MA 200 - Lecture 6

1 Recap

1. Proved some properties of derivatives (including $\nabla f \equiv 0$ implies f is a constant on each connected component).
2. Computed the derivatives of some functions of matrices using the definition.

We need an analogue of the chain rule to conclude that $\sin(x^2+y^2)$ is differentiable without using C^1 implies differentiability. More importantly, we need a formula for the derivative. This will help us answer two kinds of questions:

1. If a particle is zooming around in a room (whose temperature is $T(x,y,z)$) with trajectory $\vec{r}(t)$, then how fast does the temperature change according to it?
2. To solve partial differential equations that have radial symmetry, it helps to switch to polar coordinates. If we know the derivatives of a function $f(x,y)$ in Cartesian coordinates, how can we calculate those of $\tilde{f}(r,\theta) = f(r\cos(\theta), r\sin(\theta))$ in polar coordinates?

Intuitively, $T(x(t+h), y(t+h), z(t+h)) \approx T(x(t)+hx'(t), y(t)+hy'(t), z(t)+hz'(t)) \approx T(x(t),y(t),z(t)) + hx'(t)\frac{\partial T}{\partial x} + hy'(t)\frac{\partial T}{\partial y} + hz'(t)\frac{\partial T}{\partial z}$. Likewise, $f((r+h)\cos(\theta+k), (r+h)\sin(\theta+k)) - \tilde{f}(r,\theta) \approx f(r\cos(\theta)+h\cos(\theta)-rk\sin(\theta), r\sin(\theta)+h\sin(\theta)+rk\cos(\theta)) - \tilde{f}(r,\theta) \approx f_x(h\cos(\theta)-rk\sin(\theta)) + f_y(r\sin(\theta)+h\sin(\theta)+rk\cos(\theta)) = h(f_x x_r + f_y y_r) + k(f_x x_\theta + f_y y_\theta)$. In other words, we expect that $D\tilde{f} = Df D\vec{r}$ (as matrices). This expectation is very similar to the one-variable chain rule $(f\circ g)'(a) = f'(g(a))g'(a)$.

The rigorous statement of the chain rule is as follows. (One can easily see that the above examples and the one-variable chain rule are special cases of this general formulation.)

Theorem 1. *Let $A \subset \mathbb{R}^m, B \subset \mathbb{R}^n$ and $f: A \to \mathbb{R}^n$, $g: B \to \mathbb{R}^p$ with $f(A) \subset B$. Suppose a is an interior point of A and $f(a)$ is an interior point of B. If f is differentiable at a and g is differentiable at b, then $g\circ f$ is differentiable at a. Moreover, $D(g\circ f)_a = Dg_{f(a)}Df_a$ (as multiplication of matrices).*

Proof. Since $b = f(a)$ is an interior point, $g(y)$ is defined on $|y-b| < \epsilon_1$ for some ϵ_1. Since f is diff, it is continuous. Hence, $|f(x)-b| < \epsilon_1$ for all $x \in A$ such that $|x-a| < \delta_1$ and δ_1 can be chosen to be so small that $|x-a| < \delta_1$ is contained in A (because a is an interior point too).

Let $\Delta_1(h) = f(a+h)-f(a)-Df_a(h)$ and $\Delta_2(\tilde{h}) = g(b+\tilde{h})-g(b)-Dg_b\tilde{h}$. By definition of the differentiability of g at b, we see that for every $1 > \epsilon > 0$ there exists a $\delta_1 > \delta_2 > 0$ such that $\|\Delta_2(\tilde{h})\| < \frac{\epsilon}{2+\|Df_a(h)\|}\|\tilde{h}\|$ whenever $\|\tilde{h}\| < \delta_2$. Now $g(f(a+h))-g(f(a)) = g(b+Df_a(h)+\Delta_1(h))-g(b) = \Delta_2(Df_a(h)+\Delta_1(h))+Dg_b(Df_a(h)+\Delta_1(h))$. Therefore, $\|g(f(a+h)) - g(f(a)) - Dg_bDf_a(h)\| \leq \|\Delta_2(Df_a(h) + \Delta_1(h))\| + \|Dg_b(\Delta_1(h))\| \leq \|\Delta_2(Df_a(h) + \Delta_1(h))\| + \|Dg_b\|\|\Delta_1(h)\|$. Therefore, if $\|Df_a(h) + \Delta_1(h)\| < \delta_2 < \delta_1$, then the first term is less than $\frac{\epsilon}{2+\|Df_a(h)\|}\|Df_a(h) + \Delta_1(h)\|$. By definition of differentiability of f at a, we see that $\|\Delta_1(h)\| < \frac{\epsilon}{2+\|Df_a(h)\|}\|h\|$ whenever $\|h\| < \delta_3 < \delta_2$. In other words,

$$\|g(f(a+h)) - g(f(a)) - Dg_bDf_a(h)\| \leq \frac{\epsilon}{2+\|Df_a(h)\|}(\|Df_a(h) + \Delta_1(h)\| + \|h\|) \quad (1)$$

Now choose $\delta = \delta_3$ and we will be done (why?) □

Using the one-variable MVT, we can prove the multivariable one:

Theorem 2. *Let $U \subset \mathbb{R}^n$ be an open set and $a, b \in U$ be points such that the line segment $ta + (1-t)b \ \forall \ t \in [0,1]$ lies in U. Let $f : U \to \mathbb{R}$ be a differentiable function. Then $f(a) - f(b) = \langle \nabla f(\theta a + (1-\theta)b), a - b\rangle$ for some $\theta \in (0,1)$.*

Caution: This theorem is false for vector-valued functions!

Since we have defined differentiability for vector-valued functions, we can ask whether the derivative function Df is differentiable in its own right. We can also separately ask about each partial derivative being further partially differentiable and so on. There are examples (HW) of functions that are C^1, the second partials exist, but the mixed partials, i.e., $D_iD_jf = \frac{\partial^2 f}{\partial x_i \partial x_j}$ are not equal. To this end, we define the notion of being C^2: A function $f : U \to \mathbb{R}^m$ (where U is open) is said to be C^2 on U if Df is C^1 on U (and hence Df is continuous and hence f is C^1 and continuous). More generally, f is C^r on U if all the $r-1$ partials exist and are C^1 (thus, the $r-1$ partials are differentiable and hence continuous. By induction, all the partials upto order r are continuous and f is r-times differentiable).