MA 200 - Lecture 1

1 Logistics and about the course

Quizzes based on HW (20), a midterm (30), and a final (50) will form the evaluation methods for this course. This course will cover multivariable calculus. As mentioned in UM 102, life depends on multivariable calulus (the solution of Schrödinger equation for big molecules tells us about properties of materials for instance, building an aircraft depends on solving partial differential equations, optimising revenues of companies involves several variables, studying the molecules of air in a room depends on many variables, etc). Even in pure mathematics, there are theorems that can be stated purely using polynomials and algebra (algebraic geometry), that need multivariable analysis methods for their proofs. Analytic number theory needs complex analysis, which in turn needs multivariable calculus. To study the geometry of curved objects, we need multivariable calculus. So we need to generalise all of our usual calculus tools - functions, limits, continuity, derivatives, maxima and minima, Taylor's theorem, integrals, and even the fundamental theorem of calculus to more several variables.
Essentially, we will first go over differential calculus (some of which has been done in UM 102 but it will be done in greater detail), and then integral calculus (this part was completely glossed over in UM 102). To state the fundamental theorem of calculus in higher dimensions (in UM 102, the Stokes, Green, and Divergence theorems were stated with very little rigour), one would have to use the language of manifolds. If time permits, we can do more interesting things (like an analysis of gradient descent, abstract manifolds, de Rham cohomology, etc).

2 Review of (real) linear algebra

Linear functions are the simplest examples of functions. So linear algebra is like multivariable calculus-lite.
Recall that a real vector space V is a set where you can add elements and multiply by real numbers. These operations satisfy the usual properties (commutativity, distributivity, axioms about zero, etc). A fundamental example of a vector space is $\mathbb{R}^n$ and we shall deal almost exclusively with it. A finite-dimensional vector space is one where there exists a finite set of vectors such that every element is a linear combination of these finitely many vectors. A finite-dimensional vector space has a basis: A set of linearly independent vectors spanning the space. Every such set has the same size called the dimension of V. In fact, any linearly independent set of $dim(V)$ vectors is a basis. A subspace is a subset of the vector space that is closed under addition and scalar

multiplication. Any basis for a subspace can be extended to a basis of the entire vector space.

A linear map $T : V \to W$ is a function such that $T(av + bw) = aT(v) + bT(w)$. Examples are dilations, and rotations. Given bases e_i and f_j of V, W, we get a matrix $[T]$: $Te_i = \sum_j [T]_{ji} f_j$. If v_i are the components of a vector $v = \sum_i v_i e_i$ in V, then $w = T(v) = \sum_j w_j f_j$ has components $w_j = \sum_i [T]_{ji} v_i$, i.e., $\vec{w} = [T]\vec{v}$. Crucially, composition of linear maps translates into multiplying the corresponding matrices. Matrix multiplication satisfies the usual properties. Also, $(AB)^T = B^T A^T$.

Linear maps take squares to parallelograms. Translations are *not* linear maps. Linear maps that are also bijections are called linear isomorphisms. Secretly, every f.d. vector space is linearly isomorphic to $\mathbb{R}^n$ (using a basis).

Recall that the row rank of a matrix A is the number of linearly independent rows. Likewise, the column rank is the number of linearly independent columns. One can prove that the row rank=column rank and hence the notion of rank of a matrix is well-defined. It can be easily calculated by bringing to the row-reduced echelon form (by the Gauss-Jordan elimination algorithm). A square matrix is invertible iff it has full rank iff the RREF is Identity. Invertibility can also be tested by calculating the "volume of" (will make this precise later) an n-dimensional parallelopiped formed out of the columns and checking if it is zero or not. This motivates the definition of the determinant $\det(A)$ as an alternating multilinear normalised function of the columns. It turns out that $\det(A^T) = \det(A)$ and $\det(AB) = \det(A)\det(B)$. Moreover, $\det$ of an upper-triangular matrix is the product of the diagonal entries. In general, one has a recursive expansion of a determinant along any row or column. Lastly, a matrix is invertible iff its determinant is non-zero.

Inner products: A function $\langle , \rangle : V \times V \to \mathbb{R}$ is an inner product if it is symmetric, bilinear, and satisfies $\langle v, v \rangle \geq 0$ with equality iff $v = 0$. The usual inner product on $\mathbb{R}^n$ is $\langle x, y \rangle = \sum_i x_i y_i$. Every f.d vector space with an inner product has an orthonormal basis, i.e., a basis such that $\langle e_i, e_j \rangle = \delta_{ij}$. In this basis, the inner product (in terms of components) is precisely the usual one in $\mathbb{R}^n$. Define $\|x\|^2 = \langle x, x \rangle$. Crucially, we have the Cauchy-Schwarz inequality: $|\langle x, y \rangle| \leq \|x\|\|y\|$. Using this inequality we can show the triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$. (We also see that $\|x\| \leq \|x - y\| + \|y\|$. Reversing x, y we see the reverse triangle inequality: $\|x - y\| \geq |\|x\| - \|y\||$.)

Norms: A function $\|.\| : V \to \mathbb{R}_{\geq 0}$ is called a norm (a distance function/a metric if you will) if $\|v\| = 0$ iff $v = 0$, $\|cv\| = |c|\|v\|$ and $\|x + y\| \leq \|x\| + \|y\|$. An inner product defines a norm $\|.\|_{l^2}$ (but not all norms arise out of inner products). The function $\|v\|_{l^1} = \sum_i |v_i|$ is a norm as is $\|v\|_{l^\infty} = \max_i |v|_i$. Note that $\|v\|_{l^\infty} \leq \|v\|_{l^1} \leq \sqrt{n}\sqrt{\langle v, v \rangle} \leq n\|v\|_{l^\infty}$. Two norms are said to be equivalent if $c\|v\|_2 \leq \|v\|_1 \leq C\|v\|_2$ for some constants c, C (independent of v). As an exercise you will show that any two norms on a f.d. vector space are equivalent. Here are two (of course equivalent) norms on the space of matrices:

1. The operator norm: $\|A\|_{op} := \sup_{\|x\|_{l^2}=1} \|Ax\|_{l^2}$. (Why is this a norm?) Note that if A is diagonalisable with *orthonormal* eigenvectors e_i forming a basis, then $\|Ax\|_{l^2} = \|\sum_i \lambda_i x_i e_i\| = \sqrt{\sum_i |\lambda_i|^2 x_i^2} \leq \max_i |\lambda_i|$. In this case, the operator norm is $\max_i |\lambda_i|$. In general, it may not be the case even if A is diagonalisable! For instance, take $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

2. The Frobenius/Hilbert-Schmidt norm: $\|A\|_{HS}^2 := \sum_{i,j} ||a_{ij}|^2 = tr(A^T A)$. This norm is the usual inner product norm pretending that the space of matrices is $\mathbb{R}^{mn}$.

Both of these matrix norms satisfy $\|AB\| \leq \|A\|\|B\|$. As a consequence, if A is invertible, $1 \leq \|A\|\|A^{-1}\|$.