Inverse Function Theorem Proof

Ameer Qaqish

Lemma 1: Suppose U is an open convex subset of $\mathbb{R}^n$ and that $F\colon U \to \mathbb{R}^m$ is C^1 and that for every $x \in U$, $\|DF(x)\| \leq M$. Then for any $x_1, x_2 \in U$,

$$\|F(x_2) - F(x_1)\| \leq M \|x_2 - x_1\|.$$

Proof: Define $\phi\colon [0,1] \to \mathbb{R}^m$ by $\phi(t) = F(x_1 + t(x_2 - x_1))$. By the chain rule, $\phi'(t) = DF(x_1 + t(x_2 - x_1))(x_2 - x_1)$. By the fundamental theorem of calculus,

$$\begin{aligned}
\|F(x_2) - F(x_1)\| &= \|\phi(1) - \phi(0)\| \\
&= \left\| \int_0^1 DF(x_1 + t(x_2 - x_1))(x_2 - x_1)\, dt \right\| \\
&\leq \int_0^1 \|DF(x_1 + t(x_2 - x_1))(x_2 - x_1)\| \, dt \\
&\leq \int_0^1 \|DF(x_1 + t(x_2 - x_1))\| \|x_2 - x_1\| \, dt \\
&\leq \int_0^1 M \|x_2 - x_1\| \, dt \\
&= M \|x_2 - x_1\|.
\end{aligned}$$

$\square$

Lemma 2: Suppose Ω is an open subset of $\mathbb{R}^n$ and $F\colon \Omega \to \mathbb{R}^n$ is C^1. Suppose $p_0 \in \Omega$ is such that $DF(p_0)$ is invertible. Then there exist open sets $U \ni p_0$, $V \ni F(p_0)$ such that $F\colon U \to V$ is a bijection and F^{-1} is differentiable at $F(p_0)$.

Proof: By a C^∞ change of coordinates (shifting F and scaling F by $DF(p_0)^{-1}$), we may assume that $p_0 = 0$, $F(0) = 0$, and $DF(0) = I$. Write $F(x) = x + R(x)$, where $R(x) = F(x) - x$. Note that $R(0) = 0$ and $DR(0) = 0$ and R is C^1 and $R(x) = o(\|x\|)$ as $x \to 0$. Choose $r > 0$ such that $B(0,r) \subset \Omega$ and

$$x \in B(0,r) \implies \|DR(x)\| < \frac{1}{2}.$$

We will prove that $B(0, \frac{r}{2}) \subset F(B(0,r))$, i.e. that for every $y \in B(0, \frac{r}{2})$, there exists $x \in B(0,r)$ such that $F(x) = y$. Let $y \in B(0, \frac{r}{2})$ be arbitrary. Motivated by Newton's method for solving $F(x) = y$, define a function $G\colon B(0,r) \to \mathbb{R}^n$ by

$$\begin{aligned} G(x) &= x - DF(0)^{-1}(F(x) - y) \\ &= x - F(x) + y \\ &= y - R(x). \end{aligned}$$

Note that $G(x) = x$ if and only if $F(x) = y$. Note that for all $x \in B(0,r)$, $\|DG(x)\| = \|DR(x)\| \leq \frac{1}{2}$. Thus for $x_1, x_2 \in B(0,r)$ we have

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

This implies that if $x_1, x_2 \in B(0,r)$ with $G(x_1) = x_1$ and $G(x_2) = x_2$, then $\|x_2 - x_1\| \leq \frac{1}{2} \|x_2 - x_1\|$, which means $x_1 = x_2$. Thus if a solution $x \in B(0,r)$ to $F(x) = y$ exists, then it is unique. We claim that G maps $\overline{B(y, 2\|R(y)\|)}$ to itself. Note that since $\|R(y)\| = \|R(y) - R(0)\| \leq \frac{1}{2}\|y\|$, $\overline{B(y, 2\|R(y)\|)} \subset \overline{B(y, \|y\|)} \subset \overline{B(0, 2\|y\|)} \subset B(0,r)$. Thus for all $x \in \overline{B(y, 2\|R(y)\|)}$ we have

$$\begin{aligned} \|G(x) - y\| &= \|R(x)\| \\ &\leq \|R(x) - R(y)\| + \|R(y)\| \\ &\leq \frac{1}{2}\|x - y\| + \|R(y)\| \\ &\leq 2\|R(y)\| . \end{aligned}$$

Thus G maps $\overline{B(y, 2\|R(y)\|)}$ to itself. By the contraction mapping theorem, there exists $x \in \overline{B(y, 2\|R(y)\|)}$ such that $F(x) = y$. Thus we have established that for every $y \in B(0, \frac{r}{2})$, there exists a unique $x \in B(0,r)$ such that $F(x) = y$, and that we actually have $x \in \overline{B(y, 2\|R(y)\|)}$. If we set $V = B(0, \frac{r}{2})$, $U = F^{-1}(V)$, then $F\colon U \to V$ is a bijection, and U is open since F is continuous and V is open. Note that for $y \in V$ we have established that $F^{-1}(y)$ is within $2\|R(y)\|$ of y. Since $R(y) = o(\|y\|)$ as $y \to 0$, this means $F^{-1}(y) = y + o(\|y\|)$. Thus $DF^{-1}(0) = I$. This proves the claim $\square$

Inverse Function Theorem: Suppose Ω is an open subset of $\mathbb{R}^n$ and $F\colon \Omega \to \mathbb{R}^n$ is C^1. Suppose $p_0 \in \Omega$ is such that $DF(p_0)$ is invertible. Then there exist open sets $U \ni p_0$, $V \ni F(p_0)$ such that $F\colon U \to V$ is a C^1 diffeomorphism.

Proof: Since F is C^1 and the set of invertible $n \times n$ matrices is open, we can choose $r > 0$ such that $B(p_0, r) \subset \Omega$ and for each $x \in B(p_0, r)$, $DF(x)$ is invertible. Applying lemma 2 to $F\colon B(p_0, r) \to \mathbb{R}^n$, we get open sets $U \subset B(p_0, r)$ containing p_0, V containing $F(p_0)$ such that $F\colon U \to V$ is a bijection. Applying lemma 2 to every $x \in U$ shows that F^{-1} is differentiable at every point in V. Using the identity $F(F^{-1}(y)) = y$ and the chain rule, we get $DF^{-1}(y) = DF(F^{-1}(y))^{-1}$. Since F^{-1}, DF, and the map $A \mapsto A^{-1}$ are continuous, DF^{-1} is continuous. Thus $F\colon U \to V$ is a C^1 diffeomorphism. $\square$