[Lecture Note 1]

Contraction Mapping Principles and Implicit Function Theorem

Definition 1. *A normed vector space X is a Banach space if it is complete, i.e., every Cauchy sequence converges.*

Let X, Y be Banach spaces with norms $|\cdot|$. Let $L(X, Y)$ denote the set of all bounded linear operators T from X to Y with the induced operator norm

$$|T| = \sup_{|x|\leq 1} |Tx|,$$

where $|x|$ is the norm of x in X and $|Tx|$ is the norm of $y = Tx$ in Y. Then it can be proved that $L(X, Y)$ is a Banach space.

Lemma 1. *Let X be a Banach space with norm $|\cdot|$. Let $T \in L(X, X)$. If $|T| \leq \theta < 1$, then the linear operator $I - T$ is invertible, and the inverse is*

$$[I - T]^{-1} = I + T + T^2 + \cdots = \sum_{n=0}^{\infty} T^n$$

with bound

$$|[I - T]^{-1}| \leq \frac{1}{1 - \theta}.$$

Proof. It is left as an exercise. □

Definition 2. *Let X, Y be Banach spaces. A function $f : X \to Y$ is said to be differentiable at a point $x \in X$ if there is a bounded linear map $T : X \to Y$ so that for $\Delta(x, h) = f(x + h) - f(x) - Th$,*

$$|\Delta(x, h)| = o(|h|), \text{ as } h \to 0,$$

where $o(\epsilon)$ denotes any higher order term satisfying $o(\epsilon)/\epsilon \to 0$ as $\epsilon \to 0$. In such a case, T is called the derivative of f at x and is denoted by $T = Df(x)$. Also, $f \in C^1$ if f is differentiable at every point of X and the derivative $Df(x)$ is continuous in x.

Lemma 2. *Let X, Y be Banach spaces and $V \in Y$ be an open set. Let $T : V \to L(X, X)$. Assume $T(\cdot)$ is in $C^k(V, L(X, X))$, or $C^{k,1}(V, L(X, X))$, $k \geq 0$, and is uniformly contractive, $\sup_{y\in V} |T(y)| \leq \theta < 1$. Then the inverse $[I - T(y)]^{-1}$ is in $C^k(V, L(X, X))$, or $C^{k,1}(V, L(X, X))$.*

Proof. It is left as an exercise. (Hint: Let $f, g \in C^1(V, L(X, X))$. Prove first the product rule: $[D(f(y)g(y))]h = [Df(y)h]g(y) + f(y)[Dg(y)h]$ for $y \in V$ and $h \in Y$. Then apply the product rule to $T(y)^n$ to obtain the power-rule.) □

Lemma 3. *Let X, Y be Banach spaces and $f : X \to Y$ be differentiable at a point x. Then there is a bound $0 < K(x,h) < \infty$ so that for sufficiently small $|h|$ with $h \in X$*

$$|f(x+h) - f(x)| \leq K(x,h)|h|$$

and $K(x,h) \to |Df(x)|$ as $h \to 0$.

Proof. By assumption,

$$\begin{aligned} |f(x+h) - f(x)| &= |f(x+h) - f(x) - Df(x)h + Df(x)h| \\ &\leq |f(x+h) - f(x) - Df(x)h| + |Df(x)h| \\ &\leq (|Df(x)| + o(|h|))|h| \end{aligned}$$

This proves the result with $K(x,h) = |Df(x)| + o(|h|)$. □

Theorem 1 (Contraction Mapping Theorem)**.** *Let $\{X, d\}$ be a complete metric space. Assume $f : X \to X$ is a contraction mapping in the sense that there is a constant $0 < \theta < 1$ so that for every $x, y \in X$,*

$$d(f(x), f(y)) \leq \theta d(x,y).$$

Then f has a unique fixed point $\bar{x} \in X$, $f(\bar{x}) = \bar{x}$, and for any $x \in X$ and $n \geq 0$,

$$d(f^n(x), \bar{x}) \leq \frac{\theta^n}{1-\theta} d(x, f(x)).$$

Proof. Notice first that f is Lipschitz continuous by the contraction mapping assumption. Now by recursion, for any $x \in X$ and integers $n, k \geq 0$,

$$\begin{aligned} d(f^n(x), f^{n+1}(x)) &\leq \theta d(f^{n-1}(x), f^n(x)) \\ &\leq \theta^n d(x, f(x)) \\ d(f^n(x), f^{n+k}(x)) &\leq d(f^n(x), f^{n+1}(x)) + d(f^{n+1}(x), f^{n+2}(x)) + \cdots \\ &\qquad + d(f^{n+k-1}(x), f^{n+k}(x)) \\ &\leq (\theta^n + \theta^{n+1} + \cdots + \theta^{n+k-1}) d(x, f(x)) \\ &= \tfrac{\theta^n(1-\theta^k)}{1-\theta} d(x, f(x)) \\ &\leq \tfrac{\theta^n}{1-\theta} d(x, f(x)) \to 0 \text{ as } n \to \infty. \end{aligned}$$

Hence, $\{f^n(x)\}$ is a Cauchy sequence, and by the completeness of X, the limit $\lim_{n\to\infty} f^n(x) = \bar{x}$ exists for some $\bar{x} \in X$. We conclude first that $\bar{x}$ is a fixed point because by the continuity of f we have

$$f(\bar{x}) = f(\lim_{n\to\infty} f^n(x)) = \lim_{n\to\infty} f^{n+1}(x) = \bar{x}.$$

Also, the fixed point is unique because if x^* is also a fixed point, then

$$d(x^*, \bar{x}) = d(f(x^*), f(\bar{x})) \leq \theta d(x^*, \bar{x})$$

forcing $d(x^*, \bar{x}) = 0$ because $\theta < 1$, and $x^* = \bar{x}$ for the uniqueness of fixed point. Last the estimate follows from taking the limit $k \to \infty$ in the inequality above. □

Theorem 2 (Uniform Contraction Principle I)**.** *Let X, Y be two metric spaces with X being complete. Assume $f : X \times Y \to X$ is continuous and uniformly contractive with a contraction constant $0 < \theta < 1$. Then the unique fixed point $\bar{x}(y)$ is continuous and*

$$d(\bar{x}(z), \bar{x}(y)) \leq \frac{1}{1-\theta} d(f(\bar{x}(y), z), f(\bar{x}(y), y)).$$

Proof. Let $0 < \theta < 1$ be the uniform contraction constant. Then for any $z \in Y$

$$\begin{aligned} d(\bar{x}(z), \bar{x}(y)) &= d(f(\bar{x}(z), z), f(\bar{x}(y), y)) \\ &\leq d(f(\bar{x}(z), z), f(\bar{x}(y), z)) + d(f(\bar{x}(y), z), f(\bar{x}(y), y)) \\ &\leq \theta d(\bar{x}(z), \bar{x}(y)) + d(f(\bar{x}(y), z), f(\bar{x}(y), y)) \end{aligned}$$

implies

$$d(\bar{x}(z), \bar{x}(y)) \leq \frac{1}{1-\theta} d(f(\bar{x}(y), z), f(\bar{x}(y), y)) \tag{1}$$

which goes to 0 as $z \to y$. This shows $\bar{x}(\cdot)$ is continuous in y. □

Theorem 3 (Uniform Contraction Principle II)**.** *Let X, Y be two Banach spaces, and let $U \subset X,\ V \subset Y$ be open subsets. Let $f \in C^k(\bar{U} \times V, \bar{U}), 1 \leq k < \infty$. Assume $f : \bar{U} \times V \to \bar{U}$ is a uniform contraction mapping, and $|D_x f(x, y)|$ is uniformly bounded by a constant $\theta < 1$ in $\bar{U} \times V$. Let $\bar{x}(y)$ be the unique fixed point of $f(\cdot, y)$ in $\bar{U}$ for $y \in V$. Then $\bar{x}(\cdot) \in C^k(V, \bar{U})$ and the first derivative is*

$$D\bar{x}(\cdot) = \sum_{n=0}^{\infty} [D_x f(\bar{x}(\cdot), \cdot)]^n D_y f(\bar{x}(\cdot), \cdot). \tag{2}$$

If f is $C^{k,1}$, then $\bar{x}(\cdot)$ is $C^{k,1}$, and if f is analytic in $U \times V$, then $\bar{x}(\cdot)$ is analytic from V to X.

Proof. Without loss of generality, let $0 < \theta < 1$ be the uniform contraction constant as well. Formally, differentiating $\bar{x}(y) = f(\bar{x}(y), y)$, the linear operator $D\bar{x}(y)$ should be a solution of the following operator equation in T

$$[I - D_x f(\bar{x}(y), y)]T = D_y f(\bar{x}(y), y). \tag{3}$$

Since $|D_x f(\bar{x}(y), y)| \leq \theta < 1$, this equation has a unique solution $T(y)$ by Lemma 1. It is left to show $D\bar{x}(y) = T(y)$, namely

$$|\Delta| := |\bar{x}(y+h) - \bar{x}(y) - T(y)h| = o(|h|), \text{ as } h \to 0, \tag{4}$$

where $o(|h|)$ denotes an higher order term than h, i.e., $o(|h|)/|h| \to 0$ as $h \to 0$.

From (1) of the proof for Theorem 2 and Lemma 3 we have

$$|\bar{x}(y+h) - \bar{x}(y)| \leq \frac{1}{1-\theta} |D_y f(\bar{x}(y), y)h + o(|h|)| \leq K|h| \tag{5}$$

for some constant K and all $y,\ y+h$ in V. From (3) we have

$$\begin{aligned}
&|[I - D_x f(\bar{x}(y), y)]\Delta| = |[I - D_x f(\bar{x}(y), y)](\bar{x}(y+h) - \bar{x}(y) - T(y)h)| \\
&\quad = |\bar{x}(y+h) - \bar{x}(y) - D_x f(\bar{x}(y), y)(\bar{x}(y+h) - \bar{x}(y)) - D_y f(\bar{x}(y), y)h| \\
&\quad = |f(\bar{x}(y+h), y+h) - f(\bar{x}(y), y) \\
&\qquad\qquad\qquad - D_x f(\bar{x}(y), y)(\bar{x}(y+h) - \bar{x}(y)) - D_y f(\bar{x}(y), y)h| \\
&\quad = o(|\bar{x}(y+h) - \bar{x}(y)| + |h|)
\end{aligned}$$

because $f \in C^1(\bar{U} \times V, \bar{U})$. Because of (5), we have

$$|[I - D_x f(\bar{x}(y), y)]\Delta| = o(|h|).$$

Last by Lemma 1 we have

$$\begin{aligned}
|\Delta| &= |[I - D_x f(\bar{x}(y), y)]^{-1}[I - D_x f(\bar{x}(y), y)]\Delta| \\
&\qquad\qquad \leq \frac{1}{1-\theta}|[I - D_x f(\bar{x}(y), y)]\Delta| = o(|h|).
\end{aligned}$$

This proves $\bar{x}(\cdot)$ is differentiable in V and $D\bar{x}(y) = T(y)$. Using identity (3) and Lemma 1 we obtain identity (2). From (2) we can conclude that $D\bar{x}$ is continuous in V because $f \in C^1$ and $\bar{x}(\cdot)$ is continuous in V. This shows $D\bar{x} \in C^1$.

Suppose f is C^k for $k > 1$. From identity (2) and the same argument above we can derive recursively that $\bar{x}(\cdot)$ is C^2, C^3, etc., until that $\bar{x}(\cdot)$ is C^k.

If f is $C^{k,1}$, from identity (2) and the fact that $\bar{x}(\cdot)$ is C^k we can see easily that $\bar{x}(\cdot)$ is also $C^{k,1}$.

In the analytic case, there is a complex neighborhood of $(\bar{x}(y), y)$ in which f is differentiable and uniformly contracting. The argument above shows that $\bar{x}(y)$ is also differentiable in the corresponding complex neighborhood, and hence analyticity of $\bar{x}(y)$. □

In applications it is often the case that the uniform contraction of a mapping is proved by some bound of its derivative. The following is such a typical approach.

Lemma 4. *Let X, Y be two Banach spaces, and let $U \subset X$ be a convexed open set. If $f \in C^1(U, Y)$, then for any $x, y \in U$*

$$|f(y) - f(x)| \leq \sup_{z \in U} |Df(z)||y - x|.$$

Proof. Let $x, y \in U$. Since U is convexed, $x + th \in U$ for $t \in [0, 1]$ where $h = y - x$. Thus

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} f(x + th)dt = \int_0^1 Df(x + th)dt(y - x).$$

and

$$|f(y) - f(x)| \leq \int_0^1 |Df(x + th)|dt|y - x| \leq \sup_{z \in U} |Df(z)||y - x|$$

□

Theorem 4 (Implicit Function Theorem I)**.** *Let X, Y, Z be Banach spaces, $U \subset X$, $V \subset Y$ be open sets. Assume $F : U \times V \to Z$ is differentiable in $x \in U$ and both F and D_xF are continuous in $(x, y) \in U \times V$. If there is a point $(x_0, y_0) \in U \times V$ such that $F(x_0, y_0) = 0$ and $D_xF(x_0, y_0)$ has a bounded inverse, then there is a neighborhood $U_1 \times V_1 \subset U \times V$ of (x_0, y_0) and a continuous function $f : V_1 \to U_1$ with $f(y_0) = x_0$ such that $F(x, y) = 0$ for $(x, y) \in U_1 \times V_1$ iff $x = f(y)$.*

Proof. Let $T = [D_xF(x_0, y_0)]^{-1}$ and $G(x, y) = x - TF(x, y)$. Then x is a fixed point of G iff (x, y) is a solution of $F = 0$. The function G is as smooth as F is, and $G(x_0, y_0) = x_0$, $D_xG(x_0, y_0) = 0$. Therefore we can find a neighborhood $U_1 \times V_1 \subset U \times V$ of (x_0, y_0) with $U_1 = N_{\delta_1}(x_0)$ convexed, $V_1 = N_{\delta_2}(y_0)$ and a constant $0 < \theta < 1$ so that $\sup_{\bar{U}_1 \times V_1} |D_xG(x, y)| \le \theta < 1$. By Lemma 3, $G(\cdot, y)$ is a uniform contraction in $\bar{U}_1$ for all $y \in V_1$. To show $G : \bar{U}_1 \times V_1 \to \bar{U}_1$, we note first that for $x \in \bar{U}_1$, $|G(x, y_0) - x_0| = |G(x, y_0) - G(x_0, y_0)| \le \theta|x - x_0| \le \theta\delta_1 < \delta_1$. Hence by the continuity of G we have $|G(x, y) - x_0| \le \delta_1$ for $(x, y) \in \bar{U}_1 \times V_1$ by making δ_2 smaller if necessary. Then the result follows from Theorem 2 with fixed point $x = f(y)$ for $G(\cdot, y)$. □

Theorem 5 (Implicit Function Theorem II)**.** *Let X, Y, Z be Banach spaces, $U \subset X$, $V \subset Y$ be open sets, and $F : U \times V \to Z$ be continuously differentiable in both variables. If there is a point $(x_0, y_0) \in U \times V$ such that $F(x_0, y_0) = 0$ and $D_xF(x_0, y_0)$ has a bounded inverse, then there is a neighborhood $U_1 \times V_1 \subset U \times V$ of (x_0, y_0) and a continuously differentiable function $f : V_1 \to U_1$ with $f(y_0) = x_0$ such that $F(x, y) = 0$ for $(x, y) \in U_1 \times V_1$ iff $x = f(y)$. Also,*

$$Df(y) = -[D_xF(f(y), y)]^{-1} D_yF(f(y), y).$$

Moreover, if $F \in C^k(U \times V, Z)$, $k \ge 1$ or $C^{k,1}$ or analytic in a neighborhood of (x_0, y_0), then $f \in C^k(V_1, U_1)$ or $C^{k,1}$ or is analytic in a neighborhood of y_0.

Proof. The proof is exactly the same as the previous proof except for that the Uniformly Contraction Principle II (Theorem 3) is applied at the end for the solution $x = f(y)$ for $F(f(y), y) = 0$. In addition, apply implicit differentiation to $F(f(y), y) \equiv 0$ to obtain the derivative formula for Df, which is well-defined by making V_1, U_1 smaller if necessary. □

Reference: S.-N. Chow and J.K. Hale, Methods of Bifurcation Theory, Springer-Verlag, 1982.