Module 1: Mathematical Preliminaries

This introductory module comprises of four lectures. In these four lectures, we introduce to the readers some basic concepts from multivariable calculus and some essential results from ordinary differential equations(ODEs). Some geometrical concepts necessary for the subsequent modules are also discussed.

Module 1 is organised as follows. In Lecture 1, we review some basic definitions and results from multivariable calculus. In Lecture 2, we discuss some essential formulas for solving linear first-order and second-order (with constant coefficients only) ODEs. In addition, we review the basic existence and uniqueness theorems for initial value problem (IVP) for ODEs and systems of ODEs. In Lecture 3, we discuss some geometrical concepts like surfaces, normals and integral curves and surfaces of vector fields. Finally, Lecture 4 is devoted to methods for finding the integral curves of a vector field by solving systems of ODEs.

Lecture 1 A Review of Multivariable Calulus

In this lecture, we recall some basic concepts from multivariable calculus. The concepts of limits, continuity, partial derivatives, directional derivatives, chain rules, tangent plane and normals are discussed.

For any (x, y), $(x_0, y_0) \in \mathbb{R}^2$, let us denote

$$d((x, y), (x_0, y_0)) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

for the distance between two points (x, y) and (x_0, y_0). A disk $D_r(x_0, y_0)$ of radius r centered at (x_0, y_0) is defined as

$$D_r(x_0, y_0) = \{(x, y) \mid d((x, y), (x_0, y_0)) < r\}.$$

The concept of limit now can be defined by the same ϵ, δ technique as in one variable calculus.

Definition 1. (The ϵ, δ definition of limit) *Let $f(x, y)$ be a real-valued function of two variables defined on a disk $D_r(x_0, y_0)$, except possibly at (x_0, y_0). Then*

$$\boxed{\begin{array}{l} \displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x, y) = l \ \textit{ if for every } \epsilon > 0 \textit{ there is a } \delta > 0 \textit{ such that} \\ \qquad\qquad |f(x, y) - l| < \epsilon \textit{ whenever } 0 < d((x, y), (x_0, y_0)) < \delta. \end{array}}$$

Definition 1 means that the distance between $f(x, y)$ and l can be made arbitrarily small by making the distance from (x, y) to (x_0, y_0) sufficiently small (but not 0). That is, if any small interval $(l - \epsilon, l + \epsilon)$ is given around l, then we can find a disk $D_\delta(x_0, y_0)$ with center (x_0, y_0) and radius $\delta > 0$ such that f maps all the points in $D_\delta(x_0, y_0)$ [except possibly (x_0, y_0)] into the interval $(l - \epsilon, l + \epsilon)$.

The definition of a limit can be extended to functions of three or more variables. Using vector notation the definition can be written in a compact form as follows:

Let $f : D_r(\mathbf{x_0}) \subset \mathbb{R}^n \to \mathbb{R}$. Then

$$\boxed{\begin{array}{l} \displaystyle\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = l \ \text{ if for every } \epsilon > 0 \text{ there is } \delta > 0 \text{ such that} \\ \qquad\qquad |f(\mathbf{x}) - l| < \epsilon \ \text{ whenever } \ 0 < d(\mathbf{x}, \mathbf{x_0}) < \delta. \end{array}}$$

Definition 2. (Continuity) *Let $f(x, y)$ be a real-valued function of two variables defined in a disk $D_r(x_0, y_0)$ with center (x_0, y_0). Then*

$$\boxed{f \quad \textit{is continuous at } (x_0, y_0) \textit{ if } \lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0).}$$

We say f is continuous in $D_r(x_0, y_0)$ if f is continuous at every point (x, y) in $D_r(x_0, y_0)$. The intuitive meaning of continuity is that if the point (x, y) changes by a small amount, then the value of $f(x, y)$ changes by a small amount. Geometrically, this means that a surface that is the graph of a continuous function has no holes or breaks.

Definition 3. (Partial derivatives) *Let $f : D_r(x_0, y_0) \to \mathbb{R}$. The partial derivatives of f are the functions f_x and f_y defined by*

$$\boxed{\begin{aligned} f_x(x, y) &:= \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h}, \\ f_y(x, y) &:= \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h}. \end{aligned}}$$

To find f_x, treat y as a constant and differentiate $f(x, y)$ with respect to x. Similarly, to find f_y, treat x as a constant and differentiate $f(x, y)$ with respect to y. If $z = f(x, y)$ we write

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial z}{\partial x} = z_x,$$

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial z}{\partial y} = z_y.$$

Partial derivatives can also be defined for functions of three or more variables. In general, if z is a function of n variables, $z = f(x_1, x_2, \ldots, x_n)$, its partial derivative with respect to the ith variable x_i is

$$\boxed{\frac{\partial z}{\partial x_i} := \lim_{h\to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}.}$$

We also write

$$z_{x_i} = \frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} = f_{x_i}.$$

Since the partial derivatives are themselves functions, we can take their partial derivatives to obtain higher derivatives. If $z = f(x, y)$, we may compute

$$\boxed{\begin{aligned} f_{xx}(x, y) &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2}, \quad f_{yy}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial y^2}, \\ f_{xy}(x, y) &= \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial y \partial x}, \quad f_{yx}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial x \partial y}. \end{aligned}}$$

In general, $f_{xy} \neq f_{yx}$. However, the following theorem gives condition under which we can assert that $f_{xy} = f_{yx}$.

Theorem 4. *Let* $f : D_r(x_0, y_0) \to \mathbb{R}$. *If* f_{xy} *and* f_{yx} *are both continuous at* (x_0, y_0), *then*

$$\boxed{f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).}$$

Definition 5. (Chain rule) *Let* $z_1 = f_1(x_1, \ldots, x_n), \ \ldots, \ z_m = f_m(x_1, \ldots, x_n)$ *be* m *functions of* n *variables, and let* $x_1 = g_1(t_1, \ldots, t_k), \ldots, x_n = g_n(t_1, \ldots, t_k)$ *be* n *functions of* k *variables, all with continuous partial derivatives.*

Consider the z_i'*s as functions of the* t_j*'s by*

$$z_i = f_i(g_1(t_1, \ldots, t_k), \ldots, g_n(t_1, \ldots, t_k)).$$

Then

$$\boxed{\frac{\partial z_i}{\partial t_j} = \frac{\partial z_i}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \frac{\partial z_i}{\partial x_2}\frac{\partial x_2}{\partial t_j} + \cdots + \frac{\partial z_i}{\partial x_n}\frac{\partial x_n}{\partial t_j}.}$$

Definition 6. *If* $z = f(x, y)$ *is a function of two variables, its gradient vector field* ∇f *is defined by*

$$\boxed{\nabla f(x, y) := (f_x(x, y), f_y(x, y)) = (\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}).}$$

If $u = f(x, y, z)$ *is a function of three variables, its gradient vector field* ∇f *is defined by*

$$\nabla f(x, y, z) = (f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)) = (\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}).$$

Definition 7. (Implicit differentiation) *If* $y = f(x)$ *is a function satisfying the relation* $z = F(x, y) = 0$, *then*

$$\boxed{\frac{dy}{dx} = -\frac{F_x(x, f(x))}{F_y(x, f(x))}.} \tag{1}$$

Differentiating $F(x, y) = 0$ with respect to x using the chain rule gives

$$\begin{aligned} &\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0 \\ \implies \quad &\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0, \end{aligned}$$

which yields (1).

Definition 8. (Directional derivatives) *The directional derivative of* f *at* (x_0, y_0) *in the direction of a unit vector* $\mathbf{u} = (u_1, u_2)$ *is*

$$\boxed{\mathbb{D}_{\mathbf{u}} f(x_0, y_0) := \lim_{h \to 0} \frac{f(x_0 + hu_1, y_0 + hu_2) - f(x_0, y_0)}{h}}$$

if this limit exists.

Note that if $\mathbf{u} = (1, 0)$ then $\mathbb{D}_{\mathbf{u}}f = f_x$ and if $\mathbf{u} = (0, 1)$, then $\mathbb{D}_{\mathbf{u}}f = f_y$. In other words, the partial derivatives of f with respect to x and y are just special cases of the directional derivatives.

Theorem 9. *If $f(x, y)$ is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\mathbf{u} = (u_1, u_2)$ and*

$$\mathbb{D}_{\mathbf{u}}f(x, y) = f_x(x, y)u_1 + f_y(x, y)u_2.$$

The directional derivative can be written as

$$\begin{aligned} \mathbb{D}_{\mathbf{u}}f(x, y) &= f_x(x, y)u_1 + f_y(x, y)u_2 \\ &= (f_x(x, y), f_y(x, y)) \cdot (u_1, u_2) \\ &= \nabla f(x, y) \cdot \mathbf{u}. \end{aligned} \tag{2}$$

This expresses the directional derivative in the direction of $\mathbf{u}$ as the scalar projection of the gradient vector onto $\mathbf{u}$. From (2), we have

$$\begin{aligned} \mathbb{D}_{\mathbf{u}}f(x, y) &= \nabla f(x, y) \cdot \mathbf{u} \\ &= |\nabla f||\mathbf{u}| \cos\theta \\ &= |\nabla f| \cos\theta, \end{aligned}$$

where θ is the angle between ∇f and $\mathbf{u}$. The maximum value of $\cos\theta$ is 1 and this occurs when $\theta = 0$. Therefore, the maximum value of $\mathbb{D}_u f(x, y)$ is $|\nabla f|$ and it occurs when $\theta = 0$ i.e., when $\mathbf{u}$ has the same direction as ∇f.

Similarly, the directional derivative of functions of three variables with unit vector $\mathbf{u} = (u_1, u_2, u_3)$ can be written as

$$\mathbb{D}_{\mathbf{u}}f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}.$$

We now introduce the concept of differentiability for functions of several variable, let's first recall the definition of differentiability in one variable case.

Let D be an open subset $\mathbb{R}$. The function $f : D \to \mathbb{R}$ is said to be differentiable at $x_0 \in D$ if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. The value of this limit is called the derivative of f at x_0 and is denoted by $f'(x_0)$.

The above definition may be restated as follows: The function $f : D \to \mathbb{R}$ is differentiable at $x_0 \in D$ if there is a number $f'(x_0)$ such that

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - f'(x_0)(x - x_0)|}{|x - x_0|} = 0. \tag{3}$$

Any real number a_0 determines a linear transformation $L : \mathbb{R} \to \mathbb{R}$ defined by

$$Lx = a_0 x.$$

In particular, $f'(x_0)$ determines a linear transformation $L : \mathbb{R} \to \mathbb{R}$ given by $Lx = f'(x_0)x$. Therefore, with this linear transformation, we may rewrite (3) as

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - L(x - x_0)|}{|x - x_0|} = 0. \tag{4}$$

We now use (3) to define differentiability of a function $f : \mathbb{R}^n \to \mathbb{R}^m$.

Definition 10. (Differentiability) *Let $D \subset \mathbb{R}^n$ be an open subset and let $f : D \to \mathbb{R}^m$. We say that f is differentiable at $\mathbf{x_0} \in D$ if there is a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\boxed{\lim_{\mathbf{x} \to \mathbf{x_0}} \frac{\|f(\mathbf{x}) - f(\mathbf{x_0}) - L(\mathbf{x} - \mathbf{x_0})\|}{\|\mathbf{x} - \mathbf{x_0}\|} = 0.} \tag{5}$$

The linear transformation L of (5) is called the derivative of f at $\mathbf{x_0}$. We denote it by $f'(\mathbf{x_0})$.

We say that f is differentiable in D if it is differentiable at each every point of D.

Definition 11. (Jacobian matrix) *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is defined by*

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})),$$

where $f_i : D \to \mathbb{R}$, $1 \le i \le m$. For each $\mathbf{x} \in D$, we define the Jacobian matrix of f at $\mathbf{x}$ by

$$J_f(\mathbf{x}) := (a_{ij}),$$

where $a_{ij} = (\partial f_i / \partial x_j)(\mathbf{x})$, $1 \le i \le m$, $1 \le j \le n$.

Example 12.

Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by

$$f(x_1, x_2) = (x_1^2, x_1 x_2, x_2^2).$$

Here, $f_1(x_1, x_2) = x_1^2$, $f_2(x_1, x_2) = x_1 x_2$, $f_3(x_1, x_2) = x_2^2$. Then

$$\frac{\partial f_1}{\partial x_1} = 2x_1, \quad \frac{\partial f_2}{\partial x_1} = x_2, \quad \frac{\partial f_3}{\partial x_1} = 0.$$

$$\frac{\partial f_1}{\partial x_2} = 0, \quad \frac{\partial f_2}{\partial x_2} = x_1, \quad \frac{\partial f_3}{\partial x_2} = 2x_2$$

Therefore,

$$J_f(x_1, x_2) = \begin{bmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{bmatrix}.$$

The following theorem gives a formula for computing derivative.

Theorem 13. (Computing derivative) *Let D be an open subset of $\mathbb{R}^n$ and $f : D \to \mathbb{R}^m$ be defined by*

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})),$$

where $f_i : D \to \mathbb{R}$, $1 \le i \le m$. If f is differentiable at a point $\mathbf{x_0}$ in D, then each of the partial derivatives $(\partial f_i/\partial x_j)(\mathbf{x_0})$ exists, $1 \le i \le m$, $1 \le j \le n$. Furthermore,

$$\boxed{f'(\mathbf{x_0}) = J_f(\mathbf{x_0}).}$$

Note that the linear transformation L is defined by the Jacobian matrix of f at $\mathbf{x_0}$. In particular, for $m = 1$, we have

$$L = f'(\mathbf{x_0}) = \nabla f(\mathbf{x_0}).$$

The following theorem gives the sufficient condition for differentiability of f.

Theorem 14. (Sufficient condition for differentiability) *Let $D \subset \mathbb{R}^n$ be an open set and $f : D \to \mathbb{R}^m$ be defined by*

$$f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})),$$

where $f_i : D \to \mathbb{R}$, $1 \le i \le m$. Suppose that $(\partial f_i/\partial x_j)(\mathbf{x_0})$ exists and continuous on D, $1 \le i \le m$, $1 \le j \le n$. Then f is differentiable on D.

We shall conclude this lecture by stating some results on maxima and minima in the case of a function of several variables. We restrict our discussion to functions of two variables only.

Definition 15. (Maxima and Minima) *Let $f(x, y)$ be a function of two variables. A point (x_0, y_0) is a local minimum point for f if there is a disk $D_\delta(x_0, y_0)$ about (x_0, y_0) such that*

$$f(x, y) \ge f(x_0, y_0) \text{ for all } (x, y) \in D_\delta(x_0, y_0).$$

The number $f(x_0, y_0)$ is called a local minimum value.

Similarly, if there is a disk $D_\delta(x_0, y_0)$ about (x_0, y_0) such that

$$f(x,y) \le f(x_0,y_0) \;\; \textit{for all}\; (x,y) \in D_\delta(x_0,y_0)$$

then the point (x_0, y_0) a local maximum point for f.

A point which is either a local maximum or minimum point is called a local extremum.

The following is the analog in two variables of the first derivative test for one variable.

First Derivative Test:
If (x_0, y_0) is a local extremum of f and the partial derivatives of f exist at (x_0, y_0), then

$$\boxed{f_x(x_0,y_0) = f_y(x_0,y_0) = 0.}$$

Such point (x_0, y_0) is also called a critical point of f.

Second Derivative Test:
Let $f(x, y)$ have continuous second-order partial derivatives, and suppose that (x_0, y_0) is a critical point for f. Then

$$f_x(x_0,y_0) = 0 \quad \text{and} \quad f_y(x_0,y_0) = 0.$$

Let $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$, and $C = f_{yy}(x_0, y_0)$. Then the following statements are true.

(a) If $A > 0$, $AC - B^2 > 0$ then (x_0, y_0) is a local minimum.
(b) If $A < 0$, $AC - B^2 > 0$ then (x_0, y_0) is a local maximum.
(c) If $AC - B^2 < 0$ then (x_0, y_0) is a saddle point.
(d) If $AC - B^2 = 0$ then the test is inconclusive.

Practice Problems

1. Show that $\lim_{(x,y)\to(0,0)} \frac{\partial}{\partial x}\sqrt{x^2+y^2}$ does not exist.

2. Using ϵ and δ definition prove that $f(x, y) = |x|$ is continuous at $(0, 0)$.

3. Let
$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2}, & (x,y) \ne (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$
   (a) If $(x, y) \ne (0, 0)$, compute f_x and f_y.
   (b) What is the value of $f(x, 0)$ and $f(0, y)$?
   (c) Show that $f_x(0, 0) = 0 = f_y(0, 0)$.
   (d) Show that $f_{yx}(0, 0) = 1$ and $f_{xy}(0, 0) = -1$.
   (e) What went wrong? why are the mixed partial not equal?

4. Find the derivative of the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$f(x, y) = (x^2 + xy, x - y^2).$$

5. Find the maxima, minima and saddle points of $f(x, y) = (x^2 - y^2)e^{(-x^2-y^2)/2}$.