MA2104 - Multivariable Calculus

Liew Zhao Wei

Semester 1, 2023-2024

---

1 Vectors

Multivariable calculus is the study of *scalar fields* and *vector fields*.

Definition 1.1 (Scalar Field)
A *scalar field* is a map $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$. It is also called a *scalar-valued* function.

Definition 1.2 (Vector Field)
A *vector field* is a map $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$. It is also called a *vector-valued* function.

2 Coordinate Systems

Cylindrical and spherical coordinates are both natural extensions of the polar coordinates in $\mathbb{R}^2$, but in slightly different ways – cylindrical coordinates tack on a z-coordinate, while spherical coordinates tack on an angle ρ from the positive z-axis.

2.1 Cylindrical Coordinates

Cylindrical coordinates are useful for describing objects with an axis of symmetry.

Definition 2.1 (Cylindrical Coordinate Conversions)
We can convert between cylindrical and cartesian coordinates like so:

$$\text{Cylindrical to Cartesian: } \begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ z = z \end{cases} \quad \text{Cartesian to Cylindrical: } \begin{cases} r^2 = x^2 + y^2 \\ \tan\theta = y/x \\ z = z \end{cases}$$

We typically impose the constraints $r \geq 0$ and $0 \leq \theta < 2\pi$ to ensure that all points on $\mathbb{R}^3$, except those on the z-axis, are uniquely represented.

2.2 Spherical Coordinates

Spherical coordinates are useful for describing objects with a centre of symmetry.

Definition 2.2 (Spherical Coordinate Conversions)
We can convert between sphericla and cartesian coordinates like so:

$$\text{Spherical to Cartesian: } \begin{cases} x = \rho\sin\varphi\cos\theta \\ y = \rho\sin\varphi\sin\theta \\ z = \rho\cos\varphi \end{cases} \quad \text{Cartesian to Spherical: } \begin{cases} \rho^2 = x^2 + y^2 + z^2 \\ \tan\theta = y/x \\ \cos\varphi = \sqrt{x^2+y^2}/z \end{cases}$$

We typically impose the constraints $\rho \geq 0$, $0 \leq \theta < 2\pi$ and $0 \leq \varphi \leq \pi$ to ensure that all points on $\mathbb{R}^3$, except the origin, are uniquely represented.

2.3 Hyperspherical Coordinates

We can even generalize spherical coordinates to higher dimensions in the form of *hyperspherical coordinates*. The idea is to tack on more angles to the spherical coordinates, so that we can describe objects with more axes of symmetry.

Definition 2.3 (Hyperspherical Coordinate Conversions)

$$\text{Spherical to Cartesian: } \begin{cases} x_1 = \rho \sin\varphi_1 \cdots \sin\varphi_{n-2} \cos\varphi_{n-1} \\ x_2 = \rho \sin\varphi_1 \cdots \sin\varphi_{n-2} \sin\varphi_{n-1} \\ x_3 = \rho \sin\varphi_1 \cdots \sin\varphi_{n-3} \cos\varphi_{n-2} \\ x_4 = \rho \sin\varphi_1 \cdots \sin\varphi_{n-4} \cos\varphi_{n-3} \\ \vdots \\ x_n = \rho \cos\varphi_1 \end{cases}$$

Notice the similarity with spherical coordinates.

Just like for spherical coordinates, we typically impose the constraints

$$\begin{aligned} 0 \leq &\rho \\ 0 \leq &\varphi_i \leq \pi && \text{for } i = 1, \ldots, n-2 \\ 0 \leq &\varphi_{n-1} < 2\pi \end{aligned}$$

to ensure that all points on $\mathbb{R}^n$, except the origin, are uniquely represented.

3 Basic Topology

We begin with some basic topological definitions and results that we'll use later on.

Definition 3.1 (Neighbourhood)
A *neighbourhood* of a point p is a set $N_r(p)$ consisting of all points q such that $\|q - p\| < r$ for some $r > 0$.

An *open ball* with center $p \in \mathbb{R}^n$ and radius r is a neighbourhood $N_r(p)$ in $\mathbb{R}^n$. A *closed ball* is similarly defined, but with $\|q - p\| \leq r$ instead.

Definition 3.2 (Limit point)
A point p is a *limit point* of a set S if every neighbourhood of p contians a point $q \in X$ such that $q \neq p$. Such a point is also called an *accumulation point*.

Definition 3.3 (Isolated point)
An *isolated point* of a set S is a point $p \in S$ such that p is not a limit point of S. That is, there is some neighbourhood of p that contains no other points of S.

Definition 3.4 (Interior point)
An *interior point* of a set S is a point $p \in S$ such that there is some neighbourhood of p that is contained in S.

Definition 3.5 (Open set)
A set S is *open* if every point of S is an interior point of S.

Definition 3.6 (Closed set)
A set S is *closed* if it contains all of its limit points.

Theorem 3.7 (Complement of open set is closed)
A set is open if and only if its complement is closed.

4 Limits and Continuity

We begin by extending the definitions of limits and continuity to vector fields.

4.1 Limits

Definition 4.1 (Limit)
Consider a function $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$. We write

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{L} \tag{1}$$

to mean that given any $\epsilon > 0$, there is a $\delta > 0$ such that if $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$, then $\|\mathbf{f}(\mathbf{x}) - \mathbf{L}\| < \epsilon$.

Intuitively, this means that $\mathbf{f}(\mathbf{x})$ can be made arbitrarily close to $\mathbf{L}$ by taking $\mathbf{x}$ sufficiently close to $\mathbf{a}$. The geometric interpretation is that given an open ball B_ϵ centred at $\mathbf{L}$, there is an open ball B_δ centred at $\mathbf{a}$ such that $\mathbf{f}(\mathbf{x})$ is contained in B_ϵ whenever $\mathbf{x}$ is contained in B_δ.

The usual theorems concerning limits of sums, products, and quotients hold for scalar fields. For vector fields, we have natural extensions of these theorems, but without quotients.

Theorem 4.2 (Algebraic Properties of Limits)
Suppose that $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{L}$ and $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{g}(\mathbf{x}) = \mathbf{M}$. Then

1. $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) = \mathbf{L} + \mathbf{M}$
2. $\lim_{\mathbf{x}\to\mathbf{a}} c\mathbf{f}(\mathbf{x}) = c\mathbf{L}$ for any scalar c
3. $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x}) = \mathbf{L} \cdot \mathbf{M}$
4. $\lim_{\mathbf{x}\to\mathbf{a}} \|\mathbf{f}(\mathbf{x})\| = \|\mathbf{L}\|$

(3) is proved by rewriting

$$\mathbf{f}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x}) - \mathbf{L} \cdot \mathbf{M} = (\mathbf{f}(\mathbf{x}) - \mathbf{L}) \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{M}) + \mathbf{L} \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{M}) + \mathbf{M} \cdot (\mathbf{f}(\mathbf{x}) - \mathbf{L})$$

(4) is proved by using (3) and the fact that $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Theorem 4.3 (Uniqueness of Limits)
If $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x})$ exists, then it is unique.

Theorem 4.4 (Limit Exists Iff Limit of Components Exist)
Consider a function $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Then $\lim\limits_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x})$ exists if and only if $\lim_{\mathbf{x}\to\mathbf{a}} f_i(\mathbf{x})$ exists for each $i = 1, \ldots, m$. In this case,

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \left(\lim_{\mathbf{x}\to\mathbf{a}} f_1(\mathbf{x}), \ldots, \lim_{\mathbf{x}\to\mathbf{a}} f_m(\mathbf{x})\right) \tag{2}$$

This is proved by observing that $f_i = \mathbf{f} \cdot \mathbf{e}_i$, and applying the algebraic properties of limits.

4.2 Continuity

Definition 4.5 (Continuity)
Consider a function $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and let $a \in X$. $\mathbf{f}$ is *continuous at* $\mathbf{a}$ if either $\mathbf{a}$ is an isolated point of X or if

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) \tag{3}$$

If $\mathbf{f}$ is continuous at every point of X, then we simly say that $\mathbf{f}$ is *continuous*.

Observe that many of the properties of limits also hold for continuity.

Theorem 4.6 (Continuity of Composite Functions)
If $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $\mathbf{a}$ and $\mathbf{g}\colon Y \subseteq \mathbb{R}^m \to \mathbb{R}^p$ is continuous at $\mathbf{f}(\mathbf{a})$, then $\mathbf{g} \circ \mathbf{f}$ is continuous at $\mathbf{a}$.

5 Derivatives

Now, we extend the definition of derivatives to vector fields.

5.1 Partial Derivatives

Definition 5.1 (Directional Derivative)
Consider a scalar field $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ and let $\mathbf{a} \in X$. If $\mathbf{v}$ is a unit vector, then the *directional derivative of f at $\mathbf{a}$ in the direction of $\mathbf{v}$* is

$$D_{\mathbf{v}} f(\mathbf{a}) = \lim_{h\to 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h} \tag{4}$$

if the limit exists.

Definition 5.2 (Partial Derivative)
Consider a scalar field $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ and let $\mathbf{a} \in X$. The *partial derivative of f with respect to x_i at $\mathbf{a}$* is the directional derivative of f at $\mathbf{a}$ in the direction of $\mathbf{e}_i$ and is denoted by $\dfrac{\partial f}{\partial x_i}$, $f_{x_i}(\mathbf{a})$, or $D_{x_i} f(\mathbf{a})$.

Definition 5.3 (Smoothness of a Function)
Let $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a scalar field whose partial derivatives up to order k exist and are continuous. Then, f is said to be *of class C^k*. If f is of class C^k for all $k \geq 1$, then f is said to be *of class C^∞* or *smooth*.

Theorem 5.4 (Order of Continuous Partial Derivatives Does not Matter)
Let $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a scalar field of class C^k. Then the order in which we compute any kth-order partial derivative does not matter. That is, for any $i_1, \ldots, i_k \in \{1, \ldots, n\}$, we have

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{\sigma(i_1)} \cdots \partial x_{\sigma(i_k)}} \tag{5}$$

5.2 Total derivatives

The definitions of differentiability for scalar fields generalize naturally to vector fields, so I will simply state the ones for vector fields.

Definition 5.5 (Jacobian Matrix)
Consider a vector field $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Define $D\mathbf{f}$ to be

$$D\mathbf{f} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{6}$$

$D\mathbf{f}$ is called the *Jacobian matrix* of $\mathbf{f}$. The ith row of $D\mathbf{f}$ is the gradient vector ∇f_i.

Observe that $D\mathbf{f}$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. We will see that if a derivative exists at $\mathbf{a}$, then it must be equal to $D\mathbf{f}(\mathbf{a})$.

Definition 5.6 (Differentiability)
Consider a vector field $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and let $\mathbf{a} \in X$. $\mathbf{f}$ is *differentiable at* $\mathbf{a}$ if the linear map $D\mathbf{f}(\mathbf{a})$ exists and

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - (\mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}))\|}{\|\mathbf{x} - \mathbf{a}\|} = 0 \tag{7}$$

$\mathbf{f}$ is also said to be *differentiable at* $\mathbf{a}$.

If $\mathbf{f}$ is differentiable at $\mathbf{a}$, then the function $\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$ is the *tangent hyperplane* of $\mathbf{f}$ at $\mathbf{a}$ and is a good linear approximation of $\mathbf{f}$ near $\mathbf{a}$.

The results for scalar fields generalize naturally to vector fields too.

Theorem 5.7 (Differentiability Implies Continuity)
If $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\mathbf{a}$, then it is also continuous at $\mathbf{a}$.

Theorem 5.8 (Continuity and Existence of Partial Derivatives Imply Differentiability)
If $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $\mathbf{a}$ and if all the partial derivatives of $\mathbf{f}$ exist in some neighbourhood of $\mathbf{a}$, then $\mathbf{f}$ is differentiable at $\mathbf{a}$.

Theorem 5.9 (Differentiable Iff Components are Differentiable)
Consider a vector field $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$. $\mathbf{f}$ is differentiable at $\mathbf{a}$ if and only if each component function f_i is differentiable at $\mathbf{a}$.

Some familiar algebraic properties of derivatives also hold for vector fields.

Theorem 5.10 (Linearity of Differentiation)
If $\mathbf{f}, \mathbf{g}: X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are differentiable at $\mathbf{a}$ and c is a scalar, then

1. $D(c\mathbf{f})(\mathbf{a}) = cD\mathbf{f}(\mathbf{a})$
2. $D(\mathbf{f} + \mathbf{g})(\mathbf{a}) = D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a})$

For scalar-valued functions, product and quotient properties hold too.

Theorem 5.11 (Product and Quotient of Derivatives)
If $f, g: X \subseteq \mathbb{R}^n \to \mathbb{R}$ are differentiable at $\mathbf{a}$, then

1. $D(fg)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$
2. $D\left(\frac{f}{g}\right)(\mathbf{a}) = \frac{g(\mathbf{a})Df(\mathbf{a}) - f(\mathbf{a})Dg(\mathbf{a})}{g(\mathbf{a})^2}$

5.3 Chain Rule

We use the chain rule so much that it deserves its own section.

Theorem 5.12 (Chain Rule)
Consider the vector fields $\mathbf{f}: X \subseteq \mathbb{R}^m \to \mathbb{R}^p$ and $\mathbf{g}: Y \subseteq \mathbb{R}^n \to \mathbb{R}^m$. If $\mathbf{g}$ is differentiable at $\mathbf{x} \in Y$ and $\mathbf{f}$ is differentiable at $\mathbf{g}(\mathbf{x})$, then $\mathbf{f} \circ \mathbf{g}$ is differentiable at $\mathbf{x}$ and

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = D\mathbf{f}(\mathbf{g}(\mathbf{x}))D\mathbf{g}(\mathbf{x}) \tag{8}$$

The chain rule is also used in implicit differentiation under a condition stipulated by the following theorem.

Theorem 5.13 (Implicit Function Theorem)
Consider the scalar field $F: X \subseteq \mathbb{R}^n \to \mathbb{R}$ of class C^1. Let $\mathbf{a}$ be a point of the level set $S = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x} = c\}$. If $F_{x_n}(\mathbf{a}) \neq 0$, then there is a neighbourhood U of $(a_1, \ldots, a_{n-1}) \in \mathbb{R}^{n-1}$, a neighbourhood V of $a_n \in \mathbb{R}$, and a function $f: U \subseteq \mathbb{R}^{n-1} \to V$ of class C^1 such that if $(x_1, \ldots, x_{n-1}) \in U$ and $x_n \in V$ satisfy $F(x_1, \ldots, x_{n-1}, x_n) = c$, then $x_n = f(x_1, \ldots, x_{n-1})$.

6 Maxima, Minima, and Saddle Points

Calculus is largely useful in solving optimization problems involving the extrema of scalar fields.

Definition 6.1 (Stationary Point)
A point $\mathbf{a}$ is a *stationary point* of a differentiable scalar field f if $\nabla f(\mathbf{a}) = \mathbf{0}$.

Definition 6.2 (Critical Point)
A point $\mathbf{a}$ is a *critical point* of a differentiable scalar field f if it is stationary or if $\nabla f(\mathbf{a})$ does not exist.

Definition 6.3 (Extrema)
Let $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a differentable scalar field. A point $\mathbf{a}$ is a *local/relative minimum* of f if there is some neighbourhood N of $\mathbf{a}$ such that $f(\mathbf{a}) \leq f(\mathbf{x})$ for all $\mathbf{x} \in N$. Similarly, $\mathbf{a}$ is a *local/relative maximum* of f if there is some neighbourhood N of $\mathbf{a}$ such that $f(\mathbf{a}) \geq f(\mathbf{x})$ for all $\mathbf{x} \in N$. If $\mathbf{a}$ is either a local minimum or a local maximum, then it is an *extremum* of f.

Theorem 6.4 (Extremum Implies Stationary)
Let $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar field. If $\mathbf{a} \in X$ is a local extremum of f, then it is stationary. That is, $\nabla f(\mathbf{a}) = \mathbf{0}$.

Note that the converse is not true, which is the case for saddle points.

Definition 6.5 (Saddle Point)
Let $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar field. A stationary point $\mathbf{a}$ is a *saddle point* if $\mathbf{a}$ is not a local extremum of f. That is, every neighbourhood of $\mathbf{a}$ contains points $\mathbf{x}$ such that $f(\mathbf{x}) < f(\mathbf{a})$ and others such that $f(\mathbf{x}) > f(\mathbf{a})$.

6.1 Second Derivative Test

To determine the nature of a critical point, we can use the *second partial derivative test.*

Definition 6.6 (Hessian Matrix)
Consider a differentiable scalar field $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$. The *Hessian matrix* of f is

$$Hf = \begin{bmatrix} f_{x_1x_1} & \cdots & f_{x_1x_n} \\ \vdots & \ddots & \vdots \\ f_{x_nx_1} & \cdots & f_{x_nx_n} \end{bmatrix} \tag{9}$$

Theorem 6.7 (Second Partial Derivative Test)
Let $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar field. Given a stationary point $\mathbf{a} \in X$ of f, consider the sequence of leading principal minors d_k of the Hessian matrix $Hf(\mathbf{a})$ where

$$d_k = \begin{vmatrix} f_{x_1x_1}(\mathbf{a}) & \cdots & f_{x_1x_k}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ f_{x_kx_1}(\mathbf{a}) & \cdots & f_{x_kx_k}(\mathbf{a}) \end{vmatrix} \tag{10}$$

for all $1 \leq k \leq n$. Then, the following cases hold:

1. If $d_k > 0$ for all k, then $\mathbf{a}$ is a local minimum of f.
2. If $d_k > 0$ for even k and $d_k < 0$ for odd k, then $\mathbf{a}$ is a local maximum of f.
3. Otherwise, $\mathbf{a}$ is a saddle point of f.

If $\det Hf(\mathbf{a}) = 0$, we say that $\mathbf{a}$ is *degenerate.*

The cases above can also be expressed in terms of the eigenvalues of $Hf(\mathbf{a})$:

1. If all eigenvalues are positive (equivalently, the Hessian is positive definite), then $\mathbf{a}$ is a local minimum of f.
2. If all eigenvalues are negative (equivalently, the Hessian is negative definite), then $\mathbf{a}$ is a local maximum of f.
3. Otherwise, if $\det Hf(\mathbf{a}) \neq 0$, then $\mathbf{a}$ is a saddle point of f.

6.2 Langrange's Multipliers

Lagrange's Multipliers come in handy when solving extremum problems with constraints. The method stems from the following theorem.

Theorem 6.8 (Lagrange's Multipliers)
Let $f, g \colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ be scalar fields of class C^1. Let $S = \{x \in X \mid g(\mathbf{x}) = c\}$ denote the level set of g at height c. Then if f restricted to S has an extremum at point $\mathbf{a} \in S$ such that $\nabla g(\mathbf{a}) \neq \mathbf{0}$, then

$$\nabla f(\mathbf{a}) = \lambda \nabla g(\mathbf{a}) \tag{11}$$

for some scalar λ.