INROADS

Real Analysis Exchange
Vol. 43(2), 2018, pp. 429–444

Oswaldo de Oliveira, Department of Mathematics, Universidade de São Paulo, São Paulo, SP 05508-090, Brasil. email: oliveira@ime.usp.br

THE IMPLICIT FUNCTION THEOREM FOR MAPS THAT ARE ONLY DIFFERENTIABLE: AN ELEMENTARY PROOF

Abstract

This article shows a very elementary and straightforward proof of the Implicit Function Theorem for differentiable maps $F(x,y)$ defined on a finite-dimensional Euclidean space. There are no hypotheses on the continuity of the partial derivatives of F. The proof employs the mean-value theorem, the intermediate-value theorem, Darboux's property (the intermediate-value property for derivatives), and determinants theory. The proof avoids compactness arguments, fixed-point theorems, and Lebesgue's measure. A stronger than the classical version of the Inverse Function Theorem is also shown. Two illustrative examples are given.

1 Introduction

The aim of this article is to present a very elementary and straightforward proof of a version of the Implicit Function Theorem that is fairly stronger than the classical version. We prove the implicit function theorem for differentiable maps $F(x,y)$, defined on a finite-dimensional Euclidean space, assuming that all the leading principal minors of the partial Jacobian matrix $\frac{\partial F}{\partial y}(x,y)$ are nowhere vanishing (these hypotheses are already enough to show the existence of an implicit solution) plus an additional non-degeneracy condition on the

Mathematical Reviews subject classification: Primary: 26B10, 26B12
Key words: Implicit function theorems, Jacobians, Transformations with several variables, Calculus of vector functions
Received by the editors December 12, 2017
Communicated by: Emma D'Aniello

matrix $\frac{\partial F}{\partial y}$ to ensure the uniqueness of the implicit solution. There are no hypotheses on the continuity of the partial derivatives of the map F.

The results in this article extend de Oliveira [1] and [2]. In de Oliveira [1] are proven the classical versions (enunciated for maps of class C^1 on an open set) of the implicit and inverse function theorems. In de Oliveira [2] is proven the implicit function theorem for maps $F(x, y)$ such that the partial Jacobian matrix $\frac{\partial F}{\partial y}(x, y)$ is only continuous at the base point.

The proof of the implicit function theorem developed in this article follows Dini's inductive approach (see [3]). Moreover, the proofs of the implicit and the inverse function theorems that we present avoid compactness arguments, fixed-point theorems, and Lebesgue's theories of measure and integration. The elementary proofs that follow rely on the intermediate-value and the mean-value theorems, both on $\mathbb{R}$, the intermediate-value property for derivatives on $\mathbb{R}$ (also known as Darboux's property), and some basic results of determinants theory.

As a corollary of the implicit function theorem shown in this article we obtain a version of the inverse function theorem that is stronger than the classical one. Two illustrative examples are given.

The inverse function theorem proved in this article is valid for differentiable maps whose derivatives are not necessarily continuous. It is worth to point out that such maps are not *strongly differentiable* (also called *strictly differentiable*). This follows from the fact that differentiability on an open set plus strong differentiability at a particular point p implies continuity of the derivative at the point p. See Nijenhuis [6] for this implication and for a proof of the inverse function theorem for strongly differentiable maps.

As is well-known, most proofs of the classical implicit and inverse function theorems start by showing the inverse function theorem and end by proving the implicit function theorem as a rather trivial consequence. In general, these proofs employ either Weierstrass's theorem on minima or the contraction mapping principle, see Krantz and Parks [5, pp. 41–52] and Rudin [7, pp. 221–228].

Regarding maps that are everywhere differentiable (their differentials may be everywhere discontinuous), a proof of the implicit function theorem is shown in Hurwicz and Richter [4], and a proof of the inverse function theorem is given in Saint Raymond [8]. While these two results are quite general, they also have proofs that are quite technical and not that easy to follow. The first of these proofs employs Brouwer's fixed-point theorem while the second relies on a good amount of topological arguments.

Much more information about the implicit function theorem, and its history, can be found in Krantz and Parks [5]. See also de Oliveira [1, 2].

Henceforth, we shall freely assume that all the functions are defined on a subset of a finite-dimensional Euclidean space.

2 Notations and Preliminaries

Apart from the intermediate-value and the mean-value theorems, both on the real line, we assume the intermediate-value theorem for derivatives on $\mathbb{R}$ (also called Darboux's property) stated right below.

Lemma 1. (Darboux's Property). *Given $f : [a, b] \to \mathbb{R}$ differentiable, the image of the derivative function is an interval.*

Given a $n \times n$ real matrix $A = (a_{ij})$, let $\det A$ be its determinant. The determinant of the $k \times k$ sub-matrix of A arising by deleting the last $n-k$ rows and the last $n-k$ columns of A is the kth order leading principal minor of A. The leading principal minors of orders 1 and n are a_{11} and $\det A$, respectively.

Given a nonempty subset X of $\mathbb{R}^n$ and a nonempty subset Y of $\mathbb{R}^m$, it is well-known that the Cartesian product $X \times Y = \{(x, y) : x \in X \text{ and } y \in Y\}$ is open in $\mathbb{R}^n \times \mathbb{R}^m$ if and only if X and Y are open sets.

Let us consider n and m, both in $\mathbb{N}$, and fix the canonical bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$, of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Given $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, both in $\mathbb{R}^n$, we have the inner product $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$ and the norm $|x| = \sqrt{\langle x, x\rangle}$. We denote the open ball centered at a point x in $\mathbb{R}^n$, with radius $r > 0$, by $B(x; r) = \{y \text{ in } \mathbb{R}^n : \ |y - x| < r\}$.

We identify a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ with the $m \times n$ matrix $M = (a_{ij})$, where $T(e_j) = a_{1j}f_1 + \cdots + a_{mj}f_m$ for each $j = 1, \ldots, n$.

In this section, Ω denotes a nonempty open subset of $\mathbb{R}^n$, where $n \geq 1$. Given a map $F\colon \Omega \to \mathbb{R}^m$ and a point p in Ω, we put $F(p) = \big(F_1(p), \ldots, F_m(p)\big)$. Let us suppose that F is differentiable at p. The Jacobian matrix of F at p is

$$JF(p) = \left(\frac{\partial F_i}{\partial x_j}(p)\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(p) & \cdots & \frac{\partial F_m}{\partial x_n}(p) \end{pmatrix}.$$

If F is a real function, then we have $JF(p) = \nabla F(p)$, the gradient of F at p.

Given p and q, both in $\mathbb{R}^n$, we denote the linear segment with endpoints p and q by $\overline{pq} = \{p + t(q - p) : 0 \leq t \leq 1\}$. The next result is a trivial corollary of the mean-value theorem on the real line and thus we omit the proof.

Lemma 2. (The mean-value theorem in several variables). *Let us consider a differentiable real function $F : \Omega \to \mathbb{R}$, with Ω open in $\mathbb{R}^n$. Let p*

and q be points in Ω such that the segment $\overline{pq}$ is within Ω. Then, there exists c in $\overline{pq}$, with $c \neq p$ and $c \neq q$, that satisfies

$$F(p) - F(q) = \langle \nabla F(c), p - q \rangle .$$

Given $F : \Omega \to \mathbb{R}$, a short computation shows that the following definition of differentiability is equivalent to the one most commonly employed. We say that F is differentiable at p in Ω if there are an open ball $B(p; r) \subset \Omega$, with $r > 0$, a vector $v \in \mathbb{R}^n$, and a vector-valued map $E : B(0; r) \to \mathbb{R}^n$ satisfying

$$F(p + h) = F(p) + \langle v, h \rangle + \langle E(h), h \rangle , \text{ for all } h \in B(0; r),$$
$$\text{where } E(0) = 0 \text{ and } E(h) \to 0 \text{ as } h \to 0.$$

3 Example and Motivation

Right below we give an example of a function $F : \mathbb{R}^2 \to \mathbb{R}^2$ so that

(i) F is differentiable everywhere,

(ii) the Jacobian matrix JF is not continuous at the origin,

(iii) the two leading principal minors of JF do not vanish near the origin,

(iv) F is invertible near the origin (proven in section 4 and again in section 5).

Example 1. *Let us consider the function*

$$F(x, y) = \begin{cases} \left(8x + x^3 \cos \frac{1}{x^2+y^2}, 8y + y^3 \sin \frac{1}{x^2+y^2}\right) & \textit{outside the origin} \\ (0, 0) & \textit{at the origin.} \end{cases}$$

The Jacobian matrix of F outside the origin is given by

$$\begin{pmatrix} 8 + 3x^2 \cos \frac{1}{x^2+y^2} + \frac{2x^4}{(x^2+y^2)^2} \sin \frac{1}{x^2+y^2} & \frac{2x^3 y}{(x^2+y^2)^2} \sin \frac{1}{x^2+y^2} \\ -\frac{2xy^3}{(x^2+y^2)^2} \cos \frac{1}{x^2+y^2} & 8 + 3y^2 \sin \frac{1}{x^2+y^2} - \frac{2y^4}{(x^2+y^2)^2} \cos \frac{1}{x^2+y^2} \end{pmatrix}.$$

On the other hand, a short computation shows that

$$JF(0, 0) = \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}.$$

Let us show that F is differentiable at the origin (and thus over the plane). Let T be the linear map associated to the matrix $JF(0,0)$. Given a non-null vector $v = (h,k)$ in the plane we have

$$\frac{F(v)-F(0)-Tv}{|v|} = \frac{\left(8h + h^3\cos\frac{1}{h^2+k^2}, 8k + k^3\sin\frac{1}{h^2+k^2}\right) - (8h, 8k)}{\sqrt{h^2+k^2}}$$

$$= \frac{\left(h^3\cos\frac{1}{h^2+k^2}, k^3\sin\frac{1}{h^2+k^2}\right)}{\sqrt{h^2+k^2}} \xrightarrow{(h,k)\to(0,0)} (0,0).$$

Thus, F is differentiable at the origin.

We claim that the four entries of $JF(x,y)$ are discontinuous at the origin. For instance, let us look the (usually called) first entry, which has three terms. The first two terms are continuous at the origin. However, the third term is not. In fact, by polar coordinates and writing $(x,y) = (r\cos\theta, r\sin\theta)$ we find

$$\frac{2x^4}{(x^2+y^2)^2}\sin\frac{1}{x^2+y^2} = 2(\cos^4\theta)\sin\frac{1}{r^2}.$$

Thus, the Jacobian matrix JF is not continuous at the origin.

At last, let us fix (x,y) with $x^2+y^2 \leq 1$. There exist six real numbers a, b, c, d, e, and f, all in $[-1,1]$, so that the absolute value of the first and the second leading principal minors of JF respectively have the format and satisfy

$$\left|\frac{\partial F_1}{\partial x}(x,y)\right| = |8 + 3a + 2b| \geq 8 - 3 - 2 \text{ and}$$

$$|\det JF(x,y)| = \left|\det\begin{pmatrix} 8+3a+2b & 2c \\ 2d & 8+3e+2f \end{pmatrix}\right| \geq 3^2 - 2^2.$$

Therefore, the two leading principal minors of JF do not vanish in the unit disk centered at the origin.

In the last section we prove that F is invertible on a neighborhood of $(0,0)$.

4 The Implicit Function Theorem

The first implicit function result we prove concerns one equation with several real variables and a differentiable real function. In its proof, we denote the variable in $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ by (x,y), with $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ and y in $\mathbb{R}$.

In the next theorem, Ω denotes a nonempty open set within $\mathbb{R}^n \times \mathbb{R}$.

The proof of Theorem 3 is taken from de Oliveira [2]. We give it here to make this presentation self-contained.

Theorem 3. *Let $F : \Omega \to \mathbb{R}$ be differentiable, with $\frac{\partial F}{\partial y}$ nowhere vanishing, and (a,b) a point in Ω such that $F(a,b) = 0$. Then, there exists an open set $X \times Y$, within Ω and containing the point (a,b), that satisfies the following.*

- *There exists a unique $g : X \to Y$ satisfying $F\big(x, g(x)\big) = 0$ for all $x \in X$.*
- *We have $g(a) = b$. The function $g : X \to Y$ is differentiable and satisfies*

$$\frac{\partial g}{\partial x_j}(x) = -\frac{\frac{\partial F}{\partial x_j}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))}, \text{ for all } x \text{ in } X, \text{ where } j = 1, \ldots, n.$$

Yet, if $\nabla F(x,y)$ is continuous at (a,b) then $\nabla g(x)$ is continuous at $x = a$.

PROOF. By considering the function $F(x + a, \frac{y}{c} + b)$, with $c = \frac{\partial F}{\partial y}(a,b)$, we may assume that $(a,b) = (0,0)$ and $\frac{\partial F}{\partial y}(0,0) = 1$. We split the proof into three parts: existence and uniqueness, continuity at the origin, and differentiability.

- **Existence and Uniqueness.** Let us choose a non-degenerate $(n+1)$-dimensional parallelepiped $X \times [-r, r]$, centered at $(0,0)$ and within Ω, whose edges are parallel to the coordinate axes and X is open. Then, the function $\varphi(y) = F(0, y)$, where y runs over $[-r, r]$, is differentiable with φ' nowhere vanishing and $\varphi'(0) = 1$. Thus, by Darboux's property we have $\varphi' > 0$ everywhere and we conclude that φ is strictly increasing. Hence, by the continuity of F and shrinking X (if necessary) we may assume that F is strictly negative at the bottom of the parallelepiped and F is strictly positive at the top of the parallelepiped. That is,

  $$F\Big|_{X \times \{-r\}} < 0 \quad \text{and} \quad F\Big|_{X \times \{r\}} > 0.$$

  As a consequence, having fixed an arbitrary x in X, the function

  $$\psi(y) = F(x, y), \text{ where } y \in [-r, r],$$

  satisfies $\psi(-r) < 0 < \psi(r)$. Hence, by the mean-value theorem there exists a point η in the open interval $Y = (-r, r)$ such that $\psi'(\eta) = \frac{\partial F}{\partial y}(x, \eta) > 0$. Therefore, by Darboux's property we have $\psi'(y) > 0$ at every y in Y. Thus, ψ is strictly increasing and the intermediate-value theorem yields the existence of a unique y, we then write $y = g(x)$, in the open interval Y such that $F(x, g(x)) = 0$.

- **Continuity at the origin.** Let δ satisfy $0 < \delta < r$. From above we have $F(0, -\delta) < 0 < F(0, \delta)$. By the continuity of F there exists a non-degenerate $(n+1)$-dimensional parallelepiped $\mathcal{X} \times [-\delta, \delta]$ centered at $(0,0)$ and contained in the parallelepiped $X \times [-r, r]$, where X is the domain of the map g, satisfying the conditions

$$F\Big|_{\mathcal{X}\times\{-\delta\}} < 0 \text{ and } F\Big|_{\mathcal{X}\times\{\delta\}} > 0.$$

Thus, given an arbitrary $x \in \mathcal{X}$ we obtain that $F(x, -\delta) < 0 < F(x, \delta)$. By employing the intermediate-value theorem and the definition of the map $g : X \to Y$ we may conclude that $g(x) \in (-\delta, \delta)$. This shows that g is continuous at the origin.

- **Differentiability.** From the differentiability of the real function F at $(0,0)$, and writing $\nabla F(0,0) = (v, 1) \in \mathbb{R}^n \times \mathbb{R}$ for the gradient of F at $(0,0)$, it follows that there are functions $E_1 : \Omega \to \mathbb{R}^n$ and $E_2 : \Omega \to \mathbb{R}$ satisfying

$$F(h,k) = \langle v, h\rangle + k + \langle E_1(h,k), h\rangle + E_2(h,k)k,$$
$$\text{where } \lim_{(h,k)\to(0,0)} E_j(h,k) = 0 = E_j(0,0), \text{ for } j = 1, 2.$$

Hence, substituting [we already proved that $g(h) \xrightarrow{h\to 0} g(0) = 0$]

$$k = g(h),$$
$$E_j\big(h, g(h)\big) = \epsilon_j(h), \text{ with } \lim_{h\to 0} \epsilon_j(h) = \epsilon_j(0) = 0 \text{ for } j = 1, 2,$$

and noticing the identity $F\big(h, g(h)\big) = 0$, for all small enough h, we find

$$\langle v, h\rangle + g(h) \ + \langle \epsilon_1(h), h\rangle + \epsilon_2(h)g(h) = 0.$$

Thus,

$$[1 + \epsilon_2(h)]g(h) = -\langle v, h\rangle - \langle \epsilon_1(h), h\rangle .$$

If $|h|$ is small enough, then we have $1 + \epsilon_2(h) \neq 0$ and we may write

$$g(h) = \langle -v, h\rangle + \langle \epsilon_3(h), h\rangle ,$$

where

$$\epsilon_3(h) = \frac{\epsilon_2(h)}{1+\epsilon_2(h)}v - \frac{\epsilon_1(h)}{1+\epsilon_2(h)} \text{ and } \lim_{h\to 0} \epsilon_3(h) = 0.$$

Therefore, g is differentiable at 0 and $\nabla g(0) = -v$.

Now, given any a' in X, we put $b' = g(a')$. Then, $g : X \to Y$ solves the problem $F\big(x, h(x)\big) = 0$, for all x in X, with the condition $h(a') = b'$. From what we have just done it follows that g is differentiable at a'.

□

Next, we prove the implicit function theorem for a finite number of equations. Let us denote the variable in $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ by (x, y), where $x = (x_1, \ldots, x_n)$ is in $\mathbb{R}^n$ and $y = (y_1, \ldots, y_m)$ is in $\mathbb{R}^m$. Given Ω an open subset of $\mathbb{R}^n \times \mathbb{R}^m$ and a differentiable map $F : \Omega \to \mathbb{R}^m$ we write $F = (F_1, \ldots, F_m)$, with F_i the ith component of F and $i = 1, \ldots, m$, and

$$\frac{\partial F}{\partial y} = \left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq m}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}.$$

We also write $\frac{\partial F}{\partial x} = \left(\frac{\partial F_i}{\partial x_k}\right)$, where $1 \leq i \leq m$ and $1 \leq k \leq n$.

Theorem 4. (The Implicit Function Theorem). *Let $F : \Omega \to \mathbb{R}^m$ be differentiable, with Ω a non-degenerate open ball within $\mathbb{R}^n \times \mathbb{R}^m$ and centered at (a, b). Let us suppose that $F(a, b) = 0$ and that all the m leading principal minors of the matrix $\frac{\partial F}{\partial y}$ are nowhere vanishing. The following are true.*

- *There exists an open set $X \times Y$, within Ω and containing (a, b), and a differentiable function $g : X \to Y$ satisfying*

$$F\big(x, g(x)\big) = 0, \text{ for all } x \in X, \quad \text{and} \quad g(a) = b.$$

- *We have*

$$Jg(x) = -\left[\frac{\partial F}{\partial y}(x, g(x))\right]^{-1}_{m\times m}\left[\frac{\partial F}{\partial x}(x, g(x))\right]_{m\times n}, \text{ for all } x \text{ in } X.$$

Let us suppose that we also have $\det\left(\frac{\partial F_i}{\partial y_j}(\xi_i)\right)_{1\leq i,j\leq m} \neq 0$, *for every point $(\xi_1, \ldots, \xi_m)$ in Ω^m. Then, the following uniqueness is true.*

- *If $h : X \to Y$ satisfies $F\big(x, h(x)\big) = 0$ for all $x \in X$, then we have $h = g$.*

PROOF. Let us split the proof into three parts: existence and differentiability, differentiation formula, and uniqueness.

- **Existence and differentiability.** We claim that the system

$$\begin{cases} F_1(x, y_1, \ldots, y_m) = 0, \\ F_2(x, y_1, \ldots, y_m) = 0, \\ \quad\vdots \\ F_m(x, y_1, \ldots, y_m) = 0, \end{cases} \quad \text{with the conditions} \quad \begin{cases} y_1(a) = b_1 \\ y_2(a) = b_2 \\ \quad\vdots \\ y_m(a) = b_m, \end{cases}$$

has a differentiable solution $g(x) = \big(g_1(x), \dots, g_m(x)\big)$ on some open set X containing a [i.e., we have $F\big(x, g(x)\big) = 0$ for all x in X and $g(a) = b$].

Let us employ induction on m, the number of equations. The case $m = 1$ follows immediately from Theorem 3.

Assuming that the claim holds for $m - 1$, let us examine the case m.

Given a pair $(x, y) = (x, y_1, \dots, y_m)$ we introduce the helpful notations $y' = (y_2, \dots, y_m)$, $y = (y_1, y')$, and $(x, y) = (x, y_1, y')$.

As a first step, we consider the equation

$$F_1(x, y_1, y') = 0, \text{ with the condition } y_1(a, b') = b_1,$$

where x and y' are independent variables and y_1 is the dependent one. Since $\frac{\partial F_1}{\partial y_1}(x, y_1, y')$ is nowhere vanishing, by Theorem 3 it follows that there exists a differentiable function $\varphi(x, y')$ on some open set [let us say, $\mathcal{X} \times \mathcal{Y}'$] containing (a, b') that satisfies

$$F_1[x, \varphi(x, y'), y'] = 0 \text{ (on } \mathcal{X} \times \mathcal{Y}') \text{ and the condition } \varphi(a, b') = b_1.$$

From Theorem 3 we see that $\varphi(x, y')$ also satisfies the $m - 1$ equations

$$-\frac{\partial \varphi}{\partial y_j}(x, y') = \frac{\frac{\partial F_1}{\partial y_j}[x, \varphi(x, y'), y']}{\frac{\partial F_1}{\partial y_1}[x, \varphi(x, y'), y']}, \text{ for all } j = 2, \dots, m. \tag{1}$$

As a second step, we look at solving the system with $m - 1$ equations

$$\begin{cases} F_2[x, \varphi(x, y'), y'] = 0 \\ \qquad\vdots \\ F_m[x, \varphi(x, y'), y'] = 0 \end{cases}, \text{ with the condition } \ y'(a) = b'.$$

Here, x is the independent variable while y' is the dependent variable. Let us define $\mathcal{F}_i(x, y') = F_i[x, \varphi(x, y'), y']$, with $i = 2, \dots, m$, and write

$$\mathcal{F} = (\mathcal{F}_2, \dots, \mathcal{F}_m).$$

Evidently, the map $\mathcal{F}$ is differentiable. In order to employ the induction hypothesis, let us show that all the $m-1$ leading principal minors of the partial Jacobian matrix $\frac{\partial \mathcal{F}}{\partial y'}$ are nowhere vanishing.

Let us consider the $(k-1)$th order leading principal minor (a general one)

$$\begin{vmatrix} \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_2}+\frac{\partial F_2}{\partial y_2} & \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_3}+\frac{\partial F_2}{\partial y_3} & \cdots & \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_k}+\frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_2}+\frac{\partial F_3}{\partial y_2} & \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_3}+\frac{\partial F_3}{\partial y_3} & \cdots & \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_k}+\frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_2}+\frac{\partial F_k}{\partial y_2} & \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_3}+\frac{\partial F_k}{\partial y_3} & \cdots & \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_k}+\frac{\partial F_k}{\partial y_k} \end{vmatrix}.$$

It is clear that we have $2 \leq k \leq m$ and thus $1 \leq k-1 \leq m-1$.

Developing this determinant by the columns and then canceling the everywhere vanishing determinants we are left with (a sum of k determinants)

$$\begin{aligned} \det\left(\frac{\partial \mathcal{F}_i}{\partial y_j}\right)_{2\leq i,j\leq k} &= \begin{vmatrix} \frac{\partial F_2}{\partial y_2} & \frac{\partial F_2}{\partial y_3} & \cdots & \frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_2} & \frac{\partial F_3}{\partial y_3} & \cdots & \frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_2} & \frac{\partial F_k}{\partial y_3} & \cdots & \frac{\partial F_k}{\partial y_k} \end{vmatrix} \\ &+ \begin{vmatrix} \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_2} & \frac{\partial F_2}{\partial y_3} & \cdots & \frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_2} & \frac{\partial F_3}{\partial y_3} & \cdots & \frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_2} & \frac{\partial F_k}{\partial y_3} & \cdots & \frac{\partial F_k}{\partial y_k} \end{vmatrix} \\ &+ \begin{vmatrix} \frac{\partial F_2}{\partial y_2} & \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_3} & \frac{\partial F_2}{\partial y_4} & \cdots & \frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_2} & \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_3} & \frac{\partial F_3}{\partial y_4} & \cdots & \frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_2} & \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_3} & \frac{\partial F_k}{\partial y_4} & \cdots & \frac{\partial F_k}{\partial y_k} \end{vmatrix} \\ &+ \cdots + \begin{vmatrix} \frac{\partial F_2}{\partial y_2} & \cdots & \frac{\partial F_2}{\partial y_{k-1}} & \frac{\partial F_2}{\partial y_1}\frac{\partial \varphi}{\partial y_k} \\ \frac{\partial F_3}{\partial y_2} & \cdots & \frac{\partial F_3}{\partial y_{k-1}} & \frac{\partial F_3}{\partial y_1}\frac{\partial \varphi}{\partial y_k} \\ \vdots & & \vdots & \vdots \\ \frac{\partial F_k}{\partial y_2} & \cdots & \frac{\partial F_k}{\partial y_{k-1}} & \frac{\partial F_k}{\partial y_1}\frac{\partial \varphi}{\partial y_k} \end{vmatrix} \end{aligned}$$

Thus, we obtain (keeping track of $\frac{\partial\varphi}{\partial y_j}$ for j even and also for j odd)

$$\det\left(\frac{\partial \mathcal{F}_i}{\partial y_j}\right)_{2\le i,j\le k} = \begin{vmatrix} 1 & -\frac{\partial\varphi}{\partial y_2} & -\frac{\partial\varphi}{\partial y_3} & \cdots & -\frac{\partial\varphi}{\partial y_k} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \frac{\partial F_2}{\partial y_3} & \cdots & \frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_1} & \frac{\partial F_3}{\partial y_2} & \frac{\partial F_3}{\partial y_3} & \cdots & \frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1} & \frac{\partial F_k}{\partial y_2} & \frac{\partial F_k}{\partial y_3} & \cdots & \frac{\partial F_k}{\partial y_k} \end{vmatrix}.$$

The already remarked identity $-\frac{\partial\varphi}{\partial y_j} = \frac{\partial F_1}{\partial y_j}/\frac{\partial F_1}{\partial y_1}$ [see formula (1)] leads to the collection of identities

$$\det\left(\frac{\partial \mathcal{F}_i}{\partial y_j}\right)_{2\le i,j\le k} = \frac{1}{\frac{\partial F_1}{\partial y_1}} \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \frac{\partial F_1}{\partial y_3} & \cdots & \frac{\partial F_1}{\partial y_k} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \frac{\partial F_2}{\partial y_3} & \cdots & \frac{\partial F_2}{\partial y_k} \\ \frac{\partial F_3}{\partial y_1} & \frac{\partial F_3}{\partial y_2} & \frac{\partial F_3}{\partial y_3} & \cdots & \frac{\partial F_3}{\partial y_k} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1} & \frac{\partial F_k}{\partial y_2} & \frac{\partial F_k}{\partial y_3} & \cdots & \frac{\partial F_k}{\partial y_k} \end{vmatrix}, \ k = 2, \ldots, m.$$

Hence, all the $m-1$ leading principal minors of $\frac{\partial\mathcal{F}}{\partial y'}$ are nowhere vanishing. It is worth mentioning that the right above collection of identities depends on all the m leading principal minors of the matrix $\frac{\partial F}{\partial y}$.

Thus, by induction hypothesis there exists a differentiable function ψ defined on an open set X containing a [with $\psi(X)$ within $\mathcal{Y}'$] that satisfies

$$F_i[x, \varphi\big(x, \psi(x)\big), \psi(x)\big] = 0, \text{ for all } x \text{ in } X, \text{ for all } i = 2, \ldots, m,$$
$$\text{and the condition } \psi(a) = b'.$$

Clearly, we also have $F_1\big[x, \varphi\big(x, \psi(x)\big), \psi(x)\big] = 0$, for all x in X. Defining

$$g(x) = \big(\varphi(x, \psi(x)), \psi(x)\big), \text{ where } x \in X,$$

we obtain $F[x, g(x)] = 0$, for every x in X, with g differentiable on X, and also the identity $g(a) = \big(\varphi(a, b'), b'\big) = (b_1, b') = b$.

- **Differentiation formula.** Differentiating $F[x, g(x)] = 0$ we find

$$\frac{\partial F_i}{\partial x_k} + \sum_{j=1}^{m} \frac{\partial F_i}{\partial y_j}\frac{\partial g_j}{\partial x_k} = 0, \text{ with } 1 \le i \le m \text{ and } 1 \le k \le n.$$

In matricial form, we write $\frac{\partial F}{\partial x}\big(x,g(x)\big)+\frac{\partial F}{\partial y}\big(x,g(x)\big)Jg(x)=0$.

- **Uniqueness.** If $h: X \to Y$ and x in X satisfy $F(x,h(x))=0$, by employing the mean-value theorem in several variables (Lemma 2) to each component of the map F, we conclude that there exist m points $c_1,\ldots,c_m$, all in the open ball Ω (a convex set) but possibly distinct points, satisfying

$$0=F\big(x,h(x)\big)-F\big(x,g(x)\big)$$
$$=\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(c_1) & \cdots & \frac{\partial F_1}{\partial y_m}(c_1) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1}(c_m) & \cdots & \frac{\partial F_m}{\partial y_m}(c_m) \end{pmatrix}\begin{pmatrix} h_1(x)-g_1(x) \\ \vdots \\ h_m(x)-g_m(x) \end{pmatrix}.$$

The hypothesis $\det\big(\frac{\partial F_i}{\partial y_j}(\xi_i)\big)\neq 0$ for all m-tuple $(\xi_1,\ldots,\xi_m)$ inside Ω^m evidently implies the inequality $\det\big(\frac{\partial F_i}{\partial y_j}(c_i)\big)\neq 0$. Thus, $h(x)=g(x)$.

□

Example 2. *(This example illustrates the implicit function theorem and the proof of the inverse function theorem, which is proven in the next section.) Let us take the map $F:\mathbb{R}^2\to\mathbb{R}^2$ given in Example 1. Then we define*

$$G(u,v,x,y)=F(x,y)-(u,v),\ \textit{where } (u,v,x,y)\in\mathbb{R}^4.$$

We already saw that F is differentiable. Thus $G:\mathbb{R}^4\to\mathbb{R}^2$ is differentiable and, following a standard Calculus notation, we may write

$$\frac{\partial G}{\partial(x,y)}(u,v,x,y)=\begin{pmatrix} \frac{\partial G_1}{\partial x} & \frac{\partial G_1}{\partial y} \\ \frac{\partial G_2}{\partial x} & \frac{\partial G_2}{\partial y} \end{pmatrix}=\begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix}=JF(x,y).$$

From Example 1 it immediately follows that $G(0,0,0,0)=(0,0)$ and

(i) the partial Jacobian matrix $\frac{\partial G}{\partial(x,y)}$ is not continuous at the origin,

(ii) the two leading principal minors of $\frac{\partial G}{\partial(x,y)}$ do not vanish near the origin.

Let (ξ_1,ξ_2) in $\mathbb{R}^4\times\mathbb{R}^4$ be such that $|\xi_1|\leq 1$ and $|\xi_2|\leq 1$. Once more following Example 1, it is not difficult to see that the determinant

$$\mathcal{D}=\det\begin{pmatrix} \frac{\partial G_1}{\partial x}(\xi_1) & \frac{\partial G_1}{\partial y}(\xi_1) \\ \\ \frac{\partial G_2}{\partial x}(\xi_2) & \frac{\partial G_1}{\partial x}(\xi_2) \end{pmatrix}$$

has the format

$$\det \begin{pmatrix} 8+3a+2b & 2c \\ 2d & 8+3e+2f \end{pmatrix},$$

with all the six real numbers a, b, c, d, e, and f belonging to $[-1,1]$, and satisfies

$$|\mathcal{D}| \geq 3^2 - 2^2 \text{ for all } (\xi_1, \xi_2) \in D(0,1) \times D(0,1),$$

where $D(0,1)$ is the unit (closed) disk centered at the origin of $\mathbb{R}^4$.

Therefore, by Theorem 4 we may conclude that there exists a neighborhood $U \times Y$ of $(0,0,0,0) \in \mathbb{R}^4$, with $U \subset \mathbb{R}^2$ and $Y \subset \mathbb{R}^2$, such that there exists a unique differentiable map $H : U \to Y$ satisfying

$$\begin{aligned} &G\big(u, v, H(u,v)\big) = (0,0) \text{ for all } (u,v) \in U, \\ &H(0,0) = (0,0). \end{aligned}$$

We also have $F\big(H(u,v)\big) = (u,v)$ for all $(u,v) \in U$.

5 The Inverse Function Theorem

Theorem 5. (The Inverse Function Theorem). *Let $F : \Omega \to \mathbb{R}^n$ be a differentiable map, with Ω a non-degenerate open ball within $\mathbb{R}^n$ and centered at the point x_0. Let us suppose that all the n leading principal minors of $JF(x)$ are nowhere vanishing. We also suppose $\det\big(\frac{\partial F_i}{\partial x_j}(\xi_i)\big)_{1\leq i,j\leq n} \neq 0$ for every point $(\xi_1, \ldots, \xi_n)$ inside Ω^n. Under such conditions, there exist an open set X containing x_0, an open set Y containing $y_0 = F(x_0)$, and a differentiable map $G : Y \to X$ satisfying*

$$F\big(G(y)\big) = y \text{ for all } y \in Y, \text{ and } G\big(F(x)\big) = x \text{ for all } x \in X.$$

In addition, we have

$$JG(y) = JF\big(G(y)\big)^{-1} \text{ for all } y \text{ in } Y.$$

PROOF. Let us split it into two parts: injectivity of F and existence of G.

- **Injectivity of F.** Let us suppose that $F(p) = F(q)$, with p in Ω and q in Ω. The mean-value theorem in several variables (Lemma 2) yields n points $c_1, \ldots, c_n$, all in the ball Ω but possibly distinct points, satisfying

$$0 = F(p) - F(q) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(c_1) & \cdots & \frac{\partial F_1}{\partial x_n}(c_1) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(c_n) & \cdots & \frac{\partial F_n}{\partial x_n}(c_n) \end{pmatrix} \begin{pmatrix} p_1 - q_1 \\ \vdots \\ p_n - q_n \end{pmatrix}.$$

  The hypotheses imply $\det\big(\frac{\partial F_i}{\partial x_j}(c_i)\big) \neq 0$. Thus, $p = q$.

- **Existence of G.** The map

$$\Phi(y,x) = F(x) - y, \text{ where } (y,x) \in \mathbb{R}^n \times \Omega,$$

is differentiable and $\Phi(y_0, x_0) = 0$. From the hypotheses it follows that all the n leading principal minors of $\frac{\partial \Phi}{\partial x}(y,x) = JF(x)$ are nowhere vanishing in $\mathbb{R}^n \times \Omega$ and

$$\det\left(\frac{\partial \Phi_i}{\partial x_j}(\eta_i, \xi_i)\right) = \det\left(\frac{\partial F_i}{\partial x_j}(\xi_i)\right) \neq 0,$$

for every n-tuple $\big((\eta_1,\xi_1),\ldots,(\eta_n,\xi_n)\big)$ inside $(\mathbb{R}^n \times \Omega)^n$. The Implicit Function Theorem (Theorem 4) guarantees an open set Y containing y_0 and a differentiable map $G : Y \to \Omega$ satisfying

$$F\big(G(y)\big) = y, \text{ for all } y \text{ in } Y.$$

Thus, G is bijective from Y to $X = G(Y)$ and F is bijective from X to Y. We also have $X = F^{-1}(Y)$. Since F is continuous, the set X is open (and contains x_0).

Putting $F(x) = \big(F_1(x),\ldots,F_n(x)\big)$ and $G(y) = \big(G_1(y),\ldots,G_n(y)\big)$ and differentiating $\big(F_1(G(y)),\ldots,F_n(G(y))\big)$ we find

$$\sum_{k=1}^{n} \frac{\partial F_i}{\partial x_k}\frac{\partial G_k}{\partial y_j} = \frac{\partial y_i}{\partial y_j} = \begin{cases} 1, \text{ if } i = j, \\ 0, \text{ if } i \neq j. \end{cases}$$

□

6 Remarks

Remark 1. *It is clear that the map in Example 1, section 3 (Example and Motivation), satisfy the conditions of the above inverse function theorem and is thus invertible, with differentiable inverse map, on a neighborhood of the origin.*

Remark 2. *It is not difficult to see that Theorem 4 implies the implicit function theorem for a differentiable function $F : \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, with $F(a,b) = 0$ and Ω an open set, whose partial Jacobian matrix $\frac{\partial F}{\partial y}(x,y)$ is continuous at the base point (a,b) and $\det \frac{\partial F}{\partial y}(a,b) \neq 0$. In fact, by a linear change of coordinates in the y variable, we may assume $\frac{\partial F}{\partial y}(a,b) = I$, with I the $m \times m$ identity matrix. Thus, on some open neighborhood of (a,b), we have $\det\big(\frac{\partial F_i}{\partial y_j}(\xi_{ij})\big)_{1 \le i,j \le k} \neq 0$ for all ξ_{ij} in this neighborhood, where $1 \le i,j \le k$, for each $k = 1,\ldots,m$.*

Remark 3. *Similarly, Theorem 5 implies the inverse function theorem for a differentiable function* $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$*, with* $F(x_0) = y_0$ *and* Ω *an open set in* $\mathbb{R}^n$*, whose Jacobian matrix* $JF(x)$ *is continuous at* x_0 *and* $\det JF(x_0) \neq 0$*.*

7 Conclusion

The author hopes that these elementary proofs of the implicit and the inverse function theorems, both proofs quite dependable on "linear type arguments" and fairly free of "the handling of inequalities", may contribute to a better understanding of these two fundamental results.

Acknowledgments. The author is grateful to Professor P. A. Martin for discussions that lead to the Example 1 given in this article, and to Professor O. O. Luciano for his comments on [6]. The author is very thankful to the referee for his/her many suggestions and kind words.

References

[1] O. R. B. de Oliveira, *The implicit and the inverse function theorems: easy proofs*, Real Anal. Exchange, **39(1)** (2013-14), 207–218.

[2] O. R. B. de Oliveira, *The implicit function theorem when the partial Jacobian matrix is only continuous at the base point*, Real Anal. Exchange, **41(2)** (2016), 377–388.

[3] U. Dini, *Lezione di Analisi Infinitesimale*, **1**, Pisa, 1907, 197–241. (Italian).

[4] L. Hurwicz and M. K. Richter, *Implicit functions and diffeomorphisms without* C^1, Adv. Math. Econ., **5** (2003), 65–96.

[5] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem - History, Theory, and Applications*, Birkhäuser, Boston, 2002.

[6] A. Nijenhuis, *Strong derivatives and inverse mappings*, Amer. Math. Monthly, **81(9)** (1974), 969–980.

[7] W. Rudin, *Principles of Mathematical Analysis*. Third edition. McGraw-Hill, Singapore, 1976.

[8] J. Saint Raymond, *Local inversion for differentiable functions and the Darboux property*, Mathematika, **49** (2002), 141–158.