AN INFINITESIMAL PROOF OF THE IMPLICIT FUNCTION THEOREM

by NIGEL J. CUTLAND and FENG HANQIAO

(Received 11 November, 1991)

We give a short and constructive proof of the general (multi-dimensional) Implicit Function Theorem (IFT), using infinitesimal (i.e. nonstandard) methods to implement our basic intuition about the result. Here is the statement of the IFT, quoted from [**4**];

THEOREM. *Let $A \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set and let $F: A \to \mathbb{R}$ be a function of class C^p $(p \geq 1)$. Suppose that $(x_0, y_0) \in A$ with $F(x_0, y_0) = 0$ $(x_0 \in \mathbb{R}^n, y_0 \in \mathbb{R}^m)$ and that the Jacobian determinant $J = \dfrac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}$ is not zero at (x_0, y_0). Then there is an open neighbourhood U of x_0 and a unique function $f: U \to \mathbb{R}^m$ with*

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

for all $x \in U$. Moreover, f is of class C^p.

First let us give an intuitive informal description of f; we need some notation. Points $x, y \in \mathbb{R}^n$, $\mathbb{R}^m$ will be regarded as column vectors; we write $\partial F/\partial y$ for the $m \times n$ Jacobian matrix $\partial F/\partial y = (\partial F_i/\partial y_j)$, where we have $F = (F_1, \ldots, F_m)'$ and $F_i = F_i(x, y)$. Then $J = |\partial F/\partial y|$. Similarly $\partial F/\partial x = (\partial F_i/\partial x_j)$, an $m \times n$ matrix.

Intuitively, a recipe for f is given as follows. Writing $dx = (dx_1, \ldots, dx_n)'$ etc., we have, informally

$$0 = dF(x, f(x)) = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} df.$$

If $\partial F/\partial y$ is invertible (which it is in a neighbourhood of (x_0, y_0)) then

$$df(x) = f(x + dx) - f(x) = -\frac{\partial F^{-1}}{\partial y} \frac{\partial F}{\partial x} dx. \tag{1}$$

Using infinitesimal techniques we can implement this recipe for f, by discretizing the space $\mathbb{R}^n$ and using (1) as a recursive definition for f. We assume the basics of nonstandard analysis, which may be found in [**1**] or [**3**].

Pick a positive infinitesimal $\Delta \neq 0$ and let $T = \{k\Delta : k \in {}^*\mathbb{Z}\}$. We will consider $\tau = (t_1, \ldots, t_n)$ taking values in the lattice $T^n \subseteq {}^*\mathbb{R}^n$.

We shall need the following elementary lemma [**2**].

LEMMA. *Let $\psi: T \to {}^*\mathbb{R}$ be internal, and let $D\psi$ be the difference function:*

$$D\psi(t) = \frac{\psi(t + \Delta) - \psi(t)}{\Delta}.$$

If $\psi(0)$ is finite and $D\psi$ is S-continuous for $|t| \leq c$ then there is a unique standard function $g: [-c, c] \to \mathbb{R}$ given by

$$g({}^\circ t) = {}^\circ\psi(t).$$

Glasgow Math. J. **35** (1993) 163–166.

https://doi.org/10.1017/S001708950000971X Published online by Cambridge University Press

Moreover, g is C^1, *and* $Dg(°t) = °D\psi(t)$. (Recall that ψ is *S-continuous* if $\psi(t) \approx \psi(t')$ whenever $t \approx t'$.)

Proof of the IFT Without loss of generality we may assume that $x_0 = 0$ and $y_0 = 0$. Define an internal function $\varphi : T^n \to {}^*\mathbb{R}^m$ recursively as follows.

(i) $\varphi(0) = 0$

(ii) for each $0 < k \le n$ and $\sigma \in T^{k-1}$, if $\varphi(\sigma, 0, \ldots, 0) = \varphi(\sigma_1, \ldots, \sigma_{k-1}, 0, \ldots, 0)$ has been defined, then define $\varphi(\sigma, t, \ldots, 0)$ for $t \in T$ by:

$$\varphi(\sigma, t + \Delta, \ldots, 0) = \varphi(\sigma, t, \ldots, 0) - \frac{\partial F^{-1}}{\partial y} \frac{\partial F}{\partial x_k} \Delta \quad \text{if} \quad t \ge 0,$$

$$\varphi(\sigma, t - \Delta, \ldots, 0) = \varphi(\sigma, t, \ldots, 0) + \frac{\partial F^{-1}}{\partial y} \frac{\partial F}{\partial x_k} \Delta \quad \text{if} \quad t \le 0.$$

Note that by Cramer's rule, this explicit recipe is given by

$$\varphi_i(\sigma, t \pm \Delta, \ldots, 0) = \varphi_i(\sigma, t, \ldots, 0) \mp \Delta J^{-1} \frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_{i-1}, x_k, y_{i+1}, \ldots, y_m)}.$$

The matrices $\partial F/\partial y$ and $\partial F/\partial x$ are evaluated at $x = (\sigma, t, \ldots, 0)$ and $y = \varphi(x)$. The hypotheses on $\partial F/\partial x$ and $\partial F/\partial y$ ensure that on some rectangle $-a \le x_i, y_i \le a$ (where a is positive standard) there is a standard $M > 0$ with $\left| \left(\frac{\partial F^{-1}}{\partial y} \frac{\partial F}{\partial x} \right)_{j,k} \right| \le M$ for all j, k. It is easy to check that this ensures that for $\tau = (t_1, \ldots, t_n)$ with each $|t_i| \le \frac{a}{Mn}$ the above definition gives $|\varphi_j(\tau)| \le a$. (This is done by induction, as in the definition of φ: in fact we show that if each $|t_i| \le \frac{a}{Mn}$ then for each $k \le n$ we have

$$|\varphi_j(t_1, \ldots, t_k, 0, \ldots, 0)| \le \frac{ka}{n}. \tag{2}$$

If (2) holds for k, the definition of φ ensures that if $|t| \le \frac{a}{Mn}$ then

$$|\varphi_j(t_1, \ldots, t_k, t, \ldots, 0) - \varphi_j(t_1, \ldots, t_k, 0, \ldots, 0)| \le M\,|t| \le \frac{a}{n}$$

which is sufficient to establish (2) for $k + 1$.)

Let $b = \frac{a}{Mn}$ and for $\tau = (t_1, \ldots, t_n) \in T^n$ write $|\tau| \le b$ to mean $|t_i| \le b$ for all i. It is clear from the definition of φ that

$$|\varphi(t_1, \ldots, t_k, t, 0, \ldots, 0) - \varphi_j(t_1, \ldots, t_k, t', 0, \ldots, 0)| \le M\,|t - t'| \tag{3}$$

for $|\tau| \le b$ and $|t|$, $|t'| \le b$. In particular $\varphi(t_1, \ldots, t_n)$ is S-continuous in t_n for $|t_i| \le b$. We will show later that it is S-continuous in all its arguments.

We now show that

$$F(\tau, \varphi(\tau)) \approx 0 \quad \text{for} \quad |\tau| \le b, \qquad \tau \in T^n. \tag{4}$$

https://doi.org/10.1017/S001708950000971X Published online by Cambridge University Press

This is again done by induction as in the definition of φ. Let $\tau=(\sigma,t,\ldots,0)$ and $\tau'=(\sigma,t+\Delta,\ldots,0)$. Then by the mean value theorem

$$F_j(\tau',\varphi(\tau'))-F_j(\tau,\varphi(\tau))=\frac{\partial F_j}{\partial x_k}(\bar\tau,\bar\eta)\Delta+\frac{\partial F_j}{\partial y}(\bar\tau,\bar\eta)(\varphi(\tau')-\varphi(\tau))$$

for some $\bar\tau$ between τ and τ', and $\bar\eta$ between $\varphi(\tau)$ and $\varphi(\tau')$. Now use the definition of φ to see that

$$F_j(\tau',\varphi(\tau'))-F_j(\tau,\varphi(\tau))=\left[\frac{\partial F_j}{\partial x_k}(\bar\tau,\bar\eta)-\frac{\partial F_j}{\partial y}(\bar\tau,\bar\eta)\left(\frac{\partial F^{-1}}{\partial y}\frac{\partial F}{\partial x_k}\right)(\tau,\varphi(\tau))\right]\Delta=\Delta\epsilon$$

where $\epsilon\approx 0$ by the continuity of all derivatives of F, and the fact that $\tau\approx\tau'$ and $\varphi(\tau)\approx\varphi(\tau')$ by (3). Now ϵ depends on τ, but we may take $\epsilon_0=$ maximum of all ϵ as τ varies in $|\tau|\leqslant b$, and then it is easy to see that $F_j(\tau,\varphi(\tau))\approx F_j(0,\varphi(0))=0$ for all such τ.

We now see that φ is essentially unique with the property (4). We show that

$$F(\tau,y)\approx F(\tau,y')\Rightarrow y\approx y' \tag{5}$$

for $|\tau|$, $|y|$, $|y'|\leq b$. By the mean value theorem

$$0\approx F_j(\tau,y')-F_j(\tau,y)=\frac{\partial F_j}{dy}(\tau,y^j)(y'-y)$$

for some $y^j\in\mathbb{R}^m$ between y and y'. Now the assumption $J(0,0)\neq 0$ and continuity of derivatives means that for small enough a, and $|\tau|$, $|y|$, $|y'|\leq a$ the matrix $\left(\frac{\partial F_j}{\partial y}(\tau,y_j)\right)$ is non-singular, and so $y'\approx y$.

To show that φ is S-continuous in all its arguments, fix $k<n$ and consider another function $\bar\varphi$ defined like φ but with indices $1,\ldots,n$ permuted so that k is the last. Then the above all applies to $\bar\varphi$: in particular, from (4)

$$F(\tau,\bar\varphi(\tau))\approx 0 \quad \text{for} \quad |\tau|\leq b, \qquad \tau\in T^n$$

and so from (5)

$$\varphi(\tau)\approx\bar\varphi(\tau) \quad \text{all} \quad \tau=(t_1,\ldots,t_k), \qquad |\tau|\leq b.$$

Moreover, $\bar\varphi(t_1,\ldots,t_n)$ is S-continuous in t_k, and hence φ is S-continuous in t_k. Thus φ is S-continuous on $|\tau|\leq b$ and we can define a standard continuous function f by

$$f({}^\circ\tau)={}^\circ\varphi(\tau) \qquad |\tau|\leq b, \qquad \tau\in T^n.$$

From (4) and the continuity of F, we have

$$F(x,f(x))=0 \quad \text{for} \quad |x|\leq b.$$

The uniqueness of f for $|x|\leq b$ is given by the argument used to give (5).

To see that f is continuously differentiable, the definition of φ together with the lemma shows that

$$\frac{\partial f}{\partial x_k}(x_1,\ldots,x_k,0,\ldots,0)=-\frac{\partial f^{-1}}{\partial y}\frac{\partial F}{\partial x_k}(x,f(x))$$

https://doi.org/10.1017/S001708950000971X Published online by Cambridge University Press

(where $x=(x_1,\ldots,x_k,0,\ldots,0)$). A simple symmetry argument shows that this is valid for all x with $|x|\le b$; i.e.

$$\frac{\partial f}{\partial x}=-\frac{\partial F^{-1}}{\partial y}\frac{\partial F}{\partial x}(x,f(x))$$

for all x with $|x|\le b$. If F is C^p, repeated differentiation shows that f is also C^p.

REFERENCES

1. N. J. Cutland, (Editor), *Nonstandard analysis and its applications* (Cambridge University Press, 1988).

2. Feng Hanqiao, D. F., St. Mary and F. Wattenberg, Applications of nonstandard analysis to partial differential equations—the diffusion equation, *Mathematical Modelling*, **7** (1986) 507–523.

3. A. E. Hurd and P. A. Loeb, *An introduction to nonstandard real analysis* (Academic Press, 1985).

4. J. E. Marsden, *Elementary classical analysis* (W. H. Freeman, 1974).

Department of Pure Mathematics
University of Hull
Hull HU6 7RX
England

Department of Computer Science
Shaanxi University
Xian, 710062
China

https://doi.org/10.1017/S001708950000971X Published online by Cambridge University Press