1 Logistics

Email: vamsipingali@iisc.ac.in. Course webpage: `http://math.iisc.ac.in/vamsipingali/ma235manifoldsspring2023/ma235.html`. HW : 20% (Roughly once in two weeks. Copying (from each other or the internet) is strictly not allowed.), Midterm - 30%, and Final - 50%. Text book: Introduction to Smooth Manifolds by John Lee.

2 What is this course about and why should you care?

We will review multivariable calculus and study these objects called smooth manifolds. Smooth manifolds are generalisations of objects like circles, spheres, paraboloids, etc. Here are some questions that almost force us to study such objects:

1. Is the parallel postulate of Euclid a consequence of the other axioms? (Riemannian Geometry)
2. Maximise a function subject to some constraints (Optimisation).
3. Model the motion of a robotic arm (Classical Mechanics).
4. Predict the orbit of Mercury to high precision (General Relativity).
5. Land Chandrayaan-3 on the moon (Control theory).
6. How can we classify cubic curves? (Algebraic Geometry).
7. Prove that there are at most finitely many solutions to $a^n + b^n = c^n$ when $n \geq 4$ (a consequence of the Mordell conjecture of Number Theory).

A quick history lesson (copied liberally from Wikipedia):

- Indus-Valley, Babylonian, Egyptian, and Chinese civilisations all knew some geometric figures and a little bit of measurement.
- Euclid in the 3rd century BC axiomatised geometry. Fifth postulate.
- Perspective art during the 14th century Renaissance lead to projective geometry (no focus on distances. Only on intersections).
- Analytic/Coordinate geometry of Fermat and Descartes (17th century) and calculus due to Newton and Leibniz.
- Euler's solution of the seven bridges problem.
- Non-Euclidean geometries and Theorema Egregium (Gauss, 19th century). Riemann and curvature in higher dimensions.
- Poincaré (20th century) and algebraic topology.
- Einstein and General Relativity. Whitney and Whitehead define manifolds.
- 2006: Perelman proved the Poincaré conjecture.

3 Review of multivariable calculus

We denote coordinates in $\mathbb{R}^n$ with *superscripts*, i.e., $x^1, x^2, \ldots, x^n$. An open ball $B_r(a) \in \mathbb{R}^n$ is $|x-a| < r$. A closed ball is denoted as $\bar{B}_r(a)$. An open set $U \subset \mathbb{R}^n$ is one where every $a \in U$ has an open ball $B_a(r) \subset U$, i.e., open balls form a basis. (So do open rectangles.) $\mathbb{R}^n$ is first countable, i.e., it has a countable local basis around every point. It is also second countable, i.e., it has a countable basis. A closed set is one whose complement is open. $S \subset \mathbb{R}^n$ is closed iff it contains all of its limit points. A closed bounded subset of $\mathbb{R}^n$ is compact.

$F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous at a iff given $\epsilon > 0$, there exists $\delta > 0$ such that $|F(x) - F(a)| < \epsilon$ whenever $|x - a| < \delta$. If F, F^{-1} are continuous, F is said to be a homeomorphism. F is continuous at a iff for every sequence $x_n \to a$, $F(x_n) \to F(a)$. As a consequence, $f(x,y) = \frac{xy}{x^2+y^2}$ when $(x,y) \neq (0,0)$ and $f(0,0) = 0$ is discontinuous at $(0,0)$ inspite of being continuous in each variable taken separately. The usual laws of continuity hold. Hence, rational functions are continuous wherever the denominator is not zero.
Invariance of domain: If U is open, and $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ is $1-1$ and continuous, then $F(U)$ is open and U is homeomorphic to it. As a consequence, an open subset of $\mathbb{R}^n$ cannot be homeomorphic to an open subset of $\mathbb{R}^m$ when $m \neq n$. The proof uses the Brouwer fixed point theorem, which in turn uses some algebraic topology (degree theory).

Let $U \subset \mathbb{R}^n$ be open. $F : U \to \mathbb{R}^m$ is said to be differentiable at a if the linear approximation holds, i.e., there exists a linear map $DF_a : \mathbb{R}^n \to \mathbb{R}^m$ such that $\lim_{h \to 0} \frac{F(a+h) - F(a) - DF_a(h)}{|h|} = 0$. If F is differentiable at a, it is *partially* differentiable w.r.t each coordinate and $DF_a = \begin{bmatrix} \frac{\partial F^1}{\partial x^1} & \frac{\partial F^1}{\partial x^2} & \cdots \\ \vdots & \ddots & \vdots \end{bmatrix}$. $F : U \to \mathbb{R}^m$ is said to be *directionally* differentiable along $v \in \mathbb{R}^n$ with directional derivative $L_{v,a}$ if $\lim_{h \to 0} \frac{F(a+hv) - F(a) - L_{v,a}h}{h} = 0$. If F is differentiable, then it is directionally so, with $L_{v,a} = DF_a(v)$, i.e., $L^i_{v,a} = \sum_j \frac{\partial F^i}{\partial x^j}(a) v^j$.

Differentiability implies continuity. Unfortunately, even being directionally differentiable along all directions is not good enough! Let $f(x,y) = \frac{xy^2}{x^2+y^4}$ if $x \neq 0$ and $f(0,y) = 0$. It turns out (HW) that f is directionally differentiable along all directions at $(0,0)$ but f is NOT continuous at $(0,0)$ and hence NOT differentiable there! $f(x,y) = ||x| - |y|| - |x| - |y|$. The partials exist and f is continuous at $(0,0)$ but it isn't differentiable (HW). If the partials exist in a neighbourhood of a AND they are continuous THEN F is differentiable at a. Such functions are called C^1. If F is C^1 and F^{-1} is C^1, then F is said to be a C^1 diffeomorphism. Thus, rational functions are differentiable (in fact C^1) away from the zeroes of their denominators.

Tangent plane: If $f : U \to \mathbb{R}$ is differentiable, then we can talk of the tangent plane to the graph of $y = f(x)$ in $\mathbb{R}^{n+1}$: $y = f(a) + [Df_a][x - a]$.
All the usual rules for derivatives hold. Very importantly, here is the chain rule: Let

$f : U \to \mathbb{R}^m$ be differentiable at a and $g : V \subset \mathbb{R}^m \to \mathbb{R}^p$ be differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and $D(g \circ f)_a = Dg_{f(a)} Df_a$.
Corollary: If $f : U \to \mathbb{R}$ is differentiable, and $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}$ is differentiable, then $f \circ \gamma$ is so with $\frac{d}{dt}(f \circ \gamma) = [Df]_{\gamma(t)} \gamma'(t) = \sum_i \frac{\partial f}{\partial x^i} \frac{dx^i}{dt}$.
Corollary: Let $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ be open. Suppose $f : U \to V$ is bijective and differentiable, and so is $f^{-1} : V \to U$, then $m = n$.
Proof: Indeed, $f \circ f^{-1}(x) = x$. Thus, $Df_{f^{-1}(a)} Df_a^{-1} = I$. Likewise, $Df_a^{-1} Df_{f^{-1}(a)} = I$. Thus $m = n$ and Df_a is an isomorphism for all $a \in U$.

Gradient and steepest ascent: Suppose f is differentiable on an open set $U \subset \mathbb{R}^n$. The gradient Df_a is the direction of steepest increase : Indeed, the directional derivative $L_{a,v} f$ along v is $[Df_a][v]$. Now the famous/infamous Cauchy-Schwarz inequality says that $-|\vec{a}||\vec{b}| \leq \vec{a}.\vec{b} \leq |\vec{a}||\vec{b}|$ with equality holding only when $\vec{a}$ and $\vec{b}$ are parallel. Thus, $-|Df_a||v| \leq L_{a,v} f \leq |Df_a||v|$ with equality provided v is parallel/antiparallel to Df_a. For instance, if $f(x, y) = x^2 + 2y^2$, then at $(1, 1)$, $Df_{(1,1)} = (2, 4)$. Indeed, moving more in the y-direction would cause a greater increase in the height.

Local extrema - the first derivative test: Let $U \subset \mathbb{R}^n$ be open and $f :\to \mathbb{R}$ be a function. $a \in U$ is said to be a local minimum if $f(x) \geq f(a) \,\forall\, x \in B$ where $a \in B \subset U$ is some open ball centred at a. Likewise for a local maximum. If a is a local extremum and f is differentiable at a, then $Df_a = (0, 0, 0, \ldots)$: Indeed, fix v and consider $g_v(t) = f(a + tv)$. It attains a local minimum at $t = 0$ and is differentiable there. By one-variable calculus, $g_v'(0) = Df_a v = 0$. Since this holds for all v, $Df_a = [0\ 0\ 0 \ldots]$.

Finding global extrema: Recall that since closed and bounded sets are compact, a continuous function on them achieves global maxima and minima. How does one find them? Here is a simple example where one can find global extrema : Consider $f(x, y) = 2x^3 - 3y^2$ on the disc $x^2 + y^2 \leq 4$. First we find local extrema : $Df = [6x^2\ \ -6y] = [0\ 0]$ when $x = y = 0$ which is inside the unit disc. The value of f is $f(0, 0) = 0$. But global extrema can potentially occur on the boundary $x^2 + y^2 = 4$ (and there the gradient is not necessarily zero). On the boundary, the function is $g(x) = 2x^3 - 3(4 - x^2) = 2x^3 + 3x^2 - 12$ and $-2 \leq x \leq 2$. Now we solve a one-variable global extrema problem. Now $g'(x) = 6x^2 + 6x = 0$ precisely when $x = 0$ or $x = -1$ both of which are inside the interval. $g(0) = -12, g(-1) = -11$. Moreover, $g(-2) = -16, g(2) = 16$. Thus the maximum occurs at $x = 2, y = 0$ and the minimum at $x = -2, y = 0$.

Higher derivatives: It is of course possible that a function is differentiable and the derivative is continuous, and all the second partials exist but it is not twice-differentiable. Here is an even weirder example : $f(x, y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ for $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. By properties, f is differentiable on $\mathbb{R}^2 - \{(0, 0)\}$. At $(0, 0)$ we claim that it is differentiable and the gradient is 0 : $|f(x, y) - 0 - 0| = |xy| \frac{|x^2-y^2|}{x^2+y^2} \leq x^2 + y^2$ (why ?) Therefore, if $\sqrt{x^2 + y^2} < \epsilon$, we are done (Why ?) Now $f_x = \frac{y(x^2-y^2)(x^2+y^2)+2x^2y(x^2+y^2)-2x^2y(x^2-y^2)}{(x^2+y^2)^2} = y\frac{x^4-y^4-4x^2y^2}{(x^2+y^2)^2}$, $f_y = -x\frac{y^4-x^4-4x^2y^2}{(x^2+y^2)^2}$ away from $(0, 0)$. Thus f_x, f_y are continuous throughout. However, $f_{xy} = \lim_{h\to 0} \frac{f_y(h,0)-f_y(0,0)}{h} = 1$ and

$f_{yx} = \lim_{h\to 0} \frac{f_x(0,h) - f_x(0,0)}{h} = -1$. Therefore, $f_{xy} \neq f_{yx}$ in general !
Theorem (Clairaut) : If all the second partials exist and are continuous are a, then the mixed partials are equal.

A function is called C^2 if DF is differentiable and D^2F is continuous (and likewise for C^k). Notation: If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is a multi-index, often $\alpha! := \alpha_1!\alpha_2!\ldots$, and $|\alpha| = \alpha_1 + \alpha_2 + \ldots$. $\partial_\alpha f = \frac{\partial^{|\alpha|} f}{\partial (x^1)^{\alpha_1} \partial (x^2)^{\alpha_2} \ldots}$. If a function is C^k for all k, it is said to be C^∞ or smooth.
Proof of Clairaut: We will prove it for two-variable functions. The general case follows from this special case (why?) Let $u(h,k) = f(a+h, b+k) - f(a+h, b)$, $v(h,k) = f(a+h, b+k) - f(a, b+k)$, and $w(h,k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) - f(a,b)$. Now, $w(h,k) = u(h,k) - u(0,k) = h\partial_x u(c_1, k)$ by MVT for $c_1 \in [0,h]$. This equals $h(\partial_x f(a+c_1, b+k) - \partial_x f(a+c_1, b)) = hk\partial_y \partial_x f(a+c_1, b+c_2)$ where $c_2 \in [0,k]$. Likewise, $w(h,k) = v(h,k) - v(h,0) = hk\partial_x\partial_y f(a+d_1, b+d_2)$ where $d_1 \in [0,h], d_2 \in [0,k]$. Dividing by hk, we see that $\partial_y\partial_x f(a+c_1, b+c_2) = \partial_x \partial_y f(a+d_1, b+d_2)$. By assumption of continuity of the second partials, there exists a $\delta > 0$ such that $|\partial_y\partial_x f(a+c_1, b+c_2) - \partial_y\partial_x f(a,b)| < \frac{\epsilon}{2}$, and $|\partial_x\partial_y f(a+d_1, b+d_2) - \partial_x\partial_y f(a,b)| < \frac{\epsilon}{2}$ whenever $|(h,k)| < \delta$ (why ?). Thus, $|f_{xy}(a,b) - f_{yx}(a,b)| < \epsilon$ for all $\epsilon > 0$. We are done. $\square$
Using the C^2 Clairaut, we can prove the C^k Clairaut for any k.

Taylor's theorem (Peano form of remainder): Let $f : U \to \mathbb{R}$ be C^k at $a \in U$. Then there exists functions h_α on a neighbourhood of a such that $f(x) = \sum_{|\alpha| \leq k} \frac{\partial_\alpha f(a)}{\alpha!}(x^1 - a^1)^{\alpha_1}(x^2 - a^2)^{\alpha_2} \ldots + o(\|x-a\|^k)$.
If the Taylor series converges AND the function equals its Taylor series, it is said to be real-analytic.

Proof of a version of Taylor's theorem:

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 is easy and left as an exercise. 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 $\sum_{|\beta|=m} \frac{D^\beta f(a+t\frac{h}{\|h\|})h^\beta}{\beta!}$.
We apply the induction hypothesis to $q^{(m-1)}(t)$ to conclude that $\frac{q^{(m)}(t)\|h\|^m}{m!} = \frac{\|h\|}{m}\sum_{|\alpha|=m-1} \frac{d}{dt} \frac{D^\alpha f(a+t\frac{h}{\|h\|})h}{\alpha!}$
$\sum_{|\alpha|=m-1}\sum_i \frac{\partial_{x_i} D^\alpha f(a+t\frac{h}{\|h\|})h_i h^\alpha}{\alpha! m} = \sum_i \sum_{|\alpha|=m-1} \frac{\partial_{x_i} D^\alpha f(a+t\frac{h}{\|h\|})h_i h^\alpha}{\alpha! m}$. 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, i.e., the following theorem, completes the proof.

Theorem 2. *Let $U \subset \mathbb{R}$ be an open set and $f : U \to \mathbb{R}$ a C^k function on U. Let $a \in U$ and $|h| < \epsilon$ such that $(a-\epsilon, a+\epsilon) \in U$. Then the polynomial $p_{k,a}(h) = f(a) + f'(a)h + \frac{f''(a)h^2}{2!} + \ldots + \frac{f^{(k)}(a)h^k}{k!}$ is the unique polynomial of degree $\leq k$ 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 $f(a+h) = p_{k,a}(h) + \frac{f^{(k+1)}(\eta)h^{k+1}}{(k+1)!}$, where η lies between a and $a+h$.*

□

The second derivative test: Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ be a C^1 function. Suppose f is C^2 at $a \in U$ and $Df_a = 0$.
Theorem: If $\sum_{i,j} f_{ij}(a)v^i v^j > 0$ for all $v \neq 0$, then a is a local minimum. If $\sum_{i,j} f_{ij}(a)v^i v^j < 0$ for all $v \neq 0$, then a is a local maximum. Conversely, if a is a local minimum $\sum_{i,j} f_{ij}(a)v^i v^j \geq 0$ for all v (and likewise for a local max).
Proof: Taylor's theorem (with the Peano form of the remainder) implies that if $\sum_{i,j} f_{ij}(a)v^i v^j > 0$, $f(a+v) - f(a) > 0$ for all small v. Thus a is a local minimum. Likewise for local maxima. Conversely, if $\sum_{i,j} f_{ij}(a)v^i v^j < 0$ for some v, then $f(a+v) - f(a) < 0$ and hence a cannot be a local minimum. (Likewise for local maxima.) □

4 Bump functions

Unfortunately, even if the Taylor series converges, it need NOT be equal to the function itself! Let $E(t) = e^{-1/t}$ when $t > 0$ and 0 when $t \leq 0$. It turns out that $E(t)$ is C^{∞} everywhere (an exercise/Lee's book).
Theorem: There exists a smooth function $\chi : \mathbb{R}^n \to \mathbb{R}$ that satisfies

1. $0 \leq \chi \leq 1$.
2. $\chi = 1$ on $[-1,1] \times [-1,1] \ldots [-1,1]$.
3. $supp(\chi)$ is contained in $[-2,2] \times [-2,2] \ldots$.

We can find a function that satisfies similar properties but is instead radially symmetric. Such functions are called bump functions.
Proof: We first construct χ for $n = 1$. Let $\zeta(t) = \frac{E(t)}{E(t)+E(1-t)}$. Now $\zeta(t) = 0$ for $t \leq 0$ and $\zeta(t) = 1$ for $t \geq 1$. Let $\eta(t) = \zeta(2+t)\zeta(2-t)$. This does the job when $n = 1$. If we want a spherically symmetric bump function in $\mathbb{R}^m$, simply define it as

$\chi(x) = \eta(r)$. If we want a cylindrically symmetric bump function in $\mathbb{R}^m$, simply define it as $\chi(x) = \eta(x^1)\eta(x^2)\eta(x^3)\ldots$. $\square$