MA 200 - Lecture 15

1 Recap

1. Definition of manifolds and examples/non-examples.
2. Taylor's theorem and the second derivative test in one variable (correction in the proof of Taylor: Use Cauchy's MVT to $g(t)$ and t^k).

2 Taylor's theorem and the second derivative test

To state Taylor's theorem in multivariable calculus, we need some notation. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a "multi-index". We typically denote as follows: $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$, $\alpha! = \alpha_1!\alpha_2!\ldots$, $h^\alpha = h_1^{\alpha_1} h_2^{\alpha_2}\ldots$, and $D^\alpha f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \ldots f$ (note that the order does not matter thanks to Clairaut if f is $C^{|\alpha|}$).

Theorem 1. *Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be a C^k function on U. Let $a \in U$ and $|h| < \epsilon$ such that $B_a(\epsilon) \subset U$. Then the polynomial $p_{a,k}(h) = f(a) + \sum_i \frac{\partial f}{\partial x_i}(a)h + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + \ldots + \sum_{\|\alpha\|=k} \frac{D^\alpha f(a)}{\alpha!} h^\alpha$ is the unique polynomial of degree $\leq k$ (degree meaning the maximum sum of powers) such that $\lim_{h\to 0} \frac{f(a+h) - p_{k,a}(h)}{|h|^k} = 0$. Moreover, if f is C^{k+1}, then $f(a+h) = p_{k,a}(h) + \sum_{|\alpha|=k+1} \frac{D^\alpha f(\eta) h^\alpha}{\alpha!}$, where η lies in $B_a(h)$.*

Proof. Uniqueness will be left as a HW problem.

Let $h \neq 0$ (if it is equal to 0, we are done). Consider the one-variable function $q(t) = f(a + t\frac{h}{\|h\|})$ on $|t| < \epsilon$. This function is C^k (because it is a composition of C^k functions). Thus we can apply the one-variable Taylor theorem to it to conclude that $q(\|h\|) = q(0) + q'(0)\|h\| + \ldots$.

Now we claim inductively that $\frac{q^{(m)}(t)\|h\|^m}{m!} = \sum_{|\alpha|=m} \frac{D^\alpha f(a+t\frac{h}{\|h\|})h^\alpha}{\alpha!}$:
Indeed, for $m = 1$ we are done by the Chain rule. Assume the truth of this statement for $1, 2 \ldots, m-1$. We apply the induction hypothesis to $q^{(m-1)}(t)$ to conclude that $\frac{q^{(m)}(t)\|h\|^m}{m!} = \frac{\|h\|}{m} \sum_{|\alpha|=m-1} \frac{d}{dt} \frac{D^\alpha f(a+t\frac{h}{\|h\|})h^\alpha}{\alpha!} = \sum_{|\alpha|=m-1} \sum_i \frac{\partial_{x_i} D^\alpha f(a+t\frac{h}{\|h\|}) h_i h^\alpha}{\alpha! m} = \sum_i \sum_{|\alpha|=m-1} \frac{\partial_{x_i} D^\alpha f(a+t\frac{h}{\|h\|}) h_i h^\alpha}{\alpha! m}$. We want to compare the last expression to $\sum_{|\beta|=m} \frac{D^\beta f(a+t\frac{h}{\|h\|})h^\beta}{\beta!}$.
Given i such that $\alpha_i \geq 1$, every multi-index vector β can be written uniquely as $\beta = \alpha_i + e_i$ where $|\alpha_i| = m-1$. However, this can be done for each such i. Hence, if we fix β, then $\frac{1}{\alpha! m} = \frac{\alpha_i + 1 = \beta_i}{\beta! m}$ and if we sum over all i giving rise to the same β, then we get $\sum_i \frac{\beta_i}{\beta! m} = \frac{1}{\beta!}$. Hence these two expressions are the same, and we are done.

Now the one-variable Taylor theorem completes the proof (why?) □

Now we want to state the second derivative test in multivariable calculus. To this end, we first make a definition: The Hessian of a C^2 function f is the symmetric (by Clairaut) matrix $H(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$. Now we notice a strange phenomenon that does not occur in 1-variable: Let $f(x, y) = x^2 - y^2$. Note that $\nabla f = (2x, -2y) = (0, 0)$ at the origin. The Hessian is $H(0) = diag(2, -2)$. In other words, the Hessian is invertible at $(0, 0)$ and yet the origin is neither a local maximum nor a local minimum (why?) Such points (that is, points where $\nabla f = 0$ but it is neither a local max nor a local min) are called Saddle points.

Theorem 2. *Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be a C^k function on U. Let $a \in U$ and $\nabla f(a) = 0$. If*

1. *H(a) is positive-definite, then a is a local minimum.*
2. *H(a) is negative-definite, then a is a local maximum.*
3. *H(a) is invertible but neither positive- nor negative-definite, then it is a saddle point.*

Conversely, if a is a local minimum, $H(a)$ is positive-semidefinite (and likewise for local minima).

Proof. By Taylor, $f(a+h) = f(a) + \langle \nabla f(a), h \rangle + \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\theta) \frac{h_i h_j}{2} = f(a) + \frac{h^T H(\theta) h}{2}$. Now we need a linear-algebraic lemma:

Lemma 2.1. *A real symmetric matrix is positive-definite if and only if all of its eigenvalues are positive. (It is positive-semidefinite iff all of its eigenvalues are non-negative.) Moreover, if $H : U \subset \mathbb{R}^n \to Sym_{n \times n}(\mathbb{R})$ is a continuous function and if $H(a)$ is positive-definite, then $H(\theta)$ is positive-definite for all $\theta \in B_a(\epsilon)$ for some $\epsilon > 0$.*

Proof. The next time. □

Since f is C^2, $x \to H(x)$ is continuous. By the above lemma, if $H(a)$ is positive-definite, then $f(a + h) - f(a) > 0$ whenever $h \neq 0$. Hence a is a local minimum. If $H(a)$ is negative-definite, apply the result to $-f$. If $H(a)$ is invertible but neither positive- nor negative-definite, then since the determinant of $H(a)$ is the product of its eigenvalues, there is at least one positive eigenvalue and one negative eigenvalue. Therefore, $f(a + h) > f(a)$ for some h and $f(a + h) < f(a)$ for some other h. Thus it is neither a local max nor a local min and hence a saddle point.
If a is a local minimum, and $H(a)$ is not positive-semidefinite, i.e., there exists a v so that $v^T H(a) v < 0$, then $f(a + tv) < f(a)$ for small enough t (Indeed, $f(a + tv) = f(a) + \frac{1}{2} t^2 v^T H(\theta) v$ and by continuity of H, for small enough t, we see that $v^T H(\theta) v < 0$ if $v^T H(a) v < 0$) and hence we have a contradiction. □