Handout 4. The Inverse and Implicit Function Theorems

Recall that a linear map $L : \mathbf{R}^n \to \mathbf{R}^n$ with $\det L \neq 0$ is one-to-one. By the next theorem, a continuously differentiable map between regions in $\mathbf{R}^n$ is *locally one-to-one* near any point where its differential has nonzero determinant.

Inverse Function Theorem. *Suppose U is open in $\mathbf{R}^n$ and $F : U \to \mathbf{R}^n$ is a continuously differentiable mapping, $p \in U$, and the differential at p, dF_p, is an isomorphism. Then there exist neighborhoods V of p in U and W of $F(p)$ in $\mathbf{R}^n$ so that $F : V \to W$ has a continuously differentiable inverse $F^{-1} : W \to V$ with*

$$d(F^{-1})_y \;=\; \left[dF_{F^{-1}\{y\}}\right]^{-1} \text{ for } \; y \in W \;.$$

Moreover, F^{-1} is smooth (infinitely differentiable) whenever F is smooth.

Thus, the equation $y = F(x)$, written in component form as a system of n equations,

$$y_i \;=\; F_i(x_1, \ldots, x_n) \text{ for } \; i = 1, \ldots, n \;,$$

can be solved for $x_1, \ldots, x_n$ in terms of $y_1, \ldots, y_n$ provided we restrict the points x and y to small enough neighborhoods of p and $F(p)$. The solutions are then unique and continuously differentaible.

Proof : Let $L = dF_p$, and note that the number

$$\lambda \;=\; \frac{1}{2} \inf_{|v|=1} |L(v)| \;=\; \frac{1}{2 \sup_{|w|=1} |L^{-1}(w)|}$$

is positive. Since dF_x is continuous in x at $x = p$, we have the inequality

$$\sup_{|v|=1} |dF_x(v) - L(v)| \;\leq\; \lambda$$

true for all x in some sufficiently small ball V about p in U. Thus, by linearity,

$$|dF_x(v) - L(v)| \;\leq\; \lambda\, |v| \;\text{ for all }\; v \in \mathbf{R}^n \text{ and } \; x \in V \;.$$

With each $y \in \mathbf{R}^n$, we associate the function

$$A^y(x) \;=\; x \;+\; L^{-1}\big(y \;-\; F(x)\big) \;.$$

Then

$$F(x) = y \;\; if \; and \; only \; if \;\; x \; is \; a \; fixed \; point \; of \; A_y \;.$$

Since $dA^y = \mathrm{Id} - L^{-1}\big(dF_x\big) = L^{-1}\big(L - dF_x\big)$, the above inequalities imply that

$$|dA^y_x(v)| \;\le\; \frac{1}{2}|v| \;\text{ for }\; x \in V \;\text{ and }\; v \in \mathbf{R}^n \;.$$

Thus, for $w, x \in V$,

$$\begin{aligned} |A^y(w) - A^y(x)| \;&=\; \Big|\int_0^1 \frac{d}{dt} A^y\big(x + t(w-x)\big)\,dt\,\Big| \\ &\le\; \int_0^1 |dA^y_{x+t(w-x)}(w-x)|\,dt \;\le\; \frac{1}{2}|w-x| \;. \qquad (*) \end{aligned}$$

It follows that A^y has at most one fixed point in V, and there is at most one solution $x \in V$ for $F(x) = y$.

Next we verify that $W = F(V)$ is open. To do this, we choose, for any point $\tilde{w} = F(\tilde{x}) \in W$ with $\tilde{x} \in V$, a sufficiently small positive r, so that the ball $B = \mathbf{B}_r(\tilde{x})$ has closure $\overline{B} \subset V$. We will show that $\mathbf{B}_{\lambda r}(\tilde{w}) \subset W$. This will give the openness of W.

For any $y \in \mathbf{B}_{\lambda r}(\tilde{w})$, and A^y as above,

$$|A^y(\tilde{x}) - \tilde{x}| \;=\; |L^{-1}(y - \tilde{w})| \;<\; \frac{1}{2\lambda}\lambda r \;=\; \frac{r}{2} \;.$$

For $x \in \overline{B}$ it follows that

$$|A^y(x) - \tilde{x}| \;\le\; |A^y(x) - A^y(\tilde{x})| + |A^y(\tilde{x}) - \tilde{x}| \;<\; \frac{1}{2}|x - \tilde{x}| + \frac{r}{2} \;\le\; r \;.$$

So $A^y(x) \in B$. By (*) A^y thus gives a contraction of $\overline{B}$. So A_y has fixed point x in $\overline{B}$, and $y = F(x) \in F(\overline{B}) \subset F(V) = W$. Thus $\mathbf{B}_{\lambda r}(\tilde{w}) \subset W$.

Next we show that $F^{-1} : W \to V$ is differentiable at each point $y \in W$ and that

$$d(F^{-1})_y \;=\; M^{-1} \;\text{ where }\; M = dF_x \;\text{ with }\; x = F^{-1}(y) \in V \;.$$

Suppose $y + k \in W$ and $x + h = F^{-1}(y+k) \in V$. Then, with our previous notations,

$$|h - L^{-1}(k)| \;=\; |h - L^{-1}\big(F(x+h) - F(x)\big)\,| \;=\; |A^y(x+h) - A^y(x)| \le \frac{1}{2}|h| \;,$$

which implies that

$$\frac{1}{2}|h| \;\le\; |L^{-1}(k)| \;\le\; \big(\frac{1}{2\lambda}\big)|k| \;.$$

We now obtain the desired formula for $d(F^{-1})_y$ by computing that

$$\begin{aligned}\frac{|F^{-1}(y+k)-F^{-1}(y)-M^{-1}k|}{|k|} &= \frac{|h-M^{-1}k|}{|k|} \\ &= |M^{-1}\big(\frac{F(x+h)-F(x)-Mh}{|h|}\big)|\frac{|h|}{|k|} \\ &\le \frac{1}{\lambda}|M^{-1}\big(\frac{F(x+h)-F(x)-Mh}{|h|}\big)|\ ,\end{aligned}$$

which approaches 0 as $|k| \to 0$ because $M = dF_x$.

Finally, since the inversion of matrices is, by Cramer's rule, a continuous, in fact, smooth, function of the entries, we deduce from our formula that F^{-1} is continuously differentiable. Moreover, repeatly differentiating the formula shows that F^{-1} is a smooth mapping whenever F is. ■

Next we turn to the *Implicit Function Theorem.* This important theorem gives a condition under which one can locally solve an equation (or, via vector notation, system of equations)

$$f(x,y) = 0$$

for y in terms of x. Geometrically the solution locus of points (x,y) satisfying the equation is thus represented as the graph of a function $y = g(x)$. For smooth f this is a smooth *manifold.*

Let $(x,y) = \big((x_1,\ldots,x_m),(y_1,\ldots,y_n)\big)$ denote a point in $\mathbf{R}^m \times \mathbf{R}^n$, and, for an $\mathbf{R}^n$-valued function $f(x,y) = (f_1,\ldots,f_n)(x,y)$, let $d_x f$ denote the partial differential represented by the $n \times m$ matrix $\big[\frac{\partial f_i}{\partial x_j}\big]$ and $d_y f$ denote the partial differential represented by the $n \times n$ matrix $\big[\frac{\partial f_i}{\partial y_j}\big]$.

Implicit Function Theorem. *Suppose $f(x,y)$ is a continuously differentiable $\mathbf{R}^n$-valued function near a point $(a,b) \in \mathbf{R}^m \times \mathbf{R}^n$, $f(a,b) = 0$, and $\det d_y f|_{(a,b)} \ne 0$. Then*

$$\{(x,y) \in W \ :\ f(x,y) = 0\} = \{\big(x, g(x)\big) \ :\ x \in X\}$$

for some open neighborhood W of (a,b) in $\mathbf{R}^m \times \mathbf{R}^n$ and some continuously differentiable function g mapping some $\mathbf{R}^m$ neighborhood X of a into $\mathbf{R}^n$. Moreover,

$$(d_x g)_x = -(d_y f)^{-1}|_{(x,g(x))} d_x f|_{(x,g(x))}\ ,$$

and g is smooth in case f is smooth.

Proof : Define $F(x,y) = \big(x, f(x,y)\big)$, and compute that

$$\det dF_{(a,b)} \;=\; \det(d_y f)_{(a,b)} \neq 0 \ .$$

The Inverse Function Theorem thus gives a continuously differentiable inverse $F^{-1} : W \to V$ for some open neighborhoods V of (a,b) and W of $(a,0)$ in $\mathbf{R}^m \times \mathbf{R}^n$.

The set $X = \{x \in \mathbf{R}^m \ : \ (x,0) \in W\}$ is open in $\mathbf{R}^m$, and, for each point $x \in X$, $F^{-1}(x,0) = \big(x, g(x)\big)$ for some point $g(x) \in \mathbf{R}^n$. Moreover,

$$\begin{aligned}\{(x,y) \in W \ : \ f(x,y) = 0\} \;&=\; (F^{-1} \circ F)\big(W \cap f^{-1}\{0\}\big) \\ &=\; F^{-1}\big(W \cap (\mathbf{R}^m \times \{0\})\big) \;=\; \{\big(x, g(x)\big) \ : \ x \in X\} \ .\end{aligned}$$

One readily checks that g is continuously differentiable with

$$\frac{\partial g_i}{\partial x_j}(x) \;=\; \frac{\partial (F^{-1})_{m+i}}{\partial x_j}(x,0)$$

for $i = 1, \ldots, n$, $j = 1, \ldots, m$, and $x \in W$. The formula for $(d_x g)_x$ follows from differentiating the identity

$$f\big(x, g(x)\big) \;\equiv\; 0 \quad \text{on} \quad W \ ,$$

and using the chain rule. Smoothness of g follows from smoothness of f by repeatedly differentiating this identity. ∎