Inverse Function Theorem

The contraction mapping theorem is a convenient way to prove existence theorems such as the Inverse Function Theorem in multivariable calculus.

Recall that a map $f\colon U \to \mathbb{R}^n$ (where U is open in $\mathbb{R}^n$) is *differentiable* at a point $x \in U$ if we can write

$$f(x+h) = f(x) + Ah + e(h), \tag{1}$$

where $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation (equivalently, an $n \times n$ matrix) and $\|e(h)\|/\|h\| \to 0$ as $h \to 0$. The *derivative* of f at x is $f'(x) = A$. If $f'(x)$ is a continuous function of x, f is said to have *class* C^1. The *norm* $\|A\|$ of a linear transformation A is defined to satisfy $\|Ax\| \le \|A\|\|x\|$ for all x.

We need the following standard result.

LEMMA 2 *Suppose $\|f'(x)\| \le M$ on some disk D. Then for any points $x, x+h \in D$, $\|f(x+h) - f(x)\| \le M\|h\|$.* □

THEOREM 3 *Let U be an open set in $\mathbb{R}^n$, and $f\colon U \to \mathbb{R}^n$ a differentiable function of class C^1. Suppose $x_0 \in U$ is a point where $f'(x_0)$ is invertible. Then there exist neighborhoods $U' \subset U$ of x_0 and V of $y_0 = f(x_0)$ such that $f|U'\colon U' \to V$ is a bijection. Further, the inverse $g\colon V \to U'$ is differentiable of class C^1, with derivative $g'(y) = f'(g(y))^{-1}$.* (That is, f is a local *diffeomorphism* near x_0.)

Proof We first simplify. We may assume $x_0 = y_0 = 0$. Further, we may assume $f'(0) = I$; otherwise we replace f by $B^{-1} \circ f$, where $B = f'(0)$ (in effect, a linear change of coordinates in one copy of $\mathbb{R}^n$).

We restrict attention to the disk $D = \{x \in \mathbb{R}^n : \|x\| \le \delta\}$, where δ is chosen small enough to make $D \subset U$ and $\|f'(x) - I\| \le \frac{1}{2}$ for all $x \in D$ (possible because f has class C^1). Put $w(x) = f(x) - x$; then $w'(x) = f'(x) - I$, and by Lemma 2, $\|w(x+h) - w(x)\| \le \frac{1}{2}\|h\|$. In terms of f, this is

$$\|f(x+h) - f(x) - h\| \le \frac{1}{2}\|h\|. \tag{4}$$

We take V as the open disk $\{y \in \mathbb{R}^n : \|y\| < \frac{1}{2}\delta\}$. We assert that given any $y \in V$, there is a unique $x \in D$ such that $f(x) = y$; we then define $g(y) = x$.

We define $u\colon D \to D$ by $u(x) = x + (y - f(x))$; then $f(x) = y$ if and only if x is a fixed point of u, i.e. $u(x) = x$. (The motivation for the definition of u is that if x is an approximation to the desired value, with error $y - f(x)$, $u(x)$ should be a better approximation, because f is in some sense close to the identity function.)

We must check that $u(x)$ does in fact lie in D. If we set $x = 0$ in (4), we obtain $\|f(h) - h\| \le \frac{1}{2}\|h\| \le \frac{1}{2}\delta$ for any $h \in D$. Replacing h here by x and using the triangle inequality, we deduce that

$$\|u(x)\| \le \|y\| + \|f(x) - x\| < \frac{1}{2}\delta + \frac{1}{2}\delta = \delta.$$

Next, we check that u is a contraction mapping (and hence automatically continuous). From the definition,

$$u(x+h) - u(x) = -f(x+h) + f(x) + h.$$

Then by (4),

$$\|u(x+h) - u(x)\| = \|f(x+h) - f(x) - h\| \le \frac{1}{2}\|h\|,$$

as required (with contracting factor $\frac{1}{2}$).

We took D closed in order to make it a complete metric space. Then the Contraction Mapping Theorem yields a unique fixed point x of u; in fact, we saw that $x = u(x)$ is an *interior* point of D. We take U' as the open set $f^{-1}(V) \cap \operatorname{Int} D$.

To see that g is continuous, take any two points $y, y+k \in V$ and put $x = g(y)$ and $x+h = g(y+k)$. Then (4) becomes $\|k-h\| \le \frac{1}{2}\|h\|$. By the triangle inequality, $\|h\| \le \|h-k\| + \|k\| \le \frac{1}{2}\|h\| + \|k\|$. So $\frac{1}{2}\|h\| \le \|k\|$ or, more conveniently,

$$\|h\| \le 2\|k\|. \tag{5}$$

We can ensure $\|h\| < \epsilon$ simply by taking $\|k\| < \frac{1}{2}\epsilon$.

As for differentiability of g, (1) becomes

$$y + k = y + A[g(y+k) - g(y)] + e(h). \tag{6}$$

Now $\|Ax - x\| = \|(A-I)x\| \le \frac{1}{2}\|x\|$, from which it is clear that A has zero kernel and is therefore invertible. We apply A^{-1} to (6) and rearrange, to obtain

$$g(y+k) = g(y) + A^{-1}k - A^{-1}e(h).$$

With the help of (5), we see that

$$\frac{\|A^{-1}e(h)\|}{\|k\|} \le \|A^{-1}\| \frac{\|e(h)\|}{\|h\|}\frac{\|h\|}{\|k\|} \le 2\|A^{-1}\| \frac{\|e(h)\|}{\|h\|} \longrightarrow 0$$

as $k \to 0$. Comparison with (1)) shows that g is differentiable at y with derivative $g'(y) = A^{-1}$. Finally, the formula for $g'(y)$ is visibly continuous in y. □

Remark If we know that g is differentiable, the chain rule gives the derivative $g'(y)$ easily. The point is to prove that g *is* differentiable.

Remark The above proof works in any Banach space, except that to see that A^{-1} exists, one writes it as the sum of the convergent series

$$A^{-1} = I - T + T^2 - T^3 + \dots,$$

where $A = I + T$ with $\|T\| \le \frac{1}{2}$.