MA2104 (Multivariable Calculus)

Jia Cheng

January 2021

1 Definitions and Formula

2 Vectors, Lines and Planes

2.1 The dot product

Here, I reference MA1101R's textbook, (Linear Algebra, Concepts and Techniques in Euclidean Space). Theorems in geometry and trigonometry(cosine rule) gives rise to the equalities

$$\cos(\theta) = \frac{x_1 y_1 + x_2 y_2}{||\mathbf{x}||||\mathbf{y}||} \qquad \text{in } \mathbb{R}^2$$

$$\cos(\theta) = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3}{||\mathbf{x}||||\mathbf{y}||} \qquad \text{in } \mathbb{R}^3$$

Hence, we define the dot product in $\mathbb{R}^2$ and $\mathbb{R}^3$. Now, we extend this definition to $\mathbb{R}^n$, such that $\mathbf{x} \cdot \mathbf{y} = \sum_{1 \leq i \leq n} x_i y_i$. We also extend the definition of angle to $\mathbb{R}^n$, such that

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||||\mathbf{y}||}$$

By the Cauchy Schwarz inequality, it turns out that such a generalisation is consistent, as $\frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}$ is always bounded between -1 and 1.

Note that the angle θ between vectors is by definition the smaller angle, i.e. $0 \leq \theta \leq \pi$.

In fact, we can completely "ditch" all geometrical sources of "proof" and simply define $\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||||\mathbf{y}||}$ axiomatically. This is the direction that Prof Chin Chee Whye takes.
To make things more efficient by this approach, my personal interpretation that we can simply "assume" that this algebraic definition of angle is equivalent with the geometrical definition when it comes to 2-space and 3-space. By thinking heuristically in this manner, we can then use intuitive geometrical arguments, then fall back to algebraic axioms and definitions to proof our results formally.
If we do take the cosine rule (geometric) approach, then I am curious how length is defined there. (since we are not working from an algebraic standpoint)

2.2 Distance between parallel planes in $\mathbb{R}^3$

Let P_1 and P_2 be 2 parallel planes. Let n be the normal vector of both planes.
Choose any point r_1 on P_1.

Now let r_2 be a point on P_2. Then

$$\begin{aligned}
|(r_2 - r_1)\cos\theta| &= \frac{|(r_2 - r_1)||n||\cos\theta|}{|n|} \\
&= \frac{|(r_2 - r_1)\cdot n|}{|n|} \\
&= \frac{|r_2\cdot n - r_1\cdot n|}{|n|}
\end{aligned}$$

is the shortest distance from r_1 to plane P_2. (Here, θ is the angle between $(r_2 - r_1)$ and n.)
Regardless of which r_1, r_2 we choose, this shortest distance is the same since $r_2 \cdot \frac{n}{|n|}, r_1 \cdot \frac{n}{|n|}$ are constant.
Hence, we can interpret this shortest distance as the distance between the 2 planes.

2.3 Distance between point and a line

We can prove algebraically that the shortest distance between a point with position vector $\mathbf{u}$ and a vector space V is the $\mathbf{u} - proj_V(\mathbf{u})$, where $proj(u)$ is the orthogonal projection of $\mathbf{u}$ on V.

This result means that we can show that this shortest distance is equal to $|mathbfu|\sin\theta$ *without* geometry (i.e. without drawing a diagram and relying on our intuitive/visual understanding of geometry).
Note that angle θ is defined as

$$\cos(\theta) = \frac{\mathbf{x}\cdot\mathbf{y}}{||\mathbf{x}||||\mathbf{y}||}$$

The algebra goes:

$$\begin{aligned}
|\mathbf{u} - proj(\mathbf{u})|^2 &= (\mathbf{u} - proj(\mathbf{u}))\cdot\mathbf{u} - proj(\mathbf{u}) \\
&= |\mathbf{u}|^2 - \frac{(\mathbf{u}\cdot\mathbf{v})^2}{|\mathbf{v}|^2} \\
&= |\mathbf{u}|^2\sin^2\theta
\end{aligned}$$

Note the last equality is due to
$$|\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2 = |\mathbf{u}|^2|\mathbf{v}|^2\sin^2\theta$$
Hence,
$$|\mathbf{u} - proj(\mathbf{u})| = |\mathbf{u}|\sin\theta$$

2.4 Normal of a level set

The gradient is a special case of the general notion of a total derivative. (see later sections in MA2104)

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a function in xyz-space. For any $c \in \mathbb{R}$, the set of points $S = \{(x, y, z) : f(x, y, z) = c\}$ is a level set of f. Then for each point $(x, y, z) \in S$, $(\nabla f)(x, y, z)$ is the normal vector to S at (x, y, z).

Tangent of intersection of two surfaces in $\mathbb{R}^3$ Note that this only works in $\mathbb{R}^3$ precisely because it is 3 dimensional, given 2 linearly independent vectors, the orthogonal complement of their span has to be one-dimensional.

Let E be the intersection of 2 level sets (generated by)/surfaces given by f, g. Let $p = (x, y, z) \in E$. Then from above we know that $\nabla f(p), \nabla g(p)$ give the normal vectors of level sets (generated

by)/surfaces f, g.

This allows us to write

$$\mathbf{n}_f(p) = (\nabla f)(p) \text{ and } \mathbf{n}_g(p) = (\nabla g)(p)$$

Note that $\mathbf{n}_f, \mathbf{n}_g$ may not be unit vectors, but that's easy to rectify.

Assume that $\mathbf{n}_f, \mathbf{n}_g$ are not parallel. Using the cross product , we can then obtain the tangent vector at p

$$\mathbf{T}_E(p) = \mathbf{n}_f(p) \times \mathbf{n}_g(p)$$

3 Functions of Two Variables, Quadric Surfaces, Limits and Continuity

3.1 Point Set Topology

Definitions given in lecture are that of open, closed, interior points, (open) ball, boundary points. Define limit points as in Baby Rudin chapter 2. Define the following:

- Metric space S, Region $R \subseteq S$

- $B(p, r) = \{q \in S : d(p, q) < r\}$

- $bd(R) = \{p \in S : \forall r \in \mathbb{R}^+, B(p, r) \cap R \neq \emptyset \wedge B(p, r) \cap R^c \neq \emptyset\}$ (boundary of region R)

- $int(R) = \{p \in R : \exists r \in \mathbb{R}^+, B(p, r) \subseteq R\}$ (interior of R)

- $lim(R) = \{p \in S : \forall r \in \mathbb{R}^+, \exists q \in S, q \neq p, q \in B(p, r)\}$ (limit points of R)

3.2 Surfaces

A surface is described by a predicate. For e.g. a surface in 3D can be described by the predicate $P(x, y, z) = "x + y + 2z = k"$. The surface itself is then the set $\{(x, y, z) : P(x, y, z)\}$.

Let $f(x, y)$ be a function of 2 variables. To visualise this as a surface, we consider the equation $z = f(x, y)$. The predicate describing the surface is then $P(x, y, z) = "z = f(x, y)"$. Hence, the surface is the set $\{(x, y, z) : P(x, y, z)\} = \{(x, y, z) : z = f(x, y)\}$.

3.3 Quadric Surfaces

3.3.1 Naming

Notice that quadric surfaces have 2 parts to their name: x-ic y-loid

- Elliptic Paraboloid: The horizontal traces are ellipses, vertical traces are parabolas.

- Hyperbolic Paraboloid: The horizontal traces are hyperbolas, vertical traces are parabolas.

- n-sheet Hyperboloid (where $n = \pm 1$): Vertical traces are hyperbolas.

  Note that the "vertex" is only relevant to the elliptic paraboloid. The vertex can be interpreted to be the point of intersection between the elliptic paraboloid and the line of symmetry.

3.4 Limits

3.4.1 Squeeze Theorem

The following are equivalent statements of the Squeeze Theorem. (Either can be used to prove the other.)

MA1102R Style:
Suppose $g(x) \leq f(x) \leq h(x)$ in a neighborhood of a, and $\lim_{x \to a} g(x) = L = \lim_{x \to a} h(x)$. Then $\lim_{x \to a} f(x) = L$.

MA2104 Style:
Suppose $|f(x) - L| \leq g(x)$ in a neighborhood of a, and $\lim_{x \to a} g(x) = 0$. Then $\lim_{x \to a} f(x) = L$.

4 Differentiation of multivariable functions

See lecture notes for definitions.

Unless otherwise stated, function $f : R \to R^n$, where $R \subset R^n$. R is an open region.

4.1 Higher Order Derivatives (Lecture 8)

Def A function f is said to be of class C^0 if f is continuous on region R.
A function f is said to be of class C^r, $r \in \mathbb{Z}^+$, if f' is of class C^{r-1}.
For example, functions of class C^1 are continuously differentiable.

Thm 1 f is of class C^1 $\iff$ $\frac{\partial f_i}{\partial x_j}$ is defined and continuous on R for all $1 \leq i \leq n$, $1 \leq j \leq m$.

f is of class C^r $\iff$ $\frac{\partial f_i}{\partial x_j}$ is of class C^{r-1}.

Proof by induction on r:
Suppose $P(r)$, and let $f : R \to \mathbb{R}^n$ be of class C^{r+1}. By definition, $f' : R \to L(\mathbb{R}^m, \mathbb{R}^n) \cong M_{n \times m}(\mathbb{R}) = \mathbb{R}^{n,m}$ is of class C^r.

Note that in particular, f is of class C^1, hence the components of f' are indeed the partial derivatives of f.

Next, we will show that the partial derivatives of the component functions of f' are, in fact, second order partial derivatives of the component functions of f.

By $P(r)$,

$$\forall (i_1, i_2) \in \{1, \ldots, n\} \times \{1, \ldots, m\}, \forall j \in \{1, \ldots, m\}, \frac{\partial f'_{(i_1, i_2)}}{\partial x_j} \text{is of class } C^{r-1}$$

Note that the vector space of linear maps $L(\mathbb{R}^m, \mathbb{R}^n)$ is isomorphic to n by m matrices.
Hence, if we view $f'_{(i_1, i_2)}(\mathbf{p})$ as a matrix, then we have

$$\forall \mathbf{p} \in R, \; f'_{(i_1, i_2)}(\mathbf{p}) = \frac{\partial f_{i_1}}{\partial x_{i_2}}(\mathbf{p})$$

Hence, in general we have

$$f'_{(i_1, i_2)} = \frac{\partial f_{i_1}}{\partial x_{i_2}}$$

4

Hence,

$$\frac{\partial f'_{(i_1, i_2)}}{\partial x_j} = \frac{\partial^2 f_{i_1}}{\partial x_j \partial x_{i_2}}$$

But saying that $\frac{\partial^2 f_{i_1}}{\partial x_j \partial x_{i_2}}$ is of class C^{r-1} is the same as saying $\frac{\partial f_{i_1}}{\partial x_{i_2}}$ is of class C^r. And we are done since $1 \leq i_1 \leq n, 1 \leq i_2 \leq m$.

4.2   Directional and partial derivatives

Directional derivatives, and in particular, partial derivatives, are really just derivatives of single variable (composite) functions.

For instance, consider the directional derivative of a scalar-valued, multivariable function $f : \mathbb{R}^m \to \mathbb{R}$.

$$D_u f(p) = \lim_{t \to 0} \frac{f(p + tu) - f(p)}{t}$$

Define the function

$$f_u : \mathbb{R} \to \mathbb{R}$$
$$t \mapsto f(p + tu)$$

Usually, the domains of f and f_u are a subset of $\mathbb{R}^m$ and $\mathbb{R}$ respectively, with the image of $p + tu$ falling within the domain of f. However for simplicity, we will ignore this detail.

Then, $D_u f(p) = f'_u(0)$.
For a product of functions f, g, we have

$$(fg)_u(t) = (fg)(p + tu) = f(p + tu)g(p + tu) = f_u(t)g_u(t)$$

Differentiating, we obtain $(fg)'_u(t) = f_u(t)g'_u(t) + f'_u(t)g_u(t)$ by the product rule for single-variable functions. In particular, this holds for $t = 0$.

Writing this for partial derivatives, we have,

$$(fg)'_u(0) = f_u(0)g'_u(0) + f'_u(0)g_u(0) \qquad \text{where } u = e_i$$
$$\frac{\partial fg}{\partial x_i}(p) = f(p)\frac{\partial g}{\partial x_i}(p) + \frac{\partial f}{\partial x_i}(p)g(p)$$

This is used to prove the product and quotient rules for gradients.

4.3   Lemma on constant functions

The following lemma is used implicitly in Lecture 14, when finding the potential function when given a gradient field.

For simplicity we will work with 3 variables x, y, z.
Lemma Given a multivariable scalar-valued function $f : \mathbb{R}^3 \to \mathbb{R}$, suppose $\frac{\partial f}{\partial x}(x, y, z) = 0$ for all $x, y, z \in \mathbb{R}^3$. Then, $f(x, y, z) = g(y, z)$ for some function $g : \mathbb{R}^2 \to \mathbb{R}$.

Proof Fix some $y_0, z_0 \in \mathbb{R}$. Then $\frac{\partial f}{\partial x}(x, y_0, z_0) = 0$. Mean-value theorem from single variable calculus says that the composite function $x \mapsto f(x, y_0, z_0)$ is some constant C (which depends on choice of

y_0, z_0). Hence, we can define $g(y_0, z_0) = C = f(x, y_0, z_0)$ for any x. Clearly, g is well-defined for each $y_0, z_0 \in \mathbb{R}$.
And the proof is complete.

In other words, this lemma allows us to remove "redundant" variables from a function $f(x, y, z)$ to form a function $g(y, z)$ that is clearly independent of x.
While this may seem intuitive, I feel that it is important to write out the details.

5   Properties of the Iverson bracket

Suppose we define a real, scalar-valued function case-wise. Let A, B be disjoint sets in $\mathbb{R}$ and that $5 \notin A, B$.

$$f(x) = \begin{cases} y_1 & x \in A \\ y_2 & x \in B \\ y_3 & x = 5 \end{cases}$$

Then,

$$f(x) = y_1[x \in A] + y_2[x \in B] + y_3[x = 5]$$

Suppose $y_1 = g(x), y_2 = h(x), y_3 = l(x)$ for some functions g, h, l. Then, we have,

$$f(x) = g(x)[x \in A] + h(x)[x \in B] + l(x)[x = 5]$$
$$= g(x)[x \in A] + h(x)[x \in B] + l(5)[x = 5]$$

In particular, notice the following equality.

$$l(x)[x = 5] = l(5)[x = 5]$$

If we are to view the LHS and RHS as functions of x, then consider cases:

- If $x \neq 5$, then $LHS(x) = 0 = RHS(x)$.

- If $x = 5$, then $LHS(5) = l(5)[5 = 5] = l(5) = RHS(5)$.

- Since for all values of x, $LHS(x) = RHS(x)$, we have LHS = RHS.

Note In general, we cannot write $l(5)[x = 5] = l(5)$, since RHS here is identically $l(5)$, whereas LHS is only $l(5)$ when $x = 5$, and equals 0 for all other values of x.