Implicit and Inverse Function Theorem

Quick note on the implicit and inverse function theorem and the application to immersions and submersions. We follow [Bre93].

1 Pre-Requisites

We need two results before we prove the main results one from linear algebra and one from analysis.

Lemma 1.1 (Cramer's Rule)**.** *Consider the linear equation with $Ax = b$ for $\det(A) \neq 0$. Then the unique solution to the equation is given by $x_i = \frac{\det(A_i|b)}{\det(A)}$ where $A_i|b$ is the matrix that replaces the i^{th} column of A with b.*

Proof. Consider the map $y \mapsto c_A(y) \in \mathbb{K}^n$ where $c_A(y)_i = \frac{\det(A_i|y)}{\det(A)}$. This map is linear by the properties of the determinant and we see that it takes the i^{th} column of A to the i^{th} unit vector. It follows that $c_A = A^{-1}$ and hence the lemma. □

By Cramer's rule we see that the coefficients solutions are algebraic functions of the coefficients of A and b.

Theorem 1.2 (Banach Contraction Principle)**.** *Let E be a complete metric space, and $T : E \to E$ a map such that $\exists 0 \leq K < 1$ with $d(Tx, Ty) \leq Kd(x, y)$ for all $x, y \in E$. Then there exists a unique solution $Tx_0 = x_0$, and $x_0 = \lim T^n x$ for any $x \in E$.*

Proof. Let $x \in E$ and consider the sequence $\{x, Tx, T^2x, \ldots\}$. Then we have that

$$d(x, T^n x) \leq d(x, Tx) + \cdots + d(T^{n-1}x, T^n x) \leq (1 + K + \cdots + K^{n-1})d(x, Tx) \leq \frac{d(x, Tx)}{1 - K}.$$

It follows that $d(T^n x, T^m x) \leq \frac{K^{\min(n,m)}d(x,Tx)}{1-K}$ and thus the sequence is Cauchy by the condition on K and so has a limit ξ. We have $T\xi = T \lim T^n x = \lim T^{n+1}x = \xi$.

For uniqueness let η be another fixed point then $d(\xi, \eta) = d(T\xi, T\eta) \leq Kd(\xi, \eta)$ hence $d(\xi, \eta) = 0$. □

2 Main Results

We are now ready to prove the eponymous theorems. The first ingredient is the following technical lemma. Unless otherwise specified the norm shall be the sup norm

Lemma 2.1. *Let $\xi \in U \subset \mathbb{R}^n$, $\eta \in V \subset \mathbb{R}^m$, and suppose we have a C^1 function $f : U \times V \to \mathbb{R}^m$ such that $f(\xi, \eta) = \eta$ and $\frac{\partial f_i}{\partial y_j}(\xi, \eta) = 0$ where y_i are the coordinates on $\mathbb{R}^m$. Then there exists a continuous function ϕ defined on some neighbourhood $\xi \in U' \subset U$ such that $\phi(\xi) = \eta$ and $\phi(x) = f(x, \phi(x))$ $\forall x \in U'$.*

Proof. As $\left[\frac{\partial f}{\partial y}\right](\xi, \eta) = 0$ and by assumption it is continuous given any $\epsilon > 0$ we can find $U' \times V' \subset U \times V$ such that $\left|\left|\left[\frac{\partial f}{\partial y}\right]\right|\right| < \epsilon$ for $(x, y) \in U' \times V'$. We now apply the mean value theorem to the function $y \in V' \mapsto f(x, y)$ for each $x \in U'$ so that $\forall y_1, y_2 \in V'$ $\exists \overline{y} \in [\eta, y]$ such that

$$f(x, y_1) - f(x, y_2) = \sum \frac{\partial f}{\partial y_j}(x, \overline{y})(y_1 - y_2)_j.$$

Then we find $||f(x, y_1) - f(x, y_2)|| < \epsilon\, ||y - \eta||$.

We next choose $B = \overline{B_b(\eta)} \subset V'$. Then, using continuity of f, we can choose $A = \overline{B_a(\xi)} \subset U'$ such that $||f(x, \eta) - \eta|| \leq (1 - \epsilon)b$. The neighbourhood $U'' \times V''$ will be the one on which we shall prove the result.

Let $F = \{\phi : A \to B \mid \phi(\xi) = \eta\}$ with the sup norm (as B is bounded this is well defined) which is a complete metric space. We define the following operator $T\phi(x) = f(x, \phi(x))$. We have that

$$||f(x,\phi(x)) - \eta|| = ||f(x,\phi(x)) - f(x,\eta) + f(x,\eta) - \eta)|| \le \epsilon\, ||\phi(x) - \eta|| + (1-\epsilon)b \le b$$

hence it lies in B. Secondly we have that $T\phi(\xi) = f(\xi, \phi(\xi)) = f(\xi, \eta) = \eta$, so $TF \subset F$. Next we have

$$||T\phi(x) - T\psi(x)|| = ||f(x,\phi(x)) - f(x,\psi(x))|| < \epsilon\, ||\phi(x) - \psi(x)||$$

by our choice V'. Hence T is a contraction and thus has a unique limit by 1.2 and such a function satisfies $\phi(x) = f(x, \phi(x))$.

To show that this limit is continuous note that, again by 1.2, it is the limit of the sequence $\{T^n[x \mapsto \eta]\}$ which is a uniform limit of continuous functions hence continuous. $\square$

Before we give the formal proof of the implicit function theorem lets consider the statement we have a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ with $f(x_0, y_0) = 0$. If we expand then, to first order, we have

$$0 = -f(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)\cdot(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\cdot(y-y_0) + O(...^2).$$

So we see if the Jacobian of the map $y \mapsto f(x_0, y)$ is of full rank at y_0 then we can invert and solve. This shows that the conditions of the theorem are necessary in general and also sufficient to first order. The above equation says that if we move in the direction of $x - x_0$ then we can, using the invertability of the Jacobian, find $\phi(x)$ that counteracts the change in this direction.

Theorem 2.2 (The Implicit Function Theorem)**.** *Let $\xi \in U \subset \mathbb{R}^n$, $\eta \in V \subset \mathbb{R}^m$ and suppose we have a C^1 function $g : U \times V \to \mathbb{R}^m$ such that $g(\xi, \eta) = 0$. Then if the map $y \mapsto g(\xi, y)$ has non-zero Jacobian at η we can find a neighbourhood of (ξ, η) and a function ϕ defined on a neighbourhood of (ξ, η) such that $g(x, \phi(x)) = 0$. Moreover if g is C^p then so is ϕ.*

Proof. Let T denote the tangent of the map $y \mapsto g(\xi, y)$ at η. By assumption it is invertible so consider the function

$$f(x,y) := y - T^{-1}g(x,y).$$

We have that $f(\xi, \eta) = \eta$, and

$$\frac{\partial f_i}{\partial y_j}(\xi,\eta) = \delta_{ij} - T^{-1}_{ik}T_{kj} = 0.$$

Hence the conditions of 2.1 are satisfied so we can conclude that there exists a neighbourhood of (ξ, η) and C^0 function ϕ defined on this neighbourhood $W \subset U \times V$, such that $\phi(x) = f(x, \phi(x)) \Leftrightarrow g(x, \phi(x)) = 0$.

We now turn to the question of differentiability, first lets show that ϕ is C^1. As the Jacobian $J = J(g; y)$ is non-zero we can find a neighbourhood of (ξ, η) on which it is non-zero. We can now apply the mean value theorem to any two points in this neighbourhood (intersected with W defined above) to get that we can find some $(a, b) \in [(x_1, \phi(x_1)), (x, \phi(x)]$ such that

$$0 = g(x_1,\phi(x_1)) - g(x,\phi(x)) = \sum \frac{\partial g}{\partial x_i}(a,b)\cdot(x_1-x)_i + \sum \frac{\partial g}{\partial y_j}(a,b)\cdot(\phi(x_1)-\phi(x))_j.$$

We now set $x_1 = x + h^{(i)}$ where $h^{(i)}$ has h in the i^{th} coordinate and zero elsewhere. We then get

$$-\frac{\partial g_k}{\partial x_i}(z_h) = J(z_h)_{k,j}\frac{(\phi(x+h^{(i)}) - \phi(x))_j}{h}$$

We then have, using g is C^1, and the fact that J is non-zero on a closed neighbourhood of $(x, \phi(x))$, that there is a unique limit of as h goes to 0 which is given by $J^{-1}(x, \phi(x))\frac{\partial g}{\partial x^i}$ and so it follows that ϕ is C^1 as this function is continuous.

We now apply the note after Cramer's rule 1.1: we have that $\frac{\partial \phi}{\partial x_i} = H(x, \phi(x))$ where H is an algebraic function of the terms $\frac{\partial g}{\partial x_i}(x, \phi(x))$ and $\frac{\partial g}{\partial y_j}(x, \phi(x))$. Now assume g is C^p for $p \le \infty$, then the function $H(x, \phi(x))$ has differentiability at least $\min(p-1, \operatorname{ord}(\phi))$ where ord is the level of differentiability. Hence $\operatorname{ord}(\phi) - 1 \ge \min(p-1, \operatorname{ord}(\phi))$ which leads to a contradiction if $\operatorname{ord}(\phi) < p$ so we can differentiate ϕ p-times and by induction on p we see that ϕ is indeed C^p.

We finally turn to the case where g is analytic. We then can use so that we immediately conclude that the solution is analytic. $\square$

Theorem 2.3 (The Inverse Function Theorem). *Let $\psi : U \subset \mathbb{R}^m \to \mathbb{R}^m$ be a C^1 function that the Jacobian of g has rank m at y_0. Then there is a neighbourhood of $x_0 := \psi(y_0) \in V$ and a function $\phi : V \to U' \subset U'$ such that $\phi\psi(y) = y \forall y \in U'$. Moreover if ψ is C^p then so is ϕ*

Proof. We apply the implicit function theorem to the function $f(x, y) = x - \psi(y)$ which is zero at (x_0, y_0) and is clearly C^1. We thus have a neighbourhood of x_0, V, and a function $\phi : V \to U'$ such that $x = \psi\phi(x)$. It follows that ϕ must be 1-1 and ψ covers V from U'.

Now the equation $x = \psi\phi(x)$ implies that the $T\phi(x_0) = T\psi(y_0)^{-1}$ which has full rank. We may thus apply the implicit function theorem again to ϕ to get $\psi' : U'' \subset U' \to V' \subset V$ such that $y = \phi\psi'(y)$ for all $y \in U''$. Let $A = U''$ and B $= V'$; then we have on A $\psi(y) = \psi\phi\psi'(y) = \psi'(y)$, i.e. $\psi' = \psi$. It follows that ϕ and ψ are inverse functions on these two sets. ☐

3 Immersions and Submersions

Proposition 3.1. *Let $f : N \to M$ be an immersion at $p \in N$. Then we can find co-ordinate patches around $p \in U \subset \mathbb{R}^n$ and $f(p) \subset V$ such that f has the form $y \mapsto (y, 0)$.*

Proof. Taking any co-ordinate patch around p and $f(p)$, we can reduce to the case of open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. At the point p we have that the Jacobian of f has rank n $(\leq m)$. We can then find a matrix $A = A_p$ such that the matrix

$$\begin{bmatrix} \frac{\partial f}{\partial y} & A \end{bmatrix}$$

is invertible where A has dimension $m \times (m-n)$. It follows that if we consider $U \times \mathbb{R}^k \to \mathbb{R}^m$, where $k = m-n$, defined by

$$(y, z) \mapsto \phi(y, z) = f(y) + Az$$

where we consider A as a linear operator $\mathbb{R}^k \to \mathbb{R}^m$. This function has invertible Jacobian at the point $(p, 0)$ (which maps to $f(p)$) and hence by the inverse function theorem we have that we can find a neighbourhood of $(p, 0)$ and one of $f(p)$ on which ϕ is a diffeomorphism. Then under the co-ordinate transformation ϕ^{-1} we get $\phi^{-1} \circ f$ has the form $y \mapsto (y, 0)$ hence the result. ☐

Corollary 3.2. *Let $\iota : M \to N$ be an embedding; then for each $x \in M$ we can find a co-ordinate patch around $\iota(x)$, U, such that $\iota(N) \cap U$ hyperplane defined by $x_{m+1} = \cdots = x_n = 0$.*

Proof. Using 3.1 we can find a neighbourhood of x and $\iota(x)$, U and V respectively, such that ι has the form $\iota(y) = (y_1, \ldots, y_m, 0, \ldots 0)$. As $\iota(N)$ is an embedding we can find an open neighbourhood W of $\iota(x)$ such that $\iota^{-1}W = U$. Then $W \cap V$ is the required neighbourhood. ☐

Proposition 3.3. *Let $f : N \to M$ be a submersion at $p \in N$. Then we can find co-ordinate patches around $p \in U \subset \mathbb{R}^n$ and $f(p) \subset V$ such that f has the form $(y_1, \ldots, y_n) \mapsto (y_1, \ldots, y_m)$, i.e. it is the canonical projection onto the first m co-ordinates.*

Proof. Taking any co-ordinate patch around p and $f(p)$, we can reduce to the case of open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. At the point p we have that the Jacobian of f has rank m and so we can find a $(n-m) \times n$ matrix $A = A_p$ such that

$$\begin{bmatrix} \frac{\partial f}{\partial y} \\ A \end{bmatrix}$$

has rank m. Then the map

$$y \mapsto \phi(y) = \begin{bmatrix} f(y) \\ Ay \end{bmatrix}$$

has invertible Jacobian at $y = p$ and so we can find a local inverse there. Then we have $f \circ \phi^{-1}$ is as required by the proposition. ☐

Note that we actually do not need to alter the coordinate patch of the codomain (although we may need to shrink it) in the case of submersions (for immersions we do not need to alter the coordinate patch in the domain).

Corollary 3.4. *Let $f : M \to N$ be a smooth function. If $p \in N$ is a regular value then $f^{-1}(p)$ is a smooth manifold of dimension $m - n$.*

Proof. Let $Z = f^{-1}(p)$, and $x \in Z$. Then by 3.3 we can find a neighbourhood of x, U, and $p = f(x)$ such that f has the form $(y_1, \ldots, y_m) \mapsto (y_1, \ldots, y_n)$. It follows that $Z \cap U = U \cap \{0\} \times \mathbb{R}^k$, where we have assumed p corresponds to 0 in the given coordinate system. Hence we have locally given Z the structure of an open subset of $\mathbb{R}^k$ with $k = m - n$. On a patch of $U_x \cap U_y$ for $x, y \in Z$ by the comment at the end of the proof of 3.3 we can let $p \in V = V_x \cap V_y$ then $U_x \cap U_y = f^{-1}V$ and f has the same functional form on both coordinate patches. We thus have that, letting $\phi : U_x \cap U_y \to U_y \cap U_x$ be the coordinate transformation that

$$(y_1, \ldots, y_n) = f(y_1, \ldots, y_n) = f\phi(y_1, \ldots, y_n) = (\phi_1(y), \ldots, \phi_n(y))$$

so that ϕ acts as the identity on the first n-coordinates. It follows that the remaining k ϕs must be smooth and a diffeomorphism. This shows the transition functions are also smooth. Z is Hausdorff and second countable as these are inherited from M hence the result. $\square$

Example 3.5. *Let us consider $f(x, y, z) = x^p + y^p + z^p$ on $\mathbb{R}^3$ for $p \geq 2$. We have that this has one critical point, the origin, and hence one critical value* 0*. It follows that $f^{-1}(R^p)$ is a smooth manifold, in particular for $p = 2$ we see the sphere is smooth.*

Example 3.6. *Consider a complex affine algebraic variety, that is a subset of $\mathbb{C}^n$ cut out by k polynomials $f_1, \ldots, f_k$. We can think of this as a function $f : \mathbb{C}^n \to \mathbb{C}^k$ and the variety is $f^{-1}(0)$. If 0 is a regular value then we can put a smooth structure on the algebraic variety.*

References

[Bre93] Glen E. Bredon. *Topology and Geometry*. Springer Verlag, 1993.