1 Inverse Function Theorem

From calculus we have the result:

Theorem: (1D) Inverse Function Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function with $f'(x_0) \neq 0$, there exists an interval I containing x_0 and an interval J containing $f(x_0)$ so that $f : I \to J$ is a diffeomorphism.

Generalising to arbitrary dimensions:

Theorem: Inverse Function Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ a smooth function such that df_{x_0} is invertible at x_0. Then there is an open set U containing x_0 and an open set V containing $f(x_0)$ such that $f|_U : U \to V$ is a diffeomorphism. Moreover

$$df^{-1}_{f(x_0)} = (df_{x_0})^{-1}$$

Note that if f is a diffeomorphism, then $f^{-1} \circ f(x) = x$. That is, $f^{-1} \circ f = \mathrm{Id}_x$. Since $d\,\mathrm{Id}_x = \mathrm{Id}_n$, by the chain rule we have

$$\mathrm{Id}_n = d\,\mathrm{Id}_x = d(f^{-1} \circ f)_{x_0} = df^{-1}_{f(x_0)} \cdot df_{x_0}.$$

That is df_{x_0} is invertible and

$$(df_{x_0})^{-1} = df^{-1}_{f(x_0)}.$$

Thus $d(f^{-1})$ at $y_0 = f(x_0)$ is necessarily equal to $(df)^{-1}$ at x_0. In one dimension $df = f'$ and $d(f^{-1}) = 1/f'$.

The basic idea is that if df is invertible, then f is invertible to first order. Writing

$$f(x) = f(x_0) + df_{x_0} \cdot (x - x_0) + \mathrm{o}(|x - x_0|).$$

Let us ignore the $\mathrm{o}(|x - x_0|)$ term (after all, it's insignificant compared with everything else for x near x_0!) and assume

$$f(x) = f(x_0) + df_{x_0} \cdot (x - x_0).$$

Then we can rearrange to solve for x to get

$$x = x_0 + df^{-1}_{x_0}(f(x) - f(x_0)).$$

Writing $y = f(x)$ and $y_0 = f(x_0)$ we obtain the inverse,

$$f^{-1}(y) = f^{-1}(y_0) + df^{-1}_{x_0} \cdot (y - y_0).$$

The task then is to work out how to deal with the presense of the $\mathrm{o}(|x - x_0|)$ term. The approach is to construct a suitable *contraction* map (i.e. a map that strictly shrinks distances - see more below). We will prove the inverse function theorem below after considering some consequences. As a prview, to prove the Inverse Function Theorem, given y we need a uniquely solution of $f(x) = y$. Define

$$T_y(x) = x - df^{-1}_{x_0}(f(x) - y).$$

Then we show that for suitable $r > 0$, T_y is a contraction map $\bar{B}_r(x_0) \to \bar{B}_r(x_0)$ and a cornerstone result in analysis (namely the Banach Fixed Point Theorem) implies that T_y posses a unique fixed point x_y^*. That is, there is a unique point $x_y^* \in \bar{B}_r(x_0)$ such that $T_y(x_y^*) = x_y^*$. Observe then that

$$\begin{aligned} T_y(x_y^*) = x_y^* &\Leftrightarrow df_{x_0}^{-1}(f(x_y^*) - y) \\ &\Leftrightarrow f(x_y^*) = y. \end{aligned}$$

The last equivalence follows from the assumption that df_{x_0} is invertible. Thus $f(x_y^*) = y)$ if and only if T_y has a fixed point x_y^*. By showing this fixed point is unique we then may unambiguously define

$$f^{-1}(y) = x_y^*.$$

Here is an example application of the Inverse Function Theorem.

Example

Consider

$$\begin{cases} x - y^2 & = a \\ x^2 + y + y^3 & = b \end{cases}$$

For $(a, b) = (0, 0)$: $(x, y) = (0, 0)$ is a solution.
Question: For what (a, b) is the system solvable?
To answer the question, let $F(x, y) = (x - y^2, x^2 + y - y^3)$. Then

$$dF = \begin{pmatrix} 1 & -2y \\ 2x & 1 - 3y^2 \end{pmatrix}$$

We have $dF_{(0,0)} = \mathrm{Id}$ hence by the IFT there is a neighbourhood of $(x, y) = (0, 0)$ for which F maps diffeomorphically onto a neighbourhood of $(a, b) = (0, 0)$. Therefore, for (a, b) in a neighbourhood of $(0, 0)$, there is a neighbourhood of $(0, 0)$ containing a unique solution of $F(x, y) = (a, b)$.
Note that given $(a, b))$, there is not generally a unique solution. In fact, even for $(a, b) = (0, 0)$ there is not a unique solution since if y is a real root of $y^3 + y^2 + 1$, then $F(y^2, y) = (0, 0)$. Such a root always exists since $y^3 + y^2 + 1$ is an odd-degree polynomial.

2 Implicit Function Theorem

Using the isomorphism, $\mathbb{R}^n \oplus \mathbb{R}^k \simeq \mathbb{R}^{n+k}$ we may write a point in $\mathbb{R}^{n+k}$ as (x, y) with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^k$. Then for a function $F = F(x, y) : \mathbb{R}^{n+k} \to \mathbb{R}^k$ we also split the differential into x, y parts:

$$dF = \begin{pmatrix} d_x F & d_y F \end{pmatrix}.$$

Theorem: Implicit Function Theorem

Let $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$ be smooth with (x_0, y_0) such that $d_y F|_{(x_0,y_0)}$ is invertible. Then there is an open neighbourhood U of x_0 and a unique smooth function $g : U \to \mathbb{R}^n$ such that

$$F(x, g(x)) = F(x_0, y_0).$$

The Implicit Function Theorem is equivalent to the Inverse Function Theorem. Here we show how to derive the Implicit Function Theorem from the Inverse Function Theorem.

Proof

Define
$$\bar{F}(x,y) = (x, F(x,y)) \in \mathbb{R}^{n+k}$$
Then
$$d\bar{F} = \begin{pmatrix} \mathrm{Id}_n & 0 \\ * & d_yF \end{pmatrix}$$
is invertible at (x_0, y_0) since the assumption is that d_yF is invertible at (x_0, y_0). Hence by the inverse function theorem, $\bar{F}$ is locally invertible.
Since $\bar{F}(x,y) = (x, F(x,y))$,
$$\bar{F}^{-1}(x,y) = (x, G(x,y))$$
for a smooth function $G : \mathbb{R}^{n+k} \to \mathbb{R}^k$. This follows by writing $\bar{F}^{-1} = (H, G)$, from which we claim that necessarily $H(x,y) = x$. By the definition of inverse functions,
$$\begin{aligned}(x,y) &= \bar{F} \circ \bar{F}^{-1}(x,y) \\ &= \bar{F}(H(x,y), G(x,y)) \\ &= (H(x,y), F(G(x,y))).\end{aligned}$$
Comparing the first component of the left and right hand sides we see that $x = H(x,y)$ as claimed.
Now let $c = F(x_0, y_0)$ and
$$g(x) = G(x,c)$$
from which it follows that
$$\begin{aligned}(x, F(x, g(x))) &= \bar{F}(x, g(x)) \\ &= \bar{F}(x, G(x,c)) \\ &= \bar{F} \circ \bar{F}^{-1}(x,c) \\ &= (x,c)\end{aligned}$$
and $F(x, g(x)) = c = F(x_0, y_0)$ as required.

Exercise

Assuming the Implicit Function Theorem is true, prove the Inverse Function Theorem.

We may interpret the Implicit Function Theorem as follows: consider the level set
$$F^{-1}(c) = \{(x,y) : F(x,y) = c\}.$$
If d_yF is invertible for each $(x,y) \in F^{-1}(c)$, then the level set is locally the graph of a smooth function.

Exercise

Let $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$ be a smooth function such that dF has rank k at $z_0 \in \mathbb{R}^{n+k}$. By permuting the indices, use the Implicit Function Theorem to show that for z in a neighbourhood of z_0, we may parametrise the level set $F(z) = F(z_0)$ as the graph of a smooth function $g : \mathbb{R}^n \to \mathbb{R}^k$.

Example

Let $F(x, y) = x^2 + y^2$
Here $n = k = 1$

$$dF = \begin{pmatrix} 2x & 2y \end{pmatrix}$$

For $x \neq \pm 1$

$$F(x, \sqrt{1 - x^2}) = 1$$

3 Immersion and Submersion Theorems

Here are some further statements equivalent to the Inverse Function Theorem, and hence also equivalent to the Implicit Function Theorem.

Exercise

Prove that the theorems below are equivalent to the Inverse Function Theorem. You may find it easier to prove equivalence with the Implicit Function Theorem which is equivalent anyway.

Definition

Let $F : \mathbb{R}^{n+k} \to \mathbb{R}^k$ be a smooth map. Then F is an *submersion* if dF is surjective.

Note that

$$\begin{aligned} dF \text{ surjective} &\Leftrightarrow dF \text{ has maximal rank} \\ &\Leftrightarrow \operatorname{rnk} dF = k = \dim \operatorname{coDom}(dF) \\ &\Leftrightarrow \dim \ker dF = n \end{aligned}$$

Definition

An projection of $\mathbb{R}^{n+k}$ onto $\mathbb{R}^k$ is a map of the form

$$\pi : x \in \mathbb{R}^{n+k} \mapsto (x^{n+1}, \dots, x^{n+k}) \in \mathbb{R}^k$$

Note that $d\pi = \begin{pmatrix} \mathrm{Id}_n & 0_k \end{pmatrix}$ is surjective.

We may also change the order: eg. $\pi(x_1, x_2, x_3) = (x_2, x_3)$

Theorem

Let F be a submersion. Then F is locally a projection up to diffeomorphism.

There are diffeomorphisms

$$\begin{aligned} \varphi &: U \subseteq \mathbb{R}^{n+k} \to V \subseteq \mathbb{R}^{n+k} \\ \psi &: W \subseteq \mathbb{R}^{k} \to Z \subseteq \mathbb{R}^{k} \end{aligned}$$

such that $F|_U = \psi^{-1} \circ \pi \circ \varphi$

Dual to the notion of submersion is the notion of immersion.

Definition

Let $F : \mathbb{R}^n \to \mathbb{R}^{n+k}$ be a smooth map. Then F is an *immersion* if dF is injective.

$$\begin{aligned} dF \text{ injective} &\Leftrightarrow dF \text{ has maximal rank} \\ &\Leftrightarrow \operatorname{rnk} dF = n = \dim \operatorname{Dom}(dF) \\ &\Leftrightarrow \dim \ker dF = 0 \end{aligned}$$

Definition

An inclusion of $\mathbb{R}^n$ into $\mathbb{R}^{n+k}$ is a map of the form

$$\iota : x \in \mathbb{R}^n \mapsto (x, 0_k)$$

where $0_k = (0, \ldots, 0) \in \mathbb{R}^k$.

Note that $d\iota = \begin{pmatrix} \mathrm{Id}_n \\ 0_k \end{pmatrix}$ is injective.

We may also change the order: eg. $\iota(x_1, x_2) = (0, x_1, x_2, 0)$

Theorem

Let F be an immersion. Then F is locally an inclusion up to diffeomorphism.

There are diffeomorphisms

$$\begin{aligned} \varphi &: U \subseteq \mathbb{R}^{n} \to V \subseteq \mathbb{R}^{n} \\ \psi &: W \subseteq \mathbb{R}^{n+k} \to Z \subseteq \mathbb{R}^{n+k} \end{aligned}$$

such that $F|_U = \psi^{-1} \circ \iota \circ \varphi$

4 Contraction Mappings

[12pt]article [all]pabmacros [utf8]inputenc [T1]fontenc fixltx2e graphicx longtable float wrapfig [normalem]ulem textcomp marvosym [nointegrals]wasysym latexsym amssymb amstext hyperref amsmath

Definition

A map $T : \bar{B}_r(p) \to \bar{B}_r(p)$ is a *contraction map* if there exists a constant $0 \leq L < 1$ such that

$$|T(x) - T(y)| \leq L\,|x - y|\,.$$

A contraction map strictly decreases the distance between two points. The primary significane of the definition is the following:

Theorem: Banach fixed point theorem

Let T be a contraction map. Then there exists a unique *fixed point* $x^* \in B_r(p)$ of T. That is, there exists a unique point x^* such that $T(x^*) = x^*$.

Proof

We have

$$\begin{aligned}|x-y| &\leq |x-T(x)|+|T(x)-y| \\ &\leq |x-T(x)|+|T(x)-T(y)|+|T(y)-y| \\ &\leq |x-T(x)|+L\,|x-y|+|T(y)-y|\,.\end{aligned}$$

and hence

$$|x-y| \leq \frac{|x-T(x)|+|T(y)-y|}{1-L}$$

Thus if $T(x)=x$ and $T(y)=y$) then $x=y$ and hence fixed points are unique.
To prove existence, pick any x_0 and define $x_n = T^n(x_0) = \underbrace{T\circ\cdots\circ T}_{n \text{ times}}(x_0)$. Supposing first that the limit exists, then using $x_n = T(x_{n-1})$ we have

$$x_* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T(\lim_{n\to\infty} x_{n-1}) = T(x^*)$$

Thus x_* is a fixed point and we just need to prove the limit exists. We do this by showing that $x_n = T^n(x_0)$ is a Cauchy sequence:

$$\begin{aligned}&|T^n(x_0)-T^m(x_0)| \\ &\quad\leq \frac{|T(T^n(x_0))-T^n(x_0)|+|T(T^m(x_0))-T^m(x_0)|}{1-L} \\ &\quad= \frac{|T^n(T(x_0))-T^n(x_0)|+|T^m(T(x_0)-T^m(x_0)|}{1-L} \\ &\quad\leq \frac{L^n\,|T(x_0)-x_0|+L^m\,|T(x_0)-x_0|}{1-L} \to 0\end{aligned}$$

as $m,n\to\infty$. Note here that we used $|T^n(x)-T^n(y)| \leq L^n\,|x-y|$ which follows by induction. Completness now ensures the limit exists and the proof is complete.