Lecture 15

15.1 Inverse function theorem: Example

Recall that the inverse function theorem (14.2.1) states that if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a C^1 function and there is $a \in \mathcal{O}_n$ such that $Df(a)$ is invertible, then there exists $\widetilde{\mathcal{O}_n}$ such that $f(a) \in \widetilde{\mathcal{O}_n}$ and $f^{-1} : \widetilde{\mathcal{O}_n} \to \mathcal{O}_n$ exists and is a C^1 function and further, $D(f^{-1}) = (Df)^{-1}$ at each point. Thus, given that f satisfies the conditions, the theorem guarantees that f is locally invertible with a differentiable inverse. We discuss an important example where the theorem is used.

Consider the polar coordinate transformation,

$$x = r\cos\theta \quad y = r\sin\theta$$

We can rephrase this with the function,

$$\begin{aligned} F : \mathbb{R}_{>0} \times \mathbb{R} &\to \mathbb{R}^2 \\ F(r,\theta) &= (x,y) \end{aligned}$$

F is defined on an open set, and it is C^1 because we have,

$$\begin{aligned} \frac{\partial x}{\partial r} &= \cos\theta & \frac{\partial y}{\partial r} &= \sin\theta \\ \frac{\partial x}{\partial \theta} &= -r\sin\theta & \frac{\partial y}{\partial \theta} &= r\cos\theta \end{aligned}$$

$$\implies J_F(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \implies \det J_F(r,\theta) = r$$

and hence, $\det J_F(r,\theta)$ is non-zero in the domain we chose. The inverse function theorem then guarantees that we can express (r,θ) as a C^1 function of (x,y), locally. Further, we also have

$$DF^{-1}(x,y) = \left(DF(F^{-1}(x,y))\right)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

In other words,

$$\begin{aligned} \frac{\partial r}{\partial x} &= \cos\theta & \frac{\partial r}{\partial y} &= \sin\theta \\ \frac{\partial \theta}{\partial x} &= -\frac{1}{r}\sin\theta & \frac{\partial \theta}{\partial y} &= \frac{1}{r}\cos\theta \end{aligned}$$

which can also be verified directly.

15.2 Implicit Function Theorem

In one variable, we could differentiate functions given as $y = f(x)$. We have developed several variable calculus far enough that we can now do the same for functions given as $x_n = f(x_1, \ldots, x_{n-1})$. But what about functions that are given as $f(x, y) = 0$, that is, implicitly in both variables? We will now discuss an important theorem that will allow us to take derivatives of such functions as well without needing to solve the equation for the dependent variable.

Example 15.2.1

Let $F(x, y) = ax + by + c$. We have,

$$F(x,y) = 0 \iff ax + by + c = 0 \iff y = -\frac{a}{b}x - \frac{c}{b}$$

for all $b \neq 0$. Hence, given that $F(x, y) = 0$, we get

$$y = f(x) \text{ for a differentiable } f \iff b \neq 0 \iff \frac{\partial F}{\partial y} \neq 0$$

In other words, $\frac{\partial F}{\partial y} \neq 0 \iff F(x, f(x)) = 0$ for some differentiable f.

Example 15.2.2

Let $F(x, y) = x^2 + y^2 - 1$. We have,

$$F(x,y) = 0 \iff x^2 + y^2 = 1 \iff y^2 = 1 - x^2 \iff y = \pm\sqrt{1 - x^2}$$

but the last expression is not a function! More precisely,

$$\begin{aligned} y \geq 0 &\implies y = f_1(x) = \sqrt{1 - x^2} \\ y \leq 0 &\implies y = f_2(x) = -\sqrt{1 - x^2} \end{aligned}$$

Note that $\frac{\partial F}{\partial y} = 2y \neq 0$ for all $y \neq 0$. Hence, for $F(x_0, y_0) \neq 0$, $\frac{\partial F}{\partial y} \neq 0$, there exists a $C^1 f$ defined in a neighbourhood of x_0 such that $F(x, f(x)) = 0$ for all x in that neighbourhood.

Note

Consider $F(x, y) = 0$ and $y = f(x)$ for some $C^1 f$ such that $F(x, f(x)) = 0$. We have,

$$\begin{aligned} F(x, f(x)) &= 0 \\ \implies \frac{\mathrm{d}F}{\mathrm{d}x} &= 0 \\ \implies \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} &= 0 && \text{(Chain rule)} \\ \implies \frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \end{aligned}$$

Hence, the condition $\frac{\partial F}{\partial y} \neq 0$ is necessary for differentiating functions defined implicitly. Keeping the feeling of the above examples in mind, we now discuss the full theorem that shows that it is also sufficient.

15.2.1 Proof of the theorem

We first introduce some notation. Throughout this section, $X \in \mathbb{R}^n, Y \in \mathbb{R}^m$ so that $(X,Y) \in \mathbb{R}^{n+m}$. The point $(a,b) \in \mathbb{R}^{n+m}$ is defined with $a \in \mathbb{R}^n, b \in \mathbb{R}^m$. $\mathcal{O}$ denotes an open set in $\mathbb{R}^{n+m}$, and $F : \mathcal{O} \to \mathbb{R}^m$ has the coordinate functions $f_i(X,Y)$. Assuming $F \in C^1(\mathcal{O})$, its jacobian is

$$DF = \left(\frac{\partial f_i}{\partial x_j} \;\middle|\; \frac{\partial f_i}{\partial y_k} \right)$$

where $i,k \in \{1,\dots,m\}$ and $j \in \{1,\dots,n\}$.

Theorem 15.2.1 (Implicit Function Theorem)

Let $F \in C^1(\mathcal{O})$ and $F(a,b) = 0$. If $\det \left(\frac{\partial f_i}{\partial y_k}\right)_{m\times m} \neq 0$, then there exists an open neighbourhood $U \subset \mathbb{R}^n$ of a and a C^1 function $f : U \to \mathbb{R}^m$ such that $f(a) = b$ and $F(x, f(x)) = 0$ for all $x \in U$.

Proof. We define the function $\tilde{F} : \mathcal{O} \to \mathbb{R}^n \times \mathbb{R}^m$ as

$$\widetilde{F}(X,Y) = (X, F(X,Y))$$

As F is C^1, so is $\widetilde{F}$, and we have

$$J_{\widetilde{F}} = \begin{pmatrix} I_n & 0 \\ \dfrac{\partial f_i}{\partial x_j} & \dfrac{\partial f_i}{\partial y_k} \end{pmatrix}$$

Hence, $\det J_{\widetilde{F}}(a,b) \neq 0$ as $\det \left(\frac{\partial f_i}{\partial y_k}\right) \neq 0$, and so we can use the inverse function theorem!
By Inverse Function Theorem (14.2.1), there exists a neighbourhood $U_0 \subseteq \mathcal{O}$ of (a,b) and a neighbourhood $V_0 \subseteq \mathbb{R}^{n+m}$ of $(a,0)$ such that $\widetilde{F} : U_0 \to V_0$ has a C^1 inverse.

We shrink U_0 (also shrinking V_0 accordingly) so that $U_0 = A \times B$ where $a \in U \subseteq \mathbb{R}^n$ and $b \in U' \subseteq \mathbb{R}^m$, for open sets A, B. We also have,

$$\widetilde{F}^{-1}(X,Y) = (X, g(X,Y))$$

for some C^1 function g, from the definition of $\widetilde{F}$. Now, consider the map

$$\begin{aligned} \Pi_2 : \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^m \\ (X,Y) &\mapsto Y \end{aligned}$$

Then, $\Pi_2 \circ \widetilde{F} = F$. So we get,

$$\begin{aligned} (X, g(X,Y)) &= \widetilde{F}^{-1}(X,Y) \\ \implies F(X, g(X,Y)) &= \Pi_2(X,Y) = Y \end{aligned}$$

where $(X,Y) \in V_0, (X, g(X,Y)) \in U_0$. Hence, for $Y = 0$ and $X \in U$,

$$F(X, g(X,0)) = 0$$

Therefore, $f(X) = g(X,0), X \in U$, works. □

15.3 Solving systems of equations

Example 15.3.1

Consider the system of equations,

$$x^2 + 2y^2 + z^2 + w = 6$$
$$2x^3 + 4y^2 + z + w^2 = 9$$

We wish to know whether (z, w) can be expressed as a function of (x, y) locally. Construct the function $F : \mathbb{R}^4 \to R^2$ as $F = (f_1, f_2)$ where

$$f_1(x, y, z, w) = x^2 + 2y^2 + z^2 + w - 6$$
$$f_2(x, y, z, w) = 2x^3 + 4y^2 + z + w^2 - 9$$

Consider $\alpha = (1, -1, -1, 2) \in \mathbb{R}^4$. We have,

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial z} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial z} & \frac{\partial f_2}{\partial w} \end{pmatrix} = \begin{pmatrix} 2z & 1 \\ 1 & 2w \end{pmatrix}$$
$$\implies \det J(\alpha) = -9 \neq 0$$

Hence, by the implicit function theorem, there is a neighbourhood of $(1, -1)$ such that on that neighbourhood, $(z, w) = f(x, y)$ for some $C^1 f$.