Exercise Sheet – Advanced Calculus III

Taiyang Xu*

25/09/2025, Week 3

Exercise 1 (本练习涉及凸区域、凸函数、中值定理和多元 Taylor 公式等内容)**.** 完成以下题目

(1) 设 $D \subseteq \mathbb{R}^n$ 为凸的有界闭区域, f 在 D 上有连续的一阶偏导数。试证明 f 在 D 上满足如下的 Lipschitz 条件: 存在常数 $M > 0$, 使得对任意的 $x, y \in D$, 有

$$|f(x) - f(y)| \leq M|x - y|.$$

(2) 设 $F(x, y, z)$ 在 $\mathbb{R}^3$ 中有连续的一阶偏导数 $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, $\frac{\partial F}{\partial z}$, 并满足不等式:

$$x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z} \geq \alpha > 0, \quad \forall (x, y, z) \in \mathbb{R}^3.$$

证明: 当 (x, y, z) 沿着曲线 $\Gamma = (x(t), y(t), z(t))$, $x(t) = -\cos t$, $y(t) = \sin t$, $z(t) = t$, $t \geqslant 0$ 趋向 ∞ 时, $F(x, y, z)$ 趋向 $+\infty$.

(3) 设 f 在凸区域 $D \in \mathbb{R}^n$ 上具有连续的一阶偏导数, 则 f 在 D 上为凸函数的充分必要条件是:

$$\forall x, y \in D, \quad f(y) - f(x) \geqslant \nabla f(x) \cdot (y - x).$$

(4) 设 $f(\vec{x})$, $\vec{x} = (x_1, x_2, \ldots, x_n)$ 在凸区域 D 上具有连续的二阶偏导数, 则 f 在 D 上为凸函数的充分必要条件是 Hesse 矩阵 $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^n$ 在 D 上的任意点均为半正定矩阵.

Solution 2. (1) 由于 D 是凸的有界闭区域，且 f 在 D 上有连续的一阶偏导数，根据微分中值定理，对任意 $x, y \in D$，存在 $\theta \in [0, 1]$，使得

$$f(y) - f(x) = \langle \nabla f(x + \theta(y - x)), (y - x) \rangle.$$

由于 D 有界闭，∇f 在 D 上连续, 故根据 Bolzano-Weierstrass 定理, ∇f 在 D 上一致有界。设

$$M := \max_{z \in D} |\nabla f(z)|,$$

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

则对任意 $x, y \in D$，由微分中值定理，有

$$f(y) - f(x) = \langle \nabla f(x + \theta(y - x)), (y - x) \rangle,$$

其中 $\theta \in [0, 1]$。因此

$$|f(y) - f(x)| = |\langle \nabla f(x + \theta(y - x)), y - x \rangle| \leq |\nabla f(x + \theta(y - x))| \cdot |y - x| \leq M|y - x|$$

故 f 满足 Lipschitz 条件。

(2) 定义

$$\Phi(t) := F(-\cos t, \sin t, t).$$

利用微分中值定理, 我们有

$$\Phi(t) = \Phi(0) + \Phi'(\tau)t, \quad \tau \in (0, t).$$

直接计算

$$\begin{aligned}
&\Phi(0) = F(-1, 0, 0), \\
&\Phi'(\tau) = \frac{\partial F}{\partial x}\frac{dx}{dt}\Big|_{t=\tau} + \frac{\partial F}{\partial y}\frac{dy}{dt}\Big|_{t=\tau} + \frac{\partial F}{\partial z}\frac{dz}{dt}\Big|_{t=\tau} = \sin t\frac{\partial F}{\partial x}\Big|_{t=\tau} + \cos t\frac{\partial F}{\partial y}\Big|_{t=\tau} + 1 \cdot \frac{\partial F}{\partial z}\Big|_{t=\tau}, \\
&= x\frac{\partial F}{\partial x}\Big|_{t=\tau} - y\frac{\partial F}{\partial y}\Big|_{t=\tau} + \frac{\partial F}{\partial z}\Big|_{t=\tau} \geq \alpha > 0.
\end{aligned}$$

故而 $\Phi(t) = F(x, y, z) \geqslant F(1, 0, 0) + \alpha t \to +\infty$, $t \to +\infty$.

(3) $\Rightarrow$: 假设 f 是凸函数。对任意 $x, y \in D$，对任意 $\lambda \in [0, 1]$，由凸性有

$$f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda)f(x).$$

即

$$f(x + \lambda(y - x)) - f(x) \leqslant \lambda(f(y) - f(x)).$$

由于 f 在 D 上具有连续的一阶偏导数, 根据多元函数的中值定理,

$$f(x + \lambda(y - x)) - f(x) = \nabla f(x + \theta\lambda(y - x)) \cdot (\lambda(y - x)), \quad \theta \in (0, 1).$$

因此, 对 $\lambda \neq 0$, 有

$$\nabla f(x + \theta\lambda(y - x)) \cdot (y - x) \leqslant f(y) - f(x).$$

令 $\lambda \to 0^+$, 则 $x + \theta\lambda(y - x) \to x$, 由 ∇f 的连续性, 有最终结果。
也可以利用 Taylor 公式 (而非中值定理) 来证明. 对 $f(x + \lambda(y - x))$ 在 x 做 Taylor 展开:

$$f(x + \lambda(y - x)) = f(x) + \nabla f(x) \cdot (\lambda(y - x)) + \mathcal{O}(||\lambda(y - x)||^2), \quad \theta \in (0, 1).$$

$\Leftarrow$: 假设对任意 $x,y\in D$，都有 $f(y)-f(x)\geqslant \nabla f(x)\cdot(y-x)$。我们需要证明 f 是凸函数。取任意 $x,y\in D$，任意 $\lambda\in[0,1]$，令 $z=\lambda y+(1-\lambda)x\in D$。由假设，对 x,z 有

$$f(x)-f(z)\geqslant \nabla f(z)\cdot(x-z).$$

同理，对 y,z 有

$$f(y)-f(z)\geqslant \nabla f(z)\cdot(y-z).$$

将两式分别按乘上 λ 和 $1-\lambda$ 再相加得到

$$\begin{aligned}\lambda f(x)+(1-\lambda)f(y)&\geqslant \lambda f(z)+\lambda\nabla f(z)\cdot(x-z)+(1-\lambda)f(z)+(1-\lambda)\nabla f(z)\cdot(y-z)\\&=f(z)+\nabla f(z)\cdot(\lambda(x-z)+(1-\lambda)(y-z))\\&=f(z)+\nabla f(z)\cdot(\lambda x+(1-\lambda)y-z)\\&=f(z).\end{aligned}$$

因此 f 是凸函数。

(4) $\Leftarrow$: $\forall x,y\in D$, $\exists z=x+\theta(y-x)$, 使得 (Lagriange 余项型的 Taylor 展开)

$$f(y)=f(x)+\nabla f(x)\cdot(y-x)+\frac{1}{2}\left(\sum_{k=1}^{n}(y_k-x_k)\frac{\partial}{\partial x_k}\right)^2 f(z).$$

事实上, 你会发现

$$\begin{aligned}&\left(\sum_{k=1}^{n}(y_k-x_k)\frac{\partial}{\partial x_k}\right)^2 f(z)\\&=\sum_{i,j=1}^{n}(y_i-x_i)(y_j-x_j)\frac{\partial^2 f}{\partial x_i\partial x_j}(z)\\&=(y-x)^{\mathrm{T}}H_f(z)(y-x).\end{aligned}$$

由于 $H_f(z)$ 为半正定矩阵, 则

$$(y-x)^{\mathrm{T}}H_f(z)(y-x)\geqslant 0.$$

因此 $f(y)-f(x)\geqslant \nabla f(x)\cdot(y-x)$. 由 (3) 可知 f 是凸函数。

$\Rightarrow$: 如若不然. 假设 H_f 在某点 $x\in D$ 不是半正定矩阵, 则 $\exists v\in\mathbb{R}^n$, 使得

$$v^{\mathrm{T}}H_f(x)v<0,$$

这里 $v=(v_1,\ldots,v_n)^{\mathrm{T}}\neq 0$. 另一方面, 由 Taylor 公式，对 $f(x+tv)$ 在 x 做 Taylor 展开 $(t\to 0)$, 则

$$\begin{aligned}f(x+tv)&=f(x)+\nabla f(x)\cdot(tv)+\frac{1}{2}(tv)^{\mathrm{T}}H_f(x)(tv)+o(||tv||^2)\\&=f(x)+t\nabla f(x)\cdot v+\frac{t^2}{2}\left[v^{\mathrm{T}}H_f(x)v+o(1)\right].\end{aligned}$$

当 t 充分小时, $v^{\mathrm{T}} H_f(x) v + o(1) < 0$, 故

$$f(x+tv) \leqslant f(x) + \nabla f(x) \cdot tv.$$

这与 (3) 中的充分条件矛盾.

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.

Definition 3 (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) $d(x,y) \geqslant 0$, 且 $d(x,y) = 0 \Leftrightarrow x = y$ (正定性);

(2) $d(x,y) = d(y,x)$ (对称性);

(3) $d(x,z) \leqslant d(x,y) + d(y,z)$ (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.

Example 4. 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, $1 \leqslant p < +\infty$, 其中 $d(x,y) = (\sum_{i=1}^{\infty} |x_i - y_i|^p)^{\frac{1}{p}}$, $x = (x_1, x_2, \ldots)$, $y = (y_1, y_2, \ldots)$;
- $[a,b]$ 上连续函数的全体 $C[a,b]$, 其中 $d(f,g) = \max_{x \in [a,b]} |f(x) - g(x)|$.

Definition 5 (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 $m, n \geqslant N$, 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).

Example 6. 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, $1 \leqslant p < +\infty$, 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 $d(f,g) = \max_{x \in [a,b]} |f(x) - g(x)|$.

Proposition 7. *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 $\lim_{n\to\infty} y_n = x$. 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. □

Remark 8. 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 $(d(x_m, x_n) \leqslant d(x_m, x) + d(x_n, x))$, 但 Cauchy 列不一定是收敛列. 你能找出一个例子吗?

Definition 9 (压缩映射 (Contraction map))**.** 设 (X,d) 为度量空间, $T: X \to X$ 是 X 上的映射, 若存在常数 $0 \leqslant \alpha < 1$, 使得对任意的 $x, y \in X$, 有

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

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.

Exercise 10. 压缩映射是连续映射.
(只需注意到 $0 \leqslant d(Tx,Ty) \leqslant \alpha d(x,y)$, 由此可知在 x, y 足够接近时 $d(Tx, Ty) \to 0$)

Theorem 11 (Banach 不动点定理)**.** *设 (X,d) 为完备度量空间, $T: X \to X$ 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. $x_n = Tx_{n-1}$, $n \geqslant 1$. 即任取起始点 $x_0 \in X$, 我们有 $x_1 = Tx_0$,..., $x_n = Tx_{n-1}$, ...

因为 T 是压缩映射, 则 T 连续. 因此我们只需要说明

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

就有

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

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

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

则

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

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

下证唯一性. 假设 $y^*\in X$ 也是 T 的不动点, 则

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

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

Theorem 12 (Banach 不动点定理 II)**.** *我们要求:*

- *(X,d) 是完备度量空间.*
- *$T: X\to X$.*
- *$T^n: X\to X$ 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.

Exercise 13 (热身运动: 一元函数的隐函数存在定理). 设 $F:\mathbb{R}^2\to\mathbb{R}$ 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

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

Solution 14. 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

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

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

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

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 $T:C(I)\to C(I)$, 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

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

则 $d(Tf_1,Tf_2)\leqslant(1-\frac{m}{M})d(f_1,f_2)$. 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.

Exercise 15 (一元函数的反函数定理). 设 $f:\mathbb{R}\to\mathbb{R}$ 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

- $f:U\to V$ 是双射;

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

Solution 16. 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.

Exercise 17 (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 $f : X \to X$ 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

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

  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.

Exercise 18 (逆映射定理)**.** 设 $f : \Omega \to \mathbb{R}^n$, 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

- $f : U \to V$ 是双射 (既单且满);
- $f^{-1} : V \to U$ 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.]

Solution 19. 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 $g(x) = [Df(x_0)]^{-1}(f(x_0 + x) - f(x_0))$.

试图转化为不动点问题. 先说明是满射, 即 $\forall y \in V, \exists x \in U$, 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 $\exists \delta > 0$, s.t. $\forall x \in B(0, \delta)$,

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

我们有如下的估计:

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

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), $T : X \to X$ 自身的映射. 另一方面, $\forall x, x' \in X$, 由 Lagrange 中值定理,

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

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 $f : U \to V$ 上是**满射**. $f : U \to V$ 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 $y, y' \in V$,

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

则

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

这表明 g 是连续的.

我们考虑如下的估计:

$$||f(x)|| \geqslant ||Df(0)x|| - ||x|| \cdot ||\phi(x)|| \geqslant c||x||.$$

因此

$$\frac{||Df(0)x - f(x)||}{||f(x)||} = \frac{||x|| \cdot ||\phi(x)||}{f(x)} \leqslant \frac{||\phi(x)||}{c} \to 0, \quad x \to 0.$$

利用 g 的连续性: 当 $y \to 0$, $x = g(y) \to 0$ 时,

$$\begin{aligned}\frac{||Df(0)x - f(x)||}{||f(x)||} &= \frac{||Df(0)g(y) - y||}{||y||} = \frac{||(Df(0))\left[g(y) - (Df(0))^{-1}y\right]||}{||y||} \\ &\leqslant M \cdot \frac{||g(y) - (Df(0))^{-1}y||}{||y||} \to 0, \quad y \to 0.\end{aligned}$$

注意到: 如果 $g(y)$ 在 $y = 0$ 处亦可微, 那么

$$\frac{||g(y) - Dg(0)y||}{||y||} \to 0, \quad y \to 0.$$

则选取 $Dg(0) = Df(0)^{-1}$ 即可. 对非 0 点的证明类似, 这里不再赘述, 事实上你可以通过平移和拉伸将问题转化为 0 点的情况.

Example 20. 考虑投影映射

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

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

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

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》].