McGill University
Math 354: Honors Analysis 3

Inverse Function theorem in $\mathbf{R}^n$.

Our exposition follows that in Rudin's book.

Theorem 1 (Inverse Function Theorem). Let $\Omega \subset \mathbf{R}^n$ be an open set, and let $F : \Omega \to \mathbf{R}^n$ be continuously differentiable. Let $a \in \Omega$, let $b = F(a)$, and let the Jacobian $DF(a)$ be invertible. Then (a) there exists open sets U, V with $a \in U$ and $b \in V$ such that F is a bijection of U onto V; and (b) its inverse G defined by $G(F(x)) = x, x \in U$ is continuously differentiable.

Proof of Theorem 1. Let $A = DF(a)$ be the Jacobian an a. Let $\epsilon = 1/(4||A^{-1}||_{op})$; it is finite since A is invertible by assumption. Since $x \to DF(x)$ is continuous, we can choose an open ball U centered at a such that

$$||DF(x) - A||_{op} < 2\epsilon. \tag{1}$$

Let $x \in U, x + h \in U$. Let $f(t) = F(x + th) - tAh$, where $0 \le t \le 1$. Since U is convex, $x + th \in U$ for all $0 \le t \le 1$. Also,

$$||f'(t)|| = ||f'(x + th)h - Ah|| \le 2\epsilon||h||.$$

Since

$$2\epsilon||h|| = 2\epsilon||A^{-1}Ah|| \le 2\epsilon||A^{-1}||_{op}||Ah|| = ||Ah||/2,$$

by definition of ϵ. It follows that

$$||f'(t)|| \le (1/2)||Ah||.$$

By the generalization of the Intermediate Value theorem, we find (using the previous bound) that

$$||f(1) - f(0)|| \le (1/2)||Ah||,$$

or

$$||F(x + h) - F(x) - Ah|| \le (1/2)||Ah||.$$

It follows that

$$||F(x + h) - F(x)|| > (1/2)||Ah|| \ge 2\epsilon||h||. \tag{2}$$

It follows that F is 1-to-1 on U.

Let $x_0 \in U$, and let $r > 0$ be such that $\overline{B(x_0, r)} \subset U$.

Claim. $B(F(x_0), \epsilon r) \subset F(\overline{B(x_0, r)})$.

Proof of the Claim. Let $S = \overline{B(x_0, r)}$, and let $||y - F(x_0)|| < \epsilon r$. For $x \in S$, we define $\psi(x)$ to be

$$\psi(x) = ||y - F(x)||^2.$$

The function ψ is a continuous function of x, so it attains a minimum at $x_1 \in S$ (since S is compact). Moreover, ψ is differentiable at x_1 with derivative equal to $DF(x) \cdot (y - F(x))$, therefore by Lemma 132 in Drury we have

$$DF(x_1) \cdot (y - F(x_1)) = 0.$$

But $DF(x_1)$ is invertible by (1), so it follows that $y - F(x_1) = 0$ or $y = F(x_1)$. This finishes the proof of the Claim. QED

We have shown that every point in $F(U)$ is its interior point, so $F(U)$ is open, and if we let $V = F(U)$ we prove part (a) of Theorem 1.

To prove part (b) of Theorem 1, let $y \in V$, and $y + k \in V$. Let G be the inverse map of F, and let $x = G(y), h = G(y + k) - G(y)$. By the choice of V, we have

$x \in U$ and $x+h \in U$. It then follows from (1) that $DF(x)$ has an inverse map, that we denote $B = B(x)$.

Consider the identity

$$k = F(x+h) - F(x) = DF(x)h + r(h),$$

where $||r(h)||/||h|| \to 0$ as $||h|| \to 0$. Apply B to both sides of the previous identity to get $Bk = h + Br(h)$, or

$$G(y+k) - G(y) = Bk - Br(h). \tag{3}$$

It follows from (2) that $2\epsilon||h|| \leq ||k||$. It follows that $h \to 0$ as $k \to 0$, proving that G is continuous at y. Moreover,

$$\frac{||Br(h)||}{||k||} \leq \frac{||B||_{op} \cdot ||r(h)||}{2\epsilon||h||} \to 0, \qquad as\ k \to 0. \tag{4}$$

Finally, it follows from (3) and (4) that G is differentiable at y and its derivative is equal to B, i.e.

$$DG(y) = [DF(G(y))]^{-1},$$

for $y \in V$. Moreover, DG is continuous, since it is a composition of three continuous mappings (matrix inverse is continuous in a neighborhood of any invertible matrix, say by Cramer's rule). This finishes the proof of (b) and thus of Theorem 1.

QED

Implicit Function theorem in $\mathbf{R}^n$.

We let $x = (x_1, \dots, x_n) \in \mathbf{R}^n$, and $y = (y_1, \dots, y_m) \in \mathbf{R}^m$. We will denote by (x, y) the obvious vector in $\mathbf{R}^{n+m}$.

Let $M : \mathbf{R}^{n+m} \to \mathbf{R}^n$ be a linear map, represented by an $n \times (n+m)$ matrix, such that

$$M \cdot (h, 0)^t = 0 \quad \Leftrightarrow \quad h = 0, \tag{5}$$

where $(h,0)^t$ is the column vector equal to the transpose of $(h,0)$. This is equivalent to the statement that the $n \times n$ submatrix of M, consisting of the first n columns of M, is invertible, or that its rows/columns are linearly independent.

In that case, by standard results in linear algebra, $h \to M(h,0)^t : \mathbf{R}^n \to \mathbf{R}^n$ is a bijection. Also, for any fixed $k \in \mathbf{R}^n$ and $b \in \mathbf{R}^n$, the equation $M \cdot (x,k)^t = b$ has a unique solution. Indeed, letting $d = b - M \cdot (0,k)^t$, there exists x such that $M \cdot (x,0)^t = d$, or $M \cdot (x,k)^t = b$, and uniqueness of x follows easily from (5).

Theorem 2 (Implicit Function Theorem). Let $E \subset \mathbf{R}^{n+m}$ be an open set, and let $F : E \to \mathbf{R}^n$ be continuously differentiable. Let $(a,b) \in E, F(a,b) = 0$, let $M = DF((a,b))$, and let M satisfy the condition (5). Then there exists a neighborhood $W \subset \mathbf{R}^m$ of the point b, and a unique function $G : W \to \mathbf{R}^n$ such that $G(b) = a$ and $F(G(y), y) = 0$ for all $y \in W$. Moreover, G is continuously differentiable.

Remark. Denote by D_xF, the matrix $\partial(F_1, \dots, F_n)/\partial(x_1, \dots, x_n)$. Theorem 2 assumes that the matrix D_xF, is invertible at a. If we denote by D_yF the matrix $\partial(F_1, \dots, F_n)/\partial(y_1, \dots, y_m)$, then it follows from Theorem 2 and the Chain rule that

$$\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_m)} = -(D_xF)^{-1}(D_yF).$$

Proof of Theorem 2. Define $H : E \subset \mathbf{R}^{n+m} \to \mathbf{R}^{n+m}, H(x,y) = (z,w)$ by

$$z = F(x,y), w = y. \tag{6}$$

Then F is a continuously differentiable on E. Since $F(a,b)=0$, we have

$$F(a+h,b+k) = A\cdot(h,k)^t + r(h,k),$$

where $r(h,k)$ is little o norm at (a,b) and A is a linear operator represented by an $n\times(n+m)$ matrix $(D_xF|D_yF)$. Since

$$H(a+h,b+k) - H(a,b) = (F(a+h,b+k),k),$$

we find that

$$DH(a,b) = \begin{pmatrix} D_xF(a,b) & D_yF(a,b) \\ 0 & Id \end{pmatrix},$$

where 0 and Id denote the $m\times n$ zero matrix and the $m\times m$ identity matrix respectively.

Since $D_xF(a,b)$ is invertible by assumption, it follows from the standard properties of matrices that $DH(a,b)$ is invertible, and so we can apply Theorem 1 to the function H. Thus, there exist two open sets: U containing (a,b) and V containing $(0,b)$ such that $H: U\to V$ is a bijection, and has a continuously differentiable inverse $H^{-1}: V\to U$.

It follows from (6) that $H^{-1}:(z,w)\to(x,y)$ has the form

$$x=\Phi(z,w), y=w, \tag{7}$$

where $\Phi\in C^1(V)$, or

$$F(\Phi(z,w),w)=z, \qquad (z,w)\in V. \tag{8}$$

Now, let W be a neighborhood of b such that $(0,w)\in V, \forall w\in W$. We define $G(y)=\Phi((0,y))$ for $y\in W$. Substituting into (8), we get $F(G(y),y)=0$, which is the formula claimed in Theorem 2. Since $\Phi(0,b)=a$, we have $G(b)=a$, and uniqueness follows easily from the fact that H is one-to-one. This finishes the proof of Theorem 2.

QED

Example. Let (u,v,x,y) satisfy

$$\begin{cases} u^3+xv-y=0; \\ v^3+yu-x=0. \end{cases} \tag{9}$$

Prove that u,v are continuously differentiable functions of x,y near $x=0, y=1, u=1, v=-1$, and compute $\partial u/\partial x$ and $\partial v/\partial x$.

Solution: We denote by $F_1(x,y,u,v)$ the left-hand side of the 1st equation in (9); and by $F_2(x,y,u,v)$ the left-hand side of the 2nd equation in (9). Then

$$\left.\frac{\partial(F_1,F_2)}{\partial(u,v)}\right|_{(0,1,1,-1)} = \begin{pmatrix} 3u^2 & x \\ y & 3v^2 \end{pmatrix}_{(0,1,1,-1)} = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$$

and is thus invertible, so by Implicit Function theorem, (u,v) are C^1 functions of (x,y). To compute the x-derivatives, we differentiate (9) with respect to x and use the Chain rule; note that $\partial F_1/\partial x = v = -1$ at $(0,1,1,-1)$, and $\partial F_2/\partial x = -1$ at the same point. We get:

$$\begin{cases} (F_1)_u\cdot u_x + (F_1)_v\cdot v_x + (F_1)_x = 3u_x - 1 = 0; \\ (F_2)_u\cdot u_x + (F_2)_v\cdot v_x + (F_2)_x = u_x + 3v_x - 1 = 0. \end{cases}$$

The solution is $u_x = 1/3, v_x = 2/9$.