A PROOF OF THE INVERSE FUNCTION THEOREM

SUPPLEMENTAL NOTES FOR MATH 703, FALL 2005

First we fix some notation. For $x \in \mathbb{R}^n$ we denote by $\|x\| = \sqrt{\sum_{i=1}^n |x_i|^2}$ the Euclidean norm of x. Let $G \subset \mathbb{R}^n$ be an open set and let $f : G \to \mathbb{R}^m$ be differentiable at $x_0 \in G$, i.e., there exists a unique linear map $Df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(x_0 + h) - f(x_0) - Df(x_0)(h) = \epsilon(h)\|h\|,$$

where $\epsilon(h) \to 0$ in $\mathbb{R}^m$ as $h \to 0$ in $\mathbb{R}^n$. If we write $f(x) = (f_1(x), \cdots, f_m(x))$, then $Df(x_0)$ has with respect to the standard basis the matrix $\left(\frac{\partial f_j}{\partial x_i}(x_0)\right)$.

Lemma 1. *Let $G \subset \mathbb{R}^n$ be an open and convex set and let $f : G \to \mathbb{R}^m$ be differentiable on G. Let $M = \sup\left\{\left|\frac{\partial f_j}{\partial x_i}(x)\right| : x \in G, 1 \le i \le n, 1 \le j \le m\right\}$. Then*

$$\|f(x) - f(y)\| \le mnM\|x - y\|$$

for all $x, y \in G$.

Proof. Let $x, y \in G$. For $j = 1, \cdots, m$ define $g_j(t) = f_j(x + t(x - y))$. Then by the Mean Value Theorem there exists $c_j \in (0, 1)$ such that

$$f_j(y) - f_j(x) = g_j(1) - g_j(0) = g_j'(c_j).$$

Now by the chain rule we have $g_j'(c_j) = \nabla f_j(x + c_j(y - x)) \cdot (y - x)$, so

$$\begin{aligned} |g_j(c_j)| &= \left| \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x + c_j(y - x))(y_i - x_i) \right| \\ &\le M \sum_{i=1}^n |y_i - x_i| \le nM\|y - x\|. \end{aligned}$$

Hence $\|f(x) - f(y)\| \le \sum_{j=1}^m |f_j(x) - f_j(y)| \le mnM\|x - y\|$. $\square$

Now recall that if $A : \mathbb{R}^n \to \mathbb{R}^n$ is linear, then there exists a C such that $\|Ax\| \le C\|x\|$ for all x. If A is non-singular, then this implies that there exists also a C_1 such that $C_1\|x\| \le \|Ax\|$ for all x (take $y = Ax$ in the inequality $\|A^{-1}(y)\| \le C'\|y\|$). This observation can be used to show that linear maps satisfy the following theorem.

Theorem 2 (Inverse Function Theorem)**.** *Let $G \subset \mathbb{R}^n$ be an open set and let $f : G \to \mathbb{R}^m$ be continuously differentiable on G (i.e., all the partials of f are continuous on G). Let $a \in G$ be such that $\det Df(a) \neq 0$. Then there exist open sets $U \subset G$, $V \subset \mathbb{R}^m$ with $a \in U$ such that $f(U) = V$ and $f^{-1} : V \to U$ exists and is differentiable and $Df^{-1}(y) = (Df(f^{-1}(y)))^{-1}$ for all $y \in V$.*

Proof. Using the remark from above, that linear maps satisfy the theorem, we can replace f by $(Df(a))^{-1}f$ and assume that $Df(a) = I$. Then put $g(x) = f(x) - x$. Then $Dg(a) = 0$, so there exists $r > 0$ such that $\overline{B(a;r)} \subset G$ and $\sup_{x \in \overline{B(a;r)}} \left| \frac{\partial g_j}{\partial x_i}(x) \right| < \frac{1}{2n^2}$ for all $1 \le i \le n$ and all $1 \le j \le m$. Now apply the above lemma to g to get

$$\|f(x) - x - (f(y) - y)\| \le Mn^2 \|x - y\| < \frac{1}{2}\|x - y\|$$

for all $x \neq y \in \overline{B(a;r)}$. This implies that $\|f(x) - f(y)\| \ge \frac{1}{2}\|x - y\|$ for all $x, y \in \overline{B(a;r)}$. This implies that f is one-to-one on $\overline{B(a;r)}$ and has a continuous inverse on the range of this closed ball. By making r smaller if necessary, we can assume that $\det Df(x) \neq 0$ on $B(a;r)$. We shall now show that $f(B(a;r)) \supset V$, where V is open with $f(a) \in V$. Let S be the boundary of $B(a;r)$. Then S is compact, so $f(S)$ is compact and $f(a) \notin f(S)$ implies that $d = \text{dist }(f(S), f(a)) > 0$. Let $V = \{y : \|y - f(a)\| < \frac{d}{2}\}$. Then V is open and we claim that $V \subset f(B(a;r))$. To see this, let $y \in V$ and put $h(x) = \|f(x) - y\|^2$ for $x \in \overline{B(a;r)}$. Then h has a minimum on $\overline{B(a;r)}$, but this can't be on the boundary S of $\overline{B(a;r)}$ by the definition of d and V. Hence h has a minimum at $x_0 \in B(a;r)$. This implies that the derivative $Dh(x_0) = 0$, so

$$\sum_{i=1}^{n} -2(f_i(x_0) - y_i)\frac{\partial f_i}{\partial x_j}(x_0) = 0$$

for $j = 1, \cdots, n$. As $\det Df(x_0) \neq 0$ this implies that $f_i(x_0) - y_i = 0$ for $i = 1, \cdots, n$. Hence $f(x_0) = y$. This shows $V \subset f(B(a;r))$. Let $U = f^{-1}(V)$. Then U open and $f : U \to V$ is one-to-one and onto. Remains to show that f^{-1} is differentiable on V. Let $u, v \in V$ and put $x = f^{-1}(u)$, $y = f^{-1}(v)$. For v close enough to u we can write

$$f(y) - f(x) - Df(x)(y - x) = \epsilon(y - x)\|x - y\|,$$

where $\epsilon(h) \to 0$ as $h \to 0$. Put $A = Df(x) = Df(f^{-1}(u))$. Then we have

$$v - u - A(f^{-1}(v) - f^{-1}(u)) = \epsilon(f^{-1}(v) - f^{-1}(u))\|f^{-1}(v) - f^{-1}(u)\| = \epsilon'(v - u)\|v - u\|,$$

where we used that $\|f^{-1}(u) - f^{-1}(v)\| \le 2\|u - v\|$. Multiplying the above equation by A^{-1} we get

$$f^{-1}(v) - f^{-1}(u) - A^{-1}(v - u) = -A^{-1}(\epsilon'(v - u))\|v - u\|.$$

Now $A^{-1}(\epsilon'(v - u)) \to 0$ as $v \to u$ implies that f^{-1} is differentiable at u with derivative A^{-1}. □

Corollary 3. *Let $G \subset \mathbb{C}$ be open and $f \in H(G)$ such that $f'(z_0) \neq 0$. Then there exists open sets $U \subset G$ and $V \subset \mathbb{C}$ such that $z_0 \in U$ and $f : U \to V$ is one-to-one and onto and $f^{-1} : V \to \mathbb{C}$ is holomorphic.*

Proof. Consider G as an open subset of $\mathbb{R}^2$ and define $F : G \to \mathbb{R}^2$ by $F(x, y) = (u(x, y), v(x, y))$ where $f = u + iv$. Then $\det DF(x_0, y_0) = |f'(z_0)|^2 \neq 0$ (where $z_0 = x_0 + iy_0$). The matrix of $DF(x_0, y_0)$ is given by the skew symmetric matrix

$$\begin{bmatrix} \frac{\partial u}{\partial x}(x_0, y_0) & -\frac{\partial v}{\partial x}(x_0, y_0) \\ \frac{\partial v}{\partial x}(x_0, y_0) & \frac{\partial u}{\partial x}(x_0, y_0) \end{bmatrix}.$$

By the inverse function theorem there exist open sets $U \subset G$, $V \subset \mathbb{R}^2$ with $(x_0, y_0) \in U$ such that $F(U) = V$ and $F^{-1} : V \to U$ exists and is differentiable and $DF^{-1}((x, y)) = (DF(F^{-1}((x, y))))^{-1}$ for all $(x, y) \in V$. Define $g(z) = F^{-1}(x, y)$ for $z = x + iy$ with $(x, y) \in V$. Then g is the inverse of f and the real part and imaginary part of g are continuous and satisfy the Cauchy-Riemann equations on V, since the matrix $DF^{-1}((x, y)) = (DF(F^{-1}((x, y))))^{-1}$ is again skew symmetric for all $(x, y) \in V$. Hence g is holomorphic. $\square$

We conclude with the implicit function theorem. Assume we are given a system of m equations

$$\begin{aligned} f_1(x_1, \cdots, x_n, y_1, \cdots, y_m) &= 0 \\ f_2(x_1, \cdots, x_n, y_1, \cdots, y_m) &= 0 \\ \vdots \qquad\qquad\qquad &\ \ \vdots \\ f_m(x_1, \cdots, x_n, y_1, \cdots, y_m) &= 0 \end{aligned}$$

in $n + m$ variables $x_1, \cdots, x_n, y_1, \cdots, y_m$. Then the implicit function theorem will give sufficient conditions for solving $y_1, \cdots, y_m$ in terms of $x_1, \cdots, x_n$.

Theorem 4 (Implicit Function Theorem)**.** *Let $E \subset \mathbb{R}^{n+m}$ be open and $f : E \to \mathbb{R}^m$ a continuously differentiable map. Let $(x_0, y_0) \in E$ such that $f(x_0, y_0) = 0$ and $\det \left(\frac{\partial f_j}{\partial y_i}\right) \neq 0$. Then there exists an open set $U \subset \mathbb{R}^n$ with $x_0 \in U$, and a continuously differentiable map $g : U \to \mathbb{R}^m$ such that $f(x, g(x)) = 0$ for all $x \in U$.*

Proof. Define $F : E \to \mathbb{R}^{n+m}$ by $F(x, y) = (x, f(x, y))$. Then F is continuously differentiable in a neighborhood of (x_0, y_0) and $\det DF(x_0, y_0) = \det \left(\frac{\partial f_j}{\partial y_i}\right) \neq 0$. Hence by the Inverse Function Theorem there exists open $U_0 \subset E$ containing (x_0, y_0) and open $V_0 \subset \mathbb{R}^{m+n}$ containing $(x_0, 0)$ such that $F : U_0 \to V_0$ is one-to-one and onto and such that $F^{-1} : V_0 \to U_0$ is differentiable. Let $F^{-1}(x, y) = (G_1(x, y), G_2(x, y))$. Then for all $(x, 0) \in V_0$ we have $(x, 0) = F(G_1(x, 0), G_2(x, 0)) = (G_1(x, 0), f(G_1(x, 0), G_2(x, 0)))$. Hence $G_1(x, 0) = x$ and thus $f(x, G_2(x, 0)) = 0$ for all $(x, 0) \in V_0$. Let $U = \{x \in \mathbb{R}^m : (0, x) \in V_0\}$ and define $g(x) = G_2(x, 0)$ for $x \in U$. It is clear then that U and g satisfy the conclusion of the theorem. $\square$