MA 200 - Lecture 8

1 Recap

1. Inverse function theorem (

   Theorem 1. *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ be a C^r ($\infty \geq r \geq 1$) function on an open set U. If Df_a is invertible, then there is a neighbourhood V of a such that $f(V)$ is open, $f : V \to f(V)$ is 1-1, onto, and $f^{-1} : f(V) \to V$ is C^r.*

)

2 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 2. *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 3 (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?) $\square$

More generally, we have

Theorem 4 (Implicit function theorem)**.** *Let $f(x,y) : \tilde{U} \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be a C^r function (and $\tilde{U}$ be an open set). Suppose $f(a,b)=c$ and $D_y f_{a,b}$ is invertible. Then there exists a connected neighbourhood A of a and a connected neighbourhood B of b such that $A \times B \subset \tilde{U}$, and a* unique *$C^r$ function $g: A \to \mathbb{R}^m$ such that $f(x,y)=c$ and $(x,y) \in A \times B$ if and only if $y = g(x)$.*

Proof. As before, consider $H(x,y) = (x, f(x,y))$. This function is also C^r (why?). Moreover, $DH_{(a,b)} = \begin{pmatrix} I & 0 \\ D_x f_{a,b} & D_y f_{a,b} \end{pmatrix}$. This is invertible (why?) Now $H(a,b) = (a,c)$. The IFT shows that there are (connected) neighbourhoods $U \times V \subset \tilde{U}$ and W of (a,b) and (a,c) respectively such that $H : U \times V \to W$ is a C^r diffeomorphism. This means that $H^{-1}(p,q) : W \to U \times V$ exists and is a C^r map. Thus $x = h_1(p,q)$ and $y = h_2(p,q)$ such that $H(h_1,h_2) = (p,q)$, i.e., $h_1(p,q) = p$ and $f(h_1(p,q), h_2(p,q)) = q$. So consider small enough neighbourhoods (which are connected) $A \subset U$ of a, B of b, C of c such that $A \times C \subset W$, and $B \subset h_2(A \times C)$. Now $y = h_2(x,c)$ does the job. (In fact, it also shows that y is C^r function of x,c taken together.)
Uniqueness is straightforward because of the existence of H^{-1}. $\square$

Remark: There is nothing special about the last few coordinates. You are allowed to permute them, i.e., solve for some in terms of the others.
Examples/Non-examples:

1. Let $f(x, y) = x^2 - y^3$. Then $\nabla f(0, 0) = (0, 0)$. Therefore, it appears that the origin is a problematic point. (The graph looks rather singular.) Indeed, there is no C^1 way to solve for x in terms of y or vice-versa (we can solve for y in terms of x uniquely but the expression is not C^1!)

2. $f(x, y) = y^2 - x^4$. Here, we cannot solve for y in terms of x uniquely, but the two solutions are smooth. Elsewhere, there is no such problem.

3. If $A \in Mat_{2\times 2}(\mathbb{R})$, when is there B such that $B^2 = A$? The answer is NO in general even if the eigenvalues are non-negative (standard non-diagonalisable matrix). On the other hand, there exists a neighbourhood of the identity such that every matrix in this neighbourhood has a square root: $F(A) = A^2$ is C^1 and $DF_I(H) = 2H$, which is invertible and hence by IFT, we are done.