Chapter 4: Submersions, Immersions and Embeddings

List of Definitions

Let F be a smooth map between two smooth manifolds M and N, with dimensions m and n respectively.

Definition 1.1

The rank of F at $p \in M$ is the rank of the linear map:

$$dF_p : T_pM \to T_{F(p)}N$$

Definition 1.2

A smooth map $F \in C^\infty MN$ has constant rank if its differential $dF_p : T_pM \to T_{F(p)}N$ has the same rank at every point $p \in M$.

There are three types of constant rank maps that are of interest.

Definition 1.3

F is a smooth submersion if dF_p is a surjection onto $T_{F(p)}N$ at p-everywhere. That is, $\operatorname{rank} dF_p = \dim T_{F(p)}N = \dim N$

Definition 1.4

F is a smooth immersion if dF_p is an injection onto $T_{F(p)}N$ at p-everywhere. That is, $\operatorname{rank} dF_p = \dim T_pM = \dim M$

Definition 1.5

F is a smooth embedding if it is a smooth immersion, and it s a homeomorphism onto its range $F(M) \subseteq N$.

Definition 1.6

F is a local diffeomorphism if every $p \in M$ in its domain induces a neighbourhood $U \subseteq M$ with $F|U : U \to F(U)$ is a diffeomorphism (in the sense of two open sub-manifolds).

Pre-requisites for Chapter 4

Example 1.28 (Matrices of Full Rank)
Let $A \in \mathcal{M}(m \times n, \mathbb{R})$ be the set of $m \times n$ matrices with real entries. A has rank m iff there exists some $m \times m$ sub-matrix of A, denoted by S st S is invertible. We wish to show the set of rank-m matrices is invertible. Indeed, let

$$F : \mathcal{M}(m \times n, \mathbb{R}) \to \mathbb{R},\ \Delta_{m\times m}(A) = \sum_{\substack{S \text{ is a } m\times m \\ \text{sub-matrix of } A}} |\det\{S\}|$$

Since $S \mapsto \det\{S\}$ is continuous in the entries of S, hence continuous in the entries of A, $\Delta_{m\times m}$ is continuous.
So the set $\left\{A \in \mathcal{M}(m \times n,\ \mathbb{R}),\ \text{rank}\, A = m\right\} = F^{-1}(\mathbb{R} \setminus \{0\})$ is open.

Before proving the inverse function theorem, we will need several Lemmas

Proposition 2.1

If A and B are in $L(\mathbf{X}, \mathbf{Y})$, then

$$\|BA\| \leq \|B\|\|A\|$$

Proof. Let $\|x\| = 1$, and

$$\|B(Ax)\| \leq \|B\|\|Ax\| \leq \|B\|\|A\|\|x\|$$

this holds for every $\|x\| = 1$, hence

$$\|BA\| \leq \|B\|\|A\|$$

■

Proposition 2.2

Let f map a convex open set $U \subseteq \mathbb{R}^n$ into $\mathbb{R}^m$, if f is differentiable (pointwise) in U, and there exists some M st its derivative its founded (in the operator norm)

$$\|Df(x)\| \leq M \quad x \in U$$

then, for every pair of elements x_1, x_2 in U,

$$\|f(x_1) - f(x_2)\| \leq M\|x_1 - x_2\|$$

Proof. This proof 'passes the argument' to the scalar-valued version, in short: if x_1 and x_2 are in U. Define

$$c(t) = (1-t)x_1 + tx_2$$

as the convex combination of x_1 and x_2. The takeaway intuition here is that it suffices to check on the line joining the two points', to obtain an estimate for $\|f(x_1) - f(x_2)\|$. Indeed, define

$$g(t) = f(c(t)) \text{ is a curve } g : \mathbb{R} \to \mathbb{R}^m$$

Recall: Theorem 5.19

Proposition 2.3

Let $g : [0,1] \to \mathbb{R}^m$, and g be differentiable on $(0,1)$, then there exists some $x \in (0,1)$ with

$$|f(b) - f(a)| \leq (b-a)|f'(x)|$$

Proof. Read from Rudin Theorem 5.19. ■

Since $Dg(t) = Df(c(t)) \circ Dc(t)$ by the Chain Rule, and $Dc(t) = b - a$ by inspection,

$$\|Dg(t)\| = \|Df(c(t)) \circ Dc(t)\| \leq \|Df\|\|Dc\| = \|Df\|(b-a)$$

This holds for every $t \in [0,1]$. Applying Theorem 5.19 gives

$$\underbrace{\|g(1) - g(0)\|}_{\text{curve endpoints}} \leq M\|b-a\|$$

Replacing $\|g(1) - g(0)\| = \|f(x_1) - f(x_2)\|$ and $\|Df\| \leq M$ we get

$$\|f(x_1) - f(x_2)\| \leq M\|x_1 - x_2\|$$

■

Rudin Inverse Function Theorem 9.24

Proposition 2.4

Suppose $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, and $Df(a)$ is invertible for some $a \in \mathbb{R}^n$, and define $b = f(a)$. Then,

(a) there exist open sets U and V in $\mathbb{R}^n$ such that $a \in U$, $b \in V$, and f is one-to-one on U, and $f(U) = V$.

(b) if g is the inverse of f (which exists, by Part a), defined in V by $g(f(x)) = x$ for every $x \in U$ then $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$

Proof of Part A. We define $Df(a) = A \in \mathbb{R}^{n\times n}$, so A is invertible, and $\|A^{-1}\| \neq 0$, where $\|\cdot\|$ denotes the operator norm. Recall all norms on finite-dimensional vector spaces are equivalent, this will be useful later.
Choose $\lambda > 0$ st

$$\lambda = \left\|A^{-1}\right\|^{-1} 2^{-1} \tag{1}$$

By continuity of $Df(x)$ at the point a, let $\lambda > 0$, this induces a $B(\delta, a)$ with $x \in B(\delta, a)$ means

$$\underbrace{\|Df(x) - Df(a)\|}_{\text{operator norm}} < \lambda \tag{2}$$

as $Df : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^n)$ takes a point in $\mathbb{R}^n$ and returns a linear map., with $L(\mathbb{R}^n, \mathbb{R}^n)$ endowed with the usual vector space structure. Fix $y \in \mathbb{R}^n$, and define

$$\phi(x) = x + \underbrace{A^{-1}(y}_{\text{offset}} - f(x))$$

this is now a function solely in x, and $\phi(x) = x \iff f(x) = y$ is clear, but such a fixed point is not necessarily unique. We claim that it is unique in $B(\delta, a)$. We will use the contractive mapping principle.

Differentiating $\phi(x)$ reads

$$D\phi(x) = \underbrace{I}_{I=A^{-1}A} - A^{-1}Df(x) = A^{-1}(A - Df(x))$$

Proposition 2.1 tells us the norm of a product is bounded above by the product of the norms. Using eqs. (1) and (2), if $x \in U$ we have

$$\|D\phi(x)\| = \left\|A^{-1}(A - Df(x))\right\| \leq \left\|A^{-1}\right\| \|A - Df(x)\| \leq 2^{-1}$$

The total derivative of ϕ is uniformly bounded in U, applying Proposition 2.2 tells us that ϕ is a contractive mapping

$$\|D\phi(x)\| \leq 2^{-1} \implies \|\phi(x_1) - \phi(x_2)\| \leq 2^{-1}\|x_1 - x_2\|$$

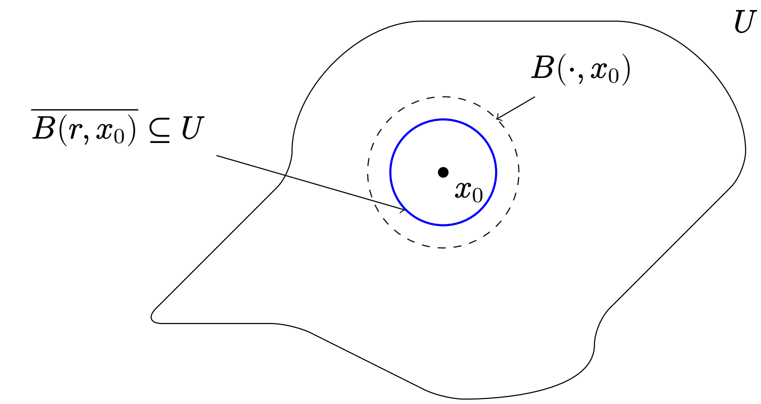

Figure 1: Every point x_0 in an open set U admits an open ball that hides in U

for x_1, x_2 in U.

To show $f|U$ is indeed a bijection, fix $y \in f(U)$ so $y = f(x)$ for some $x \in U$, and there can only be one fixed point stemming from $\phi|U$, with $\phi(z) = z + A^{-1}(y - f(z))$ being the 'fixed point detector'. Write $(f|U)^{-1}(y) = \lim\{(\phi|U)(x_n)\}_n$ and every point in $f(U)$ has a unique inverse.

For the last part of the proof, we wish to show $V = f(U)$ is open. Let $y_0 \in V$ and we can 'hone into' the inverse of y_0 using the same construction as earlier. So $f(x_0) = y_0$ for some unique $x_0 \in U$.

If x_0 is in U, it induces an open ball (see fig. 1) st

$$x_0 \in B(r, x_0) \subseteq \overline{B(r, x_0)} \subseteq U, \quad r > 0$$

We claim the open ball $B(\lambda r, y_0) \subseteq V$. Indeed, suppose $y \in \mathbb{R}^n$ with

$$d(y, y_0) < \lambda r$$

If ϕ is the 'fixed-point detector' with respect to y (the point we are trying to prove that is in $f(U)$), in fact: we will prove $y \in f(\overline{B(r, x_0)}) \subseteq f(U)$.

$$\underbrace{\phi(x_0) - x_0}_{\text{removing the offset from } \phi(x_0)} = A^{-1}(y - f(x_0)) = A^{-1}(y - y_0)$$

using the operator norm on $A^{-1}(y - y_0)$ reads

$$\|\phi(x_0) - x_0\| = \left\|A^{-1}(y - y_0)\right\| \leq \left\|A^{-1}\right\|\|y - y_0\| \leq \left\|A^{-1}\right\|\lambda r = r2^{-1}$$

We will drag y into the image of the closed ball as follows: suppose x is another point that lies in the closed ball, ϕ is contractive on $\overline{B} \subseteq U$ regardless of the point y that induces ϕ. But $\overline{B}$ is closed, hence it is complete. So the Cauchy sequence (from the contractive mapping theorem) produces exactly one point in $\overline{B}$. It remains to show that if we start our sequence at some point $x \in \overline{B}$, then $\phi(x) \in \overline{B}$ as well, and a simple induction will produce our contractive sequence.

To this, fix $x \in \overline{B}$, and

$$\begin{aligned} |\phi(x) - x_0| &\leq |\phi(x) - \phi(x_0)| + |\phi(x_0) - x_0| \\ &\leq \overbrace{2^{-1}|x - x_0|}^{\text{contraction on } \overline{B}\subseteq U} + \overbrace{r2^{-1}}^{\text{earlier}} \\ &= r \end{aligned}$$

therefore ϕ contracts to a fixed point $x^* \in \overline{B}$, and $f(x^*) = y$. So $y \in f(\overline{B}) \subseteq f(U)$ as desired. ■

Proof of Part B. The proof is quite long, and we will only focus on the important bits. Rudin uses the technique of approximating smooth functions using first-order terms. He writes

$$\begin{cases} f(x) &= y \\ f(x+h) &= y + k \end{cases} \implies k = f(x+h) - f(x)$$

Furthermore, if $x \in U$, then the derivative $Df(x)$ is invertible, this is from Theorem 9.8, obtains an estimate on the open ball in $GL(n, \mathbb{R})$. Roughly speaking, this open ball 'drags' other matrices into $GL(n, \mathbb{R})$. If A is invertible, and B is a conformable matrix with A, then

$$\underbrace{\|B - A\|}_{\substack{\text{distance}\\ \text{between}\\ A,B}} \left\|A^{-1}\right\| < 1 \implies B \in GL(n, \mathbb{R})$$

If $x \in B(\delta, a)$, then Equation (2) reads

$$\|Df(x) - A\| < \lambda \implies \|Df(x) - A\| \left\|A^{-1}\right\| < 2^{-1} < 1$$

so $Df(x)$ is invertible with inverse T.

And we estimate the deviation $|k|^{-1} \leq \lambda|h|^{-1}$ by using the contraction inequality with y as the basepoint for ϕ. Skipping a few lines ahead (to the confusing part), we see that

$$|h| \leq |h - A^{-1}k| + |A^{-1}k| \leq 2^{-1}|h| + |A^{-1}k|$$

subtracting over, and multiplying across gives a upper bound on $|k|^{-1}$

$$2^{-1}|h| \leq |A^{-1}k| \implies 2^{-1}|h| \leq \left\|A^{-1}\right\||k| \implies |k|^{-1} \leq \frac{2}{\underbrace{\|A^{-1}\|}_{\lambda}}|h|^{-1}$$

Notice $2\lambda\|A^{-1}\| = 1$, so $2/\|A^{-1}\| = \lambda$. Finally, we 'factor out' $-T$ on the line just before the difference quotient.

$$\begin{aligned}
\overbrace{g(y+k) - g(y) - Tk}^{\substack{\text{numerator in}\\ \text{difference quotient}}} &= h - Tk \\
&= -T\Big(\underbrace{f(x+h) - f(x)}_{=k} - \underbrace{Df(x)h}_{=T^{-1}h}\Big)
\end{aligned}$$

We see that $T = Dg(y)$, indeed:

$$\begin{aligned}
\frac{|g(y+k) - g(y) - Tk|}{|k|} &\leq \frac{\|T\|}{\lambda}\frac{|f(x+h) - f(x) - Df(x)h|}{|h|} \\
&\lesssim \frac{|f(x+h) - f(x) - Df(x)h|}{|h|} \\
= \underbrace{o(h) = o(k)}_{|h| \lesssim |k|} \to 0
\end{aligned}$$

Finally, $Df|U : U \to GL(n, \mathbb{R})$ is a continuous mapping. By Theorem 9.8, $(Df|U)^{-1} : U \to GL(n, \mathbb{R})$ is continuous as well. Therefore $g \in C^1(U, U)$, and $f|U$ is a C^1-diffeomorphism. ■

Commentary

Proposition 4.1 roughly states that, if the differential of F at some point p is injective or surjective, then there exists a neighbourhood U about p such that $dF|U(p)$ is an injection or surjection. The continuity of the map $dF|U(p) \mapsto \Delta_{m\times m}(dF|U(p))$, induces a neighbourhood in the vector space of matrices about the differential $dF|U(p)$. This vector space is endowed with any of the equivalent norms on $\mathcal{M}(m \times n, \mathbb{R})$, which is equivalent to the entrywise 2-norm. Since all partials of the form $\frac{\partial \hat{F}^k}{\partial x^j}\Big|_{\hat{p}}$ are continuous, we take the intersection over all $n \times m$ partials such that $dF|U(p)$ is an injection or surjection. Finally, send this neighbourhood about $\hat{p}$ through to p by using the continuity of ϕ.