MA 200 - Lecture 12

1 Recap

1. Further examples of manifolds.
2. Taylor's theorem and the second derivative test for one-variable calculus.
3. Statement of Taylor for multivariable. Reached a point in the proof where combinatorics plays a role.

2 Taylor's theorem and the second derivative test

Proof. 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 $\beta_i \geq 1$, every multi-index vector β can be written uniquely as $\beta = \alpha + e_i$ where $|\alpha| = m-1$. 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 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 $\nabla f(a) = 0$. 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 $\epsilon > 0$.*

Lemma 2.2. *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. If A is symmetric, then by the spectral theorem, there exists an orthogonal matrix O such that $O^T A O = D$ where D is the diagonal matrix consisting of eigenvalues of A. As a consequence, if A is positive-definite (semidefinite), then $v^T A v > 0$ ($v^T A v \geq 0$) for all $v \neq 0$. Hence $(Ov)^T A (Ov) > 0$ (non-negative) for all $v \neq 0$. Thus, $v^T D v > 0$ (non-negative) and thus the eigenvalues are positive. Tracing the argument backwards (indeed, O is invertible), we see that A is positive-definite if its eigenvalues are positive. Now assume $H(a)$ is positive-definite. Suppose no ϵ works, i.e., for every n there is a $\theta_n \in B_a(1/n)$ such that $v_n^T H(\theta_n) v_n \leq 0$ (and $\|v_n\| = 1$). Then since the unit sphere is compact, there exists a convergent subsequence v_{n_k} converging to v on the unit sphere. Moreover, $\theta_{n_k} \to a$ (why?) By continuity of the function H, $v^T H(a) v \leq 0$ but $v \neq 0$. Therefore we have a contradiction. □

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. □

When does the Taylor series converge? For instance, $1 + x + x^2 + \ldots$ is the formal Taylor series of $\frac{1}{1-x}$ around $x = 0$. It certainly does not converge if $x > 1$ for instance. When $-1 < x < 1$, it does converge to $\frac{1}{1-x}$ (why?)
Here is a supremely strange example:
Consider $f(x) = e^{-1/x^2}$ when $x > 0$ and $f(x) = 0$ when $x < 0$. This function is

C^∞ everywhere: Indeed, it is so away from $x = 0$ (compositions of smooth functions are smooth). The claim is that it is smooth at the origin too (with all derivatives equal to 0). Indeed, we claim that there exists a polynomial $p_k(t)$ of degree $3k$ such that $f^{(k)}(x) = p_k(1/x)e^{-1/x^2}$ when $x > 0$ and 0 when $x \leq 0$: For $k = 0$ this is true. Assume truth for $0, 1, 2, \ldots, k-1$. Then $f^{(k-1)}(x) = p_{k-1}(1/x)e^{-1/x^2}$ when $x > 0$ and 0 when $x \leq 0$. So $f^{(k)}(0) = \lim_{h \to 0} \frac{p_{k-1}(1/h)e^{-1/h^2}}{h} = 0$ (by the squeeze rule and the fact that $g(x)e^{-x}$ goes to 0 when g is a polynomial and $x \to \infty$). When $x > 0$, $f^{(k)}(x) = \left(\frac{-1}{x^2}p'_{k-1}(1/x) + \frac{2}{x^3}p_{k-1}(x)\right)e^{-1/x^2}$. $\square$

This means that the Taylor series converges (it is identically zero!) but NOT to the original function! (By the way, a variant of this function plays a role in physics: Look up the KT-phase transition).

This function leads to a very interesting phenomenon: By reflecting, i.e., $g(x) = f(-x)$, and translating, i.e., $h(x) = g(x-a)$ (where $a > 0$), and multiplying, i.e., $k(x) = h(x)f(x)$, we get a smooth function that is *identically* zero outside a compact set! This leads to a definition:

Def: Let $f : X \to \mathbb{R}^n$ be a continuous function and X be a metric space. Then the closure of the set $\{x \in X | f(x) \neq 0\}$ is called the *support* of f. (It is a closed set by definition.) f is said to have compact support if its support is compact. (Note that if X is itself compact, every continuous function has compact support.)

In other words, we have found a smooth function on $\mathbb{R}$ having compact support! (It is easy to come up with continuous compactly supported functions.) Not just that, we can have some more fun: Note that by further translation (and scaling if necessary), we can easily make sure that the support is *any* compact interval of our choice!

Now, upon integration, i.e., $l(x) = \int_{-\infty}^{x} k(t)dt$, we obtain a smooth function that is a constant on $x \leq 0$ and $x \geq a$. The same tricks as before allow us to produce a function that has compact support on a compact interval of our choice and is identically 1 on a sub-interval of our choice!

3 Integration in more than one-variable

Let $Q \subset \mathbb{R}^n$ be a closed rectangle $[a_1, b_1] \times [a_2, b_2] \ldots [a_n, b_n]$. The volume of Q is defined to be $v(Q) = (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n)$. The "width" is the maximum of $b_i - a_i$ and the intervals $[a_i, b_i]$ are called the component intervals of Q.

Let $f : Q \to \mathbb{R}$ be a bounded function. We want to define $\int_Q f dV$. To this end, roughly speaking, we split Q into sub-rectangles where f is roughly a constant, and add up the resulting numbers. We are led to some definitions:

Def: Given $[a, b] \subset \mathbb{R}$, a partition P is a set of points $a = t_0 < t_1 < \ldots < t_k = b$. The sub-intervals of the partition are $[t_i, t_{i+1}]$. Given a rectangle Q, a partition of Q is the set $P_1 \times P_2 \ldots P_n$ where P_i are partitions of $[a_i, b_i]$. The Cartesian products of the sub-intervals yield several subrectangles of the partition. The maximum width of all these subrectangles is called the mesh of the partition (so the smaller the mesh, the more the number of sub-rectangles we are dividing into). A partition P' is said to be finer than a partition P (or P' is said to be a refinement of P) if $P'_i \subset P_i$ for every i. Given any two partitions $P = P_1 \times P_2 \times \ldots$ and $P' = P'_1 \times \ldots$, the partition $C = (P_1 \cup P'_1) \times (P_2 \cup P'_2) \ldots$ is finer than P and P' and is called their common refinement.

Def: Let P be a partition of Q. For every subrectangle R, let $m_R(f)$ be the infimum of f and $M_R(f)$ be the supremum of f over R. The lower Riemann sum $L(P, f) = \sum_R m_R(f)v(R)$ and the upper Riemann sum $U(P, f) = \sum_R M_R v(R)$.
The key point is

Lemma 3.1. *Let P be a partition of a rectangle Q and $f : Q \to \mathbb{R}$ be a bounded function. If P' is a refinement of P, then $L(f, P) \leq L(f, P')$ and $U(f, P') \leq U(f, P)$.*