MA 200 - Lecture 5

1 Recap

1. Proved the formula for $L(h)$ if f is differentiable.

2. Proved that C^1 implies differentiability.

There are some properties for differentiability: Suppose $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ are differentiable at an interior point a. Then the following hold.

1. $f+g$ is diff at a with derivative $\nabla f(a)+\nabla g(a)$: $\frac{\|(f+g)(a+h)-(f+g)(a)-(Df+Dg)_a(h)\|}{\|h\|} \leq$ $\frac{\|f(a+h)-f(a)-Df_a(h)\|}{\|h\|} + \frac{\|g(a+h)-g(a)-Dg_a(h)\|}{\|h\|} \to 0$ as $h \to 0$. $\square$

2. fg is diff at a with derivative $g(a)\nabla f(a) + f(a)\nabla g(a)$: Let $\Delta_1 = f(a+h) - f(a) - Df_a(h)$ and $\Delta_2 = g(a+h) - g(a) - Dg_a(h)$. Note that $\frac{\Delta_i}{\|h\|} \to 0$ as $h \to 0$.

$$\frac{|(fg)(a+h)-(fg)(a)-(g(a)Df_a(h)+f(a)Dg_a(h))|}{\|h\|} =$$
$$\frac{|(\Delta_1+f(a)+Df_a(h))(\Delta_2+g(a)+Dg_a(h))-(fg)(a)-(g(a)Df_a(h)+f(a)Dg_a(h))|}{\|h\|}$$
$$= \frac{|\Delta_1\Delta_2 + Df_a(h)Dg_a(h)|}{\|h\|} \leq \frac{|\Delta_1|}{\|h\|}\frac{|\Delta_2|}{\|h\|}\|h\| + \|\nabla f_a\|\|\nabla g_a\|\|h\| \to 0 \quad (1)$$

and hence we are done by the squeeze rule. $\square$

3. If $g(a) \neq 0$, then $\frac{f}{g}$ is diff at a with derivative $\frac{g(a)\nabla f(a) - f(a)\nabla g(a)}{g^2(a)}$: WLog $f = 1$ (why?). Now for δ small enough, we see that for all $\|h\| < \delta$, $g(a+h) \neq 0$.

$$\frac{|\frac{1}{g(a+h)} - \frac{1}{g(a)} + \frac{Dg_a(h)}{g^2(a)}|}{\|h\|} = \frac{|\frac{g(a)-g(a)-Dg_a(h)-\Delta_2}{g(a+h)g(a)} + \frac{Dg_a(h)}{g^2(a)}|}{\|h\|}$$
$$\leq \frac{|\Delta_2|}{\|h\|}\frac{1}{|g(a+h)g(a)|} + \|\nabla g(a)\|\frac{1}{|g(a)|}\left|\frac{1}{g(a+h)} - \frac{1}{g(a)}\right| \to 0. \quad (2)$$

$\square$

4. If f is constant, $\nabla f = 0$ (trivial). Conversely, if U is a connected open set and f is differentiable on all of U with $\nabla f = 0$ identically, then f is a constant.
   Indeed, fix $a \in U$ and let $S = \{x \in U \mid f(x) = f(a)\}$. By continuity of f, S is closed in U. It is clearly not empty. If S is proven to be open, then $S = U$ by connectedness (why?).
   S is also open: To this end, consider an open ball B around $b \in S$ that is wholly contained in U. We shall prove that $B \subset S$. Indeed, let $x \in B$. Then $b + t(x-b) \in B$ for all $t \in [0,1]$. The function $g(t) = f(b + t(x-b))$ is continuous on $[0,1]$ (because it is a composition of continuous functions). It is differentiable on $(0,1)$ and $g' = 0$: $\lim_{h \to 0} \frac{g(t+h) - g(t)}{h} = \lim_{h \to 0} \frac{f(b + t(x-b) + h(x-b)) - f(b + t(x-b))}{h} = \nabla_{x-b} f(b + t(x-b)) = \langle \nabla f(b + t(x-b)), x - b \rangle = 0$. By Lagrange's MVT, $g(1) = g(0)$ and hence $x \in S$. □

5. If $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, then F is differentiable and $DF(h) = F(h)$: $\frac{F(a+h) - F(a) - F(h)}{\|h\|} = 0 \; \forall \; h$. □

Here is a concrete example where using the definition is as efficient as blindly calculating. Prove that $F : Mat_{n \times n} \to Mat_{n \times n}$ given by $F(A) = A^2$ is differentiable everywhere and calculate its derivative: The components of F are polynomials (quadratic ones) in the entries of A and are hence C^1 and therefore differentiable. Thus, so is F. We can calculate the derivative in two ways:

1. Blindly: $(DF_A(H))_{ij} = \sum_{k,l} \frac{\partial F_{ij}}{\partial A_{kl}} H_{kl} = \sum_{k,l,m} \frac{\partial A_{im} A_{mj}}{\partial A_{kl}} H_{kl} = \sum \delta_{ik} \delta_{ml} A_{mj} H_{kl} + \delta_{mk} \delta_{jl} A_{im} H_{kl} = \sum_{k,l} A_{lj} H_{il} + A_{ik} H_{kj} = \{H, A\}_{ij}$.

2. Using the definition: $F(A+H) - F(A) = (A+H)^2 - A^2 = \{A, H\} + H^2$ and hence $\frac{\|F(A+H) - F(A) - \{A, H\}\|_{Frob}}{\|H\|_{Frob}} \leq \|H\|_{Frob} \to 0$.

Here is another interesting example. Prove that $\det : Mat_{n \times n} \to \mathbb{R}$ is differentiable everywhere and that $\nabla_H \det(I) = tr(H)$: The determinant is a polynomial in the entries of the matrix and hence C^1 (and differentiable). This can be proven using properties of determinants. (Optional: A fun way to do this is to notice that $\det(I + tH) = (1 + t\lambda_1)(1 + t\lambda_2) \ldots$ where λ_i are (possibly complex) eigenvalues of H.)