Today, we discuss the Inverse and Implicit Function Theorems.

Theorem (Inverse Function Theorem):

Let $U \subset \mathbb{R}^n$ be open, $f: U \to \mathbb{R}^n$ be $C^1(U)$, and let $Df|_p$ be invertible for some $p \in U$.

Then:

a) There are open neighborhoods V of p and W of $q := f(p)$ s.t. $f|_V : V \to W$ is bijective (i.e., $f|_V$ has an inverse).

b) The inverse $(f|_V)^{-1}$ is $C^1(W)$.

Note:

- $Df|_p$ invertible $\Longleftrightarrow$ The Jacobian matrix $J_{ij}(p) = \frac{\partial f_i(p)}{\partial x_j}$ is invertible.
- Using the chain rule we find: $\underset{\substack{\uparrow \\ \text{identity} \\ \text{on } \mathbb{R}^n}}{1} = \underbrace{D(f^{-1} \circ f)}_{\substack{\text{derivative of} \\ f^{-1}(f(x)) = x}}|_p \overset{\text{chain rule}}{=} Df^{-1}|_{f(p)} Df|_p$

$$\Rightarrow \boxed{Df^{-1}|_{f(p)} = (Df|_p)^{-1}}$$

- The inverse fct. thm. implies: The system of equations $f_i(x_1, \ldots, x_n) = y_i$, $i = 1, \ldots, n$ can be solved for $x_1, \ldots, x_n$ in terms of $y_1, \ldots, y_n$, if x and y are in small enough neighborhoods of p and q.
- If $f: V \to W$ is C^k, and f^{-1} exists and is C^k then f is called a C^k diffeomorphism.
- If any $p \in V$ has a neighborhood $\tilde{V}$ s.t. $f|_{\tilde{V}} : \tilde{V} \to f(\tilde{V})$ is a diffeomorphism, then f is called a local diffeomorphism. Note: A fct. that is a local diffeo. is not necessarily a global diffeo., see HW.

Closely related to the Inverse Function Theorem is the following question:

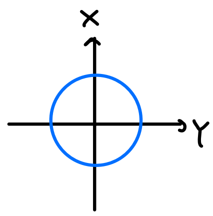
Let $f: \mathbb{R}^{n+m} \to \mathbb{R}^n$. Under which conditions can we solve $f(x,y)=0$ for $x \in \mathbb{R}^n$ in terms of $y \in \mathbb{R}^m$?

In other words: In the system of equations

$$\begin{aligned} f_1(x_1,\dots,x_n,y_1,\dots,y_m) &= 0 \\ &\vdots \\ f_n(x_1,\dots,x_n,y_1,\dots,y_m) &= 0, \end{aligned}$$

can we solve for $x_1(y_1,\dots,y_m),\dots,x_n(y_1,\dots,y_m)$, at least locally?

Ex.: $f(x,y) = x^2+y^2-1$, $x,y \in \mathbb{R}$.



$\Rightarrow$ $f(x,y)=0$ has two local solutions $x_\pm(y) = \pm\sqrt{1-y^2}$.

More precisely: Solution possible in an open neighborhood except when $x=0$ $(y=\pm 1)$.

At $x=0$, we have $\frac{\partial f}{\partial x}\Big|_{x=0} = 2x\Big|_{x=0} = 0$, i.e., $\frac{\partial f}{\partial x}\Big|_{x=0}$ not invertible.

$\Rightarrow$ It seems we require $\frac{\partial f}{\partial x}$ to be invertible

This is generalized in the following theorem:

Theorem (Implicit Function Theorem):

Let $U \subset \mathbb{R}^{n+m}$ be open, $f: U \to \mathbb{R}^n$ be $C^1(U)$, and $f(p,q)=0$ for some $(p,q)\in U$.

We assume that

$$\frac{\partial f}{\partial x}(p,q) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}\Bigg|_{(p,q)}$$

is invertible.

Then there are open sets $V \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$ with $(p,q)\in V$, $q \in W$ s.t. to every $y \in W$ corresponds a unique x s.t. $(x,y)\in V$ and $f(x,y)=0$. If this $x := g(y)$, then $g: W \to \mathbb{R}^n$ is C^1, $g(q)=p$, $f(g(y),y)=0$, and $Dg|_q = -\left(\frac{\partial f}{\partial x}\right)^{-1}\Big|_{(p,q)} \frac{\partial f}{\partial y}\Big|_{(p,q)}$.

Note: The formula for the derivative follows again from the chain rule:

$$0 = Df(g(\cdot))|_q = Df|_{(p,q)}\, D(g(y),y)|_q = \left(\frac{\partial f}{\partial x} \ \ \frac{\partial f}{\partial y}\right)\Big|_{(p,q)} \begin{pmatrix} \frac{\partial g}{\partial y} \\ 1 \end{pmatrix}\Big|_q$$

$$= \frac{\partial f}{\partial x}\Big|_{(p,q)} Dg|_q + \frac{\partial f}{\partial y}\Big|_{(p,q)}$$

In our example above: $f(x,y) = x^2+y^2-1$, $x \neq 0$.

$\Rightarrow g(y) := \sqrt{1-y^2}$ for $x>0$ $\Rightarrow f(g(y),y) = 0$, and $\frac{\partial g}{\partial y} = \frac{-\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}} = \frac{-2y}{2x} = \frac{-y}{\sqrt{1-y^2}}$. ✓

More generally:

A surface $M \subset \mathbb{R}^3$ can be defined via $F(x,y,z)=0$, $F: U \to \mathbb{R}^3$, i.e.,

$M = \{(x,y,z) \in U : F(x,y,z) = 0\}$. Then the implicit fct. thm. tells us that if $F \in C^1(U)$ and $\frac{\partial F}{\partial z} \neq 0$, then locally the surface can be defined via the explicit equation $z = \phi(x,y)$.

Surfaces are special cases of manifolds, a concept that will be introduced in Analysis III.

For the proofs, we need an important theorem.

First, on a metric space (X,d), a map $f: X \to X$ is called a contraction if there is $0 \le c < 1$ s.t. $d(f(x),f(y)) \le c\, d(x,y)$.

A point $x^* \in X$ is called fixed point if $f(x^*) = x^*$.

Note: Suppose f is a contraction and it has two fixed points: $f(x_1) = x_1$, $f(x_2) = x_2$.

Then $d(x_1,x_2) = d(f(x_1),f(x_2)) \le c\, d(x_1,x_2)$ with $0 \le c < 1$, which implies $d(x_1,x_2) = 0$, i.e., $x_1 = x_2$.

So if a contraction has a fixed point, then it is unique.

Moreover:

Banach Fixed-Point Theorem (or: Contraction Mapping Principle):

If X is a complete metric space, then any contraction $f: X \to X$ has a unique fixed point.

Proof: See homework 5.

(Define $x_{n+1} := f(x_n)$ $\forall n$ and show that (x_n) is Cauchy $\Rightarrow$ limit x^* exist since X is complete $\Rightarrow$ $f(x^*) = f(\lim_{n\to\infty}(x_n)) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x^*$.) $\square$

f continuous. Why?

Note: This proof gives us an explicit way to construct the fixed point:

It is the limit of the sequence $x_{n+1} = f(x_n)$.

(I.e., choose some x_0, then $x_1 = f(x_0)$, $x_2 = f(x_1) = f(f(x_0))$, i.e., $x_n = f^{\circ n}(x_0)$.)

Next: Extra example not covered in the in-person class

Example: Newton's method for finding zeroes of $f(x)$.

We guess/hope that the iteration $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} =: F(x_n)$ converges to a zero of f. Suppose $f(x^*)=0$, $f'(x^*) \neq 0$. Then $F(x^*) = x^*$, i.e., x^* is a fixed point of the map F.

With the Banach Fixed-Point Theorem we could now find sufficient conditions for Newton's method to converge by constructing a suitable complete metric space X on which F maps $X \to X$ and is a contraction.

E.g.: For $f(x) = x^2 - 3$, we have $F(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2-3}{2x} = \frac{1}{2}\left(x + \frac{3}{x}\right)$.

Here $F: [\sqrt{3}, \infty) \longrightarrow [\sqrt{3}, \infty)$, i.e., we can choose $X = [\sqrt{3}, \infty)$ (which is closed, so X with the standard metric (absolute value) is indeed complete).

Is F a contraction on X?

$$d(F(x), F(y)) = |F(x) - F(y)| = \frac{1}{2}\left|\left(x + \frac{3}{x}\right) - \left(y + \frac{3}{y}\right)\right|$$

$$= \frac{1}{2}\Big|x - y + 3\underbrace{\left(\frac{1}{x} - \frac{1}{y}\right)}_{= \frac{y-x}{xy}}\Big| = \frac{1}{2}\left|(x-y)\left(1 - \frac{3}{xy}\right)\right|$$

$$\leq \frac{1}{2}|x-y| \underbrace{\left|1 - \frac{3}{xy}\right|}_{\leq 1} \leq \frac{1}{2}|x-y|$$, so yes, F is a contraction.

So, by Banach's fixed point thm., $x_{n+1} = F(x_n)$ converges to the unique fixed point $x^* = \sqrt{3}$ for any initial $x_0 \in [\sqrt{3}, \infty)$.

(In fact, for $x_0 \in (0, \sqrt{3})$, we have $x_1 = F(x_0) = \frac{1}{2}\left(x_0 + \frac{3}{x_0}\right) > \sqrt{3}$, so we could use any $x_0 > 0$ as initial point for the iteration.)