x, y) \geq 0$, 且 $d(x, y) = 0 \Leftrightarrow x = y$ (正定性);
- (2) $d(x, y) = d(y, x)$ (对称性);
- (3) $d(x, z) \leq d(x, y) + d(y, z)$ (三角不等式).

则称 d 为 X 上的 (距离) 度量, (X, d) 称为度量空间 (Metric space). 在不致于引起歧义的情况下, 我们可将之简记为 X .

例子 2. 常见的度量 (空间) 有:

- 距离定义是实数中绝对值概念的一般化或抽象: (\mathbb{R}, d) , 其中 $d(x, y) = |x - y|$;
- 离散度量空间: (X, d) , 其中 $d(x, y) = 0$ 当且仅当 $x = y$, 否则 $d(x, y) = 1$;
- (\mathbb{R}, d) , 其中 $d = \frac{|x-y|}{1+|x-y|}$;
- n 维欧氏空间 (\mathbb{R}^n, d) , 其中 $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$;
- l^p 空间, $1 \leq p < +\infty$, 其中 $d(x, y) = (\sum_{i=1}^{\infty} |x_i - y_i|^p)^{\frac{1}{p}}$, $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$;
- $[a, b]$ 上连续函数的全体 $C[a, b]$, 其中 $d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$.

*School of Mathematical Sciences, Fudan University, Shanghai 200433, China. Email: tyxu19@fudan.edu.cn

定义 3 (Cauchy 列 (基本列) 与完备空间). 设 (X, d) 为度量空间, 则称 $\{x_n\} \subset X$ 为 Cauchy 列 (基本列), 若对任意的 $\epsilon > 0$, 存在正整数 N , 使得对任意的 $m, n \geq N$, 有

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

若任意 Cauchy 列在 X 中都有极限, 则称 X 为完备的 (Completed).

例子 4. 常见的完备空间有:

- (\mathbb{R}, d) , 其中 $d(x, y) = |x - y|$;
- n 维欧氏空间 (\mathbb{R}^n, d) , 其中 $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$;
- l^p 空间, $1 \leq p < +\infty$, 其中 $d(x, y) = (\sum_{i=1}^{\infty} |x_i - y_i|^p)^{\frac{1}{p}}$, $x = (x_1, x_2, \dots)$, $y = (y_1, y_2, \dots)$;
- $[a, b]$ 上连续函数的全体 $C[a, b]$, 其中 $d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$.

命题 5. 完备度量空间的闭子空间是完备的.

证明. 设 (X, d) 为完备度量空间, $Y \subseteq X$ 为闭子空间. 设 $\{y_n\} \subseteq Y$ 是 Y 中的 Cauchy 列, 则 $\{y_n\}$ 也是 X 中的 Cauchy 列. 由 X 完备, $\exists x \in X$, 使得 $\lim_{n \rightarrow \infty} y_n = x$. 由于 Y 是闭集, 则 $x \in Y$. 因此 $\{y_n\}$ 在 Y 中收敛. 故而 Y 是完备的. \square

注记 6. 在 \mathbb{R} 中 Cauchy 列与收敛列是等价的. 但在一般的度量空间中, 容易知道收敛列是 Cauchy 列 ($d(x_m, x_n) \leq d(x_m, x) + d(x_n, x)$), 但 Cauchy 列不一定是收敛列. 你能找出一个例子吗?

定义 7 (压缩映射 (Contraction map)). 设 (X, d) 为度量空间, $T: X \rightarrow X$ 是 X 上的映射, 若存在常数 $0 \leq \alpha < 1$, 使得对任意的 $x, y \in X$, 有

$$d(Tx, Ty) \leq \alpha d(x, y),$$

则称 T 为 X 上的压缩映射. 若存在 x^* , 使得 $Tx^* = x^*$, 则称 x^* 为 T 的不动点.

练习 8. 压缩映射是连续映射.

(只需注意到 $0 \leq d(Tx, Ty) \leq \alpha d(x, y)$, 由此可知在 x, y 足够接近时 $d(Tx, Ty) \rightarrow 0$)

定理 9 (Banach 不动点定理). 设 (X, d) 为完备度量空间, $T: X \rightarrow X$ 为压缩映射, 则 T 存在唯一的不动点 $x^* \in X$.

证明. 作迭代序列 x_n , s.t. $x_n = Tx_{n-1}$, $n \geq 1$. 即任取起始点 $x_0 \in X$, 我们有 $x_1 = Tx_0, \dots, x_n = Tx_{n-1}$, ...

因为 T 是压缩映射, 则 T 连续. 因此我们只需要说明

$$x^* = \lim_{n \rightarrow \infty} x_n,$$

就有

$$x^* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} T x_{n-1} = T(\lim_{n \rightarrow \infty} x_{n-1}) = T x^*.$$

下证 $\lim_{n \rightarrow \infty} x_n$ 存在. 注意到 (X, d) 完备, 我们只需说明 $\{x_n\}$ 是 Cauchy 列.

$$\begin{aligned} d(x_1, x_2) &= d(Tx_0, Tx_1) \leq \alpha d(x_0, x_1), \\ d(x_2, x_3) &= d(Tx_1, Tx_2) \leq \alpha d(x_1, x_2) \leq \alpha^2 d(x_0, x_1), \\ &\dots\dots\dots \\ d(x_{n-1}, x_n) &= d(Tx_{n-2}, Tx_{n-1}) \leq \alpha d(x_{n-2}, x_{n-1}) \leq \dots \leq \alpha^{n-1} d(x_0, x_1). \end{aligned}$$

则

$$\begin{aligned} d(x_n, x_{n+p}) &\leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \dots + d(x_{n+1}, x_n) \\ &\leq \alpha^{n+p-1} d(x_0, x_1) + \dots + \alpha^n d(x_0, x_1) \\ &= \alpha^n \frac{1 - \alpha^p}{1 - \alpha} d(x_0, x_1) \\ &\leq \frac{\alpha^n}{1 - \alpha} d(x_0, x_1) \rightarrow 0, \text{ as } n \rightarrow \infty. \end{aligned}$$

故而 $\{x_n\}$ 是 Cauchy 列. 注意到 X 完备, $\exists x^* \in X$, 使得 $\lim_{n \rightarrow \infty} x_n = x^*$.

下证唯一性. 假设 $y^* \in X$ 也是 T 的不动点, 则

$$d(x^*, y^*) = d(Tx^*, Ty^*) \leq \alpha d(x^*, y^*).$$

注意到 $\alpha < 1$, 我们有 $d(x^*, y^*) = 0$, 即 $x^* = y^*$. □

定理 10 (Banach 不动点定理 II). 我们要求:

- (X, d) 是完备度量空间.
- $T: X \rightarrow X$.
- $T^n: X \rightarrow X$ 压缩映射, n 为正整数.

则 T 在 X 上存在唯一不动点.

证明. 由 Banach 不动点定理, T^n 存在唯一不动点 $x^* \in X$, 即 $T^n x^* = x^*$. 下证 x^* 也是 T 的不动点. 注意到

$$Tx^* = T(T^n x^*) = T^{n+1} x^* = T^n(Tx^*),$$

所以 Tx^* 也是 T^n 的不动点. 由 T^n 不动点的唯一性, 我们有 $Tx^* = x^*$.

(注意: T 本身是否是压缩映射则是不清楚的.) □

下面的两个练习我们先关心一元函数的情况, 之后再推广到多元函数的情况.

练习 11 (热身运动: 一元函数的隐函数存在定理). 设 $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ 上的连续函数, 且处处有关于 F_y 的一阶连续偏导数. $\exists(x_0, y_0) \in \mathbb{R}^2$, 使得 $F_y(x_0, y_0) \neq 0$, $F(x_0, y_0) = 0$. 那么 $\exists \delta > 0$, 使得方程 $F(x, y)$ 在区间 $(x_0 - \delta, x_0 + \delta)$ 上有唯一的连续函数 $y = f(x)$ 作为解

$$\begin{aligned} F(x, f(x)) &= 0, \quad \forall x \in (x_0 - \delta, x_0 + \delta), \\ f(x_0) &= y_0. \end{aligned}$$

解答 12. 由于 F_y 连续且在点 (x_0, y_0) 处不等于零, 因此存在 $\delta' > 0$ 使得它在带状区域

$$D = \{(x, y) : |x - x_0| \leq \delta', |y - y_0| \leq \delta'\}$$

上恒不等于零. 不妨假设 $F_y(x, y) > 0, \forall (x, y) \in D$. 注意到 F 在 D 上连续, 则 F 在紧集 D 上有界, 因此存在 $m, M > 0$, 使得

$$0 < m \leq F(x, y) \leq M, \quad \forall (x, y) \in D.$$

下面我们考虑连续函数空间 $C(I)$, 它关于度量 $d(f, g) = \max_{t \in I} |f(t) - g(t)|$ 是完备的. 定义映射

$$(Tf)(x) = f(x) - \frac{1}{M} F(x, f(x)).$$

由于 $F(x, f(x))$ 也是连续的, 因此 Tf 也是连续的, 故而 $T: C(I) \rightarrow C(I)$, 即自身到自身的映射.

下面说明 T 是压缩映射. $\forall f_1, f_2 \in C(I)$, 由 Lagrange 中值定理, $\exists \xi \in (0, 1)$, 使得

$$\begin{aligned} & |Tf_2(x) - Tf_1(x)| \\ &= |f_2(x) - \frac{1}{M} F(x, f_2(x)) - f_1(x) + \frac{1}{M} F(x, f_1(x))| \\ &= |f_2(x) - f_1(x) - \frac{1}{M} f'_y(x, f_1(x) + \theta(f_2(x) - f_1(x)))(f_2(x) - f_1(x))| \\ &\leq (1 - \frac{m}{M}) |f_2(x) - f_1(x)|. \end{aligned}$$

则 $d(Tf_1, Tf_2) \leq (1 - \frac{m}{M}) d(f_1, f_2)$. 注意到 $0 < 1 - \frac{m}{M} < 1$, 于是 T 是压缩映射. 由 Banach 不动点定理, T 存在唯一不动点 $f \in C(I)$, 使得

$$f(x) \equiv Tf(x) = f(x) - \frac{1}{M} F(x, f(x)).$$

因此 $F(x, f(x)) = 0$. 由 $F(x_0, y_0) = 0$, 可知 $f(x_0) = y_0$.

利用上述的一元隐函数定理, 很快就能得到一元函数的反函数定理.

练习 13 (一元函数的反函数定理). 设 $f: \mathbb{R} \rightarrow \mathbb{R}$ 是一个连续可微函数并且对于 $x_0 \in \mathbb{R}$ 有 $f'(x_0) \neq 0$. 那么, 存在 x_0 的某个邻域 $U = [x_0 - \epsilon, x_0 + \epsilon]$, 以及 $f(x_0)$ 的某个邻域 V 使得

- $f: U \rightarrow V$ 是双射;
- $f^{-1}: V \rightarrow U$ 也是连续可微函数;
- $\forall y \in V, (f^{-1})'(y) = \frac{1}{f'(x)}$, 这里 $x = f^{-1}(y)$.

解答 14. 定义 $F(x, y) = y - f(x)$, 则 $F_y(x, y) = 1 \neq 0$. 由隐函数定理, $\exists \delta > 0$, 使得方程 $F(x, y) = 0$ 在区间 $[y_0 - \delta, y_0 + \delta]$ 上有唯一的连续函数 $x = g(y)$ 作为解.

取 $\epsilon > 0$ 足够小, 使得 $f([x_0 - \epsilon, x_0 + \epsilon]) \subseteq [y_0 - \delta, y_0 + \delta]$, 可以验证 $g|_V$ 就是所需要的 f^{-1} .
更多的细节留给大家自己验证.

练习 15 (Picard-Lindelöf 定理). 之后我们会介绍如何通过不动点定理证明常微分方程的解的存在唯一性, 在这里暂时先略过了.

下面我们试图得到高维的隐函数定理以及反函数定理, 为此我们还需要更多的努力.

如果我们给度量空间 (X, d) 赋予拓扑结构 τ , 则 (X, τ, d) 称为拓扑度量空间. 那么如果称 $f: X \rightarrow X$ 是连续的, 当且仅当对任意的开集 $U \in \tau$, $f^{-1}(U)$ 也是开的.

在下面证明高维的反函数定理的过程中, 我们考虑的空间是 (\mathbb{R}^n, d) , 这里的 d 是由一个合适的范数诱导的度量, 比如 2-范数诱导的度量就是一个合适的选择:

$$d(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

对于某一个线性算子 $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (这里的符号表示 L 是从 \mathbb{R}^n 到 \mathbb{R}^n 的线性映射, 在我们设定的欧式空间下 L 是 $\mathbb{R}^{n \times n}$ 的矩阵的全体).

- 我们定义算子 T 在某点 x_0 的 **Fréchet 导数**使得

$$\lim_{h \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - Lh\|}{\|h\|} = 0.$$

这个极限本质上还是个正常定义的极限, 我们当然可以写成如下的一阶展开:

$$f(x_0 + h) = f(x_0) + Lh + o(h).$$

记该导数为 $L := Df(x_0) \in \mathbb{R}^{n \times n}$.

- 如果 Df 这一矩阵中的各个元素都是 \mathbb{R} 上的连续函数, 我们就称 $f \in C^1(\mathbb{R}^n)$. 记 $Jf(x_0) := \det Df(x_0)$ 为 f 在 x_0 处的 **Jacobian 行列式**. 我们称 f 在 x_0 处是非退化的 (non-degenerate) 如果 $Jf(x_0) \neq 0$.

下面我们来证明多元函数的逆映射定理.

练习 16 (逆映射定理). 设 $f: \Omega \rightarrow \mathbb{R}^n$, 这里 $\Omega \subseteq \mathbb{R}^n$ 是开集, 且 $f \in C^1$. 在 $x_0 \in \Omega$ 处 $Jf(x_0) \neq 0$. 证明: 存在 x_0 的某个邻域 $U \subseteq \Omega$, 以及 $f(x_0)$ 的某个邻域 $V \subseteq \mathbb{R}^n$, 使得

- $f: U \rightarrow V$ 是双射 (既单且满);
- $f^{-1}: V \rightarrow U$ 也是 C^1 函数;
- $\forall y \in V, Df^{-1}(y) = [Df(x)]^{-1}$ and $Jf^{-1}(y) = \frac{1}{Jf(x)}$.

[课后练习: 试着从这个较为抽象的定理“翻译”成教材 p. 164 定理 12.4.5 这一较简单的情形.]

解答 17. 不失一般性, 我们假设 $x_0 = 0, f(x_0) = 0$, 并且 $Df(x_0) = I$, 这里 I 是 n 维单位矩阵. 事实上, 你总是可以经过适当的平移和拉伸做到这件事情, 比如定义 $h(x) = [Df(x_0)]^{-1}(f(x) - f(x_0))$.

试图转化为不动点问题. 先说明是满射, 即 $\forall y \in V, \exists x \in U, \text{ s.t. } f(x) = y$. (注意: 这里的 U, V 还没有确定) 上面的等式等价于 $x - f(x) + y = x$, 因此我们考虑构造如下的一个函数

$$T(x) = x - f(x) + y.$$

那么 T 的不动点便是我们要寻找的 x .

说明 T 是压缩映射. 注意到 $T(0) = y, DT(0) = I - Df(0) = 0$. 由于 $DT(x)$ 在 0 处连续 (由于 $f \in C^1$), 则 $\exists \delta > 0, \text{ s.t. } \forall x \in B(0, \delta),$

$$\|DT(x)\| < \frac{1}{25}.$$

我们有如下的估计:

$$\begin{aligned} \|T(x)\| &\leq \|T(x) - T(0)\| + \|T(0)\| \leq \max_{\xi \in B(0, \delta)} \|DT(\xi)\| \cdot \|x\| + \|y\| \\ &< \frac{1}{25} \|x\| + \|y\| < \delta. \end{aligned}$$

这里我们取 $y \in B(0, \frac{1}{2}\delta)$ (于是选取了 V). 同时表明了 $T(x) \in B(0, \delta)$. 所以对 $X = \overline{B(0, \delta)}$ (作为 \mathbb{R}^n 的闭子空间是完备的), $T: X \rightarrow X$ 自身的映射. 另一方面, $\forall x, x' \in X$, 由 Lagrange 中值定理,

$$\begin{aligned} \|T(x) - T(x')\| &\leq \max_{\xi \in B(0, \delta)} \|DT(\xi)\| \cdot \|x - x'\| \\ &< \frac{1}{25} \|x - x'\|. \end{aligned}$$

于是 T 是压缩映射. 由 Banach 不动点原理, 对任意的 $y \in B(0, \frac{1}{2}\delta)$, 都存在唯一的 $x \in B(0, \delta)$, 使得 $T(x) = x$, 即 $f(x) = y$. 综上所述, 我们选取如下的两个开集

$$U := B(0, \delta) \cap T^{-1}(B(0, \frac{1}{2}\delta)), \quad V := B(0, \frac{1}{2}\delta).$$

由前面的证明过程, 说明了 $f: U \rightarrow V$ 上是**满射**. $f: U \rightarrow V$ 上也是**单射**. 事实上, 如果 $f(x_1) = f(x_2) = y$, 那么等价于 $T(x_1) = x_1 - f(x_1) + y = x_1$ 且 $T(x_2) = x_2 - f(x_2) + y = x_2$. 由 T 的不动点的唯一性, 我们有 $x_1 = x_2$.

考察 $g := f^{-1}$. g 是连续函数. 事实上, 任意取 $y, y' \in V$,

$$\begin{aligned} \|g(y) - g(y')\| &= \|x - x'\| = \|T(x) - T(x') + y - y'\| \\ &\leq \|T(x) - T(x')\| + \|y - y'\| \\ &< \frac{1}{25}\|x - x'\| + \|y - y'\| \\ &= \frac{1}{25}\|g(y) - g(y')\| + \|y - y'\|, \end{aligned}$$

则

$$\|g(y) - g(y')\| < \frac{25}{24}\|y - y'\|.$$

这表明 g 是连续的.

我们最后说明 g 是可微的. 令 $y \in V$, 取足够小的 $k \in \mathbb{R}^n$ 使得 $y + k \in V$. 设 $x = g(y)$, $x' = g(y + k)$, 以及 $h := g(y + k) - g(y) = x' - x$. 注意到 f 是可微的, 则

$$f(x + h) - f(x) = Df(x)h + o(\|h\|).$$

$$\implies k = Df(x)h + o(\|h\|).$$

注意到 $Df(x)$ 可逆 (因为 $Jf(x) \neq 0$), 则

$$h = (Df(x))^{-1}k + (Df(x))^{-1}o(\|h\|),$$

i.e.,

$$g(y + k) - g(y) = (Df(x))^{-1}k + (Df(x))^{-1}o(\|h\|).$$

由于 g 是 Lipschitz 连续的, 因此 $o(\|h\|) = o(\|k\|)$, 所以

$$g(y + k) - g(y) = (Df(x))^{-1}k + o(\|k\|).$$

这意味着 g 在 y 处可微, 且 $Dg(y) = (Df(x))^{-1}$, 即 $Df^{-1}(y) = [Df(x)]^{-1}$.

例子 18. 考虑投影映射

$$F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m : (x_1, \dots, x_n, y_1, \dots, y_m) \mapsto (y_1, \dots, y_m).$$

取 $U = \mathbb{R}^n$, $V = \mathbb{R}^m$, 则

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto 0$$

满足隐函数定理的要求. 隐函数定理本质上就是谈论选定好适当坐标系后, 这就是一个唯一的例子. [关于隐函数定理更多几何上的直观 (从子流形的角度), 可以参考清华大学于品老师的《讲义: 数学分析》].