Lecture 7

We continue our proof of the Inverse Function Theorem.

As before, we let U be an open set in $\mathbb{R}^n$, and we assume that $0 \in U$. We let $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^1$ map, and we assume $f(0) = 0$ and that $Df(0) = I$. We summarize what we have proved so far in the following theorem.

Theorem 2.18. *There exists a neighborhood U_0 of 0 in U and a neighborhood V of 0 in $\mathbb{R}^n$ such that*

1. *f maps U_0 bijectively onto V*
2. *$f^{-1} : V \to U_0$ is continuous,*
3. *f^{-1} is differentiable at 0.*

Now, we let U be an open set in $\mathbb{R}^n$, and we let $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^2$ map, as before, but we return to our original assumptions that $a \in U$, $b = f(a)$, and $Df(a) : \mathbb{R}^n \to \mathbb{R}^n$ is bijective. We prove the following theorem.

Theorem 2.19. *There exists a neighborhood U_0 of a in U and a neighborhood V of b in $\mathbb{R}^n$ such that*

1. *f maps U_0 bijectively onto V*
2. *$f^{-1} : V \to U_0$ is continuous,*
3. *f^{-1} is differentiable at b.*

Proof. The map $f : U \to \mathbb{R}^n$ maps a to b. Define $U' = U - a = \{x - a : x \in U\}$. Also define $f_1 : U' \to \mathbb{R}^n$ by $f_1(x) = f(x + a) - b$, so that $f_1(0) = 0$ and $Df_1(0) = Df(a)$ (using the Chain Rule).

Let $A = Df(a) = Df_1(0)$. We know that A is invertible.

Now, define $f_2 : U' \to \mathbb{R}^n$ by $f_2 = A^{-1} f_1$, so that $f_2(0) = 0$ and $Df_2(0) = I$. The results from last lecture show that the theorem at hand is true for f_2. Because $f_1 = A \circ f_2$, the theorem is also true for f_1. Finally, because $f(x) = f_1(x - a) + b$, the theorem is true for f. □

So, we have a bijective map $f : U_0 \to V$. Let us take $c \in U_0$ and look at the derivative

$$Df(c) \sim \left[\frac{\partial f_i}{\partial x_j}(c)\right] = J_f(c). \tag{2.97}$$

Note that

$$Df(c) \text{ is bijective} \iff \det\left[\frac{\partial f_i}{\partial x_j}(c)\right] \neq 0. \tag{2.98}$$

Because f is $\mathcal{C}^1$, the functions $\frac{\partial f_i}{\partial x_j}$ are continuous on U_0. If $\det J_f(a) \neq 0$, then $\det J_f(c) \neq 0$ for c close to a. We can shrink U_0 and V such that $\det J_f(c) \neq 0$ for all $c \in U_0$, so for every $c \in U_0$, the map f^{-1} is differentiable at $f(c)$. That is, f^{-1} is differentiable at all points of V.

We have thus improved the previous theorem. We can replace the third point with

$$3.\ f^{-1} \text{ is differentiable at all points of } V. \tag{2.99}$$

Let $f^{-1} = g$, so that $g \circ f =$ identity map. The Chain Rule is used to show the following. Suppose $p \in U_0$ and $q = f(p)$. Then $Dg(q) = Df(p)^{-1}$, so $J_g(q) = J_f(p)^{-1}$. That is, for all $x \in V$,

$$\left[\frac{\partial g_i}{\partial x_j}(x)\right] = \left[\frac{\partial f_i}{\partial x_j}(g(x))\right]^{-1}. \tag{2.100}$$

The function f is $\mathcal{C}^1$, so $\frac{\partial f_i}{\partial x_j}$ is continuous on U_0. It also follows that g is continuous, so $\frac{\partial f_i}{\partial x_j}(g(x))$ is continuous on V.

Using Cramer's Rule, we conclude that the entries of matrix on the r.h.s. of Equation 2.100 are continuous functions on V. This shows that $\frac{\partial f_i}{\partial x_j}$ is continuous on V, which implies that g is a $\mathcal{C}^1$ map.

We leave as an exercise to show that $f \in \mathcal{C}^r$ implies that $g \in \mathcal{C}^r$ for all r. The proof is by induction.

This concludes the proof of the Inverse Function Theorem, signifying the end of this section of the course.

□

3 Integration

3.1 Riemann Integral of One Variable

We now begin to study the next main topic of this course: integrals. We begin our discussion of integrals with an 18.100 level review of integrals.

We begin by defining the Riemann integral (sometimes written in shorthand as the R. integral).

Let $[a, b] \subseteq \mathbb{R}$ be a closed interval in $\mathbb{R}$, and let P be a finite subset of $[a, b]$. Then P is a *partition* if $a, b \in P$ and if all of the elements $t_i, \dots, t_N$ in P can be arranged such that $t_1 = a < t_2 < \cdots < t_n = b$. We define $I_i = [t_i, t_{i+1}]$, which are called the subintervals of $[a, b]$ belonging to P.

Let $f : [a, b] \to \mathbb{R}$ be a bounded function, and let I_i be a subinterval belonging to P. Then we define

$$\begin{aligned} m_i &= \inf f : I_i \to \mathbb{R} \\ M_i &= \sup f : I_i \to \mathbb{R}, \end{aligned} \tag{3.1}$$

from which we define the *lower* and *upper* Riemann sums

$$\begin{aligned} L(f, P) &= \sum_i m_i \times (\text{length of } I_i) \\ U(f, P) &= \sum_i M_i \times (\text{length of } I_i), \end{aligned} \tag{3.2}$$

respectively.

Clearly,

$$L(f, P) \le U(f, P). \tag{3.3}$$

Now, let P and P' be partitions.

Definition 3.1. The partition P is a *refinement* of P if $P' \supset P$.

If P' is a refinement of P, then

$$\begin{aligned} L(f, P') &\ge L(f, P), \text{ and} \\ U(f, P') &\le U(f, P). \end{aligned} \tag{3.4}$$

If P and P' are *any* partitions, then you can take $P'' = P \cup P'$, which is a refinement of both P and P'. So,

$$L(f, P) \le L(f, P'') \le U(f, P'') \le U(f, P') \tag{3.5}$$

for any partitions P, P'. That is, the lower Riemann sum is always less than or equal to the upper Riemann sum, regardless of the partitions used.

Now we can define the Lower and Upper Riemann integrals

$$\begin{aligned} \underline{\int}_{[a,b]} f &= \text{l.u.b. } \{L(f, P) | P \text{ a partition of } [a, b]\} \\ \overline{\int}_{[a,b]} f &= \text{g.l.b. } \{U(f, P) | P \text{ a partition of } [a, b]\} \end{aligned} \tag{3.6}$$

We can see from the above that

$$\underline{\int} f \le \overline{\int} f. \tag{3.7}$$

Claim. *If f is continuous, then*

$$\underline{\int} f = \overline{\int} f. \tag{3.8}$$

Definition 3.2. For any bounded function $f : [a, b] \to \mathbb{R}$, the function f is *(Riemann) integrable* if

$$\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f. \tag{3.9}$$

In the next lecture we will begin to generalize these notions to multiple variables.