# Revision Book

MATH1821: Mathematical Methods for Actuarial Science I

2024-2025 1st Semester

Course Instructor: Dr. Law Ka Ho
lawkaho@connect.hku.hk

Student Teaching Assistant: YU Xinyue, Christina
xinyue_yu@connect.hku.hk

December, 2024

# Contents

# Chapter 1

# Introduction

This revision book is a collection of materials for all revision classes in 2024-2025 1st semester of the course MATH1821. Students may use this book for knowledge review or practice. The suggested solutions are attached just after the questions.
Some of you may use this book as a reference for your further learning in course MATH2822 or other mathematics courses.

# Chapter 2

# Notes for Review

## 2.1 Chapter 1: Basic Concepts

### 2.1.1 Functions

**Definition 2.1.1** (Function)**.** For any sets $X$ and $Y$, we say that $f$ is a **function** from $X$ to $Y$, denoted $f : X \to Y$, if it provides a rule which assigns to every element $x \in X$ exactly one element $y \in Y$.
A function is also called a mapping.

**Definition 2.1.2** (Domain, codomain and range)**.** In defining a function $f$, theoretically we have to specify two sets $X$ and $Y$. They are called the **domain** (dom $f$) and **codomain** of $f$ respectively.
**Natural domain:** the largest subset of $\mathbb{R}$ for which the function formula is well defined.
**Restricted domain:** if there is a story (or some restrictions) on the function, then the restricted domain will be different from the natural domain.
**Range:** the range of a function $f : X \to Y$ is the set $\{f(x) : x \in X\}$, which is the set of all possible outputs.
*Usually, the range of a function is a subset of its codomain, not necessary to be a real subset.*

Different classification of functions:

- **piecewise-defined functions** (e.g. absolute value function)
- **odd functions** ($f(-x) = -f(x)$ for all $x$) and **even functions** ($f(-x) = f(x)$ for all $x$)
- **increasing functions** (if $a > b$ implies $f(a) \geq f(b)$) and **decreasing functions** (if $a > b$ implies $f(a) \leq f(b)$)
- **strictly increasing functions** (if $a > b$ implies $f(a) > f(b)$) and **strictly decreasing functions** (if $a > b$ implies $f(a) < f(b)$)
- **random variable**: frequently used in probability and statistics, map from the sample space, which is the set of all possible outcomes of the experiment, to $\mathbb{R}$

**Definition 2.1.3** (Image and preimage)**.** Let $f : X \to Y$ be a function. For $S \subseteq X$, the **image** of $S$ under $f$, denoted by $f(S)$, is the set $\{f(x) : x \in S\}$.
For $T \subseteq Y$, the **preimage** of $T$ under $f$, denoted by $f^{-1}(T)$, is the set $\{x \in X : f(x) \in T\}$.

**Definition 2.1.4** (Injective and surjective)**.** Let $f : X \to Y$ be a function. We say that $f$ is

- **injective** if for any $x_1 \neq x_2$ we have $f(x_1) \neq f(x_2)$.
- **surjective** if for any $y \in Y$ there exists $x \in X$ for which $f(x) = y$.
- **bijective** if the function is injective and surjective. A bijective function leads to a one-to-one correspondence between elements of $X$ and $Y$.

**Definition 2.1.5** (Composition of functions)**.** Let $f : X \to Y$ and $g : Y \to Z$ be functions. This naturally corresponds to a function from $X$ to $Z$. Such a function is called the **composition** of $g$ and $f$. Notation: $g \circ f$, normally, we have $(g \circ f)(x) = g(f(x))$.

**Definition 2.1.6** (Inverse functions)**.** Let $f : X \to Y$ be a function. We say that a function $g : Y \to X$ is the **inverse** of $f$, denoted $g = f^{-1}$, if

$$(g \circ f)(x) = x \text{ for all } x \in X \text{ and } (f \circ g)(y) = y \text{ for all } y \in Y.$$

*Remark.* Do not confuse the notation of inverse function ($f^{-1}(y)$ for $y \in Y$) and preimage ($f^{-1}(T)$ for $T \subseteq Y$).

**Theorem 2.1.1.** *Let $f : X \to Y$ be a function. Then the inverse function $f^{-1}$ exists **if and only if** $f$ is bijective.*

### 2.1.2 Trigonometric Functions

**Definition 2.1.7** (Radian measure)**.** A central angle subtended by an arc of length equal to the radius of a circle is said to have **radian measure 1**, written 1 rad. In general, if the central angle $\theta$ is subtended by an arc of length $s$, then

$$s : r = \theta : 1,$$

and hence

$$s = r\theta.$$

Similarly, the area of a sector with central angle $\theta$ is

$$\tfrac{1}{2}r^2\theta.$$

**Definition 2.1.8** (Cotangent, secant and cosecant)**.**

$$\cot\theta = \frac{\text{adjacent side}}{\text{opposite side}}, \quad \sec\theta = \frac{\text{hypotenuse}}{\text{adjacent side}}, \quad \csc\theta = \frac{\text{hypotenuse}}{\text{opposite side}}.$$

**Definition 2.1.9** (Inverse trigonometric functions)**.** For $x \in [-1, 1]$, we define the **arcsine** of $x$, denoted by $\sin^{-1}$ or $\arcsin x$, to be the (unique) value of $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ such that $\sin\theta = x$. Similarly, we can define the functions:

- $\cos^{-1} : [-1, 1] \to [0, \pi]$
- $\tan^{-1} : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$

**Compound angle formulas:**

- $\sin(A + B) = \sin A \cos B + \cos A \sin B$
- $\sin(A - B) = \sin A \cos B - \cos A \sin B$
- $\cos(A + B) = \cos A \cos B - \sin A \sin B$
- $\cos(A - B) = \cos A \cos B + \sin A \sin B$
- $\tan(A + B) = \dfrac{\tan A + \tan B}{1 - \tan A \tan B}$
- $\tan(A - B) = \dfrac{\tan A - \tan B}{1 + \tan A \tan B}$

**Double and triple angle formulas:**

- $\sin 2A = 2 \sin A \cos A$
- $\cos 2A = \cos^2 A - \sin^2 A = 2\cos^2 A - 1 = 1 - 2\sin^2 A$
- $\tan 2A = \dfrac{2 \tan A}{1 - \tan^2 A}$
- $\sin 3A = 3 \sin A - 4 \sin^3 A$
- $\cos 3A = 4 \cos^3 A - 3 \cos A$
- $\tan 3A = \dfrac{3 \tan A - \tan^3 A}{1 - 3 \tan^2 A}$

**Half angle formulas:**

- $\sin^2 \dfrac{A}{2} = \dfrac{1-\cos A}{2}$
- $\cos^2 \dfrac{A}{2} = \dfrac{1+\cos A}{2}$
- $\tan^2 \dfrac{A}{2} = \dfrac{1-\cos A}{1+\cos A}$

**Product-to-sum formulas:**

- $\sin A \cos B = \dfrac{1}{2}[\sin(A+B) + \sin(A-B)]$
- $\sin A \sin B = \dfrac{1}{2}[\cos(A-B) - \cos(A+B)]$
- $\cos A \cos B = \dfrac{1}{2}[\cos(A-B) + \cos(A+B)]$

**Sum-to-product formulas:**

- $\sin A + \sin B = 2\sin(\dfrac{A+B}{2})\cos(\dfrac{A-B}{2})$
- $\sin A - \sin B = 2\cos(\dfrac{A+B}{2})\sin(\dfrac{A-B}{2})$
- $\cos A + \cos B = 2\cos(\dfrac{A+B}{2})\cos(\dfrac{A-B}{2})$
- $\cos A - \cos B = -2\sin(\dfrac{A+B}{2})\sin(\dfrac{A-B}{2})$

### 2.1.3 Complex Numbers

**Definition 2.1.10** (Complex Numbers)**.** The set of complex numbers is $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$, where $i$ is the **imaginary unit** satisfying $i^2 = -1$.

**Definition 2.1.11** (Conjugates)**.** For a complex number of the form $z = a + bi$, we define $\bar{z} = a - bi$ to be the **conjugate** of $z$.

Properties of conjugates

- $\overline{z \pm z'} = \overline{z} \pm \overline{z}'$
- $\overline{zz'} = \overline{z} \cdot \overline{z'}$
- $\overline{\left(\dfrac{z}{z'}\right)} = \dfrac{\overline{z}}{\overline{z'}}$ provided that $z' \neq 0$
- $z + \overline{z}$ is real
- $z - \overline{z}$ is purely imaginary
- $z\overline{z} = |z|^2$, where $|z|$ denotes the **modulus** of $z$

**Theorem 2.1.2.** *Let $f(x)$ be a polynomial with real coefficients. If $f(z) = 0$ for some $z \in \mathbb{C}$, then $f(\overline{z}) = 0$.*

**Definition 2.1.12** (Polar form of complex numbers)**.** We can change the expression of a complex number from rectangular coordinates to polar coordinates $(r, \theta)$.

$$z = a + bi = r\cos\theta + i\sin\theta = r \text{ cis } \theta \quad \text{where } r = |z|.$$

Computing with polar form can simplify computations in complex numbers.

- $z_1 z_2 = r_1 r_2 \text{ cis}(\theta_1 + \theta_2)$;
- $\dfrac{z_1}{z_2} = \dfrac{r_1}{r_2} \text{ cis}(\theta_1 - \theta_2)$ provided that $z_2 \neq 0$.

**Theorem 2.1.3** (n-th roots of complex number)**.** *Let $\alpha = r\ cis\ \theta$ where $r > 0$ and $\theta \in [0, 2\pi)$. Then for any $n \in \mathbb{N} \setminus \{1\}$, the equation $z^n - \alpha = 0$ has $n$ distinct roots*

$$z_k = r^{1/n}\ cis\ \frac{\theta + 2k\pi}{n},\ k = 0, 1, \cdots, n-1.$$

**Theorem 2.1.4** (Euler's formula)**.** *For $\theta \in \mathbb{R}$, we define*

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

**Theorem 2.1.5** (Index laws of complex numbers)**.** *For any $z \in \mathbb{C}$ with $z = a + bi$, we define $e^z = e^{a+bi} = e^a e^{bi} = e^a(\cos b + i\sin b)$ It can be checked that such definition preserves the index laws*

$$e^{z+z'} = e^z \cdot e^{z'} \text{ for any } z, z' \in \mathbb{C}.$$

### 2.1.4 Mathematical Induction

**Principle of mathematical induction:**
To prove that a statement $S(n)$ holds for all positive integers n, we can proceed as follows:

1. Prove that $S(1)$ is true.
2. Assume that $S(k)$ is true for some positive integer $k$.
3. Using the induction hypothesis, prove that $S(k+1)$ is true.

*Remark.* There are a lot of variations of mathematical inductions, you will learn them systematically in course MATH2012. In general, We firstly show that the $S(n)$ is true when $n$ is some small number. Then we assume $S(n)$ is true for some $n = k$, based on this, we try to show that $S(n)$ is true when $n$ is greater than $k$ or in some situations that will occur after $n = k$. See Exercise 1F, example (d), (e) in lecture notes.

## 2.2 Chapter 2: Limits and Continuity

**Definition 2.2.1** (Limit)**.** Intuitively, we write

$$\lim_{x \to c} f(x) = L$$

to mean the value of $f(x)$ gets closer and closer to $L$ as $x$ gets closer and closer to $c$.

*Remark.* This is **not** a rigorous definition for limit, you will learn more in course MATH2241.

Limits may not exist:

- limit is taken at $\pm\infty$
- the value of $f(x)$ oscillates as $x \to c$
- $f(x)$ approaches different values from the left and from the right

**Definition 2.2.2** (One-sided limits)**.** We say that the left-hand limit of $f(x)$ as $x$ tends to $c$ is equal to $L$, denoted by $\lim_{x \to c^-} f(x) = L$, if $f(x)$ gets closer and closer to $L$ as $x$ gets closer and closer to $c$ from the left.
Similarly, we say that the right-hand limit of $f(x)$ as $x$ tends to $c$ is equal to $L$, denoted by $\lim_{x \to c^+} f(x) = L$, if $f(x)$ gets closer and closer to $L$ as $x$ gets closer and closer to $c$ from the right.

**Theorem 2.2.1** (Sandwich theorem)**.** *Suppose that $I$ is an interval with $c \in I$, and that $g(x) \le f(x) \le h(x)$ for all $x \in I \setminus \{c\}$. If $\lim_{x \to c} g(x) = \lim_{x \to c} h(x) = L$ then $\lim_{x \to c} f(x) = L$ as well.*

**Theorem 2.2.2** (Divergence to infinity)**.** *Suppose that $I$ is an interval with $c \in I$, and that $g(x) \le f(x)$ for all $x \in I \setminus \{c\}$. If $\lim_{x \to c} g(x) = \infty$, then $\lim_{x \to c} f(x) = \infty$ as well.*

*Remark.* We can replace $\infty$ by $-\infty$; True for one-sided limits; We also have a similar result for limits at infinity instead of limit at $c$.

**Special limits:**

- $\lim_{x\to 0} \frac{\sin x}{x} = 1$
- $\lim_{x\to \infty} \frac{e^x}{x} = \infty$
- $\lim_{x\to 0^+} \frac{\ln x}{x} = -\infty$

**Definition 2.2.3** (Continuous). Let $f$ be a function and $c$ be an interior point of dom $f$ (i.e. there exists an open interval $I$ containing $c$ such that $I \subseteq \text{dom } f$). We say that $f$ is **continuous at** $c$ if $\lim_{x\to c} f(x)$ exists and is equal to $f(c)$.

**Definition 2.2.4** (Boundary and isolated points). $X$ is the domain of a function $f$.

- We call $c \in \mathbb{R}$ a boundary point of $X$ if any open interval $I$ containing $c$ intersects both $X$ and $\mathbb{R} \setminus X$.
- We call $c \in X$ an isolated point of $X$ if there exists an open interval $I$ containing $c$ such that $I \cap X = \{c\}$.

**Definition 2.2.5** (Continuity at boundary and isolated points). Suppose $f$ is a function with domain $X$.

- If $c$ is a boundary point of $X$, then we say that $f$ is continuous at $c$ if either $\lim_{x\to c^-} f(x)$ or $\lim_{x\to c^+} f(x)$ exists and is equal to $f(c)$ where the choice of one-sided limit depends on the side of $c$ on which $f$ is defined.
- By definition $f$ is continuous at an isolated point.

**Definition 2.2.6** (Continuity of a function). We can say that $f$ is continuous on a set $S \subseteq \text{dom } f$ if it is continuous at every point in $S$, or simply $f$ is continuous if it is continuous at every point of its domain.

**Types of discontinuities:**

- Removable discontinuity
- Jump discontinuity
- Infinite discontinuity
- Oscillating discontinuity

**Properties of continuous functions:**
Suppose $f$ and $g$ are two functions, both of which are continuous at $c$, and $k$ be a constant, Then the following functions are also continuous at $c$:

- $f \pm g$
- $kf$
- $fg$
- $\frac{f}{g}$ provided that $g \neq 0$
- $f^k$ provided that it is defined on an interval containing $c$

It follows that all polynomials, rational functions, exponential, logarithmic, and trigonometric functions are continuous.

**Theorem 2.2.3** (Continuity of composite functions). *Suppose* $\lim_{x\to c} f(x) = L$. *If* $g$ *is a function which is continuous at* $L$, *then*

$$\lim_{x\to c} g(f(x)) = g(\lim_{x\to c} f(x)) = g(L).$$

*If* $f$ *is continuous at* $c$, *then* $f(c) = L$ *and the above result can be rephrased as if* $f$ *is continuous at* $c$ *and* $g$ *is continuous at* $f(c)$, *then* $g \circ f$ *is continuous at* $c$.

**Theorem 2.2.4** (Extreme value theorem). *Suppose that the function* $f$ *is continuous on the interval* $[a, b]$. *Then* $f$ *attains a maximum and a minimum value in this interval.*

**Theorem 2.2.5** (Intermediate value theorem). *Suppose that the function* $f$ *is continuous on the interval* $[a, b]$. *Suppose* $f(a) \neq f(b)$ *and* $y$ *is a real number lying strictly between* $f(a)$ *and* $f(b)$ *(i.e. either* $f(a) \leq y \leq f(b)$ *or* $f(a) \geq y \geq f(b)$*). Then there exists* $c \in (a, b)$ *such that* $f(c) = y$.

## 2.3 Chapter 3: Differentiation

### 2.3.1 Derivatives

**Definition 2.3.1** (Derivative)**.** A function $f$ is said to be differentiable at $c$ if the limit

$$\lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

exists. The value of the limit is know as the derivative of $f$ at $c$, denoted by $f'(c)$.
The derivative of a function is a function as well, given by

$$f'(x) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$

wherever it is well-defined.

*Remark.* $f'(x)$ is a function, hence it will have its own domain and codomain, which are not necessary to be the same as $f(x)$.

**Theorem 2.3.1** (Differentiable and Continuous)**.** *If $f$ is differentiable at $c$, then it must be continuous at $c$. Consider the contrapositive statement of this statement, we can conclude that if a function $f$ is not continuous at $c$, then it must be not differentiable at $c$.*

*Remark.* For a conditional statement, we have hypothesis and conclusion. In this statement, $f$ is differentiable at $c$ is the hypothesis, and $f$ is continuous at $c$ is the conclusion. The contrapositive of a conditional statement both swaps the hypothesis and the conclusion and negates both the hypothesis and the conclusion.
If a conditional statement is true, then its contrapositive must be true. If a conditional statement is false, then its contrapositive must be false. This can be shown by truth table.
However, we need to notice that the converse of this statement may not be true. If $f$ is continuous at $c$, $f$ may not be differentiable at $c$. A trivial example is $f(x) = |x|$ is continuous at $x = 0$, but is not differentiable at $x = 0$.

**Definition 2.3.2** (Higher order derivatives)**.** Recall that the derivative of a function $f$ is also a function, when we take the derivative of the derivative of $f$, we will have the second order derivative. Similarly, we will have higher order derivatives if we continue taking derivatives.
Remember to check the differentiability of the function.

**Definition 2.3.3** (Implicit Differentiation)**.** When we face some difficulties to express $y$ in terms of $x$, we can use the technique of implicit differentiation.
We differentiate both sides of the equality $f(x, y) = 0$ with respect to $x$ to find $\dfrac{dy}{dx}$.

**Derivatives of Trigonometric Functions:**

- $\dfrac{d}{dx}(\sin x) = \cos x$
- $\dfrac{d}{dx}(\cos x) = -\sin x$
- $\dfrac{d}{dx}(\tan x) = \sec^2 x$
- $\dfrac{d}{dx}(\cot x) = -\csc^2 x$
- $\dfrac{d}{dx}(\sec x) = \sec x \tan x$
- $\dfrac{d}{dx}(\csc x) = -\csc x \cot x$

**Theorem 2.3.2** (Derivative of Inverse Function)**.** *Suppose $f$ has an inverse function $g$. If $f$ is differentiable at $(c, d)$ with $f'(c) \neq 0$, then $g$ is differentiable at $d$ and*

$$g'(d) = g'(f(c)) = \frac{1}{f'(c)}.$$

**Theorem 2.3.3** (Recall the Chain Rule)**.** *Suppose $h(x) = f(g(x))$ for every $x$, that is, function $h$ is the composition function of $f$ and $g$, then the derivative of $h(x)$ is*

$$h'(x) = (f(g(x)))' = f'(g(x))g'(x).$$

**Derivatives of Inverse Trigonometric Functions:**

- $\dfrac{d}{dx}(\sin^{-1} x) = \dfrac{1}{\sqrt{1-x^2}}$
- $\dfrac{d}{dx}(\cos^{-1} x) = -\dfrac{1}{\sqrt{1-x^2}}$
- $\dfrac{d}{dx}(\tan^{-1} x) = \dfrac{1}{1+x^2}$

### 2.3.2 Monotonicity and Concavity

**Definition 2.3.4** (Derivatives and Monotonicity)**.** If $f'(x) \geq 0$ for all $x \in (a, b)$, then $f$ is increasing in $(a, b)$; If $f'(x) \leq 0$ for all $x \in (a, b)$, then $f$ is decreasing in $(a, b)$.

**Definition 2.3.5** (Stationary Points and Extrema)**.** Stationary points are points where $f'(x) = 0$ and are potential candidates for extremum points.

*Remark.* A global extremum must be a local extremum, while the converse may not be true.
Stationary points are only the potential candidates for extremum points. A stationary point may not be an extremum point. Consider the point $(0, 0)$ on the function $f(x) = x^3$. It is a stationary point but not an extremum point.

**Definition 2.3.6** (Critical Points and Extrema)**.** We define critical points of a function $f$ to be points at which $f'(x)$ is either zero or undefined.

**Theorem 2.3.4** (Extremum)**.** *Let $f$ be a continuous function on $[a, b]$. Then the maximum and minimum of $f$ occur at either critical points or boundary points.*

**Definition 2.3.7** (Convex and Concave)**.** A function is said to be convex in an interval $I$ if

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \text{ for all } x_1, x_2 \in I \text{ and } t \in (0, 1).$$

A function is said to be concave in an interval $I$ if

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2) \text{ for all } x_1, x_2 \in I \text{ and } t \in (0, 1).$$

*Remark.* If a function $f$ is twice differentiable on an interval $I$, then being convex (concave) on $I$ is equivalent to $f''(x) \geq 0$ ($f''(x) \leq 0$) for all $x \in I$.
A point of inflection of a function $f$ is a point at which $f''(x)$ changes sign. For nice functions, this must be a point at which $f''(x)$ is either zero or undefined.

### 2.3.3 Asymptotes

**Definition 2.3.8** (Asymptotes)**.** The graph of $y = f(x)$ has a vertical asymptote $x = c$ if either

$$\lim_{x \to c^-} f(x) = \pm\infty \text{ or } \lim_{x \to c^+} f(x) = \pm\infty.$$

The graph of $y = f(x)$ has a horizontal asymptote $y = d$ if either

$$\lim_{x \to \infty} f(x) = d \text{ or } \lim_{x \to -\infty} f(x) = d.$$

The graph of $y = f(x)$ has an oblique asymptote $y = ax + b$ if either

$$\lim_{x \to \infty} [f(x) - (ax + b)] = 0 \text{ or } \lim_{x \to -\infty} [f(x) - (ax + b)] = 0.$$

How to find the value of $a$ and $b$ of an oblique asymptote $y = ax + b$?
Compute the values of $a$ and $b$ by evaluating the following limits:

$$a = \lim_{x\to\infty} \frac{f(x)}{x} \text{ and } b = \lim_{x\to\infty}[f(x) - ax].$$

and then replace $\infty$ by $-\infty$ and repeat the calculations.

**Theorem 2.3.5** (Rolle's Theorem)**.** *Suppose the function $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$. If $f(a) = f(b)$, then*

$$f'(c) = 0 \textit{ for some } c \in (a, b).$$

**Theorem 2.3.6** (Mean Value Theorem)**.** *Suppose the function $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$. Then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

**Theorem 2.3.7** (L'Hôpital's rule)**.** *Let $f$ and $g$ be differentiable functions such that either*

$$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0 \textit{ or } \lim_{x\to c} f(x) = \lim_{x\to c} g(x) = \pm\infty.$$

*Then*

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$$

*provided that the latter limit is finite or is $\pm\infty$.*

*Remark.* The same result holds if $x \to c$ is replaced by a ond-sided limit or a limit at infinity.

## 2.4 Chapter 4: Integration

### 2.4.1 Integration Techniques

In this chapter, we mainly focused on the calculations of integration of different functions. We also learned several techniques which are normally applied to commonly-seen functions.

- Integration by substitution
- Integration of trigonometric functions
- Integration by partial fraction decomposition
- Integration by parts (reduction formula)

**Integration by substitution**
Recall the chain rule in derivatives, for differentiable functions $F$ and $g$, we have

$$\frac{d}{dx}[F(g(x))] = F'(g(x)) \cdot g'(x).$$

Then we have

$$\int F'(g(x)) \cdot g'(x)\,dx = F(g(x)) + C.$$

If $f(u) = F'(u)$, then we have

$$\int f(g(x)) \cdot g'(x)\,dx = F(g(x)) + C.$$

**Generalized Result for Some Integrations:**

$$\int \frac{1}{ax + b} = \frac{1}{a} \ln|ax + b| + C$$

$$\int \frac{dx}{(x + a)^n} = \begin{cases} -\dfrac{1}{n-1} \cdot \dfrac{1}{(x + a)^{n-1}} + C & \text{if } n \in \mathbb{N} \setminus \{1\} \\ \ln|x + a| + C & \text{if } n = 1 \end{cases}$$

$$\int \frac{x}{(x^2+a^2)^n}\,dx = \begin{cases} \frac{1}{2}\ln(x^2+a^2)+C & \text{if } n=1 \\ \frac{1}{2(1-n)}\cdot\frac{1}{(x^2+a^2)^{n-1}} & \text{if } n\in\mathbb{N}\setminus\{1\} \end{cases}$$

**Integrating Trigonometric Functions**
Basic formulas:

- $\int \sin x\,dx = -\cos x + C$
- $\int \cos x\,dx = \sin x + C$
- $\int \sec^2 x\,dx = \tan x + C$

**Commonly used trigonometric techniques:**

- $\sin^2 x + \cos^2 x = 1$
- Double/Triple angle formula
- Half angle formula
- Sum-to-product and Product-to-sum formula

**Integration by Partial Fraction Decomposition**

**Definition 2.4.1** (Partial Fraction)**.** A partial fraction is a fraction of one of the following two forms:

$$\frac{p}{(x-k)^n} \text{ or } \frac{qx+r}{(ax^2+bx+c)^n}$$

where $a, b, c, k, p, q, r$ are constants with $b^2-4ac<0$ and $n$ is a positive integer.

**Definition 2.4.2** (Partial Fraction Decomposition)**.** Given a proper rational function $f(x)$ with real coefficients, i.e.

$$f(x) = \frac{p(x)}{q(x)}$$

where $p(x)$ and $q(x)$ are polynomials with real coefficients and $\deg p(x) < \deg q(x)$, a partial fraction decomposition of $f(x)$ means rewriting $f(x)$ as a sum of partial fractions.

*Remark.* Notice that every proper rational function with real coefficients can be expressed as a sum of partial fractions. Just take it for granted now, this can be proved by some algebra techniques.

**Method of Partial Fraction Decomposition:**

$$\frac{5x-4}{x^2-x-2} = \frac{A}{x+1} + \frac{B}{x-2}$$
$$5x-4 \equiv A(x-2)+B(x+1)$$

Method 1 - Plugging in appropriate values of $x$

- We plug in $x=2$, then we will have
$$6 = 3B$$
$$B = 2$$
- We plug in $x=-1$, then we will have
$$-9 = -3A$$
$$A = 3$$

Method 2 - Comparing coefficients

$$A(x-2)+B(x+1)=(A+B)x+(-2A+B)\equiv 5x-4$$

$$\begin{aligned} A+B&=5\\ -2A+B&=-4 \end{aligned}$$

Solve the linear equations system, we have $A=3$ and $B=2$.

**Method of Partial Fraction Decomposition - Repeated Factor**

$$\frac{12x^2+1}{4x^3-4x^2+x}=\frac{A}{x}+\frac{B}{2x-1}+\frac{C}{(2x-1)^2}$$

**Integration by Parts**

We have

$$\int u(x)v'(x)\,dx=u(x)v(x)-\int u'(x)v(x)\,dx$$

or in more compact notation,

$$\int u\,dv=uv-\int v\,du$$

Applying to definite integrals

$$\int_a^b u\,dv=u(b)v(b)-u(a)v(a)-\int_x^b v\,du$$

**Integration by Parts - Reduction Formula**

**Summary**

| | | |
|---|---|---|
| **Integration by substitution** | Not easy to integrate directly | $\int \frac{1}{1-2x}\,dx \int 2x\sqrt{1+x^2}\,dx$ |
| **Integration of trigonometric functions** | See some trigonometric function terms or some terms can be substituted by trigonometric functions | $\int \cos^2 3x\,dx \int \frac{dx}{x^2+4}$ |
| **Integration by partial fraction decomposition** | Integrator is a proper rational function | $\int \frac{12x^2+1}{4x^3-4x^2+x}\,dx \int \frac{5x-4}{x^2-x-2}\,dx$ |
| **Integration by parts** | Integrator is a product of some commonly-seen functions | $\int x\sec^2 x\,dx \int \sin(\ln x)\,dx$ |

## 2.4.2 Improper Integrals

**Definition 2.4.3** (Improper Integrals)**.** Usually, a definite integral can be interpreted as the area of a region that is normally bounded. The integral becomes an improper integral if it represents the area of a region that is unbounded. There are generally two types of unboundedness: Horizontally unbounded and vertically unbounded.

**Definition 2.4.4** (Horizontal Unboundedness)**.** If a function $f$ is continous on $[a,\infty)$, we define

$$\int_a^\infty f(x)\,dx=\lim_{b\to\infty}\int_a^b f(x)\,dx$$

provided that the limit on the right hand side exists. In this case, we say that the improper integral on the left hand side is convergent.

If the limit on the right hand side does not exist, then we say that the improper integral on the left hand side is divergent.

If $f$ is continuous on $(-\infty,b]$, we define

$$\int_{-\infty}^b f(x)\,dx=\lim_{a\to-\infty}\int_a^b f(x)\,dx.$$

**Definition 2.4.5** (Horizontal Unboundedness)**.** If the function $f$ is continuous on $\mathbb{R}$. then we define

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{\infty} f(x)\,dx$$

if we can find a real number $c$ for which both improper integrals on the right hand side converge.

**Definition 2.4.6** (Vertical Unboundedness)**.**

- If the function $f$ is continuous on $[a,b)$ and $x=b$ is a vertical asymptote of $f$, we define

$$\int_{a}^{b} f(X)\,dx = \lim_{c\to b}\int_{a}^{c} f(x)\,dx$$

- If $f$ is continuous on $(a,b]$ and $x=a$ is a vertical asymptote of $f$, we define

$$\int_{a}^{b} f(X)\,dx = \lim_{c\to a}\int_{a}^{b} f(x)\,dx$$

- If $f$ is continuous on $[a,b]$ except that it has a vertical asymptote at $x=c$ where $c\in(a,b)$. Then we define

$$\int_{a}^{b} f(x)\,dx = \int_{a}^{c} f(x)\,dx + \int_{c}^{b} f(x)\,dx$$

### 2.4.3 Applications

**Application in Probability Theory**

Properties of Probability Density Function

- $f(x)\geq 0$ for any $x$
- $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ for any probability density function $f(x)$ no matter for discrete random variable or continuous random variable
- We use the area between $f(x)$ and the x-axis to denote the probability of the occurrence of an event.
- For the probability density function $f(x)$ for a continuous random variable, when we input some value $x=a$, the output of the function $f(x)=f(a)$ is not the probability when $x=a$, the probability at every single point is always 0 for a continuous random variable.

**Definition 2.4.7** (Expected Value)**.** We compute the expected value of a random variable $X$ by $E[X]=\displaystyle\int_{-\infty}^{\infty} xf(x)\,dx$.

**Application of Definite Integral - Disk Method and Shell Method**

**Definition 2.4.8** (Disk Method)**.** If the graph of $y=f(x)$ in the interval $[a,b]$ is revolved about the $x$-axis, the volume of the solid of revolution obtained is given by

$$\int_{a}^{b} \pi[f(x)]^2\,dx.$$

**Definition 2.4.9** (Shell Method)**.** If the graph of $y=f(x)$ in the interval $[a,b]$ is revolved about the $y$-axis, the volume of the solid of revolution obtained is given by

$$\int_{a}^{b} 2\pi x[f(x)]\,dx.$$

**Definition 2.4.10** (Arc Length)**.** If $f'$ is continuous on $[a,b]$, the arc length of the curve $y=f(x)$ between $x=a$ and $x=b$ is given by

$$\int_{a}^{b} \sqrt{1+[f'(x)]^2}\,dx.$$

**Definition 2.4.11** (Surface Area)**.** Suppose $f'$ is continuous on $[a, b]$. When the graph of $y = f(x)$ between $x = a$ and $x = b$ is revolved about the $x$-axis, the area of the surface generated is given by

$$\int_a^b 2\pi f(x)\sqrt{1 + [f'(x)]^2}\, dx.$$

**Integration of Odd and Even Functions**
For an odd function $f$ and $a > 0$,

$$\int_{-a}^{a} f(x)\, dx = 0.$$

For an even function $f$ and $a > 0$,

$$\int_{-a}^{a} f(x)\, dx = 2\int_0^a f(x)\, dx.$$

**Fundamental Theorem of Calculus**

- If the function $f$ and $F$ are continuous on $[a, b]$ such that $F'(x) = f(x)$ for all $x \in (a, b)$, then

$$\int_a^b f(x)dx = F(b) - F(a).$$

- Let $f$ be a continuous function on $[a, b]$ and $c \in [a, b]$. Define the function $F : [a, b] \to \mathbb{R}$ by

$$F(x) = \int_c^x f(x)\, dt.$$

  Then $F$ is continuous on $[a, b]$ and $F'(x) = f(x)$ for all $x \in (a, b)$.

## 2.5 Chapter 5: Numerical Methods

### 2.5.1 Method of Bisection

**Steps of the method of bisection**
Suppose we want to solve the equation $f(x) = 0$ where $f$ is continuous. We do the following steps until we have obtained an interval containing the root that is within a desired error tolerance.

- Identify an interval $[a, b]$ for which $f(a)f(b) < 0$.
- Split the interval into two equal subintervals $[a, c]$ or $[c, b]$.
- Determine whether the root lies in $(a, c)$ or $(c, b)$.
- Repeat the above until the length of the interval that contains the root is within the error tolerance.

*Remark.*

- The rationale of the method of bisection is the Intermediate Value Theorem. Recall the theorem here: Suppose that the function $f$ is continuous on the interval $[a, b]$. Suppose $f(a) \neq f(b)$ and $y$ is a real number lying strictly between $f(a)$ and $f(b)$. Then there exists $c \in (a, b)$ such that $f(c) = y$.
- Method of bisection is a way of finding an interval containing a root. When applying this method, a higher error tolerance means less cost in calculation, indicating that only a few repetitions can achieve the desired interval. If we are pursuing a more accurate interval (a lower error tolerance), then we need to repeat more times with a higher calculation cost. This is commonly seen in approximations and other numerical methods.
- The bisection method can only be applied on a continuous function or on an interval that the function is continuous. We can consider the function $f(x) = \frac{1}{x}$, the function is not continuous on the interval $[-1, 2]$. If we try to use the method of bisection to find a root in this interval, no matter how many times we attempt, we will eventually fail since there is no root within this interval but every step we can find the desired subintervals. On the other hand, by the first remark, we know that the rationale of the bisection method is the I.V.T. and continuity is also a condition of I.V.T..

### 2.5.2 Taylor Approximation

**Series expansion of functions** Let $f$ be a function that has continuous derivatives of all orders on an interval $I$ containing $a$. The $n$-th Taylor polynomial of $f$ at $a$ is the polynomial

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

whereas the power series

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots,$$

is called the Taylor series of $f$ at $a$.
A Taylor series at 0 is sometimes called a Maclaurin series.

| | |
|---|---|
| **Taylor series** | $f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$ |
| **$n$-th Taylor polynomial** | $f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$ |
| **Maclaurin series** | $f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$ |

**Theorem 2.5.1** (Error analysis)**.** *Let $f$ be a function that has continuous derivatives of all orders on an interval $I$ containing $a$, and let $P_n(x)$ be the $n$-th Taylor polynomial of $f$ at $a$. Then*

$$f(x) - P_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

*where $c$ is a real number between $x$ and $a$.*

How to find the $c$ to determine the upper bound of the error?
Look at the error term $\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$, $c$ only appears in the $(n+1)$-th derivative. We can pay attention to the monotonic trends of $f^{(n+1)}(x)$ between $x$ and $a$.

### 2.5.3 Newton's Method

#### Procedure

1. We begin with an initial guess $x_0$ as the root of function $f(x)$. Set $n = 0$.
2. Replace $x_n$ by $x_n - \frac{f(x_n)}{f'(x_n)}$ and increase $n$ by 1.
3. Repeat the above and stop when the stopping condition met.

#### Stopping condition

1. The difference between $x_n$ and $x_{n+1}$ is small enough, i.e. $|x_{n+1} - x_n|$ is small enough.
2. The value of $f(x_n)$ is close enough to 0, i.e. $|f(x_n)|$ is small enough.

Newton's method does not always work.
Refer to the Assignment 5 Question 3.
Possible obstacles:

- The derivative may be difficult to compute.
- The formula fails if $f'(x_n)$ is undefined for some $n$.
- The formula fails if $f'(x_n)$ is equal to 0 for some $n$.
- The iterates may not be convergent.

### 2.5.4 Trapezoidal Rule

Suppose $f$ is continuous on $[a, b]$ and we divide $[a, b]$ into $n$ equal subintervals with endpoints $x_0, x_1, \cdots, x_n$ (with $x_0 = a$ and $x_n = b$). Then

$$\int_a^b f(x)\,dx \approx \frac{h}{2}\{f(x_0) + 2[f(x_1) + f(x_2) + \cdots + f(x_{n-1})] + f(x_n)\}$$

where $h = \frac{1}{n}(b-a)$ is the width of each subinterval.
In general, $\int_a^b f(x)\,dx$ is overestimated/underestimated if $f$ is convex/concave on the interval $[a, b]$.
Let $f$ be twice differentiable on $[a, b]$ and $K$ be a real number such that

$$|f''(x)| \leq K \text{ for all } x \in [a, b].$$

If we approximate $\displaystyle\int_a^b f(x)\,dx$ by the trapezoidal rule with $n$ subintervals, the error is at most

$$\frac{(b-a)^3 K}{12n^2}.$$

### 2.5.5 Simpson's Rule

Suppose $f$ is continuous on $[a, b]$ and we divide $[a, b]$ into $2m$ equal subintervals with endpoints $x_0, x_1, \cdots, x_{2m}$ (with $x_0 = a$ and $x_{2m} = b$). Then

$$\begin{aligned}\int_a^b f(x)\,dx \approx \frac{h}{3}\{&f(x_0) + 4[f(x_1) + f(x_3) + \cdots + f(x_{2m-1})] \\ &+ 2[f(x_2) + f(x_4) + \cdots + f(x_{2m-2})] + f(x_{2m})\}\end{aligned}$$

where $h = \frac{1}{2m}(b-a)$ is the width of each subinterval.
Let $f$ be a function for which $f^{(4)}$ exists and is continuous on $[a, b]$, and $K$ be a real number such that

$$|f^{(4)}(x)| \leq K \text{ for all } x \in [a, b].$$

If we approximate $\displaystyle\int_a^b f(x)\,dx$ by the Simpson's rule with $2m$ subintervals, the error is at most

$$\frac{(b-a)^5 K}{180(2m)^4}.$$

### 2.5.6 Summary of Numerical Methods

| | |
|---|---|
| **Method of Bisection** | Find an interval that contains a root of a continuous function |
| **Taylor Approximation** | Find the approximated value of the function at some point |
| **Newton's Method** | Find an approximated root given some initial guess |
| **Trapezoidal Rule** | Find the approximation of an integration using the trapezoids |
| **Simpson's Rule** | Find the approximation of an integration using the interpolation in every subintervals |

## 2.6 Chapter 6: Differential Equations

### 2.6.1 Basics

If the unknown function is a function of one variable, the resulting DE is called an ordinary differential equation (ODE). If the unknown function is a function of two or more variables, the resulting DE is called a partial differential equation (PDE). More details will be elaborated in course MATH3405 and MATH4406 respectively. **Order of DE** The order of a differential equation is the highest order derivative that appears in the equation.
**Classifications of ODEs** An ODE is said to be linear if it is of the form

$$F_n(x)y^{(n)} + \cdots + F_2(x)y'' + F_1(x)y' + F_0(x)y = G(x).$$

And the ODE is said to be

- homogeneous if $G(x) = 0$;
- with constant coefficients if $F_k(x)$ is a constant for all $k$.

**Initial Value Problem (IVP)** The problem in solving ODE with appropriate initial condition(s) provided. Usually, the unknown function can be found without any constant and is uniquely determined.

### 2.6.2 First Order ODEs

**Definition 2.6.1** (Separable First Order ODEs)**.** A first order ODE is said to be seperable if it can be written in the form

$$M(x) = N(y)\frac{dy}{dx}$$

**General solution to the separable first order ODE**
The solution is given by

$$\int M(x)\,dx = \int N(y)\,dy$$

**Definition 2.6.2** (Homogeneous of degree $n$)**.** Let $f(x, y)$ be a function of two variables. We say that $f$ is homogeneous of degree $n$ if

$$f(\lambda x, \lambda y) = \lambda^n f(x, y) \text{ for all } x, y.$$

Intuitively, this means each term of $f(x, y)$ is of degree $n$.

**Transforming into separable first order ODEs**
If a first order ODE is of the form

$$M(x, y) = N(x, y)\frac{dy}{dx}$$

where both $M(x, y)$ and $N(x, y)$ are homogeneous of the same degree, then we may apply the substitution $y = vx$ to transform the ODE to a separable one.
Here, $v := v(x)$ is a function of $x$.
**Integrating factor**
We can solve a general linear first order ODE of the form

$$y' + P(x)y = Q(x)$$

by multiplying both sides by the integrating factor $e^{\int P(x)\,dx}$.

### 2.6.3 Second Order ODEs

We mainly discuss the solution of linear second order ODEs with constant coefficients, i.e. an ODE of the form

$$y'' + py' + qy = r(x),$$

where $p$ and $q$ are real numbers, and $r(x)$ is a function of $x$.
**Initial Conditions and Uniqueness of the Solution**
For a second order ODE, we usually need two initial conditions to determine a unique solution, and the initial conditions are usually in the form of

$$y(x_0) = y_1 \text{ and } y'(x_0) = y_2$$

where $x_0, y_1, y_2$ are real numbers.

**Homogeneous second order ODEs**

For a second order ODE in the form of

$$y'' + py' + qy = r(x),$$

it is homogeneous when $r(x) = 0$.

**Definition 2.6.3** (Characteristic Equation)**.** The function $y = e^{kx}$ satisfies the ODE

$$y'' + py' + q = 0$$

if and only if $k$ is a root of the equation

$$\lambda^2 + p\lambda + q = 0,$$

which is called the characteristic equation associated with the ODE.

**Distinct real roots**
Consider the ODE

$$y'' + py' + qy = 0.$$

If the characteristic equation has two distinct real roots $m$ and $n$, then the ODE has general solution

$$y = Ae^{mx} + Be^{nx}.$$

**Double root (repeated root)**
Consider the ODE

$$y'' + py' + qy = 0.$$

If the characteristic equation has a double real roots $m$ then the ODE has general solution

$$y = Ae^{mx} + Bxe^{mx}.$$

**Complex roots**
Consider the ODE

$$y'' + py' + qy = 0.$$

If the characteristic equation has two complex roots $m \pm ni$, then the ODE has general solution

$$y = e^{mx}(A\cos nx + B\sin nx).$$

**Non-homogeneous second order ODEs**

Every solution to the ODE

$$y'' + py' + qy = r(x)$$

is of the form $y = y_c(x) + y_p(x)$, where

- $y_c(x)$ satisfies the associated homogeneous part $y'' + py' + qy = 0$, called the homogeneous solution of complementary solution;
- $y_p(x)$ is one solution to the original ODE, called a particular solution.

| **Form of $r(x)$** | **Try for particular solutions of the form** |
|---|---|
| $p(x)$ | $a_n x^n + \cdots + a_1 x + a_0$ |
| $e^{kx}$ | $a_0 e^{kx}$ |
| $\sinh x$ | $a_0 \sinh x + b_0 \cosh x$ |
| $e^{kx}p(x)$ | $e^{kx}(a_n x^n + \cdots + a_1 x + a_)$ |

*Remark.*

- If $r(x)$ is a linear combination of the forms listed above, we may try a linear combination of the corresponding suggested forms of the particular solutions.
- Other forms like $\cosh x$ or $p(x)\sinh x$ can be handled in similar ways.

**Steps for solving a non-homogeneous ODE with two initial conditions**

1. Solve the homogeneous solution with two different constants.
2. Guess the particular solution based on the form of $r(x)$ with some constants.
3. Evaluate whether the particular solution is in the same form of the homogeneous solution. If yes, multiply $x$ on the particular solution; if no, keep the particular solution for the next step.

4. Plug the particular solution into the non-homogeneous ODE and evaluate every unknown constant in the particular solution. (After plugging in, we should be able to evaluate all the constants in the particular solution.)

5. Add the homogeneous solution and the particular solution together to get the general solution. Use the two initial conditions to evaluate the two constants in the general solution.

## 2.7 Chapter 7: Matrices

### 2.7.1 Basic Terminology and Concepts

- row
- column
- size (dimension)
- $(i, j)$-entry
- submatrix
- equal matrices (size and every entry should be equal)
- square matrix
- main diagonal (collection of all diagonal entries)
- diagonal matrix (all off-diagonal entries are 0)
- triangular matrix
  - upper triangular matrix ($p_{ij} = 0$ whenever $i > j$)
  - lower triangular matrix ($p_{ij} = 0$ whenever $i < j$)

*Remark.*

- The *diagonal entries* of diagonal matrix are not necessary to be non-zero. The zero matrix is also a diagonal matrix.
- A diagonal matrix is default to be a square matrix, but triangular matrices are not necessary to be square matrices.

### 2.7.2 Matrix Operation

**Addition**

- defined entrywise and only defined on same size matrices;
- commutative;
- associative.

**Definition 2.7.1.** The **additive identity** $O$ is the matrix satisfying $A + O = A$ for any matrix $A$, and $O$ should be in the same size of $A$.
By the definition we know that the additive identity is indeed the zero matrix with same size.
The **additive inverse** of matrix $A = [a_{ij}]$ is defined by $-A = [-a_{ij}]$ such that $A + (-A) = O$.

*Remark.* We can define the matrix subtraction by considering $A - B$ as $A + (-B)$.

#### Scalar Multiplication

**Definition 2.7.2.** Let $A$ be a matrix and $k$ be a scalar (which can be taken as a real number or a constant number in this course). Then $kA$ is the matrix obtained by multiplying every entry of $A$ by $k$.

#### Distributive laws between addition and scalar multiplication

- $s(A+B) = sA + sB$
- $(s+t)A = sA + tA$
- $(st)A = s(tA)$

#### Transpose

**Definition 2.7.3.** Let $A = [a_{ij}]$ be a matrix. The **transpose** of $A$, denoted by $A^T$, is a matrix whose $(i,j)$-entry is the $(j,i)$-entry of $A$.
Remark: If $A$ is a $m \times n$ matrix, then $A^T$ is a $n \times m$ matrix.

**Definition 2.7.4.** A matrix $A$ is said to be

- symmetric if $A^T = A$;
- skew-symmetric if $A^T = -A$.

If $A, B$ are matrices with the same size and $s$ is a scalar, then we have

- $(A+B)^T = A^T + B^T$;
- $(sA)^T = s(A^T)$;
- $(A^T)^T = A$.

#### Matrix Multiplication

**Definition 2.7.5** (Dot Product)**.** Given two ordered $n$-tuples $(a_1, a_2, \cdots, a_n)$ and $(b_1, b_2, \cdots, b_n)$, we define their dot product to be

$$(a_1, a_2, \cdots, a_n) \cdot (b_1, b_2, \cdots, b_n) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

**Definition 2.7.6.** Let $A$ be an $m \times n$ matrix and $B$ be an $n \times k$ matrix. We define the product $AB$ to be the $m \times k$ matrix $C = [c_{ij}]$ whose $(i,j)$-entry is the dot product of the $i$-th row of $A$ and the $j$-th column of $B$, i.e.

$$c_{ij} = a_{i1}b_{ij} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

*Remark.* The matrix multiplication is not commutative. $BA$ may not be well defined as $AB$, and may have different size or unequal entry with $AB$.

Properties of matrix products
Let $A, B, C$ be matrices and $s$ be a scalar. We have the following properties whenever the operations are well-defined:

- $A(BC) = A(BC)$;
- $A(B+C) = AB + AC$;
- $(A+B)C = AC + BC$;
- $s(AB) = (sA)B = A(sB)$;
- $(AB)^T = B^TA^T$.

**Definition 2.7.7.** Let $A$ be a square matrix and $n$ be a positive integer. Naturally, we define

$$A^n = \underbrace{AA\cdots A}_{n \text{ times}}.$$

Alternative views of matrix multiplication
Suppose the matrix product $AB$ is well-defined.

- The $i$-th row of $AB$ is a linear combination of the rows of $B$, with the coefficients coming from the $i$-th row of $A$.
- The $j$-th column of $AB$ is a linear combination of the columns of $A$, with the coefficients coming from the $j$-th row of $B$.

Matrix Inverse

**Definition 2.7.8.** The identity matrix $I$ is a square matrix with all diagonal entries 1 and all other entries 0. We may write $I_n$ for identity matrix with size $n \times n$. The product of identity matrix with any other matrix returns to the matrix itself (remember to check the size to well-define the matrix multiplication).

**Definition 2.7.9.** Let $A$ be a square matrix. If $f(x)$ is a general polynomial, then we can naturally define $f(A)$ to be the matrix polynomial with same polynomial terms.

**Definition 2.7.10.** Let $A$ be a square matrix. We say that $B$ is an inverse of $A$ if $AB = BA = I$. Notice that the inverse of a matrix is unique.

*Remark.*

- The inverse of $A$ is usually denoted by $A^{-1}$.
- We can also define the left inverse and right inverse for a non-square matrix.
- $B$ should be called a multiplicative inverse of $A$, but we usually call it the inverse of $A$.

**Definition 2.7.11.** A matrix with inverse is said to be invertible (non-singular). A matrix without inverse is said to be singular.

**Theorem 2.7.1.** *Let $A$ and $B$ be square matrices with the same size. If $AB = I$, then $BA = I$.*

**Properties of matrix inverse**
Let $A, B, C$ be invertible matrices of the same size, $n$ be a positive integer and $s$ be a scalar. Then we have

- $(A^{-1})^{-1} = A$;
- $(AB)^{-1} = B^{-1}A^{-1}$;
- $(A^n)^{-1} = (A^{-1})^n$;
- $(A^T)^{-1} = (A^{-1})^T$;
- $(sA)^{-1} = \frac{1}{s}A^{-1}$

### 2.7.3 Determinants

**Definition 2.7.12.** For a $2 \times 2$ matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we define the determinant of $A$, denoted by $\det A$, by

$$\det A = ad - bc.$$

*Remark.*

- Alternative notations for the determinant include $|A|$ and $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.
- The determinant is a scalar.
- It is easy to check that if $\det A = 0$, then $A$ is not invertible.

**Definition 2.7.13.** For a $3 \times 3$ matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, we define

$$\det A = aei + bfg + cdh - ceg - afh - bdi.$$

*Remark.*

- Beware that $4 \times 4$ and higher-order determinants are defined in a different way.
- It is still true that if $\det A = 0$, then $A$ is not invertible.

**Definition 2.7.14.** Let $A = [a_{ij}]$ be a square matrix.

- $A_{ij}$ denotes the submatrix of $A$ obtained from $A$ by deleting row $i$ and column $j$, and is still a square matrix with smaller size.
- The $(i, j)$-cofactor of $A$, denoted by $c_{ij}(A)$ or simply $c_{ij}$, is defined by $(-1)^{i+j} \det A_{ij}$.

**Definition 2.7.15.** The cofactor matrix of $A$ is the matrix whose $(i, j)$-entry is the $(i, j)$-cofactor of $A$. The **transpose** of the cofactor matrix of $A$ is called the adjoint of $A$, denoted by adj$A$.

**Theorem 2.7.2.** *Let $A$ be a square matrix. If $\det A \neq 0$, then*

$$A^{-1} = \frac{1}{\det A}(adjA).$$

Properties of determinants Let $A$ and $B$ be square matrices with the same size. Then

1. $A$ is invertible if and only if $\det A \neq 0$;
2. $\det AB = \det A \det B$;
3. $\det A^T = \det A$;
4. $\det A^{-1} = \dfrac{1}{\det A}$ if $A$ is invertible.

## 2.8 Chapter 8: Vectors

### 2.8.1 Basics

**Definition 2.8.1.** A vector is a quantity with both **magnitude** and **direction**.

Notation: We usually draw an arrowed lined segment from an initial point $A$ to a terminal point $B$. So we can denote the vector by $\overrightarrow{AB}$. Or we can use $\boldsymbol{v}$ or simply $\overrightarrow{v}$ in handwriting.

**Definition 2.8.2.** The length of a vector $\boldsymbol{v}$ is called the norm of $\boldsymbol{v}$ and is denoted by $\|\boldsymbol{v}\|$.

- A vector with norm 1 is called a unit vector;
- A vector with norm 0 is called the zero vector and is denoted by $\mathbf{0}$.
- Two vectors are equal if they are same in both magnitude and direction, parallel if they have the same or opposite direction(s).

**Vector Operation**

**Scalar Multiplication**
Let $\boldsymbol{v}$ be a non-zero vector, for a positive constant $k$, the vector $k\boldsymbol{v}$ is the vector in the same direction and $k$ times as long as $\boldsymbol{v}$. If $k$ is negative, then $k\boldsymbol{v}$ is the vector in the opposite direction and $k$ times as long as $\boldsymbol{v}$. If $k$ is 0, then $k\boldsymbol{v} = \mathbf{0}$.
**Addition**
To add two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, we move the initial point of $\boldsymbol{v}$ to the terminal point of $\boldsymbol{u}$. The sum $\boldsymbol{u}+\boldsymbol{v}$ is the vector whose initial point is the initial point of $\boldsymbol{u}$ and whose terminal point is the terminal point of $\boldsymbol{v}$. (tip-to-tail method)
It is easy to see that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$, known as the parallelogram rule.
Naturally, we define the subtraction by $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-\boldsymbol{v})$.
**Properties of vector operations**
Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ be vectors and $s, t$ be scalars. We have:

- $(s+t)\boldsymbol{u} = s\boldsymbol{u} + t\boldsymbol{u}$;
- $(st)\boldsymbol{u} = s(t\boldsymbol{u})$;
- $\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$;
- $\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w}) = (\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}$;
- $s(\boldsymbol{u}+\boldsymbol{v}) = s\boldsymbol{u}+s\boldsymbol{v}$.

Algebraically, a vector may be defined as

- a matrix with only one row (row vector); or
- a matrix with only one column (column vector).

We usually regard a vector as a column vector, use $\mathbb{R}^n$ to denote the set of all such vectors with $n$ entries.

**Vectors in Coordinate System**

**Two-Dimensional**
Every vector $\begin{bmatrix} a \\ b \end{bmatrix}$ in $\mathbb{R}^2$ can be associated with the vector on the rectangular coordinate plane with initial point at the origin and terminal point at the point $P(a,b)$, and hence can be identified with the point $P$. Such a vector is also known as the position vector of $P$.
**Three-Dimensional**
Every vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathbb{R}^3$ can be associated with the vector in the three-dimensional rectangular coordinate system with initial point at the origin and terminal point at the point $P(a,b,c)$, and hence can be identified with the point $P$. This is known as the position vector of $P$. (the labelling obeys the right hand rule)
In $\mathbb{R}^2$, we usually use $\boldsymbol{i}$ and $\boldsymbol{j}$ to denote the unit vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ respectively.
In $\mathbb{R}^3$, we usually use $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ to denote the unit vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ respectively.
Every vector in $\mathbb{R}^2$ and $\mathbb{R}^3$ can be expressed as a linear combination of the standard unit vectors (also called coordinate vectors) defined above.

### 2.8.2 Dot Product

**Definition 2.8.3.** Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be two vectors. We define

$$\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta$$

where $\theta$ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. This is known as the dot product or scalar product of $\boldsymbol{u}$ and $\boldsymbol{v}$.

*Remark.*

- Two vectors are orthogonal (perpendicular) if and only if their dot product is zero.
- The zero vector is by definition orthogonal to any other vector.
- The dot product of $\boldsymbol{u}$ and $\boldsymbol{v}$ can be interpreted as the product of the lengths of $\boldsymbol{u}$ and $\boldsymbol{v}$ after one is projected onto the other.

**Properties of dot product** For any vector $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ and scalar $k$, we have:

- $\boldsymbol{u}\cdot\boldsymbol{v} = \boldsymbol{v}\cdot\boldsymbol{u}$;
- $\boldsymbol{u}\cdot(k\boldsymbol{v}) = (k\boldsymbol{u})\cdot\boldsymbol{v}$;
- $\boldsymbol{u}\cdot(\boldsymbol{v}+\boldsymbol{w}) = \boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{u}\cdot\boldsymbol{w}$;
- $\boldsymbol{u}\cdot\boldsymbol{u} = \|\boldsymbol{u}\|^2$.

For two vectors $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j}$ and $\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j}$ in $\mathbb{R}^2$, we have

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2.$$

For two vectors $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j} + u_3\boldsymbol{k}$ and $\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}$ in $\mathbb{R}^3$, we have

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

**Angle between two vectors**
For any two vectors $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)^T$ and $\boldsymbol{v} = (v_1, v_2, \cdots, v_n)^T$, we can define their dot product to be

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

We can define the angle $\theta$ between $\boldsymbol{u}$ and $\boldsymbol{v}$ by

$$\cos\theta = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}$$

where $\theta \in [0, \pi]$.

### 2.8.3 Cross Product

**Definition 2.8.4.** Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^3$, we define the cross product (vector product) $\boldsymbol{u} \times \boldsymbol{v}$ to be the vector $\boldsymbol{w}$ such that

- $\|\boldsymbol{w}\| = \|\boldsymbol{u}\|\|\boldsymbol{v}\| \sin\theta$, where $\theta$ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$;
- $\boldsymbol{w}$ is orthogonal to both $\boldsymbol{u}$ and $\boldsymbol{v}$ given by the right-hand rule.

**Properties 1**

- The cross product is a vector.
- We have $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{0}$ if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel, in which case the orientation does not matter.
- The cross product is skew-commutative, i.e.

$$\boldsymbol{u} \times \boldsymbol{v} = -(\boldsymbol{v} \times \boldsymbol{u}) \quad \text{for any } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3.$$

**Properties 2**
Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ be vectors in $\mathbb{R}^3$ and $s$ be a scalar. Then we have:

- $\boldsymbol{u} \times \boldsymbol{0} = \boldsymbol{0} \times \boldsymbol{u} = \boldsymbol{0}$;
- $\boldsymbol{u} \times \boldsymbol{u} = \boldsymbol{0}$;
- $\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w}$;
- $(\boldsymbol{u} + \boldsymbol{v}) \times \boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{w} + \boldsymbol{v} \times \boldsymbol{w}$;
- $s(\boldsymbol{u} + \boldsymbol{v}) = (s\boldsymbol{u}) \times \boldsymbol{v} = \boldsymbol{u} \times (s\boldsymbol{v})$.

Recall the unit vectors in the three-dimensional coordinate system. We can see that

$$\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}, \boldsymbol{i} \times \boldsymbol{k} = -\boldsymbol{j}, \boldsymbol{j} \times \boldsymbol{i} = -\boldsymbol{k}, \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}, \boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}, \boldsymbol{k} \times \boldsymbol{j} = -\boldsymbol{i}.$$

Given two vectors $\boldsymbol{u} = (u_1, u_2, u_3)^T$ and $\boldsymbol{v} = (v_1, v_2, v_3)^T$ in $\mathbb{R}^3$, we have

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

**Geometric meaning of cross product**

- Area of parallelogram ABCD $= \|\overrightarrow{AB} \times \overrightarrow{AC}\|$

- Area of $\triangle ABC = \frac{1}{2}\|\overrightarrow{AB} \times \overrightarrow{AC}\|$

**Scalar Triple Product**

**Definition 2.8.5.** Given three vectors $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ in $\mathbb{R}^3$, the quantity

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$$

is called the scalar triple product of these vectors.

Given three vectors $\boldsymbol{u} = (u_1, u_2, u_3)^T$, $\boldsymbol{v} = (v_1, v_2, v_3)^T$ and $\boldsymbol{w} = (w_1, w_2, w_3)^T$ in $\mathbb{R}^3$, their scalar triple product is given by

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \det \begin{bmatrix} \boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w} \end{bmatrix}.$$

The absolute value of $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ is the volume of the parallelepiped formed by these three vectors.
The scalar triple product $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ is equal to 0 if and only if these three vectors are coplanar.

**Properties**

Notice that $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$, the scalar triple product is sometimes written as

$$(\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}).$$

**Definition 2.8.6.** We also define the following vector triple product

$$\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}).$$

- $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}$
- $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) + \boldsymbol{v} \times (\boldsymbol{w} \times \boldsymbol{u}) + \boldsymbol{w} \times (\boldsymbol{u} \times \boldsymbol{v}) = \mathbf{0}$

# Chapter 3

# Practice Questions

## 3.1 Chapter 1: Basic Concepts

1. Consider a function $f : X \to Y$, where $X = \{1, 2, 3\}$, $Y = \{4, 5, 6\}$, and $f(1) = f(2) = 4$, $f(3) = 5$. $A = \{1, 3\}$, $B = \{5, 6\}$.

   (a) What are the domain, codomain, and range of $f$?

   (b) Is $f$ injective, surjective, and bijective?

   (c) Find $f(A)$ and $f^{-1}(f(A))$.

   (d) Find $f^{-1}(B)$ and $f(f^{-1}(B))$.

   (e) What is the relationship between $A$ and $f^{-1}(f(A))$, $B$ and $f(f^{-1}(B))$?

   (f) Consider a general function $f : X \to Y$, for what condition should $f$ satisfies so that for any $A \subseteq X$, we always have $A = f^{-1}(f(A))$?

   (g) Consider a general function $f : X \to Y$, for what condition should $f$ satisfies so that for any $B \subseteq Y$, we always have $B = f(f^{-1}(B))$?

   **Suggested solution:**

   (a) Domain: $\{1, 2, 3\}$, codomain: $\{4, 5, 6\}$, range: $\{4, 5\}$.

   (b) $f$ is not injective, not surjective, and not bijective.

   (c) $f(A) = \{4, 5\}$, $f^{-1}(f(A)) = \{1, 2, 3\}$.

   (d) $f^{-1}(B) = \{3\}$, $f(f^{-1}(B)) = \{5\}$.

   (e) $A \subseteq f^{-1}(f(A))$, $B \subseteq f(f^{-1}(B))$.

   (f) $f : X \to Y$ is injective if and only if $A = f^{-1}(f(A))$ for any $A \subseteq X$.

   (g) $f : X \to Y$ is surjective if and only if $B = f(f^{-1}(B))$ for any $B \subseteq Y$.

2. Express the following complex numbers in polar form:

   (a) $z = 1 + i$

   (b) $(i^{25})^3$

   (c) $z = -2 - 2\sqrt{3}i$

   **Suggested solution:**

   (a) $z = \sqrt{2}(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4})$

   (b) $z = \cos \frac{\pi}{2} - i \sin \frac{\pi}{2}$

   (c) $z = 4(\cos \frac{2\pi}{3} - i \sin \frac{2\pi}{3})$

3. Prove that $S(n): \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} \leq 2\sqrt{n} - 1$ is true for every positive integer $n$.

   **Suggested solution:**
   When $n = 1$, we have $\frac{1}{\sqrt{1}} = 1 \leq 2\sqrt{1} - 1 = 1$, hence $S(n)$ is true when $n = 1$.
   Let $k$ be an arbitrary integer such that $k \geq 1$. Assume that $S(k)$ is true.
   When $n = k + 1$,
   LHS $= (\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{k}}) + \frac{1}{\sqrt{k+1}}$
   $\leq 2\sqrt{k} - 1 + \frac{1}{\sqrt{k+1}}$
   $= 2\sqrt{k} - 1 + \frac{2}{2\sqrt{k+1}}$
   $\leq 2\sqrt{k} - 1 + \frac{2}{\sqrt{k+1} + \sqrt{k}}$
   $= 2\sqrt{k} - 1 + \frac{2(\sqrt{k+1} - \sqrt{k})}{(\sqrt{k+1} + \sqrt{k})(\sqrt{k+1} - \sqrt{k})}$
   $= 2\sqrt{k} - 1 + \frac{2(\sqrt{k+1} - \sqrt{k})}{(k+1) - k}$
   $= 2\sqrt{k} - 1 + 2(\sqrt{k+1} - \sqrt{k})$
   $= 2\sqrt{k+1} - 1.$
   Hence, $S(k+1)$ is true.
   By the principle of mathematical induction, $S(n)$ holds for every positive integer $n$.

4. Let $-\pi < \theta < \pi$. Show that $\theta = 2\arctan(\frac{\sin\theta}{1+\cos\theta})$.
   Hint: try to use $\sin\theta$ and $\cos\theta$ to express $\tan\frac{\theta}{2}$.
   **Suggested Solutions:**

   *Proof.* Consider $\tan\frac{\theta}{2} = \frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}} = \frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}} \cdot \frac{2\cos\frac{\theta}{2}}{2\cos\frac{\theta}{2}} = \frac{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{2\cos^2\frac{\theta}{2}}$.
   Apply the double angle formulas, we get $\tan\frac{\theta}{2} = \frac{\sin\theta}{1+\cos\theta}$
   Since $-\pi < \theta < \pi$, $-\frac{\pi}{2} < \frac{\theta}{2} < \frac{\pi}{2}$, $\theta = 2\arctan(\frac{\sin\theta}{1+\cos\theta})$. □

## 3.2 Chapter 2: Limits and Continuity

1. Let $f(x) = \frac{x^2-1}{x-1}$, find $\lim_{x\to 2} f(x)$, $\lim_{x\to 1} f(x)$.
   **Suggested solution:**
   Since $\lim_{x\to 2}(x-1) = 1 \neq 0$, by the algebraic properties of limit,

   $$\lim_{x\to 2} f(x) = \frac{\lim_{x\to 2}(x^2-1)}{\lim_{x\to 2}(x-1)} = 3.$$

   Since $\lim_{x\to 1}(x-1) = 0$, we cannot apply algebraic properties of limit immediately.
   Note that for $x$ close to 1, $f(x) = \frac{(x+1)(x-1)}{x-1} = x + 1$ provided that $x \neq 1$. It follows that

   $$\lim_{x\to 1} f(x) = \lim_{x\to 1}(x+1) = 2.$$

2. Consider

   $$f(x) = \begin{cases} \frac{x^2-1}{x-1} & \text{if } x < 1, \\ 2x & \text{if } x \geq 1. \end{cases}$$

Evaluate its limit at 1.
**Suggested solution:**
Note that $f(x)$ is given by different formulas on different sides of 1.
For any $x$ slightly greater than 1, $f(x)$ is given the formula $f(x) = 2x$.

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} 2x = 2.$$

For any $x$ slightly less than 1, $f(x)$ is given the formula $f(x) = \dfrac{x^2-1}{x-1}$.

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} (x+1) = 2.$$

Since

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^-} f(x) = 2$$

we arrive at

$$\lim_{x\to 1} f(x) = 2.$$

3. Let $f(x) = g(x)\ln(2 + \sin\ln x)$ where $g$ is a function whose domain contains $(0,1)$. If $\lim_{x\to 0^+} |g(x)| = 0$, then evaluate $\lim_{x\to 0^+} |f(x)|$.
   **Suggested solution:**
   Note that for any $x \in (0,1)$,

   $$-1 \le \sin\ln x \le 1$$

   $$1 \le 2 + \sin\ln x \le 3$$

   Since $\ln x$ is strictly increasing,

   $$0 = \ln 1 \le \ln(2 + \sin\ln x) \le \ln 3.$$

   As $|g(x)| \ge 0$, we have

   $$0 \le |f(x)| = |g(x)| \ln(2 + \sin\ln x) \le |g(x)| \ln 3.$$

   As $\lim\limits_{x\to 0^+} |g(x)| = 0$, $\lim\limits_{x\to 0^+} |f(x)| = 0$ by sandwich theorem.

4. Let

   $$f(x) = \begin{cases} \sqrt[3]{3x-1} & \text{if } x > 3 \\ \sin\frac{\pi x}{2} & \text{if } 1 < x \le 3 \\ e^{1-x} & \text{if } x < 1 \end{cases}$$

   Show that $f$ is discontinuous at $x = 3$.
   **Suggested solution:**
   We need to calculate the limit of $f$ at $x = 3$.
   Notice that $f$ has different formulas on left and right hand of $x = 3$, we need to calculate the one-sided limit at $x = 3$ separately.

   $$\lim_{x\to 3^-} f(x) = \lim_{x\to 3^-} \sin\frac{\pi x}{2} = \sin\frac{3\pi}{2} = -1$$

   $$\lim_{x\to 3^+} f(x) = \lim_{x\to 3^-} \sqrt[3]{3x-1} = \sqrt[3]{8} = 2$$

   Since $\lim\limits_{x\to 3^-} f(x) \ne \lim\limits_{x\to 3^+} f(x)$, $f$ is discontinuous at $x = 3$.

## 3.3 Chapter 3: Differentiation

1. Find the slope of the tangent line to the graph of $x^2 + y^2 = 9$ at point $(1, -2\sqrt{2})$.
   **Suggested solution:** $f(x,y) = x^2 + y^2 - 9 = 0$
   We take the derivative respect to $x$ (consider $y$ as a constant), hence we have

   $$\frac{df(x,y)}{dx} = \frac{d(x^2+y^2-9)}{dx} = 2x + 2y\frac{dy}{dx} = 0$$

which means

$$\frac{dy}{dx} = -\frac{x}{y} \text{ if } y \neq 0$$

At point $(1, -2\sqrt{2})$, the slope is $\dfrac{\sqrt{2}}{4}$.

2. Find $f'(x)$ of $f(x) = x^x$.
   **Suggested solution:**
   Consider $y = f(x) = x^x$. take logarithms respect on the both sides

$$\ln y = \ln x^x = x \ln x$$

   Then take the derivative respect to $x$ on both sides, we will have

$$\begin{aligned}\frac{dy}{dx}(\ln y) &= \frac{dy}{dx}(x \ln x)\\ \frac{1}{y}\frac{dy}{dx} &= \ln x + 1\\ \frac{dy}{dx} &= y(\ln x + 1) = x^x(\ln x + 1)\end{aligned}$$

3. Find the point(s) of inflection of function $f(x) = -\dfrac{1}{6}x^4 + 7x$.
   **Suggested solution:**
   Notice that $f''(x) \leq 0$ for any $x$, hence the sign of $f''(x)$ has no change over any intervals. $f(x)$ has no point of inflection,

   *Remark.* Some may argue that $f(0) = 0$ so that $(0, 0)$ is a point of inflection, however, please remember to check whether the sign of $f''(x)$ has changed at that point.

4. By applying the mean value theorem to the function $f(x) = \sqrt{x}$ over $[3, 4]$, prove that $12/7 < \sqrt{3} < 7/4$.
   **Suggested solution:**
   Since $f$ is continuous on $[3, 4]$ and is differentiable on $(3, 4)$, the mean value theorem asserts that there is $c \in (a, b)$ such that

$$2 - \sqrt{3} = \frac{f(4) - f(3)}{4 - 3} = f'(c) = \frac{1}{2\sqrt{c}}$$

   As $\sqrt{x}$ is increasing, we have

$$\begin{aligned}\sqrt{3} < \sqrt{c} < \sqrt{4} &= 2\\ \frac{1}{4} < \frac{1}{2\sqrt{c}} < \frac{1}{2\sqrt{3}} &= \frac{1}{6}\sqrt{3}\end{aligned}$$

   As a result,

$$\begin{gathered}\frac{1}{4} < 2 - \sqrt{3} < \frac{1}{6}\sqrt{3}\\ 2 < (1 + \frac{1}{6})\sqrt{3} \text{ and } \sqrt{3} < 2 - \frac{1}{4}\\ \frac{12}{7} < \sqrt{3} < \frac{7}{4}\end{gathered}$$

5. Given a curve with equation

$$x^2y + y^3 + x^2 = 8$$

   (a) Find $\dfrac{dy}{dx}$ by implicit differentiation.
   (b) Find all points on the curve where its tangent line is horizontal.
   (c) Find $\dfrac{d^2y}{dx^2}$ at the point(s) found in previous part.

   **Suggested solution:**

(a) Implicit differentiate the given equation

$$\begin{aligned}\frac{d}{dx}(x^2y+y^3+x^2) &= \frac{d}{dx}8\\ 2xy+x^2\frac{dy}{dx}+3y^2\frac{dy}{dx}+2x &= 0\\ (x^2+3y^2)\frac{dy}{dx} &= -2x(1+y)\\ \frac{dy}{dx} &= -\frac{2x(1+y)}{x^2+3y^2}\end{aligned}$$

(b) To find a horizontal tangent line, we proceed to solve

$$\frac{dy}{dx} = -\frac{2x(1+y)}{x^2+3y^2} = 0$$

and its solutions are $x = 0$ or $y = -1$. Substitute $x = 0$ into the original equation and we have $y = 2$. Substitute $y = -1$ into the original euqation and we have $-1 = 8$, which is impossible, indicating that there is no point with $y = -1$ on the curve. Therefore, $y = -1$ is rejected.
Hence $(0, 2)$ is the unique point on the curve at which the tangent line is horizontal, and the tangent line is $y = 2$.

(c) We implicit differentiate the equation we obtained in the first part to get

$$\begin{aligned}\frac{d}{dx}[(x^2+3y^2)\frac{dy}{dx}] &= \frac{d}{dx}[-2x(1+y)]\\ (x^2+3y^2)\frac{d^2y}{dx^2}+\frac{d}{dx}(x^2+3y^2)\frac{dy}{dx} &= -2\left[(1+y)+x\frac{dy}{dx}\right]\end{aligned}$$

Since $\left.\frac{dy}{dx}\right|_{(x,y)=(0,2)} = 0$, we have

$$(0+12)\left.\frac{d^2y}{dx^2}\right|_{(x,y)=(0,2)} + 0 = -2(1+2+0) = -6$$

hence the $\frac{d^2y}{dx^2}$ at (0.2) is $-\frac{1}{2}$.

6. If $y = \cos 2x$, show that $\frac{d^ny}{dx^n} = 2^n \cos\left(2x+\frac{n\pi}{2}\right)$, where $n$ is a positive integer. (Hint: use mathematical induction on $n$.)

**Suggested solution:**

We start with the case when $n = 1$, we have

$$\frac{d}{dx}\cos 2x = -2\sin 2x = 2^1\cos\left(2x+\frac{\pi}{2}\right)$$

Next, we assume that the quality holds when $n = k$ for some positive integer $k$. That is

$$\frac{d^ky}{dx^k} = 2^k\cos\left(2x+\frac{k\pi}{2}\right)$$

For $n = k + 1$, we have

$$\begin{aligned}\frac{d^{k+1}y}{dx^{k+1}} &= \frac{d}{dx}2^k\cos\left(2x+\frac{k\pi}{2}\right)\\ &= 2^k\left[-2\sin\left(2x+\frac{k\pi}{2}\right)\right]\\ &= -2^{k+1}\sin\left(2x+\frac{k\pi}{2}+\frac{\pi}{2}-\frac{\pi}{2}\right)\\ &= -2^{k+1}\sin\left(2x+\frac{k+1}{2}\pi-\frac{\pi}{2}\right)\\ &= 2^{k+1}\cos\left(2x+\frac{k+1}{2}\pi\right)\end{aligned}$$

By the principle of mathematical induction, we can conclude as desired.

7. Let $f(x) = xe^{1/x^2}$, $x > 0$. Determine all the vertical, horizontal, and oblique asymptotes of the graph of $f$, if any.
**Suggested solution:**
Since $f$ is continuous on $(0, \infty)$, the only possible vertical asymptote is the $y$-axis. Let's test whether the $y$-axis is a vertical asymptote or not. (In fact, the $y$-axis is a vertical asymptote.) By l'Hôpital's rule

$$\lim_{x\to 0^+} xe^{1/x^2} = \lim_{x\to 0^+} \frac{e^{1/x^2}}{\frac{1}{x}} = \lim_{x\to 0^+} \frac{-2e^{1/x^2}\frac{1}{x^3}}{-\frac{1}{x^2}} = \lim_{x\to 0^+} 2e^{1/x^2}\frac{1}{x} = \infty$$

Next let's find the oblique asymptote.

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} e^{\frac{1}{x^2}} = 1$$

Then we compute

$$\lim_{x\to\infty} (f(x) - x) = \lim_{x\to\infty} x(e^{1/x^2} - 1) = \lim_{x\to\infty} \frac{e^{1/x^2} - 1}{1/x} = \lim_{x\to\infty} \frac{-(2/x^3)e^{1/x^2}}{-1/x^2} = \lim_{x\to\infty} \frac{2e^{1/x^2}}{x} = 0$$

So $y = x$ is an oblique asymptote of $f$, therefore $f$ does not have any horizontal asymptote.

8. Let $f$ be a differentiable function and $g(x) = 1 + \sin f(x)$. Given that $g$ has an inverse function $h = g^{-1}$, $f(0) = 0$ and $f'(0) = 3$. Explain why $h$ is differentiable at 1 and compute $h'(1)$.
**Suggested solution:**
Since $g$ is a composition of differentiable functions, $g$ is differentiable and

$$g'(x) = f'(x)\cos f(x)$$

Note that $g(0) = 1 + \sin f(0) = 1$ and $g'(0) = f'(0)\cos f(0) = 3 \neq 0$. By the inverse function theorem, $g^{-1}$ is differentiable at $g(0) = 1$ and $h'(1) = (g^{-1})'(1) = 1/g'(0) = 1/3$.

9. Let $g(x) = (x^2+1)^{\sin x}$, $x \in \mathbb{R}$. First show that $g$ is differentiable and find the derivative. Determine whether the following claims are true or false.

   (a) $g$ has absolute maximum.

   (b) $g$ has infinitely many critical points.

**Suggested solution:**
Notice that we cannot discuss the differentiability for a function as the composition of a power function with the power and the base are two functions. Hence, we try to transfer the function to some other forms. Note that $1 + x^2$ is always positive and hence $g$ is the same as

$$g(x) = e^{\ln g(x)} = e^{(\sin x)\ln(1+x^2)}$$

As $1 + x^2 > 0$, $1 + x^2$ and $\ln u$ are differentiable, $\ln(1 + x^2)$ is well-defined and is differentiable by the chain rule. As $\sin x$ is differentiable, the product $(\sin x)\ln(1 + x^2)$ is differentiable by the product rule. As $e^u$ is differentiable, $g$ is differentiable by the chain rule.
Derivative of $g$ can be obtained by logarithmic differentiation since directly differentiating is hard to deal with.

$$\frac{g'(x)}{g(x)} = (\ln g(x))' = \cos x \ln(x^2+1) + \frac{\sin x}{x^2+1}2x$$
$$g'(x) = \left(\cos x \ln(x^2+1) + \frac{2x\sin x}{x^2+1}\right)(x^2+1)^{\sin x}$$

(a) False. For any $k \in \mathbb{N}$, consider $g(2k\pi + \pi/2)$, we have

$$g(2k\pi + \pi/2) = ((2k\pi + \pi/2)^2 + 1)^{\sin(2k\pi+\pi/2)} = (2k\pi + \pi/2)^2 + 1$$
$$\lim_{k\to\infty} g(2k\pi + \pi/2) = \infty$$

So $g$ does not have any absolute maximum.

(b) True. Note that for any $k \in \mathbb{N}$,

$$g(k\pi) = (k^2\pi^2 + 1)^{\sin(k\pi)} = 1 = ((k+1)^2\pi^2 + 1)^{\sin((k+1)\pi)} = g((k+1)\pi)$$

By the Rolle's Theorem, $g'(c_k) = 0$ for some $c_k \in (k\pi, (k+1)\pi)$.

## 3.4 Chapter 4: Integration

1. Find $\displaystyle\int \frac{1}{1-2x}\,dx$.
**Suggested solution:**
We make the substitution $u = 1 - 2x$ and $\dfrac{du}{dx} = -2$. Then

$$du = \left(\frac{du}{dx}\right) dx = -2\,dx.$$

So we have

$$\begin{aligned}\int \frac{1}{1-2x}\,dx &= \int \frac{1}{u}\left(-\frac{1}{2}\,du\right) = -\frac{1}{2}\int \frac{1}{u}\,du \\ &= -\frac{1}{2}\ln|u| + C = -\frac{1}{2}\ln|1-2x| + C.\end{aligned}$$

2. Find $\displaystyle\int 2x\sqrt{1+x^2}\,dx$.
**Suggested solution:**
We make the substitution $u = 1 + x^2$ and $\dfrac{du}{dx} = 2x$. Then

$$du = \left(\frac{du}{dx}\right) dx = 2x\,dx.$$

So we have

$$\int 2x\sqrt{1+x^2}\,dx = \int \sqrt{u}\,du = \frac{2}{3}u^{3/2} + C = \frac{2}{3}(1+x^2)^{3/2} + C.$$

3. Evaluate $\displaystyle\int \tan^2 x\,dx$.
**Suggested solution:**
We can consider that $\tan x = \dfrac{\sin x}{\cos x}$. Now we transfer the original integration to $\displaystyle\int \frac{\sin^2 x}{\cos^2 x}\,dx$. Here we have the power forms of $\sin x$ and $\cos x$, we can recall that $\sin^2 x + \cos^2 x = 1$, let's use this to substitute $\sin^2 x$ to see what will happen.

$$\int \frac{\sin^2 x}{\cos^2 x}\,dx = \int \frac{1-\cos^2 x}{\cos^2 x}\,dx = \int \left(\frac{1}{\cos^2 x} - 1\right) dx = \int \left(\sec^2 x - 1\right) dx = \tan x - x + C.$$

4. Evaluate $\displaystyle\int \sin 2x\,dx$.
**Suggested solution:**
Since this integration is in power 1 and have no other special features, let's try to integrate it directly.

$$\int \sin 2x\,dx = -\frac{1}{2}\cos 2x + C.$$

*Remark.* We have following two formulas for this type of integration:

- $\displaystyle\int \sin(ax+b)\,dx = -\frac{1}{a}\cos(ax+b) + C$

- $\displaystyle\int \cos(ax+b)\,dx = \frac{1}{a}\sin(ax+b) + C$

5. Evaluate $\displaystyle\int \cos^2 3x\,dx$.
   **Suggested solution:**
   We see some power term, let's consider the double/triple angle formula. We know that

   $$\cos 2x = \cos^2 x + \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x.$$

   Here we have the term we want and $\cos^2 x = \dfrac{\cos 2x + 1}{2}$. So

   $$\int \cos^2 3x\,dx = \int \frac{\cos 6x + 1}{2}\,dx = \frac{1}{12}\sin 6x + \frac{1}{2}x + C.$$

6. Evaluate $\displaystyle\int \sin 3x \cos 5x\,dx$.
   **Suggested solution:**
   Noticed that the integrator is a product of two trigonometric functions, we can consider the product-to-sum formula.

   $$\sin\alpha\sin\beta = \frac{1}{2}[\sin(\alpha+\beta) + \sin(\alpha-\beta)].$$

   So we have

   $$\int \sin 3x\cos 5x\,dx = -\frac{1}{16}\cos 8x + \frac{1}{4}\cos 2x + C.$$

7. Evaluate $\displaystyle\int \frac{dx}{x^2+4}$.
   **Suggested solution:**
   Motivation: We know that $\displaystyle\int \frac{1}{x^2+1}\,dx = \tan x + C$. Then

   $$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{\frac{1}{4}x^2+1} = \frac{1}{2}\int \frac{du}{u^2+1} (\text{ for } u = \frac{x}{2}) = \frac{1}{2}\tan^{-1}\left(\frac{x}{2}\right) + C.$$

   Hence we have a generalized result for $a > 0$:

   $$\int \frac{b}{x^2+a}\,dx = \frac{b}{\sqrt{a}}\tan^{-1}\left(\frac{x}{\sqrt{a}}\right) + C.$$

   We have two extra similar practices:

   - $\displaystyle\int \frac{dx}{\sqrt{9-x^2}} = \sin^{-1}\left(\frac{x}{3}\right) + C$
   - $\displaystyle\int \frac{dx}{x^2+4x+13} = \frac{1}{3}\tan^{-1}\left(\frac{x+2}{3}\right) + C$

8. Let $I_n = \int (\ln t)^n\,dt$ for any integer $n \geq 1$.
   (a) Use integration by parts to obtain a formula relating $I_n$ and $I_{n-1}$ for any $n \geq 2$.
   (b) Find $I_1$. Use your formula in the previous part to compute $I_2$ and $I_3$.

   **Suggested solution:**

   (a) Integration by parts to get

   $$\begin{aligned} I_n &= t(\ln t)^n - \int t\,d((\ln t)^n) \\ &= t(\ln t)^n - n\int (\ln t)^{n-1}\,dt \\ &= t(\ln t)^n - nI_{n-1} \end{aligned}$$

(b) It is known that $I_1 = t\ln t - t + C$. Using the formula in previous part,

$$\begin{aligned} I_2 &= t(\ln t)^2 - 2I_1 = t(\ln t)^2 - 2t\ln t + 2t + C \\ I_3 &= t(\ln t)^3 - 3I_2 = t(\ln t)^3 - 3[t(\ln t)^2 - 2t(\ln t - 1)] + C \\ &= t(\ln t)^3 - 3t(\ln t)^2 + 6t(\ln t) - 6t + C \end{aligned}$$

9. We have two functions $y = \ln x$ and $y = \dfrac{1}{x}$. They have an intersection in the first quadrant, suppose the x-coordinate of the intersection is $a$. Find the area of the region between these two functions and between the two straight lines $x = a$ and $x = e$. (final answer should only contains terms without any logarithm)
**Suggested solution:**
We should check the relationship between $a$ and $e$ to confirm with straight line is on the left.
We know that the function value of $y = lnx$ is always smaller than the function value of $y = \frac{1}{x}$ when $x < a$, and greater when $x > a$. Test the function value when $x = e$. $\ln e = 1$ and $\frac{1}{e} < 1 < \ln e$. Hence we can conclude that the straight line $x = e$ is on the right hand side of the straight line $x = a$.
Now we can evaluate the area by

$$\int_a^e \ln x - \frac{1}{x}\, dx = [x(\ln x - 1) - \ln|x|]_a^e = -1 - a(\ln a - 1) + \ln a$$

In order to escape some terms of logarithm, we noticed that $\ln a = \dfrac{1}{a}$ since $a$ is the x-coordinate of the intersection. So we can substitute $\ln a$ by $\dfrac{1}{a}$. We eventually have the area is $-2 + a + \dfrac{1}{a}$.
*Remark.* Is there any method that we can confirm this answer makes sense?
We know that the area of the region cannot be negative, so let's test the relationship between $-2+a+\dfrac{1}{a}$ and 0. We know that $a$ is positive, by the basic inequality, we always have

$$-2 + a + \frac{1}{a} \geq -2 + 2\sqrt{a \cdot \frac{1}{a}} = -2 + 2 = 0$$

So the area is always non-negative, indicating this answer makes sense.

10. Differentiate $\ln(x + \sqrt{3 + x^2})$ and $x\sqrt{3 + x^2}$. Hence find $\int \sqrt{3 + x^2}\, dx$.
**Suggested solution:**

$$\begin{aligned} (\ln(x + \sqrt{3 + x^2}))' &= \frac{1}{x + \sqrt{3 + x^2}} \left(1 + \frac{1}{2} \cdot 2x \frac{1}{\sqrt{3 + x^2}}\right) \\ &= \frac{1}{x + \sqrt{3 + x^2}} \frac{x + \sqrt{3 + x^2}}{\sqrt{3 + x^2}} \\ &= \frac{1}{\sqrt{3 + x^2}} \\ (x\sqrt{3 + x^2})' &= \sqrt{3 + x^2} + x \cdot \frac{1}{2} \cdot 2x \frac{1}{\sqrt{3 + x^2}} \\ &= \sqrt{3 + x^2} + \frac{x^2}{\sqrt{3 + x^2}} \\ &= \frac{3 + 2x^2}{\sqrt{3 + x^2}} \end{aligned}$$

We can write

$$\sqrt{3 + x^2} = \frac{1}{2}\left(\frac{3 + 2x^2}{\sqrt{3 + x^2}} + \frac{3}{\sqrt{3 + x^2}}\right)$$

Then by previous results, we have

$$\begin{aligned} \int \sqrt{3 + x^2}\, dx &= \frac{1}{2}\int \frac{3 + 2x^2}{\sqrt{3 + x^2}}\, dx + \frac{3}{2}\int \frac{1}{\sqrt{3 + x^2}}\, dx \\ &= \frac{1}{2}x\sqrt{3 + x^2} + \frac{3}{2}\ln(x + \sqrt{3 + x^2}) + C. \end{aligned}$$

11. Find a reduction formula of $I_n = \displaystyle\int \frac{1}{(x^2+a^2)^n}\,dx$ where $n > 1$ and $a > 0$.
Recall that $I_1$ is
$$\int \frac{dx}{x^2+a^2} = \frac{1}{a}\arctan\frac{x}{a} + C$$

**Suggested solution:**

$$\begin{aligned}
I_n &= \frac{1}{a^2}\int \frac{(x^2+a^2)-x^2}{(x^2+a^2)^n}\,dx = \frac{1}{a^2}\left(I_{n-1} - \int \frac{x^2}{(x^2+a^2)^n}\,dx\right) \\
a^2 I_n &= I_{n-1} + \frac{1}{2(n-1)}\left(\frac{x}{(x^2+a^2)^{n-1}} - \int \frac{dx}{(x^2+a^2)^{n-1}}\right) \\
I_n &= \frac{2n-3}{2a^2(n-1)} I_{n-1} + \frac{1}{2a^2(n-1)} \cdot \frac{x}{(x^2+a^2)^{n-1}}
\end{aligned}$$

12. Determine whether $\displaystyle\int_a^\infty \frac{1}{x} e^{2+\sin 1/x}\,dx$ exists or not.
**Suggested solution:**
Note that for any $x$,
$$2 + \sin(1/x) \geq 2 + (-1) = 1$$
$$e^{2+\sin(1/x)} \geq e^1 = e > 0$$

It follows for any $x > 1$,
$$0 \leq \frac{e}{x} \leq \frac{1}{x} e^{2+\sin(1/x)}$$

For any $b > 1$, by the fundamental theorem of calculus,
$$\int_1^b \frac{e}{x}\,dx = e[\ln x]_1^b = e\ln b \to \infty \text{ as } b \to \infty$$

Hence the given improper integral does not exist and diverges to $\infty$.

## 3.5 Chapter 5: Numerical Methods

1. Consider the graph of the equation
$$xy^2 - e^{-(x^2+y)} = 0.$$
Given that it passes through the point $P(1,-1)$ and $y = f(x)$ is the local solution of this equation at $P$.

   (a) Find the equation of the tangent line to the graph of the equation at $P$.

   (b) Let $Q$ be a point of the graph near $P$ with coordinates $(1.1, \beta)$. Use the 2nd Taylor polynomial of $f$ at the reference point $P(1,-1)$ to approximate $\beta$.

**Suggested solution:**

(a) Implicit differentiate the equation with respect to $x$ to get
$$y^2 + 2xyy' + (2x + y')e^{-(x^2+y)} = 0.$$
Substitute $(x,y) = (1,-1)$ to get
$$1 - 2y' + 2 + y' = 0.$$
So that $y' = 3$. It follows that the equation of the tangent line at point $P$ is
$$y = -1 + 3(x-1) = 3x - 4.$$

(b) Implicit differentiate the equation

$$y^2 + 2xyy' + (2x + h')e^{-(x^2+y)} = 0$$

with respect to $x$ to get

$$2yy' + 2(yy' + x(y')^2 + xyy'') + e^{-(x^2+y)}(-(2x + y')^2 + (2 + y'')) = 0.$$

Substitute $(x, y, y') = (1, -1, 3)$ and get

$$-6 + 2(-3 + 9 - y'') + (-(2 + 3)^2 + (2 + y'')) = 0$$

$$-6 + 12 - 2y'' - 25 + 2 + y'' = 0.$$

So that $y'' = -17$. It follows that the second Taylor polynomial at 1 is

$$P_2(x) = f(1) + f'(1)(x - 1) + \frac{1}{2}f''(1)(x - 1)^2 = -1 + 3(x - 1) - \frac{17}{2}(x - 1)^2$$

and thus $\beta \approx P_2(1.1) = -0.785$.

2. Let $f(x) = x \arctan x - \displaystyle\int_1^x \arctan t\, dt$.

(a) Without computing $\displaystyle\int \arctan t\, dt$, find $f'$ and $f''$.

(b) Find the second degree Taylor polynomial of $f$ at the reference point $\alpha = 1$.

(c) Use the polynomial in part (b) to estimate $f(1.1)$ and find an upper bound for the approximation error.

**Suggested solution:**

(a) By the fundamental theorem of calculus,

$$f'(x) = x \cdot (\arctan x)' + \arctan x - \arctan x = \frac{x}{1 + x^2}.$$

Differentiate $f'$ to get

$$f''(x) = \frac{1 - x^2}{(1 + x^2)^2}.$$

(b) Using the formula of $f$ and part (a), $f(1) = 1 \cdot \arctan 1 - 0 = \pi/4$, $f'(1) = 1/2$ and $f''(1) = 0$. As a result, the second degree Taylor polynomial of $f$ is

$$P_2(x) = \frac{\pi}{4} + \frac{1}{2} \cdot (x - 1).$$

(c) Firstly, $f(1.1) \approx P_2(1.1) = \frac{\pi}{4} + \frac{1}{20}$. We then proceed to find an upper bound for the error committed when $f(1.1)$ is approximated by $P_2(1.1)$. Further differentiation yields

$$f^{(3)}(x) = \frac{2x(x^2 - 3)}{(x^2 + 1)^3}$$

and hence the error committed is given by

$$|f(1.1) - P_2(1.1)| = \frac{1}{3!}|f^{(3)}(c)|0.1^3 = \frac{2c(3 - c^2)}{6000(c^2 + 1)^3}$$

for some $c \in (1, 1.1)$. If follows that an upper bound is

$$\frac{2(1.1)(3 - 1^2)}{6000(1^2 + 1)^3} = \frac{11}{120000}$$

as $0 \leq 3 - c^2 \leq 3 - 1^2$ and $0 \leq (c^2 + 1)^{-3} \leq (1^2 + 1)^{-3}$ for all $c \in (1, 1.1)$.

3. Use Newton's method to find the root of the following equation

$$xe^{-x} = 0,$$

with initial points at 0.5 and 1.1 respectively. (Stopping condition: $|x_{n+1} - x_n| < 0.01$)
**Suggested solution:**
Here, we have $f(x) = xe^{-x}$ and $f'(x) = (1-x)e^{-x}$. With initial points 0.5 and 1.1, we can have the following tables:

| $n$ | $x_n$ | $f(x_n)$ | $f'(x_n)$ | $\|x_{n+1} - x_n\|$ |
|---|---|---|---|---|
| 0 | 0.5000 | 0.3033 | 0.3033 | |
| 1 | -0.5000 | -0.8244 | 2.4731 | 1.0000 |
| 2 | -0.1667 | -0.1969 | 1.3783 | 0.3333 |
| 3 | -0.0238 | -0.0244 | 1.0485 | 0.1429 |
| 4 | -0.0006 | -0.0006 | 1.0011 | 0.0233 |
| 5 | -0.0000 | -0.0000 | 1.0000 | 0.0006 |

| $n$ | $x_n$ | $f(x_n)$ | $f'(x_n)$ | $\|x_{n+1} - x_n\|$ |
|---|---|---|---|---|
| 0 | 1.1000 | 0.3662 | -0.0333 | |
| 1 | 12.1000 | 0.0001 | -0.0001 | 11.0000 |
| 2 | 13.1901 | 0.0000 | -0.0000 | 1.0901 |
| 3 | 14.2721 | 0.0000 | -0.0000 | 1.0820 |
| 4 | 15.3475 | 0.0000 | -0.0000 | 1.0753 |
| 5 | 16.4172 | 0.0000 | -0.0000 | 1.0697 |
| $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ |

Observing the above tables, when the initial point is 0.5, our final $x_n$ is around 0, which is a good approximation of the root. But it seems not the case when the initial point is 1.1. In fact, when the initial point is 1.1, the sequence $\{x_n\}$ is divergent. So we cannot use the Newton's method with initial point 1.1 to approximate the root.
For $f(x) = xe^{-x}$, in course MATH3904 or MATH3601, we may learn that the $x_n$ obtained by Newton's method will converge to the root $x^* = 0$ if we have an initial point $x_0 < 1$ and diverge to infinity if $x_0 > 1$.

## 3.6 Chapter 6: Differential Equations

1. $y' = \dfrac{xy\sin x}{y+1}$, $y(0) = 1$.

2. $xy' = y + xe^{y/x}$, $y(1) = 0$.

3. $x^2y' + 2xy = \ln x$, $y(1) = 2$.

4. $(x^2+1)y' + 3x(y-1) = 0$, $y(0) = 2$.

5. $y' = \dfrac{y^2}{1+x^2}$, $y(0) = 1$.

6. $2y' - y = 4\sin 3x$, $y(0) = 1$.

7. $y'\sin x + 2y\cos x = \cos x$.

8. $y' + 2y = xy^{-2}$.

9. $y'' - y' - 6y = 0, y(0) = 1, y'(0) = -7$.

10. $y'' + 4y' + 5y = 0, y(0) = -2, y'(0) = 5$.

11. $y'' + 6y' + 9y = 0, y(0) = 2, y'(0) = -7$.
12. $x'' - 4x = 3e^{-2t}$.
13. $x'' + 3x' = -2t$.
14. $x'' - 4x' = 0$.

**Suggested solution:**

1. The ODE is a first order separable ODE.
$$\int \frac{y+1}{y}\,dy = \int x\sin x\,dx$$
$$\int (1+\frac{1}{y})\,dy = \int x\,d(-\cos x)$$
$$y + \ln|y| = -x\cos x + \int \cos x\,dx = -x\cos x + \sin x + C$$
Since $y(0) = 1$, we have $1 + \ln 1 = 1 = -0\cos 0 + \sin 0 + C$, so $C = 1$. The solution is $y + \ln|y| = -x\cos x + \sin x + 1$.

2. Consider substitute $y$ as $y = vx$ for $v$ is a function of $x$. Then $y' = v + v'x$. Substitute into the original ODE, we have
$$v + xv' = \frac{y}{x} + e^{y/x} = v + e^v$$
$$xv' = e^v$$
$$\int \frac{dv}{e^v} = \int \frac{dx}{x}$$
$$-e^{-v} = \ln|x| + C$$
Since $y(1) = 0$, we have $-1 = -e^0 = \ln 1 + C$, so $C = -1$. The solution in terms of $y$ is $e^{-y/x} = 1 - \ln|x|$.

3. Notice that the LHS of the ODE is the derivative of $x^2y$, so we have
$$x^2y = \int \ln x\,dx = x\ln x - \int x\,d(\ln x) = x\ln x - \int dx = x\ln x - x + C$$
Since $y(1) = 2$, we have $2 = 1\ln 1 - 1 + C$, so $C = 3$. The solution is $x^2y = x\ln x - x + 3$.

4. We try to use the integrating factor to solve this ODE. Simplify the ODE, we have
$$(x^2+1)y' + 3xy = 3x$$
$$y' + \frac{3x}{x^2+1}y = \frac{3x}{x^2+1}$$
Take the integrating factor as $e^{\int \frac{3x}{x^2+1}\,dx} = e^{\ln(x^2+1)^{3/2}} = (x^2+1)^{3/2}$. Thus we have
$$((x^2+1)^{\frac{3}{2}}y)' = 3x(x^2+1)^{\frac{1}{2}}$$
$$(x^2+1)^{\frac{3}{2}}y = \int 3x(x^2+1)^{\frac{1}{2}}\,dx = \int \frac{3}{2}u^{\frac{1}{2}}\,du = u^{\frac{3}{2}} + C = (x^2+1)^{\frac{3}{2}} + C \text{ with } u = x^2+1$$
Since $y(0) = 2$, we have $2 = 1 + C$, so $C = 1$. The solution is $y = 1 + (x^2+1)^{-\frac{3}{2}}$.

5. This is a first order separable ODE.
$$\int \frac{dy}{y^2} = \int \frac{dx}{1+x^2}$$
$$-\frac{1}{y} = \arctan x + C$$
Since $y(0) = 1$, we have $-1 = 0 + C$, so $C = 1$. The solution is $y = \dfrac{1}{1-\arctan x}$.

6. Simplifying the ODE, we have
$$y' - \frac{1}{2}y = 2\sin 3x.$$
The integrating factor is $e^{\int -\frac{1}{2}\,dx} = e^{-\frac{x}{2}}$. Then
$$(e^{-\frac{x}{2}}y)' = 2e^{-\frac{x}{2}}\sin 3x$$
$$e^{-\frac{x}{2}}y = \int 2e^{-\frac{x}{2}}\sin 3x\,dx = 2\int e^{-\frac{x}{2}}\sin 3x\,dx = -\frac{24}{37}e^{-\frac{x}{2}}\cos 3x - \frac{4}{37}e^{-\frac{x}{2}}\sin 3x + C.$$
Then
$$y = \frac{24}{37}\cos 3x - \frac{4}{37}\sin 3x + Ce^{\frac{x}{2}}.$$
Since $y(0) = 1$, we have $-\frac{24}{37} + C = 1$, so $C = \frac{61}{37}$. The solution is $y = \frac{24}{37}\cos 3x - \frac{4}{37}\sin 3x + \frac{61}{37}e^{\frac{x}{2}}$.

7. Simplifying the ODE, we have
$$y' + 2y\cot x = \cot x.$$
The integrating factor is $e^{\int 2\cot x\,dx} = e^{2\ln\sin x} = \sin^2 x$. Then
$$y\sin^2 x = \int \sin^2 x\cot x\,dx = \int \sin x\cos x\,dx = -\frac{1}{4}\cos 2x + C.$$
The solution is $y = \dfrac{C - \cos 2x}{4\sin^2 x}$.

8. This is a Bernoulli Equation. Consider $z = y^{1-(-2)} = y^3$, then $z' = 3y^2y'$. Substitute into the ODE, we have
$$3y^2y' + 2(3y^2)y = xy^{-2}(3y^2)$$
$$z' + 6z = 3x.$$
The integrating factor is $e^{\int 6\,dx} = e^{6x}$. So we have
$$e^{6x}z = \int 3xe^{6x}\,dx = e^{6x}(\frac{1}{2}x - \frac{1}{12}) + C.$$
The solution is $y^3 = \frac{1}{2}x - \frac{1}{12} + Ce^{-6x}$.

9. Solve the characteristic equation
$$\lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2) = 0,$$
a general solution of the ODE is
$$y(x) = Ae^{-2x} + Be^{3x}.$$
Impose the initial condition to get
$$y(0) = A + B = 1$$
$$y'(0) = -2A + 3B = -7$$
Solve the system of equations to get $A = 2$ and $B = -1$. The solution is $y(x) = 2e^{-2x} - e^{3x}$.

10. Solve the characteristic equation
$$\lambda^2 + 4\lambda + 5 = (\lambda + 2 + \sqrt{-1})(\lambda + 2 - \sqrt{-1}) = 0,$$
a general solution of the ODE is
$$y(x) = e^{-2x}(A\cos x + B\sin x).$$
Impose the initial condition to get
$$y(0) = A = -2$$
$$y'(0) = -2A + B = 5$$
Solve the system of equations to get $A = -2$ and $B = 1$. The solution of the ODE is $y(x) = e^{-2x}(-2\cos x + \sin x)$.

11. Solve the characteristic equation
$$\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2 = 0,$$
a general solution of the ODE is
$$y(x) = Ae^{-3x} + Bxe^{-3x}.$$
Impose the initial condition to get
$$y(0) = A = 2$$
$$y'(0) = -3A + B = -7$$
Solve the system of equations to get $A = 2$ and $B = -1$. The solution is $y(x) = e^{-3x}(2 - x)$.

12. First, solve the homogeneous solution.
$$x_c(t) = Ae^{2t} + Be^{-2t}.$$
Next, we guess the particular solution as the form of
$$x_p(t) = Cte^{-2t}.$$
(We can try $x_p(t) = Ce^{-2t}$ first and find that it is in the form already seen in the homogeneous solution, so we need to adjust it to the above form.)
Plug the particular solution into the ODE and solve for $C$.
$$C(-4e^{-2t} + 4te^{-2t}) - 4Cte^{-2t} = -4Ce^{-2t} = 3e^{-2t}$$
So we have $C = -\frac{3}{4}$. The general solution is $x(t) = Ae^{2t} + Be^{-2t} - \frac{3}{4}te^{-2t}$.

13. First, we solve the homogeneous solution.
$$x'' + 3x' = (x' + 3x)' = 0$$
Then we have
$$x' + 3x = \int 0\,dt = A.$$
Calculate the integrating factor $e^{\int 3\,dt} = e^{3t}$. Hence
$$e^{3t}x = \int Ae^{3t}\,dt = A'e^{3t} + B.$$
So
$$x_c(t) = A' + Be^{-3t}.$$
Next, we guess the particular solution as the form of
$$x_p(t) = Ct^2 + Dt + E.$$
Plug the particular solution into the ODE and solve for $C, D$ and $E$.
$$2C + 3(2Ct + D) = 2C + 6Ct + 3D = -2t$$
Then we have $C = -\frac{1}{3}$ and $D = \frac{2}{9}$. The general solution is $x(t) = A' + Be^{-3t} - \frac{1}{3}t^2 + \frac{2}{9}t + E = F + Be^{-3t} - \frac{1}{3}t^2 + \frac{2}{9}t$.

14. Notice that
$$x'' - 4x' = (x' - 4x)' = 0.$$
Then
$$x' - 4x = \int 0\,dt = A.$$
The integrating factor $e^{\int -4\,dt} = e^{-4t}$. So
$$e^{-4t}x = \int Ae^{-4t}\,dt = Ce^{-4t} + B.$$
The solution is $x(t) = C + Be^{4t}$.

## 3.7 Chapter 7: Matrices

1. Let $A = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}$. Find the $(3,2)$-entry of $\left(\frac{1}{3}A\right)^{-1}$.

   **Suggested Solutions:**
   $\det A = 2\cdot 3 + 3\cdot 4 + (-2) - 4\cdot 3 - 2\cdot 2 - (-3) = 3$. Next, note that
   $$\left(\frac{1}{3}A\right)^{-1} = 3A^{-1}.$$
   So we proceed to find the $(3,2)$ entry of $A^{-1}$:
   $$\frac{1}{\det A}C_{23} = \frac{1}{3}(-1)^{2+3}\begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = -1.$$
   As a result, the $(3,2)$ entry is $3\cdot(-1) = -3$.

2. Compute the determinants of the following matrices:
   $$A = \begin{bmatrix} 5 & 7 \\ -3 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & -1 \\ 2 & 5 & 7 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 & 3 \\ -3 & 1 & -3 \\ 3 & -1 & 1 \end{bmatrix}.$$
   Are they invertible? Find the inverse whenever the matrix is invertible.

   **Suggested Solutions:**
   Since $\det A = \begin{vmatrix} 5 & 7 \\ -3 & -4 \end{vmatrix} = 1 \neq 0$, $A$ is invertible. In addition,
   $$A^{-1} = \begin{bmatrix} -4 & -7 \\ 3 & 5 \end{bmatrix}.$$
   Since $\det B = \begin{vmatrix} 1 & 4 & 2 \\ 0 & 1 & -1 \\ 2 & 5 & 7 \end{vmatrix} = 7 - 8 + 5 - 4 = 0$, B is singular.
   Since $\det P = \begin{vmatrix} 1 & 1 & 3 \\ -3 & 1 & -3 \\ 3 & -1 & 1 \end{vmatrix} = 1 - 9 + 9 - 3 + 3 - 9 = -8 \neq 0$, $P$ is invertible. Note that the cofactors of $P$ are
   $C_{11} = \begin{vmatrix} 1 & -3 \\ -1 & 1 \end{vmatrix} = -2, C_{12} = -\begin{vmatrix} -3 & -3 \\ 3 & 1 \end{vmatrix} = -6, C_{13} = \begin{vmatrix} -3 & 1 \\ 3 & -1 \end{vmatrix} = 0, C_{21} = -\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = -4, C_{22} = \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} = -8,$
   $C_{23} = -\begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} = 4, C_{31} = \begin{vmatrix} 1 & 3 \\ 1 & -3 \end{vmatrix} = -6, C_{32} = -\begin{vmatrix} 1 & 3 \\ -3 & -3 \end{vmatrix} = -6, C_{33} = \begin{vmatrix} 1 & 1 \\ -3 & 1 \end{vmatrix} = 4.$
   It follows that
   $$P^{-1} = \frac{1}{\det A}(C_{ij})^T = -\frac{1}{8}\begin{bmatrix} -2 & -4 & -6 \\ -6 & -8 & -6 \\ 0 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/2 & 3/4 \\ 3/4 & 1 & 3/4 \\ 0 & -1/2 & -1/2 \end{bmatrix}.$$

3. Let $A$ and $B$ be square matrices in $3\times 3$. Suppose that $\det A = 2, \det B = -1$. Find

   (a) $\det(3A)$,

   (b) $\det(2B^{-1})$,

   (c) $\det(B^T A^{-1})$.

   **Suggested Solutions:**

   (a) $\det(3A) = 3^3 \det A = 27\cdot 2 = 54$.

(b) $\det(2B^{-1}) = 2^3 \det(B^{-1}) = 8 \cdot \dfrac{1}{\det B} = 8 \cdot \dfrac{1}{-1} = -8.$

(c) $\det(B^T A^{-1}) = \det(B^T)\det(A^{-1}) = \det B \dfrac{1}{\det A} = (-1) \cdot \dfrac{1}{2} = -\dfrac{1}{2}.$

*Remark.* If $A$ is a $n \times n$ square matrix, $\det(kA) = k^n \det A$ for any scalar $k$ and $n \in \mathbb{N}$. This proposition can be proved by the mathematical induction using the cofactor expansion of calculating the determinant of high-dimension matrix.

4. Let $A$ be a $2 \times 2$ matrix which satisfies $2A^2 - 3A - 5I = O$. If $\det(I + A) \neq 0$, find $\det A$.
**Suggested Solutions:**
The equation is equivalent to $(2A - 5I)(A + I) = O$. As $\det(A + I) \neq 0$, $A + I$ is invertible and we can use $(A + I)^{-1}$ to represent the inverse of $A + I$. So we have

$$2A - 5I = (2A - 5I)(A + I)(A + I)^{-1} = O(A + I)^{-1} = O,$$

$$A = (5/2)I.$$

Thus $\det A = \det((5/2)I) = (5/2)^2 \det I = 25/4$.

5. In Assignment 7, we have the definition of nilpotent matrix given by $A^k = O$ for some positive integer $k$.

(a) Show that the determinant of a nilpotent matrix is 0.

(b) By part(a), if a matrix is singular (not invertible), can we conclude the matrix is nilpotent?

**Suggested Solutions:**

(a) Given a nilpotent matrix $A$ and $A^k = O$, we have $\det(A^k) = \det O = 0$. Since $\det(A^k) = (\det A)^k$, we have $(\det A)^k = 0, \det A = 0$. So the determinant of a nilpotent matrix is 0.

(b) No. If a matrix is singular, then its determinant is 0, but we cannot conclude whethr it is nilpotent or not. We can consider an example $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. $\det B = 0$ but $B^k = B$ for any $k \in \mathbb{N}$ so $B$ is not nilpotent.
**Remarks:**
Some may argue that a singular matrix have determinant of 0, and the determinant of its power should be 0 as well, which is equal to the determinant of the zero matrix, so that we can conclude its power is the zero matrix, and the matrix is nilpotent. The problem of this argument is that, we cannot conclude that two matrices are equal if they have equal determinant, so we cannot conclude the power of singular matrix is just the zero matrix. So for singular matrices, some of them are nilpotent while some of them are not. But for nilpotent matrices, all of them are singular.

6. Let $A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ -1 & x & 2 & -1 \\ 2 & 0 & 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & y & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ where $x$ and $y$ are real non-negative constants. If $AB$ is singular, find the value of $x$ and $y$.
**Suggested Solutions:**
Compute that $AB = \begin{bmatrix} 0 & y & 1 \\ 4 & x - y - 4 & x - 1 \\ 3 & 2y - 7 & -2 \end{bmatrix}$. Then $\det(AB) = (3x + 16)(y - 1) = 0$ since $AB$ is singular. Since $x, y > 0$, the solution for the equation is $y = 1$ and $x$ can be any non-negative numbers.

7. Let $A$ and $B$ are square matrices with $AB = O$. Prove or disprove the following statements.

(a) $BA = O$.

(b) Either $A = O$ or $B = O$.

(c) If $A$ is invertible, then $B = O$.

**Suggested Solutions:**

(a) *Disproof.* Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$. Then $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ but $BA = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ which is not the zero matrix.

(b) *Disproof.* By the same counterexample in part (a).

(c) *Proof.* Since $A$ is invertible, we have $A^{-1}AB = A^{-1}O$, that is $B = O$.

## 3.8 Chapter 8: Vectors

1. Suppose $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in $\mathbb{R}^3$. Determine whether the statement is true or false: if $\boldsymbol{u} \times \boldsymbol{x} = \boldsymbol{v} \times \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^3$, then $\boldsymbol{u} = \boldsymbol{v}$.
**Suggested Solutions:**
The statement is *true*. Since the equality holds for any $\boldsymbol{x} \in \mathbb{R}^3$, we can substitute $\boldsymbol{x}$ by $\boldsymbol{u}$, then we get $\boldsymbol{0} = \boldsymbol{v} \times \boldsymbol{u}$. This means $\boldsymbol{v}$ and $\boldsymbol{u}$ are parallel. Without loss of generality, let $\boldsymbol{v} = k\boldsymbol{u}$ for some $k \in \mathbb{R}$. Then $\boldsymbol{u} \times \boldsymbol{x} = (k\boldsymbol{u}) \times \boldsymbol{x} = k(\boldsymbol{u} \times \boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^3$.

   - If $\boldsymbol{u} = \boldsymbol{0}$, then $\boldsymbol{v} = k\boldsymbol{u} = \boldsymbol{0} = \boldsymbol{u}$.
   - If $\boldsymbol{u} \neq \boldsymbol{0}$, then $\boldsymbol{u} \times \boldsymbol{x}$ is nonzero for those $\boldsymbol{x}$ not parallel to $\boldsymbol{u}$. So we must have $k = 1$, and hence $\boldsymbol{v} = \boldsymbol{u}$.

   This shows $\boldsymbol{u} = \boldsymbol{v}$ in any case.

2. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ and $\boldsymbol{x}$ be vectors in $\mathbb{R}^3$. Prove the following identities:

   (a) $\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2$.

   (b) $(\boldsymbol{u} \times \boldsymbol{v}) \cdot (\boldsymbol{w} \times \boldsymbol{x}) = \begin{vmatrix} \boldsymbol{u} \cdot \boldsymbol{w} & \boldsymbol{u} \cdot \boldsymbol{x} \\ \boldsymbol{v} \cdot \boldsymbol{w} & \boldsymbol{v} \cdot \boldsymbol{x} \end{vmatrix}$.

   **Suggested Solutions:**

   (a) Let $\theta$ be the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. Then we have

$$\begin{aligned} \|\boldsymbol{u} \times \boldsymbol{v}\|^2 &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 \sin^2\theta \\ &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2(1 - \cos^2\theta) \\ &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2\cos^2\theta \\ &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2. \end{aligned}$$

   (b) Use the property of vector triple product: $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}$ and consider $(\boldsymbol{w} \times \boldsymbol{x})$ as a whole, we have

$$\begin{aligned} (\boldsymbol{u} \times \boldsymbol{v}) \cdot (\boldsymbol{w} \times \boldsymbol{x}) &= \boldsymbol{u} \cdot (\boldsymbol{v} \times (\boldsymbol{w} \times \boldsymbol{x})) \\ &= (\boldsymbol{v} \times (\boldsymbol{w} \times \boldsymbol{x})) \cdot \boldsymbol{u} \\ &= ((\boldsymbol{v} \cdot \boldsymbol{x})\boldsymbol{w} - (\boldsymbol{v} \cdot \boldsymbol{w})\boldsymbol{x}) \cdot \boldsymbol{u} \\ &= (\boldsymbol{v} \cdot \boldsymbol{x})(\boldsymbol{w} \cdot \boldsymbol{u}) - (\boldsymbol{v} \cdot \boldsymbol{w})(\boldsymbol{x} \cdot \boldsymbol{u}) \\ &= \begin{vmatrix} \boldsymbol{u} \cdot \boldsymbol{w} & \boldsymbol{u} \cdot \boldsymbol{x} \\ \boldsymbol{v} \cdot \boldsymbol{w} & \boldsymbol{v} \cdot \boldsymbol{x} \end{vmatrix}. \end{aligned}$$

# References


1. Thomas, G. (1952). *Thomas' Calculus Early Transcendentals* (Weir, M. & Hass, J., Eds.). (2009). Pearson Education, Inc..
2. Spence, I. & Friedberg S. (2014). *Elementary Linear Algebra – A Matrix Approach*. Pearson Education, Inc..
3. MATH1821 (2024-2025 Sem1) Dr. Law Ka Ho, Lecture Slides
4. MATH1013 (2021-2022 Sem1) Course Materials
5. MATH1013 (2022-2023 Sem1) Mr. Huang Jiasheng (STA), Revision Class Materials
6. MATH3904 (2023-2024 Sem1) Prof. Zang Wenan, Course Materials