Jovo Jaric

Implicit function theorem

The reader knows that the equation of a curve in the xy - plane can be expressed either in an "explicit" form, such as $y = f(x)$, or in an "implicit" form, such as $F(x,y) = 0$. However, if we are given an equation of the form $F(x,y) = 0$, this does not necessarily represent a function. Take, for example

$$F(x,y) = y^2 - 2xy - x^5 = 0. \qquad (1)$$

The equation $F(x,y) = 0$ *does* always represent a *relation*, namely, that set of all pairs (x,y) which satisfy the equation. The following question therefore presents itself quite naturally:

When is the relation defined by $F(x,y) = 0$ also a function? In other words, when can the equation $F(x,y) = 0$ be solved explicitly for y in terms of x, yielding a *unique* solution.

In the above example we can solve the quadratic for y and obtain

$$y = f(x) = x \pm x\left(1 + x^2\right)^{1/2}; \qquad (2)$$

here a certain care is needed, since we really have two explicit functions, corresponding to the positive and negative signs, from the one implicit relation. Accordingly we have to restrict attention to the vicinity of a particular point. For example, the values $x = 2$, $y = 8$ satisfy the relation $F(x,y) = 0$ but only one of the explicit formulas. Generally, what is needed is the assertion that there exists a function $y = f(x)$, which satisfies $F(x, f(x)) = 0$ and for which $f(2) = 8$, even though we could not obtain the expression

$$f(x) = x + x\left(1 + x^2\right)^{1/2}. \qquad (3)$$

This is useful if we want to treat x as a function of y, say $g(y)$, for here no explicit formula exists. At the point $x = -1$, $y = -1$ (or $x = 0$, $y = 0$) we run into trouble for here both formulas are valid. We shall obtain criteria for such points.
For that purpose we shall prove several general theorems that are required. The proof of the simplest theorem will be given in detail.

The implicit function theorem 1.
Let $x = x_0$, $y = y_0$ be a pair of values satisfying $F(x,y) = 0$ and let F and its first derivatives be continuous in the neighborhood of this point. Then, if F_y does not vanish at $x = x_0$, $y = y_0$ there exists one and only one continuous function $y = f(x)$ such that

$$F(x, f(x)) = 0 \ \text{and} \quad y_0 = f(x_0). \qquad (4)$$

Proof. By the hypothesis, $F_y(x_0, y_0) \neq 0$. Then we can assume without loss of generality that F_y at $x = x_0$, $y = y_0$ is positive, i.e. $F_y(x_0, y_0) > 0$. (If $F_y(x_0, y_0)$ were negative we could consider $G(x, y) = -F(x, y)$ instead, and G_y at $x = x_0$, $y = y_0$ would be positive.)

Moreover, since F, F_x, F_y are continuous we can find a box about point (x_0, y_0)

$$|x - x_0| \leq \delta, \quad |y - y_0| \leq \delta \tag{5}$$

within which $F(x, y)$, $F_x(x, y)$, $F_y(x, y)$ are continuous and $F_y(x, y) > 0$. Then $F(x_0, y)$, considered as a function of y, is an increasing function for $|y - y_0| \leq \delta$.

But $F(x_0, y_0) = 0$, by hypothesis, and so it must be that $F(x_0, y) < 0$ for $y_0 - \delta \leq y < y_0$, and that $F(x_0, y) > 0$ for $y_0 < y \leq y_0 + \delta$. In particular, $F(x_0, y_0 - \delta) < 0$ and $F(x_0, y_0 + \delta) > 0$. Further, by hypothesis, $F(x, y)$ is continuous at points $(x_0, y_0 - \delta)$ and $(x_0, y_0 + \delta)$. Hence, $F(x, y_0 - \delta) < 0$ and $F(x, y_0 + \delta) > 0$ for every x sufficiently close to x_0, i.e. for every $|x - x_0| \leq l$, where $l > 0$ is sufficiently small number. Let η be the smaller of two numbers l and δ, i.e. $\eta = \min(l, \delta)$. Then,

$$F_y(x, y) > 0, \quad F(x, y_0 - \delta) < 0, \quad F(x, y_0 + \delta) > 0 \tag{6}$$

for every x and y within intervals

$$|x - x_0| \leq \eta, \quad |y - y_0| \leq \delta. \tag{7}$$

Now, for any x from the interval $|x - x_0| \leq \eta$ function $F(x, y)$, as a function of y, will be increasing function for any y from the interval $|y - y_0| \leq \delta$, since then $F_y(x, y) > 0$. Thus, by the Intermediate value theorem for continuous functions, there must be a unique y in the interval $|y - y_0| \leq \delta$ for which $F(x, y) = 0$. Also when $x \to x_0$ then $y \to y_0$.

To summarize, for each x in the interval $|x - x_0| \leq \eta$ we have shown that there is one and only one y in the interval $|y - y_0| \leq \delta$ such that $F(x, y) = 0$; this association of y's with x's is a function $y = f(x)$ with the domain $|x - x_0| \leq \eta$ such that $F(x, f(x)) = 0$. Moreover, As $x \to x_0$ $\quad y \to y_0 = f(x_0)$ so that the function is continuous at x_0.

Note that we have only gotten an element and perhaps only a small element of the function, namely, the part within $|x - x_0| \leq \eta$. Therefore, the implicit function theorem deals the question *locally*. However, we notice that the condition of the theorem are satisfied for any point x_1 within this interval. We can therefore begin again to construct the function $y = f(x)$ from the point (x_1, y_1) and hope to extend the interval. In fact, we shall always be able to extend the interval until we reach a point where $F_y = 0$ and then there is no unique solution to be found.

Finally, since the function $F(x, y)$ is continuously differentiable so is the function $y = f(x)$. In order to show this we start from the identity

$$F(x,y)=F(x,y)-F(x_0,y)+F(x_0,y)-F(x_0,y_0), \tag{8}$$

since $F(x_0,y_0)=0$ .Then applying the first mean value theorem to each of these differences , within the intervals $|x-x_0|\le\eta$ and $|y-y_0|\le\delta$, we have

$F(x,y)=(x-x_0)F_x\left[x_0+\theta(x-x_0),y\right]+(y-y_0)F_y\left[x_o,y_0+\theta(y-y_0)\right],\ 0\le\theta\le 1.$

But $F(x,f(x))=0$ for $y=f(x)$ so that

$0=(x-x_0)F_x\left[x_0+\theta(x-x_0),f(x)\right]+\left(f(x)-y_0\right)F_y\left[x_o,y_0+\theta\left(f(x)-y_0\right)\right].$

If we write $x=x_0+\Delta x$, $y=y_0+\Delta y$ for any x and y in the above intervals , we get

$F_x(x_0+\theta\Delta x,y)\Delta x+F_y(x_0,y_0+\theta\Delta y)\Delta y=0.$

Now $y\to y_0$ and $\Delta y\to 0$ as $\Delta x\to 0$, because $y=f(x)$ is continuous ; also $F_y(x_0,y_0)\neq 0$. Then the following limit exists

$$\frac{dy}{dx}=f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=-\frac{F_x(x_0,y_0)}{F_y(x_0,y_0)}.\ \square \tag{9}$$

The theorem may be extended in a number of ways. For example the same technique can be applied to solving $f(x,y,z)=0$ for one of variables in terms of the others. Now we proceed in solving two simultaneous equations

$$\varphi(x,y,z)=0\ ,\ \psi(x,y,z)=0 \tag{10}$$

Theorem 2. *Let* $\varphi(x_0,y_0,z_0)=0$ *and* $\psi(x_0,y_0,z_0)=0$. *Let* $\varphi(x,y,z)$ *and* $\psi(x,y,z)$ *be differentiable in a neighborhood* R *of the point* (x_0,y_0,z_0) *and let the Jacobian*

$$\frac{\partial(\varphi,\psi)}{\partial(y,z)}=\begin{vmatrix}\frac{\partial\varphi}{\partial y} & \frac{\partial\varphi}{\partial z}\\ \frac{\partial\psi}{\partial y} & \frac{\partial\psi}{\partial z}\end{vmatrix} \tag{11}$$

be nonsingular there. Then, there is one and only one set of solutions

$$y=y(x),\quad z=z(x)$$

which are continuous, satisfy the equations $\varphi(x,y,z)=0$, $\psi(x,y,z)=0$ *and for which* $y_0=y(x_0)$, $z_0=z(x_0)$. *Furthermore,* $y=y(x)$, $z=z(x)$ *are differentiable.*

` *Proof.* Since The Jacobian (11) does not vanish at (x_0,y_0,z_0) then at least one of partial derivatives $\frac{\partial\psi}{\partial y}$ or $\frac{\partial\psi}{\partial z}$ must not vanish there. Let us assume that $\frac{\partial\psi}{\partial z}$ does not vanish at (x_0,y_0,z_0). Then according to the **Theorem 1** $\psi(x,y,z)=0$ defines a unique differentiable function $z=\phi(x,y)$. Now if we substitute this function into (10)1 we obtain $F(x,y)=\varphi(x,y,\phi(x,y))=0$.

Now, in order to prove the **theorem 2** it is sufficient to show that

$$\frac{\partial F}{\partial y}=\frac{\partial\varphi}{\partial y}+\frac{\partial\varphi}{\partial z}\frac{\partial\phi}{\partial y}\neq 0 \tag{12}$$

at $\left(x_0,y_0\right)$. To do this we eliminate $\dfrac{\partial\phi}{\partial y}$ from this expression by making use of the identity $\psi\left(x,y,\phi\left(x,y\right)\right)=0$ .Then

$$\frac{\partial\psi}{\partial y}+\frac{\partial\psi}{\partial z}\frac{\partial\phi}{\partial y}=0\,. \tag{13}$$

which can be solved for $\dfrac{\partial\phi}{\partial y}$ in the neighborhood of the point $\left(x_0,y_0,z_0\right)$. In fact, from (12) and (13) we have

$$\frac{\partial F}{\partial y}=\frac{\partial\left(\varphi,\psi\right)}{\partial\left(y,z\right)}\Big/\frac{\partial\psi}{\partial z}\,. \tag{14}$$

But $\dfrac{\partial\left(\varphi,\psi\right)}{\partial\left(y,z\right)}$ and $\dfrac{\partial\psi}{\partial z}$ at $\left(x_0,y_0,z_0\right)$ do not vanish by hypothesis and our assumption, respectively. Therefore, the same holds for $\dfrac{\partial F}{\partial y}$. Hence (12) defines a unique function $y=y\left(x\right)$. Then, substituting this function into $z=\phi\left(x,y\right)$ we obtain $z=\phi\left(x,y\left(x\right)\right)$, i.e. $z=z\left(x\right)$. □

A more general theorem follows.

Theorem 3. *Let*

$$F_i\left(x_1,....,x_m,y_1,....,y_n\right),\qquad i=1,....,n$$

be differentiable in a neighborhood of the point $\left(\boldsymbol{x}^0,\boldsymbol{y}^0\right)=\left(x_1^0,....,x_m^0,y_1^0,....,y_n^0,\right)$. *Further, let* $F_i\left(\underset{0}{\boldsymbol{x}},\underset{0}{\boldsymbol{y}}\right)=0$ *and let the Jacobian*

$$\frac{\partial\left(F_1,....,F_n\right)}{\partial\left(y_1,....,y_n\right)}=\begin{vmatrix}\dfrac{\partial F_1}{\partial y_1} & \cdot & \cdot & \dfrac{\partial F_1}{\partial y_n}\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ \dfrac{\partial F_n}{\partial y_1} & & & \dfrac{\partial F_n}{\partial y_n}\end{vmatrix} \tag{15}$$

be nonsingular at $\left(\boldsymbol{x}^0,\boldsymbol{y}^0\right)$. *Then there exists neighborhood* $\Re$ *of* $\left(\boldsymbol{x}^0,\boldsymbol{y}^0\right)$ *and a unique set of solutions*

$$y_i=f_i\left(x_1,....,x_m\right) \tag{16}$$

of the equations

$$F_i\left(x_1,....,x_m,y_1,....,y_n\right)=0\,. \tag{17}$$

Furthermore, f_i *are differentiable.*

Proof. If we assume the theorem true for $\left(n-1\right)$ equations and prove its truth for n equations, then by induction we can rise to the general case from $n=1,2$ (already proved by the **theorem 1** and **theorem 2**). Since the Jacobian does not vanish at least one of cofactors of its last row must not vanish and for convenience we may take this to be

$$\frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)} \neq 0 .$$

Then since the theorem is assumed for the case $(n-1)$, we can solve the first $(n-1)$ equations in the form

$$y_\alpha = \varphi_\alpha\left(x_1,....,x_m ; y_n\right), \qquad \alpha = 1,....,n-1,$$

and the φ_α are differentiable. Then substituting in the last equation we have

$$\Phi\left(x_1,....,x_m ;\, y_n\right) = F_n\left(x_1,....,x_m,\varphi_1,....,\varphi_{n-2} ; y_n\right) = 0 .$$

If the derivative $\frac{\partial\Phi}{\partial y_n}$ does not vanish, then this can be solved to give $y_n = \varphi_n\left(x_1,....,x_m\right)$ and then substitution gives

$$y_\alpha = f_\alpha\left(x_1,....,x_m\right) = \varphi_\alpha\left(x_1,....,x_m ; \varphi_n\right), \qquad \alpha = 1,....,n-1,$$

$$y_n = f_n\left(x_1,....,x_m\right) \equiv \varphi_n\left(x_1,....,x_m\right) . \tag{18}$$

However,

$$\frac{\partial\Phi}{\partial y_n} = \frac{\partial F_n}{\partial y_\alpha}\frac{\partial\varphi_\alpha}{\partial y_n} + \frac{\partial F_n}{\partial y_n} .$$

The $\frac{\partial\varphi_\alpha}{\partial y_n}$ are calculated from the first $(n-1)$ of (17):

$$\frac{\partial F_\beta}{\partial y_\alpha}\frac{\partial\varphi_\alpha}{\partial y_n} + \frac{\partial F_\beta}{\partial y_n} = 0, \qquad \alpha,\beta = 1,.....,n-1.$$

From this

$$\frac{\partial\varphi_\alpha}{\partial y_n} = -\frac{\partial\left(F_1,..,..,F_{n-1}\right)}{\partial\left(y_1,..,y_{\alpha-1},y_n,y_{\alpha+1},...,y_{n-1}\right)} \Big/ \frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)} =$$

$$= \left(-1\right)^{n-\alpha}\frac{\partial\left(F_1,..,..,F_{n-1}\right)}{\partial\left(y_1,..,y_{\alpha-1},y_{\alpha+1},...,y_{n-1},y_n\right)} \Big/ \frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)}$$

so that

$$\frac{\partial\Phi}{\partial y_n} = \sum_{\alpha=1}^{n-1} \left(-1\right)^{n-\alpha}\frac{\partial F_n}{\partial y_\alpha}\frac{\partial\left(F_1,..,..,F_{n-1}\right)}{\partial\left(y_1,..,y_{\alpha-1},y_{\alpha+1},...,y_{n-1},y_n\right)} + \frac{\partial F_n}{\partial y_n}\frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)} \Big/ \frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)}$$

$$= \sum_{\alpha=1}^{n} \left(-1\right)^{n-\alpha}\frac{\partial F_n}{\partial y_\alpha}\frac{\partial\left(F_1,..,..,F_{n-1}\right)}{\partial\left(y_1,..,y_{\alpha-1},y_{\alpha+1},...,y_{n-1},y_n\right)} \Big/ \frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)} .$$

Finally,

$$\frac{\partial\Phi}{\partial y_n} = \frac{\partial\left(F_1,..,..,F_n\right)}{\partial\left(y_1,.....,y_n\right)} \Big/ \frac{\partial\left(F_1,....,F_{n-1}\right)}{\partial\left(y_1,....,y_{n-1}\right)} \tag{19}$$

and since neither term on the right-hand side is zero, $\frac{\partial \Phi}{\partial y_n} \neq 0$. □

The particular form in which the theorem is most needed is the inversion of a functional transformation. If $m = n$ and F_i has the form

$$F_i = g_i(y_1, \ldots\ldots\ldots, y_n) - x_i \,,$$

where g_i are continuous differentiable functions then the **theorem 3** takes the following form.

Theorem 4. *If*

$$x_i = g_i(y_1, \ldots\ldots\ldots, y_n)\,, \qquad i = 1, \ldots., n\,, \tag{20}$$

are n continuous functions of the variables $y_1, \ldots\ldots\ldots, y_n$ with continuous first partial derivatives, and if the Jacobian

$$J = \frac{\partial\left(x_1, \ldots\ldots\ldots, x_n\right)}{\partial\left(y_1, \ldots\ldots\ldots, y_n\right)} \tag{21}$$

does not vanish, then the transformation from y to x can be uniquely inverted to give

$$y_i = f_i(x_1, \ldots\ldots\ldots, x_n). \tag{22}$$

Proof. We simply apply the **teorem 3.** But now the Jacobian (15) reads

$$J = \frac{\partial\left(g_1, \ldots\ldots\ldots, g_n\right)}{\partial\left(y_1, \ldots\ldots\ldots, y_n\right)}$$

because of the form of functions F_i. Then (21) follows from (20). □

References

[1] Tom M. Apostol, Mathematical Analysis, Second edition, Addison-Wesley Publishing Company, Reading, Massachusetts, Amsterdam, London, Manila, Singapore, Sydney, Tokyo, pp. 373, 1974

[2] Howard Anton, Calculus, Second edition, John Wiley & Sons& Sons, New York, Chichester, Brisbane, Toronto, Singapore , pp. 177, 1984

[3] Abraham Schwartz, Calculus and Analytic Geometry, Third edition, pp. 828-831 1974

[4] Rutherford Aris, Vectors, Tensors and the Basic Equations of Fluid Mechanics, Dover Publication. Inc. New York, 1989