Introduction to Manifolds

Lecture 5

Balázs Szendrői, University of Oxford, Trinity term 2020

The Implicit Function Theorem in $\mathbb{R}^n$: setup

We write

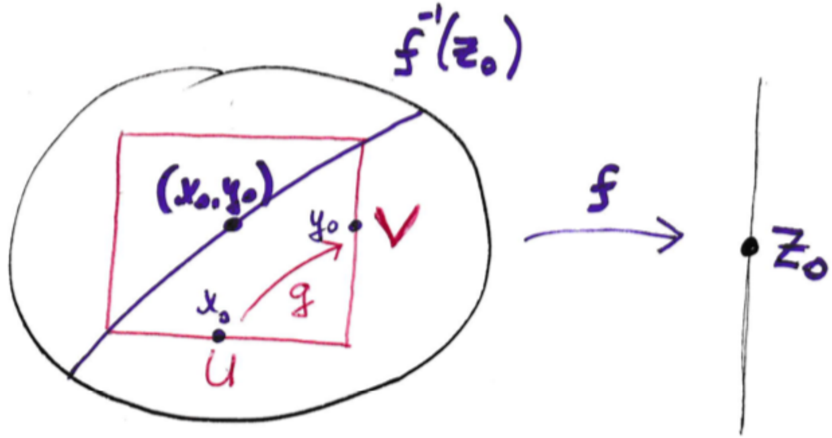
- $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m \ni (x, y) = (x_1, \ldots, x_k, y_1, \ldots, y_m)$
- $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m, (x_0, y_0) \in \Omega \subseteq \mathbb{R}^n$,
- $z_0 = f(x_0, y_0)$

We are looking for open neighbourhoods U of x_0 and V of y_0, as well as a function $g\colon U \to V$, such that

$$\forall (x, y) \in U \times V : \boxed{f(x, y) = z_0} \iff \boxed{y = g(x)}$$

The Implicit Function Theorem in $\mathbb{R}^n$: setup

$$\mathbb{R}^k \times \mathbb{R}^m \supset \Omega \xrightarrow{\ f\ } \mathbb{R}^m$$



$$f(x,y) = z_0 \iff y = g(x)$$

inside $U \times V$

The Implicit Function Theorem in $\mathbb{R}^n$: the linear case

Consider first the linear case.

Let

$$f(x, y) = Ax + By$$

with

$$A \in M_{m\times k}(\mathbb{R}), \quad B \in M_{m\times m}(\mathbb{R}).$$

Our target value is at (x_0, y_0) is

$$z_0 = Ax_0 + By_0.$$

Then the "level set" equation becomes

$$Ax + By = z_0.$$

This can be solved for y as

$$y = B^{-1}\big(z_0 - Ax\big)$$

if B is invertible.

The Implicit Function Theorem in $\mathbb{R}^n$: the linear case

So indeed we get

$$\forall (x, y) : \boxed{Ax + By = z_0} \iff \boxed{y = B^{-1}\big(z_0 - Ax\big)}$$

as long as B is invertible.

Notice that the matrix B is precisely the **matrix of partial derivatives of f with respect to the y variables**, a certain $m \times m$ submatrix of Df.

The general Theorem will be a generalization of this observation for a non-linear situation.

The Implicit Function Theorem in $\mathbb{R}^n$: separating derivatives

Consider the nonlinear case. Let $\Omega \subset \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m$.
For $f \in C^1(\Omega, \mathbb{R}^m)$, write

$$Df(x,y) = \left(D_x f(x,y), D_y f(x,y)\right),$$

where

$$D_x f(x,y) = \left(\frac{\partial f_j}{\partial x_i}\right) \in M_{m\times k}(\mathbb{R}) \qquad (j = 1,\dots,m;\ i = 1,\dots,k)$$

and

$$D_y f(x,y) = \left(\frac{\partial f_j}{\partial y_i}\right) \in M_{m\times m}(\mathbb{R}) \qquad (j = 1,\dots,m;\ i = 1,\dots,m).$$

The Implicit Function Theorem in $\mathbb{R}^n$: approximation

Then by the definition of derivative, we can write

$$\boxed{\begin{aligned} f(x,y) &= f(x_0,y_0) + D_x f(x_0,y_0)(x-x_0) + D_y f(x_0,y_0)(y-y_0) \\ &\quad + o(|(x,y)-(x_0,y_0)|) \end{aligned}}$$

where $o(h)$ is a quantity such that $o(h)/h$ tends to 0 as $h \to 0$.

If the remainder term in the second line were zero, then we would have

$$f(x,y) = z_0$$

if and only if

$$D_x f(x_0,y_0)(x-x_0) = -D_y f(x_0,y_0)(y-y_0)\,.$$

If $D_y f(x_0, y_0)$ is invertible, this is equivalent to

$$\boxed{y = y_0 - \Big(D_y f(x_0,y_0)\Big)^{-1} D_x f(x_0,y_0)(x-x_0)}$$

Hence there would exist a function $g(x)$ such that $f(x,y) = z_0$ iff $y = g(x)$, as desired.

The Implicit Function Theorem in $\mathbb{R}^n$: full statement

Theorem (The Implicit Function Theorem) Let

$$f\colon \Omega \subseteq \mathbb{R}^n = \mathbb{R}^{k+m} \to \mathbb{R}^m,$$

$f \in C^1(\Omega, \mathbb{R}^m)$. Let $(x_0, y_0) \in \Omega$ with $z_0 = f(x_0, y_0)$.

Assume that the $m \times m$ matrix $D_y f(x_0, y_0)$ is invertible.

Then there exist open neighbourhoods $U \subset \mathbb{R}^k$ of x_0 and $V \subset \mathbb{R}^m$ of y_0, and a function $g \in C^1(U, V)$ such that

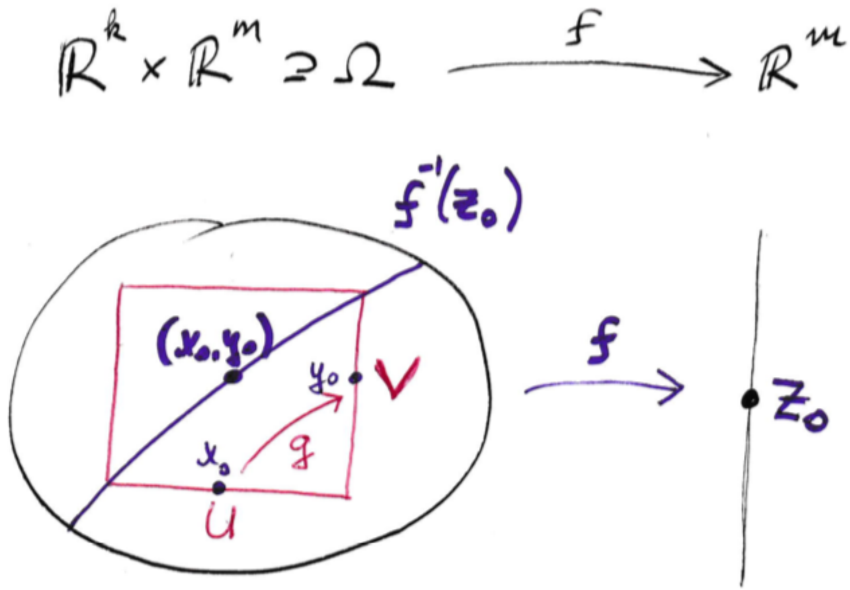
$$\{(x, y) \in U \times V \mid f(x, y) = z_0\} = \{(x, y) \mid x \in U, y = g(x)\}.$$

Furthermore

$$Dg(x_0) = -\Big(D_y f(x_0, y_0)\Big)^{-1} D_x f(x_0, y_0)\,.$$

Summary: $y \in V \subset \mathbb{R}^m$ is defined **implicitly** as a function of $x \in U \subset \mathbb{R}^k$ via the relation $f(x, y) = 0$, equivalently **explicitly** as $y = g(x)$.

The Implicit Function Theorem in $\mathbb{R}^n$



$f(x,y) = z_0 \iff y = g(x)$

inside $U \times V$

The Implicit Function Theorem in $\mathbb{R}^n$: an example

Let $h\colon \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ for $\Omega = \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}$ be given by

$$h(x, y, z) = xy - z \log y + e^{xz} - 1.$$

Consider

$$S = \{(x, y, z) \in \Omega \mid h(x, y, z) = 0\} \subset \mathbb{R}^3.$$

This is a surface in $\mathbb{R}^3$, given by a single equation.

We would like to parametrize a piece of this surface, by "two parameters". In other words, we would like to choose one coordinate, say x, and write

$$(x, y, z) \in S \iff x = g(y, z).$$

This is precisely what the Implicit Function Theorem achieves,

- in a neighbourhood of a point $(x_0, y_0, z_0) \in S$ on the surface,
- under the non-degeneracy assumption

$$\partial_x h(x_0, y_0, z_0) \neq 0.$$

The Implicit Function Theorem in $\mathbb{R}^n$: an example

$h\colon \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ be $h(x, y, z) = xy - z \log y + e^{xz} - 1$, and

$$S = \{(x, y, z) \in \Omega \,|\, h(x, y, z) = 0\} \subset \mathbb{R}^3.$$

The Jacobian matrix of h is

$$Dh(x, y, z) = (y + ze^{xz}, x - \frac{z}{y}, -\log y + xe^{xz}).$$

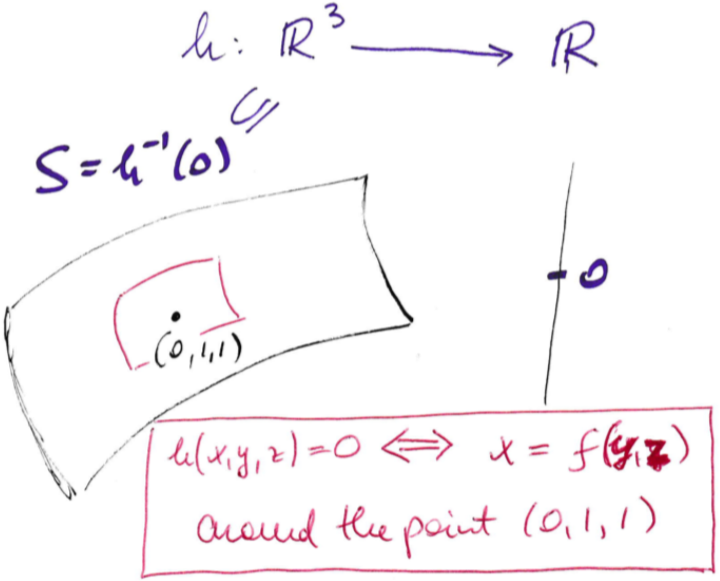
We have $(0, 1, 1) \in S$ and thus $Dh(0, 1, 1) = (2, -1, 0)$.

The Implicit Function Theorem tells us that we can represent the surface locally as $x = f(y, z)$ or $y = g(x, z)$ using a differentiable parametrization. **No formula for f or g.**

The Implicit Function Theorem in $\mathbb{R}^3$: an example

$h\colon \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ be $h(x,y,z) = xy - z\log y + e^{xz} - 1$, and

$$S = \{(x,y,z) \in \Omega \,|\, h(x,y,z) = 0\} \subset \mathbb{R}^3.$$

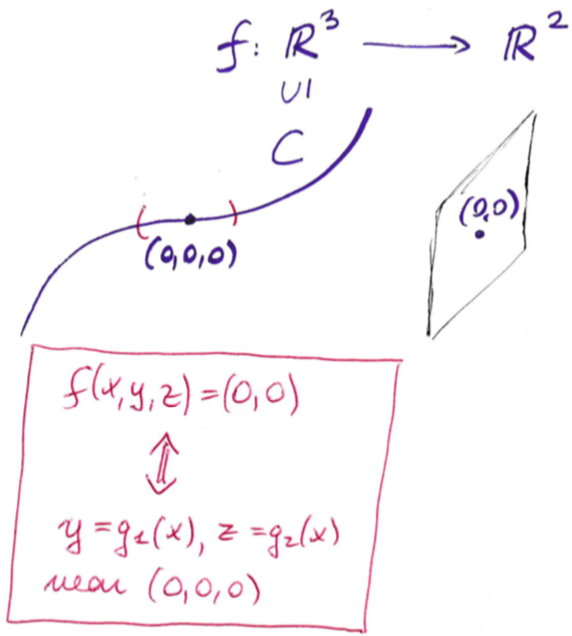


The Implicit Function Theorem in $\mathbb{R}^3$: another example

For $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by $f(x,y,z) = (x^2 - y + z^2, x^2 - 2y + z)$ and

$$C = f^{-1}(0,0) \subset \mathbb{R}^3,$$

a differentiable parametrisation $g : I \to \mathbb{R}^2$ for C exists near $(0,0,0)$.



The Inverse Function Theorem: introduction

We consider the following problem. Given a function

$$f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n,$$

we would like to understand when does there exist, locally around a point x_0, an inverse function

$$g = f^{-1}.$$

Once again, consider the linear case first. This is trivial now.
We have

$$f(x) = Ax$$

with

$$A \in M_{n\times n}(\mathbb{R}).$$

Then

$$y = Ax$$

is **invertible if and only if A is invertible**.

The Inverse Function Theorem: linearizing the problem

Consider a general

$$f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n.$$

Let $x_0 \in \Omega$, $y_0 = f(x_0)$ and assume that the Jacobian matrix $Df(x_0)$ is invertible. Then we find for general x that

$$f(x) = y_0 + Df(x_0)(x - x_0) + o(|x - x_0|)\,.$$

Now, if the remainder term were not present, then we could just invert the function by

$$y = f(x) \quad \Longleftrightarrow \quad x = x_0 + Df(x_0)^{-1}(y - y_0).$$

Example Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$.

- For $x_0 > 0$ or $x_0 < 0$ we have that f is invertible in a neighbourhood of x_0.
- For $x_0 = 0$ there is no neighbourhood of x_0 where f has an inverse. Indeed, $f'(0) = 0$ is not invertible.

Diffeomorphisms

The following concept will be convenient to use.

Definition (Diffeomorphism) Let

$$f\colon U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^n$$

with U, V domains in $\mathbb{R}^n$. We say that f is a **diffeomorphism** if f is bijective, that is there exists $f^{-1} : V \to U$, and if $f \in C^1(U, V)$ and $f^{-1} \in C^1(V, U)$.

Important caveat: $f \in C^1(U, V)$ and f invertible do not imply that f is a diffeomorphism!

Example Let $f\colon (-1, 1) \to (-1, 1)$ be given by $f(x) = x^3$.
f is bijective with inverse $g\colon (-1, 1) \to (-1, 1)$ given by $g(y) = y^{\frac{1}{3}}$.
Furthermore, $f \in C^\infty(-1, 1)$.
However, f^{-1} is not differentiable in any neighbourhood of 0.
Hence, f is **not a diffeomorphism**.

The Inverse Function Theorem

Theorem (The Inverse Function Theorem in $\mathbb{R}^n$) Let $\Omega \subseteq \mathbb{R}^n$ be open, let $f \in C^1(\Omega, \mathbb{R}^n)$ and let $x_0 \in \Omega$.

Assume that $Df(x_0)$ is invertible.

Then there exists an open neighbourhood U of x_0 such that $f(U)$ is open and

$$f|_U \colon U \to f(U)$$

is a diffeomorphism, i.e. it has a differentiable inverse $g \colon f(U) \to U$.

Furthermore, for this local inverse g of f, we have

$$Dg(f(x_0)) = \big(Df(x_0)\big)^{-1}.$$

Remark Notice that the last conclusion is identical to the conclusion of our earlier result about the derivative of the inverse, discussed in the section on the Chain Rule. The big difference is that in that result, we needed to **assume** existence and differentiability of the inverse, whereas here the existence of g **follows** from a much weaker assumption.