UNIT 6 IMPLICIT AND INVERSE FUNCTION THEOREM

Structure

6.1 INTRODUCTION

In this unit, we state two very important theorems — inverse function theorem and implicit function theorem. The inverse function theorem discusses the invertibility of transformations from $\mathbf{R}^n$ to $\mathbf{R}^n$ with the help of Jacobians. The invertibility of transformations from one Euclidean space $\mathbf{R}^n$ to another Euclidean space $\mathbf{R}^m$ plays an important role in finding solutions to a system of equations. In many situations it happens that it is difficult to find a solution to the system of equations explicitly. Now the inverse function theorem gives us certain criteria to tackle this problem.

The implicit function theorem in its simplest form deals with an equation of the form $f(x, y) = 0$. The problem is to decide whether this equation determines x as a function of y uniquely or vice-versa. The problem assumes a more general form when we have a system of several equations involving several variables and we ask whether we can solve these equations for some of the variables in terms of the remaining variables. The implicit function theorem gives some conditions by which we can find a solution to this problem.

Objectives

After studying this unit, you should be able to

- identify locally invertible functions;
- state and apply inverse function theorem;
- state and apply implicit function theorem.

6.2 INVERSE FUNCTION THEOREM

In this section, we shall discuss Inverse Function Theorem for functions from $\mathbf{R}^n$ to $\mathbf{R}^n$.

We shall begin this by reviewing the situation from the single variable case.

Suppose f is a real-valued, continuously differentiable function defined on some open subset E of R. If for any point $x_0 \in E, f'(x_0) \neq 0$, then f' is not zero in $I =]x_0 - \delta, x_0 + \delta[\subset E$ for a suitable $\delta > 0$. In fact, $f'(x)$ has the sign of $f'(x_0)$

in I. If $f'(x_0) > 0$, then f(x) is strictly increasing in I, and if $f'(x_0) < 0$ then f(x) is strictly decreasing in I. In any case, f is one-one on I. Clearly, f(I) is an open interval containing $f(x_0)$. Thus, the function $f : I \to f(I)$ is one-one and onto and hence is invertible on I. Moreover, you may recall that the function $f^{-1} : f(I) \to I$ is differentiable at $f(x_0)$ and

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}. \qquad (1)$$

Thus, if $f'(x) \neq 0$ for every $x \in D$, then Eqn.(1) given above holds for every point of E.

A similar result is true for functions from $\mathbf{R}^n$ to $\mathbf{R}^n$ and is known as "the inverse function theorem". Before we state the theorem, we give some preliminaries.

Definition 1: A function f with domain $E \subset \mathbf{R}^n$ and range $E^* \subset \mathbf{R}^n$ is said to be **invertible** on E if there exists a function $g : E^* \to E$ such that $g(f(\mathbf{x})) = \mathbf{x}$ and $f(g(\mathbf{u})) = \mathbf{u}$ for every $\mathbf{x} \in E$ and $\mathbf{u} \in E^*$.

Recall that $f : E \to E^*$ is invertible on E if and only if f is one-one (it is already onto). The function g in the definition above is called the **inverse** of f. It is usually denoted by f^{-1}. We also note that if in addition, f is linear, then f^{-1} is also linear. This follows from the result that if a linear transformation on $\mathbf{R}^n$ is invertible, then the inverse transformation is also linear.

You might have seen examples of functions on $\mathbf{R}^2$ which are invertible. One such example is given below:

Example 1: Let E be a subset of $\mathbf{R}^2$ consisting of all pairs (r, θ) with $r > 0$ and $0 < \theta < \pi$. Define a function f on E by

$$f(r, \theta) = (f_1(r, \theta), f_2(r, \theta)).$$

where $f_1(r, \theta) = r\cos\theta, f_2(r, \theta) = r\sin\theta.$

We first note that the image of E, say E^*, under this map is the upper half-plane consisting of all (x, y) such that $y > 0 (y > 0$ because $r > 0$ and $0 < \theta < \pi)$. See Fig.1.

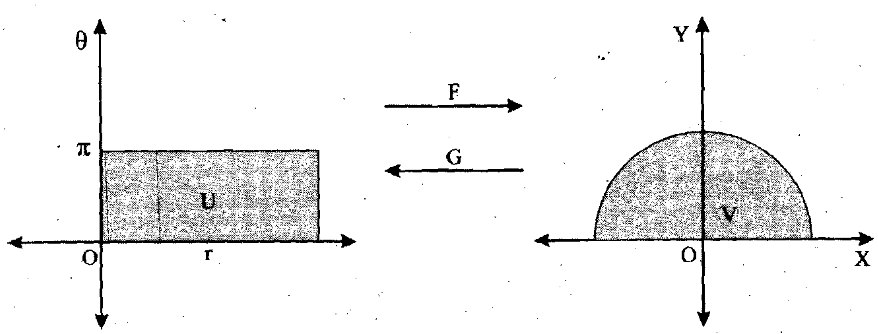

Fig. 1

Solving for r and θ, we get

$$r = \sqrt{x^2 + y^2} \text{ and } \theta = \cos^{-1}\frac{x}{r}$$

Then the inverse map $g : E^* \to E$ is given by

$$g(x, y) = \left(\sqrt{x^2 + y^2}, \cos^{-1}\frac{x}{r}\right)$$

This shows that f is invertible on E. But you can note that f is not invertible on the whole of $\mathbf{R}^2$ since f is not one-one on $\mathbf{R}^2$.

* * *

Next, we shall consider another example.

Example 2: Let f be the function defined from $\mathbf{R}^2$ to $\mathbf{R}^2$ by

$$f(x, y) = (x^2 - y^2, 2xy).$$

Note that this function is not one-one in $\mathbf{R}^2$ since $f(1, 1) = f(-1, -1) = (2, 0)$. Therefore, it is not invertible.

Now we shall check whether this function is invertible on any subsets of $\mathbf{R}^2$.

For that let us consider the open set $E \subset \mathbf{R}^2$ given by

$$E = \{(x, y) | x > 0\}$$

Now we shall check the invertibility of f on E. We shall first note that f restricted to E denoted by $f|_E$ is one-one. Next we shall show that the range of $f|_E$ is E^* which is given by

$$\begin{aligned} E^* &= \{(u, v) | v > 0 \text{ if } u = 0\} \\ &= \mathbf{R}^2 - \text{negative y-axis} \end{aligned}$$

For that we observe that if $u = 0$, then y has to be zero, because $u = 2xy$ and $x > 0$ in E. In that case $v = x^2 > 0$. Thus, no point on the negative y-axis can be the image of an element of E under f.

Now we shall try to find an inverse of f on E^*

Suppose that $u \neq 0$. This implies that $2xy \neq 0$ which in turn implies that $x \neq 0, y \neq 0$.

Also $y = \dfrac{u}{2x}$.

$\therefore v = x^2 - \dfrac{u^2}{4x^2}$. This implies that

$$4x^4 - 4x^2v - u^2 = 0$$

$$\Rightarrow x^2 = \frac{4v + \sqrt{16v^2 + 16u^2}}{8}$$

We do not consider the other root as it will mean that $x^2 < 0$.

$$\therefore x^2 = \frac{v + \sqrt{v^2 + u^2}}{2}$$

$$\therefore x = \left[\frac{v + \sqrt{v^2 + u^2}}{2}\right]^{1/2}$$

$$\Rightarrow y = u\left[2v + 2\sqrt{v^2 + u^2}\right]^{-1/2}$$

This shows that the inverse function f^{-1} is given by

$$f^{-1}(u, v) = (f_1^{-1}(u, v), f_2^{-1}(u, v))$$

where

$$f_1^{-1}(u,v) = \left[\frac{v+\sqrt{u^2+v^2}}{2}\right]^{1/2}$$

$$f_2^{-1}(u,v) = u\left[2v+2\sqrt{u^2+v^2}\right]^{-1/2}$$

Similarly, if we consider E′ given by

$$E' = \{(x,y) | x < 0\},$$

then we can show that f restricted to E is also invertible on E′.

* * *

Let us now closely look at Example 1 and 2. In the case of Example 1, you can notice that f is invertible at all points (r, θ) such that $r > 0$ and $0 < \theta < \pi$. But when we take the point (0, 0), then any neighbourhood of this point contains points $(0, \theta), \theta \neq 0$ which are mapped to (0, 0) i.e. f is not one-one in any neighbourhood of (0, 0) and, therefore, not invertible. In this case, we say that f is locally invertible. Same phenomena you can observe for Example 2 also.

In the light of this we give the following definition.

Definition 2: Let $f : E \to \mathbf{R}^n, E \subset \mathbf{R}^n$. We say that f is locally invertible at a point $a \in E$ if there exist a neighbourhood N of a contained in E and a neighbourhood N^* of f(a), such that

i) $f(N) = N^*$

ii) f is $1-1$ on N.

From Example 1 you can observe that the function f given there is locally invertible at all points of $\mathbf{R}^2$ except (0, 0).

We are now ready to state the inverse function theorem which provides a sufficient criterion for the local invertibility of a function.

Before that we recall that for any differentiable function f defined on an open subset $E \subseteq \mathbf{R}^n$ to $\mathbf{R}^m$, the Jacobian matrix is given by

$$Jf = \begin{bmatrix} D_1f_1 & \dots & D_nf_n \\ \vdots & \dots & \vdots \\ D_1f_n & \dots & D_nf_n \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \dots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

is called the Jacobian matrix of f. Its determinant at any point $x \in \mathbf{R}^n$ is called the Jacobian of f and denoted by $J_f(x)$. Sometimes the notation

$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}.$$

is also used to denote the Jacobian of f especially when the point at which it is calculated is not referred.

Theorem 1: (Inverse Function Theorem): Let f be a C^1 function defined from an open set $E \subseteq \mathbf{R}^n$ to $\mathbf{R}^n$. If for some $\mathbf{x_0} \in E$, the Jacobian $J_f(x_0) \neq 0$, then there exists an open set E_0 containing $\mathbf{x_0}$ such that

i) The restriction of f to E_0 denoted by $f|_{E_0}$ is one-one.

ii) The set $f(E_0)$ is open.

iii) The inverse of the function $(f|_{E_0})$ denoted by $(f|_{E_0})^{-1}$ is also a C^1-function.

iv) $(f|_{E_0}^{-1})'(u) = [f'(x)]^{-1}(u)$, if $u = f(x), x \in E_0$.

We have omitted the proof of this theorem. Now we shall illustrate the theorem with some examples and also discuss some consequences of the theorem.

Remark: From your undergraduate linear algebra course you might be already knowing that the determinant of the inverse L^{-1} of a linear transformation L is $(\det L)^{-1}$ where "det L" denotes the determinant of L. Therefore, the Part (iv) of the Inverse Function Theorem implies that

$$J_{f^{-1}_{|E_0}}(u) = \frac{1}{J_f(x)}, u = f|_{E_0}(x)$$

This means that given any point x_0, there exists a neighbourhood of that point such that the Jacobian of the inverse function $f|_{E_0}^{-1}$ agrees with the inverse of the Jacobian of the derivative f' for all points in the neighbourhood.

We shall omit the proof of this theorem. But we shall illustrate the theorem by giving some examples.

Example 3: Let f be the function defined from $\mathbf{R}^2$ to $\mathbf{R}^2$ by

$$f(x, y) = (x^2 - y^2, 2xy).$$

We shall now check the local invertibility of this function by applying the inverse function theorem.

For that let us consider the open set $E \subset \mathbf{R}^2$ given by

$$E = \{(x, y) | x > 0\}.$$

Let $(x_0, y_0) \in E_0$

Now we shall check the theorem for the set E.

We shall first note that the function f is C^1 on E since its component functions are continuously differentiable on E. Also $f|_E$ is one-one.

Now we compute $J_f(x, y)$. For that we write $f(x, y) = (f_1(x, y), f_2(x, y))$ where $f_1(x, y) = x^2 - y^2$ and $f_2(x, y) = 2xy$. Then

$$J_f(x, y) = \begin{vmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{2f_2}{\partial y} \end{vmatrix} = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4(x^2 + y^2)$$

Therefore, $J_f(x, y) \neq 0 \; \forall \; (x, y) \in E$.

This shows that f satisfies all the requirements of the inverse function theorem for all points of E. Therefore, by the theorem f is locally invertible at all points of E. Also f satisfies (i), (ii), (iii) and (iv) of the theorem. Now we shall verify (i), (ii), (iii) for the given function f. We shall first note that $f|_E$ is one-one. Next we have to show that the range set f(E) is open. We have already shown that the range f(E) given by E^*

$$\begin{aligned} E^* &= \{(u,v) | v > 0 \text{ if } u = 0\} \\ &= \mathbf{R}^2 - \text{ negative y-axis} \end{aligned}$$

is the range of f.

We also observe that E^* is open. Thus, condition (ii) of the inverse function theorem is verified.

To check the condition (iii) we note from Example 2 that the inverse functions of f is given by $f^{-1}(u,v) = (f_1^{-1}(u,v), f_2^{-1}(u,v))$ where f_1 and f_2 are given by

$$f_1^{-1}(u,v) = \left[\frac{v + \sqrt{u^2+v^2}}{2}\right]^{1/2}$$

$$f_2^{-1}(u,v) = u\left[2v + 2\sqrt{u^2+v^2}\right]^{-1/2}$$

Since f_1 and f_2 are C^1-functions, we get that f^{-1} is also a C^1-function. Thus the condition (iii) is satisfied.

Now the inverse function also asserts that the derivative of the inverse function f^{-1} agrees with the inverse of the derivative of f on E.

* * *

Here is another example.

Example 4: Let us check that the function defined on $\mathbf{R}^4 \to \mathbf{R}^4$ by $f(x,y,z,w) = (x+2y, x^2-y^2, wz, y+w)$ is locally invertible at the point $(1,1,1,1)$.

Choose $E = \{(x,y,z,w) : x > 0, y > 0, z > 0 \text{ and } w > 0\}$.

Now E is an open subset of $\mathbf{R}^4$ containing the point $(1,1,1,1)$.

Let

$$f_1(x,y,z,w) = x + 2y$$

$$f_2(x,y,z,w) = x^2 - y^2$$

$$f_3(x,y,z,w) = wz$$

$$f_4(x,y,z,w) = y + w$$

Then f is continuously differentiable on E as each of the functions f_1, f_2, f_3, f_4 is a polynomial and so are C^1-function.

We also have

$$J_f(x,y,z,w) = \frac{\partial(f_1,f_2,f_3,f_4)}{\partial(x,y,z,w)} = \begin{vmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} & \frac{\partial f_2}{\partial w} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} & \frac{\partial f_3}{\partial w} \\ \partial f_4 & \partial f_4 & \partial f_4 & \partial f_4 \end{vmatrix}$$

$f_1(x, y, z, w) = x + 2y$

$f_2(x, y, z, w) = x^2 - y^2$

$f_3(x, y, z, w) = wz$

$f_4(x, y, z, w) = y + w$

$$\text{Then } J_f(x, y, z, w) = \begin{vmatrix} 1 & 2 & 0 & 0 \\ 2x & -2y & 0 & 0 \\ 0 & 0 & w & z \\ 0 & 1 & 0 & 1 \end{vmatrix} = -2w(x + y)$$

Also $J_f(1, 1, 1, 1) \neq 0$.

Therefore, the function is to be locally invertible at the point $(1, 1, 1, 1)$ and therefore there exists an openset E_0 containing the point $(1, 1, 1, 1)$ such that the conditions (i), (ii), (iii) and (iv) of Theorem 1 is satisfied.

Here we also note that for points for which $w = 0$, inverse function does not provide us with any information as J_f becomes zero.

This means at points like $(1, 3, 7, 0)$ we cannot infer any thing regarding local invertibility of the function f. Similarly, for points like $(4, -4, 5, 1)$ for which $x + y = 0$ also the inverse function theorem does not provide any information.

* * *

Next we shall prove a theorem which is a corollary of the inverse function theorem.

Theorem 2: Let f be a C^1 function defined from an open set $E \subseteq \mathbf{R}^n$ to $\mathbf{R}^n$, and suppose that $Jf(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in E$. Then the image f(B) of any open set $B \subset E$ is an open set in $\mathbf{R}^n$.

Proof: Let $B \subset E$ be open, and consider any $x_1 \in B$ such that $f(x_1) \in f(B)$. We apply the inverse function theorem, with E replaced by B. Then there exists an open set E_1 containing x_1, with $E_1 \subset B$, such that $f(E_1)$ is open. Since $f(x_1) \in f(E_1)$ and $f(E_1)$ is open. $f(x_1)$ has a neighbourhood U_1 such that $U_1 \subset f(E_1)$. But $f(E_1) \subset f(B)$, and hence $f(x_1)$ is an interior point of f(B). This shows that all the points of f(B) are interior points. Hence f(B) is open. □

The above theorem says that a map which satisfies conditions of inverse function theorem, carry open sets onto open sets. Geometrically this implies that the basic structure of an open set is preserved by the transformation given by the function f. To put it in other words the basic structure of an open set is preserved by such transformations. The transformations that can preserve a particular structure of a given set are studied in great detail because of its importance in mathematical modelling of different situations. In view of this the inverse function theorem has got lot of significance in applications.

Now we want you to try some exercises.

E1) Apply the inverse function theorem to check the local invertibility of the function $\mathbf{R}^3 \to \mathbf{R}^3$ given by

$$f(x, y, z) = (x + y + z, e^x \cos z, e^x \sin z).$$

E2) Check whether the conditions of the inverse function theorem are satisfied for the function $f : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $f(x, y) = (x, \phi(x, y))$ where

$$\phi(x,y) = \begin{cases} y - x^2 & \text{if } x^2 \le y, \\ \dfrac{(y^2 - x^2y)}{x^2} & \text{if } 0 \le y \le x^2 \\ \dfrac{(y^2 + x^2y)}{x^2} & \text{if } y \le 0. \end{cases}$$

E3) Show by an example that the condition that the Jacobian vanishes at a point is not necessary for a function to be locally invertible at that point.

In the next section, we shall discuss another important theorem in mathematical analysis.

6.3 IMPLICIT FUNCTION THEOREM

In this section we discuss implicit function theorem.

Before we consider the theorem we look into some preliminaries. Let f be a real valued function of two variables defined on a subset $E \subseteq \mathbf{R}^2$. Then $f(x, y) = 0$ represents the cross section of the set $\{(x, y, f(x, y)) : x, y \in E\}$ with the plane $z = 0$. We would like to investigate if this set can be explicitly represented as a curve of the form $y = f_1(x)$ or $x = f_2(y)$ and if yes do f_1 and f_2 inherit any good properties of f like continuity on differentiability. You must have studied answer to some of these questions in your undergraduate courses to some extent. This section is devotted to a study of such question or related matter in the general context, say $\mathbf{R}^n$.

Let us now consider a function $f : E \subset \mathbf{R}^n \to \mathbf{R}^m$.

Associated with any function $f : U \subset \mathbf{R}^n \to \mathbf{R}^m$ we have two types of sets which play a special role in the development of the subject — graphs and level sets. The graph of f is the subset of $\mathbf{R}^n \times \mathbf{R}^m$ defined as follows

$$\begin{aligned} \text{graph}(f) &= \{(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in U \text{ and } \mathbf{y} = f(\mathbf{x})\} \\ &= \{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in U\}. \end{aligned}$$

If $\mathbf{c} = (c_1, \ldots, c_m)$ is a point in $\mathbf{R}^m$ we define the level set of f, $f^{-1}(\mathbf{c})$, by the formula

$$f^{-1}(\mathbf{c}) = \{\mathbf{x} \in U : f(\mathbf{x}) = \mathbf{c}\}.$$

[You are already familiar with level curves and level surfaces for real-valued functions of 2 or 3 variables which are also called level sets in $\mathbf{R}^2$ and $\mathbf{R}^3$ respectively, from your undergraduate course (see IGNOU course MTE-07, Block 1, Unit 3).]

In terms of the coordinate functions $f_1, \ldots, f_m$ of f and the coordinates $C_1, \ldots, C_n$ of C we have and thus

$$f^{-1}(c) = \bigcap_{i=1}^{m}\{x \in U : f_i(x) = c_i\} = \bigcap_{i=1}^{m} f_i^{-1}(c_i).$$

Thus, a level set of a vector-valued function is the intersection of level sets of real-valued coordinate functions of the vector-valued functions. This is frequently useful in arriving at a geometrical interpretation of level sets as the following example shows.

Example 5: Let $f(x,y,z) = (x^2+y^2+z^2-1, 2x^2+2y^2-1)$. We have $f=(f_1,f_2)$ where $f_1(x,y,z) = x^2+y^2+z^2-1$ and $f_2(x,y,z) = 2x^2+2y^2-1$. The set $f_1^{-1}(0)$ is a sphere of radius 1 while $f_2^{-1}(0)$ is a cylinder parallel to the z-axis built on a circle with centre the origin and radius $1/\sqrt{2}$. If $\mathbf{0} = (0,0)$ is the origin in $\mathbf{R}^2$ then

$$f^{-1}(\mathbf{0}) = f_1^{-1}(0) \cap f_2^{-1}(0)$$

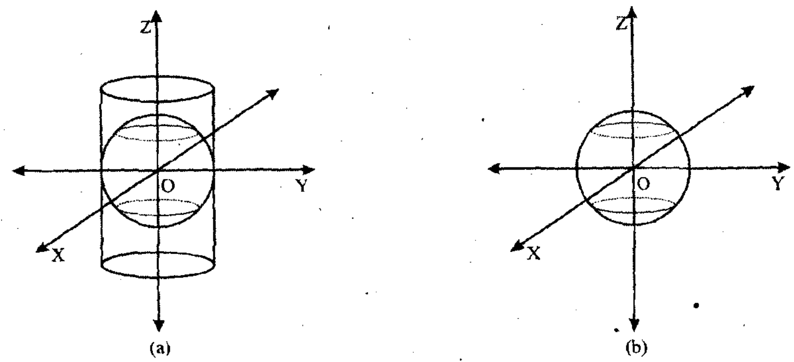

Fig. 2

which is the intersection of a sphere and a cylinder in $\mathbf{R}^3$ (Fig.2).

* * *

The relationship between graphs and level sets plays an important role in our study. The easy part of this relationship — every graph is a level set — is given in the next example while the difficult part — every level set "good in certain sense" is locally a graph — is the implicit function theorem.

Before we go further we introduce some notations.

Notation 1: If $x = (x_1,\ldots,x_n) \in \mathbf{R}^n$ and $y = (y_1,\ldots,y_m) \in \mathbf{R}^m$, let us write (x,y) for the point (or vector)

$$(x_1,\ldots,x_n,y_1,\ldots,y_m) \in \mathbf{R}^{n+m}.$$

The first entry in (x,y) or in a similar symbol will always be a vector in $\mathbf{R}^n$, the second will be a vector in $\mathbf{R}^m$.

Then any system of n equations in n + m variables can be expressed as

$$f_1(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0$$

....

....

....

$$f_n(x_1,\ldots,x_n,y_1,\ldots,y_m) = 0.$$

This can be written as $f(x, y) = 0$

where $f_1, f_2, \ldots, f_n$ are coordinate functions of the vector-valued function f i.e. $f = (f_1, \ldots, f_n)$.

Example 6: Let $f : U \subset R^n \to R^m$. We define $g : U \times R^n \to R^m$ by $g(x, y) = f(x) - y$. If **0** is the origin in R^m then

$$(x, y) \in g^{-1}(0) \iff g(x, y) = 0$$
$$\iff f(x) - y = 0$$
$$\iff (x, y) \in \text{ graph } (f).$$

Hence $g^{-1}(0) = \text{graph}(f)$ and every graph is a level set.

* * *

Let us now look at Example 5. In this case we obtain more information by solving the equation $f_1(x, y, z) = f_2(x, y, z) = 0$. We have $x^2 + y^2 = 1 - z^2 = 1/2$. This gives $z^2 = 1/2, z = \pm 1/\sqrt{2}$ and the level set consists of two circles on the unit sphere (see Fig.2). But it is not possible always to solve equation explicitly as in this case. Let us, for example, consider the following case.

Consider the equation $f(x, y) = x^2 + y^2 - 1 = 0$ $x, y \in R$. Here we cannot find a single value of y for a given value of x.

However, there are functions such as

i) $y = f_1(x) = +\sqrt{1 - x^2}, x \in [-1, 1]$

ii) $y = f_2(x) = -\sqrt{1 - x^2}, x \in [-1, 1]$

iii) $y = f_3(x) = \begin{cases} \sqrt{1 - x^2}, & x \in [-1, 0[\\ -\sqrt{1 - x^2}, & x \in [0, \frac{1}{2}[\\ \sqrt{1 - x^2}, & x \in [\frac{1}{2}, 1], \end{cases}$

which satisfy the equation $f(x, f(x)) = 0$. Fig. 3 shows the graphs of these functions.

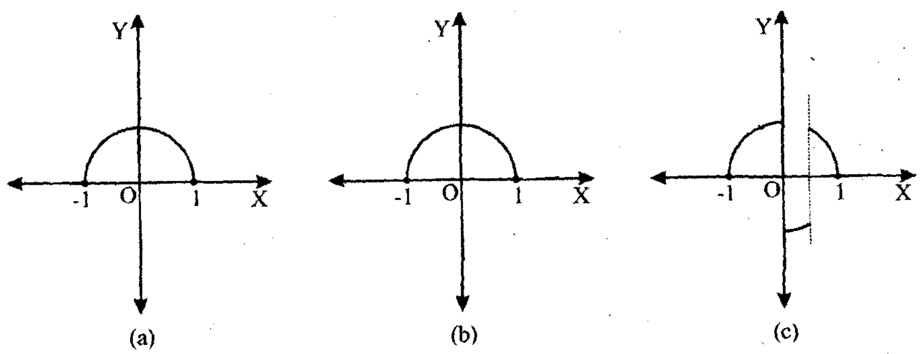

Fig. 3

In this context we can raise the following question.

"When can the equation $f(x, y) = 0$ be solved explicitly for y in terms of x, or solved explicitly for x in terms of y, yielding a function $y = f_1(x)$ or $x = f_2(y)$?" The implicit function theorem deals with this question.

Now, look at the three functions corresponding to the equation $f(x, y) = x^2 + y^2 - 1 = 0$. You will agree that f_1 and f_2 are continuous on

$[-1,1]$ and differentiable in $]-1,1[$, whereas f_3 is not even continuous. In the light of this we can ask whether the functions obtained by solving the equation $f(x,y)=0$ have nice properties like continuity, differentiability, etc. So we would like a simple criterion for deciding when, in general, such a function can be found. The implicit function theorem deals with this question locally. For the simple case of a real-valued functions of two variables, it tells us that, given a point (x_0, y_0) such that $f(x_0,y_0)=0$, under certain conditions there will be a neighbourhood of (x_0, y_0) such that in this neighbourhood the relation defined by $f(x,y)=0$ is also a function. The conditions are that f and its partial derivatives $\frac{\partial f}{\partial y}$ are continuous in some neighbourhood of (x_0, y_0) and that $\partial f(x_0, y_0) \neq 0$.

Theorem 3: (Implicit function theorem): Let E be an open subset of $\mathbf{R}^{n+1}$, let $f: E \to \mathbf{R}$ be continuously differentiable, and let $a=(a_1,\ldots,a_{n+1})$ be a point in E such that $f(a)=0$ and $\frac{\partial f}{\partial x_{n+1}}(a) \neq 0$. Then there exists an open subset U of $\mathbf{R}^n$ containing $(a_1,\ldots,a_n)$, an open subset V of E containing y, and a function $g: U \to \mathbf{R}$ such that $g(a_1,\ldots,a_n)=a_{n+1}$, and

$$\{(x_1,\ldots,x_{n+1}) \in V : f(x_1,\ldots,x_{n+1})=0\}$$
$$=\{(x_1,\ldots,x_n, g(x_1,\ldots,x_n)) : (x_1,\ldots,x_n) \in U\}.$$

In other words, the set $\{x \in V : f(x)=0\}$ is a graph of a function over U. Moreover, g is differentiable at $(a_1,\ldots,a_n)$, and we have

$$\frac{\partial g}{\partial x_j}(a_1,\ldots,a_n) = -\frac{\partial f}{\partial x_j}(a) \Big/ \frac{\partial f}{\partial x_{n+1}}(a) \qquad (2)$$

for all $1 \leq j \leq n$.

Remark : The equation (2) is sometimes derived using implicit differentiation. Basically, the point is that if you know that

$$f(x_1,\ldots,x_{n+1})=0$$

then (as long as $\frac{\partial f}{\partial x_{n+1}} \neq 0$) the variable x_{n+1} is "implicitly" defined in terms of the other n variables, and one can differentiate the above identity in, say, the x_j direction using the chain rule to obtain

$$\frac{\partial f}{\partial x_j} + \frac{\partial f}{\partial x_{n+1}}\frac{\partial x_n}{\partial x_j} = 0$$

Thus, the implicit function theorem allows one to define a dependence implicitly, by means of a constraint rather than by a direct formula of the form $x_{n+1} = g(x_1,\ldots,x_n)$.

We shall omit the proof of the implicit function. But we shall illustrate this theorem with some examples in some particular cases.

Let us look at some examples.

Example 7: Let f be the function defined from $\mathbf{R}^3$ to $\mathbf{R}$ given by

$$f(x,y,z) = x^2 + y^3 - xy\sin z$$

Let us show that the equation $f(x, y, z) = 0$ defines a unique continuously differentiable function g in a neighbourhood of the point $(1, -1)$ such that $g(1, -1) = 0$.

Let us apply the implicit function theorem to the function f.

We first note that $F(1, -1, 0) = 0$ and f is continuously differentiable on the whole of $\mathbf{R}^2$. Also

$$\frac{\partial f}{\partial x} = 2x - y \sin z$$

$$\frac{\partial f}{\partial y} = 3y - x \sin z$$

$$\frac{\partial f}{\partial z} = -xy \cos z$$

Then,

$$\frac{\partial f}{\partial z}(1, -1, 0) = 1 \neq 0.$$

Thus, f satisfies all the conditions of the implicit function theorem.

Therefore, by the implicit function theorem there exists a unique real valued function g defined on $\mathbf{R}^2$ such that $g(1, -1) = 0$.

We also have

$$\frac{\partial g}{\partial x} = -\frac{\partial f/\partial x}{\partial f/\partial z} = -\frac{(2x - y \sin z)}{xy \cos z}$$

and

$$\frac{\partial g}{\partial y} = -\frac{-\partial f/\partial y}{\partial f/\partial z} = -\frac{(3y - x \sin z)}{xy \cos z}$$

* * *

In its more general form, the implicit function theorem treats, instead of one equation in two variables, a system of n equations in n + m variables:

$$f_r(x_1, \ldots, x_n; t_1, \ldots, t_m) = 0 \quad (r = 1, 2, \ldots, m).$$

This system can be solved for $t_1, \ldots, t_m$ in terms of $x_1, \ldots, x_n$, provided that certain partial derivatives are continuous and provided that the $m \times m$ Jacobian determinant $\partial(f_1, \ldots, f_n)/\partial(t_1, \ldots, t_m)$ is not zero.

Theorem 4: Let $f = (f_1, \ldots, f_n)$ be a vector-valued function defined on an open set E in $\mathbf{R}^{n+m}$ with values in $\mathbf{R}^n$ such that

i) f is continuously differentiable on E.

ii) There exists a point $(\mathbf{x_0}, \mathbf{t_0}) \in E$ such that $f(\mathbf{x_0}, \mathbf{t_0}) = 0$ and for which $\dfrac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \neq 0.$

Then there exists an open set $T_0 \subseteq \mathbf{R}^m$ containing t_0 and one, and only one vector valued function g defined on T_0 satisfying

a) $g(t_0) = x_0$.

b) g is continuously differentiable on T_0.

c) $f(g(t), t) = 0$ for every $t \in T_0$.

Note: If f and g are as in Theorem 5, then we say that the function $f = (f_1, \ldots, f_n)$ defined from $\mathbf{R}^{n+m}$ to $\mathbf{R}^m$ is implicitly defined by g from $\mathbf{R}^m \to \mathbf{R}^n$

Example 8: Let $f = (f_1, f_2)$ be a vector-valued function defined from $\mathbf{R}^5$ to $\mathbf{R}^2$ where f_1 and f_2 are real-valued functions defined on $\mathbf{R}^5$ by

$$f_1(x_1, x_2, y_1, y_2, y_3) = 2e^{x_1} + x_2y_1 - 4y_2 + 3$$

$$f_2(x_1, x_2, y_1, y_2, y_3) = x_2 \cos x_1 - 6x_1 + 2y_1 - y_3$$

Let us show that f defines a unique function g from $\mathbf{R}^3$ to $\mathbf{R}^2$ in a neighbourhood T of the point $(3, 2, 7)$ such that $g(3, 2, 7) = (0, 1)$ and

$$f(x_1, x_2, y_3, g(x_1, x_2, x_3)) = 0$$

for $y = (y_1, y_2, y_3) \in T$

Note that here $n = 2$ and $m = 3$. We shall first observe that f_1 and f_2 are continuously differentiable and, therefore, f is also continuously differentiable. Also

$$f_1(0, 1, 3, 2, 7) = 0$$

and

$$f_2(0, 1, 3, 2, 7) = 0$$

Therefore $f(3, 2, 7, 0, 1) = 0$. Thus, the condition (i) is satisfied

Now we calculate the Jacobian $\dfrac{\partial(f_1, f_2)}{\partial(x_1, x_2)}$ which is given by

$$\frac{\partial(f_1, f_2)}{\partial(x_1, x_2)} = \begin{vmatrix} D_1f_1 & D_2f_1 \\ D_1f_2 & d_2f_2 \end{vmatrix}$$

$$= \begin{vmatrix} 2e^{x_1} & y_1 \\ -x_2 \sin x_1 - 6 & \cos x_1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 3 \\ -6 & 1 \end{vmatrix} \neq 0 \text{ at } (0, 1, 3, 2, 7)$$

This shows that the condition (ii) is satisfied at $(0, 1, 3, 2, 7)$. Thus, f satisfies all the conditions of the implicit function theorem.

* * *

Now we shall look at the implication of the implicit function theorem from a geometrical point of view. To understand this, we shall consider functions from $\mathbf{R}^n$ to $\mathbf{R}$.

If f is function from $\mathbf{R}^n$ to $\mathbf{R}$, then its graph given by

$$\{(x, f(x) : x \in \mathbf{R}^n\}$$

in general looks like some sort of n-dimensional surface in $\mathbf{R}^{n+1}$. This surface is technically known as a hypersurface. So, the question arises which hypersurfaces are actually graphs of some function and whether that function is continuous or differentiable. Now, the Implicit Function Theorem says that if the hypersurface is given by an algebraic equation, say for instance $\{(x, y, z) \in \mathbf{R} : xy + yz + zx = -1\}$ then it is possible to say whether the hypersurface is a graph, locally.

You can try some exercises now.

E4) Can the surface whose equation is

$$x + y + z - \sin(xyz) = 0$$

be described by an equation of the form $z = f(x, y)$ in a neighbourhood of the point (0, 0) satisfying $f(0, 0) = 0$? Justify your answer.

E5) Consider the function f from $\mathbf{R}^3 \to \mathbf{R}$ given by

$$f(x, y_1, y_2) = x^2 y_1 - e^x + y_2$$

Show that $f(0, 1, -1) = 0, (D_1 f)(0, 1, -1) \neq 0$ and that there exists, therefore, a differentiable function g in some neighbourhood of (1, 1) in $\mathbf{R}^2$, such that $g(1, -1) = 0$ and

$$f(g(y_1, y_2), y_1, y_2) = 0.$$

With this we come to an end of this unit.

6.4 SUMMARY

In this unit, we have covered the following points:

1. Defined the local invertibility for function from $\mathbf{R}^n \to \mathbf{R}$.
2. Stated the inverse function theorem and explained its importance for solution of system of equations.
3. Stated implicit function theorem and explain its significance.

6.5 HINTS/SOLUTIONS

E1) Here the function f is given by

$$f(x, y, z) = (x + y + z, e^x \cos z, e^x \sin z)$$

We note that the function f is continuously differentiable. Also

$$J_f = \begin{vmatrix} 1 & 1 & 1 \\ e^x \cos z & 0 & -e^x \sin z \\ e^x \sin z & 0 & e^x \cos z \end{vmatrix} = -e^{2x} \sin^2 z - e^{2x} \cos^2 z$$

$J_f = -e^{2x} \neq 0$ at any (x, y, z).

Therefore, the result follows by the implicit function theorem.

E2) **Hint:** f is differentiable at any point but the derivative of f is not coninuous at (0, 0).

E3) **Hint:** Take the function $f : \mathbf{R}^2 \to \mathbf{R}^2$

$$f(x, y) = (x^3, y^3)$$

E4) **Hint:** Apply implicit function theorem to the function $f(x, y, z) = x + y + z - \sin(xyz)$.

E5) **Hint:** Apply implicit function theorem.