MA 200 - Lecture 10

1 Recap

1. Proved that the image of an open neighbourhood of a is open using three methods.

2 Inverse function theorem

Proof. We first prove for $r = 1$:

1. f is locally $1 - 1$.

2. The image of a neighbourhood of f is open.

3. f^{-1} is continuous: Note that on U, Df is invertible at every point $x \in U$. Hence, applying the previous step to each x, we conclude that the image of every open subset of U is open (how?) Hence $f^{-1} : f(U) \to U$ is open and hence f^{-1} is continuous (why?)

4. f^{-1} is C^1: Firstly, f^{-1} is differentiable with derivative $[Df_{f^{-1}(b)}]^{-1}$. Indeed, $f^{-1}(b+h) - f^{-1}(b) = k$. Thus $b + h = f(f^{-1}(b) + k)$. Thus, for each component $h_i = \langle \nabla f_i(\theta_i), k\rangle$ (by MVT). In other words, $h = Bk$ where the i^{th} row of B is $\nabla f_i(\theta_i)$ (where $\theta_i \to f^{-1}(b)$ as $k \to 0$. Thus by continuity of f^{-1}, as $h \to 0$, $B \to Df_{f^{-1}(b)}$).
   This implies that $\frac{\|k-[Df_{f^{-1}(b)}]^{-1}h\|}{\|h\|} = \frac{\|B^{-1}h-[Df_{f^{-1}(b)}]^{-1}h\|}{\|h\|} \leq \|B^{-1} - [Df_{f^{-1}(b)}]^{-1}\| \to 0$ as $h \to 0$.
   Since $Df^{-1}(y) = [Df_{f^{-1}(y)}]^{-1}$, by properties of continuity f is C^1.

Now we shall prove it for general r: Assume it is true for $1, 2, \ldots, r-1$. Then $D(f^{-1})_y = [Df_{f^{-1}(y)}]^{-1}$. This is C^{r-1} (why?) Hence, f^{-1} is C^r. □

The IFT motivates a definition: Let $U, V \subset \mathbb{R}^n$ be open subsets. A function $f : U \to V$ that is $1 - 1$, onto, C^r, and whose inverse is also C^r (where $1 \leq r \leq \infty$) is called a C^r-diffeomorphism between U and V. (If $r = 0$, it is called a homeomorphism.)
So IFT states that if Df_a is invertible, then f is a local C^r diffeomorphism. Now $(r, \theta) \in (0, \infty) \times (0, 2\pi) \to (r\cos(\theta), r\sin(\theta)) \in \mathbb{R}^2 - \{(x, y)|x = 0, y \geq 0\}$ is an example of a C^∞ diffeomorphism. These are the famous polar coordinates. The terminology raises a question? Are there other coordinates? What are coordinates after all? (They are called “frames of reference" in physics) We will answer these questions some time but basically, diffeomorphisms are used to get new coordinate systems. (Einstein wanted the laws of physics to remain invariant under change of frame, i.e., under

diffeomorphisms. Newton merely wanted them to be invariant under *linear* changes with time remaining unchanged, i.e., inertial frames.)

3 Implicit function theorem

Recall that one of the points of the IFT is to solve *certain* systems of n *nonlinear* equations with n unknowns $f(x) = b$ where b is near $f(a)$. The IFT shows that not only do solutions exist near a, but also, they vary nicely (in a C^r manner) as the RHS varies. What if (like in linear algebra) we want to solve m equations with n unknowns? (where $m \leq n$), then just like in linear algebra, if the equations are "independent" in some sense, then we ought to have $n - m$ "free parameters". Let's look at an example: $x^2 + y^2 = 1$. Of course $y = \pm\sqrt{1 - x^2}$. This example tells us that

1. The solution need not be unique.
2. *Locally*, one can hope for a unique solution. But even this need not be true at some points (like $x = 1$).
3. The solution can fail to be differentiable at some points.
4. It might be prudent to interchange the roles of the "independent/free variables" and the "dependent variables", i.e., $x = \pm\sqrt{1 - y^2}$ makes more sense near $y = 0$.

Here is another example: $x^2 + y^2 + e^{y^4}\sin^2(x^3) = 1$. Obviously it is hard to solve for (if it is possible at all) for one variable in terms of another. Even if we manage to do so, let's say $y = g(x)$, this function is not going to be as *explicit* (that is, a combination of known functions) as the previous one (one can attempt to make this sort of a thing precise using Galois theory). Note that when $x = 0, y = \pm 1$. Near $(0, 1)$, the "bad term" is roughly of the order of x^6 and hence it is not shocking to claim that we can perhaps solve for y in terms of x near $(0, 1)$ in perhaps a smooth manner.

So we arrive at this question: Suppose $f(x, y) : U \subset \mathbb{R}^2 \to \mathbb{R}$ is a C^1 function, and $f(a, b) = c$, then when can we expect that y can locally (near (a, b)) be solved for in a C^1 manner in terms of x? The answer as usual is obtained by looking at the linear approximation of f: $f(x, y) \approx f(a, b) + (x - a)f_x(a, b) + (y - b)f_y(a, b)$. That is, $f(x, y) = c$ when $(x - a)f_x(a, b) + (y - b)f_y(a, b) \approx 0$. Thus, if $f_y(a, b) \neq 0$, we expect y to be solvable in terms of x. Here is an easy proposition (Chain rule) that fortifies this expectation:

Theorem 1. *Let $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ be a C^1 function (and U be an open set). Suppose $f(a, b) = c$. Assume that there is a C^1 function $g : (a - \epsilon, a + \epsilon) \to \mathbb{R}$ such that $(x, g(x)) \in U \ \forall\ x$ such that $f(x, g(x)) = c \ \forall\ x$. Also assume that $f_y(a, b) \neq 0$. Then $g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))}$.*

This technique is called implicit differentiation. However, the drawback is that we already needed to know that $g(x)$ existed. Ideally, we would want an IFT-type existence theorem:

Theorem 2 (Implicit function theorem in two variables)**.** *Let $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ be a C^1 function (and U be an open set). Suppose $f(a, b) = c$ and $f_y(a, b) \neq 0$. Then there exists* *a neighbourhood* $(a-\epsilon, a+\epsilon)$ *of* a *and a* C^1 *function* $g : (a-\epsilon, a+\epsilon) \to \mathbb{R}$ *such that* $(x, g(x)) \in U \; \forall \, x$ *and* $f(x, g(x)) = c \; \forall \, x$.

Proof. Basically, given x, we want to solve for y from $f(x, y) = c$. Recall that the IFT does something like this but the RHS is allowed to change in IFT (not the LHS). So what if we want to convert it into two equations by not fixing x but solving for it trivially? That is, consider $h(x, y) = (x, f(x, y))$. IFT then states that if $Dh_{(a,b)}$ is invertible, then h is a local C^1 diffeomorphism, that is, $(x, f(x, y)) = (p, c)$ can be solved for x, y in terms of p and c in a C^1 manner locally. Since $x = p$, we see that y is a local C^1 function of x, c. (In particular, if you fix c, it is a C^1 function of x.) (So why is $Dh_{a,b}$ invertible?) $\square$