Proof of Inverse Function Theorem

(following N. Wallach's expansion of M. Spivak's proof from "Calculus on manifolds")

Recall we're given \( f : U^\text{open} \to \mathbb{R}^n \), \( f \in C^1(U) \), \( \det J_f(a) \neq 0 \) for some \( a \in U \)

and want to exhibit open sets \( \tilde{a} \in V(c, U) \subset \mathbb{R}^n \) and \( \tilde{g} : W \to \mathbb{R}^n \)

such that (i) \( \tilde{f} : V \to W \) are inverses

(ii) \( \tilde{g} \) is differentiable on \( W \).

**First**, a reduction to ease computations: we can assume WLOG that \( Df(a) = 1_{\mathbb{R}^n} \) since if \( L = J_f(a) \) we can replace \( f \) with the composite \( U^\text{open} \xrightarrow{f} \mathbb{R}^n \xrightarrow{L^{-1}} \mathbb{R}^n \) having \( D\hat{f}(a) = L^{-1} \circ L = 1_{\mathbb{R}^n} \).

If we then find \( \hat{g} \) inverses \( \hat{g} \) for \( \hat{f} \) with

we can check that \( \tilde{g} = L \circ \hat{g} \) does the job:

**Second**, we can shrink \( U \) to a ball \( B_{\epsilon}(\tilde{a}) \) of small radius \( \delta \) about \( \tilde{a} \) so that to make these things both happen:

- \( |\partial_{xi}(\tilde{x}) - \partial_{xj}(\tilde{a})| < \frac{1}{2n^2} \) \( \forall i,j \in \{1, \ldots, n\} \) \( \forall \tilde{x} \in U \) (as \( \frac{\partial h_i}{\partial x_j} \) are continuous)

- \( \det J_f(a) \neq 0 \) \( \forall \tilde{x} \in U \) (as \( \det J_f(a) \) is continuous as a composite

\[ U \xrightarrow{\phi} \text{Mat}(m,n) \xrightarrow{det} \mathbb{R} \]

so pick \( \delta \) small enough that \( |\det J_f(\tilde{x}) - \det J_f(\tilde{a})| < \frac{1}{2} \det J_f(\tilde{a}) \)
The point of that \( \frac{1}{n^2} \) was anticipating a Lipschitz bound we'll use:

**Lemma:** If \( g: U \to \mathbb{R} \) has \( g \in C^1(U) \) and \( \left| \frac{\partial g_i}{\partial x_j}(x) \right| \leq M \) \( \forall x \in U \) open, then \( g \) satisfies a Lipschitz condition with Lipschitz constant \( n^2 M \).

i.e., \( |g(y) - g(x)| \leq n^2 M |y - x| \) \( \forall x, y \in U \).

**Proof:** Just like our proof of \( f \) is \( C^p \Rightarrow f \) is \( \text{diff} \).

Introduce intermediate points \( x_0, x_1, x_2, \ldots, x_n \) inside \( U \) such that \( x_j - x_{j-1} = \frac{1}{n} \).

and use 1-variable MVT to find \( c_{j-1}^{(i)}, c_j^{(i)}, \ldots, c_{j-1}^{(n)} \) on the segments between them having

\[
\frac{f_i(x_j) - f_i(x_{j-1})}{x_j - x_{j-1}} = \frac{\partial f_i^j(x_j)}{\partial x_j}(y_j - x_j)
\]

and so

\[
\frac{f_i(g_j) - f_i(x)}{g_j - x} = \sum_{j=1}^{n} \left( f_i(x_j) - f_i(x_{j-1}) \right) = \sum_{j=1}^{n} \frac{\partial f_i^j(x_j)}{\partial x_j}(y_j - x_j)
\]

and

\[
|f(x) - f(g)| \leq \sum_{i=1}^{n} |f_i(x) - f_i(x_{i-1})| \\
\leq \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \frac{\partial f_i^j(x_j)}{\partial x_j} \right| |y_j - x_j| \\
\leq n^2 M |y - x|
\]

This shows that on our shrunken ball \( U \), far away points have far away \( f \) values:

**Claim 1:** \( |f(x) - f(g)| \geq \frac{1}{2} |x - y| \) for \( x, y \in U \).

**Proof:** We'll apply the Lemma to \( f(x) = f(x) - x \), which has

\[
\left| \frac{\partial g_i}{\partial x_j}(x) \right| = \left| \frac{\partial f_i}{\partial x_j}(x) - \delta_{ij} \right| = \left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(x) \right| \leq \frac{1}{2n^2} = M \quad \forall x, y \in U.
\]

Now

\[
|f(x) - f(y)| = |(f(x) - x) - (f(y) - y)| \\
\leq |f(x) - f(y)| + |x - y| = \frac{1}{2} |f(x) - f(y)| + \frac{1}{2} |x - y|
\]

**Lemma** \( \leq n^2 \cdot \frac{1}{2n^2} |x - y| = \frac{1}{2} |x - y| \).
Now shrink $U$ to an even smaller radius ball $B_{\delta}(a)$ so as to make \[
abla \left| \frac{f(a+h) - f(a) - \nabla f(a) \cdot h}{|h|} < 1 \text{ \ for } |h| \leq \delta \quad \text{(using } \Delta f(a) = \frac{d}{dx} f(x) = f'(a) \text{)}\]

which then forces $f(a+h) \neq f(a)$, else \[
\left| \frac{f(a+h) - f(a) - \nabla f(a) \cdot h}{|h|} = \frac{|0-h|}{|h|} = 1. \]

Thus we now have $\exists \ x \in U$ for all $x \in U$

If we consider the continuous function $\nabla \left| \nabla f(x) - (x) \rightarrow \mathbb{R} \right|

on the compact set $\mathbb{R}^n$, it achieves some minimum value $d > 0$, so we can bound $|f(x) - f(a)| \geq d \ \forall x \in \mathbb{R}^n$.

This finally lets us define $W := B_{\delta/2}(a) = B_{\delta/2}(f(a))$

$= \left\{ y \in \mathbb{R}^n : |y - f(a)| < \frac{d}{2} \right\}$

By construction $\forall y \in W, x \in U \ \ |y - f(a)| < |f(x) - f(a)|$

Claim 2: $\forall y \in W \exists \ x \in U$ with $f(x) = y$

proof: In fact, $\exists \ x \in \overline{U}$ (closed ball) that achieves the minimum value of the continuous function $h: \overline{U} \rightarrow \mathbb{R} \ \ x \mapsto h(x) = \sum_{i=1}^{n} |y_i - f_i(x)|^2$. 
To see this, note that \( h \) cannot achieve its minimum on \( \mathcal{U} \) by the
inequality (\( *) \), so it achieves it at some \( \bar{x} \in \mathcal{U} \)
and then one must have

\[
0 = \frac{\partial h}{\partial x_j}(\bar{x}) = \sum_{i=1}^m 2(y_i - f_i(\bar{x})) \frac{\partial f_i}{\partial x_j}(\bar{x}) \quad \forall j = 1, \ldots, n
\]

\[\Rightarrow \quad \bar{y} = \left[ \nabla f(\bar{x}) \right] (\bar{y} - f(\bar{x}))\]

\[\forall \bar{x} \in \mathcal{U} \quad \Rightarrow \quad \bar{y} = \bar{y} - f(\bar{x}) \quad \text{i.e.} \quad \bar{f}(\bar{x}) = \bar{y}.
\]

\[12/3/2016 \Rightarrow \text{Uniqueness of } \bar{x} \text{ follows because we showed } \forall \bar{x}, \bar{x} \in \mathcal{U} \text{- that } \]

\[|\bar{f}(\bar{x}) - f(\bar{x})| \geq \frac{1}{2} |\bar{x} - \bar{x}'| \quad \text{so if } f(\bar{x}') = \bar{y} = f(\bar{x}) \text{ then } |\bar{x} - \bar{x}'| \leq \frac{1}{\varepsilon} \quad \text{i.e. } \bar{x}' = \bar{x}.
\]

So now if we define

\[\mathcal{V} := \{ \bar{x} \in \mathcal{U} : f(\bar{x}) \in \mathcal{W} \}\]

then as maps of sets, we have

\[\forall \bar{x} \in \mathcal{U} \quad \bar{f} \xrightarrow{\text{from }} \mathcal{V} \quad \xrightarrow{f} \mathcal{W} \quad \text{are } 2 \text{-sided inverses}
\]

\[f \circ \bar{f} = 1_\mathcal{V} \quad \bar{f} \circ f = 1_\mathcal{W}
\]

Also, it is an easy exercise to check that since

\( \bar{f} \) is continuous and \( \mathcal{W} \) is open, \( \mathcal{V} \) will also be open.

Claim 1 also shows \( \bar{f} \) is continuous, since \( \forall \varepsilon > 0 \), if we choose \( \delta = \frac{\varepsilon}{2} \)

then we find that for \( \bar{y}_1, \bar{y}_2 \in \mathcal{W} \) with \( |\bar{y}_1 - \bar{y}_2| < \delta = \frac{\varepsilon}{2} \)

the elements \( \bar{x}_1 = \bar{f}(\bar{y}_1) \) have \( \bar{y}_1 = \bar{f}(\bar{x}_1) \), so by Claim 1,

\[|f(\bar{x}_1) - f(\bar{x}_2)| \geq \frac{1}{2} |\bar{x}_1 - \bar{x}_2|
\]

\[\Rightarrow \quad |\bar{y}_1 - \bar{y}_2| = \frac{1}{2} |f(\bar{y}_1) - f(\bar{y}_2)| < \varepsilon.
\]

It remains to show \( \bar{f} \) is differentiable at every \( \bar{y} \in \mathcal{W} \).

In fact, we'll check that if \( \bar{y}(\bar{y}) = \bar{x} \in \mathcal{V} \)

\[\text{(so } \bar{f}(\bar{x}) = \bar{y} \text{) , and if } A := \nabla \bar{f}(\bar{x})
\]

then \( \bar{A}^1 = D\bar{f}(\bar{x}) \) , as we'd expect from chain rule.