Inverse Function Theorem and its Applications

Sachchidanand Prasad

Outline

Statement of the Theorem

Theorem 1 (Inverse Function Theorem)**.** Suppose $U \subseteq \mathbb{R}^n$ is open, $\mathbf{x}_0 \in U$, $\mathbf{f} : U \to \mathbb{R}^n$ is $\mathscr{C}^1$ and $D\mathbf{f}(\mathbf{x}_0)$ is invertible. Then there is a neighborhood $V \subseteq U$, $W \subseteq \mathbb{R}^n$ of $\mathbf{x}_0$ and $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$ respectively and a $\mathscr{C}^1$ function $\mathbf{g} : W \to V$ (see Figure 1) such that

$$\mathbf{f}(\mathbf{g}(\mathbf{y})) = \mathbf{y} \text{ and } \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}, \forall\, \mathbf{x} \in V \text{ and } \forall\, \mathbf{y} \in W.$$

Moreover,

$$D\mathbf{g}(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}$$

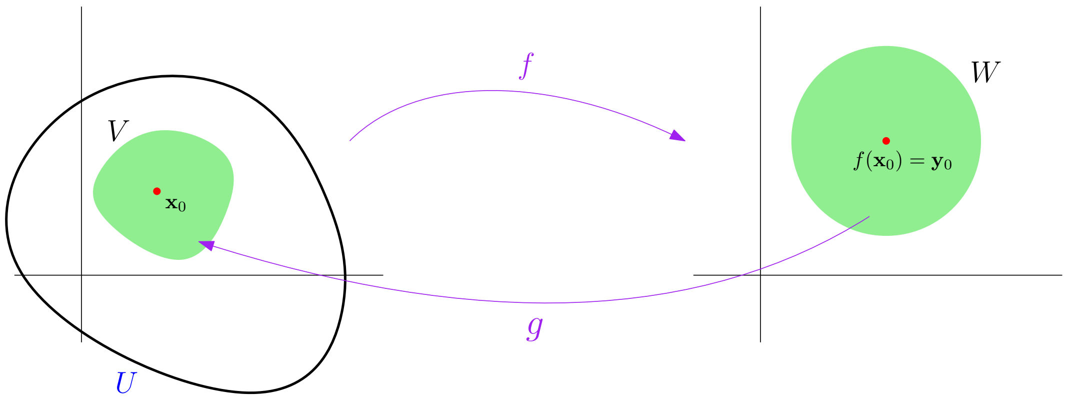

Figure 1

Before seeing the proof of the above theorem we will recall some of the theorems that we are going to use in our proof.

Mean Value Inequality

Theorem 2. Let $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be differentiable on the open set U. Let $\mathbf{x}, \mathbf{y} \in U$ such that $[\mathbf{x}, \mathbf{y}]$[1] $\subseteq U$. Then

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \max_{\mathbf{c} \in [\mathbf{x},\mathbf{y}]} \|D\mathbf{f}(\mathbf{c})\| \cdot \|\mathbf{x} - \mathbf{y}\|. \tag{1}$$

Proof. See [3, p. 248]. □

Contraction Mapping Theorem

Definition 1 (Contraction Map)**.** Let (X, d) be a metric space and $\phi : X \to X$ be any function. If there exists a number $c \in (0, 1)$ such that for all $x, y \in X$

$$d(\phi(x), \phi(y)) \leq c\, d(x, y), \tag{2}$$

then ϕ is said to be a *contraction map* of X.

[1]line joining $\mathbf{x}$ and $\mathbf{y}$

Theorem 3. If X is a complete metric space, and if ϕ is a contraction map of X, then there exists a unique $x \in X$ such that $\phi(x) = x$, i.e. ϕ has a unique fixed point.

Proof. See [2, p. 220]. □

Proof of the Theorem 1

Without loss of generality we can assume the following: (For an explicit definition see [1, p. 185])

1. $\mathbf{x}_0 = \mathbf{0}$ (By setting $\mathbf{x}_{\text{new}} = \mathbf{x} - \mathbf{x}_0$).
2. $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ (by setting $\mathbf{f}_{\text{new}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_{\text{new}} + \mathbf{x}_0) - \mathbf{f}(\mathbf{x}_0)$).[2]
3. $D\mathbf{f}(\mathbf{0}) = I$ $\Big($ By setting $\mathbf{f}_{\text{new}}(\mathbf{x}) = (D\mathbf{f}(\mathbf{0}))^{-1}\mathbf{f}(\mathbf{x})\Big)$. [3]

Let $\psi(\mathbf{x}) = \mathbf{x} - \mathbf{f}(\mathbf{x})$. Clearly $\psi \in \mathscr{C}^1$. Observe that, $D\psi(\mathbf{0}) = \text{O}$. Since $\psi \in \mathscr{C}^1$ and $D\psi(\mathbf{0}) = \text{O}$, implies there exists an $r > 0$ such that,

$$\|\mathbf{x}\| \leq r \implies \|D\psi(\mathbf{x})\| \leq \frac{1}{2}. \tag{3}$$

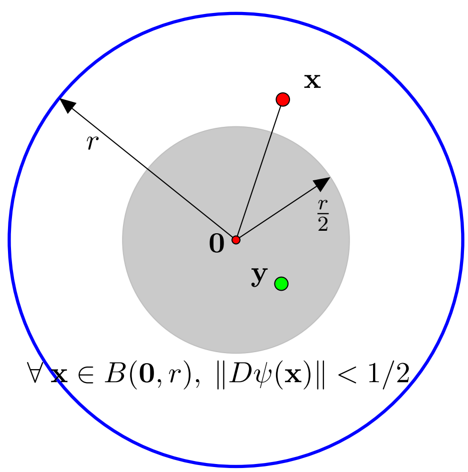

Figure 2

Using Theorem 2,

$$\|\mathbf{x}\| \leq r \implies \|\psi(\mathbf{x})\| \leq \max_{\mathbf{c}\in[\mathbf{0},\mathbf{x}]} \|D\psi(\mathbf{c})\| \, \|x\| \leq \frac{r}{2} \tag{4}$$

Fix $\mathbf{y} \in \mathbb{R}^n$ with $\|\mathbf{y}\| < \frac{r}{2}$ and define

$$\phi_{\mathbf{y}}(\mathbf{x}) = \underbrace{\mathbf{x} - \mathbf{f}(\mathbf{x})}_{\psi(\mathbf{x})} + \mathbf{y} \tag{5}$$

[2] Observe that $\mathbf{x}_{\text{new}} + \mathbf{x}_0 \in U$.

[3] Here $\mathbf{f}$ is the new function that we have defined in step 2.

Using Equation (4),

$$\|\mathbf{x}\| \le r \implies \|\phi_{\mathbf{y}}(\mathbf{x})\| = \|\mathbf{x} - \mathbf{f}(\mathbf{x}) + \mathbf{y}\| \le \|\mathbf{x} - \mathbf{f}(\mathbf{x})\| + \|\mathbf{y}\| < r.$$

So, due to continuity of $\phi_{\mathbf{y}}$, we can say that $\phi_{\mathbf{y}} : \mathbb{B}[\mathbf{0}, r] \to \mathbb{B}[\mathbf{0}, r]$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{B}[\mathbf{0}, r]$,

$$\|\phi_{\mathbf{y}}(\mathbf{x}) - \phi_{\mathbf{y}}(\mathbf{y})\| = \|\psi(\mathbf{x}) - \psi(\mathbf{y})\| \le \frac{1}{2}\|\mathbf{x} - \mathbf{y}\| \qquad \text{(follows from Theorem 2)}$$

Now using Theorem 3, $\phi_{\mathbf{y}}$ is a contraction map. So, $\phi_{\mathbf{y}}$ has a unique fixed point, say $\mathbf{x}_{\mathbf{y}} \in \mathbb{B}[\mathbf{0}, r]$. So, there exists a unique $\mathbf{x}_{\mathbf{y}} \in \mathbb{B}[\mathbf{0}, r]$ such that,

$$\mathbf{f}(\mathbf{x}_{\mathbf{y}}) = \mathbf{y}.$$

Since $\|\mathbf{y}\| < \frac{r}{2} \implies \mathbf{x}_{\mathbf{y}} \in \mathbb{B}(\mathbf{0}, r)$. Now define,

$$W = \mathbb{B}\left(\mathbf{0}, \frac{r}{2}\right), \ V = \mathbf{f}^{-1}(W) \cap \mathbb{B}(\mathbf{0}, r) \text{ and } \mathbf{g} : W \to V \text{ as } \mathbf{g}(\mathbf{y}) = \mathbf{x}_{\mathbf{y}}. \text{ (See Figure 3)}$$

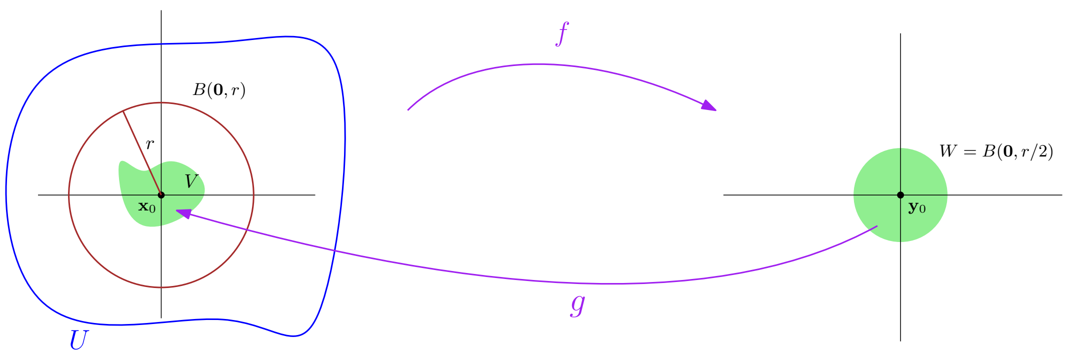

Figure 3

Clearly, $(\mathbf{f} \circ \mathbf{g})(\mathbf{y}) = \mathbf{y}; (\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{x}, \forall \mathbf{x} \in V$ and $\forall \mathbf{y} \in W$. Now we will prove that $\mathbf{g} \in \mathscr{C}^1$.

g is continuous : Let $\mathbf{y}, \mathbf{z} \in W$. Let $\mathbf{g}(\mathbf{y}) = \mathbf{u}$ and $\mathbf{g}(\mathbf{z}) = \mathbf{v}$. Consider,

$$\begin{aligned}
&\|\psi(\mathbf{u}) - \psi(\mathbf{v})\| \le \frac{1}{2}\|\mathbf{u} - \mathbf{v}\| \\
\implies &\|(\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v})) - (\mathbf{u} - \mathbf{v})\| \le \frac{1}{2}\|\mathbf{u} - \mathbf{v}\| \\
\implies &\|\mathbf{u} - \mathbf{v}\| \le 2\|\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v})\| \\
\implies &\|\mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{z})\| \le 2\|\mathbf{y} - \mathbf{z}\| < \varepsilon
\end{aligned}$$

g is differentiable : Let $\mathbf{y} \in W$ and $\mathbf{g}(\mathbf{y}) = \mathbf{x}$. We will prove that $D\mathbf{g}(\mathbf{y}) = (D\mathbf{f}(\mathbf{x}))^{-1}$. Let $D\mathbf{f}(\mathbf{x}) = A$. Since W is open, $\exists \mathbf{k}$ such that $\mathbf{y} + \mathbf{k} \in W$. Set $\mathbf{g}(\mathbf{y}+\mathbf{k}) = \mathbf{x}+\mathbf{h} \implies \mathbf{h} = \mathbf{g}(\mathbf{y}+\mathbf{k}) - \mathbf{g}(\mathbf{y})$.

We want

$$\begin{aligned}
\frac{\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y})-A^{-1}\mathbf{k}}{\|\mathbf{k}\|} &\to \mathbf{0} \text{ as } \mathbf{k}\to\mathbf{0}. \\
\iff \frac{A(\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y}))-\mathbf{k}}{\|\mathbf{k}\|} &\to \mathbf{0} \text{ as } \mathbf{k}\to\mathbf{0}. \\
\iff \frac{A\mathbf{h}-\mathbf{k}}{\|\mathbf{k}\|} = \frac{A\mathbf{h}-(\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x}))}{\|\mathbf{k}\|} &\to \mathbf{0} \text{ as } \mathbf{k}\to\mathbf{0}. \\
\iff -\frac{(\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x}))-A\mathbf{h}}{\|\mathbf{h}\|}\frac{\|\mathbf{h}\|}{\|\mathbf{k}\|} &\to \mathbf{0} \text{ as } \mathbf{k}\to\mathbf{0}. \qquad (6)
\end{aligned}$$

Since, $\|\mathbf{h}\| \leq 2\|\mathbf{k}\|,\ \mathbf{h}\to\mathbf{0}$ if $\mathbf{k}\to\mathbf{0}$. So, the limit in (6) is $\mathbf{0}$ as $\mathbf{h}\to\mathbf{0}$. Hence, $\mathbf{g}$ is differentiable. And,

$$D\mathbf{g}(\mathbf{y}) = (D\mathbf{f}(\mathbf{g}(\mathbf{y})))^{-1} \qquad (7)$$

Now, since $D\mathbf{g}$ is a composition of the function $\mathbf{y} \rightsquigarrow D\mathbf{f}(\mathbf{g}(\mathbf{y}))$ and $A \rightsquigarrow A^{-1}$ on the space of invertible matrices. Since $\mathbf{g}$ is continuous and $\mathbf{f}\in\mathscr{C}^1,\ D\mathbf{f}$ is continuous. As $D\mathbf{g} = i\circ D\mathbf{f}\circ g$, where i denotes the inverse of a matrix and the composition of continuous functions is continuous hence $\mathbf{y}\rightsquigarrow D\mathbf{g}(\mathbf{y})$ is continuous.

Remark: In the above theorem, if $\mathbf{f}\in\mathscr{C}^k$ then $\mathbf{g}\in\mathscr{C}^k$.

Corollary 1. If $\mathbf{f}:\mathbb{R}^n\to\mathbb{R}^n$ is $\mathscr{C}^1$ and has the property that for every $\mathbf{a}\in\mathbb{R}^n,\ D\mathbf{f}(\mathbf{a})$ is non-singular. Then $\mathbf{f}$ is ana open map.

Proof. Let $U\subset\mathbb{R}^n$ be any open set. We will show that $\mathbf{f}(U)$ is open. Let $\mathbf{y}\in\mathbf{f}(U) \implies \exists\, \mathbf{a}\in U$ such that $\mathbf{f}(\mathbf{a})=\mathbf{y}$. According to the hypothesis, $D\mathbf{f}(\mathbf{a})$ is non-singular, so using Theorem 1, there exist neighborhoods $V\subset U, W$ of $\mathbf{a}$ and $\mathbf{f}(\mathbf{a})$ respectively and a $\mathscr{C}^1$ function $g: W\to V$ such that $\mathbf{f}\circ\mathbf{g} = id_W$. So, $W = f(g(W))$ which is open and $\mathbf{x}\in W \implies \mathbf{g}(\mathbf{x})\in V \implies \mathbf{x} = \mathbf{f}(\mathbf{g}(x))\in U \implies W\subset f(U)$. Hence, $\mathbf{f}$ is an open map. □

Holomorphic Inverse Function Theorem in one Complex Variable

Theorem 4. Let $f:\mathbb{C}\to\mathbb{C}$ be a holomorphic map such that $f(0)=0$ and $Df(0)$ is non-singular. Then there exists open sets U,V around the origin and $g: V\to U$ such that $f\circ g = id_V,\ g\circ f = id_U$ and g is holomorphic.

Proof. Let $f(z) = u(x,y)+iv(x,y)$ where $u,v:\mathbb{R}^2\to\mathbb{R}$. So we can write the function $f:\mathbb{R}^2\to\mathbb{R}^2$ given by $f(x,y) = (u(x,y), v(x,y))$. Given that,

- $f(0,0) = (0,0)$
- f is holomorphic implies $f\in\mathscr{C}^\infty$.
- $Df(0,0) = \begin{pmatrix} u_x(0,0) & -v_x(0,0) \\ v_x(0,0) & u_x(0,0)\end{pmatrix}$ and $\det Df(0,0) = u_x(0,0)^2 + v_x(0,0)^2 \neq 0$.

Hence from the real inverse function theorem there exists neighborhoods U, V around the origin and $g : V \to U$ such that $f \circ g = id_V$ and $g \circ f = id_U$. From inverse function theorem g will be $\mathscr{C}^\infty$. Now we will show that g is holomorphic function. Let $f'(0) = \alpha$. Consider

$$Dg(0,0) = \begin{pmatrix} u_x(0,0) & -v_x(0,0) \\ v_x(0,0) & u_x(0,0) \end{pmatrix}^{-1} = \frac{1}{|\alpha|^2} \begin{pmatrix} \text{Re}(\alpha) & \text{Im}(\alpha) \\ -\text{Im}(\alpha) & \text{Re}(\alpha) \end{pmatrix} = \begin{pmatrix} \text{Re}(\alpha^{-1}) & -\text{Im}(\alpha^{-1}) \\ \text{Im}(\alpha^{-1}) & \text{Re}(\alpha^{-1}) \end{pmatrix}.$$

As, $Dg(0,0)$ is $\mathbb{C}-$linear, g is holomorphic $\square$

Implicit Function Theorem

Theorem 5. Suppose $U \subset \mathbb{R}^{n+m}$ is open and $\mathbf{F} : U \to \mathbb{R}^m$ is $\mathscr{C}^1$. Writing a vector in $\mathbb{R}^{n+m}$ as $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$, with $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, suppose that $\mathbf{F}\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} = \mathbf{0}$ and the $m \times m$ matrix $\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix}$ is non-singular. Then there are neighborhoods V of $\mathbf{x}_0$ and W of $\mathbf{y}_0$ and a $\mathscr{C}^1$ function $\phi : V \to W$ so that

$$\mathbf{F}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{0}, \mathbf{x} \in V \text{ and } \mathbf{y} \in W \iff \mathbf{y} = \phi(\mathbf{x}).$$

Moreover,

$$D\phi(\mathbf{x}) = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix}\right)^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix}.$$

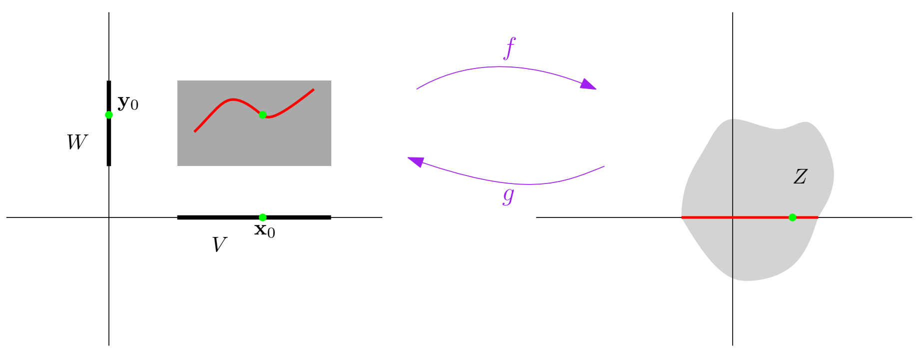

Figure 4

Proof. Define $\mathbf{f} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ by

$$\mathbf{f}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{bmatrix} \mathbf{x} \\ \mathbf{F}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \end{bmatrix}.$$

Since the linear map

$$D\mathbf{f}\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} = \left[\begin{array}{c:c} I & \text{O} \\ \hdashline \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} & \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} \end{array}\right]$$

is invertible. Using Theorem 1 there are neighborhoods $V \subset \mathbb{R}^n$ of $\mathbf{x}_0$, $W \subset \mathbb{R}^m$ of $\mathbf{y}_0$, and $Z \subset \mathbb{R}^{n+m}$ of $\mathbf{0}$ and a $\mathscr{C}^1$ function $\mathbf{g} : Z \to V \times W$ so that $\mathbf{g}$ is the inverse of $\mathbf{f}$ on $V \times W$ (See Figure 4). Now we define $\phi : V \to W$ by

$$\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} = \mathbf{g} \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix};$$

Since $\mathbf{g}$ is $\mathscr{C}^1$ so is $\phi(\mathbf{x})$. Now,

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{F}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} \end{bmatrix} = \mathbf{f}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} = \mathbf{f}\left(\mathbf{g}\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}\right) = \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \implies \mathbf{F}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} = \mathbf{0}.$$

On the other hand, if $\mathbf{F}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{0},\ x \in V$ and $\mathbf{y} \in W$, then $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{g}\left(\mathbf{f}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}\right) = \mathbf{g}\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{F}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \end{bmatrix}\right) =$ $\mathbf{g}\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \implies \mathbf{y} = \phi(\mathbf{x})$.

Now we will calculate the derivative of ϕ. Since, $\mathbf{F}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} = \mathbf{0}$, define $\mathbf{h} : V \to \mathbb{R}^m$ by $\mathbf{h}(\mathbf{x}) = \mathbf{F}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix}$. Then $\mathbf{h}$ is $\mathscr{C}^1$ and

$$0 = D\mathbf{h}(\mathbf{x}) = \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} + \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix} D\phi(\mathbf{x}) \implies D\phi(\mathbf{x}) = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix}\right)^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{x}}\begin{pmatrix} \mathbf{x} \\ \phi(\mathbf{x}) \end{pmatrix}.$$

□

While proving the inverse function theorem we observe that the theorem only tells about the existence of the inverse function, but the natural question arises is how will we find out the inverse. So we will discuss the situation in one dimension.

An approximation of inverse function in one dimension

We saw that the inverse function theorem (Theorem 1) only guarantees the existence of the inverse but it did not say anything about "How to compute the inverse?". As we have used the contraction mapping theorem (Theorem 3) which is basically a limit of the sequence

$$\mathbf{x}_{k+1} = \phi_{\mathbf{y}}(\mathbf{x}_k), k \geq 1,\ \mathbf{x}_0 = \mathbf{0}.$$

We will see, the local inverse mapping $\mathbf{g} : W \to V$ is the limit of the sequence of successive approximations $\{\mathbf{g}_k\}_0^\infty$ defined inductively on V by

$$\mathbf{g}_0(\mathbf{y}) = \mathbf{0},\ \mathbf{g}_{k+1}(\mathbf{y}) = \mathbf{g}_k(\mathbf{y}) - \mathbf{f}(\mathbf{g}(\mathbf{y})) + \mathbf{y} \tag{6}$$

for all $\mathbf{y} \in W$. Let us formalize it in one dimension.
Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function which is real analytic at 0 and $f'(0) = 1$. We also assume that $f(0) = 0$. We construct a local inverse g of f by defining $g(y)$ to be the fixed point point of the sequence given in Equation (6). f is real analytic at 0 and using the given conditions we can

write $f(x) = x + a_2x^2 + a_3x^3 + \dots, |x| < \varepsilon$, for some $\varepsilon > 0$. Now computing the sequence given in Equation (6) around 0, gives

$$\begin{aligned} g_1(y) &= y \\ g_2(y) &= y - a_2y^2 \\ g_3(y) &= y - a_2y^2 + (2a_2^2 - a_3)y^3 + o(y^4) \\ g_4(y) &= y - a_2y^2 + (2a_2^2 - a_3)y^3 + (5a_2a_3 - a_4 - 5a_2^3)y^4 + o(y^5) \end{aligned} \tag{7}$$

Theorem 6. If the input function f for inverse function theorem is real analytic then the local inverse g is also real analytic.

Proof. See [5, p 47] $\square$

Now from Theorem 6, there exists a function g which is real analytic at $f(0) = 0$. So we can write $g(y) = b_0 + b_1y + b_2y^2 + b_3y^3 + \dots$. Again, $g(0) = 0 \implies b_0 = 0$. Since g is the inverse of f, so for every y in the neighborhood of 0, $f(g(y)) = y$. Using this we obtain,

$$\begin{aligned} b_1 &= 1 \\ b_2 &= -a_2 \\ b_3 &= 2a_2^2 - a_3 \\ b_4 &= 5a_2a_3 - a_4 - 5a_2^3 \end{aligned} \tag{8}$$

Comparing eq. (7) and eq. (8) we observe that our approximation by sequence is correct up to n terms.

Application of Inverse Function Theorem in Manifold

In this section, we will give one of the major application of inverse function theorem which will be useful for proving something is a submanifold. Before going ahead let us define some terminology.

Definition 2 (Submanifold)**.** We say $\mathscr{M} \subset \mathbb{R}^n$ is a *k-dimensional submanifold* if for any $\mathbf{a} \in M$, there is a neighborhood $W \subset \mathbb{R}^n$ of $\mathbf{a}$ such that $\mathscr{M} \cap W$ is diffeomorphic[4] to $\mathbb{R}^k$.

Definition 3 (Regular Value)**.** Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ be smooth map. Then a point $\mathbf{p} \in \mathbb{R}^m$ is called a *regular value of f* if for every $\mathbf{a} \in \mathbf{f}^{-1}\{\mathbf{p}\}$, the map

$$D\mathbf{f}(\mathbf{a}) : T_{\mathbf{a}}\mathbb{R}^n \to T_{\mathbf{p}}\mathbb{R}^m$$

is surjective.

Lemma 1. Let $\mathbf{f} : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a smooth map with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. Assume that $\mathbf{0}$ is a regular value of $\mathbf{f}$. Then $f^{-1}\{\mathbf{0}\}$ is a submanifold of dimension n.

Proof. Let $\mathbf{p} = \begin{pmatrix} \mathbf{p}_x \\ \mathbf{p}_y \end{pmatrix} \in f^{-1}\{\mathbf{0}\}$ be given, where $\mathbf{p}_x \in \mathbb{R}^n$ and $\mathbf{p}_y \in \mathbb{R}^m$. So $\mathbf{f}(\mathbf{p}) = \mathbf{0}$ and $\mathbf{0}$ is a regular value of $\mathbf{f}$, $D\mathbf{f}(\mathbf{p})$ is a surjective map and hence by multiplying with permutation matrices we can assume that $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{p})$ is invertible. So using Implicit function theorem (theorem 5), there exist neighborhoods V, W and Z of $\mathbf{p}_x, \mathbf{p}_y$ and $\mathbf{0}$ respectively and a $\mathscr{C}^1$ function $\phi : V \to W$

[4] Let $f : A \to B$ be a homeomorphism. Then it is called a *diffeomorphism* if f and f^{-1} are differentiable.

such that for any $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in f^{-1}\{\mathbf{0}\}, \; \mathbf{y} = \phi(\mathbf{x})$ (see Figure 4). Define $U = f^{-1}\{\mathbf{0}\} \bigcap (V \times W)$ and $O = Z \cap \mathbb{R}^n \times \{\mathbf{0}\}$ and $\psi : U \to O$ by $(\mathbf{x}, \phi(\mathbf{x})) \mapsto (\mathbf{x}, \mathbf{0})$ which is a diffeomorphism. (Observe it is just a projection). $\square$

Theorem 7. Let $GL(n, \mathbb{R})$ be the set of all invertible matrices. Then $SL(n, \mathbb{R}) := \{A \in GL(n, \mathbb{R}) : \det A = 1\}$ is a submanifold of dimension $n^2 - 1$.

Proof. Define the map

$$sl : \mathbb{R}^{n \times n} \cong M(n, \mathbb{R}) \to \mathbb{R}$$

by $sl(A) = \det A$, which is a smooth map because $\det A$ is a polynomial in n^2 variables. Since $SL(n, \mathbb{R}) = sl^{-1}\{1\}$ so it is enough to prove that 1 is a regular value. We will prove that for given $A \in sl^{-1}\{1\}, \; Dsl(A) : T_A\mathbb{R}^{n \times n} \to T_1\mathbb{R}$ is a surjective map. Since the image is contained in $\mathbb{R}$ so enough to prove that the directional derivative of sl is non zero at A in some direction. Consider the directional derivative in the direction of A

$$\lim_{t \to 0} \frac{sl(A + At) - sl(A)}{t} = \lim_{t \to 0} \frac{\det (A + At) - 1}{t} = n$$

$\square$

References

References

[1] Edwards, C.H.: *Advanced Calculus of Several Variables,*, Academic Press, New York and London 1973.

[2] Rudin, W.: *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill Book Company, New York, 1976.

[3] Shifrin, T: *Multivariable Mathematics: Linear Algebra, Multivariable Calculus and Manifolds,*, John Wiley and Sons Publishing, 2005.

[4] Spivak, Michael: *Calculus on Manifolds: A Modern Approach To Classical Theorems of Advanced Calculus,*, Mathematics monograph series, Avalon Publishing, 1971.

[5] Krantz, Steven G., Parks, Harold R.: *A Primer of Real Analytic Functions*, Birkhäuser Boston, 2002.