Mathematical Analysis II

Felix Chen

Contents

For a function $f: \mathbb{R}^d \to \mathbb{R}$, $df(x_0) = (\nabla f)(x_0)$ is a function $\mathbb{R}^d \to \mathbb{R}^d$, hence $d(df)(x_0) = J(\nabla f)$ is a matrix. If we look at the higher derivatives, it will become an n dimensional array, which is hard to represent.

When we have multiple functions to deal with, the differentiation is almost the same as 1 dimensional case:

Proposition 0.0.1 (Chain rule)

Let $\Omega_i \subset \mathbb{R}^{n_i}, 1 \leq i \leq 3$ be open sets, and $f: \Omega_1 \to \Omega_2, g: \Omega_2 \to \Omega_3$ be differentiable functions. Then $g \circ f: \Omega_1 \to \Omega_3$ is differentiable, and

$$d(g \circ f)(x) = dg\big|_{f(x)} \cdot df(x).$$

where dg is a $n_3 \times n_2$ matrix, df is a $n_2 \times n_1$ matrix, so d $(g \circ f)$ is a $n_3 \times n_1$ matrix, as defined above.

Proof. Let $f(x_0) = y_0$,

$$f(x_0 + v) = y_0 + df(x_0)v + o(|v|),$$

and

$$g(y_0 + w) = g(y_0) + dg(y_0)w + o(|w|).$$

Now we compute

$$g(f(x_0 + v)) = g(y_0 + df(x_0)v + o(|v|))$$

$$= g(y_0) + dg(y_0)(df(x_0)v + o(|v|)) + o(|df(x_0)v + o(|v|))$$

$$= g(y_0) + dg(y_0) df(x_0)v + dg(y_0)o(|v|) + o(|df(x_0)v + o(|v|)),$$

so we only need to verify that

$$\lim_{|v|\to 0} \frac{|\operatorname{d}g(y_0)o(|v|) + o(\operatorname{d}f(x_0)v + o(v))|}{|v|} = 0.$$

Note that $|A \cdot v| \leq ||A|| |v|$, where the norm of a matrix is defined as $(\sum A_{ij}^2)^{\frac{1}{2}}$, so it's clear the above limit holds.

Corollary 0.0.2

Let $\Omega_1 \subset \mathbb{R}^{n_1}$, $\Omega \subset \mathbb{R}^{n_2}$, let f be a differentiable map $\Omega_1 \to \Omega_2$. If f is a bijection and f^{-1} is differentiable, then:

- $n_1 = n_2$;
- $df^{-1}(y) = (df)^{-1}(x)$, where $x = f^{-1}(y)$.

Proof. Consider the composite function id = $f \circ f^{-1} : \Omega_2 \to \Omega_2$, by chain rule,

$$I_{n_2} = \mathrm{d}(f \circ f^{-1}) = \mathrm{d}f \cdot \mathrm{d}f^{-1}.$$

since I_{n_2} has rank n_2 , we know that $n_1 \ge n_2$. Similarly $n_2 \ge n_1$, so $n_1 = n_2$. Hence the inverse of df exists and is equal to df^{-1} .

Example 0.0.3

Consider the exponential map:

$$\exp: M_n(\mathbb{R}) \to M_n(\mathbb{R}), A \mapsto \sum_{k=0}^{\infty} \frac{A^k}{k!} =: e^A.$$

then $d \exp(A)$ is a linear map $M_n(\mathbb{R}) \to M_n(\mathbb{R})$.

By definition,

$$e^{A+V} - e^A = \operatorname{d}\exp(A) \cdot V + o(|V|).$$

The left hand side is equal to

$$\sum_{l=0}^{\infty} \frac{(A+V)^k - A^k}{k!} = \sum_{l=0}^{\infty} \frac{\sum_{l=0}^{k-1} A^l V A^{k-1-l} + O(|V|^2)}{k!}.$$

since $||AB|| \le ||A|| ||B||$, the $O(|V|^2)$ part has norm at most $2^k ||V||^2 ||A||^{k-2}$.

$$\implies e^{A+V} - e^A = \sum_{k=0}^{\infty} \frac{\sum_{l=0}^{k-1} A^l V A^{k-1-l}}{k!} + o(\|V\|).$$

In particular,

- $\operatorname{d}\exp(I)(V) = \sum_{k=0}^{\infty} \frac{kV}{k!} = eV;$
- $d \exp(0)(V) = V$;
- If A and V is commutative, $d \exp(A)(V) = \exp(A)V$.

Theorem 0.0.4 (Substitution formula)

Let $\phi: U \to V$ be a bijection, ϕ, ϕ^{-1} are C^1 functions, and Jacobi determinant

$$J_{\phi}(x) := \det(J(\phi)(x)) \neq 0, \quad \forall x \in U.$$

If f is Lebesgue integrable on V, then

$$\int_{V} f(y) \, \mathrm{d}y = \int_{U} f(\phi(x)) |J_{\phi}(x)| \, \mathrm{d}x.$$

Remark 0.0.5 — In fact we only need to check for cuboid I,

$$m(\phi(I)) = \int_{I} |J_{\phi}(x)| \, \mathrm{d}x.$$

and ϕ maps null sets to null sets.

Proof. Since $\phi \in C^1$, exists constant M s.t.

$$M^{-1} < \|\mathrm{d}\phi\|, \|\mathrm{d}\phi^{-1}\|, |J_{\phi}| < M.$$

 $\forall \varepsilon > 0$, divide I into sufficiently small cuboids I_j , such that

$$\phi(x) - \phi(x_i) - d\phi(x_i)(x - x_i) \le M\varepsilon |x - x_i|, \quad \forall x \in I_i,$$

where x_i is the center of I_i , because

$$\phi(x) - \phi(x_j) = \int_0^1 \frac{d}{dt} \phi(tx + (1 - t)x_j) dt$$

$$= \int_0^1 d\phi(tx + (1 - t)x_j)(x - x_j) dt$$

$$= d\phi(x_j)(x - x_j) + \int_0^1 (d\phi(tx + (1 - t)x_j) - d\phi(x_j)) dt \cdot (x - x_j)$$

Hence there exists K independent of ε .

$$m(\phi(I_i)) \le (|J_{\phi}(x_i)| + MK\varepsilon)m(I_i).$$

since the image $\phi(I_j)$ is a subset of $d\phi(x_j)(I_j)$ (which is a parallogram) extending $M\varepsilon|x-x_j|$ on each side.

By taking sufficiently small ε ,

$$m(\phi(I)) \le \sum_{j} (|J_{\phi}(x_j)| + MK\varepsilon) m(I_j) = 2MK\varepsilon m(I) + \int_{I} |J_{\phi}(x)| dx.$$

Therefore

$$\int_{V} f(y) \, \mathrm{d}y \le \int_{U} f(\phi(x)) |J_{\phi}| \, \mathrm{d}x.$$

apply this to ϕ^{-1} we'll get the equality:

$$m(E) \le \int_{\phi^{-1}(E)} |J_{\phi}(x)| \, \mathrm{d}x \le \int_{E} |J_{\phi}(\phi^{-1}(x))| |J_{\phi^{-1}}(x)| \, \mathrm{d}x = m(E).$$

Example 0.0.6

Consider the spherical coordinates $x = r \sin \theta \sin \varphi$, $y = r \sin \theta \cos \varphi$, $z = r \cos \theta$. Let $F: (r, \theta, \varphi) \mapsto (x, y, z)$.

$$J_F = \begin{pmatrix} \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \cos \theta & -\sin \theta & 0 \end{pmatrix}$$

So $det(J_F) = r^2 \sin \theta$. Thus

$$\int_{\mathbb{R}^3} f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \int_{(0, 2\pi)^2} \int_0^{+\infty} f(r, \theta, \phi) r^2 \sin \theta \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\varphi.$$

Theorem 0.0.7 (Clairaut-Schwarz)

Given an open set $\Omega \subset \mathbb{R}^n$ and $f: \Omega \to \mathbb{R}$. Assume $\frac{\partial f}{\partial x_i}(x), \frac{\partial f}{\partial x_j}(x), \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(x)$ exists and are continuous, then $\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(x)$ exists and

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (x) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_j} \right) (x).$$

Proof. WLOG n = 2. We'll just expand and compute:

$$\frac{\partial}{\partial y} \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{s \to 0} \frac{1}{s} \left(\frac{\partial f}{\partial x}(x_0, y_0 + s) - \frac{\partial f}{\partial x}(x_0, y_0) \right) \\
= \lim_{s \to 0} \frac{1}{s} \lim_{t \to 0} \frac{1}{t} (f(x_0 + t, y_0 + s) - f(x_0, y_0 + s) - f(x_0 + t, y_0) + f(x_0, y_0)).$$

Since

$$(f(x_0 + t, y_0 + s) - f(x_0, y_0 + s) - f(x_0 + t, y_0) + f(x_0, y_0)) = \int_0^s \int_0^t \frac{\partial}{\partial x} \frac{\partial f}{\partial y} (x_0 + \tilde{t}, y_0 + \tilde{s}) d\tilde{t} d\tilde{s}.$$

So by Fubini's theorem,

Notation: Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be a multiple index, where $\alpha_i \geq 0$ are integers. define

$$\partial^{\alpha} f = \left(\frac{\partial f}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial f}{\partial x_n}\right)^{\alpha_n} f.$$

or we can write

$$\partial^{\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad |\alpha| = \sum_{i=1}^n \alpha_i.$$

Theorem 0.0.8 (Multi-dimensional Taylor expansion)

Let $\Omega \subset \mathbb{R}^n$ be a convex open set. Let $f \in C^{k+1}(\Omega)$, for all $x, y \in \Omega$, then $\exists \theta \in (0, 1]$ s.t.

$$f(y) = \sum_{|\alpha| \le k} \frac{\partial^{\alpha} f(x)}{\alpha!} (y - x)^{\alpha} + \sum_{|\alpha| = k+1} \frac{\partial^{\alpha} f(x + \theta(y - x))}{\alpha!} (y - x)^{\alpha}.$$

where $(y-x)^{\alpha} = \prod_{i=1}^{n} (x_i - y_i)^{\alpha_i}$, $\alpha! = \prod_{i=1}^{n} \alpha_i!$.

Proof. Let g(t) = f(ty + (1-t)x), $g \in C^{k+1}((-1,1))$. By Taylor expansion, there exists $\theta \in [0,1]$,

$$g(1) = \sum_{l=0}^{k} \frac{g^{(l)}(0)}{l!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

so it's just a differential formula of composite function, which can be easily proved by induction, and I don't bother to write it down. \Box

§0.1 Implicit function theorem

As usual let $C^k(\Omega)$ denote the k times continuously differentiable functions on Ω .

Definition 0.1.1 (Differential homeomorphisms). Let $U, V \subset \mathbb{R}^n$, if there exists a bijection $f: U \to V$, such that f, f^{-1} are smooth, then we say U and V are **smoothly homeomorphic**. Denoted by $C^{\infty}(U, V)$ the smooth homeomorphisms from U to V.

Example 0.1.2

Let $f: \mathbb{R} \to \mathbb{R}$ by $x \mapsto x^3$, then f is a smooth bijection, but f^{-1} is not differentiable at 0.

Recall that in \mathbb{R} we have the following results:

- If f is strictly increasing and continuous, then f^{-1} is continuous.
- If f is strictly increasing and C^1 , $f' \neq 0$, then $f^{-1} \in C^1$.

Theorem 0.1.3

Let Φ be an differential homeomorphism $U \to V$, $f \in C^k(V)$. Then $f \circ \Phi =: \Phi^* f \in C^k(\Omega)$, this is called the **pullback** of f by Φ .

Proof. We proceed by induction on k. When k = 0, this is just the continuity of composite functions.

Assume k = n holds, then for k = n + 1,

$$\frac{\partial \Phi^* f}{\partial x_j} = \frac{\partial f(\Phi(x))}{\partial x_j} = \sum_{i=1}^n \frac{\partial f}{\partial y_i}(\Phi(x)) \cdot \frac{\partial \Phi_i(x)}{\partial x_j}.$$

Since $f \in C^{n+1} \implies \frac{\partial f}{\partial y_i} \in C^n$, and $\frac{\partial \Phi_i}{\partial x_j}$ is smooth, so $\frac{\partial \Phi^* f}{\partial x_j} \in C^n$.

Note that the condition $f' \neq 0$ grants that f is indeed a bijection locally. In higher dimensional spaces, the derivatives are more complex, so let's look at some simple cases first.

Lemma 0.1.4

Let $U, V \subset \mathbb{R}^d$ be open regions. Let $f: U \to V$ be a C^1 bijection, and J(f) is non-degenerate (i.e. det $J(f) \neq 0$). Then $f^{-1}: V \to U$ is continuously differentiable.

Proof. Let $x_0 \in U$, $y_0 = f(x_0) = V$,

$$f(x_0 + v) = y_0 + A \cdot v + o(|v|).$$

Let $E(\delta)$ be a function which satisfies

$$f^{-1}(y_0 + \delta) = x_0 + A^{-1}\delta + E(\delta).$$

this can be derived from taking f^{-1} on both sides of the above equation.

$$y_0 + \delta = f(x_0 + A^{-1}\delta + E(\delta)) = y_0 + \delta + AE(\delta) + o(|A^{-1}\delta + E(\delta)|)$$

 $\implies AE(\delta) + o(A^{-1}\delta + E(\delta)) = 0.$

From this we can calculate

$$\frac{|E(\delta)|}{|\delta|} = \frac{|o(A^{-2}\delta + A^{-1}E(\delta))|}{|\delta|} \le \frac{|o(A^{-2}\delta + A^{-1}E(\delta))|}{|A^{-2}\delta + A^{-1}E(\delta)|} \cdot \frac{|A^{-2}\delta + A^{-1}E(\delta)|}{|\delta|} \le o(1) \left(C + C\frac{|E(\delta)|}{|\delta|}\right).$$

Hence $\lim_{|\delta| \to 0} \frac{|E(\delta)|}{|\delta|} = 0$.

In this case we are given f^{-1} exists, but generally we need to prove this existence.

Theorem 0.1.5 (Inverse function theorem)

Let $f: \Omega \to \mathbb{R}^d$ be a C^1 map, and $df(x_0)$ is non-degenerate, then f is a C^1 differential homeomorphism in some neighborhood of x_0 .

This is to say, $\exists U \ni x_0, V \ni f(x_0)$ s.t. f is a bijection from U to V and $f^{-1}: V \to U$ is a C^1 map.

Proof. WLOG $x_0 = 0$, $f(x_0) = 0$, also we can apply a linear transformation such that $df(x_0) = I$. There exists $\delta > 0$, s.t.

$$|f(v) - v| < \varepsilon_0 |v|, \quad ||J(f)(v) - I|| < \varepsilon_0, \quad \forall |v| < \delta.$$

note that

$$f(v) - f(u) = \int_0^1 \frac{d}{dt} f(tv + (1-t)u) dt = \int_0^1 df(tv + (1-t)u) \cdot (v-u) dt$$
$$= \int_0^1 (Jf(tv + (1-t)u) - I) \cdot (v-u) dt + (u-v).$$

but when $|u|, |v| < \delta, |f(v) - f(u) - (v - u)| < \varepsilon_0 |v - u|$.

Hence $f(u) = f(v) \implies u = v$, f is injective in $B_{\delta}(0)$.

As for surjectivity, it's sufficient to prove $f(B_{\delta}(0))$ contains a neighborhood of f(0) = 0. i.e. $\forall |v| < \delta_1, \exists |u| < \delta \text{ s.t. } f(u) = u + o(u) = v.$

Since we don't know the non-linear term o(u), we'll iterate to get a solution u: let $u_0 = v$. Define $u_{k+1} = v - (f(u_k) - u_k)$. When δ_1 is sufficiently small,

$$|u_{k+1}| \le |v| + |f(u_k) - u_k| \le |v| + \varepsilon_0 |u_k| \le \delta_1 + \varepsilon_0 \delta \le \delta.$$

Now we prove the convergency:

$$|u_{k+2} - u_{k+1}| = |f(u_{k+1}) - u_{k+1} - f(u_k) + u_k|$$

$$= |\int_0^1 (J(f)(tu_{k+1} + (1-t)u_k) - I) dt(u_{k+1} - u_k)|$$

$$\leq \varepsilon_0 |u_{k+1} - u_k|.$$

by contraction mapping principle we're done.

Remark 0.1.6 — This theorem holds for any Banach space.

Corollary 0.1.7

Let $k \geq 2$ be an integer, when $f \in C^k$ in the above theorem, we can imply that $f^{-1} \in C^k(V)$.

Proof. Since

$$df^{-1}(u) = (df)^{-1}(f^{-1}(u)),$$

so $df \in C^{k-1} \implies (df)^{-1} \in C^{k-1}$.

Theorem 0.1.8 (Implicit function theorem)

Let $f: \Omega \subset \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$ be a continuously differentiable function. If $\exists (x^*, y^*) \in \Omega$ s.t. $f(x^*, y^*) = 0$, and $d_y f(x^*, y^*)$ is inversible, then there exists an open neighborhood $U \subset \mathbb{R}^n$ of x^* , $V \subset \mathbb{R}^p$ of y^* , and a C^1 map $\phi: U \to V$ such that:

$$f(x,\phi(x)) = 0$$
, $d\phi(x) = -(d_u f(x,\phi(x)))^{-1} \cdot d_x f(x,\phi(x))$.

Also if $x \in U$ and f(x, y) = 0, we must have $y = \phi(x)$.

Remark 0.1.9 — This is to say, if f(x,y) = 0, $x \in U, y \in V$, then $y = \phi(x)$. Also remember that $d_u f$ is a $p \times p$ matrix, $d_x f$ is a $p \times n$ matrix.

Proof. By the inverse function theorem, let $F(x,y) := \mathbb{R}^{n+p} \to \mathbb{R}^{n+p}$ with

$$(x,y) \mapsto (x,f(x,y))$$

So $F(x^*, y^*) = (x^*, 0)$, and

$$dF(x^*, y^*) = \begin{pmatrix} I_n & 0 \\ d_x f(x^*, y^*) & d_y f(x^*, y^*) \end{pmatrix}.$$

Since $d_y f(x^*, y^*)$ is inversible, $dF(x^*, y^*)$ is inversible as well. Hence there exists neighborhoods of (x^*, y^*) and $(x^*, 0)$, say $\widetilde{\Omega}$ and $\widetilde{\Omega}_1$, such that F is a C^1 homeomorphism $\widetilde{\Omega} \to \widetilde{\Omega}_1$.

We can find $U \ni x^*, V \ni y^*$ s.t. $U \times V \subset \widetilde{\Omega}$. Let T be the C^1 map s.t.

$$F^{-1}(x,z) = (x,T(x,z)).$$

Let $\phi(x) = T(x,0)$, we have

$$F(x, \phi(x)) = F(x, T(x, 0)) = (x, 0) \implies f(x, \phi(x)) = 0.$$

Since F is a bijection, clearly $f(x,y)=0 \implies y=\phi(x)$. By taking the differentiation of $f(x,\phi(x))=0$,

$$(\mathrm{d}_x f, \mathrm{d}_y f) \cdot \begin{pmatrix} I_n \\ \mathrm{d}\phi(x) \end{pmatrix} = 0 \implies \mathrm{d}_x f(x, \phi(x)) + \mathrm{d}_y f(x, \phi(x)) \cdot \mathrm{d}\phi(x) = 0.$$

§0.2 Submanifolds in Euclid space

The implicit function theorem actually gives an example of manifolds: the preimage of f(x,y) = 0 is an *n*-dimensional manifold in \mathbb{R}^{n+p} .

Definition 0.2.1 (Manifolds). Let $M \subset \mathbb{R}^n$ be a nonempty set. If $\exists d \geq 0, \forall x \in M$ exists open sets U and V in \mathbb{R}^n , and a differential homeomorphism $\Phi: U \to V$, such that

$$\Phi(U\cap M)=V\cap\{\mathbb{R}^d\times\{0\}\},$$

we say M is a d-dimensional differential manifold. Denote dim M = d, and n - d is called the **codimension** of M.

Corollary 0.2.2 (Regular value theorem)

Let $f: \Omega \to \mathbb{R}^p$ be a smooth map, where $\Omega \subset \mathbb{R}^n$ is an open set, $n \geq p$. For all $c \in \mathbb{R}^p$, we call the **fibre** of c to be its preiamge:

$$f^{-1}(c) = \{ x \in \mathbb{R}^n \mid f(x) = c \}.$$

If $\forall x \in f^{-1}(c)$, rank df(x) = p, then $f^{-1}(c)$ is a manifold with codimension p.

Example 0.2.3

Let S^{n-1} be the unit sphere in \mathbb{R}^n . Let $f: \mathbb{R}^n \to \mathbb{R}$ with $x \mapsto |x|^2 - 1$, then $S^{n-1} = f^{-1}(0)$. Since $\mathrm{d}f = (2x_1, 2x_2, \dots, 2x_n)$, clearly rank $\mathrm{d}f = 1$ for all $x \in S^{n-1}$, so S^{n-1} is a manifold with codimension 1.

Example 0.2.4

Consider a surface in $\mathbb{R}^4 = \mathbb{C}^2$:

$$T^2 = \{(z_1, z_2) \mid |z_1| = 1, |z_2| = 1\}.$$

Let
$$f(x, y, z, w) = x^2 + y^2 - 1$$
, $g(x, y, z, w) = z^2 + w^2 - 1$, then $T^2 = \begin{pmatrix} f \\ g \end{pmatrix}^{-1}$ (0).

The differentiation is

$$d\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} 2x & 2y & 0 & 0 \\ 0 & 0 & 2z & 2w \end{pmatrix},$$

so T^2 is a manifold with codimension 2.

Definition 0.2.5. Let $M \subset \mathbb{R}^n$ be a manifold. If dim M = 1, we say M is a curve; if dim M = 2, M is a surface; and if dim M = n - 1, we say M is a hyperplane.

Lemma 0.2.6

Let $f: \mathbb{R}^n \to \mathbb{R}$ is a smooth function, if $\forall x_0 \in f^{-1}(0)$, $\mathrm{d}f(x_0) \neq 0$, then $f^{-1}(0)$ is a smooth hyperplane, it is called the hyperplane globally determined by an equation.

Example 0.2.7

In \mathbb{R}^3 , f, g are smooth functions. If for all $x \in \mathbb{R}^3$ with f(x) = g(x) = 0 we have $\nabla f, \nabla g$ are linearly independent, then $\{f = g = 0\}$ is a smooth curve.

Theorem 0.2.8 (Parametrization of manifolds)

Let Ω be an open set in \mathbb{R}^n , $f: \Omega \to \mathbb{R}^{n+p}$ is a smooth map. If $\forall x^* \in \Omega$, rank $\mathrm{d} f(x^*) = n$, then there exists an open set $U, x^* \in U$ s.t. $f(U) \subset \mathbb{R}^{n+p}$ is an n-dimensional manifold.

Proof. Let x_i be a coordinate in \mathbb{R}^n , y_j be a coordinate in \mathbb{R}^p .

WLOG $(\frac{\partial f_i}{\partial x_j})_{1 \leq i,j \leq n}$ is non-degenerate, let $F = (f_1,\ldots,f_n), G = (f_{n+1},\ldots,f_{n+p})$ and apply inverse function theorem on F,