MATH 395 Notes

Faye Jackson

October 14, 2020

More Inverse Function Theorem

Theorem (Inverse Function Theorem, IFT)**.** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a C^r-function for $r \geq 1$ and suppose $Df(x_0)$ is invertible where $x_0 \in A$. Then f is a local C^r-diffeomorphism around x_0. In other words there are open neighborhoods U of x_0 and V of $f(x_0)$ such that:*

1) f is a bijection from U to V

2) The inverse function $g : V \to U$ is C^r and $Dg(y) = [Df(x)]^{-1}$ where $x \in U$ and $y = f(x)$.

Lemma. *If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is C^1 and $Df(x_0)$ is non-singular. Then f is locally one-to-one around x_0. More strongly there is an open neighborhood U around x_0 such that for some $\alpha > 0$ we have that for all $x, y \in U$:*

$$|f(x) - f(y)| \geq \alpha \, |x - y|$$

Great!

Lemma. *Suppose $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ (where A is open) is differentiable. If f admits a local minimum (or maximum) at $x_0 \in A$, then $Df(x_0) = 0$.*

Proof. Let $u \in \mathbb{R}^n$ be arbitrary and set $\phi(t) = f(x_0 + tu)$ where $t \in (-\delta, \delta)$ for δ small enough so that $x_0 + tu$ is always in A. Since f has an extremum at x_0, then so does ϕ at 0. By the chain rule ϕ is differentiable on $(-\delta, \delta)$. Therefore $\phi'(0) = 0$, but:

$$\begin{aligned} \phi'(t) &= Df(x_0 + tu) \cdot u \\ 0 = \phi'(0) &= Df(x_0) \cdot u \end{aligned}$$

And this is true for any $u \in \mathbb{R}^n$, so $Df(x_0) = 0$.

Proof of the Inverse Function Theorem, IFT. By the first lemma there exists a neighborhood U of x_0 on which f is one-to-one. By shrinking U if necessary we may also assume that $Df(x)$ is non-singular for every $x \in U$. We may do this because $f \in C^1$ and so Df varies continuously, meaning that since $\det Df(x_0) \neq 0$ we can shrink U to get nonzero determinant all across U. Let $V = f(U)$.

Step 1 We must show V is open in $\mathbb{R}^n$. Take $y \in V$, we want to show that there exists an $\varepsilon > 0$ such that $B(y, \varepsilon) \subseteq V$. Write $y = f(x)$ for some $x \in U$. Since U is open there is some $\delta > 0$ so that $\overline{B(x,\delta)} \subseteq U$. Note that the boundary $\partial B(x, \delta) = \{z \in \mathbb{R}^n \mid |z - x| = \delta\}$ is a compact set, and so if we let $\Gamma = f(\partial B(x, \delta))$ we know that this is compact since f is continuous. Note that $y \notin \Gamma$ because f is one-to-one. Thus there is an $\varepsilon > 0$ such that $B(y, 2\varepsilon) \subseteq \Gamma^c$. We claim that $B(y, \varepsilon) \subseteq V$. To show that, let $c \in B(y, \varepsilon)$ and set:

$$
\begin{aligned}
\phi : \overline{B(x,\delta)} &\to \mathbb{R} \\
z &\overset{\phi}{\mapsto} |f(z) - c|^2
\end{aligned}
$$

Now since ϕ is a continuous function on a compact set it achieves its minimum value at some point $z_\star \in \overline{B(x,\delta)}$. We claim that $z_\star \notin \partial B(x, \delta)$, and so $z \in B(x, \delta)$. Why? Well if $z_\star \in \partial B(x, \delta)$ then $f(z_\star) \in \Gamma$ and so:

$$
\begin{aligned}
\phi(z_\star) = |f(z_\star) - c|^2 &= |f(z_\star) - y + y - c|^2 \\
&\geq (|f(z_\star) - y| - |y - c|)^2 > (2\varepsilon - \varepsilon)^2 = \varepsilon^2
\end{aligned}
$$

This is a problem since $\phi(x) = |y - c|^2 < \varepsilon^2$, but this contradicts the fact that ϕ has its minimum at $z_\star$. Therefore $z_\star \in B(x, \delta)$ since $z \in \overline{B(x,\delta)}$ and $z \notin \partial B(x, \delta)$. By Lemma 2 we must have that $D\phi(z_\star) = 0$ so we calculate the derivative.

Claim. *To justify the above we look at the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by $g(x) = |x|^2$*

Consider that:

$$\begin{aligned} g(x_1, \dots, x_n) &= \sum_{i=1}^{n} x_i^2 \\ \partial_i g(x_1, \dots, x_n) &= 2x_i \\ Dg(x_1, \dots, x_n) &= (2x_1, \dots, 2x_n) = 2x \end{aligned}$$

So then setting $F(z) = f(z) - c$ and so:

$$D\phi(z) = Dg(F(z)) \cdot DF(z) = 2F(z) \cdot Df(z) = 2(f(z) - c) \cdot Df(z)$$

This gives that:

$$0 = D\phi(z_\star) = 2(f(z_\star) - c)Df(z_\star)$$

Since $Df(z_\star)$ is invertible, this implies that $f(z_\star) - c = 0$, and so $f(z_\star) = c$. Therefore $c \in f(B(x, \delta)) \subseteq f(U)$. And so $B(y, \varepsilon) \subseteq f(U) = V$ as desired.

Great! The conclusion of Step 1 is that $f : U \to V$ is one-to-one, onto, and U, V are open. Therefore there exists an inverse function $g : V \to U$ such that $f \circ g = \mathrm{Id}_V$and $g \circ f = \mathrm{Id}_U$.

Step 2: We must show g is continuous. We need to show that $g^{-1}(U')$ is open for every open $U' \subseteq U$. This is equivalent to showing that $f(U')$ is open for any open $U' \subseteq U$. But wait! This is exactly what we did in Step 1 by replacing U by U'.

Step 3: We show that g is differentiable. To do this. Let $y \in V$ where $y = f(x)$ for some $x \in U$. Now let $E = Df(x)$, by hypothesis E is invertible. We will show that:

$$\frac{g(y + k) - g(y) - E^{-1}(k)}{|h|} \to 0 \text{ as } k \to 0$$

This result implies that g is differentiable at y and $Dg(y) = [Df(x)]^{-1}$ where $y = f(x)$. We know that if $|k|$ is small enough then $\overline{B(y, |k|)} \subseteq V$ by openness. Thus there exists some h such that $y + k = f(x + h)$ for some $x + h \in U$. And so we know $k = f(x + h) - f(x)$. Now note that $h = g(y + k) - g(y)$ and so $h \to 0$ as $k \to 0$ by continuity of g. By the differentiability of f at x we know that:

$$\begin{aligned} r(h) &:= f(x+h) - f(x) - Eh \\ &= k - Eh \\ \frac{r(h)}{|h|} &\to 0 \text{ as } |h| \to 0 \end{aligned}$$

Now we know that:

$$\begin{aligned} E^{-1}r(h) &= E^{-1}k - h = E^{-1}k - g(y+k) + g(y) \\ \frac{-E^{-1}r(h)}{|k|} &= \frac{g(y+k) - g(y) - E^{-1}k}{|k|} \end{aligned}$$

It then suffices to show that $\lim_{k\to 0} \frac{E^{-1}r(h)}{|k|} = 0$. It suffices to show that $\lim_{k\to 0} \frac{r(h)}{|k|} = 0$, since E^{-1} is linear. Writing then:

$$\frac{r(h)}{|k|} = \frac{r(h)}{|h|}\frac{|h|}{|k|}$$

Since $\frac{r(h)}{|h|} \to 0$ as $|h| \to 0$ and since $|h| \to 0$ as $|k| \to 0$ it suffices to show that $\frac{|h|}{|k|}$ is bounded by some $C > 0$ for nonzero but small enough k. Recall that:

$$\begin{aligned} h &= E^{-1}k - E^{-1}r(h) \\ |h| &= \left|E^{-1}(k - r(h)\right| \\ &\leq \|E^{-1}\| \cdot |k - r(h)| \\ &\leq \|E^{-1}\| \cdot (|k| + |r(h)|) \end{aligned}$$

Now since $\frac{r(h)}{|h|} \to 0$ as $|h| \to 0$ if $|h|$ is small enough we get:

$$\frac{|r(h)|}{|h|} \leq \frac{1}{2\|E^{-1}\|}$$

Therefore if $|k|$ is small enough then $|h|$ is small enough so that $|r(h)| \leq \frac{|h|}{2\|E^{-1}\|}$.

And therefore:

$$\begin{aligned}
|h| &\leq \|E^{-1}\| \left(|k| + \frac{|h|}{2\|E^{-1}\|} \right) \\
&= \|E^{-1}\| \, |k| + \frac{|h|}{2} \\
|h| &\leq 2\|E^{-1}\| \, |k| \\
\frac{|h|}{|k|} &\leq 2\|E^{-1}\|
\end{aligned}$$

Pulling this all together:

$$\left| \frac{g(y+k) - g(y) - E^{-1}k}{|k|} \right| = \left| \frac{E^{-1}r(k)}{|h|} \right|$$

And we know that:

$$\left| \frac{r(h)}{|k|} \right| = \frac{|r(h)|}{|h|} \cdot \frac{|h|}{|k|} \leq 2\|E^{-1}\| \frac{|r(h)|}{|h|} \to 0 \text{ as } h \to 0$$

And so since $h \to 0$ as $k \to 0$ and E^{-1} is linear, we are done, g is differentiable.

Step 4: We need to check that $g \in C^r(V)$. We have shown that $Dg(y) = [Df(x)]^{-1}$ where $y = f(x)$. We can write this as:

$$Dg = [Df]^{-1} \circ g$$

By Cramer's rule $[Df]^{-1}$ is a rational function (a polynomial over a polynomial) of the partials $\frac{\partial f_i}{\partial x_j}$, and this rational function has nonzero denominator

Recall. Cramer's rule gives you a formula for the inverse of a matrix C, namely:

$$C^{-1} = \frac{1}{\det C} \cdot [\operatorname{Adj} C]$$

We have that $\det C$ is a polynomial in entries of C and:

$$(\operatorname{Adj} C)_{ij} = \det(C_i^j)$$

Where C_i^j is the same as C except that we replace the i-th column with $\vec{e}_j$. Of course these are all polynomials in terms of the entries of C.

This implies that $[Df]^{-1}$ belongs to C^{r-1} if $f \in C^r$ because Df belongs to C^{r-1}.

Now consider that:

$$Dg = [Df]^{-1} \circ g \tag{$\star$}$$

Now we know that $g \in C^0$ and so since $[Df]^{-1} \in C^0$ we get $Dg \in C^0$. But then $g \in C^1$. Feeding this into $(\star)$ again we get that $Dg \in C^1$ if $r \geq 2$, and so $g \in C^2$. We may do this r times to obtain that $g \in C^r$.