Analysis 2, Chapter 9

Matthijs Muis

December 2023

1 Implicit function theorem

The **implicit function theorem** enables one to study the shape of a solution set to a system of equations. In particular, we have $C^1(R^{N-k}\times\mathbb{R}^k,\mathbb{R}^k)$ functions $\Phi_1,...,\Phi_k:(x,y)\mapsto\Phi_1(x,y),...,\Phi_k(x,y)$ with $k\leq N$ and we consider S, the solution set of the system:

$$\begin{cases}\Phi_1(x,y)=0\\ \quad ...\\ \Phi_k(x,y)=0\end{cases}$$

We expect that locally, for $(a,b)\in S\subset R^{N-k}\times\mathbb{R}^k$, S could be approximated by the affine space formed by the solution set of the following system:

$$\begin{cases}\langle\nabla\Phi_1(a,b),(x-a,y-b)\rangle+\Phi_1(a,b)=0\\ \quad ...\\ \langle\nabla\Phi_k(a,b),(x-a,y-b)\rangle+\Phi_k(a,b)=0\end{cases}$$

Notice that this system also writes as

$$D_{(a,b)}\Phi(x-a,y-b)+\Phi(a,b)=0$$

This means that if we consider the k-subset of the variables $\{y_1,...,y_k\}\subset\{x_1,...,x_{N-k},y_1,...,y_k\}$ which correspond to a $k\times k$ matrix B built from a subset of the k final columns of $J_{(a,b)}f$, and B is **invertible**, we can argue that $D_{(a,b)}\Phi(x-a,y-b)+\Phi(a,b)=0$ if and only if

$$B^{-1}(J_{(a,b)}\Phi)(x-a,y-b)+B^{-1}\Phi(a,b)=0$$

, and writing $J_{(a,b)}\Phi=[A|B]$, $y=(y_1,...,y_k)$, $x=(x_1,...,x_{N-k})$, we find

$$y=-B^{-1}Ax-B^{-1}\Phi_k(a,b)$$

This intuition leads us to the **implicit function theorem**: first, we introduce some helpful notation for a **partial Jacobi matrix**:

Definition 1. *The **partial Jacobi matrix** of a function $f : \Omega \to \mathbb{R}^M$ with partial derivatives, where $\{i_1, ..., i_m \subset [M]$, $\{x_1, ..., x_n\} \subset \mathbb{R}^N \backslash \{0\}$, is defined as follows:*

$$\frac{\partial f_{i_1}, ..., f_{i_m}}{\partial x_1, ..., x_n} := \begin{pmatrix} \partial_{x_1} f_{i_1}(x) & ... & \partial_{x_n} f_{i_1}(x) \\ ... & ... & ... \\ \partial_{x_1} f_{i_m}(x) & ... & \partial_{x_n} f_{i_m}(x) \end{pmatrix}$$

*This is an $m \times n$ matrix. If it is clear from the context (which is the notation used below, in the implicit function theorem, where we name the first $N-k$ variables as x_i and the last k variables as y_j), we can also take $x_1, ..., x_n$ to be **variables** rather than **direction vectors** $x_{i_1}, ..., x_{i_n}$ (which can also be written with vectors $e_{i_1}, ..., e_{i_n}$, the standard directions).*

Theorem 1. ***Implicit Function Theorem*** *Let $f : \mathbb{R}^{N-k} \times \mathbb{R}^k \to \mathbb{R}^k$ be of class $C^1(\Omega)$, let $f(a, b) = c$ and let $\det\left(\frac{\partial f_1, ..., f_k}{\partial y_1, ..., y_k}(a, b)\right) \neq 0$ That means, equivalently, that the vectors*

$$\begin{pmatrix} \partial_{y_1} f_1(x) \\ ... \\ \partial_{y_1} f_k(x) \end{pmatrix}, ..., \begin{pmatrix} \partial_{y_k} f_1(x) \\ ... \\ \partial_{y_k} f_k(x) \end{pmatrix}$$

*Are **linearly independent**, i.e. a basis for $\mathbb{R}^k$. Then, there are **open** $X \subset \mathbb{R}^{N-k}$ and **open** $Y \subset \mathbb{R}^k$ and a C^1 function $g : X \to Y$ such that $a \in X, b \in Y$ and*

$$x \in X, y \in Y, f(x, y) = c \iff y = g(x), x \in X$$

Or,

$$\{f = c\} \cap (X \times Y) = \{(x, g(x)) : x \in X\}$$

Proof. We can show that the implicit function theorem follows from the **inverse function theorem**. In fact, the two are *equivalent*: the inverse function theorem also follows from the implicit function theorem.

Since f is of class C^1, and we have that

$$(x, y) \mapsto \det\left(\frac{\partial f_1, ..., f_k}{\partial y_1, ..., y_k}(x, y)\right)$$

Is a polynomial in the partial derivatives of f, it follows that this map is continuous, hence there is an open neighbourhood in $\mathbb{R}^N$, which we will narrow to a product of an open $A \subset \mathbb{R}^{N-k}$, and an open $B \subset \mathbb{R}^k$ such that

$$\forall (x, y) \in A \times B : \det\left(\frac{\partial f_1, ..., f_k}{\partial y_1, ..., y_k}(x, y)\right) \neq 0$$

Next, we define a new function $F : A \times B \to \mathbb{R}^{N-k} \times \mathbb{R}^k$, argue that it is a diffeomorphism and use its local inverse to define g. The only thing left to do is to pick a suitable K... We show that $F(x, y) := (x, f(x, y))$ will do (it might be a first guess, if we simply look at the dimension that we need to add to the codomain of K, but in fact it also works). Note that

$$J_{(x,y)}F = \left(\begin{array}{c|c} (Id)_{N-k} & O \\ \hline \frac{\partial f_1,...,f_k}{\partial x_1,...,x_{N-k}}(a,b) & \frac{\partial f_1,...,f_k}{\partial y_1,...,y_k}(a,b) \end{array} \right)$$

So that we get

$$\det(J_{(x,y)}F) = \det\left(\frac{\partial f_1,...,f_k}{\partial y_1,...,y_k}(x,y)\right) \neq 0$$

This means K is indeed locally a diffeomorphism (because the premises of the inverse function theorem hold) with a C^1 inverse, say that there are $X \subset A$, $Y \subset B$ open, such that $a \in X$, $b \in Y$, and $h : F(X \times Y) \to X \times Y$ such that $h \circ F = F \circ h = (Id)_{\mathbb{R}^N}$, but this means that $(h_1, ..., h_{N-k})(x, f(x, y)) = x$, in other words $h(x, y) = x$ So denote

$$h(x, y) = (x, \varphi(x, y))$$

For some function $\varphi : X \times Y \to Y$. Now define $g(x) = \varphi(x, c)$, then we get for free that g is C^1, since φ is, since h is (by the inverse function theorem). Next, we show

$$\{f = c\} \cap (X \times Y) \supset \{(x, g(x)) : x \in X\}$$

For this, simply note that $(x, g(x)) = h(x, c) \in X \times Y$ and because $f(x, g(x)) = (F_{N-k+1}, ..., F_N)(h(x, c)) = c$ To prove

$$\{f = c\} \cap (X \times Y) \subset \{(x, g(x)) : x \in X\}$$

, just note that if $(x, y) \in X \times Y$ and $f(x, y) = c$, then $y = (h_{N-k+1}, ..., h_N)(F(x, y)) = \varphi(x, f(x, y)) = \varphi(x, c) = g(x)$. □

1.1 The derivative of the implicit function

Even though the proof of the implicit function theorem does not tell us how to construct the implicit function, we can obtain additional information on its derivatives.

For any $i = 1, ..., k$ and $j = 1, ..., N-k$, we can take a partial derivative with respect to x_j on both sides of the equality

$$f_i(x, g(x)) = c_i$$

We know that the left has to be differentiable since it equals a constant function, but since on this occasion it is also a composition of two C^1-functions, we can do this using the chain rule, getting equations for the partial derivatives of g:

$$\partial_{x_j} f_i(x, g(x)) + \sum_{r=1}^{k} \partial_{y_r} f_i(x, g(x)) \partial_{x_j} g_r(x) = 0$$

This gives $k(N-k)$ equations in $k(N-k)$ unknowns, namely $J_x g$ (which is a linear map $\mathbb{R}^{N-k} \to \mathbb{R}^k$). More compactly, we can write the chain rule as follows:

$$\frac{\partial f_1, ..., f_k}{\partial x_1, ..., x_{N-k}}(x, y) + \left(\frac{\partial f_1, ..., f_k}{\partial y_1, ..., y_k}(x, y)\right) J_x g = 0$$

By the assumption $\det\left(\frac{\partial f_1,...,f_k}{\partial y_1,...,y_k}(x, y)\right) \neq 0$ in $X \times Y$, this system has a unique solution that fixes the partial derivatives of g:

$$J_x g = -\left(\frac{\partial f_1, ..., f_k}{\partial y_1, ..., y_k}(x, y)\right)^{-1} \left(\frac{\partial f_1, ..., f_k}{\partial x_1, ..., x_{N-k}}(x, y)\right)$$

1.2 Equivalence of the Implicit and Inverse Function Theorem

We used the inverse function theorem to prove the implicit function theorem. We will now show that we can also do the converse:

Proposition 1. *The inverse function theorem follows from the implicit function theorem.*

Proof. Assume the implicit function theorem and let $h : \Omega \to \mathbb{R}^N$ be such that it is of class C^1 in Ω and that $a \in \Omega$ such that $\det(J_a h) \neq 0$.

We now need to define a suitable function f with an equality in a such that its implicit function around a can be used to construct a $\mathbb{C}^1$ inverse of h. An obvious guess, since we need $h^{-1} : h(X) \to X$, which are open subsets of $\mathbb{R}^N$, is to define $f : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$. An implicit function would express the first N components in the last N, so...

$$f(x, y) = h(x) - y$$

And the system is

$$f(x, y) = 0$$

With $b = h(a)$. Now, f satisfies the assumptions of the implicit function theorem:

$$J_(a, b) f = \left(\begin{array}{c|c} J_x h & -(Id)_k \end{array} \right)$$

Which has the same determinant as $J_a h$, which is nonzero by assumption. This means that we receive a C^1 function $g : X \to Y$, where $X \ni a$, $Y \ni h(x)$, such that on $X \times Y$,

$$f(x, y) = 0 \iff x = g(y)$$

Or in other words,

$$h(x) = y \iff x = g(y)$$

This makes g a left- and right-inverse. Thus, it is the inverse function we were looking for. $\square$