1.4 The Jacobian of a map

Derivative of a differentiable map

Let $F : M^n \to N^m$ be a differentiable map between two $C^1$ manifolds. Given a point $p \in M$ we define the derivative of $F$ at $p$ by

$$dF_p : T_p M \to T_{F(p)} N,$$

where for $X \in T_p M$,

$$dF_p (X) (\varphi) \equiv X (f \circ F)$$

for any function $f : N \to \mathbb{R}$. It is easy to check that the map $dF_p (X) : T_{F(p)} N \to \mathbb{R}$ is a derivation (the product rule follows from $X$ being a derivation and equality $(fg) \circ F = (f \circ F) \cdot (g \circ F)$). Another notation for the derivative of $F$ is $F^*$.

We also say that $F^*(X)$ is the pushforward of $X$.

We have the following properties of the derivative (see Lemma 3.5 on p. 66 of Lee [5]).

**Lemma 1** Suppose that $F : M \to N$ and $G : N \to P$ are differentiable maps between $C^1$ manifolds and let $p \in M$. We have the following.

1. $dF_p : T_p M \to T_{F(p)} N$ is a linear map.
2. $d(G \circ F)_p = dG_{F(p)} \circ dF_p$.
3. $d(id_M)_p = id_{T_p M}$.
4. If $F$ restricted to some neighborhood of $p$ is a diffeomorphism, then $dF_p$ is an isomorphism.

It is useful to understand the derivative with respect to local coordinates. The use of local coordinates to describe objects and morphisms in geometric analysis is prevalent. We have the following (see (3.6) on p. 70 of Lee [5]).

**Lemma 2** If $\{x^i\}_{i=1}^n$ is a local coordinate system defined in an open neighborhood of a point $p \in M$ and if $\{y^\alpha\}_{\alpha=1}^m$ is a local coordinate system defined in an open neighborhood of $F(p) \in N$, then we have

$$dF_p (X) = \sum_{i,\alpha} X^i \frac{\partial (y^\alpha \circ F)}{\partial x^i} \frac{\partial}{\partial y^\alpha},$$

where $X^i = X (x^i)$ and $X = \sum_{i} X^i \frac{\partial}{\partial x^i}$.

Note that

$$\frac{\partial (y^\alpha \circ F)}{\partial x^i} = e_i \left( y^\alpha \circ F \circ x^{-1} \right).$$

Recall that given a $C^\infty$ path $\gamma : (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$, we have the derivation $\dot{\gamma}(0) \in T_p M$. Using definition (1), we then may compute the derivative of a differentiable map $F : M \to N$ by

$$dF_p (\dot{\gamma}(0)) (f) = \dot{\gamma}(0) (f \circ F)$$

$$= \frac{d}{ds} \bigg|_{s=0} (f \circ F \circ \gamma) (s)$$

$$= (F \circ \gamma)' (0) (f)$$

for $f : \mathcal{V} \to \mathbb{R}$, where $\mathcal{V}$ is an open neighborhood of $F(p)$; see Proposition 3.12 on p. 76 of Lee [5]).

**Inverse and implicit function theorems**

In this section we discuss the inverse function theorem and the implicit function theorem. These results are basic to understanding immersions and submersions.
Basic functional analysis

We first recall some basic definitions and general results from functional analysis. We refer the reader to Rudin’s ‘Functional Analysis’ book [6] for further background and proofs. A topological vector space is a vector space 𝑉 with a topology 𝜓 such that

1. for every 𝑥 ∈ 𝑉 the set {𝑥} is a closed set,
2. the vector space operations of addition + : 𝑉 × 𝑉 → 𝑉 and scalar multiplication ⋅ : ℝ × 𝑉 → 𝑉 are continuous maps with respect to the topologies induced by 𝜓 and the standard topology on ℝ.

A topological vector space (𝑉, 𝜓) is an 𝐹-space if the topology 𝜓 is induced by a complete translation invariant metric 𝑑. By translation invariant we mean the for any 𝑥, 𝑦, 𝑧 ∈ 𝑉 we have

\[ 𝑑(𝑥 + 𝑧, 𝑦 + 𝑧) = 𝑑(𝑥, 𝑦). \]

A Banach space is vector space 𝑋 with a norm ∥ ⋅ ∥ such that the induced metric space with distance

\[ 𝑑(𝑥, 𝑦) = ∥𝑥 − 𝑦∥ \]

is complete, i.e., every Cauchy sequence converges.\(^1\) Note that a Banach space is an 𝐹-space.

Given topological vector spaces (𝑋, 𝜓𝑋) and (𝑌, 𝜓𝑌), let 𝐿(𝑋, 𝑌) denote the vector space of continuous linear maps from 𝑋 to 𝑌. If (𝑋, ∥ ⋅ ∥𝑋) and (𝑌, ∥ ⋅ ∥𝑌) are Banach spaces, then we may define a norm on 𝐿(𝑋, 𝑌) by

\[ ∥𝐴∥ = \sup_{∥𝑥∥=1} ∥𝐴(𝑥)∥_𝑌 = \sup_{𝑥 ≠ 0} \frac{∥𝐴(𝑥)∥_𝑌}{∥𝑥∥_𝑋}. \]

(𝐿(𝑋, 𝑌), ∥ ⋅ ∥) is a Banach space.

If a continuous linear operator between Banach spaces 𝐴 : 𝑋 → 𝑌 is bijective with continuous inverse 𝐴⁻¹ : 𝑌 → 𝑋, then we say that 𝐴 a Banach space isomorphism. We denote the set of Banach space isomorphisms from 𝑋 to 𝑌 by Iso(𝑋, 𝑌). The subset Iso(𝑋, 𝑌) ⊂ 𝐿(𝑋, 𝑌) is open (see Proposition 2.5 on p. 5 of [4]).

The open mapping theorem says the following (see Theorem 2.11 on p. 47 of [6]).

**Theorem 3 (Open mapping)** If 𝐴 : 𝑋 → 𝑌 is a continuous linear map from an 𝐹-space to a topological vector space and 𝐴(𝑋) is of the second category\(^2\) in 𝑌, then 𝐴 is onto, 𝐴 is an open mapping, and 𝑌 is an 𝐹-space.

In particular, a surjective continuous linear map 𝐴 : 𝑋 → 𝑌 between Banach spaces is an open mapping.

A consequence of the open mapping theorem is the closed graph theorem (see Theorem 2.15 on p. 50 of [6]).

**Theorem 4 (Closed graph)** Let 𝐴 : 𝑋 → 𝑌 be a linear map between 𝐹-spaces. The graph

\[ \{(𝑥, 𝐴(𝑥)) : 𝑥 ∈ 𝑋\} ⊂ 𝑋 × 𝑌 \]

is a closed subset if and only if the map 𝐴 is continuous.

This in turn implies the following.

**Theorem 5 (Inverse of a continuous bijection is continuous)** If 𝐴 : 𝑋 → 𝑌 is a bijective continuous linear operator between Banach spaces (for example, Hilbert spaces), then its inverse 𝐴⁻¹ : 𝑌 → 𝑋 is continuous.\(^3\) That is, 𝐴 a Banach space isomorphism.

\(^1\) A norm ∥ ⋅ ∥ : 𝑋 → [0, ∞) is defined to satisfy

\[ ∥𝑥 + 𝑦∥ ≤ ∥𝑥∥ + ∥𝑦∥, \]

\[ ∥𝑐𝑥∥ = |𝑐| ∥𝑥∥ \]

for all 𝑥, 𝑦 ∈ 𝑋 and 𝑐 ∈ ℝ with ∥𝑥∥ = 0 if and only if 𝑥 = 0.

\(^2\) A set is of the second category if it cannot be written as the union of a countable collection of nowhere dense sets.

\(^3\) An elementary fact is that being bounded is equivalent to being continuous (see Theorem 1.32 in [6] for example).
Proof. By the closed graph theorem it suffices to show that
\[ Y \times X \supset \{ (y, A^{-1}(y)) : y \in Y \} = \{ (A(x), x) : x \in X \} \]
is a closed set. This follows from the continuity of \( A \).

Let \( A : X \to Y \) be a linear map between topological vector spaces. The cokernel of \( A \) is
\[ \text{coker}(A) = Y/\text{image}(A). \]
Hence, by definition, the cokernel of \( A \) is trivial if and only if \( A \) is onto. Note that if \( X \) and \( Y \) are Hilbert spaces, then \( \text{coker}(A) \cong (\text{image}(A))^\perp \), where an isomorphism is induced by the projection map.

Let \( A : X \to X \) be a self-adjoint linear map of a Hilbert space. We have \( y \in (\text{image}(A))^\perp \) if and only if for every \( x \in X \)
\[ 0 = \langle y, Ax \rangle = \langle A^*y, x \rangle = \langle Ay, x \rangle, \]
i.e., \( Ay = 0 \). If \( A \) is one-to-one, then we conclude \( (\text{image}(A))^\perp = \{0\} \). That is, a one-to-one self-adjoint linear map of a Hilbert space is onto.

Implicit function theorem

First we recall the contraction mapping principle.

Lemma 6 (Contraction maps have unique fixed points) Let \( (X, d) \) be complete metric space. If \( f : X \to X \) is such that there exists \( \lambda \in [0, 1) \) such that
\[ d(f(x), f(y)) \leq \lambda d(x, y) \]
for all \( x, y \in X \), then there exists a unique point \( x_\infty \in X \) such that
\[ f(x_\infty) = x_\infty. \]

Proof. The idea is to iterate the map \( f \) (starting at any point) and obtain a Cauchy sequence which converges to the fixed point. Let \( x_0 \in X \) and define \( x_i \in X \), where \( i \in \mathbb{N} \), inductively by
\[ x_{i+1} = f(x_i). \quad (2) \]
We have
\[ d(x_i, x_{i+1}) \leq \lambda d(x_{i-1}, x_i) \leq \cdots \leq \lambda^i d(x_0, x_1) \]
for all \( i \in \mathbb{N} \) and hence
\[ d(x_i, x_j) \leq d(x_i, x_{i+1}) + \cdots + d(x_j-1, x_j) \leq (\lambda + \cdots + \lambda^{j-1}) d(x_0, x_1) \leq \frac{\lambda^i}{1-\lambda} d(x_0, x_1) \]
for any \( j > i \). Since \( \lambda \in [0, 1) \), we have \( \{x_i\}_{i=0}^\infty \) is a Cauchy sequence. Since \( X \) is complete, the limit
\[ x_\infty = \lim_{i \to \infty} x_i \]
exists. Taking the limit as \( i \to \infty \) of \( (2) \) and using the continuity of \( f \), we have
\[ x_\infty = \lim_{i \to \infty} x_{i+1} = \lim_{i \to \infty} f(x_i) = f(\lim_{i \to \infty} x_i) = f(x_\infty). \]
The fixed point \( x_\infty \) is unique since if \( x'_\infty \) is also a fixed point of \( f \), then
\[ d(x_\infty, x'_\infty) = d(f(x_\infty), f(x'_\infty)) \leq \lambda d(x_\infty, x'_\infty), \]
which implies \( d(x_\infty, x'_\infty) = 0 \).

Next we state the inverse function theorem (see Theorem 5.2 on p. 13 of [4] or Dieudonné [1]).
Theorem 7 (Banach space IFT) Let \((X, \| \cdot \|_X)\) and \((Y, \| \cdot \|_Y)\) be Banach spaces and let \(U \subset X\) be an open set. If \(f : U \to Y\) is a \(C^1\) map and \(x_0 \in U\) is such that \(Df(x_0) : X \to Y\) is a bijection, then there exists an open neighborhood \(W\) of \(x_0\) such that \(f|_W : W \to f(W)\) is a bijection onto an open neighborhood of \(f(x_0)\) and the inverse of \(f|_W\), i.e.,

\[(f|_W)^{-1} : f(W) \to W,\]

is a \(C^1\) map.

**Proof.** Without loss of generality we may assume, by replacing \(f\) by the map

\[x \mapsto (Df(x_0))^{-1}(f(x + x_0) - f(x_0)),\]

that \(X = Y\), \(x_0 = 0\), \(f(0) = 0\) and \(Df(x_0) = \text{id}_X\). First define \(g : U \to X\) by

\[g(x) = x - f(x),\]

which is the difference of \(f\) and its linearization at 0. We have \(Dg(0) \equiv 0\) and since \(g\) is \(C^1\), there exists \(\delta > 0\) such that

\[\|Dg(x)\| \leq \frac{1}{2}\]

for \(x \in \overline{B}(0, \delta) \equiv \{x \in X : \|x\| \leq \delta\}\). Note that this implies for \(x \in \overline{B}(0, \delta),\)

\[\|Df(x)(v)\| \geq \|v\| - \|Dg(x)(v)\| \geq \frac{1}{2}\|v\|\]

for all \(v \in X\).

Since \(g(0) = 0\), by the fundamental theorem of calculus, we have for any \(x \in \overline{B}(0, \delta),\)

\[g(x) = \int_0^1 Dg(tx)(x)\, dt \leq \|x\| \int_0^1 \|Dg(tx)\|\, dt \leq \frac{1}{2}\|x\|,\]

so that \(g(x) \in \overline{B}(0, \delta/2)\).

Now given \(y \in X\), define \(g_y : U \to X\) by

\[g_y(x) = y + x - f(x) = y + g(x).\]

In particular, \(f(x) = y\) if and only if \(g_y(x) = x\). For \(y \in \overline{B}(0, \delta/2)\) we have by the triangle inequality,

\[g_y(\overline{B}(0, \delta)) \subset \overline{B}(0, \delta).\]

Now if \(x_1, x_2 \in \overline{B}(0, \delta),\) then

\[\|g_y(x_1) - g_y(x_2)\| = \left\|\int_0^1 Dg_y(tx_1 + (1-t)x_2)(x_1 - x_2)\, dt\right\| \leq \frac{1}{2}\|x_1 - x_2\|\]

since \(Dg_y = Dg\) and by (3). Hence for any \(y \in \overline{B}(0, \delta/2)\) the map

\[g_y|_{\overline{B}(0, \delta)} : \overline{B}(0, \delta) \to \overline{B}(0, \delta)\]

is a contraction mapping of a complete metric space. By Lemma 6 there exists a unique fixed point \(x \in \overline{B}(0, \delta)\) of \(g_y|_{\overline{B}(0, \delta)}\), that is, there exists a unique point \(x \in \overline{B}(0, \delta)\) such that \(f(x) = y\). In other words, we have a map

\[f^{-1} : \overline{B}(0, \delta/2) \to \overline{B}(0, \delta)\]

such that \((f \circ f^{-1})(y) = y\) for all \(y \in \overline{B}(0, \delta/2)\).
Now for any \(x_1, x_2 \in B(0, \delta)\),
\[
\| f(x_1) - f(x_2) \| \geq \| x_1 - x_2 \| - \| g(x_1) - g(x_2) \|
\]
\[
\geq \frac{1}{2} \| x_1 - x_2 \|.
\]
Hence, given any \(y_1, y_2 \in B(0, \delta/2)\), we have
\[
\| f^{-1}(y_1) - f^{-1}(y_2) \| \leq 2 \| y_1 - y_2 \|.
\]
Thus \(f^{-1}\) is (Lipschitz) continuous.

To see its differentiability, first note that by (4) we have
\[
\sup_{x \in B(0, \delta)} \left\| (Df(x))^{-1} \right\| \leq 2.
\]
Hence given \(y \in B(0, \delta/2)\) and \(h \in X\) with \(0 < \| h \| \leq \delta/2 - \| y \|\), we have
\[
\left\| f^{-1}(y + h) - f^{-1}(y) - (Df(f^{-1}(y)))^{-1}(h) \right\|
\leq 2 \left\| Df(f^{-1}(y)) \right\| (f^{-1}(y + h) - f^{-1}(y)) - h
\]
\[
= o \left( \| f^{-1}(y + h) - f^{-1}(y) \| \right)
\]
\[
= o (\| h \|),
\]
where we also used (5). Therefore \(f^{-1}\) is differentiable in \(B(0, \delta/2)\) and since \(Df\) and \(f^{-1}\) are continuous, we have
\[
D (f^{-1}) (y) = (Df(f^{-1}(y)))^{-1}
\]
is a continuous function of \(y\), i.e., \(f^{-1}\) is \(C^1\).

**Remark 8** Theorem 7 also holds if we replace \(C^1\) by \(C^k\) in the two instances where it occurs in the statement, for \(k \in \mathbb{N}\) (see pp. 14–15 of [4]).

### 1.5 Curves and integral curves

We say that \(\gamma : (a, b) \to M\) is an integral curve of a vector field \(X\) if
\[
\frac{d}{ds} \gamma(s) = X(\gamma(s))
\]
for all \(s \in (a, b)\). If \(X\) is a locally Lipschitz continuous vector field defined on an open set \(\mathcal{V}\), then for every \(x_0 \in \mathcal{V}\) there exists \(\varepsilon > 0\) and an integral curve \(\gamma : (-\varepsilon, \varepsilon) \to M\) of \(X\) such that \(\gamma(0) = x_0\). To see this, choose a local coordinate chart \((U, \bar{x})\) with \(\bar{x}(x_0) = 0\). Note that \(X(x) = \sum_{i=1}^{n} \overrightarrow{\partial x_i}(x) \frac{\partial}{\partial x_i}\) in \(U\). Since \(X^i \circ \bar{x}^{-1}\) are locally Lipschitz functions in \(\bar{x}(U)\), by the uniqueness and existence theorems for systems of ordinary differential equations (see Theorem 3 on p. 7 and Theorem 4 on p. 8 of Hurewicz [3] or Hartman [2]), there exists \(\varepsilon > 0\) and a unique curve \(c \in \{ c^1, \ldots, c^n \} : (-\varepsilon, \varepsilon) \to \bar{x}(U)\) such that
\[
\frac{dc^i}{ds}(s) = (X^i \circ \bar{x}^{-1})(c(s)),
\]
\[
c(0) = 0.
\]
We leave it to the reader to check that \(\gamma \circ \bar{x}^{-1} \circ c : (-\varepsilon, \varepsilon) \to M\) is an integral curve of \(X\) with \(\gamma(0) = x_0\). If \(\mathcal{M}^n\) is closed and if \(X\) is locally Lipschitz continuous, then \(X\) is globally Lipschitz continuous (with a finite Lipschitz constant). In this case, for any \(x \in \mathcal{M}\), there exists an integral curve \(\gamma : \mathbb{R} \to M\) of \(X\) such that \(\gamma(0) = x\). In general (with the noncompact case in mind), we say that a vector field \(X\) on a \(C^\infty\) manifold is **complete** if for any \(x \in \mathcal{M}\) there exists an integral curve \(\gamma : \mathbb{R} \to M\) of \(X\) such that \(\gamma(0) = x\).

We say that a 1-parameter group of diffeomorphisms \(\{ \varphi_s : \mathcal{M} \to \mathcal{M} \}_{s \in \mathbb{R}}\) is generated by a vector field \(X\) if
\[
\frac{\partial}{\partial s} \varphi_s(x) = X(\varphi_s(x))
\]
for all \(x \in \mathcal{M}\) and \(s \in \mathbb{R}\). If \(\mathcal{M}^n\) is a closed differentiable manifold and if \(X\) is a Lipschitz continuous vector field, then there exists a 1-parameter group of diffeomorphisms \(\{ \varphi_s : \mathcal{M} \to \mathcal{M} \}_{s \in \mathbb{R}}\) generated by \(X\).\footnote{Clearly, we may replace the condition of \(\mathcal{M}\) being closed by the condition of \(X\) being complete.}
This follows from the existence of globally defined integral curves on closed manifolds stated in the previous paragraph.

References


