Department of Mathematics
Guru Nanak Dev University, Amritsar-143005
Lecture notes[1] on 'The inverse function theorem'

1 The inverse function theorem

Recall the calculus of functions of one variable where a differentiable function $f : (a, b) \to \mathbb{R}$. If for some $\alpha \in (a, b)$, $f'(\alpha) \neq 0$. Then either $f'(\alpha) > 0$ or $f'(\alpha) < 0$. W.l.o.g. let $f'(\alpha) < 0$; then f is decreasing at α which means, there is an open interval $(\alpha-\epsilon, \alpha+\epsilon)$, $\epsilon > 0$ such that $f'(x) < 0$ for all $x \in (\alpha-\epsilon, \alpha+\epsilon)$. Then the restriction $g : (\alpha - \epsilon, \alpha + \epsilon) \to g(\alpha - \epsilon, \alpha + \epsilon)$ such that $g(x) = f(x)$ for all $x \in (\alpha - \epsilon, \alpha + \epsilon)$ is bijective (why? use mean value theorem to conclude). Question arises, is the inverse function g^{-1} differentiable over its domain? Answer to this is yes and we have

$$(g^{-1}(y))' = \frac{1}{g'(g^{-1}(y))}, \text{ for all } y \in (g(\alpha - \epsilon, g(\alpha + \epsilon)))$$

which is well defined since $g'(g^{-1}(y)) \neq 0$ for all $g^{-1}(y) \in (\alpha - \epsilon, \alpha + \epsilon)$. We record this in the following.

Theorem 1.1. (The inverse function theorem: one variable case) *Let $f : (a, b) \to \mathbb{R}$ be continuously differentiable function, such that $f'(\alpha) \neq 0$ for some $\alpha \in (a, b)$. Then there exist open intervals $U \subset (a, b)$ containing α and V containing $f(\alpha)$ such that the restriction $g : U \to V$ defined by $g(x) = f(x)$ for all $x \in U$, is bijective. Moreover, the inverse function $g^{-1} : V \to U$ is also differentiable and is given by*

$$(g^{-1}(y))' = \frac{1}{g'(g^{-1}(y))}, \text{ for all } y \in V.$$

To our amusement, this result is true in several variables as well and is known as *the inverse function theorem.* First we prove the following lemma which will be used in the proof of the inverse function theorem.

Lemma 1.2. *Let R be an open rectangle in $\mathbb{R}^n$ and $f : \bar{R} \to \mathbb{R}^m$ be continuously differentiable. If $|D_j f_i(x)| \leq M$ for some $M > 0$ for all $1 \leq i \leq m, 1 \leq j \leq n$, then*

$$\|f(x) - f(y)\| \leq Mmn\|x - y\|, \text{ for all } x, y \in \bar{R}$$

Proof. Let $f = (f_1, ..., f_m)$. Using the mean value theorem for each f_i, there is a point $c_i \in R$ such that the following hold $f_i(x) - f_i(y) = \nabla f_i(c_i) \cdot (x - y) = \sum_{j=1}^{n} D_j f_i(c_i)(x_j - y_j)$ which gives

$$\begin{aligned} \|f_i(x) - f_i(y)\| &= \|\sum_{j=1}^{n} D_j f^i(c_i)(x_j - y_j)\| \leq \sum_{j=1}^{n} |D_j f^i(c_i)||(x_j - y_j)| \\ &\leq M \sum_{j=1}^{n} |(x_j - y_j)| \leq M \sum_{j=1}^{n} \|x - y\| = nM\|x - y\| \end{aligned} \tag{1.1}$$

[1] **Jitender Singh**; E-mail: sonumaths@gmail.com; Web page: https://sites.google.com/site/sonumaths2/

Now consider using (1.1)

$$\|f(x)-f(y)\| = \|(f_1(x)-f_1(y),...,f_m(x)-f_m(y))\| = (\sum_{i=1}^{m}\|f_i(x)-f_i(y)\|^2)^{1/2}$$
$$\leq (\sum_{i=1}^{m} n^2M^2\|x-y\|^2)^{1/2} = Mn\sqrt{m}\|x-y\|$$

from which the assertion follows. □

Definition: Let U be an open subset of $\mathbb{R}^n$ and V be an open subset of $\mathbb{R}^m$. A map $f: U \to V$ is said to be a diffeomorphism if (1) f is a continuous bijection (2) f^{-1} is continuous and, (3) f and f^{-1} are differentiable on U and V respectively.

Theorem 1.3. (The inverse function theorem) *Let $f: \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable such that $\det(f'(a)) \neq 0$ for some $a \in \mathbb{R}^n$. Then there exist an open st $U \ni a$ and an open set $V \ni f(a)$ in $\mathbb{R}^n$ such that the map $g: U \to V$ defined by $g(x) = f(x)$ for all $x \in U$ is a diffeomorphism. Moreover,*

$$(g^{-1}(y))' = [g'(g^{-1}(y))]^{-1} \text{ for all } y \in V.$$

Proof. Let $\lambda := Df(a)$ i.e. $\lambda(h) = f'(a)\cdot h$ for all $h \in \mathbb{R}^n$. Since $\det(f'(a)) \neq 0$ it follows that λ is invertible. Define $F: \mathbb{R}^n \to \mathbb{R}^n$ by $F = \lambda^{-1} \circ f$. Then using the chain rule $DF(x) = D(\lambda^{-1}\circ f)(x) = D(\lambda^{-1}(f(x))) \circ Df(x) = \lambda^{-1}\circ\lambda = I$ which is the identity map of $\mathbb{R}^n$. Note that if the theorem holds for F i.e. there exist open sets $U \ni a$ and $V' \ni F(a)$ of $\mathbb{R}^n$, such that $F|_U : U \to V'$ is a diffeomorphism, then the map $f|_U : U \to V = \lambda(V')$ is the required diffeomorphism. So it is enough to prove the theorem for the function F.

step 1: We will prove that there is an open set $U \ni a$ such that $F(a) \neq F(x)$ for all $x \in U$. To see this, if possible, let $f(x) = f(a)$ for some $x \in U$ and each open set U in $\mathbb{R}^n$. By differentiability of F, there is an open set $U' \ni a$ such that the error function $\epsilon : U' \to \mathbb{R}$ defined by

$$\epsilon(x-a) := \begin{cases} \dfrac{\|F(x)-F(a)-I(x-a)\|}{\|x-a\|} & \text{if } x \neq a \\ 0 & \text{if } x = a \end{cases}$$

satisfies $\epsilon(x-a) \to 0$ as $x \to a$. If $F(z) = F(a)$ for some $z \in U$ such that $z \neq a$, then $\epsilon(z-a) = \frac{\|I(z-a)\|}{\|z-a\|} = 1$. Take the sequence of open balls $B_d(a, \frac{1}{n})$, $n \geq N$ for some positive integer N such that $B_d(a,\frac{1}{n}) \subset U'$ for all $n \geq N$. Then by hypothesis on each open ball $B_d(a,\frac{1}{n})$ there is a $z_n \neq a$ such that $F(z_n) = F(a)$ and $\epsilon(z_n - a) = 1$. As $z_n \to a$ and ϵ is continuous at 0, it follows (using the sequence lemma!) that $1 = \epsilon(z_n - a) \to \epsilon(a) = 0$ which is a contradiction.

Choose an open rectangle S containing a such that $\bar{S} \subset U$. Since F is continuously differentiable on $\bar{S}$, the map $\mathcal{D}F : \bar{S} \to M(n,n;\mathbb{R}^n)$ sending $x \mapsto F'(x)$ is continuous such that $\det(F'(a)) = 1 \neq 0$. Also the determinant function $\det : M(n,n;\mathbb{R}^n) \to \mathbb{S}$ being compositions of $\pm$ and the usual product of reals, is continuous. Therefore the composition $\det\circ\mathcal{D}F : \bar{S} \to \mathbb{R}$ is continuous at a. There fore for each $0 < \epsilon < 1$ such that $\det(F'(x)) \in (1-\epsilon, 1+\epsilon)$ there is an open ball $B_d(a,\delta) \subset \bar{S}$ such that

$\det \circ F'(B_d(a,\delta)) \subseteq (1-\epsilon, 1+\epsilon)$. Thus $\det(F'(x)) > 0$ for all $x \in B_d(a,\delta)$. Now choose an open rectangle $R \ni a$ such that $\bar{R} \subset B_d(a,\delta)$. Then

$$\det(F'(x)) \neq 0 \text{ for all } x \in \bar{R}.$$

Since the partial derivatives of D_jF_i are continuous on $\bar{R}$, we have

$$|D_jF_i(x) - D_jF_i(a)| \leq \frac{1}{2n^2}, \text{ for all } i,j; x \in \bar{R}. \tag{1.2}$$

step 2: We will show that $F : \bar{R} \to F(\bar{R})$ is one one and onto. For this, we consider the function

$$g(x) = F(x) - x$$

which satisfies $|D_jg_i(x)| = |D_jF_i(x) - D_jx_i| = |D_jF_i(x) - D_jF_i(a)| \leq \frac{1}{2n^2}$ (by (1.2)). This means that the partial derivatives of g are bounded on R by $M := \frac{1}{2n^2}$. Therefore by lemma 1.2 for all $x_1, x_2 \in \bar{R}$

$$\|g(x_1) - g(x_2)\| \leq \frac{1}{2n^2}n^2\|x_1 - x_2\| \Rightarrow \|F(x_1) - F(x_2) - (x_1 - x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|$$
$$\Rightarrow \|x_1 - x_2\| - \|F(x_1) - F(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|$$

which yields

$$\|x_1 - x_2\| \leq 2\|F(x_1) - F(x_2)\| \tag{1.3}$$

Now if $F(x_1) = F(x_2)$ then by (1.3) we get $\|x_1 - x_2\| \leq 0$ which gives $\|x_1 - x_2\| = 0 \Leftrightarrow x_1 - x_2 = 0$ which proves injectivity of F.

To establish that F is onto, we first note that as ∂R is compact, and F is continuous, it follows that $F(\partial R)$ is compact and $F(a) \notin F(\partial R)$. As $a \in (\mathbb{R}^n - \partial R)$, $F(a) \in \mathbb{R}^n - F(\partial R)$ which is open. Choose an open ball of radius d centered at $F(a)$ which is completely contained in $\mathbb{R}^n - F(\partial R)$. Then $\|F(x) - F(a)\| \geq d$ for all $x \in \partial R$. *Let W be the open ball of radius $\frac{d}{2}$ and centered at $f(a)$ in $\mathbb{R}^n - F(\partial R)$*. Then for all $y \in W$, $\|y - F(a)\| < \frac{d}{2}$. It follows that

$$\|y - F(a)\| < \frac{d}{2} \leq \|y - F(x)\|, \text{ for all } y \in W, x \in \partial R. \tag{1.4}$$

Using (1.4) we will show that $F : R \to W$ is onto. Let $y \in W$. Define $G : \bar{R} \to \mathbb{R}$ by

$$G(x) = \|y - F(x)\|^2 = \sum_{i=1}^{n}(y_i - F_i(x))^2.$$

Using the extreme value theorem, G has a minimum on $\bar{R}$. If $x \in \partial R$ then using (1.4) $\|y - F(a)\| < \|y - F(x)\|$ which gives $G(a) < G(x)$; therefore the point $x \in \partial R$ can not be a point of minimum for G. Therefore minimum of G is attained in R. Since G is a differentiable function on R, at the point of minima $D_jG = 0$ for all $j = 1, ..., n$. This gives

$$2\sum_{i=1}^{n}(y_i - F_i(x))D_jF_i(x) = 0, \ \forall\ j = 1, 2..., n \text{ or } F'(x) \cdot (y - F(x)) = 0$$

which is a homogeneous linear system. As $\det(F'(x)) \neq 0$ on R, it follows that the only solution is the trivial solution, i.e. $y = F(x)$ and that proves surjectivity of F.

step 3: If we define $V = R \cap F^{-1}(W)$, then by the step 2, the function $F : V \to W$ has the inverse $F^{-1} : W \to V$. We will show that F^{-1} is continuous. This follows since for all $y_1, y_2 \in W$ from (1.3) we have

$$\|F^{-1}(y_1) - F^{-1}(y_2)\| \leq 2\|y_1 - y_2\| \to 0 \text{ as } y_1 \to y_2.$$

step 4: Finally, we will show that $F^{-1} : W \to V$ is differentiable. Let $\mu = DF(x)$; we will prove that F^{-1} is differentiable at $F(x)$ with derivative μ^{-1}. For any $y, y' \in W$ s.t. $y = F(x),\ y' = F(x')$, consider

$$\begin{aligned}
\frac{\|F^{-1}(y) - F^{-1}(y') - \mu^{-1}(y - y')\|}{\|y - y'\|} &= \frac{\|x - x' - \mu^{-1}(F(x) - F(x'))\|}{\|F(x) - F(x')\|} \\
&= \frac{\|x - x'\|}{\|F(x) - F(x')\|}\|\mu^{-1}\left(\frac{F(x) - F(x') - \mu(x - x')}{\|x - x'\|}\right)\| \\
&= \frac{\|F^{-1}(y) - F^{-1}(y')\|}{\|y - y'\|}\|\mu^{-1}\left(\frac{F(x) - F(x') - \mu(x - x')}{\|x - x'\|}\right)\| \\
&\leq 2\mu^{-1}\left(\frac{F(x) - F(x') - \mu(x - x')}{\|x - x'\|}\right)\|\{\text{from step 3}\} \\
&\to 2\mu^{-1}(0) = 0 \text{ as } y \to y'
\end{aligned}$$

since $y \to y'$ implies $x \to x'$ by continuity of F^{-1}. We have also used the continuity of μ^{-1} and differentiability of F.

step 5: To verify the formula to compute the Jacobian of the inverse function F^{-1}, we proceed as follows. As we have $(F \circ F^{-1})(y) = y$ for all $y \in W$. As $F = \lambda^{-1} \circ f$. This gives $f : V \to \lambda(W)$ such that $f = \lambda \circ F$ which shows that f is invertible with differentiable inverse on V. Using chain rule for all $y \in \lambda(W)$ we have

$$\begin{aligned}
& f \circ f^{-1} = Id \\
\Rightarrow\ & D(f \circ f^{-1})(y) = Id \\
\Rightarrow\ & Df(f^{-1}(y)) \circ Df^{-1}(y) = Id
\end{aligned}$$

from which the result follows. □

Remark: It should be noted that inverse function f^{-1} may exist even if $f'(a) = 0$. However in such a case f^{-1} is not differentiable at $f(a)$. A typical example is the function $f : \mathbb{R} \to \mathbb{R}$ s.t. $f(x) = x^3$ where $\det(f'(0)) = 0$.

Example: Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^n$ a continuously differentiable injective function such that $\det(f'(x)) \neq 0$ for all x. Show that $f(U)$ is open and that the map $f^{-1} : f(U) \to U$ is differentiable. Show that $f(V)$ is open for all open $V \subset U$.

Proof. Using the inverse function theorem for each $x \in U$ there is an open set $U_x \subset U$ containing x and an open set $V_x \ni f(x)$ such that the restriction $f : U_x \to V_x$ is a diffeomorphism and that $f(U_x) = V_x$ is open. Therefore $f(\cup_{x \in U} U_x) = \cup_{f(x) \in f(U)} V_x$ is open which shows that $f(U)$ is open. Infact replacing U be any open subset A of $\mathbb{R}^n$ contained in U, via similar argument, we see that $f(A)$ is open. This proves that $f^{-1} : f(U) \to U$ is also continuous and that f is a homeomorphism.

As $f^{-1} : V_x \to U_x$ is differentiable for each $f(x) \in V_x$ and that $\cup_{f(x)\in V_x} = f(U)$, it follows that f^{-1} is differentiable at each $f(x) \in f(U)$. This proves that $f : U \to f(U)$ is a diffeomorphism. $\square$

Example: Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function. Then f is not one one.
If $D_1 f(x, y) = 0 = D_2 f(x, y)$ then in every neighborhood of (x, y), f is not injective and we are done. So let $D_1 f(x, y) \neq 0$ in some neighborhood U of (x, y). Consider $g : U \to \mathbb{R}^2$ by $g(x, y) = (f(x, y), y)$. Then g is continuously differentiable with $g'(x, y) = \begin{pmatrix} D_1 f & D_2 f \\ 0 & 1 \end{pmatrix}$ such that $\det(g'(x, y)) \neq 0$. Using the inverse function theorem, there exist open sets $W \ni (f(x, y), y)$ and $V_1 \times V_2 \ni (x, y)$ such that $g : V_1 \times V_2 \to W$ is a diffeomorphism with inverse $g^{-1} : W \to V_1 \times V_2$ of the form $g^{-1}(x, y) = (\eta(x, y), y)$. For some smooth function $\eta : W \to V_1$. Then we have for all $(x, y) \in W$

$$f(\eta(x, y), y) = (\pi_1 \circ g)g^{-1}(x, y) = x.$$

This shows that for any two distinct points y_1 and y_2 in V_2, we have

$$f(\eta(x, y_1), y_1) = f(\eta(x, y_2), y_2)$$

i.e. f is not injective.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ satisfies $f'(x) \neq 0$ for all x. Then f is one one. For any $x_1, x_2 \in \mathbb{R}$, using the mean value theorem $f(x_1) - f(x_2) = (x_1 - x_2)f'(c)$ for some $c \in (x_1, x_2)$. If $f(x_1) = f(x_2)$ then we immediately obtain $x_1 = x_2$ since $f'(c) \neq 0$. This proves that f is one one. Note that, here, the continuity of derivative is not required!

1.1 Another proof of inverse function theorem

The inverse function theorem is a consequence of the contraction mapping theorem. We provide a second proof of the inverse function theorem which is based on the contraction mapping theorem. This proof is comparatively straight and easy to understand. We begin with the following.

Definition: (Contraction mapping) Let (X, d) be a metric space. A map $\varphi : X \to X$ is called a contraction if there is a $\epsilon \in (0, 1)$ such that

$$d(\varphi(x), \varphi(y)) \leq \epsilon \, d(x, y), \text{ for all } x, y \in X.$$

Recall that a complete metric space is the one in which every Cauchy sequence converges. For example, $\mathbb{R}^n$ is a complete metric space with respect to the Euclidean metric.

Theorem 1.4. (Contraction mapping theorem) *Let (X, d) be a complete metric space. Then every contraction $\varphi : X \to X$ has a unique fixed point in X i.e. there is a unique $y \in X$ such that $\varphi(y) = y$.*

Idea of the proof is to construct a suitable Cauchy sequence using the given contraction mapping.

Proof. Start with any point $x_0 \in X$ and define the sequence x_n as follows. $x_i = \varphi(x_{i-1})$ for all $i = 1, 2, ...$ Observe that $x_2 = \varphi(x_1) = (\varphi \circ \varphi)(x_0) = \varphi^2(x_0)$. By induction, it follows that $x_n = \varphi^n(x_0)$ for all

$n, m = 1, 2, ...$ Suppose that $m \leq n$. First note that $d(x_{m+1}, x_m) = d(\varphi(x_m), \varphi x_{m-1}) \leq \epsilon d(x_m, x_{m-1}) ... \leq \epsilon^m d(x_1, x_0)$. Using this estimate we obtain,

$$d(\varphi^k(x_0), x_0) \leq \sum_{i=1}^{k} d(\varphi^i(x_0), \varphi^{i-1}(x_0)) \leq \sum_{i=1}^{k} \epsilon^{i-1} d(x_1, x_0) = \frac{1-\epsilon^{k+1}}{1-\epsilon} d(x_1, x_0) < \frac{d(x_1, x_0)}{1-\epsilon} \tag{1.5}$$

Now consider for any two positive integers n and m; and, $\epsilon' > 0$, the following

$$\begin{aligned} d(x_n, x_m) &= d(\varphi^n(x_0), \varphi^m(x_0)) = d(\varphi(\varphi^{n-1}(x_0)), \varphi(\varphi^{m-1}(x_0))) \\ &\leq \epsilon d(\varphi^{n-1}(x_0), \varphi^{m-1}(x_0)) \\ &\vdots \\ &\leq \epsilon^m d(\varphi^{n-m}(x_0), x_0) < \frac{d(x_1, x_0)}{1-\epsilon} \epsilon^m < \epsilon' \end{aligned} \tag{1.6}$$

for all$n \geq m \geq 1 + \dfrac{\ln(\epsilon'(1-\epsilon)) - d(x_1, x_0)}{\ln(\epsilon)}$ □

This proves that the sequence x_n is Cauchy. By completeness of (X, d), the sequence x_n converges to some point $y \in X$. Since φ is a continuous map, it follows by an application of the sequence lemma that $y \leftarrow x_{n+1} = \varphi(x_n) \rightarrow \varphi(y)$. Since (X, d) is Hausdörff, it follows that the limit is unique i.e. $y = \varphi(y)$, which proves the result.

Remark: A closed subspace of a complete metric space is complete.

Now, the proof of the inverse function theorem goes like this. Let F be as before and choose $0 < \epsilon \leq \frac{1}{2}$. Let $0 < \delta < 1$ be such that the open ball $U = B_d(a, \delta)$ is in $\mathbb{R}^n$ and using continuity of $DF(x)$, we have

$$\|F'(x) - I\| < \epsilon$$

for all $x \in U$. Remember $F'(a) = I$ the identity matrix of order n. Let $V = F(U)$. Then check that $F : U \rightarrow V$ is the required diffeomorphism. To see this we need to generate a contraction mapping. Let us define for any $y \in V$ $\varphi_y : U \rightarrow \mathbb{R}$ by $\varphi_y(x) = x + (y - F(x))$ which is continuously differentiable on U and

$$\varphi_y'(x) = I - F'(x), \ \forall \ x \in U.$$

We see that $\|\varphi_y'(x)\| < \epsilon < 1$ for all $x \in U$. Using mean value theorem for any $x_1, x_2 \in U$ we have

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \|D\varphi_y(c)\| \|x_1 - x_2\| < \epsilon \|x_1 - x_2\|$$

for some $c \in U$. This proves that φ_y is a contraction mapping. By the contraction mapping theorem, φ_y has a fixed point. Let it be α. Then $\alpha = \alpha + y - F(\alpha)$ or $y = F(\alpha)$, this proves that F is onto. If $F(x_1) = F(x_2)$ then $\varphi_y(x_1) - \varphi_y(x_2) = x_1 - x_2$ and therefore

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| = \|x_1 - x_2\| < \epsilon \|x_1 - x_2\|$$

which is possible only if $x_1 - x_2 = 0$ and this proves that F is injective.

To see that $F(U)$ is open, take a point $y \in F(U) = V$. Then $y = F(x)$ for some $x \in U$. For $\epsilon > 0$ as above, using the continuity of F at x, there is an open ball $B_d(x,r) \subset \bar{B}_d(x,r) \subset U$ such that for all $x' \in \bar{B}_d(x,r)$

$$\|F(x) - F(x')\| \leq \epsilon.$$

Now consider the open ball $B_d(F(x), \epsilon r)$ such that for each $y' = F(x') \in \bar{B}_d(F(x), \epsilon r)$ we have

$$\|\varphi_y'(x') - x\| \leq \|\varphi_{y'}(x') - \varphi_{y'}(x)\| + \|\varphi_{y'}(x) - x\| \leq \epsilon\|x' - x\| + \|F(x) - F(x')\| \leq \epsilon r + \epsilon r \leq r$$

which shows that $\varphi_{y'}(\bar{B}_d(x,r)) \subseteq \bar{B}_d(x,r)$ and as $r \leq \delta < 1$, using the contraction mapping theorem, there is a point $a \in B_d(x,r)$ such that $\varphi_{y'}(a) = a$ or $a + y' - F(a) = a$ or $F(a) = y'$. We have proved that $B_d(F(x), r\epsilon) \subset f(U)$. This proves that V is open. Continuity and differentiabliy of F^{-1} is done in the same way as before and we omit it here.

2 The implicit function theorem

Here we will discuss an application of the inverse function theorem, namely, the implicit function theorem.

Definition: (Inclusion map) The inclusion maps $\eta_1 : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$ and $\eta_2 : \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$are defined by $\eta_1(x) = (x, 0)$ and $\eta_2(y) = (0, y)$.

If $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is a linear map such that $T_i = T \circ \eta_i$ then observe that $(T \circ \eta_1 \circ \pi_1)(x, y) = (T \circ \eta_1)(x) = T(x, 0)$. Similarly $(T \circ \eta_2 \circ \pi_2)(x, y) = (T \circ \eta_2)(y) = T(0, y)$. It follows that $T = T_1 \circ \pi_1 + T_2 \circ \pi_2$

Proposition 2.1. *Let $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be a linear map such that $T_1 = T \circ \eta_1$ is invertible. Then for every $y \in \mathbb{R}^m$ there is a unique $x \in \mathbb{R}^n$ such that $T(x, y) = 0$ such that $x = -T_1^{-1} \circ T_2(y)$ where $T_i = T \circ \eta_i$.*

Proof. Since $T = T_1 \circ \pi_1 + T_2 \circ \pi_2$ and if we take $x = T_1^{-1} \circ T_2(y)$ this means $T_1(x) = T_2(y)$ and we have

$$T(x, y) = T_1 \circ \pi_1(x, y) + T_2 \circ \pi_2(x, y) = T_1(x) + T_2(y) = T_1(x) - T_1(x) = 0$$

which proves the assertion. □

For nonlinear maps also an analogous result holds which we will prove as the next theorem.

Theorem 2.2. (Implicit function theorem) *Suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable function in an open set containing the point (a, b) such that $f(a, b) = 0$. Let M be the $m \times m$ matrix s.t. $M_{i,j} = D_{n+j} f_i(a, b),\ 1 \leq i, j \leq m$. If $\det(M) \neq 0$, then there is an open set $A \subset \mathbb{R}^n$ containing a and an open set $B \subset \mathbb{R}^m$ containing b and, a differentiable map $g : A \to B$ such that $f(x, g(x)) = 0$ for all $x \in A$.*

Proof. Define $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ by $F(x, y) = (x, f(x, y))^T = (x_1, ..., x_n, f_1(x, y), ..., f_m(x, y))^T$. Then F is also continuously differentiable and

$$F'(a, b) = \begin{pmatrix} I_{n\times n} & O_{n\times m} \\ N_{m\times n} & M_{m\times m} \end{pmatrix}$$

where $N_{r,s} = D_s f_r(a, b), 1 \leq r \leq m, 1 \leq s \leq n$, $I_{n\times n}$ and $O_{n\times m}$ are identity and zero matrices respectively with the subscripts denoting their orders. Observe that $\det(F'(a, b)) = \det(M) \neq 0$. Thus by the inverse function theorem, there exist open sets $W \ni F(a,b) = (a,0)$ and an open set $A \times B \ni (a,b)$ in $\mathbb{R}^n \times \mathbb{R}^m$ such that $F|_{A\times B} : A \times B \to W$ is a diffeomorphism with the inverse function $H : W \to A \times B$ of the form $H(x,y) = (x, k(x,y))$ for some differentiable function $k : W \to B$. Now consider

$$f(x, k(x,y)) = (f \circ H)(x,y) = (\pi_2 \circ F) \circ H(x,y) = \pi_2(x,y) = y.$$

Then it follows that $f(x, k(x,0)) = 0$ and $g(x) = k(x,0)$ is the required function. □

Theorem 2.3. (Surjective form of implicit function theorem) *Suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable function in an open set containing the point (a,b) such that $f(a,b) = 0$. If $Df(a,b)$ is surjective, then there is an open set $V \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing 0 and a diffeomorphism $\varphi : V \to \varphi(V) \subseteq \mathbb{R}^n \times \mathbb{R}^m$ such that $(f \circ \varphi)(x_1, ..., x_{n+m}) = (x_{n+1}, ..., x_{n+m})$.*

Proof. Since $Df(a,b)$ is surjective, it follows that rank of the Jacobian matrix $f'(a,b)$ is $\geq m$. Using elementary row transformations, we can transform the matrix $f'(a,b)$ into the matrix in which at least m columns are the standard basis elements of $\mathbb{R}^{n+m}$. Let these columns be $j_1, ..., j_m$ th columns for some $j_i \in \{1, ..., n+m\}$. Let $\sigma \in S_{n+m}$ such that $\sigma(x_1, ..., x_{n+m}) = (x_{i_1}, ..., x_{i_n}, x_{j_1}, ..., x_{j_m}); i_k, j_\ell \in \{x_1, ..., x_{n+m}\}$. Then $F(x,y) = (x, f \circ \sigma)$ satisfy the conditions of the inverse function theorem with $\det(M) \neq 0$ where $M_{ij} = D_{j+n}(f \circ \sigma)_i(a,b)$ for $F(x,y) = (x, f \circ \sigma(x,y))$. Therefore there exist open sets $U \times V \ni (a,b)$, $W \ni f \circ \sigma(a,b)$ and a diffeomorphism $F|_{U\times V} U \times V \to W$ with differentiable inverse $F^{-1} : W \to U \times V$ such that $F(x,y) = (x, (f \circ \sigma)(x,y))$ and $\varphi(x,y) = (x, k(x,y))$ for some differentiable function $k : W \to V$. Define $\varphi = \sigma \circ F^{-1}$. Then $f \circ \varphi(x,y) = \pi_2 \circ F \circ F^{-1}(x,y) = y = (x_{n+1}, ..., x_{n+m})$ as required. □

Theorem 2.4. (Injective form of implicit function theorem) *Suppose $f : \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ is continuously differentiable function in an open set containing the point b such that $f(b) = 0$. If $Df(b)$ is injective, then there is an open set $V \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing $(0,b)$ and a diffeomorphism $\varphi : V \to \varphi(V) \subseteq \mathbb{R}^n \times \mathbb{R}^m$, such that $(\varphi \circ f)(x_{n+1}, ..., x_{n+m}) = (0, ..., 0, x_{n+1}, ..., x_{n+m})$.*

Proof. Since f is injective, again the rank of the $(m+n) \times n$ matrix $f'(b) \geq m$. Using elementary row transformations, we can transform the matrix $f'(b)$ into the matrix in which at least m rows are the standard basis elements of $\mathbb{R}^{n+m}$. Let these rows be be $j_{n+1}, ..., j_{n+m}$ th rows for some $j_{n+i} \in \{1, ..., n+m\}$. Let $\sigma \in S_{n+m}$ s.t. $\sigma(x_1, ..., x_{n+m}) = (x_{j_1}, ..., x_{j_{n+m}})$ and M be the $m \times m$ matrix s.t. $M_{i,j} = D_{n+i,j}(\sigma \circ f)_{n+i}(b),\ i,j = 1, 2..., m$. Then $\det(M) \neq 0$. Using inverse function theorem for the function

$$F(x,y) = \sigma \circ f(y) + (x,0) = ((\sigma \circ f)_1(y) + x_1, ..., (\sigma \circ f)_n(y) + x_n, (\sigma \circ f)_{n+1}, ..., (\sigma \circ f)_{n+m}),$$

with $\det(F'(0,b)) = \det(M) \neq 0$, there exist open sets $U \times V \ni (0,b)$ and $W \ni (0,0)$ such that $F|_{U\times V} : U \times V \to W$ is a diffeomorphism. Define $\varphi = F^{-1} \circ \sigma$. Note that $F(0,y) = \sigma \circ f(y)$. Then

$$\varphi \circ f(y) = F^{-1} \circ \sigma \circ f(y) = F^{-1} \circ F(0,y) = (0,y).$$

This completes the proof. □

Example: Consider the function $f(x,y) = y^2 - x^3$. Then the derivative of f vanishes identically only at $(0,0)$. Therefore the implicit function theorem is valid for all $(a,b) \neq (0,0)$. If we define $g(x,y) =$

$2bx + 3a^2y - (2ab + 3a^2b)$ and $\psi = (f, g)$, then $\psi(a, b) = (0, 0)$ and $D\psi(a, b)$ is invertible. Therefore, by inverse function theorem, in some open set $W \ni (a, b)$ we have $\psi : U \times V \to W$ is a diffeomorphsim. Put $\varphi = \psi^{-1} : W \to U \times V$. Then

$$(x, y) = \psi \circ \varphi(x, y) = (f, g) \circ \varphi(x, y) = (f(\varphi(x, y)), g(\varphi(x, y)))$$

which implies that $f \circ \varphi(x, y) = x$.

2.1 References:

1. M. Spivak. *Calculus on Manifolds*, Addison Wesley (1965)
2. A. R. Shastri. *Elements of Differential Topology*, CRC Press (Taylor and Francis) (2011).