Math 225 class notes

Elizabeth Denne

Proof of the Inverse Function Theorem

1 Inverse Function Theorem

Theorem 1. *Suppose that* $f : \mathbb{R}^n \to \mathbb{R}^n$ *is continuously differentiable in an open set containing* a *and* $\det Df(a) \neq 0$*. Then there is an open set* V *containing* a *and an open set* W *containing* $f(a)$ *such that* $f : V \to W$ *has a continuous inverse* $f^{-1} : W \to V$ *which is differentiable, and for all* $y \in W$ *satisfies*

$$Df^{-1}(y) = [Df(f^{-1}(y))]^{-1}.$$

That is, the Jacobian matrix of $Df^{-1}(y)$ is the inverse of the matrix of $Df(f^{-1}(y))]^{-1}$. Recall that a function is continuously differentiable if it is differentiable and all partial derivatives are continuous functions.

Proof. Let λ be the linear transformation $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$.

Claim 1: Without loss of generality λ is the identity, so $\lambda(h) = h$.
Proof of Claim 1: $\det Df(a) \neq 0$, so λ is invertible. Hence

$$D(\lambda^{-1} \circ f)(a) = D(\lambda^{-1}(f(a)) \circ Df(a) = \lambda^{-1} \circ Df(a) = \text{identity}$$

So if the theorem holds for $\lambda^{-1} \circ f$ it will hold for f. $\square$

Suppose $f(a + h) = f(a)$. Then

$$\frac{||f(a+h) - f(a) - \lambda(h)||}{||h||} = \frac{||\lambda(h)||}{||h||} = \frac{||h||}{||h||} = 1.$$

But, since f is differentiable at a, $\lim_{h\to 0} \frac{||f(a+h)-f(a)-\lambda(h)||}{||h||} = 0$

Observation: Using the fact that f is continuously differentiable, we can choose a rectangle U with $a \in U$ such that

1. $f(x) \neq f(a)$ if $x \in U$ and $x \neq a$.
2. $\det f'(x) \neq 0$ for $x \in U$.

3. $||\frac{\partial f^i(x)}{\partial x_j} - \frac{\partial f^i(a)}{\partial x_j}|| < \frac{1}{2n^2}$ for all i, j and all $x \in U$.

- We construct the set W as follows. Choose $r > 0$ such that $B_r(a) \subset U$. Set $W = B_{r/2}(f(a))$. Make sure you choose r small enough so that $W \subset f(U)$. You do this by picking $r < d$, where d is the minimum distance from $f(a)$ to $f(\partial U)$ (the image of the boundary of U under f).

Fact 1 If $y \in W$ and $x \in \partial U$, then $||y - f(a)|| < ||y - f(x)||$.

The plan is to

- construct f^{-1},
- construct V,
- show f^{-1} is continuous,
- show f^{-1} is differentiable.

We need two more facts to do this.
Let $g(x) = f(x) - x$. Use Observation 3 and Lemma 2 below applied to g to show:

Fact 2: If $x_1, x_2 \in U$, then $||f(x_1) - x_1 - (f(x_2) - x_2)|| < \frac{1}{2}||x_1 - x_2||$.

Fact 3: If $x_1, x_2 \in U$, then $||x_1 - x_2|| < 2||f(x_1) - f(x_2)||$.

To see Fact 3, use Fact 2 and a property of norms:

$$||x_1 - x_2|| - ||f(x_1) - f(x_2)|| \leq ||f(x_1) - x_1 - (f(x_2) - x_2)|| \leq \frac{1}{2}||x_1 - x_2||$$

- We construct f^{-1} by fixing $y \in W$ and finding a unique x in the interior of U such that $f(x) = y$. (This turns out to be an application of the contraction mapping theorem, something you'll learn in later courses.) Define $g(x) : U \to \mathbb{R}$ by

$$g(x) = ||y - f(x)||^2 = \sum_{i=1}^{n} (y_i - f^i(x))^2.$$

By the Extreme Value Theorem, g has a minimum value. (Since $g(x)$ is a continuous function (as f and norm are continuous) and is defined on a compact set U.) Let $x \in \partial U$, then by Fact 1 $g(a) < g(x)$; thus the minimum is not on the boundary of U.

Let z be a minimum is in the interior of U. This means that $\frac{\partial g}{\partial x_j} = 0$ for all j. We now compute $\frac{\partial g(z)}{\partial x_j}$ using the chain rule.

$$0 = \frac{\partial g(z)}{\partial x_j} = \sum_{i=1}^{n} 2(y_i - f^i(z)) \frac{\partial f^i(z)}{\partial x_j}.$$

Using Observation 2, we know that $\det Df(z) \neq 0$ and so each column of the Jacobian matrix must have at least one non-zero entry. This implies that there is some i for which $\frac{\partial f^i(z)}{\partial x_j} \neq 0$, allowing us to conclude that $y_i - f^i(z) = 0$. This is true for $i = 1, \dots, n$ and so $y = f(z)$. We have thus constructed z such that $f^{-1}(y) = z$.

Is this z unique? Yes. Assume by way of contradiction that there is z_1, z_2 with $z_1 \neq z_2$ but $f(z_1) = y = f(z_2)$. Then by Fact 3 we see

$$0 \leq ||z_1 - z_2|| \leq 2||f(z_1) - f(z_2)|| = 0,$$

hence $z_1 = z_2$, a contradiction.

- Let $V = \text{int}(U) \cap f^{-1}(W)$. We have thus shown that $f : V \to W$ has an inverse $f^{-1} : W \to V$.

- Check f^{-1} is continuous. Rewrite Fact 3 as

Fact 4: For y_1, y_2 in W, $||f^{-1}(y_1) - f^{-1}(y_2)|| \leq 2||y_1 - y_2||$.

We show f^{-1} is continuous at $y \in W$. Fix $\epsilon > 0$. We want $\delta > 0$ such that $0 < ||y - y_1|| < \delta$ means that $||f(y) - f(y_1)|| < \epsilon$. Fact 4 shows that $\delta < \epsilon/2$ works.

- Check f^{-1} is differentiable. Given f differentiable at $x \in V$ with $\mu = Df(x)$, we show f^{-1} is differentiable at $y = f(x) \in W$ and $Df^{-1}(y) = \mu^{-1}$. We know

$$\lim_{h \to 0} \frac{||f(x+h) - f(x) - \mu(h)||}{||h||} = 0.$$

Let's rewrite this as $f(x+h) = f(x) + \mu(h) + \phi(h)$, where $\phi(h)$ is the error, or $\phi(h) = f(x+h) - f(x) - \mu(h)$. As f is differentiable we know $\lim_{h \to 0} \frac{||\phi(h)||}{||h||} = 0$. Now $f(x+h) - f(x) = \mu(h) + \phi(h)$ hence

$$\begin{aligned} \mu^{-1}(f(x+h) - f(x)) &= \mu^{-1}(\mu(h)) + \mu^{-1}(\phi(h)) \\ &= h + \mu^{-1}(\phi(h)) \end{aligned}$$

We know $f(x+h) = y_1$ and $f(x) = y$ and let $k = y_1 - y$, so that $y_1 = y + k$. Observe that $h = (x+h) - h$ and rewrite the equation above to get

$$\mu^{-1}(y_1 - y) = f^{-1}(x+h) - f^{-1}(x) + \mu^{-1}(\phi(h)).$$

Rearranging gives

$$f^{-1}(y_1) = f^{-1}(y) + \mu^{-1}(y_1 - y) - \mu^{-1}(\phi(f^{-1}(y_1) = f^{-1}(y)))$$

or

$$f^{-1}(y+k) = f^{-1}(y) + \mu^{-1}(k) - \mu^{-1}(\phi(f^{-1}(y+k) = f^{-1}(y))).$$

To show that f^{-1} is differentiable at y, we just need to show that

$$\lim_{k \to 0} \frac{||\mu^{-1}(\phi(f^{-1}(y+k) - f^{-1}(y)))||}{||k||} = 0.$$

As μ^{-1} is linear, $||\mu^{-1}(z)|| \leq M||z||$, so this reduces to showing

$$\lim_{k \to 0} \frac{||\phi(f^{-1}(y+k) - f^{-1}(y))||}{||k||} = 0.$$

But

$$\lim_{k \to 0} \frac{||\phi(f^{-1}(y+k) - f^{-1}(y))||}{||k||} = \lim_{k \to 0} \frac{||\phi(f^{-1}(y+k) - f^{-1}(y))||}{||f^{-1}(y+k) - f^{-1}(y)||} \frac{||f^{-1}(y+k) - f^{-1}(y)||}{||k||}.$$

This limit is indeed 0. [As f^{-1} is continuous we find $\lim_{k \to 0} \frac{||\phi(f^{-1}(y+k) - f^{-1}(y))||}{||f^{-1}(y+k) - f^{-1}(y)||} = 0$ and by Fact 4 $\lim_{k \to 0} \frac{||f^{-1}(y+k) - f^{-1}(y)||}{||k||} < 2$.]

□

Lemma 2. *Let $A \subset \mathbb{R}^n$ be a rectangle and let $f : A \to \mathbb{R}^n$ be continuously differentiable. If there exists $M > 0$ such that $||\frac{\partial f^i}{\partial x_j}|| \leq M$ for x in the interior of A, then*

$$||f(x) - f(y)|| \leq n^2 M ||x - y||.$$

Proof. The idea is that if all partial derivatives are bounded, then the function does not change very much.

Now $f(x_1, \ldots x_n) = (f^1(x_1, \ldots, x_n), \ldots, f^n(x_1, \ldots, x_n))$. For each component function f^i observe that

$$\begin{aligned} f^i(y) - f^i(x) &= f^i(y_1, \ldots, y_n) - f^i(x_1, \ldots, x_n) \\ &= f^i(y_1, x_2 \ldots, x_n) - f^i(x_1, \ldots, x_n) + f^i(y_1, y_2, x_3 \ldots, x_n) - f^i(y_1, x_2 \ldots, x_n) + \ldots \\ &\cdots + f^i(y_1, \ldots, y_n) - f^i(y_1, \ldots, y_{n-1}, x_n) \end{aligned}$$

Apply the mean value theorem to each pair in the equation to find a $z_{ij} \in [x_j, y_j]$ such that

$$\begin{aligned} f^i(y_1, \ldots, y_j, x_{j+1}, \ldots, x_n) - f^i(y_1, \ldots, y_j, x_{j+1}, \ldots, x_n) &= (y_j - x_j)\frac{\partial f(z_{ij})}{\partial x_j} \\ &\leq M||y_j - x_j|| \end{aligned}$$

Then

$$\begin{aligned} ||f^i(y) - f^i(x)|| &\leq \sum_{j=1}^{n} M||y_j - x_j|| \\ &\leq nM||y - x|| \end{aligned}$$

Hence

$$\begin{aligned} ||f(y) - f(x)|| &\leq ||f^1(y) - f^1(x)|| + \cdots + ||f^n(y) - f^n(x)|| \\ &\leq n^2 M||y - x|| \end{aligned}$$

□

2 Implicit Function Theorem

Theorem 3. *Suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable in an open set containing (a, b) and $f(a, b) = 0$. Let M be the $m \times m$ matrix whose ijth entry is $D_{n+j} f^i(a, b) = \frac{\partial f^i}{\partial x_{n+j}}(a, b)$ for $1 \leq i, j, \leq m$. If $\det M \neq 0$, there is an open set $A \subset \mathbb{R}^n$ containing a and an open set $B \subset \mathbb{R}^m$ containing b, with the following property: for each $x \in A$, there is a unique $g(x) \in B$ such that $f(x, g(x)) = 0$. The function g is differentiable.*