MAT 473 Intermediate Real Analysis II

John Quigg Spring 2009

revised February 13, 2009

Implicit functions

Given a system of m independent linear equations in n variables $x_1, \dots, x_n$, linear algebra tells us that we can solve for m of the variables in terms of the others. More precisely, if A is the coefficient matrix, then we can solve uniquely for $x_{j_1}, \dots, x_{j_m}$ if and only if the columns $j_1, \dots, j_m$ of A are linearly independent. For example, if $A = \begin{pmatrix} B & C \end{pmatrix}$ is the block matrix decomposition with C square (of size $m \times m$), and if C is invertible, then for any $c \in \mathbb{R}^m$ the equation $Bx + Cy = c$ can be solved uniquely for y as a function of x, namely $y = C^{-1}(c - Bx)$.

In the case of a nonlinear system of equations, the situation is similar, and is controlled by the derivative. To see what solving for y in terms of x entails, let f be a function from a subset of $\mathbb{R}^n$ into $\mathbb{R}^m$, with $m < n$ (the case $m = n$ is handled by the Inverse Function Theorem). Put $k = n - m$, and identify $\mathbb{R}^n$ with $\mathbb{R}^k \times \mathbb{R}^m$, so that a typical element is regarded as an ordered pair (x, y) with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^m$. We want to solve an equation of the form $f(x, y) = c$ for y in terms of x. In general, the best we can hope for is to solve for y in terms of x locally, so that near any sufficiently nice point in the domain of f, the set of solutions of $f(x, y) = c$ should be the graph $\{(x, g(x)) : x \in \operatorname{dom} g\}$ of a function g.

Here is the main result:

Theorem 1 (Implicit Function Theorem)**.** *Let $E \subset \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m$, let $f : E \to \mathbb{R}^m$ be C^1, and let $(a, b) \in E$. Write $f'(a, b) = \begin{pmatrix} A & B \end{pmatrix}$ where B is $m \times m$, and assume that B is invertible. Put $f(a, b) = c$. Then there exist open sets $U \subset E$ and $W \subset \mathbb{R}^k$ such that $(a, b) \in U$ and $U \cap f^{-1}(c)$ is the graph of a C^1 function $g : W \to \mathbb{R}^m$.*

Note that, once we have proved the above theorem, if we wanted to we could find open sets $W_0 \subset W$ and $V_0 \subset \mathbb{R}^m$ such that: (1) $W_0 \times V_0 \subset U$, (2) for all $x \in W_0$ there exists a unique $y = g_0(x) \in V_0$ such that $f(x, y) = c$, and (3) the resulting function $g_0 : W_0 \to V_0$ is C^1. This is how the Implicit Function Theorem is sometimes phrased.

Proof. Define $\phi : E \to \mathbb{R}^n$ by $\phi(x, y) = (x, f(x, y))$. Then ϕ is C^1, and

$$\phi'(a, b) = \begin{pmatrix} I & 0 \\ A & B \end{pmatrix}$$

is invertible. By the Inverse Function Theorem there exist open sets $U \subset E$ and $V \subset \mathbb{R}^n$ such that $(a, b) \in U$, ϕ maps U 1-1 onto V, and $\phi^{-1} : V \to U$ is C^1.

Put

$$W = \{x \in \mathbb{R}^k : (x, c) \in V\}.$$

Then $a \in W$, and W is open because V is. Define $g : W \to \mathbb{R}^m$ by

$$g(x) = \pi_2 \circ \phi^{-1}(x, c),$$

where $\pi_2 : \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ is the projection onto the second coordinate (i.e., $\pi_2(x, y) = y$). Then g is C^1 since ϕ^{-1} is (and π_2 is linear). By construction, for all $(x, y) \in U$ we have both

$$x \in W \quad \text{and} \quad y = g(x)$$

if and only if $(x, y) = \phi^{-1}(x, c)$, equivalently $f(x, y) = c$. Thus $U \cap f^{-1}(c)$ is the graph of g. □

In the above statement of the Implicit Function Theorem, there is nothing magical about our choice of the identification of $\mathbb{R}^n$ with $\mathbb{R}^k \times \mathbb{R}^m$, namely with the "$\mathbb{R}^k$-variable" being $(x_1, \ldots, x_k)$:

Corollary 2. *Let $E \subset \mathbb{R}^n$, $f : E \to \mathbb{R}^m$, and $a \in E$. Assume that f is C^1. If columns $j_1, \ldots, j_m$ of $f'(a)$ are linearly independent, then we can solve the equation $f(x) = c$ for $x_{j_1}, \ldots, x_{j_m}$ as a C^1-function of the remaining variables.*

Proof. Just compose f with a rearrangement of coordinates and apply the Implicit Function Theorem. □

The Implicit Function Theorem allows us to solve $f(x, y) = c$ as $y = g(x)$ for a C^1-function g. The following result shows how to compute g' in terms of f':

Proposition 3. *With the notation of the Implicit Function Theorem,*

$$g'(a) = -B^{-1}A.$$

Proof. Define $h : W \to E$ by $h(x) = (x, g(x))$. Then h is differentiable, and

$$h'(a) = \begin{pmatrix} I \\ g'(a) \end{pmatrix}.$$

Since $f \circ h$ is constant, we have

$$\begin{aligned} 0 &= (f \circ h)'(a) \\ &= f'(h(a))h'(a) \\ &= f'(a, b) \begin{pmatrix} I \\ g'(a) \end{pmatrix} \\ &= \begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} I \\ g'(a) \end{pmatrix} \\ &= A + Bg'(a), \end{aligned}$$

and the result follows by solving for $g'(a)$. □

The main idea of the above proof can be expressed loosely as follows: differentiating both sides of the equation $f(x, g(x)) = c$ with respect to x gives

$$\frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x))g'(x) = 0,$$

and we can solve for $g'(x)$. This can't be quite right, though, because we haven't bothered to define partial derivatives with respect to a "chunk" of variables. This could be put on a firm footing, but I didn't think it was worth it.

Example 4. Consider the system

$$\begin{aligned} u^5 + xv^2 - y + w &= 0 \\ v^5 + yu^2 - x + w &= 0 \\ w^4 + y^5 - x^4 &= 1 \end{aligned}$$

of 3 equations in the 5 real variables x, y, u, v, w.

Define $f : \mathbb{R}^5 \to \mathbb{R}^3$ by

$$f(x, y, u, v, w) = (u^5 + xv^2 - y + w, v^5 + yu^2 - x + w, w^4 + y^5 - x^4).$$

Then f is C^1, and

$$f'(x, y, u, v, w) = \begin{pmatrix} v^2 & -1 & 5u^4 & 2xv & 1 \\ -1 & u^2 & 2yu & 5v^4 & 1 \\ -4x^3 & 5y^4 & 0 & 0 & 4w^3 \end{pmatrix}.$$

We have

$$f(1, 1, 1, 1, -1) = (0, 0, 1),$$

and

$$f'(1, 1, 1, 1, -1) = \begin{pmatrix} 1 & -1 & 5 & 2 & 1 \\ -1 & 1 & 2 & 5 & 1 \\ -4 & 5 & 0 & 0 & -4 \end{pmatrix}.$$

Since the matrix

$$\begin{pmatrix} 5 & 2 & 1 \\ 2 & 5 & 1 \\ 0 & 0 & -4 \end{pmatrix}$$

is invertible, we can apply the Implicit Function Theorem to conclude that there exist $r > 0$ and C^1-functions $h, k, l : B_r(1, 1) \to \mathbb{R}$ such that

$$h(1, 1) = 1, \quad k(1, 1) = 1. \quad l(1, 1) = -1,$$

and

$$\begin{aligned} h(x, y)^5 + xk(x, y)^2 - y + l(x, y) &= 0 \\ k(x, y)^5 + yh(x, y)^2 - x + l(x, y) &= 0 \\ l(x, y)^4 + y^5 - x^4 &= 1 \qquad \text{for all } (x, y) \in B_r(1, 1). \end{aligned}$$

Moreover, we have

$$-\begin{pmatrix} 5 & 2 & 1 \\ 2 & 5 & 1 \\ 0 & 0 & -4 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ -4 & 5 \end{pmatrix} = \begin{pmatrix} -4/21 & 13/84 \\ 10/21 & -43/84 \\ -1 & 5/4 \end{pmatrix},$$

so

$$\begin{aligned} \frac{\partial u}{\partial x}(1, 1) &= \frac{-4}{21} & \qquad \frac{\partial u}{\partial y}(1, 1) &= \frac{13}{84} \\ \frac{\partial v}{\partial x}(1, 1) &= \frac{10}{21} & \qquad \frac{\partial v}{\partial y}(1, 1) &= \frac{-43}{84} \\ \frac{\partial w}{\partial x}(1, 1) &= -1 & \qquad \frac{\partial w}{\partial y}(1, 1) &= \frac{5}{4}. \end{aligned}$$