MATH1051 - LECTURE NOTES

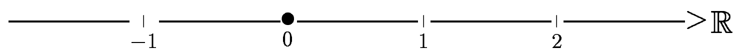
Real Numbers

$$\begin{aligned}
\mathbb{R} &= \text{set of all real numbers – represented by a decimal expansion} \\
&\qquad\qquad \text{(either truncating or not)} \\
\mathbb{Z} &= \text{set of integers} \\
\mathbb{N} &= \text{set of natural numbers: } 0,\ 1,\ 2,\ 3\ \ldots\ldots \\
\mathbb{Q} &= \text{set of rational numbers – i.e. those real numbers of form} \\
&\qquad \tfrac{p}{q} \text{ where } p, q \text{ are integers, } q \neq 0.
\end{aligned}$$

Irrational numbers: those reals which are not fractions eg. $\sqrt{2}$, π, $e = 2 \cdot 718$ ....

Ordering on $\mathbb{R}$:

Have order relation $<$ on $\mathbb{R}$. Geometrically, $a < b$ means a lies to left of b on real number line



e.g. $\frac{1}{2} < 1$ etc.
We write

$$a \leqslant b \quad \text{if} \quad a < b \quad \text{or} \quad a = b.$$

Absolute value:

Absolute value (or *modulus*) of x is given by

$$|x| = \begin{cases} x, & \text{if } x \geq 0 \\ -x, & \text{if } x < 0 \end{cases}$$

so we always have $|x| \geqslant 0$ e.g.

$$\left|-\frac{1}{2}\right| = \frac{1}{2}, \quad |-2| = 2, \quad |4| = 4 \quad \text{etc.}$$

Properties.

(i) $|c|^2 = c^2 \quad \Rightarrow \quad |c| = \sqrt{c^2}$ (positive square root)

(ii) $|a\, b| = |a| \cdot |b|$

(iii) But for $a+b$ all we can say is

$$|a+b| \leq |a| + |b| \quad \text{(triangle inequality)}$$

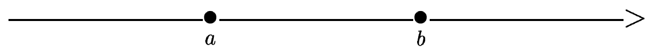
eg.

$$|2+(-3)| = |-1| = 1 \quad , \quad |2| + |-3| = 2+3 = 5.$$

Line Intervals:

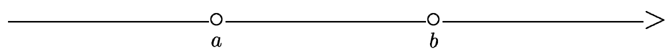
In calculus frequently wish to consider an entire segment of real number line. We have following types of intervals, where $a < b \in \mathbb{R}$:

Closed Interval:



$$\begin{aligned}[a,\, b] &= \{x \in \mathbb{R} \mid a \leq x \leq b\} \\ &= \text{set of } x \in \mathbb{R} \text{ s.t. } a \leq x \leq b.\end{aligned}$$

Open Interval:

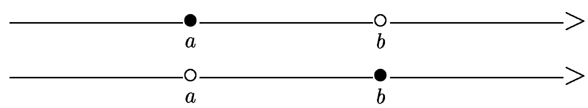


$$(a,\, b) = \{x \in \mathbb{R} \mid a < x < b\}.$$

Half-open, half closed:

$[a,b) = \{x \in \mathbb{R} \mid a \leq x < b\}$

$(a,b] = \{x \in \mathbb{R} \mid a < x \leq b\}$



Infinite intervals:

$$[a,\infty) = \{x \in \mathbb{R} \mid x \geq a\} \quad , \quad (a,\infty) = \{x \in \mathbb{R} \mid x > a\}$$

and similarly may define $(-\infty, a]$, $(-\infty, a)$. Finally we may write

$$\mathbb{R}(-\infty,\, \infty).$$

NOTE: Here $\pm\infty$ are just symbols – there is no real number ∞.

FUNCTIONS:

Let X, Y be subsets of $\mathbb{R}$. A *function* $f : X \to Y$ is a rule which assigns to every $x \in X$ *exactly one* element $f(x) \in Y$ – called the *value* of f at x. X is called the *domain* of f and

$$\{f(x) \mid x \in X\} \quad - \quad \text{denoted } f(X)$$

is called the *range* of f.

NOTE: The range of f is a subset of Y but need not equal Y.

Examples: (1) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is a function with domain $\mathbb{R}$. Here

$$\begin{aligned} \text{range of} \quad f &= \{f(x) \mid \in \mathbb{R}\} \\ &= \{x^2 \mid x \in \mathbb{R}\} = [0, \infty). \end{aligned}$$

(2) $f : (-2, 3) \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} -1 & , \quad -2 < x < 0 \\ \pi & , \quad x = 0 \\ x & , \quad 0 < x < 3 \end{cases}$$

is a perfectly good function.

Graph: May picture a function by drawing its *graph* – set of all points (x, y) in plane where $y = f(x)$. Vital function property states that to each point x in domain there must correspond *exactly* one y value on graph: i.e.

DIAGRAM 1

DIAGRAM 2

Examples: (1) Modulus function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ is a function with graph shown. In this case range of $f = [0, \infty)$

DIAGRAM 3

(2) DIAGRAM 4

$$f : \mathbb{R} \to \mathbb{R}, \quad \begin{aligned} f(x) &= x^2 - 4x + 3 \\ &= (x-2)^2 - 1 \\ &= (x-3)(x-1) \end{aligned}$$

is a function with graph shown. In this case range $f = [-1, \infty)$.

(3) DIAGRAM 5

Circle $x^2 + y^2 = 1$ is *not* a function since have two y-values $y = \pm\sqrt{1 - x^2}$ corresp. to each $x \in (-1, 1)$. However

$$y = \sqrt{1 - x^2}$$

DIAGRAM 6

is a function – top $\frac{1}{2}$ of circle.

NOTE: For $\alpha > 0$, $\sqrt{\alpha}$ always denotes the (positive) square root. Thus $\sqrt{4} = 2$ etc. For $\alpha > 0$, solutions to $x^2 = \alpha$ are $x = \pm\sqrt{\alpha}$.

CONVENTION: An expression like "the function $y = \sqrt{1 - x^2}$" means the function f with $y = f(x) = \sqrt{1 - x^2}$. When the domain is not specified it is taken to be the largest set on which rule is defined: so here, domain $= [-1, 1]$.

Exponential Function - given by function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$, which has graph shown. Note that range $f = (0, \infty)$.

DIAGRAM 7

Properties: We have the following exponent laws

(i) $e^x \cdot e^y = e^{x+y}$

(ii) $e^x / e^y = e^{x-y}$

(iii) $e^{rx} = (e^x)^r$, r any real number.

Trigonometric Functions

In calculus measure angles in *radians*:

DIAGRAM 8

1. radian = angle subtended at centre of circle of radius 1 by segment of arc length 1 so

$$2\pi \quad \text{radians} = 360^\circ.$$

DIAGRAM 9

The point on circle $x^2 + y^2 = 1$ making angle θ (radians) anti- clockwise from x-axis has coords $(\cos\theta, \sin\theta)$, so $\cos^2\theta + \sin^2\theta = 1$ – actually defined cos and sin functions

The graphs of functions $f(x) = \cos x$, $f(x) = \sin x$ are given below:

DIAGRAM 10 DIAGRAM 11

$y = \cos x$ $\qquad\qquad$ $y = \sin x$

In both cases range of $f = [-1, 1]$.

DIAGRAM 12

Have also the tan function $f : (-\pi/2, \pi/2) \longrightarrow \mathbb{R}$, $f(x) = \tan x = \dfrac{\sin x}{\cos x}$ with graph shown. In this case range $f = \mathbb{R}$. May similarly define other trig. functions: eg. $\sec x = \dfrac{1}{\cos x}$, cosec $x = \dfrac{1}{\sin x}$ etc.

Composition of functions

Given two functions f, g, the *composition* of f and g, denoted $f \circ g$, is the function defined by

$$(f \circ g)(x) = f(g(x)).$$

Example: $f(x) = x^2 + 1$, $g(x) = \dfrac{1}{x}$. Then

$$\begin{aligned} f(g(x)) &= g(x)^2 + 1 = \frac{1}{x^2} + 1 \quad , \quad x \neq 0 \\ g(f(x)) &= \frac{1}{f(x)} = \frac{1}{x^2 + 1} \end{aligned}$$

$1 - 1$ **Functions**

Definition: $f : X \to Y$ is said to be *one-to-one* $(1 - 1)$ if, for all $x_1, x_2 \in X$

$$f(x_1) = f(x_2) \Longrightarrow x_1 = x_2.$$

□

NOTE: Equivalently f is $1 - 1$ if for all $x_1 x_2 \in X$

$$x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2).$$

On the graph $y = f(x)$, $1 - 1$ means that line $y =$ const. cuts graph in at most 1 place: ie.

DIAGRAM 13 DIAGRAM 14

Examples: (1) $y = x^3$ is $1-1$ DIAGRAM 15

(2) $y = x^2$ is not $1-1$.

DIAGRAM 16

Inverse Functions

Let $f : X \to Y$ be a $1-1$ function. Then may define the *inverse* function $f^{-1} : f(X) \to X$ by

$$f^{-1}(f(x)) = x \quad , \quad \text{for all } x \in X.$$

By definition

$$\text{domain } f^{-1} = \text{ range } f \quad , \quad \text{range } f^{-1} = \text{ domain } f.$$

NOTE:

$$y = f(x) \leftrightarrow x = f^{-1}(y).$$

Need f to be $1-1$ in order that f^{-1} satisfy crucial function property.

Graph of $y = f^{-1}(x)$: Geometrically,
(a, b) lies on graph of $f(x)$

$\Leftrightarrow b = f(a)$
$\Leftrightarrow a = f^{-1}(b)$
$\Leftrightarrow (b, a)$ lies on graph of $f^{-1}(x)$.

DIAGRAM 17

Thus by interchanging x and y coords of all points on graph of $f(x)$ get graph of $f^{-1}(x)$ – ie. graph of $f^{-1}(x)$ is obtained from that of $f(x)$ by reflection about the line $y = x$.

Examples: (1)

DIAGRAM 18

$f(x) = x^3$ is $1-1$ and range $f = \mathbb{R}$. Its inverse is thus function $f^{-1} : \mathbb{R} \to \mathbb{R}$, $f^{-1}(x^3) = x$. We write $f^{-1}(x) = x^{1/3}$ or $f^{-1}(x) = \sqrt[3]{x}$ (cube root function).

(2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not $1 - 1$ $\therefore$ No inverse. But part $x \geq 0$ gives a $1-1$ function $f : [0, \infty) \longrightarrow \mathbb{R}$, $f(x) = x^2$ with range $[0, \infty)$. Corresponding inverse is then function

DIAGRAM 19

$$f^{-1} : [0, \infty) \longrightarrow [0, \infty) \quad , \quad f^{-1} = \sqrt{x}.$$

(Also part $x \leq 0$ is $1 - 1$: inverse is now $f^{-1}(x) = -\sqrt{x}$).

NOTE: This trick is often used. If function is not $1 - 1$ over all its domain, just take a part where it is $1 - 1$ and get inverse for that part.

Natural Logarithm

From graph, $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$ is $1 - 1$ and range of $f = (0, \infty)$. Thus inverse determines a function $f^{-1} : (0, \infty) \to \mathbb{R}$ with graph shown. We write

DIAGRAM 20

$$f^{-1}(x) = \ln x \qquad - \text{ called the } \textit{natural logarithm}$$

NOTES: (1) $\ln(1) = 0$

(2) $y = \ln x \leftrightarrow x = e^y$

$\therefore \quad \ln x = \log_e(x)$ (school definition of ln)

(3) By definition of inverse

$$\begin{aligned} e^{\ln x} &= x \quad , \quad x > 0 \\ \ln(e^x) &= x \quad , \quad x \in \mathbb{R} \end{aligned}$$

Properties: Using exponent laws, together with fact that ln is the inverse of exp, may prove the following: for $x, y > 0$

(1) $\ln(xy) = \ln(x) + \ln(y)$

(2) $\ln(x/y) = \ln(x) - \ln(y)$

(3) $\ln(x^r) = r \ln x$, r any real number.

In particular

$$\ln(\frac{1}{x}) = \ln(x^{-1}) = -\ln x.$$

INVERSE TRIG FUNCTIONS:

$y = \sin x$ is $1-1$ if we take just the interval $[-\pi/2,\, \pi/2]$. The inverse for this part is the function denoted

DIAGRAM 21

$$y = \text{arsin } x \qquad \text{or} \qquad y = \sin^{-1} x$$

DIAGRAM 22

Thus $\text{arsin}\, x$ is defined on interval $[-1,1]$ and takes values in the range $[-\pi/2,\, \pi/2]$. Have $\text{arsin}\, x = y \in [-\pi/2,\, \pi/2]$, with $\sin y = x$ eg.

$$\begin{aligned} \text{arsin}\,(1/\sqrt{3}) &= y \in [-\pi/2,\, \pi/2] \text{ with } \sin y = 1/\sqrt{2} \\ &= \pi/4. \end{aligned}$$

Similarly $y = \cos x$ is $1-1$ on $[0,\pi]$ and its inverse is the function denoted

DIAGRAM 23

$y = \arccos x \quad$ or $\quad y = \cos^{-1} x$

DIAGRAM 24

$\therefore$ $\arccos x$ is defined on $[-1,\, 1]$ and takes values in the range $[0,\, \pi]$.
Also $y = \tan x$ is $1-1$ on $(-\pi/2,\, \pi/2)$ with inverse

$$y = \text{artan}\, x \quad (\text{or } \tan^{-1} x)$$

DIAGRAM 25

– defined for $x \in (-\infty,\, \infty)$ with values in range $(-\pi/2,\, \pi/2)$, so $\text{artan}\, x = y \in (-\pi/2,\, \pi/2)$ with $\tan y = x$.

Limits:

$\lim_{x\to a} f(x) = \ell$, $\ell \in \mathbb{R}$, means $f(x)$ is close to ℓ for all x values sufficiently close to a, with $x \neq a$.

DIAGRAM 26

Roughly, x close to a, $x \neq a$, ensures $f(x)$ is close to ℓ. Use $\lim_{x\to a} f(x)$ to *predict* what happens to $x = a$ from looking at x values close to a, but $x \neq a$.

Examples: (1) $\lim_{x\to 1}(x^2+1) = 2$, $\lim_{x\to 0}|x| = 0$, $\lim_{x\to 0}\frac{1}{x^2}$ does not exist.

(2) Find $\lim_{x\to 3}\frac{x^2-9}{x-3}$.

Solution: $f(x) = \frac{x^2-9}{x-3}$ is not defined at $x = 3$ (gives $\frac{0}{0}$). But $f(x) = \frac{(x-3)(x+3)}{x-3} = x+3$, for $x \neq 3$.
Since when we take limit $x \to 3$, we assume $x \neq 3$,

DIAGRAM 27

$$\lim_{x\to 3} f(x) = \lim_{x\to 3}(x+3) = 6$$

(This is the reason for $x \neq a$).

ONE-SIDED LIMITS:

Consider eg.

$$f(x) = \begin{cases} 1 & , \quad x \geq 0 \\ -2 & , \quad x < 0. \end{cases}$$

Then $\lim_{x\to 0} f(x)$ does not exist.

For $x > 0$ want $\lim_{x\to 0} f(x) = 1$, while for $x < 0$ want $\lim_{x\to 0} f(x) = -2$. Have *one-sided* limits for this.

Say that limit as $x \to 0$ *from above* (or *from the right*) equals 1 and write

$$\lim_{x\to 0^+} f(x) = 1.$$

Similarly have limit as $x \to 0$ *from below* (or *from the left*) and write

$$\lim_{x\to 0^-} f(x) = -2.$$

NOTE: In general, for $\lim_{x\to a^+} f(x) = \ell$, just consider x with $x > a$ and similarly for $\lim_{x\to a^-} f(x) = \ell$, consider just $x < a$.

Example: $\lim_{x\to 2^+} \sqrt{x-2} = 0$ but $\lim_{x\to 2^-} \sqrt{x-2}$ does not exist since $\sqrt{x-2}$ not defined for $x < 2$.

DIAGRAM 29

Theorem: $\lim_{x\to a} f(x) = \ell$ if and only if

$$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = \ell.$$

DIAGRAM 30

Example:

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

Then

$$\left.\begin{array}{l} \lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} x^2 = 1 \\ \lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} (2-x) = 1, \end{array}\right\} \therefore \lim_{x\to 1} f(x) = 1.$$

Theorem: (Squeeze Principle)

Suppose

$$\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = \ell$$

and, for x close to a $(x \neq a)$

DIAGRAM 31

$$h(x) \leq f(x) \leq g(x).$$

Then $\lim_{x\to a} f(x) = \ell$.

Example: $\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right) = 0$.

Proof: For $x \neq 0$, $-1 \leq \sin\left(\frac{1}{x}\right) \leq 1 \Rightarrow -x^2 \leq x^2 \sin\left(\frac{1}{x}\right) \leq x^2$. Since $\lim_{x\to 0} x^2 = \lim_{x\to 0}(-x^2) = 0$, squeeze principle $\Rightarrow$

$$\lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right) = 0.$$

DIAGRAM 32

Limit as $x \longrightarrow \pm\infty$

Write

$$\lim_{x\to\infty} f(x) = \ell \quad , \quad \ell \in \mathbb{R}$$

if $f(x)$ approaches ℓ as x gets larger and larger; ie. $f(x)$ is close to ℓ for x sufficiently large. Similarly we write

$$\lim_{x\to -\infty} f(x) = \ell \quad , \quad \ell \in \mathbb{R}$$

if $f(x)$ approaches ℓ as x gets more and more negative.

Some Basic limits:

$$\lim_{x\to\infty}\frac{1}{x}=\lim_{x\to-\infty}\frac{1}{x}=0$$

DIAGRAM 33

DIAGRAM 34

$$\lim_{x\to\infty} e^{-x} = 0$$
$$\lim_{x\to-\infty} e^{x} = 0$$

Examples: (1) Find $\lim_{x\to\infty}\frac{2x^2+3}{3x^2+x}$

Solution: Divide top and bottom by highest power of x in denominator (so new denominator $\longrightarrow$ finite limit) – in this case x^2:

$$\therefore \quad \lim_{x\to\infty}\frac{2x^2+3}{3x^2+x}=\lim_{x\to\infty}\frac{2+\frac{3}{x^2}}{3+\frac{1}{x}}=\frac{2+0}{3+0}=\frac{2}{3}.$$

(2) $$\lim_{x\to\infty}\frac{x^2+5}{x+1}=\lim_{x\to\infty}\frac{x+\frac{5}{x}}{1+\frac{1}{x}}$$

Here denominator $\longrightarrow 1$ while numerator gets larger and larger therefore limit doesn't exist.

(3) $\lim_{x\to\infty}\sin x$ doesn't exist.

DIAGRAM 35

CONTINUITY:

Even if $\lim_{x\to a} f(x)$, $f(a)$ are both defined it is possible to have $\lim_{x\to a} f(x)\neq f(a)$.

Example:

$$f(x)=\begin{cases} x & , \quad x\neq 0 \\ 1 & , \quad x=0. \end{cases}$$

DIAGRAM 36

Here $f(0)=1$, but
$\lim_{x\to 0} f(x)=\lim_{x\to 0} x=0$. We say that f has a discontinuity at $x=0$. More precisely

Definition: Say that a function is *continuous* at a if

(i) $a\in$ domain f and (ii) $\lim_{x\to a} f(x)=f(a)$.

Examples: (1) Function $f(x)$ of previous example is not continuous at $x = 0$ since $\lim_{x \to 0} f(x) = 0 \neq f(0)$.

(2) $f(x) = \dfrac{1}{x^2}$ is not continuous at $x = 0$, since $0 \notin$ dom. f.

(3) $f(x) = \begin{cases} x+1 & , \quad x \geq 0 \\ x^2 & , \quad x < 0 \end{cases}$ DIAGRAM 37

is not continuous at $x = 0$ since $\lim_{x \to 0} f(x)$ doesn't exist.

Definition: f is continuous on the *open* interval (a, b) means: f is continuous at c, for all $c \in (a, b)$. f is continuous on the *closed* interval $[a, b]$ means f is continuous on (a, b) *and*

$$\lim_{x \to a^+} f(x) = f(a) \quad , \quad \lim_{x \to b^-} f(x) = f(b)$$

NOTE: If $\lim_{x \to a^+} f(x) = f(a)$, say f, is continuous from above (or from the right) at a. Similarly if $\lim_{x \to b^-} f(x) = f(b)$, say f is continuous from below (or from the left) at b.

Examples: (1) $f(x)$ a polynomial in x is continuous on $\mathbb{R}$ eg. $f(x) = ax^2 + bx + c$ is continuous on $\mathbb{R}$

(2) $e^{\pm x}$, $\sin x$, $\cos x$, $\operatorname{artan} x$ are continuous on $\mathbb{R}$

(3) $f(x) = \ln x$ is continuous on $(0, \infty)$

(4)

DIAGRAM 38 $\quad f(x) = \sqrt{x}$ is cont. on $[0, \infty)$. $\quad$ DIAGRAM 39

Properties: If $f(x)$, $g(x)$ are continuous at $x = a$, then also

$$f(x) \pm g(x) \quad , \quad f(x)g(x) \quad , \quad f(x)/g(x)$$

are continuous at $x = a$ (provided $g(a) \neq 0$ in the last).
Also

Theorem: If $\lim_{x \to a} g(x) = b$ and f is continuous at $x = b$, then

$$\lim_{x \to a} f(g(x)) = f(b).$$

Similarly if $\lim_{x\to\infty} g(x) = b$ and f is continuous at $x = b$, then

$$\lim_{n\to\infty} f(g(x)) = f(b).$$

Corollary: If g is continuous at a and f is continuous at $g(a)$ then $f \circ g$ is continuous at a.

Proof: Since $\lim_{x\to a} g(x) = g(a)$, Theorem $\Rightarrow \lim_{x\to a} f(g(x)) = f(g(a))$ therefore $(f \circ g)(x) = f(g(x))$ is continuous at $x = a$.

Examples: Using above properties we have

(1) $f(x) = p(x)/q(x)$ – quotient of two polynomials, is continuous everwhere *except* where $q(x) = 0$.

(2) $h(x) = \sqrt{x^2 + 3x + 1}$ is continuous on its domain since $f(x) = \sqrt{x}$, $g(x) = x^2 + 3x + 1$ are continuous and $h(x) = f(g(x))$ etc.

Differentiation

Given curve $y = f(x)$, want slope of tangent at some value $x = a$. Approx. tangent at A by chord AB where B is a pt. on curve close to A – with x value $a + h$, where h is small. If Δy is the corresponding change in y-value,have

DIAGRAM 40

$$\begin{aligned} \Delta y &= f(a+h) - f(a) \\ \therefore \quad \text{Slope of Chord} &= \frac{\Delta y}{\Delta x} = \tfrac{f(a+h)-f(a)}{h} \end{aligned}$$

and as $h \to 0$, Chord $\to$ tangent.

Definition: The *derivative* of f at x, is defined by

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \text{ – gives slope of tangent at } x$$

and we say that f is *differentiable* at x if this limit exists.

NOTE: This defines a new function f' from f – called the *derivative of* f. Often use the Leibnitz notation $\dfrac{dy}{dx}$ or $\dfrac{df}{dx}$ in place of $f'(x)$.

(Some books use D for differentiation, so $Df(x) = f'(x)$).

Example: (1) Let $f(x) = e^x$. Then $f'(x) = e^x$.

Proof: First e^x is defined so that slope of tangent at $x = 0$ equals 1; ie.

$$1 = \lim_{h\to 0} \frac{e^{h+0} - e^0}{h} = \lim_{h\to 0} \frac{e^h - 1}{h} \text{ – actually defines } e.$$

$$\begin{aligned} \therefore \quad f'(x) &= \lim_{h\to 0} \tfrac{e^{x+h} - e^x}{h} &&= \lim_{h\to 0} \tfrac{e^x \cdot e^h - e^x}{h} \\ &= \lim_{h\to 0} \tfrac{e^x(e^h - 1)}{h} &&= e^x \cdot \lim_{h\to 0} \tfrac{e^h - 1}{h} \\ & &&= e^x. \end{aligned}$$

DIAGRAM 41

(2) Other useful derivatives are

$$\begin{aligned} &\tfrac{d}{dx}(x^\alpha) = \alpha x^{\alpha-1} \quad , \quad \alpha \text{ any real number} \\ &\tfrac{d}{dx}(\sin x) = \cos x \quad , \quad \tfrac{d}{dx}(\cos x) = -\sin x \end{aligned}$$

Theorem: If $f(x)$ is differentiable at $x = a$ then $f(x)$ is continuous at a.

NOTE: However can have f continuous at a without f being differentiable at a.

Example: $f(x) = |x|$ is continuous at $x =$ 0 but is not differentiable at $x = 0$ since

DIAGRAM 42

$$f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{|h| - 0}{h} = \lim_{h \to 0} \frac{|h|}{h} \quad \text{– doesn't exist.}$$

RATES OF CHANGE

Recall that the slope of tangent to $y = f(x)$ at a point measures the *rate of change* of y w.r.t. x at that point.

Example: *Velocity* $v =$ rate at which distance s changes with time t, so

$$v = \frac{ds}{dt} = s'(t).$$

Similarly *acceleration* $a =$ rate at which velocity v changes with time, so

$$a = \frac{dv}{dt} = v'(t).$$

Higher Derivatives

Differentiating $y = f(x)$ n times gives the n^{th} *derivative* of f – denoted $f^{(n)}(x)$ or $\dfrac{d^n f}{dx^n}$ or $\dfrac{d^n y}{dx^n}$.

Thus eg., for acceleration we have

$$a = \frac{dv}{dt} = \frac{d}{dt}\left(\frac{ds}{dt}\right) = \frac{d^2 s}{dt^2} = s''(t) = s^{(2)}(t).$$

Rules for Differentiation

Suppose $f(x)$, $g(x)$ are differentiable functions. Then

(i) $\dfrac{d}{dx}((cf(x)) = c\dfrac{df}{dx}$ or $(cf)' = cf'$, $c = a$ const.

(ii) $(f \pm g)' = f' \pm g'$

(iii) $(fg)' = f'g + g'f$

(iv) $(f/g)' = \dfrac{gf' - fg'}{g^2}$. In particular $\dfrac{d}{dx}\left(\dfrac{1}{g}\right) = \dfrac{-g'}{g^2}$.

Example:

$$\tan = \frac{\sin x}{\cos x} = \frac{f}{g}$$
$$\Rightarrow \frac{d}{dx}(\tan x) = \frac{gf' - f'g}{g^2} = \frac{\cos x \sin'(x) - \sin x \cos' x}{\cos^2 x}$$
$$= (\cos^2 x + \sin^2 x)/\cos^2 x = \frac{1}{\cos^2 x} = \sec^2 x.$$

CHAIN RULE: Suppose $g(x)$ is diff. at $x = a$ and $f(x)$ is diff. at $b = g(a)$. Then $y = f(g(x))$ is diff. at $x = a$ and

$$y'(a) = f'(b)g'(a).$$

Thus we may write

$$\frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x).$$

Put another way, if $y = f(u)$, where $u = g(x)$, then $\frac{dy}{dx} = f'(u)g'(x)$ or

$$\boxed{\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}} \qquad - \qquad \text{Chain Rule.}$$

Examples: (1) Vel. $v = \frac{ds}{dt}$, $s =$ distance $\therefore$ by Chain Rule acceleration $a = \frac{dv}{dt} = \frac{dv}{ds} \cdot \frac{ds}{dt} = v\frac{dv}{ds}$.

(2) If $y = \cos^3 x$, find $\frac{dy}{dx}$.

Solution: Have $y = u^3$, where $u = \cos x$ $\therefore$ Chain Rule $\Rightarrow$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 3u^2 \cdot (-\sin x) = -3\cos^2 x \sin x.$$

(3) A rocket is launched vertically and tracked by an observation station 5 km from launch pad. The angle of elevation θ is observed to be increasing at $3°$/sec when $\theta = 60°$. What is the velocity of the rocket at that instant?

Solution: 1st convert to radians:
At $\theta = 60° = \frac{\pi}{3}$ rad., $\frac{d\theta}{dt} = 3°/\sec = \frac{3\pi}{180}$ rad./sec DIAGRAM 43

To find $\frac{dy}{dt}$,

have $y = 5\tan\theta$, with θ an unknown function of t. But can find $\dfrac{dy}{dt}$ using chain rule:

$$\frac{dy}{dt} = \frac{dy}{d\theta}\cdot\frac{d\theta}{dt} = 5\sec^2\theta\cdot\frac{d\theta}{dt} \quad \therefore \text{ at } \theta = \frac{\pi}{3} \text{ obtain}$$
$$\frac{dy}{dt} = 5\sec^2(\frac{\pi}{3})\cdot\frac{3\pi}{180} = 5\times 4\times\frac{3\pi}{180} = \frac{\pi}{3}\ \text{km/sec.}$$

Derivatives of Inverse Functions

Suppose $y = f^{-1}(x)$, where f^{-1} is the inverse of f. To obtain $\dfrac{dy}{dx}$ we use

$$x = f(f^{-1}(x)) = f(y).$$

Diff. both sides w.r.t. x using chain rule gives

$$1 = \frac{df(y)}{dx} = \frac{df}{dy}\cdot\frac{dy}{dx} = \frac{dx}{dy}\cdot\frac{dy}{dx}$$
$$\Rightarrow \frac{dy}{dx} = 1/\frac{dx}{dy}.$$

Examples: (1) To find deriv. of $y = \ln x$ have $x = e^y$

$$\Rightarrow \frac{dy}{dx} = 1/\frac{dx}{dy} = \frac{1}{e^y} = \frac{1}{x}, \qquad \text{so}$$
$$\frac{d(\ln x)}{dx} = \frac{1}{x}.$$

(2) To find deriv. of $y = \text{ arsin}\, x$ have $x = \sin y$, so

$$\frac{dy}{dx} = 1/\frac{dx}{dy} = 1/\cos y = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}$$
$$\Rightarrow \frac{d}{dx}(\text{ arsin}\, x) = \frac{1}{\sqrt{1-x^2}} \quad , \quad x\in(-1,1).$$

(3) Other useful derivatives are

$$\frac{d}{dx}(\text{ arcos}\, x) = -\frac{1}{\sqrt{1-x^2}} \quad , \quad x\in(-1,1)$$
$$\frac{d}{dx}(\text{ artan}\, x) = \frac{1}{1+x^2} \quad , \quad x\in\mathbb{R}.$$

L'Hôpitals rule for $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}\left(\dfrac{0}{0}\text{ or }\dfrac{\infty}{\infty}\right)$.

Using Taylor series (see later) can prove

Theorem: Assume $f'(x)$, $g'(x)$ exist and $g'(x)\neq 0$ in an interval about $x = a$, except possibly $x = a$, and suppose either

(i) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$

or

(ii) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \pm\infty$.

If $\lim_{x\to a} \dfrac{f'(x)}{g'(x)} = \ell \in \mathbb{R}$, then $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \ell$

– same applies to one-sided limits and $\lim_{x\to\pm\infty}$.

Examples: (1)

$$\begin{aligned} \lim_{x\to 0} \frac{\sin x}{x} &= \lim_{x\to 0} \frac{\cos x}{1} \qquad (\text{L'Hôp } \frac{0}{0}) \\ &= 1. \end{aligned}$$

(2)

$$\begin{aligned} \lim_{x\to 0} x \ln x &= \lim_{x\to 0} \frac{\ln x}{(1/x)} \\ &= \lim_{x\to 0} \frac{(1/x)}{(-1/x^2)} \qquad (\text{L'Hôp } \frac{\infty}{\infty}) \\ &= \lim_{x\to 0} (-x) = 0. \end{aligned}$$

(3)

$$\begin{aligned} \lim_{x\to\infty} xe^{-x} &= \lim_{x\to\infty} \frac{x}{e^x} \\ &= \lim_{x\to\infty} \frac{1}{e^x} \qquad (\text{L'Hôp } \frac{\infty}{\infty}) \\ &= 0. \end{aligned}$$

SEQUENCES

A *sequence* is an infinite list of numbers $a_0, a_1, a_2, a_3, \ldots$.

Examples:

(1) sequence $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\ldots$.

(2) sequence 2, 0, 2, 0, 2, ...

Formally a sequence is a function with domain given by the natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots.\}$.

NOTE: Some texts take $\mathbb{N} = \{1, 2, 3, ...\}$ and start sequence at a_1 rather than a_0 – not important to results below.

If a is a sequence, call a_n the n th **term**. May also write $\{a_n\}_{n=0}^{\infty}$ for a sequence a.

Example: The sequence of (1) above corresponds to

$$a_n = \frac{1}{n+1} \quad ; \quad n = 0, 1, 2 .... \quad (\text{or } a_n = \frac{1}{n},\ n = 1, 2, ...)$$

while sequence of (2) above corresponds to

$$a_n = 1 + (-1)^n.$$

Limits: Let $\{a_n\}_{n=0}^{\infty}$ be a sequence. Then

$$\lim_{n\to\infty} a_n = \ell \quad , \quad \ell \in \mathbb{R}, \qquad \text{means} \ :$$

a_n approaches ℓ as n gets larger and larger – ie. a_n is close to ℓ for n sufficiently large.

NOTE: This is the same as $\lim_{x\to\infty} f(x) = \ell$, except here a_n is defined only for natural numbers n.

Examples: (1) $\lim_{n\to\infty} \frac{3n^2+5}{n^2+n+1} = \lim_{n\to\infty} \frac{3+5/n^2}{1+1/n+1/n^2} = \frac{3+0}{1+0} = 3.$

(2) Consider sequence $\{(-1)^n\}_{n=0}^{\infty}$ ie. 1, −1, 1, −1, ... Then $\lim_{n\to\infty}(-1)^n$ does not exist.

CONVENTION: If sequence $\{a_n\}_{n=0}^{\infty}$ has limit $\ell \in \mathbb{R}$, say that a_n *converges* to ℓ and that sequence $\{a_n\}_{n=0}^{\infty}$ is *convergent* – otherwise sequence is *divergent*.

Example: Thus sequence $\left\{\frac{3n^2+5}{n^2+n+1}\right\}$ converges to 3 but sequence $\{(-1)^n\}$ is divergent.

Theorem: If $\lim_{x\to\infty} f(x) = \ell$

DIAGRAM 44

and $a_n = f(n)$, $n \in \mathbb{N}$, then

$$\lim_{n\to\infty} a_n = \ell.$$

Example: Evaluate $\lim_{n\to\infty} \frac{\ln(n+1)}{n+1}$.

Solution: $\lim_{x\to\infty} \frac{\ln(x+1)}{(x+1)} = \lim_{x\to\infty} \frac{1/(x+1)}{1} \quad (\text{L'Hôp } \frac{\infty}{\infty}) = 0$

$\therefore$ Theorem $\Rightarrow \lim_{n\to\infty} \frac{\ln(n+1)}{n+1} = 0$ – ie. sequence converges to 0.

Limit Theorems: Same limit Theorems hold as for functions. Thus if $\{a_n\}$, $\{b_n\}$ are convergent sequences and $\lim_{n\to\infty} a_n = \ell$, $\lim_{n\to\infty} b_n = m$ then

(i) $\lim_{n\to\infty} (a_n \pm b_n) = \ell \pm m$

(ii) $\lim_{n\to\infty} (a_n\, b_n) = \ell m$. In particular
$\lim_{n\to\infty} (c\, a_n) = c.\ell$, c const.

(iii) $\lim_{n\to\infty} (a_n/b_n) = \ell/m$, provided $m \neq 0$.

Also Squeeze Principle holds: if $b_n \le a_n \le c_n$, for all n, and

$$\begin{aligned} \lim_{n\to\infty} b_n &= \lim_{n\to\infty} c_n = \ell \in \mathbb{R}, \qquad \text{then} \\ \lim_{n\to\infty} a_n &= \ell. \end{aligned}$$

Example: Find $\lim_{n\to\infty} \dfrac{\sin(n)}{\sqrt{n+1}}$

Solution: Since $-1 \le \sin(n) \le 1$, have

$$-\frac{1}{\sqrt{n+1}} \le \frac{\sin(n)}{\sqrt{n+1}} \le \frac{1}{\sqrt{n+1}}, \qquad \text{for all } n \in \mathbb{N}.$$

But

$$\lim_{n\to\infty} \frac{1}{\sqrt{n+1}} = \lim_{n\to\infty} \left(-\frac{1}{\sqrt{n+1}}\right) = 0$$

$\therefore$ Squeeze Principle $\Rightarrow \lim_{n\to\infty} \dfrac{\sin(n)}{\sqrt{n+1}} = 0$.

Useful sequences to remember

(1) For constant c, $\lim_{n\to\infty} c^n = \begin{cases} 0, & \text{if } |c| < 1 \\ 1, & \text{if } c = 1. \end{cases}$

Sequence $\{c^n\}_{n=0}^{\infty}$ is divergent if $c = -1$ or $|c| > 1$.

(2) For constant $c > 0$, $\lim_{n\to\infty} c^{1/n} = 1$.

(3) Even $\lim_{n\to\infty} n^{1/n} = 1$

(4) For constant c, $\lim_{n\to\infty} \dfrac{c^n}{n!} = 0$ (recall: $n! = 1 \times 2 \times 3 \times .... \times n$).
$0! = 1$

Example: Evaluate $\lim_{n\to\infty}(2^n+1)^{1/n}$

Solution:

$$\begin{aligned} & 2^n \le 2^n+1 \le 2\cdot 2^n \quad , \quad \text{for all } n\in\mathbb{N} \\ \Rightarrow\ & 2 \le (2^n+1)^{1/n} \le 2\cdot 2^{1/n}. \end{aligned}$$

By (2) above

$$\lim_{n\to\infty} 2\cdot 2^{1/n} = 2\lim_{n\to\infty} 2^{1/n} = 2$$

$\therefore$ squeeze principle $\Rightarrow \lim_{n\to\infty}(2^n+1)^{1/n} = 2.$

SIGMA NOTATION

As a shorthand notation for sum $a_0+a_1+a_2+\cdots+a_n$ we write

$$\sum_{i=0}^{n} a_i \qquad \left(\sum \quad \text{stands for sum}\right).$$

Move generally, for $m<n$, we write

$$\sum_{i=m}^{n} a_i \quad \text{for } a_m + a_{m+1} + \cdots + a_n$$

– important for series.

NOTE: Any other letter is as good as i: eg.

$$\sum_{i=1}^{3} i^3 = 1^3+2^3+3^3 = \sum_{j=1}^{3} j^3 \qquad \text{etc.}$$

SERIES:

The sum of a sequence $\{a_n\}_{n=0}$ is called a *series*, written

$$\sum_{n=0}^{\infty} a_n = a_0+a_1+a_2+\cdots\cdots$$

To see whether such a sum converges we consider the sequence of *partial sums*:

$$s_n = \sum_{k=0}^{n} a_k = a_0+a_1+\cdots+a_n.$$

Then

$$\sum_{n=0}^{\infty} a_n = \lim_{n\to\infty} s_n$$

and we say series *converges* if sequence $\{s_n\}$ converges.

Examples: (1) $\sum_{n=0}^{\infty}(-1)^n = 1 - 1 + 1 - 1 + 1 \cdots\cdots$

Partial sums are

$s_0 = 1,\ s_1 = 0,\ s_2 = 1,\ s_3 = 0$ – oscillates between 0 & 1 $\therefore$ series diverges.

(2) Consider series $\sum_{n=0}^{\infty} \frac{1}{(n+1)(n+2)} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \cdots$

Now

$$\frac{1}{(n+1)(n+2)} = \frac{1}{n+1} - \frac{1}{n+2}$$

$$\Rightarrow s_n = \sum_{k=0}^{n} \frac{1}{(k+1)(k+2)} = \sum_{k=0}^{n} \left(\frac{1}{k+1} - \frac{1}{k+2} \right)$$

$$= (1 - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + (\frac{1}{3} - \frac{1}{4}) + \cdots\cdots + (\frac{1}{n+1} - \frac{1}{n+2}) \begin{pmatrix} \text{telescopic} \\ \text{cancellation} \end{pmatrix}$$

$$= 1 - \frac{1}{n+2}$$

$$\therefore \quad \sum_{n=0}^{\infty} \frac{1}{(n+1)(n+2)} = \lim_{n\to\infty} (1 - \frac{1}{n+2}) = 1.$$

NOTE: Above series may also be written

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \cdots$$

Example (p-series):

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \sum_{n=0}^{\infty} \frac{1}{(n+1)^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \cdots\cdots,\ p > 0$$

– converges for $p > 1$ but diverges for $p \leq 1$.

NOTE: A proof of this requires integration – see later.

A useful test for convergence is

n^{th}**–Term test**:

If series $\sum_{n=0}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Proof: Let $s_n = \sum_{k=0}^{n} a_k$ be sequence of partial sums and suppose $\lim_{n\to\infty} s_n = \ell \in \mathbb{R}$. Then $a_n = s_n - s_{n-1} \Rightarrow \lim_{n\to\infty} a_n = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = \ell - \ell = 0$.

NOTES: (1) Above applies also if series starts at $n = 1$ (or 2, 3 etc.).

(2) Condition $\lim_{n\to\infty} a_n = 0$ is *necessary* for series $\sum_{n=0}^{\infty} a_n$ to converge, but *not* sufficient.

eg.

$$\lim_{n\to\infty} \frac{1}{n} = 0$$

but series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots\cdots$$

is known to diverge (p-series with $p = 1$).

Examples: (1) $\sum_{n=0}^{\infty}(-1)^n$: $\lim_{n\to\infty}(-1)^n \neq 0$ (in fact limit doesn't exist)

$\therefore$ series divergent as seen previously.

(2) $$\sum_{n=1}^{\infty} \frac{n}{2n+1}: \lim_{n\to\infty} \frac{n}{2n+1} = \lim_{n\to\infty} \frac{1}{2+1/n} = \frac{1}{2} \neq 0$$

$\therefore$ series diverges.

Geometric Series: Consider series

$$\sum_{n=0}^{\infty} ax^n = a + ax + ax^2 + \cdots\cdots .$$

Sequence of partial sums is given by

$$\begin{aligned} S_n &= a + ax + \cdots\cdots + ax^n \\ \Rightarrow xS_n &= ax + ax^2 + \cdots\cdots + ax^n + ax^{n+1} \\ \therefore \quad S_n - xS_n &= a - ax^{n+1} \Rightarrow (1-x)S_n = a(1-x^{n+1}) \\ \therefore \quad S_n &= \frac{a(1-x^{n+1})}{1-x} \\ \Rightarrow \sum_{n=0}^{\infty} ax^n &= \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a(1-x^{n+1})}{1-x}. \end{aligned}$$

Hence

(1) If $|x| < 1$, $\sum_{n=0}^{\infty} ax^n = \frac{a}{1-x}$ $\quad\therefore$ converges

(2) If $|x| \geq 1$, $\lim_{n\to\infty} ax^{n-1} \neq 0$ $\quad\therefore$ series diverges by n^{th} term test.

NOTE: In the special case $a = 1$ obtain

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x} \quad , \quad \text{for } |x| < 1.$$

Examples: (1) $\displaystyle\sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n = \frac{1}{1-1/2} = 2 \quad (x=\tfrac{1}{2}<1)$

(2) $\displaystyle\sum_{n=0}^{\infty} 5\cdot\left(\frac{2}{3}\right)^n = \frac{5}{1-2/3} = 15 \quad (a=5,\ x=2/3).$

□

CONVERGENCE THEOREM

Suppose series $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ are convergent and

$$\sum_{n=0}^{\infty} a_n = \ell \quad , \quad \sum_{n=0}^{\infty} b_n = m.$$

Then series

$$\sum_{n=0}^{\infty}(a_n \pm b_n) \quad , \quad \sum_{n=0}^{\infty} c\cdot a_n \qquad (c = \text{ const.})$$

are convergent and

$$\sum_{n=0}^{\infty}(a_n \pm b_n) = \ell \pm m \quad , \quad \sum_{n=0}^{\infty} c\cdot a_n = c\ell.$$

Example: Evaluate $\sum_{n=0}^{\infty} x^n(2x+1)$, for $|x|<1$.

Proof: $\displaystyle\sum_{n=0}^{\infty} x^n(2x+1) = \sum_{n=0}^{\infty} 2x\cdot x^n + \sum_{n=0}^{\infty} x^n.$

Now $\sum_{n=0}^{\infty} 2x\cdot x^n$, $\sum_{n=0}^{\infty} x^n$ are Geom. Series Corresp. to $a=2x$, 1 resp.
$\therefore$ converge for $|x|<1$. Hence Theorem $\Rightarrow$ given series converges and

$$\begin{aligned}\sum_{n=0}^{\infty} x^n(2x+1) = \sum_{n=0}^{\infty} 2x\cdot x^n + \sum_{n=0}^{n} x^n &= \frac{2x}{1-x} + \frac{1}{1-x} \\ &= \frac{2x+1}{1-x}\ , \ \text{for } |x|<1.\end{aligned}$$

Comparison Test:

Frequently series are too difficult to evaluate directly. Here we give a useful test for convergence of a series. First note that if $\sum_{n=0}^{\infty} a_n$ is a series and $k \geq 0$ is a natural number then

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + \cdots\cdots a_{k-1} + \sum_{n=k}^{\infty} a_n$$

$\therefore$ Given series converges iff series

$$\sum_{n=k}^{\infty} a_n = a_k + a_{k+1} + \cdots\cdots \qquad \text{converges.}$$

FACT: For $k \geq 0$, consider series $\sum_{n=0}^{\infty} a_n$ and suppose $\sum_{n=k}^{\infty} b_n$ is a series of positive terms such that

$$|a_n| \leq b_n \quad , \quad \text{for all } n \geq k.$$

If $\sum_{n=k}^{n} b_n$ converges, so too does $\sum_{n=0}^{\infty} a_n$.

Examples: Determine whether following series converge.

$$\text{(i)} \quad \sum_{n=0}^{\infty} \frac{(-1)^n}{\pi + n^{3/2}} \qquad\qquad \text{(ii)} \quad \sum_{n=0}^{\infty} \frac{1}{\sqrt{n^2 - n + 1}}$$

Solution: (i) $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{\pi + n^{3/2}} = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{\pi + n^{3/2}}$.

Now $\left|\dfrac{(-1)^n}{\pi + n^{3/2}}\right| = \dfrac{1}{\pi + n^{3/2}} < \dfrac{1}{n^{3/2}}$, for all $n \geq 1$ and series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{3/2}}$ converges (p series with $p = 3/2 > 1$).

$\therefore$ By comparison test, given series must converge.

(ii) $\displaystyle\sum_{n=0}^{\infty} \frac{1}{\sqrt{n^2 - n + 1}} = 1 + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n^2 + 1 - n}}$ – series of (+ ve) terms.

Now $\dfrac{1}{\sqrt{n^2 + 1 - n}} \geq \dfrac{1}{\sqrt{n^2}} = \dfrac{1}{n}$, for all $n \geq 1$ and series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ diverges ($p = 1$ case of p-series)

$\therefore$ Given series must diverge.

Absolute convergence

DEFINITION: A series $\sum_{n=0}^{\infty}$ is said to be absolutely convergent if series $\sum_{n=0}^{\infty} |a_n|$ is convergent.

FACT: If $\sum_{n=0}^{\infty} a_n$ is absolutely convergent, then it is convergent.

Proof: Clearly $|a_n| \leq b_n = |a_n|$, $\forall n \geq 0$, and since $\sum_{n=0}^{\infty} |a_n|$ converges, Comparison Test $\Rightarrow \sum_{n=0}^{\infty} a_n$ converges.

NOTE: However it is possible the series $\sum_{n=0}^{\infty} a_n$ is convergent but not absolutely – such a series is said to be *conditionally convergent*.

Examples: (1) Can be shown that $\sum_{n=0}^{\infty} \frac{(-1)^n}{n}$ is convergent but not absolutely convergent (c.f. $p = 1$ case of p series) $\therefore$ conditionally convergent.

(2) Geom. Series $\sum_{n=0}^{\infty} ax^n$ is absolutely convergent for $|x| < 1$.

NOTE: In this course we are mainly concerned with series which are absolutely convergent.

□

The following is the main test for absolute convergence

Ratio Test: Let $\sum_{n=0}^{\infty} a_n$ be a series and suppose

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \ell \geq 0. \qquad \text{Then}$$

(i) $\ell < 1 \Rightarrow$ series converges absolutely.

(ii) $\ell > 1 \Rightarrow$ series diverges.

(iii) $\ell = 1$ – test fails and anything can happen.

Proof: Follows by comparing with Geometric Series.

NOTE:
$\therefore$ In applications, usually concentrate on the n^{th} term test and the Ratio Test for convergence – only use Comparison Test if Ratio Test fails.

Examples: Determine whether the following series converge:

$$\text{(i)} \quad \sum_{n=0}^{\infty} \frac{(-2)^n}{n!} \qquad \text{(ii)} \quad \sum_{n=0}^{\infty} n \left(\frac{1}{2}\right)^n \qquad \text{(recall : } 0! = 1)$$

Solution: (i) Here $a_n = \dfrac{(-1)^n 2^n}{n!}$ $\therefore$ Using Ratio Test

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{2^{n+1}}{(n+1)!} \frac{n!}{2^n} = \lim_{n\to\infty} \frac{2}{n+1} = 0 < 1.$$

$\therefore$ By Ratio Test, given series converges (absolutely).

(ii) Here $a_n = n\left(\dfrac{1}{2}\right)^n$ $\therefore$ Using Ratio Test

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{(n+1)(1/2)^{n+1}}{n(1/2)^n} = \frac{1}{2} \lim_{n\to\infty} \left(\frac{n+1}{n}\right) = \frac{1}{2} < 1.$$

$\therefore$ By Ratio Test, given series converges.

Power Series

If $\{c_n\}_{n=0}^{\infty}$ is a sequence of real numbers, then a series of form

$$\sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + \cdots\cdots$$

is called a *power series* in x.

Examples: (1) Obviously any polynomial in x

$$p(x) = c_0 + c_1 + \cdots\cdots + c_k x^k = \sum_{n=0}^{k} c_n x^n$$

is a power series – in this case $c_n = 0$, for $n > k$.

(2) Examples of infinite power series are

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots \qquad \text{– Geom. series}$$
$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

More generally, for any $a \in \mathbb{R}$, may consider power series of the form

$$\sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots\cdots$$

Such series have many applications eg. solution of higher order ODEs with non-const. coeffs, etc.

In general, however, such a series does not converge for all x.

Theorem: There exists $r \geq 0$, called the *radius of convergence*, such that

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

converges absolutely for $|x-a| < r$ and diverges for $|x-a| > r$. The series may or may not converge for $|x-a| = r$.

NOTE: If series converges for all x say radius of convergence is $r = \infty$. It is also possible to have $r = 0$ – then series diverges for all x except $x = a$.

FACT: The series $\sum_{n=0}^{\infty} c_n (x-a)^n$ has radius of convergence

$$r = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}.$$

If $|c_n| / |c_{n+1}| \to \infty$, as $n \to \infty$, the radius of convergence is $r = \infty$ (ie. series converges for all x).

Proof: nth term of series is $a_n = c_n (x-a)^n$

$$\begin{aligned} \Rightarrow \quad \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} &= \lim_{n\to\infty} \frac{|c_{n+1}| \cdot |(x-a)^{n+1}|}{|c_n| \cdot |(x-a)^n|} \\ &= |x-a| \cdot \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = |x-a| / r. \end{aligned}$$

$\therefore$ By Ratio Test, series converges absolutely for $|x-a|/r < 1$ and diverges for $|x-a|/r > 1$, which proves the result.

NOTE: Result does not apply when sequence $|c_n| / |c_{n+1}|$ is *not defined* – but may still get radius of convergence by Ratio Test (c.f. $\sin x$). eg. consider polynomial $\sum_{n=0}^{k} c_n (x-a)^n$ – obviously converges for all x but here $|c_n| / |c_{n+1}|$ is not defined for $n \geq k$.

Examples: (1) For Geom. Series $\sum_{n=0}^{\infty} x^n$, corresponding to $a = 1$, $c_n = 1$ for all n, have radius of convergence

$$r = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = 1 \quad , \qquad \text{as seen earlier.}$$

(2) For series $\sum_{n=0}^{\infty} \frac{(x-a)^n}{n!}$ have $c_n = \frac{1}{n!}$. In this case

$$\frac{|c_n|}{|c_{n+1}|} = \frac{(n+1)!}{n!} = n+1 \longrightarrow \infty \quad \text{as} \quad n \longrightarrow \infty.$$

$\therefore$ series converges absolutely for all x (ie. $r = \infty$).

Taylor Series:

Given a function $f(x)$ it is sometimes possible to expand $f(x)$ as a series

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \cdots.$$

– useful for summing series and many other applications.

To determine coefficients c_n, first note that

$$f(a) = c_0.$$

Now take derivatives of both sides $\Rightarrow$

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^3 + \cdots\cdots$$

so

$$f'(a) = c_1.$$

Take derivatives again and continue to give

$$\begin{aligned} f''(a) &= 2c_2 \Rightarrow c_2 = f''(a)/2 \\ f'''(a) &= 6c_3 \Rightarrow c_3 = f^{(3)}(a)/3! \end{aligned}$$

and, in general

$$c_n = f^{(n)}(a)/n! \quad , \quad \text{with } f^{(0)}(x) \equiv f(x).$$

$\therefore$ Expect $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$ – called *Taylor series* of $f(x)$ about $x = a$.

NOTE: Above equality does not hold in general. Obviously f and all of its derivatives must be defined at $x = a$ (for series to exist) – then say that f is *infinitely differentiable* at $x = a$.

Theorem: Suppose $f(x)$ is infinitely differentiable at $x = a$. Then for f "sufficiently well behaved",

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \quad , \quad \text{for } |x-a| < r$$

where r is the radius of convergence of series.

NOTES: (1) Define a sequence of numbers

$$b_n = \max_{x \in [-r,r]} f^{(n)}(x) \quad , \quad n \in \mathbb{N}.$$

Then "suff. well behaved" means

$$\lim_{n\to\infty} \frac{b_n r^n}{n!} = 0 \qquad \text{– satisfied by all functions in this course.}$$

(2) For $|x-a| < r$, have approx.

$$f(x) \simeq \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

which is useful for approximating values of functions – the smaller $|x-a|$ the better the approximatation. This approximation can be made accurate to arb. order for k suff. large – provided $|x-a| < r$.

Examples: (1) $f(x) = e^x \Rightarrow f^{(n)}(x) = e^x$, so

$$f^{(n)}(a) = e^a \quad , \quad \text{for all } n.$$

$\therefore$ Taylor series for e^x about $x = a$ is

$$e^a \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!}.$$

In previous example we have seen this series converges for all x, so

$$e^x = e^a \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} \quad , \quad \text{for all } x \in \mathbb{R}.$$

(2) $f(x) = \ln(x) \Rightarrow f'(x) = x^{-1}$, $f''(x) = -x^{-2}$, $f'''(x) = 2x^{-3}$, and, in general

$$f^{(n)}(x) = (n-1)!(-1)^{n-1}x^{n-1} \Rightarrow f^{(n)}(1) = (-1)^{n-1}(n-1)!\,, \quad n \geq 1$$

while $f^{(0)}(1) = f(1) = 0$ $\therefore$ Taylor series exp. of $\ln x$ about $x = 1$ is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(1)}{n!}(x-1)^n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n.$$

For radius of convergence, in this case have coeffs

$$c_n = (-1)^{n-1}/n$$

$$\Rightarrow \quad r = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = \lim_{n\to\infty} \left(\frac{n+1}{n}\right) = 1 \quad , \text{ so}$$

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(x-1)^n}{n} \quad , \text{for } |x-1| < 1.$$

In particular, setting $x = \frac{1}{2}$ gives

$$-\ln(2) = \ln(\frac{1}{2}) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(-\frac{1}{2})^n}{n} = -\sum_{n=1}^{\infty} \frac{1}{n2^n}$$

$$\Rightarrow \quad \sum_{n=1}^{\infty} \frac{1}{n2^n} = \ln(2).$$

This gives approx.

$$\ln(2) \simeq \sum_{n=1}^{4} \frac{1}{n2^4} = \frac{1}{2} + \frac{1}{8} + \frac{1}{24} + \frac{1}{64} \simeq \cdot 682.$$

NOTE: Even if only first k derivatives of $f(x)$ are defined at $x = a$, Taylor series still gives useful polynomial approx.

$$f(x) \simeq \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

– valid for x close to a; ie. $|x - a|$ suff. small.

MacLaurin Series

Taylor series expansion of a function about $x = 0$ is called a *MacLaurin series* for $f(x)$. Thus if $f(x)$ is infinitely diff. at $x = 0$, have MacLaurin series expansion

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \quad , \quad |x| < r$$

where r is the radius of convergence.

Examples: (1) Setting $a = 0$ into previous example for e^x gives MacLaurin series expansion

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad , \quad \text{for all } x \in \mathbb{R}.$$

eg. at $x = 1$, obtain

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots\cdots \text{ – defines } e.$$

This gives approx.

$$e \simeq 2 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} \simeq 2 \cdot 70833.$$

(2) (*binomial expansion*): Suppose $\alpha \in \mathbb{R}$ and consider $f(x) = (1+x)^\alpha \Rightarrow f'(x) = \alpha(1+x)^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2}$ and, in general,

$$\begin{aligned} f^{(n)}(x) &= \alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)(1+x)^{\alpha-n} \\ \Rightarrow \quad f(0) = 1\,, &\quad f^{(n)}(0) = \alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1) \;\; n \geq 1 \end{aligned}$$

$\therefore$ MacLaurin series expansion for $f(x)$ is

$$\sum_{n=0}^\infty \frac{f^{(n)}(0)}{n!} x^n = 1 + \sum_{n=1}^\infty \frac{\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)}{n!} x^n$$

$$\Rightarrow \quad \text{coeffs } c_0 = 1\,,\; c_n = \alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)/n!\,, \quad n \geq 1.$$

NOTES: For $\alpha = m \in \mathbb{N}$ have $f^{(n)}(x) = 0$ for $n > m$ and above series truncates to

$$\begin{aligned} (1+x)^m &= 1 + \sum_{n=1}^m \frac{m(m-1)\cdot\cdots\cdot(m-n+1)}{n!} x^n \\ &= \sum_{n=0}^m \frac{m!}{(m-n)!n!} x^n \quad \text{– usual binomial expansion} \end{aligned}$$

(valid for all x).

$\square$

For $\alpha \notin \mathbb{N}$, the radius of convergence is

$$\begin{aligned} r &= \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = \lim_{n\to\infty} \frac{|\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)|}{n!} \cdot \frac{(n+1)!}{|\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n)|} \\ &= \lim_{n\to\infty} \frac{n+1}{|\alpha-n|} = \lim_{n\to\infty} \frac{1+\frac{1}{n}}{\left|\frac{\alpha}{n}-1\right|} = 1, \end{aligned}$$

so

$$(1+x)^\alpha = 1 + \sum_{n=1}^\infty \frac{\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)}{n!} x^n \quad , \quad |x| < 1$$

valid for all α – diverges for $|x| > 1$ when $\alpha \notin \mathbb{N}$ (but converges for all x if $\alpha \in \mathbb{N}$).

Note: Replacing x with $-x$ gives

$$(1-x)^\alpha = 1 + \sum_{n=1}^\infty \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)(-1)^n x^n}{n!} \quad , \quad |x| < 1$$

so when $\alpha = -1$, this gives

$$(1-x)^{-1} = \sum_{n=0}^\infty x^n \quad , \quad |x| < 1$$

in agreement with geom. series.

(3) ($\sin x$): Consider $f(x) = \sin x \Rightarrow f'(x) = \cos x$, $f''(x) = -\sin x$, $f^{(3)}(x) = -\cos x$, $f^{(4)}(x) = \sin x$ etc. $\therefore$ In this case get $f^{(2n)}(0) = 0$, $f^{(2n+1)}(0) = (-1)^n$, $n = 0, 1, 2, ...$ so only *odd* terms survive, therefore MacLaurin series expansion is

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\cdots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.\end{aligned}$$

For radius of convergence, nth term in this case is

$$\begin{gathered} b_n = (-1)^n x^{2n+1}/(2n+1)! \\ \Rightarrow \lim_{n\to\infty} \frac{|b_{n+1}|}{|b_n|} = \lim_{n\to\infty} \frac{|x^{2n+3}|}{(2n+3)!} \cdot \frac{(2n+1)!}{|x^{2n+1}|} \\ = |x|^2 \lim_{n\to\infty} \frac{1}{(2n+3)(2n+3)} = 0 < 1\,, \quad \text{for all } x. \end{gathered}$$

$\therefore$ By Ratio Test series converges for all x (ie. $r = \infty$), so

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \quad , \quad \text{for all } x \in \mathbb{R}.$$

Thus, for x close to 0, have approx.

$$\sin x \simeq x - \frac{x^3}{3!} \qquad \text{(first 2 terms)}$$

eg. $\sin(\cdot 1) \simeq \cdot 1 - \frac{(\cdot 1)^3}{6} \simeq \cdot 0998$ – check! (**recall**: $\cdot 1 \equiv \cdot 1$ rad.)

NOTE: From above get $\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - = \frac{x^6}{7!} + \cdots$

$$\Rightarrow \lim_{x\to 0} \frac{\sin x}{x} = 1 \qquad \text{(c.f. L'Hôp. rule).}$$

(4) Similarly obtain following MacLaurin series expansion for $\cos x$ (Exercise):

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad , \quad \text{for all } x \in \mathbb{R}.$$

Monotonic Functions

DEFINITION: A function f called *strictly increasing* on an interval I if

$$f(x_1) < f(x_2) \quad , \quad \text{whenever } x_1 < x_2 \text{ in } I$$

while f is called *strictly decreasing* on I if

$$f(x_1) > f(x_2) \quad , \quad \text{whenever } x_1 < x_2 \text{ in } I$$

DIAGRAM 45 DIAGRAM 46

¿From graph, slope of tangent to curve at $x \in I$, given by $f'(x)$, is $+ve$ for f strictly increasing and $-ve$ for f strictly decreasing. We in fact have

Theorem: Suppose f is cont. on $[a, b]$ and *differentiable on* (a, b) [ie. diff. at x, for all $x \in (a, b)$]. Then

(i) $f'(x) > 0$, for all $x \in (a, b) \Rightarrow f$ strictly increasing on $[a, b]$

(ii) $f'(x) < 0$, for all $x \in (a, b) \Rightarrow f$ strictly decreasing on $[a, b]$

(iii) $f'(x) = 0$, for all $x \in (a, b) \Rightarrow f$ is const. on $[a, b]$ (ie. $f(x) = c$, for all $x \in [a, b]$).

EXAMPLES:(1)

DIAGRAM 47

$$\begin{aligned} &f(x) = x^3 + x = x(x^2+1) \\ \Rightarrow\ &f'(x) = 3x^2 + 1 > 0\,, \quad \text{for all } x \\ \therefore\ \ &f(x) \text{ is strictly increasing on } \mathbb{R} \end{aligned}$$

(2) $f(x) = e^x \Rightarrow f'(x) = e^x > 0$, for all x $\therefore f$ is strictly increasing on $\mathbb{R}$.

DIAGRAM 48

(3)

DIAGRAM 49

$$\begin{aligned} &f(x) = \ln x \quad , \quad x > 0 \\ \Rightarrow\ &f'(x) = \frac{1}{x} > 0 \quad , \quad \text{for } x > 0 \\ \therefore\ \ &f(x) \text{ is strictly increasing on } (0, \infty). \end{aligned}$$

(4)

DIAGRAM 50

$$\begin{aligned} f(x) &= \arccos x \\ \Rightarrow f'(x) &= \frac{-1}{\sqrt{1-x^2}} < 0 \quad , \quad \text{for all } x \in (-1,1) \\ \therefore \quad & f(x) \text{ strictly decreasing on } [-1,1]. \end{aligned}$$

Maxima and Minima

DEFINITION: Say f has a *relative* (or *local*) max. at $x = c$ if $f(c) \geq f(x)$ for all x near c; ie. if there is $\delta > 0$ s.t.

DIAGRAM 51

$$|x - c| < \delta \Rightarrow f(c) \geq f(x).$$

Similarly f has a relative (or local) min. at $x = c$ if $f(x) \geq f(c)$ for all x near c.

Theorem: Suppose f has relative max. (or min.) at $x = c$ and that f is diff. at c. Then $f'(c) = 0$.

Proof: Suppose eg. f has a rel. max. at $x = c$ and consider

DIAGRAM 52

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

For x near c with $x > c$, $f(x) \leq f(c)$ so

$$\frac{f(x) - f(c)}{x - c} \leq 0 \Rightarrow \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} \leq 0.$$

But for x near c with $x < c$, still have $f(x) \leq f(c)$ so

$$\frac{f(x) - f(c)}{x - c} \geq 0 \Rightarrow \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \geq 0.$$

Since f is diff. at $x = c$, must have

$$\lim_{x \to c^+} = \lim_{x \to c^-} = f'(c) \Rightarrow f'(c) = 0.$$

DIAGRAM 53

NOTES: (i) Can have $f'(c) = 0$, but neither max. or min. at c

eg. $f(x) = x^3$ at $x = 0$.

(ii) Can have rel. max. or min. at places where f is not diff. eg. $f(x) = |x|$ has a min. at $x = 0$.

DIAGRAM 54

TEST FOR RELATIVE MAX. OR MIN. (2nd Derivative Test)

Suppose f'' is cont. at $x = c$. Then

(i) $f'(c) = 0$ and $f''(c) < 0 \Rightarrow f$ has ref. max. at $x = c$

(ii) $f'(c) = 0$ and $f''(c) > 0 \Rightarrow f$ has ref. min. at $x = c$

(iii) $f'(c) = 0$ and $f''(c) = 0$ – test fails and anything can happen.

Proof: (outline): Consider Taylor series expansion to 2nd. order about $x = c$ (valid for x close to c):

$$\begin{aligned} f(x) &\simeq f(c) + f'(c)(x-c) + \frac{1}{2} f''(c)(x-c)^2 \\ &= f(c) + \frac{1}{2} f''(c)(x-c)^2 \end{aligned}$$

$$\begin{aligned} \therefore \quad & f''(c) > 0 \Rightarrow f(x) \geq f(c) \text{ for } x \text{ near } c \quad \therefore \text{ rel. min. at } c \\ & f''(c) < 0 \Rightarrow f(x) \leq f(c) \text{ for } x \text{ near } c \quad \therefore \text{ rel. max. at } c \end{aligned}$$

Examples: (1) Concerning (iii) consider

$$f(x) = x^4, \qquad g(x) = -x^4, \quad h(x) = x^3.$$

For each $f'(0) = f''(0) = 0$. But $f(x)$ has a rel. min., $g(x)$ a rel. max. and $h(x)$ neither a max. or min. at $x = 0$.

(2) Find the rel. max. and min. of

$$f(x) = x^4 - 2x^2$$

and hence sketch the graph of $y = f(x)$.

SOLUTION: $f'(x) = 4x^3 - 4x = 4x(x^2 - 1) = 4x(x-1)(x+1)$
$\therefore \quad f'(x) = 0$ at $x = 0, \pm 1$ – poss. rel. max. or min.

$$f''(x) = 12x^2 - 4$$

$\therefore \quad f''(0) = -4 < 0 \Rightarrow$ rel. max. at $x = 0$ (when $y = 0$)
$f''(\pm 1) = 8 > 0 \Rightarrow$ rel. min at $x = \pm 1$ (when $y = -1$).
Finally observe
$f(x) = x^2(x^2 - 2) = x^2(x - \sqrt{2})(x + \sqrt{2})$ – so crosses x-axis at $x = 0, \pm\sqrt{2}$. Hence graph shown.

DIAGRAM 55

(3) A manufacturer has a certain amount of material to construct a rectangular cardboard box with a square base and an open top. Find the dimensions of the box which maximize the volume.

SOLUTION: Let x be sidelength of base and y the height of the box. Therefore surface area a of box (a given const.) is

DIAGRAM 56

$$a = x^2 + 4xy \Rightarrow y = \frac{a - x^2}{4x}$$

therefore volume of box is
$V(x) = x^2 y = \frac{x}{4}(a - x^2)$.

To obtain max. vol. need graph of $V(x)$ – note that $V(x)$ crosses x– axis at $x = 0, \pm\sqrt{a}$. Now $V'(x) = \frac{a}{4} - \frac{3x^2}{4}$. Hence $V'(x) = 0$ at $x = \pm\sqrt{\frac{a}{3}}$ – poss. max. and min.

Using 2nd deriv. test have

$$V''(x) = -\frac{3x}{2} \quad \begin{array}{ll} < & 0 \text{ at } x = \sqrt{a/3} \to \text{ rel. max.} \\ > & \text{at } x = -\sqrt{a/3} \Rightarrow \text{ rel. min.} \end{array}$$

DIAGRAM 57

giving graph shown. Since we are only concerned with $x \geq 0$ it follows from graph that indeed max. vol. occurs for $x = \sqrt{a/3}$ – absolute max. on $[0, \infty)$.

$\therefore$ Box should be manfactured so that

$$x = \sqrt{\frac{a}{3}}, \quad y = \frac{a - x^2}{4x} = \frac{1}{2}\sqrt{\frac{a}{3}}$$

ie. for max volume require height to be half of base length.

□

The above problem can in fact be solved in a simpler way (without requiring a graph or the 2nd derivative test) if we use absolute max. and min.

Absolute Max. and Min.

Definition: c in the domain of f is called a *critical point* if f is not diff. at c or $f'(c) = 0$.

Critical points are important in max. and min. problems since if f has a rel. max. or min. at c then c must be a critical point of f.

Example: Find critical points of $f(x) = x\sqrt{x+1}$.

Solution: Domain $f = [-1, \infty)$ and $f'(x) = \sqrt{x+1} + \dfrac{x}{2\sqrt{x+1}} = \dfrac{(3x+2)}{2\sqrt{x+1}} = 0$ when $x = -2/3$, and undef. at $x = -1$ (both in domain of f) $\therefore x = -2/3,\ -1$ are crit. pts.

Definition: Say $f(x)$ has an *absolute* max. on interval I if there is $c \in I$ s.t. $f(c) \geq f(x)$, for all $x \in I$ then $f(c)$ is the max. value. Similarly an *absolute* min on an interval I if there is $d \in I$ s.t. $f(x) \leq f(d)$, for all $x \in I$: then $f(d)$ is the min. value.

Extreme Value Theorem

Suppose f is cont. on $[a, b]$. Then f has an absolute max. and an absolute min. on $[a, b]$ – occur either at

DIAGRAM 58

(i) end points a, b or (ii) crit. pt. in (a, b).
Calculate the value of f at each and pick the largest and smallest.

NOTE:

DIAGRAM 59

This need not happen if f *not* cont.

eg. $f(x) = \begin{cases} x, & 0 \leq x < 1 \\ 0, & 1 \leq x \leq 2 \end{cases}$ on $[0, 2]$.

Here there is no max. value. The max. "wants to be" 1, but there is no $c \in [0, 2]$ with $f(c) = 1$.

$\therefore$ Need f *cont.* on *closed* interval $[a, b]$ – since max. or min. may occur at end points which thus need to be included.

Examples: (1) Find absolute max. and min. values of $f(x) = x\sqrt{x+1}$ on $[-1, 1]$.

Solution: $f(x)$ is cont. on $[-1,1]$, so max. and min. occur at end pts ± 1 or int. crit. pt. From previous example crit. pts. are $x=-1,\ -2/3$ – both occur in $[-1,1]$ $\therefore$ Poss max. and min. at $x=\pm 1,\ -2/3$

$$\begin{array}{lcll} x=1 & : & f(x)=\sqrt{2} & \text{– absol. max.} \\ x=-1 & : & f(x)=0 & \\ x=-\frac{2}{3} & : & f(x)=-\frac{-2}{3\sqrt{3}} & \text{– absol. min.} \end{array}$$

$\therefore$ Absolute max. of $\sqrt{2}$ at end pt. $x=1$ and absolute min of $\dfrac{-2}{3\sqrt{3}}$ at int. crit. pt. $x=-\dfrac{2}{3}$.

(2) A farmer has a certain amount of fencing to build a rectangular enclosure which is to be subdivided into 3 parts by 2 fences parallel to one of the sides. How to do it so that area enclosed is a max?

Solution: Let x, y be sidelengths as shown. Then total length ℓ of fencing (a given const.) is

DIAGRAM 60

$$\ell = 4x+2y \Rightarrow y=\frac{1}{2}\ell-2x.$$

Clearly

$$0\le x, y=\frac{1}{2}\ell-2x \Rightarrow 0\le x\le \frac{\ell}{4}.$$

$\therefore$ To find absolute max. value of area

$$A(x)=xy=x(\frac{\ell}{2}-2x) \qquad \text{on } [0,\ \frac{\ell}{4}]$$

– occurs either at end pts $x=0,\ \frac{\ell}{4}$ or int. crit. pt. where $0=A'(x)=\frac{\ell}{2}-4x \Rightarrow x=\frac{\ell}{8}$ an int. crit. pt.

$\therefore$ Poss. max. and min. at $x=0,\ \frac{\ell}{8},\ \frac{\ell}{4}$:

$$\begin{array}{lcll} x=0 & : & A=0 & \\ x=\frac{\ell}{8} & : & A=\frac{2}{3}\ell & \text{– obviously the max.} \\ x=\frac{\ell}{4} & : & A=0 & \end{array}$$

$\therefore$ For absolute max. area, take $x=\frac{\ell}{8}$, $y=\frac{1}{2}\ell-2x=\frac{\ell}{4}$.

Integration

Definition: A function $F(x)$ is an *anti-derivative* of $f(x)$ on an interval I if $F'(x)=f(x)$, for $x\in I$.

Examples: (1) For $f(x) = x^2 - 2$, $F(x) = \frac{1}{3}x^3 - 2x$ is an anti-derivative on $\mathbb{R}$ since $F'(x) = f(x)$. Also $G(x) = \frac{1}{3}x^3 - 2x + 10$ is an anti-derivative.

(2) For $f(x) = \frac{1}{x+1}$, $F(x) = \ln(x+1)$ is an anti-derivative on $(-1, \infty)$.

NOTE: If $F(x)$ is an anti-derivative of $f(x)$, obviously so too is $F(x)+c$ for any constant c, since

$$\frac{d}{dx}(F(x)+c) = F'(x) = f(x).$$

Conversely if $F_1(x)$, $F_2(x)$ are both anti-derivatives of $f(x)$ on interval I, then

$$\begin{aligned} \frac{d}{dx}[F_1(x) - F_2(x)] &= F_1'(x) - F_2'(x) \\ &= f(x) - f(x) = 0 \\ \Rightarrow F_1(x) &- F_2(x) = c \quad \text{(const.)} \end{aligned}$$

or

$$F_1(x) = F_2(x) + c \quad , \quad \text{for } x \in I.$$

□

Definition: The *indefinite integral* $\int f(x)dx$ is defined by

$$\int f(x)dx = F(x) + c$$

where $F(x)$ is any anti-derivative and c is an arbitrary constant – called the *constant of integration*. As c varies this gives all anti-derivatives of f.

Examples: (1) Some basic integrals are

$$\begin{array}{l} \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + c\,, \quad \alpha \neq -1\,; \quad \int \frac{1}{x}dx = \ln x + c\,, \quad x > 0 \\ \int \cos x\, dx = \sin x + c\,, \quad \int \sin x\, dx = -\cos x + c\,, \quad \int e^x dx = e^x + c. \end{array}$$

(2) If $\operatorname{artan} x$ is the inverse tan function, have seen that $\frac{d}{dx}(\operatorname{artan} x) = \frac{1}{1+x^2}$. This gives useful integral

$$\int \frac{1}{1+x^2}dx = \arctan x + c.$$

DEFINITE Integrals

Let $F(x)$ be any anti-derivative of $f(x)$ on $[a, b]$. We define the *definite integral* of f on $[a, b]$ by

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

NOTES: (1) Above integral may also be written

$$\int_a^b f(t)dt \quad \text{– any other letter is as good as } x.$$

(2) Usually adopt notation

$$\begin{aligned} [F(x)]_a^b &= F(b) - F(a)\,, \quad \text{so} \\ \int_a^b f(x)dx &= [F(x)]_a^b. \end{aligned}$$

Since any two anti-derivs differ by a const., it doesn't matter which is used.

(3) With above definition have:

(i)

$$\begin{aligned} \int_a^b f(x)dx &= F(b) - F(a) \\ &= [F(a) - F(b)] = -\int_b^a f(x)dx. \end{aligned}$$

(ii) If $c \in [a, b]$, then

$$\begin{aligned} \int_a^b f(x)dx &= F(b) - F(a) \\ &= F(b) - F(c) + F(c) - F(a) \\ &= \int_a^c f(x)dx + \int_c^b f(x)dx. \end{aligned}$$

Examples: (1)

$$\begin{aligned} \int_0^{\pi/2} \sin x\, dx &= [-\cos x]_0^{\pi/2} \\ &= -\cos(\pi/2) + \cos(0) = 1. \end{aligned}$$

(2) $\displaystyle\int_1^2 \frac{1}{t}dt = [\ln t]_1^2 = \ln(2) - \ln(1) = \ln(2)$

Properties: Using definition of anti-derivative, we have

$$\begin{gathered} \int [f(x) \pm g(x)]dx = \int f(x)dx \pm \int g(x)dx \\ \int cf(x)dx = c\int f(x)dx\,, \quad c = \text{ const.} \end{gathered}$$

for both definite and Indefinite integrals.

AREAS

Fundamental Theorem:
Suppose $f(x)$ is cont. on $[a, b]$ and $f(x) \geq 0$ on $[a, b]$. Then $\int_a^b f(x)dx =$ area under graph above $[a, b]$

DIAGRAM 61

Proof (Outline): For $x \in [a, b]$ let $A(x)$ be area under graph above $[a, x]$ – so $A(a) = 0$. To show $A(x)$ is an anti-deriv. of $f(x)$.

DIAGRAM 62

DIAGRAM 63

Now for h small, consider

$$\begin{aligned} A(x+h) - A(x) - & \text{ area above } [a, x+h] \\ & - \text{ area above } [a, x] \\ = & \text{ area above } [x, x+h] \\ \simeq & \text{ rectangle area } h \cdot f(x) \quad \begin{pmatrix} \text{becomes exact} \\ \text{as } h \to 0 \end{pmatrix} \\ \Rightarrow \tfrac{A(x+h)-A(x)}{h} \simeq & f(x). \end{aligned}$$

Taking limit $h \to 0$ obtain

$$\begin{aligned} A'(x) = f(x) &\Rightarrow A(x) \text{ an anti-deriv of } f(x) \\ \therefore \quad \int_a^b f(x)dx = [A(x)]_a^b &= A(b) - A(a) \\ &= A(b) = \text{ area above } [a, b]. \end{aligned}$$

Corollary: If $f(x) \geq 0$ on $[a, b]$ then $\int_a^b f(x)dx \geq 0$.

NOTE: If $f(x) \leq 0$ on $[a, b]$ can similarly show

$$\begin{aligned} \int_a^b f(x)dx &= -(\text{Area above graph} \\ &\qquad \text{below interval } [a, b]) \\ &\leq 0. \end{aligned}$$

DIAGRAM 64 A

Examples: (1) Find area under graph of $y = \sqrt{x}$ above $[0, 2]$.

Solution: $f(x)$ is $+ve$ on $[0, 2]$ $\therefore$
Theorem $\Rightarrow$

$$\text{Area} = \int_0^2 x^{1/2} dx = \left[\frac{2}{3}x^{3/2}\right]_0^2 = \frac{2}{3} \cdot 2^{3/2} = \frac{4\sqrt{2}}{3}.$$

DIAGRAM 64 B

(2) Find area under graph $y = e^{-x}$ above $[0, \infty)$.
Solution: $f(x) = e^{-x}$ is $(+ve)$ on $[0, \infty)$. Hence for $a > 0$, Theorem $\Rightarrow$

$$\begin{aligned} \text{Area above } [0, a] &= \int_0^a e^{-x} dx \\ &= [-e^{-x}]_0^a = 1 - e^{-a}. \end{aligned}$$

DIAGRAM 65

Taking the limit $a \to \infty$, the area above $[0, \infty) = \lim_{a\to\infty}(1 - e^{-a}) = 1$.

NOTE: Usually define integrals over infinite intervals by

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

and say integral converges if limit exists. □

(3) Find area under graph $y = \dfrac{1}{1+x^2}$ above $[0, 1]$.

Solution: $f(x) = \dfrac{1}{x^2+1}$ is $+ve$ on $[1, 2]$ so Theorem $\Rightarrow$

$$\begin{aligned} \text{Area} = \int_0^1 \frac{dx}{x^2+1} &= [\text{Artan}\, x]_0^1 \\ &= \frac{\pi}{5}. \end{aligned}$$

DIAGRAM 66

(4) Find area under graph $y = 3x^2 - 2$ above $[1, 2]$.

Solution: $f(x) = 3x^2 - 2$ is $(+ve)$ on $[1, 2]$ $\therefore$ Theorem $\Rightarrow$

$$\begin{aligned} \text{Area} = \int_1^2 (3x^2 - 2)dx &= \left[x^3 - 2x\right]_1^2 \\ &= (8 - 4) - (1 - 2) = 5. \end{aligned}$$

Techniques of Integration

Substitution:

Theorem (Indefinite Integrals): *Suppose f, g' are cont. and set $u = g(x)$. Then*

$$\int f(g(x))g'(x)dx = \int f(u)du.$$

Proof: Substitute $u = g(x)$, using $du = \frac{du}{dx}dx \Rightarrow$

$$\int f(g(x))g'(x)dx = \int f(u)\frac{du}{dx}dx = \int f(u)du.$$

NOTE: Above shows that if $F(x)$ is an anti-derivative of $f(x)$ then $F(u) = F(g(x))$ is an anti-derivative of $f(g(x))g'(x)$. Indeed, using chain rule

$$\frac{d}{dx}F(u) = \frac{d}{du}F(u) \cdot \frac{du}{dx} = f(u)\frac{du}{dx} = f(g(x))g'(x).$$

Examples: (1) Evaluate $\int (x^3+3)^5 3x^2 dx$

Solution: Substitute $u = x^3 + 3 \Rightarrow du = \frac{du}{dx}dx = 3x^2 dx$

$$\therefore \qquad \int (x^3+3)^5 3x^2 dx = \int u^5 du = \frac{1}{6}u^6 + c = \frac{1}{6}(x^3+3)^6 + c$$

(2) Evaluate $\int \frac{\ln x}{x} dx$, $x > 0$

Solution: Substitute $u = \ln x \Rightarrow du = \frac{du}{dx}dx = \frac{1}{x}dx$

$$\therefore \qquad \int \frac{\ln x}{x} dx = \int u\, du = \frac{1}{2}u^2 + c = \frac{1}{2}(\ln x)^2 + c.$$

(3) Evaluate $\int \frac{dx}{x^2 + 2x + 2}$.

Solution: $x^2+2x+2=(x+1)^2+1 \quad \therefore$ Substitute $u=x+1 \Rightarrow du=dx$

$$\begin{aligned} \therefore \qquad \int \frac{dx}{x^2+2x+2} = \int \frac{dx}{(x+1)^2+1} &= \int \frac{du}{u^2+1} \\ &= \operatorname{artan}(u)+c \\ &= \operatorname{artan}(x+1)+c. \end{aligned}$$

Theorem (Definite Integrals): *Suppose f, g' are continuous and set $u=g(x)$. Then*

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du.$$

Proof: If $F(x)$ is an anti-derivative of $f(x)$, previous Theorem shows $F(u)=F(g(x))$ is an anti-derivative of $f(g(x))g'(x)$

$$\begin{aligned} \Rightarrow \int_a^b f(g(x))g'(x)dx &= [F(g(x))]_a^b \\ &= F(g(b))-F(g(a)) \\ &= \int_{g(a)}^{g(b)} f(u)du. \end{aligned}$$

Examples: (1) Evaluate $\displaystyle\int_0^{\pi/4} \sin(2x)\cos(2x)dx$.

Solution: Substitute $u=\sin(2x) \Rightarrow du = 2\cos(2x)dx$. Now when $x=0$, $u=0$ and when $x=\pi/4$, $u=\sin(\pi/2)=1$

$$\begin{aligned} \Rightarrow \int_0^{\pi/4} \sin(2x)\cos(2x)dx = \int_0^1 u\cdot\frac{1}{2}du &= \frac{1}{2}\left[\frac{1}{2}u^2\right]_0^1 \\ &= \frac{1}{4}. \end{aligned}$$

(2) Evaluate $\displaystyle\int_1^2 \frac{2x}{(1-2x)^2}dx$

Solution: Substitute $u=2x-1 \Rightarrow du=2dx$. When $x=1$, $u=1$ and when $x=2$, $u=3 \quad \Rightarrow$

$$\begin{aligned} \int_1^2 \frac{2x}{(1-2x)^2}dx &= \frac{1}{2}\int_1^3 \frac{(u+1)}{u^2}du \\ &= \frac{1}{2}\int_1^3 \left(\frac{1}{u}+\frac{1}{u^2}\right)du \\ &= \frac{1}{2}\left[\ln u - \frac{1}{u}\right]_1^3 = \frac{1}{2}\left[\ln 3 + \frac{2}{3}\right]. \end{aligned}$$

NOTE: Last example works for any integral of form

$$\int \frac{cx+d}{(bx-a)^2}dx \qquad - \qquad \text{subst.} u = bx - a.$$

Integrals involving ln function

For $x > 0$, have seen that for $y = \ln x$

$$\frac{dy}{dx} = \frac{1}{x} \Rightarrow \int \frac{dx}{x} = \ln x + c \quad , \quad \text{for } x > 0.$$

For $x < 0$, if $y = \ln(-x)$ have

$$\frac{dy}{dx} = \frac{-1}{-x} = \frac{1}{x},$$

so also

$$\int \frac{dx}{x} = \ln(-x) + c \quad , \quad x < 0.$$

It is usual to combine these cases and write

$$\int \frac{dx}{x} = \ln|x| + c \quad , \quad x \neq 0.$$

Example (1): Evaluate $\int \tan x \, dx$

Solution: $\tan x = \dfrac{\sin x}{\cos x}$ $\therefore$ substitute $u = \cos x \Rightarrow du = -\sin x \, dx$.

$$\begin{aligned} \therefore \quad \int \tan x \, dx = \int \frac{\sin x}{\cos x} dx = \int -\frac{du}{u} &= -\ln|u| + c \\ &= -\ln|\cos x| + c. \end{aligned}$$

Example (2) (PARTIAL FRACTIONS)

Evaluate $\displaystyle\int \frac{x+2}{x(x+1)} dx$

Solution: First note $\dfrac{1}{x} - \dfrac{1}{x+1} = \dfrac{(x+1)-x}{x(x+1)} = \dfrac{1}{x(x+1)}$

$$\Rightarrow \frac{x+2}{x(x+1)} = \frac{(x+1)+1}{x(x+1)} = \frac{1}{x} + \frac{1}{x(x+1)} = \frac{2}{x} - \frac{1}{x+1}$$

$$\therefore \quad \int \frac{x+2}{x(x+1)} dx = 2\int \frac{dx}{x} - \int \frac{dx}{x+1}$$

↑ (subst. $u = x + 1$)

$$= 2\ln|x| - \ln|x+1| + c.$$

NOTE: Above method applies to any integral of form

$$\int \frac{cx+d}{(x-a)(x-b)}dx$$

by noting that $\dfrac{1}{(x-a)} - \dfrac{1}{(x-b)} = \dfrac{(a-b)}{(x-a)(x-b)}$ (check!).
– in the case $a = b$ use substitution $u = x - a$ (as in earlier example).

Integration by parts

Given two functions $u(x)$, $v(x)$ have product rule for differential

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$$

$\Rightarrow$ uv an anti-derivative of $uv' + vu'$
ie. $uv = \int (uv' + vu')dx = \int uv'dx + \int vu'dx$, so $\int uv'dx = uv - \int vu'dx$ – integration by parts formula.

Examples: (1) Evaluate $\int x \sin x \, dx$

Solution: Put $u = x$, $v' = \sin x \Rightarrow u' = 1$, $v = -\cos x$

$$\begin{aligned} \therefore \quad \int \underset{u}{\underset{\uparrow}{x}} \, \underset{v^1}{\underset{\uparrow}{\sin x}} \, dx &= uv - \int vu'dx \\ &= -x\cos x + \int \cos x dx \\ &= -x\cos x + \sin x + c. \end{aligned}$$

NOTES: Same method works for integrals

$$\int x \cos x \, dx \quad , \quad \int xe^x \, dx.$$

(2) Evaluate $\int x^3 \ln x \, dx$.

Solution: First try $u = x^3$, $v' = \ln x \Rightarrow u' = 3x^2$, $v = \int \ln x \, dx = ?$
$\therefore$ try $u = \ln x$, $v' = x^3 \Rightarrow u'\dfrac{1}{x}$, $v = \dfrac{1}{4}x^4$

$$\therefore \quad \int \underset{v^1}{\overset{\uparrow}{x^3}} \underset{u}{\overset{\uparrow}{\ln x}}\, dx = uv - \int vu'dx$$
$$= \frac{1}{4}x^4 \ln x - \int \frac{1}{4}x^4 \cdot \frac{1}{x} dx$$
$$= \frac{1}{4}x^4 \ln x - \frac{1}{4}\int x^3 dx$$
$$= \frac{1}{4}x^4 \ln x - \frac{1}{16}x^4 + c.$$

Numerical Integration using Taylor Series

Frequently encounter integrals which cannot be evaluated directly. In that case Taylor series provide a useful approx.

Illustrative Example:

Estimate $\int_0^{1/2} e^{-x^2} dx$ to 4 decimal places. Here the integration cannot be done exactly, but using Taylor series

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\cdots \\
\Rightarrow e^{-x^2} &= 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!}\cdots\cdots \\
\therefore \quad \int_0^{1/2} e^{-x^2} dx &= \int_0^{1/2} \left(1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!}\cdots\right) dx \\
&= \left[x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \frac{x^9}{9\cdot 4!}\cdots\right]_0^{1/2} \\
&= 1 - \frac{1}{32^3} + \frac{1}{5\cdot 2!\cdot 2^5} - \frac{1}{7\cdot 3!\cdot 2^7} + \frac{1}{9\cdot 4!2^9}\cdots\cdots \\
&= 0\cdot 5 - 0\cdot 041667 + 0\cdot 003125 - 0\cdot 000186 + 0\cdot 000090\cdots.
\end{aligned}$$

This is an *alternating series*, which has following important property.

FACT: Suppose $\{a_n\}_{n=0}^{\infty}$ is a decreasing sequence of positive terms (ie. $a_n \geq a_{n+1}$, for all $n \in \mathbb{N}$) such that $\lim\limits_{n\to\infty} a_n = 0$. Then the alternating series

$$\sum_{n=0}^{\infty}(-1)^n a_n = a_0 - a_1 + a_2 - a_3 + a_4 \cdots\cdots$$

converges. If ℓ is the sum of series then

$$\left|\sum_{k=0}^{n}(-1)^k a_k - \ell\right| < a_{n+1}$$

ie. error involved in summing 1st. n terms is less than next term. $\square$

For above example we have approximate

$$\int_0^{1/2} e^{-x^2} dx \simeq \text{ sum of 1st 4 terms } \simeq 0 \cdot 4613$$

and then the

$$\text{error } < \text{ fifth term } = 9 \times 10^{-5}$$

$\therefore$ Approximately valid to 4 decimal places.

Example: Estimate $\int_0^1 \sin(x^2)dx$ to 4 decmil places.

Solution: Using Taylor series

$$\begin{aligned}
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\cdots \\
\Rightarrow \sin(x^2) &= x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \cdots\cdots \\
\therefore \quad \int_0^1 \sin(x^2)dx &= \int_0^1 \left(x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \cdots \right) dx \\
&= \left[\frac{x^3}{3} - \frac{x^7}{7 \cdot 3!} + \frac{x^{11}}{11 \cdot 5!} - \frac{x^{15}}{15.7!} + \cdot \right]_0^1 \\
&= \frac{1}{3} - \frac{1}{7 \cdot 3!} + \frac{1}{11 \cdot 5!} - \frac{1}{15.7!} + \cdot \\
&\simeq 0 \cdot 3333 - 0 \cdot 00076 - 0 \cdot 00003 + \cdots\cdots \\
\therefore \quad \int_0^1 \sin(x^2)dx &\simeq \text{ sum 1st 3 terms} \simeq 0 \cdot 310268
\end{aligned}$$

with

$$\text{error} < \text{4th term} = 3 \times 10^{-5} \quad \text{(ie. accurate to 4 decimal places).}$$

NOTE: Above gives useful method for numerically approximating integrals even when resulting sums are not alternating (but error estimates may be difficult).

Example: Estimate $\int_0^{1/2} \frac{1}{1-x^2} dx$

Solution: Using Taylor series

$$
\begin{aligned}
\frac{1}{1-x} &= 1 + x + x^2 + x^3 + \cdots\cdots && \text{(Geom. Series)} \\
\Rightarrow \frac{1}{1-x^2} &= 1 + x^2 + x^4 + x^6 + \cdots\cdots \\
\therefore \quad \int_0^{1/2} \frac{1}{1-x^2} dx &= \int_0^{1/2} \left(1 + x^2 + x^4 + x^6 + \cdots\right) dx \\
&= \left[x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots \right]_0^{1/2} \\
&\simeq \frac{1}{2} + \frac{1}{3 \cdot 2^3} + \frac{1}{5 \cdot 2^5} + \frac{1}{7 \cdot 2^7} && \text{(1st 4 terms)} \\
&\simeq \cdot 05493.
\end{aligned}
$$

Exact result: Using partial fractions

$$
\begin{aligned}
&\frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)} = \frac{1}{2}\left(\frac{1}{1-x} + \frac{1}{1+x}\right) \\
&\Rightarrow \int_0^{1/2} \frac{1}{1-x^2} dx = \frac{1}{2}\left[-\ln(1-x) + \ln(1+x)\right]_0^{1/2} \\
&= \frac{1}{2}\left[-\ln(\frac{1}{2}) + \ln(\frac{3}{2})\right] = \frac{1}{2}\ln(3) \simeq \cdot 05493
\end{aligned}
$$

in good agreement with above approx.

Volumes of revolution

Suppose $f(x)$ is $+ve$ and cont. on $[a, b]$. By rotating graph of $f(x)$ above interval $[a, b]$ about x-axis obtain

DIAGRAM 67

a cylindrical solid. Here we consider problem of obtaining vol. V of this solid – called a *volume of revolution*

DIAGRAM 68

FACT: Suppose $f(x)$ is $+ve$ and cont. on $[a, b]$. Then volume of revolution of solid obtained by rotating graph $y = f(x)$ above $[a, b]$ about x-axis is

$$V = \pi \int_a^b f(x)^2 dx.$$

Proof (Outline): For $x \in [a, b]$, let $V(x)$ be volume obtained by rotating graph of f above interval $[a, x]$ about x-axis. So $V(a) = 0$, $V(b) = V$.

DIAGRAM 69

Then for $h \geq 0$ small

DIAGRAM 70

$$\begin{aligned} V(x+h) - V(x) &= \text{volume obtained by rotating graph} \\ &\quad \text{of } f \text{ above } [x, x+h] \text{ about } x\text{-axis} \\ &\simeq \text{vol. cylinder radius } f(x) \text{ and ht. h} \\ &= \pi f(x)^2 h \quad \text{(becomes exact as } h \to 0) \\ \Rightarrow \frac{V(x+h) - V(x)}{h} &\simeq \pi f(x)^2 \end{aligned}$$

$\therefore$ Taking limit $h \to 0$ obtain

$$V'(x) = \pi f(x)^2$$

so $V(x)$ is an anti-deriv. of $\pi f(x)^2$ over $[a, b]$ $\therefore$

$$\begin{aligned} \int_a^b \pi f(x)^2 dx = [V(x)]_a^b &= V(b) - V(a) \\ &= V(b) = V. \end{aligned}$$

$\square$

Examples: (1) Let $a > 0$. Find volume V of solid obtained by rotating $y = f(x)$ over $[0, a]$ about x-axis for following choices of f:

(i) $f(x) = x$ (cone) (ii) $f(x) = \sqrt{x}$ (paraboloid) (iii) $y = e^x$.

Solution: First note that in all cases $f(x)$ is $+ve$ and continuous on $[0, a]$

(i)
$$\begin{aligned} V = \pi \int_0^a (x)^2 dx &= \pi \left[\frac{x^3}{3}\right]_0^a \\ &= \frac{1}{3}\pi a^3. \end{aligned}$$

DIAGRAM 71

(ii)
$$\begin{aligned} V = \pi \int_0^a (\sqrt{x})^2 dx &= \pi \int_0^a x dx \\ &= \left[\frac{x^2}{2}\right]_0^a = \frac{1}{2}\pi a^2 \end{aligned}$$

DIAGRAM 72

(iii)
$$\begin{aligned} V = \pi \int_0^a (e^x)^2 dx &= \pi \int_0^a e^{2x} dx \\ &= \pi \left[\frac{e^{2x}}{2}\right]_0^a = \frac{\pi}{2}(e^{2a} - 1). \end{aligned}$$

DIAGRAM 73

(2) Derive the formula for volume of sphere radius $a > 0$.

Solution:

V is volume of revolution determined by rotating $y = \sqrt{a^2 - x^2}$ above $[-a, a]$ about x-axis

$$\begin{aligned} \Rightarrow V &= \pi \textstyle\int_{-a}^{a} \left(\sqrt{a^2 - x^2}\right)^2 dx \\ &= \pi \textstyle\int_{-a}^{a} (a^2 - x^2) dx = \pi \left[a^2 x - \frac{x^3}{3}\right]_{-a}^{a} \\ &= \pi(2a^3 - \frac{2}{3}a^3) = \frac{4}{3}\pi a^3. \end{aligned}$$

Complex Numbers

Recall that a complex number is of form $z = a + bi$ where $a, b \in \mathbb{R}$ and i is the imaginary number satisfying

$$i^2 = -1$$

– introduced in 17th century to obtain roots of a polynomial equation, but has since revolutionized mathematics and applications. Recall that complex numbers can be added, multiplied and divided as for real numbers – just replace i^2 everywhere by -1: eg.

$$\begin{aligned} \frac{3 - 2i}{1 + i} &= \frac{3 - 2i}{1 + i} \times \frac{(1 - i)}{(1 - i)} = \frac{(3 - 2i)(1 - i)}{1 - i^2} \\ &= \frac{1}{2}(3 - 2i)(1 - i) = \frac{1}{2}(3 - 3i - 2i + 2i^2) \\ &= \frac{1}{2}[(3 - 2) - 5i] = \frac{1}{2}(1 - 5i) \qquad \text{etc.} \end{aligned}$$

Polar Form:

DIAGRAM 75

A complex number $z = x + iy$ may be represented by a point in *complex plane* where vertical axis is the *imagionary axis* and horizontal axis the *real axis.* We call $r = \sqrt{x^2 + y^2}$ the *modulus* of z, denoted $|z|$- gives distance of z from the origin. In terms of polar coordinates have

$$x = r\cos\theta \quad , \quad y = r\sin\theta$$

$\Rightarrow \quad z = x + iy = r(\cos\theta + i\sin\theta)$ – called *polar form* of z.

Euler's Formula: $\cos\theta + i\sin\theta = e^{i\theta}$.

Proof: (Outline): Set $f(\theta) = \cos\theta + i\sin\theta$

$$\begin{aligned}
\Rightarrow \frac{df}{d\theta} = -\sin\theta + i\cos\theta &= i^2\sin\theta + i\cos\theta \\
&= i(\cos\theta + i\sin\theta) = if(\theta) \\
\therefore \quad \frac{d\ln f}{d\theta} = \frac{d\ln f}{df}\cdot\frac{df}{d\theta} &= \frac{1}{f}\cdot if = i \\
\Rightarrow \quad \ln f = i\theta + c\,, \quad \text{so} & \\
f(\theta) = Ke^{i\theta} \quad , \quad K = e^c. &
\end{aligned}$$

But $f(0) = \cos(0) + i\sin(0) = 1$

$$\begin{aligned}
\Rightarrow \quad 1 = f(0) = Ke^0 = K \\
\therefore \quad f(\theta) = e^{i\theta}.
\end{aligned}$$

Thus every complex number $z = x + iy$ is expressible in polar form

$$z = re^{i\theta} \quad , \quad \text{where } r = |z| = \sqrt{x^2 + y^2}.$$

– has many applications.

NOTE: Euler's formula can also be proved using Taylor series.

Exponential form for cos and sin.

$$\begin{aligned}
e^{-i\theta} = \cos(-\theta) + i\sin(-\theta) = \cos\theta - i\sin\theta \\
\Rightarrow \quad e^{\pm i\theta} = \cos\theta \pm i\sin\theta
\end{aligned}$$

$$\begin{array}{rl}
\text{Adding} \ \Rightarrow & \cos\theta = \frac{1}{2}(e^{i\theta} + e^{i\theta}) \\
\text{Substracting} \ \Rightarrow & \sin\theta = \frac{1}{2i}(e^{\theta} - e^{-i\theta})
\end{array}$$

Recall: $\frac{1}{i} = \frac{1}{i} \times \frac{i}{i} = \frac{i}{i^2} = -i.$

Applications to Integration

(1) Evaluate $\int_0^{\pi} \sin^2 \theta d\theta.$

Solution:

$$\begin{aligned}
\sin^2\theta &= \left(\frac{1}{2i}\right)^2 \left(e^{i\theta} - e^{-i\theta}\right)^2 = -\frac{1}{4}\left(e^{2i\theta} + e^{-2i\theta} - 2e^{i\theta}\cdot e^{-i\theta}\right) \\
&= -\frac{1}{4}(2\cos 2\theta - 2) = \tfrac{1}{2}(1-\cos 2\theta) \\
\Rightarrow \int_0^{\pi} \sin^2\theta d\theta &= \frac{1}{2}\int_0^{\pi}(1-\cos 2\theta)d\theta \\
&= \frac{1}{2}\left[\theta - \tfrac{1}{2}\sin 2\theta\right]_0^{\pi} = \tfrac{\pi}{2}.
\end{aligned}$$

(2) Evaluate $\int e^x \cos x \, dx$

Solution: Use

$$\begin{aligned}
e^x \cos x &= \text{Re}\left[e^x(\cos x + i \sin x)\right] \quad (\text{Re} \equiv \text{ real part}) \\
&= \text{Re}\left[e^x \cdot e^{ix}\right] = \text{Re } e^{(1+i)x}. \\
\Rightarrow \int e^x \cos x \, dx &= \text{Re} \int e^{(1+i)x} dx \\
&= \text{Re}\left[\frac{e^{(1+i)x}}{1+i}\right] + c.
\end{aligned}$$

Now

$$\begin{aligned}
\frac{e^{(1+i)x}}{1+i} &= \frac{e^x(\cos x + i\sin x)}{1+i} \times \frac{(1-i)}{(1-i)} \\
&= \frac{1}{2}e^x(\cos x - i \sin x)(1-i) \\
&= \frac{1}{2}e^x\left[\underset{\substack{\Uparrow \\ \text{real part}}}{(\cos x - \sin x)} + i(\sin x - \cos x)\right] \\
\therefore \quad \int e^x \cos x \, dx &= \frac{1}{2}e^x(\cos x - \sin x) + c.
\end{aligned}$$

NOTE: Above integral can also be done using integration by parts (twice).

Vectors

In real life a *vector* quantity is something that needs a magnitude and a direction to specify it - eg. force, velocity, displacement (ie. displacement 10m ME), as distinct from a *scalar* quantity where just a magnitude is needed (eg. time, area, temperature etc.)

A vector is represented geometrically in the (x, y) plane or in (x, y, z) - space by a line segment of appropriate length (called its *magnitude*), pointed in the correct direction – indicated by an arrow. Only the length and direction of the represention are significant: it can be placed anywhere convenient in the x - y plane.

eg. vector of unit length 45° to x-axis has many representations as shown.

DIAGRAM 76

DIAGRAM 77

A vector $\boldsymbol{v} = \overrightarrow{PQ}$ in the x - y plane may be represented by a pair of numbers

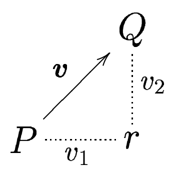
$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x_Q - x_P \\ y_Q - y_P \end{pmatrix}$$

which is the same for all representations $\overrightarrow{PQ}$ pf $\boldsymbol{v}$. We call v_1, v_2 the *components* of vector $\boldsymbol{v}$.

NOTES: (1) We call vector $\binom{0}{0}$ the *zero vector* – denoted by $\mathbf{0}$.

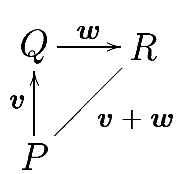
(2) If P, Q are points in 2-space or 3-space, we let $\overrightarrow{PQ}$ denote the vector from P to Q.

Definition: For vector $\boldsymbol{v} = \overrightarrow{PQ}$, the *norm* (or *length* or *magnitude*) of $\boldsymbol{v}$, written $\|\boldsymbol{v}\|$, is the distance PQ between P and Q. Thus for $\boldsymbol{v} = \binom{v_1}{v_2}$ we have



$$\|\boldsymbol{v}\| = PQ = \sqrt{v_1^2 + v_2^2}.$$

Vector Addition: Add vectors by the triangle rule:



Consider $\Delta\, PQR$ with $\boldsymbol{v} = \overrightarrow{PQ}$, $\boldsymbol{w} = \overrightarrow{QR}$. Then $\boldsymbol{v} + \boldsymbol{w} = \overrightarrow{PR}$, the diagonal.

DIAGRAM 78

In terms of components, if

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad \boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

obviously

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix}$$

– ie. add vectors component wise.

Properties: It can be shown that vector addition satisfies the following properties:

$$\begin{aligned} \boldsymbol{v} + \boldsymbol{w} &= \boldsymbol{w} + \boldsymbol{v} && \text{(commutative law)} \\ \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) &= (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} && \text{(associative law)} \\ \boldsymbol{v} + \mathbf{0} &= \mathbf{0} + \boldsymbol{u} = \boldsymbol{v} \end{aligned}$$

etc.

Scalar Multiplication

May multiply a vector $\boldsymbol{v}$ by a number α (scalar). We define $\alpha\boldsymbol{v}$ to be the vector of magnitude

$$\|\alpha\boldsymbol{v}\| = |\alpha| \cdot \|\boldsymbol{v}\|$$

in the same direction as $\boldsymbol{v}$ if $\alpha > 0$ and opposite direction if $\alpha < 0$, DIAGRAM 79

DIAGRAM 80

By similar Δ's, if $\boldsymbol{v} = \binom{v_1}{v_2}$ then $\alpha\boldsymbol{v} = \binom{\alpha v_1}{\alpha v_2}$.

NOTES: (1) We usually set

$$-\boldsymbol{v} = (-1) \cdot \boldsymbol{v}$$

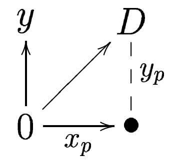
ie. vector with same magnitude as $\boldsymbol{v}$ but pointing in opposite direction. *Subtraction* of vectors is then defined by

$$\boldsymbol{v} - \boldsymbol{w} = \boldsymbol{v} + (-\boldsymbol{w})$$

.

(2) If mult. any vector $\boldsymbol{v} = \binom{v_1}{v_2}$ by zero obtain the zero vector:

$$0 \cdot \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}.$$



Position Vectors:

Let $P = (x_p, y_p)$ be point in x - y plane. Vector $\overrightarrow{OP}$, where O is the origin, is called the *position vector* of P. Obviously

$$\overrightarrow{OP} = \begin{pmatrix} x_p \\ y_p \end{pmatrix}.$$

Unit Vectors:

A *unit vector* is a vector of unit length. If $\boldsymbol{v} \neq \mathbf{0}$ is any vector, then

$$\hat{\boldsymbol{v}} = \boldsymbol{v} / \|\boldsymbol{v}\|$$

determines a unit vector in the direction of $\boldsymbol{v}$.

DIAGRAM 81

In particular

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

determine unit vectors along the x and y axes respectively.

For any vector $\boldsymbol{v} = \binom{v_1}{v_2}$ we have

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = v_1 \boldsymbol{i} + v_2 \boldsymbol{j}.$$

$v_1 v_2$ are called the components of $\boldsymbol{v}$ in the $\boldsymbol{i}$ and $\boldsymbol{j}$ directions respectively.

Vectors in 3-space

Similarly in 3-space a vector $\boldsymbol{v} = \overrightarrow{PQ}$ is represented in component form by

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} x_Q - x_P \\ y_Q - y_P \\ z_Q - z_P \end{pmatrix}$$

DIAGRAM 82

which is the same for all representations $\overrightarrow{PQ}$ of $\boldsymbol{v}$.

DIAGRAM 83

For $\boldsymbol{v} = \overrightarrow{OP} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ the *length* (or *norm*) of vector $\boldsymbol{v}$ is

$$\begin{aligned} \|\boldsymbol{v}\| &= OP \\ &= \sqrt{ON^2 + NP^2} = \sqrt{v_1^2 + v_2^2 + v_3^2}. \end{aligned}$$

As before we add vectors component wise and we may define multiplication by a scalar α. So if

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad , \quad \boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

are vectors then

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{pmatrix} \quad , \quad \alpha\boldsymbol{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \alpha v_3 \end{pmatrix} \ , \ \alpha \in \mathbb{R}.$$

The unit vectors along the x, y, z axes are respectively

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \ , \quad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \ , \quad \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

in terms of which a vector $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ may be expressed

$$\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}.$$

We call v_1, v_2, v_3 are *components* of $\boldsymbol{v}$ in the $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ directions respectively.

NOTES: (1) We similarly have the zero vector

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(2) 2-space is usually denoted by $\mathbb{R}^2$ and 3-space by $\mathbb{R}^3$.

Example: A Zepplin is directed NE at 20 km/hr into a 10 km/hr wind in direction E75°S. Find the actual speed and direction of the airship.

Solution:

DIAGRAM 84

Using vectors actual velocity is $\boldsymbol{v}$ as shown. From trigonometry speed is

$$\|\boldsymbol{v}\| = \sqrt{20^2 - 10^2} = 10\sqrt{3} \quad \text{km/hr}$$

DIAGRAM 85

while $\theta = 30°$ $\therefore$ Actual directions is E15°N.

SCALAR PRODUCT

For non zero vectors $\boldsymbol{v} = \overrightarrow{OP}$, $\boldsymbol{w} = \overrightarrow{OQ}$ the *angle between* $\boldsymbol{v}$ and $\boldsymbol{w}$ is the angle θ with $0 \leq \theta \leq 180°$ between $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ at the origin O.

DIAGRAM 86 DIAGRAM 87 DIAGRAM 88

Definition: The *scalar* (or *dot* or *inner*) product of vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, denoted by $\boldsymbol{v} \cdot \boldsymbol{w}$, is the *number* given by

$$\boldsymbol{v} \cdot \boldsymbol{w} = \begin{cases} 0\,, & \text{if } \boldsymbol{v} \text{ or } \boldsymbol{w} = \mathbf{0} \\ \|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\| \cos\theta\,, & \text{otherwise} \end{cases}$$

where θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$.

Properties:

(1) If $\boldsymbol{v}$, $\boldsymbol{w} \neq \mathbf{0}$ then

$$\boldsymbol{v} \cdot \boldsymbol{w} = 0 \Leftrightarrow \|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\| \cos\theta = 0 \Leftrightarrow = 90°$$

ie. $\boldsymbol{v} \cdot \boldsymbol{w} = 0 \Leftrightarrow \boldsymbol{v}$ and $\boldsymbol{w}$ are *perpendicular* (or *orthogonal*).

(2) $\boldsymbol{v} \cdot \boldsymbol{v} = \|\boldsymbol{v}\| \cdot \|\boldsymbol{v}\| \cos(0) = \|\boldsymbol{v}\|^2$, or

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}.$$

(3)

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{w} &= \boldsymbol{w} \cdot \boldsymbol{v} \\ \boldsymbol{v} \cdot (\boldsymbol{u} + \boldsymbol{w}) &= \boldsymbol{v} \cdot \boldsymbol{u} + \boldsymbol{v} \cdot \boldsymbol{w} \\ \alpha(\boldsymbol{v} \cdot \boldsymbol{w}) &= (\alpha\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot (\alpha\boldsymbol{w}) \quad , \quad \alpha \text{ a scalar.} \end{aligned}$$

Component form

¿From above it is easily seen that

$$\begin{aligned} \boldsymbol{i}\cdot\boldsymbol{i} &= \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = 1 \\ \boldsymbol{i}\cdot\boldsymbol{j} &= \boldsymbol{j}\cdot\boldsymbol{i} = \boldsymbol{i}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{j} = 0. \end{aligned}$$

If $\boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}$, $\boldsymbol{w} = w_1\boldsymbol{i} + w_2\boldsymbol{j} + w_3\boldsymbol{k}$ are two vectors then using properties above obtain

$$\boxed{\boldsymbol{v}\cdot\boldsymbol{w} = v_1w_1 + v_2w_2 + v_3w_3}.$$

NOTES: In particular

$$\|\boldsymbol{v}\|^2 = \boldsymbol{v}\cdot\boldsymbol{v} = v_1^2 + v_2^2 + v_3^2.$$

Examples: (1) Find the angle θ between the following pairs of vectors:

$$\text{(i)}\quad \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\quad \boldsymbol{w} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \qquad \text{(ii)}\quad \boldsymbol{v} = \begin{pmatrix} 1 \\ -5 \\ 4 \end{pmatrix},\quad \boldsymbol{w} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

Solutions: From definition of dot product, angle θ given by

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

(i) Here $\|\boldsymbol{v}\| = 1$, $\quad \|\boldsymbol{w}\| = \sqrt{0+4+4} = 2\sqrt{2}$

$$\boldsymbol{v}\cdot\boldsymbol{w} = 0 + 0 + 2 = 2$$

$$\therefore \quad \cos\theta = \frac{2}{2\sqrt{2}} = \frac{1}{\sqrt{2}} \Rightarrow \theta = 45^\circ$$

(ii) $\boldsymbol{v}\cdot\boldsymbol{w} = 3\cdot 1 - 3\cdot 5 + 3\cdot 4 = 0 \therefore$ vectors perpendicular, hence $\theta = 90^\circ$,

(2) If $P = (2,4,-1)$, $Q = (1,1,1)$, $R = (-2,2,3)$, find the angle $\theta = PQR$.

Solution:

$$\overrightarrow{QP} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$$

DIAGRAM 89

$$\overrightarrow{QR} = \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix}$$

$$\therefore \quad \left\|\overrightarrow{QP}\right\| = \left\|\overrightarrow{QR}\right\| = \sqrt{1+9+4} = \sqrt{14}$$

$$\overrightarrow{QP} \cdot \overrightarrow{QR} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = -3 + 3 - 4 = -4$$

$$\therefore \quad \cos\theta = \frac{\overrightarrow{QP} \cdot \overrightarrow{QR}}{\left\|\overrightarrow{QP}\right\| \cdot \left\|\overrightarrow{QR}\right\|} = \frac{-4}{14} \Rightarrow \theta \simeq 106°36'.$$

Application : Decomposition w.r.t. a vector

DIAGRAM 90

Given a vector $\boldsymbol{v}$, suppose we want to decompose another vector $\boldsymbol{w}$ as a vector in the direction of $\boldsymbol{v}$ and a vector perpendicular to $\boldsymbol{v}$, $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ as shown.

Call $\boldsymbol{w}_1$ compt. of $\boldsymbol{w}$ in direction of $\boldsymbol{v}$ – also called the *projection* of $\boldsymbol{w}$ on to $\boldsymbol{v}$, and $\boldsymbol{w}_2$ compt. of $\boldsymbol{w}$ perpendicular to $\boldsymbol{v}$.
Clearly $\boldsymbol{w}_1 = \alpha\boldsymbol{v}$, for some scalar α

$$\begin{aligned}
&\Rightarrow \quad \boldsymbol{w} = \alpha\boldsymbol{v} + \boldsymbol{w}_2. \quad \text{Since } \boldsymbol{w}_2 \text{ perp. to } \boldsymbol{v} \text{ have } \boldsymbol{w}_2 \cdot \boldsymbol{v} = 0 \\
&\Rightarrow \quad \boldsymbol{w} \cdot \boldsymbol{v} = (\alpha\boldsymbol{v} + \boldsymbol{w}_2) \cdot \boldsymbol{v} = \alpha\boldsymbol{v} \cdot \boldsymbol{v} = \alpha \|\boldsymbol{v}\|^2 \\
&\qquad \therefore \quad \alpha = \boldsymbol{w} \cdot \boldsymbol{v} / \|\boldsymbol{v}\|^2 \\
&\Rightarrow \quad \boldsymbol{w} = \frac{(\boldsymbol{w} \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|^2}\boldsymbol{v} \quad , \quad \boldsymbol{w}_2 = \boldsymbol{w} - \frac{(\boldsymbol{w} \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|^2}\boldsymbol{v}.
\end{aligned}$$

Example: A metal ball placed at rest on a smooth inclined plane at an angle of 45° to the horizontal, experiences a force due to gravity (recall acc. due to gravity is 9.8 m/sec^2 downward). Determine the acceleration vector of the ball.

Solution:

DIAGRAM 91

Let $\boldsymbol{g} = -9 \cdot 8\boldsymbol{j} = -9 \cdot 8\binom{0}{1}$ be acceleration due to gravity. A vector in direction of ball is

$$\boldsymbol{v} = \boldsymbol{i} + \boldsymbol{j} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then acceleration vector is $\boldsymbol{a} =$ component of $\boldsymbol{g}$ in direction of $\boldsymbol{v}$.

$$= \frac{(\boldsymbol{g} \cdot \boldsymbol{v})}{\|\boldsymbol{v}\|^2}\boldsymbol{v} = -\frac{9 \cdot 8}{2}\boldsymbol{v} = -4 \cdot 9(\boldsymbol{i} + \boldsymbol{j})$$

NOTE: Magnitude of acceleration is

$$\begin{aligned} \|\boldsymbol{a}\| = \|-4.9\boldsymbol{v}\| &= |-4.9| \cdot \|\boldsymbol{v}\| \\ &= 4 \cdot 9\sqrt{2} \text{ m/sec}^2. \end{aligned}$$

Vectors in $\mathbb{R}^n$

The above generalizes to $\mathbb{R}^n$. A vector $\boldsymbol{v}$ in $\mathbb{R}^n$ is specified by n components:

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad , \quad v_i \in \mathbb{R}.$$

We define addition of vectors and multiply by a scalar α component wise as before. Thus if $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$ is another vector then

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_2 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \boldsymbol{w} + \boldsymbol{v} \quad , \quad \alpha\boldsymbol{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{pmatrix} , \quad \alpha \in \mathbb{R}.$$

May also define the *scalar* (or *dot*) product between $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_2 w_1 + v_2 w_2 + \cdots\cdots + v_n w_n.$$

We call

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

the *length* of the vector $\boldsymbol{v}$.

As for $\mathbb{R}^3$, every vector $\boldsymbol{v}$ is expressible in terms of the coordinate vectors

$$\boldsymbol{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix}$$

as

$$\begin{aligned} \boldsymbol{v} &= \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots\cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ v_n \end{pmatrix} \\ &= v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \cdots\cdots + v_n\boldsymbol{e}_n. \end{aligned}$$

NOTES: (1) For $\mathbb{R}^3$ we have, in our previous notation,

$$\boldsymbol{i} = \boldsymbol{e}_1\,, \quad \boldsymbol{j} = \boldsymbol{e}_2\,, \quad \boldsymbol{k} = \boldsymbol{e}_3.$$

(2) We have also the zero vector

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{satisfying}$$

$\mathbf{0} + \boldsymbol{v} = \boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$, $0 \cdot \boldsymbol{v} = \mathbf{0}$, for every vector $\boldsymbol{v}$.

(3) If $\boldsymbol{v} \neq \mathbf{0}$, $\boldsymbol{w} \neq \mathbf{0}$, we say that vectors $\boldsymbol{v}$, $\boldsymbol{w}$ are perp. (or orthogonal) if

$$\boldsymbol{v} \cdot \boldsymbol{w} = 0.$$

Vector Spaces:

Definition: A non-empty subset $V \subseteq \mathbb{R}^n$ is called a *vector space* if it satisfies the following

(i) If $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in V$ then $\boldsymbol{v} + \boldsymbol{w} \in V$ (closure under vector addition).

(ii) If $\boldsymbol{v} \in V$ and $\alpha \in \mathbb{R}$ then $\alpha \boldsymbol{v} \in V$ (closure under scalar mult.)

If $w \subseteq V$ and W is a vector space, we call W a *subspace* of V.

NOTE: If $V \subseteq \mathbb{R}^n$ is a vector space and $\boldsymbol{v} \in V$, then by (ii)

$$\mathbf{0} = 0 \cdot \boldsymbol{v} \in V$$

ie. the zero vector belongs to every vector space. □

If $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m \in \mathbb{R}^n$, a vector of the form

$$\boldsymbol{v} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_m \boldsymbol{v}_m \quad , \quad \alpha_i \in \mathbb{R}$$

is called a *linear combination* of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m$. Let V be the set of all such linear combinations:

$$V = \{\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_m \boldsymbol{v}_m \mid \alpha_1, ..., \alpha_m \in \mathbb{R}\}.$$

Then it can be shown that V is a vector space. We say that the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m$ *span* V – or that $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m\}$ is a *spanning set* for V.

Examples:

(1) The coord. vectors $\boldsymbol{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$ span $\mathbb{R}^n$ since every $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ is a linear combination of these vectors:

$$\begin{aligned} \boldsymbol{v} &= v_1 \boldsymbol{e}_1 + v_2 \boldsymbol{e}_2 + \cdots + v_n \boldsymbol{e}_n. \\ \text{Therefore} \qquad \mathbb{R}^n &= \{v_1 \boldsymbol{e}_1 + v_2 \boldsymbol{e}_2 + \cdots + v_n \boldsymbol{e}_n \mid v_1, v_2, ..., v_n \in \mathbb{R}\} \end{aligned}$$

is a vector space.

(2) Consider $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^2$. Then

$$V = \left\{ \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}$$

is a vector space spanned by a single vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^2 \therefore V$ is a subspace of $\mathbb{R}^2$.

(3) The vector space spanned by the zero vector $\mathbf{0} \in \mathbb{R}^n$ is $\{\alpha \mathbf{0} \mid \alpha \in \mathbb{R}\} = \{\mathbf{0}\}$ – called the *zero vector space.*

Definition: Say that the set of vectors $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m\} \subseteq \mathbb{R}^n$ are *linearly independent* if

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \cdots + \alpha_m \boldsymbol{v}_m = \mathbf{0} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$$

otherwise S is called *linearly dependent* – ie. S is linearly depandent if there exist scalars $\alpha_1, ..., \alpha_m$ not all zero s.t.

$$\alpha_1 \boldsymbol{v}_1 + \alpha \boldsymbol{v}_2 + \cdots + \alpha_m \boldsymbol{v}_m = \mathbf{0}.$$

Examples: (1) In $\mathbb{R}^2$ vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are linearly independent since

$$\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \alpha + \beta \\ \alpha - \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\Rightarrow \left.\begin{matrix} \alpha + \beta = 0 \\ \alpha - \beta = 0 \end{matrix}\right\} \Rightarrow \alpha = \beta = 0.$$

(2) In $\mathbb{R}^n$ the coordinate vectors $\boldsymbol{e}_i$ are linearly independent since

$$\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2 + \cdots + \alpha_n \boldsymbol{e}_n = \mathbf{0}$$
$$\Rightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix} \Rightarrow \alpha_1 = \alpha_2 = \cdots\cdots = \alpha_n = 0.$$

(3) In $\mathbb{R}^3$ vectors $\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ are linearly dependent since

$$\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(4) Any set of the vectors containing the zero vector $\mathbf{0}$ is linearly dependent eg. if $S = \{\boldsymbol{v}_1, ..., \boldsymbol{v}_m, \mathbf{0}\} \subseteq \mathbb{R}^n$ then

$$0 \cdot \boldsymbol{v}_1 + \cdot + 0 \cdot \boldsymbol{v}_m + 1 \cdot \mathbf{0} = \mathbf{0}$$

$\Rightarrow$ S linearly dependent.

Definition: Let V be a vector space. We say that a set of vectors $B = \{\boldsymbol{v}_1, ..., \boldsymbol{v}_m\} \subseteq V$ is a *basis* for V if B *spans* V *and* B is linearly independent.

NOTE: Usually work with a basis for a vector space V because it contains the smallest number of vectors required to span V – but V may have many different bases.

Properties:

If $B = \{\boldsymbol{v}_1, ..., \boldsymbol{v}_m\}$ is a basis for V then

(i) Every basis for V has m vectors – we call m the *dimension* of V and write $m = \dim V$. The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero.

(ii) Any m linearly indept. vectors in V must form a basis

(iii) Any $m+1$ vectors in V must be linearly dependent.

(iv) If $\boldsymbol{v} \in V$, then the representation of $\boldsymbol{v}$ w.r.t. B is *unique*; ie. if

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots\cdots + \alpha_m\boldsymbol{v}_m \quad \text{and also} \quad \boldsymbol{v} = \beta_1\boldsymbol{v}_1 + \cdots\cdots + \beta_m\boldsymbol{v}_m$$

then $\alpha_i = \beta_i \quad , \quad i = 1, ..., m.$

Proof:

$$\begin{aligned} \mathbf{0} &= \boldsymbol{v} - \boldsymbol{v} = (\alpha_1\boldsymbol{v}_1 + \cdots + \alpha_m\boldsymbol{v}_m) - (\beta_1\boldsymbol{v}_1 + \cdots + \beta_m\boldsymbol{v}_m) \\ &= (\alpha_1 - \beta_1)\boldsymbol{v}_1 + \cdots + (\alpha_m - \beta_m)\boldsymbol{v}_m \end{aligned}$$

therefore B linearly independent $\Rightarrow \alpha_i - \beta_i = 0,\ i = 1, ..., m.$

NOTES: If $\boldsymbol{v} = \alpha\boldsymbol{v}_1 + \cdots + \alpha_m\boldsymbol{v}_m$ we call α_i the ith *component* of $\boldsymbol{v}$ w.r.t. B.

Examples: (1) Have seen that the coordinate vectors $\boldsymbol{e}_i$, $i = 1, ..., n$ are linearly independent and span $\mathbb{R}^n$ $\therefore$ form a basis for $\mathbb{R}^n$

$$\Rightarrow \dim \mathbb{R}^n = n.$$

(2) The vector space

$$V = \{\alpha\begin{pmatrix}1\\2\end{pmatrix} \mid \alpha \in \mathbb{R}\} \subseteq \mathbb{R}^2$$

is spanned by a single vector $\begin{pmatrix}1\\2\end{pmatrix}$ which thus constitutes a basis for V. Hence $\dim V = 1$.

(3) Have seen that vectors $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}$ are linearly independent in $\mathbb{R}^2$. Since $\dim \mathbb{R}^2 = 2$, these vectors must form a basis for $\mathbb{R}^2$. In fact every vector $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ may be expanded

$$\boldsymbol{v} = \frac{1}{2}(v_1 + v_2)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}(v_1 - v_2)\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

so $\boldsymbol{v}$ has components $\frac{1}{2}(v_1 + v_2)$, $\frac{1}{2}(v_1 - v_2)$ in this basis.

(4) More generally, any n linearly independent vectors in $\mathbb{R}^n$ must form a basis eg. in $\mathbb{R}^3$ vectors

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

are linear independent (Exercize) and hence must form a basis for $\mathbb{R}^3$.

Matrix Algebra

Historically matrices were introduced to solve simultaneous linear equations of the form:

$$\begin{aligned} 2x - 3y + z &= -13 \\ x + 4y &= 0 \\ x - 2y + 3z &= 2 \end{aligned}$$

– occur in many real life problems. Since then matrices have found many other applications in Abstract Algebra, Quantum Mechanics, Applied Mathematics, Chemistry, Biology and Engineering.

Definition: A rectangular array of numbers

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1n} \\ A_{21} & A_{22} & \cdots\cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mn} \end{pmatrix}$$

is called an $m \times n$ matrix – made up of m rows and n columns. Entries of jth column may be assembeled into a vector

$$\begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{pmatrix} -$$

called the jth *column vector* of A. Similarly the entries of the ith row may be assemble into a row vector

$(A_{i1}\, A_{i2} \cdots\cdots A_{in})$ – called the ith *row vector.*

A_{ij} is called the $(i,\, j)$ *entry* of A – ie. the number in the ith row and jth column. For brevity we write $A = (A_{ij})$.

Examples: (1) The 2×3 matrix with entries $A_{ij} = i - j$ is

$$A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \end{pmatrix}.$$

(2) The $m \times n$ matrix A whose entries are all zero is called the *zero matrix*, denoted O; eg. the zero 3×2 matrix is

$$O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Equality: Matrices A, B are said to be *equal*, written $A = B$, if they have the same size and $A_{ij} = B_{ij}$, for all i, j.

Addition: The *sum* of two $m \times n$ matrices A and B is defined to be the $m \times n$ matrix $A + B$ with entries

$$(A + B)_{ij} = A_{ij} + B_{ij}$$

– ie. we add matrices element wise, as for vectors.

NOTE: Addition is only defined between matrices of the *same* size.

Properties: For matrices A, B, C of the same size we have

$$\begin{aligned} A + B &= B + A \\ (A + B) + C &= A + (B + C). \end{aligned}$$

Scalar Multiplication

Let A be an $m \times n$ matrix and $\alpha \in \mathbb{R}$. We define $\alpha A = A\alpha$ to be the $m \times n$ matrix with entries

$$(\alpha A)_{ij} = (A\alpha)_{ij} = \alpha \cdot A_{ij} \quad , \quad \text{for all } i, j.$$

Throughout we write $-A$ for $(-1) \cdot A$. We may then define *subtraction* between matrices of the *same* size by

$$A - B = A + (-B).$$

Examples: (1) Let

$$A = \begin{pmatrix} -4 & 6 & 3 \\ 0 & 1 & 2 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{pmatrix}.$$

Then

$$A + B = \begin{pmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{pmatrix} \quad , \quad A - B = \begin{pmatrix} -9 & 7 & 3 \\ -3 & 0 & 2 \end{pmatrix}.$$

(2) If

$$\begin{aligned} A &= \begin{pmatrix} 6 & 0 \\ 2 & -1 \end{pmatrix} \quad \text{then} \\ A + A &= \begin{pmatrix} 12 & 0 \\ 4 & -2 \end{pmatrix} = 2 \cdot A \\ \frac{1}{2} A &= \begin{pmatrix} 6 & 0 \\ 2 & -1 \end{pmatrix} \quad \text{etc.} \end{aligned}$$

Matrix Multiplication

Let $A = (A_{ij})$ be an $m \times n$ and $B = (B_{jk})$ an $n \times r$ matrix. Then AB is the $m \times r$ matrix with entries

$$(AB)_{ik} = A_{i1}B_{1k} + A_{i2}B_{2k} + \cdots\cdots + A_{in}B_{nk} \ :$$

$$\underset{(m \times n)}{\underset{A}{\begin{pmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1n} \\ \vdots & \vdots & & \vdots \\ A_{i1} & A_{i2} & \cdots\cdots & A_{in} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mn} \end{pmatrix}}} \underset{(n \times r)}{\underset{B}{\begin{pmatrix} B_{11} & \cdots\cdots & B_{1k} & \cdots\cdots & B_{1r} \\ B_{21} & \cdots\cdots & B_{2k} & \cdots\cdots & B_{2r} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ B_{n1} & \cdots\cdots & B_{nk} & \cdots\cdots & B_{nr} \end{pmatrix}}} = \underset{m \times r}{AB}$$

where $(AB)_{ik}$ = dot product between ith row of A and jth column of B.

NOTE: AB is only defined if number columns $A =$ number rows of B.

Examples: (1) $A = \underset{2\times2}{\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}}$, $B = \underset{2\times2}{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}$ $\therefore AB$ defined and

$$AB = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0+3 & 0+4 \\ 2-3 & 4-4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -1 & 0 \end{pmatrix}$$

(2) $A = \underset{1 \times 3}{(-1 \ 0 \ 1)}$, $B = \underset{3 \times 2}{\begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & \frac{1}{2} \end{pmatrix}}$ $\therefore$ AB defined and

$$AB = (-1 \ 0 \ 1) \begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & \frac{1}{2} \end{pmatrix} = (0 + 0 + -2 \ \ 1 + 0 + \frac{1}{2}) = (-2 \ \ \frac{3}{2})$$

but $\underset{3\times 2}{B}\ \underset{1\times 3}{A}$ is *not* defined.

(4)

$$\underset{1\times 3}{(1\ 3\ 9)}\ \underset{3\times 1}{\begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix}} = 2+3+63 = 68$$

$$\underset{3\times 1}{\begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix}}\ \underset{1\times 3}{\begin{pmatrix} 1 & 3 & 9 \end{pmatrix}} = \begin{pmatrix} 2 & 6 & 18 \\ 1 & 3 & 9 \\ 7 & 21 & 63 \end{pmatrix}.$$

Properties: For matrices of appropriate size

(1) $(AB)C = A(BC)$

(2) $(A+B)C = AC + BC$

$A(B+C) = AB + AC.$

NOTES: (1) Even if AB, BA are both defined, in general

$$AB \neq BA$$

ie. matrix mult. is *not* commutative

$$\text{eg.} \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

but

$$BA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq AB.$$

(2) Another unusual property of matrices is that $AB = 0$ does *not* imply $A = 0$ or $B = 0$ ie. it is possible for the product of two non-zero matrices to be zero:

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Transposition:

The *transpose* of an $m \times n$ matrix $A = (A_{ij})$ is the $n \times m$ matrix A^T with entries

$$A^T_{ji} = A_{ij} \quad , \quad \text{for all } i,\, j$$

ie.

$$\begin{pmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1n} \\ A_{21} & A_{22} & \cdots\cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mn} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & \cdots\cdots & A_{m1} \\ A_{12} & A_{22} & \cdots\cdots & A_{m2} \\ \vdots & & & \vdots \\ A_{1n} & A_{2n} & \cdots\cdots & A_{mn} \end{pmatrix}$$

– ie. row vectors of A become column vectors of A^T and vice versa.

Examples: (1) $\begin{pmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{pmatrix}$

(2) $\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$ (3) $(7 \;\; 5 \;\; -2)^T = \begin{pmatrix} 7 \\ 5 \\ -2 \end{pmatrix}$.

Properties: For matrices of appropriate size

(1) $(\alpha A)^T = \alpha \cdot A^T \quad , \quad \alpha \in \mathbb{R}$

(2) $(A + B)^T = A^T + B^T$

(3) $(A^T)^T = A$

(4) $(AB)^T = B^T A^T$.

NOTE ON DOT PRODUCTS:

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

may be interpreted as $n \times 1$ matrices. Then their dot product may be expressed using matrix mult.:

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{w} &= v_1 w_2 + v_2 w_2 + \cdots + u_n w_n = \underset{1 \times n}{(v_1 v_2 \cdots v_n)} \underset{n \times 1}{\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}} \\ &= \boldsymbol{v}^T \boldsymbol{w}. \end{aligned}$$

Square Matrices and Inverses

An $n \times n$ matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

is called a *square matrix* of *order* n. The diagonal containing the entries

$$A_{11}, A_{22}, \cdots, A_{nn}$$

is called the *principal diagonal* of A. If the entries above (resp. below) this diagonal are all zero then A is called *lower* (resp. *upper*) triangular eg.

$$\underset{\text{lower } \Delta^{ar}}{\begin{pmatrix} A_{11} & 0 & 0 & \cdots & 0 \\ \vdots & A_{22} & 0 & \cdots & 0 \\ \vdots & & \ddots & & 0 \\ A_{n1} & \cdots & & \cdots & A_{nn} \end{pmatrix}} \qquad \underset{\text{upper } \Delta^{ar}}{\begin{pmatrix} A_{11} & \cdots & & \cdots & A_{1n} \\ 0 & A_{22} & & & \vdots \\ & & \ddots & & \vdots \\ 0 & 0 & & & \\ \vdots & & & & \\ 0 & 0\cdots & & 0 \ddots & A_{nn} \end{pmatrix}}$$

If elements above *and* below principal diagonal are zero, so

$$A_{ij} = 0 \quad , \quad i \neq j$$

then A is called a *diagonal matrix*.

Example: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ is a diagonal 3×3 matrix.

Definition: The $n \times n$ identity matrix $I = I_n$, is the diagonal matrix whose entries are all 1: ie.

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \end{pmatrix}.$$

NOTES: (1) It is easily checked that I acts as the identity for $n \times n$ matrices:

$$I \cdot A = A \cdot I = A$$

eg. for 2×2 case have

$$I \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A = A \cdot I.$$

(2) More generally if A is $m \times n$ then

$$I_m \cdot A = A \cdot I_n = A.$$

(3) The jth column vector of I is the coordinator vector

$$\boldsymbol{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j.$$

MATRIX POWERS: For a square matrix A we may define matrix powers

$$A^2 = A \cdot A \quad , \quad A^3 = A^2 \cdot A, \ldots, A^{k+1} = A^k \cdot A$$

and usually set

$$A^1 = A \quad , \quad A^0 = I.$$

Then usual exponent laws hold: eg.

$$A^k \cdot A^r = A^{k+r} \quad , \quad (A^k)^r = A^{kr}.$$

Example: $A = \begin{pmatrix} -1 & -4 \\ 1 & 3 \end{pmatrix} \quad \Rightarrow \quad A^2 = A \cdot A = \begin{pmatrix} -3 & -8 \\ 2 & 5 \end{pmatrix}$

$$A^4 = (A^2)^2 = A^2 \cdot A^2 = \begin{pmatrix} -7 & -16 \\ 4 & 9 \end{pmatrix}.$$

Definition: A square matrix A is non-*singular* (or *invertible*) of there exists a matrix B s.t.

$$AB = BA = I.$$

B is called the *inverse* if A – denoted A^{-1}.

It is important to note that not every matrix has an inverse. In fact

Theorem: A has an inverse if and only if the columns of A are linearly independent.

Proof: First note

$$A\boldsymbol{e}_j = \begin{pmatrix} A_{11} & \cdots & A_{1j} & \cdots & A_{1n} \\ A_{21} & \cdots & A_{2j} & \cdots & A_{2n} \\ \vdots & & \vdots & & \vdots \\ A_{n1} & \cdots & A_{nj} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{nj} \end{pmatrix} \quad - \text{ jth column vector of A.}$$

If A has an inverse these column vectors are linear independent. To see this consider

$$\alpha_1 A\boldsymbol{e}_1 + \alpha_2 A\boldsymbol{e}_2 + \cdots + \alpha_n A\boldsymbol{e}_n = \mathbf{0}$$
$$\Rightarrow \mathbf{0} = A(\alpha\boldsymbol{e}_1 + \alpha_2\boldsymbol{c}_2 + \cdots + \alpha_n\boldsymbol{e}_n) = A\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}.$$

Mult. on left by A^{-1} to give

$$\mathbf{0} = \underbrace{A^{-1}\cdot A}_{I}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix} = I\cdot\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}$$
$$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

so column vectors of A are linear independent. Conversely, can be shown that if columns of A are linear independent then A has an inverse.

Examples: (1) From a previous example, columns of

$$A = \begin{pmatrix}1 & 1 & 0\\ 3 & 1 & 1\\ 1 & 1 & 0\end{pmatrix}$$

are linearly dependent therefore A is singular (ie has no inverse).

(2) But columns of

$$A = \begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}$$

are linearly independent therefore A has an inverse (ie. is non-singular).

Properties:

(1) A has at most **one** inverse

(2) If A, B are non singular so too is AB and

$$(AB)^{-1} = B^{-1}A^{-1}$$

(3) A is non-singular if and only if A^T is non-singular

(4) $(A^T)^{-1} = (A^{-1})^T$.

(5) A is non-singular if and only if the rows of A are linearly independent.

Proof: (1) Suppose B, C are both inverses of A, so

$$AB = BA = I \quad , \quad AC = CA = I.$$

Then $B = B \cdot I = B \cdot (AC) = (BA)C = I \cdot C = C$
Therefore inverse, if it exists, is *unique*.

(2) Suppose A, B have inverses A^{-1}, B^{-1} respectively. Then

$$\begin{aligned}(AB) \cdot B^{-1}A^{-1} &= A \cdot (B \cdot B^{-1}) \cdot A^{-1} \\ &= A \cdot I \cdot A^{-1} = A \cdot A^{-1} = I\end{aligned}$$

and similarly

$$(B^{-1}A^{-1}) \cdot (AB) = I$$

Therefore AB is invertible with inverse

$$(AB)^{-1} = B^{-1}A^{-1}.$$

(3) and (4) are left as an exercise (Assignment 4)

(5)

$$\begin{aligned} A \text{ is non-singular} &\Leftrightarrow A^T \text{ is non-singular} \\ &\Leftrightarrow \text{columns } A^T \text{ are lin.indept.} \\ &\Leftrightarrow \text{rows } A \text{ are lin.indept.}\end{aligned}$$

2 × 2 Case:

For a 2×2 matrix $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, set

$$\Delta = A_{11}A_{22} - A_{12}A_{21}.$$

If $\Delta \neq 0$, it is easily seen by direct mult. that A is non-singular with inverse

$$A^{-1}\frac{1}{\Delta}\begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

– the problem of obtaining inverses of higher order square matrices will be considered later.

Example: $A = \begin{pmatrix} 2 & -\frac{2}{3} \\ 6 & 4 \end{pmatrix} \quad \Rightarrow \quad \begin{aligned} \Delta &= 2 \times 4 - 6 \times (-2/3) \\ &= 8 + 4 = 12 \end{aligned}$

Therefore $A^{-1} = \dfrac{1}{12}\begin{pmatrix} 4 & \frac{1}{2} \\ -6 & 2 \end{pmatrix}$.

Check: $A \cdot A^{-1} = A^{-1} \cdot A = I$.

Simultaneous Equs. and Gauss elimination.

Here we consider simultaneous equations of form

$$\begin{aligned}
A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\
A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\
\vdots \qquad\qquad\qquad &\quad \vdots \\
A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m,
\end{aligned}$$

where coefficients A_{ij} and the b_i are given real numbers, which we wish to solve for unknowns $x_1, x_2, \cdots\cdots, x_n$. Such a set of equations is most conveniently expressed in matrix form

$$A\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

where $A = (A_{ij})$ is the $m \times n$ matrix of coefficients, $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is a column vector of unknowns and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ is a given $m \times 1$ matrix (column vector) – this in fact motivates definition of matrix multiplication.

If A is $n \times n$ and non-singular, may solve (1) by mult. on left by A^{-1} to give $\underbrace{A^{-1} \cdot A}_{I}\, \boldsymbol{x} = A^{-1}\boldsymbol{b} \Rightarrow$ unique solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ – works if A^{-1} is known.

FACT: $(\boldsymbol{b} = \boldsymbol{0})$: $\boldsymbol{x} = \boldsymbol{0}$ is the unique solution to system of equations

$$\underset{n\times n}{A}\boldsymbol{x} = \boldsymbol{0}$$

iff A is non-singular.

Proof: Above equation is expressible

$$A(x_1\boldsymbol{e}_1 + \cdots\cdots + x_n\boldsymbol{e}_n) = \boldsymbol{0}$$

or

$$x_1A\boldsymbol{e}_1 + x_2A\ \boldsymbol{e}_2 + \cdots\cdots + x_nA\boldsymbol{e}_n = \boldsymbol{0}$$

where $A\boldsymbol{e}_j = j$th column vector of A. Therefore $\boldsymbol{x} = \boldsymbol{0}$ unique solution $\Leftrightarrow$ column vectors of A linear independent $\Leftrightarrow$ A is non-singular. $\square$

In the case $\boldsymbol{b} \neq \boldsymbol{0}$ this method fails if A^{-1} is not known or if A is singular. Moreover it does not apply to the general case if A is $m \times n$.

The most efficient way of solving the system (1) is by *Gaussian elimination.*

Illustrative Example

Recall Ohm's law:

DIAGRAM 92

$$\underset{\text{(volts)}}{V} = \underset{\text{(amps)}}{I}\,\underset{\text{(ohms)}}{R}$$

and the following rules for voltages:

DIAGRAM 93 DIAGRAM 94

Now consider the following electrical network

DIAGRAM 95

so that $I_1 = I_2 + I_3$.

To find unknown currents I_1, I_2, I_3 have

$$\left.\begin{array}{lrl} (r_1) & I_1 - I_2 - I_3 & = 0 \\ (r_2) & 4I_1 + 2I_2 & = 12 \\ (r_3) & 2I_2 - 2/3I_3 & = 0 \end{array}\right\} \qquad (1)$$

To solve system (1) we eliminate I_1 from the 2nd. equation by subtracting appropriate multiple of r_1 to give:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 \\ 6I_2 + 4I_3 &= 12 \qquad (r_2 \to r_2 - 4r_1) \\ 2I_2 - 2/3I_3 &= 0. \end{aligned}$$

Next divide r_2 by 6 to make coefficients of leading term I_2 1:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 \\ I_2 + 2/3I_3 &= 2 \qquad (r_2 \to 1/6r_2) \\ 2I_2 - 2/3I_3 &= 0 \end{aligned}$$

We then eliminate I_2 from last equation by subtracting appropriate multiple of r_2 to give:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 \\ I_2 + 2/3I_3 &= 2 \\ -2I_3 &= -4 \qquad (r_3 \to r_3 - 2r_2). \end{aligned}$$

Finally divide last equation by -2 to make coefficient of leading term I_3 1:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 \\ I_2 + 2/3 I_3 &= 2 \\ I_3 &= 2 \end{aligned}$$

which completes the Gaussian elimination. We call this last set of equations, which is equivalent to the original set (1), the *reduced form* of the equations. Once in reduced form the set of equations is easily solved.

For the above example we obtain $I_3 = 2$ (amps), and back substituting into 2nd equation gives $I_2 = 2 - 2/3 I_3 = 2/3$ (amps) and finally back substituting into 1st equation gives

$$I_1 = I_2 + I_3 = 8/3 \quad \text{(amps)}.$$

Gaussian elimination using Matrices

The above proceedure may be described most conveniently in terms of matrices. Given a set of simultaneous equations

$$\underset{m\times n}{A}\ \underset{n\times 1}{\boldsymbol{x}} = \underset{m\times 1}{\boldsymbol{b}}$$

we construct the new matrix

$$(A \vdots \boldsymbol{b})$$

called the *Augmented matrix.* Instead of manipulating equations we manipulate the corresponding Augmented matrix.

The Augmented matrix for the set of equations (1) may be written

$$\begin{pmatrix} 1 & -1 & -1 & \vdots & \\ 4 & 2 & 0 & \vdots & 12 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix}$$

The 1st step is to divide r_1 by appropriate number to make leading entry 1 – in this case leading entry of 1st row is already 1.

Next step is to eliminate leading entries of last 2 rows by subtracting appropriate multiple of r_1 – called *eliminating the first column*: we write

$$\begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 4 & 2 & 0 & \vdots & 12 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 6 & 4 & \vdots & 12 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix} \quad (r_2 \to r_2 - 4r_1)$$

which indicates the two matrices are related by the *elementary row operation*: $r_2 \longrightarrow r_2 - 4r_1$. Continuing the elimination as before we have

$$\begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 6 & 4 & \vdots & 12 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 1 & \frac{2}{3} & \vdots & 2 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix} \quad (r_2 \to \tfrac{1}{6}r_2)$$

$$\sim \begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 1 & \frac{2}{3} & \vdots & 2 \\ 0 & 0 & -2 & \vdots & -4 \end{pmatrix} \quad (r_3 \to r_3 - 2r_2) \quad \text{(elimination of 2nd column)}$$

$$\sim \begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 1 & \frac{2}{3} & \vdots & 2 \\ 0 & 0 & 1 & \vdots & 2 \end{pmatrix} \quad (r_3 \to -\frac{1}{2}r_2)$$

– called the *Gauss reduced form* of the original Augmented matrix. The solution can now be read off directly by back substitution as before: $I_3 = 2$ and back substitution into r_2 and r_1 gives $I_2 = \frac{2}{3}$, $I_1 = \frac{8}{3}$ (amps).

Elementary Row Operations (EROs)

In general we say that $m \times n$ matrices A, B are *row equivalent*, written $A \sim B$, if and only if they are related by a series of EROs where an ERO is one of the following operations on the rows of A:

(1) Interchange rows i and j written $r_i \leftrightarrow r_j$.

(2) Replace r_i with αr_i, $0 \neq \alpha \in \mathbb{R}$: written $r_i \to \alpha r_i$

(3) Replace r_i with $r_i + \alpha r_j$, $\alpha \in \mathbb{R}$, where r_j is *another* row of A: written $r_i \longrightarrow r_i + \alpha r_j$.

EROs applied to a system of equations have no effect on the solution. So the idea is to apply EROs to Augmented matrix to reduce it to Gauss reduced form from which solution can be read off – most efficient proceedure for solving equations and is used in most computer codes.

Example: Every matrix A is row equivalent to its Gauss reduced form – in Gauss reduced form an $m \times n$ matrix A has diagram entries A_{11}, A_{22}, ... which are all zeros or 1's with zeros below this diagonal

NOTE: Matrix A may have more than one Gauss reduced form – doesn't matter which is used.

In the last example we never made use of row interchanges – often simplify elimination and are sometimes necessary.

Examples (1): Solve system of equations

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 2 \\ -x_1 - 3x_2 + 2x_3 &= -2 \\ 2x_1 + 4x_2 - 2x_3 &= 2 \end{aligned}$$

Solution: Applying Gauss elimination to Aug. Matrix gives

$$\left(\begin{array}{rrr:r} 1 & 3 & -2 & 2 \\ -1 & -3 & 2 & -2 \\ 2 & 4 & -2 & 2 \end{array}\right) \sim \left(\begin{array}{rrr:r} 1 & 3 & -2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 2 & -2 \end{array}\right) \quad \left.\begin{array}{l} \\ (r_2 \to r_2 + r_1) \\ (r_3 \to r_3 - 2r_1) \end{array}\right\} \text{elimination of 1st col.}$$

In this case cannot eliminate 2nd column by subtracting appropriate multiple of r_2 from r_3 since have 0 in (1, 2) position. Thus need to interchange rows 2 and 3 to complete elimination:

$$\sim \left(\begin{array}{rrr:r} 1 & 3 & -2 & 2 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) (r_2 \longrightarrow r_3) \sim \left(\begin{array}{rrr:r} 1 & 3 & -2 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) (r_2 \to -\frac{1}{2}r_2) - \text{Gauss reduced form.}$$

Last row is equivalent to

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 0 \quad \text{– gives no information.}$$

Therefore effectively have just 2 equations in 3 unknowns

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 2 \\ x_2 - x_3 &= 1 \end{aligned}$$

so expect ∞ number of solutions. To solve set

$$x_3 = \alpha\,, \qquad \text{arbitrary}$$

and back substitute into r_2 and r_1 to give

$$\begin{aligned} x_2 &= 1 + x_3 = 1 + \alpha\,, \\ x_2 &= 2 - 3x_2 + 2x_3 = -1 - \alpha \end{aligned}$$

$\Rightarrow$ general solution

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1-\alpha \\ 1+\alpha \\ \alpha \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

$\alpha \in \mathbb{R}$ – ie. ∞ numbers of solutions.

(2) Find General solution to

$$\begin{aligned} x_1 + 2x_2 - x_3 + 2x_4 &= 4 \\ 2x_1 + 7x_2 + x_3 + x_4 &= 14 \\ 3x_1 + 8x_2 - x_3 + 4x_4 &= 17 \end{aligned}$$

Solution: Aug. Matrix is

$$\begin{pmatrix} 1 & 2 & -1 & 2 & \vdots & 4 \\ 2 & 7 & 1 & 1 & \vdots & 14 \\ 3 & 8 & -1 & 4 & \vdots & 17 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 2 & \vdots & 4 \\ 0 & 3 & 3 & -3 & \vdots & 6 \\ 0 & 2 & 2 & -2 & \vdots & 5 \end{pmatrix} \begin{matrix} (r_2 \to r_2 - 2r_1) \\ (r_3 \to r_3 - 3r_1) \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 2 & -1 & 2 & \vdots & 4 \\ 0 & 1 & 1 & -1 & \vdots & 2 \\ 0 & 2 & 2 & -2 & \vdots & 5 \end{pmatrix} (r_2 \to \tfrac{1}{3} r_2) \sim \begin{pmatrix} 1 & 2 & -1 & 2 & \vdots & 4 \\ 0 & 1 & 1 & -1 & \vdots & 2 \\ 0 & 0 & 0 & 0 & \vdots & 1 \end{pmatrix}_{(r_3 \to r_3 - 2r_1)} \quad \text{– Gauss reduces form}$$

Bottom line is equivalent to equation

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 1$$

which is impossible therefore *no solutions.*

NOTE: In general a system of equations

$$\underset{m \times n}{A} \; \underset{n \times 1}{\boldsymbol{x}} = \underset{m \times 1}{\boldsymbol{b}}$$

has following possibilities for solution:

(1) A unique solution

(2) An ∞ number of solutions

(3) No solution.

Example (3): Determine whether following vectors are linearly independent

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}.$$

Solution: Given vectors form columns of matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 1 & 0 \end{pmatrix}$$

$\therefore$ Vectors linear independent $\Leftrightarrow \boldsymbol{x} = \mathbf{0}$ is unique solution to $A\boldsymbol{x} = \mathbf{0}$
$\therefore$ Apply Gauss elimination to $A\boldsymbol{x} = \mathbf{0}$.

Augmented matrix: $\begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 3 & \vdots & 0 \\ 1 & 1 & 0 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 3 & \vdots & 0 \\ 0 & -1 & 0 & \vdots & 0 \end{pmatrix}_{(r_3 \to r_3 - r_1)}$

$$\sim \begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 3 & \vdots & 0 \\ 0 & 0 & 3 & \vdots & 0 \end{pmatrix}_{(r_3 \to r_3 + r_2)} \sim \begin{pmatrix} 1 & 2 & 0 & \vdots & 0 \\ 0 & 1 & 3 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 0 \end{pmatrix}$$

$\Rightarrow x_3 = 0$ and back substitution gives unique solution

$$\boldsymbol{x} = \mathbf{0}.$$

$\therefore$ Given vectors linearly independent.

Gauss-Jordan elimination

Find solution of

$$\begin{aligned} x_1 - x_2 - x_3 &= 0 \\ 4x_1 + 2x_2 &= 12 \\ 2x_2 - \frac{2}{3}x_3 &= 0. \end{aligned}$$

Solution: Augmented matrix is

$$\begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 4 & 2 & 0 & \vdots & 12 \\ 0 & 2 & -\frac{2}{3} & \vdots & 0 \end{pmatrix} \quad \text{– same as in electrical network example}$$

$$\sim \begin{pmatrix} 1 & -1 & -1 & \vdots & 0 \\ 0 & 1 & \frac{2}{3} & \vdots & 2 \\ 0 & 0 & 1 & \vdots & 2 \end{pmatrix} \quad \text{–Gauss reduced form.}$$

In this case elimination can be carried further:

$$\sim \begin{pmatrix} 0 & -1 & 0 & \vdots & 2 \\ 0 & 1 & 0 & \vdots & \frac{2}{3} \\ 0 & 0 & 1 & \vdots & 2 \end{pmatrix} \begin{matrix} (r_1 \to r_1 + r_3) \\ (r_2 \to_2 -\frac{2}{3} r_3) \end{matrix} \sim \begin{pmatrix} 1 & 0 & 0 & \vdots & \frac{8}{3} \\ 0 & 1 & 0 & \vdots & \frac{2}{3} \\ 0 & 0 & 1 & \vdots & 2 \end{pmatrix} \quad \begin{matrix} (r_1 \to r_1 + r_2) \\ \text{– Jordan} \\ \text{reduced form} \end{matrix}$$

Has advantage that solution $\boldsymbol{x} = \begin{pmatrix} 8/3 \\ 2/3 \\ 2 \end{pmatrix}$ can be read off directly. However Jordan reduction involves more arithmetic than Gaussian reduction which is usually preferred. More over Jordan reduction cannot always be completed – eg. if have a row of zeros in bottom line.

Example: Augmented matrix for Example (1) was

$$\begin{pmatrix} 1 & 3 & -2 & \vdots & 2 \\ -1 & -3 & 2 & \vdots & -2 \\ 2 & 4 & -2 & \vdots & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -1 & \vdots & 2 \\ 0 & 1 & -1 & \vdots & 1 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix} \quad \text{(Gauss red. form)}$$

Cont. elimination:

$$\sim \begin{pmatrix} 1 & 1 & 0 & \vdots & 0 \\ 0 & 1 & -1 & \vdots & 1 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix} \quad (r_1 \to r_1 - 2r_3) \; -$$

Jordan reduction cannot be completed in this case.

NOTE: However Gauss–Jordan reduction is useful for obtaining inverses of matrices.

Inverses using Gaiss–Jordan reduction

Let A be an $n \times n$ matrix. The most efficient way of obtaining the inverse of A is to solve the matrix equation.

$$AX = I \tag{1}$$

where $X = (x_{ij})$ is an $n \times n$ matrix of unknowns. This system actually represents n sets of simultaneous equations

$$A\boldsymbol{x}_j = \boldsymbol{e}_j$$

where $\boldsymbol{x}_j$ is the jth column of X and $\boldsymbol{e}_j$ (coordinate vector) is the jth column of I. Thus may solve (1) by applying Gauss–Jordan elimination to Augmented matrix

$$(A \vdots I).$$

Reduce to the form

$$(I \vdots B) - \text{ then } B = A^{-1} \text{ gives inverse of } A.$$

If this reduction cannot be completed, A has no inverse.

Example: Find inverse of $A = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & -2 \\ 1 & 4 & -1 \end{pmatrix}$

Solution: Augmented Matrix is

$$(A \,\vdots\, I) = \left(\begin{array}{ccc|ccc} -1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 1 & 4 & -1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 6 & 0 & 1 & 0 & 1 \end{array}\right) \begin{matrix} (r_1 \to -r_1) \\ \\ (r_3 \to r_3 + r_1) \end{matrix}$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 12 & 1 & -6 & 1 \end{array}\right)_{(r_3 \to r_3 - 6r_2)} \sim \left(\begin{array}{ccc|ccc} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{12} & -\frac{1}{12} & \frac{1}{12} \end{array}\right) \text{(Gauss red. form)}$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & -\frac{11}{12} & -\frac{1}{2} & \frac{1}{12} \\ 0 & 1 & 0 & \frac{1}{6} & 0 & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{12} & -\frac{1}{2} & \frac{1}{12} \end{array}\right)^{(r_1 \to r_1 + 3)}_{(r_2 \to r_2 + 2r_3)} \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{7}{12} & -\frac{1}{2} & \frac{5}{12} \\ 0 & 1 & 0 & \frac{1}{6} & 0 & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{12} & -\frac{1}{2} & \frac{1}{12} \end{array}\right)^{(r_1 \to r_1 + 2r_2)}$$

–Jordan reduced form.

Solution can now be read off directly to give

$$A^{-1} = \begin{pmatrix} -7/12 & -1/2 & 5/12 \\ 1/6 & 0 & 1/6 \\ 1/12 & -1/2 & 1/12 \end{pmatrix}$$

Check: $A \cdot A^{-1} = A^{-1} \cdot A = I$.

Determinants

For a square matrix A, define the *number* $|A|$, also denoted $\det(A)$, called the *determinant* of A by:

(i) If $A = (A_{ij})$ is 2×2 then

$$|A| = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{12} - A_{12}A_{21}$$

(ii) In general if $A = (A_{ij})$ is $n \times n$ we define $|A|$ as follows: first set

$$C_{ij} = (-1)^{i+j} \underbrace{\begin{vmatrix} A_{11} & A_{12} & \dots & A_{1j} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2j} & \dots & A_{2n} \\ \vdots & \vdots & & & & \vdots \\ A_{i1} & A_{i2} & \dots & A_{ij} & \dots & A_{in} \\ \vdots & \vdots & & & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nj} & \dots & A_{nn} \end{vmatrix}} \quad - \text{ called the } \textit{cofactor} \text{ of} A_{ij} \text{ in } A$$

– determinant obtained by omitting ith row and jth column from A.

We then define

$$|A| = A_{11}C_{11} + A_{12}C_{12} + \cdots\cdots + A_{1n}C_{1n}$$

– gives a *recursive* definition of determinant.

NOTE: $(-1)^{i+j}$ gives pattern $\begin{pmatrix} + & - & + & - & \cdot & \cdot & \cdot & \cdot \\ - & + & - & + & \cdot & \cdot & \cdot & \cdot \\ + & - & + & - & \cdot & \cdot & \cdot & \cdot \\ \cdot \\ \cdot\cdot \end{pmatrix}$

Thus for a 3×3 matrix A we have

$$|A| = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13}$$

where

$$C_{11} = \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix}, \quad C_{12} = -\begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix}, \quad C_{13} = \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}$$

Example:

$$\begin{vmatrix} 3 & 5 & 7 \\ 1 & 0 & 2 \\ 0 & 3 & 0 \end{vmatrix} = 3\begin{vmatrix} 0 & 2 \\ 3 & 0 \end{vmatrix} - 5\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} + 7\begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix}$$

$$\begin{aligned} &= 3(0-6) - 5(0-0) + 7(3-0) \\ &= 3. \end{aligned}$$

Properties of Determinants

(1) $\underline{|A| = |A^T|}$
Consider for example the 2×2 case. Then

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix}$$

$\therefore |A| = A_{11}A_{22} - A_{12}A_{21} = |A^T|$ – holds in general.

NOTES: This means that any results about the *rows* in a general determinant is also true about the columns (since rows of A^T are columns of A).

(2) The determinant may be expanded along any row or down any column eg. for a 3×3 matrix A

$$\begin{aligned} |A| &= A_{11}C_{11} + A_{12}C_{12} + A_{13}C_{13} && \text{(defn.- exp. along 1st row)} \\ &= A_{21}C_{21} + A_{22}C_{22} + A_{23}C_{23} && \text{(expansion along } r_2\text{)} \\ &= A_{13}C_{13} + A_{23}C_{23} + A_{33}C_{33} && \text{(exp. down 3rd col.) etc.} \end{aligned}$$

Examples:

(i)

$$\begin{vmatrix} 3 & 5 & 6 \\ 1 & 0 & 2 \\ 0 & 3 & 0 \end{vmatrix} = 7\begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix} - 2\begin{vmatrix} 3 & 5 \\ 0 & 3 \end{vmatrix} = 21 - 18 = 3 \quad \begin{pmatrix} \text{exp. down} \\ \text{3rd. col.} \end{pmatrix}$$

$$= -3\begin{vmatrix} 3 & 7 \\ 1 & 2 \end{vmatrix} = -3(6-7) = 3 \quad \begin{pmatrix} \text{exp. along} \\ r_3 \end{pmatrix}$$

(ii)

$$\begin{aligned} \begin{vmatrix} 3 & 6 & 9 \\ 0 & 0 & 0 \\ 2 & 4 & 5 \end{vmatrix} &= 0 \cdot C_{21} + 0 \cdot C_{22} + 0 \cdot C_{23} \\ &= 0 \qquad \text{(exp. along } r_2\text{)} \end{aligned}$$

– in general any det. with a row or column of zeros vanishes.

(3) From (2), a common factor in *all* entries of a row (or column) can be taken out as a factor of the determinant.

Example:

$$\begin{aligned} \begin{vmatrix} 3 & 6 & 9 \\ 0 & 1 & 2 \\ 3 & 4 & 5 \end{vmatrix} &= 3 \cdot C_{11} + 6C_{12} + 9C_{13} \qquad \text{(exp. along } r_1\text{)} \\ &= 3[C_{11} + 2C_{12} + 3C_{13}] \\ &= 3 \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 4 & 5 \end{vmatrix} \qquad \text{– } \textit{check}! \end{aligned}$$

(4) If 2 rows (or columns) of A interchanged, $|A|$ changes sign.

Example: Have seen $\begin{vmatrix} 3 & 5 & 7 \\ 1 & 0 & 2 \\ 0 & 3 & 0 \end{vmatrix} = 3$. Interchanging columns 2 and 3 gives $\begin{vmatrix} 3 & 7 & 5 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} =$ $3\begin{vmatrix} 3 & 7 \\ 1 & 2 \end{vmatrix}$ (exp. along $r_3 = -3$).

NOTE: Property (4) implies that any determinant with 2 equal rows (or 2 equal columns) must vanish eg. interchanging c_1 and c_3 gives

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & -1 \\ 3 & 2 & 3 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & -1 \\ 3 & 2 & 3 \end{vmatrix} = 0 \qquad \text{(check!)}$$

– holds in general.

(5) If a multiple of 1 row (or column) is added to another, determinant is unchanged.

Above 5 properties give shortcut for evaluating determinants – use a sequence of last 3 operations to produce a determinant with a row (or column) with at most one non-zero entry and then expand along this row (or column).

Examples: (1)

$$\begin{vmatrix} 2 & 5 & 6 & 1 \\ 1 & 2 & 3 & -2 \\ 3 & 2 & 4 & 4 \\ 0 & 2 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & 5 \\ 1 & 2 & 3 & -2 \\ 0 & -4 & -5 & 10 \\ 0 & 2 & 4 & 6 \end{vmatrix} \quad \begin{matrix} (r_1 \to r_1 - 2r_2) \\ \\ (r_3 \to r_3 - 3r_2) \end{matrix}$$

$$= -1 \cdot \begin{vmatrix} 1 & 0 & 5 \\ -4 & -5 & 10 \\ 2 & 4 & 6 \end{vmatrix} \quad \text{(exp. down col. 1)}$$

$$= -1 \cdot \begin{vmatrix} 1 & 0 & 0 \\ -4 & -5 & 30 \\ 2 & 4 & -4 \end{vmatrix} \quad (c_3 \to c_3 - 5c_1) = -\begin{vmatrix} -5 & 30 \\ 4 & -4 \end{vmatrix} \quad \text{(exp. along } r_1\text{)}$$

$$= -5 \times 4 \cdot \begin{vmatrix} -1 & 6 \\ 1 & -1 \end{vmatrix} = -20(1-6) = 100.$$

(2) Evaluate $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}$ (van der Monde's determinant).

Solution:

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-1 \\ 0 & b^2-a^2 & c^2-a^2 \end{vmatrix} \begin{matrix} \\ (r_2 \to r_2 - ar_1) \\ (r_2 \to r_3 - a^2 r_1) \end{matrix}$$

$$= \begin{vmatrix} b-a & c-a \\ b^2-a^2 & c^2-a^2 \end{vmatrix} \quad (\text{exp. down } c_1)$$

$$= (b-a)(c-a) \begin{vmatrix} 1 & 1 \\ b+a & c+a \end{vmatrix} \begin{matrix} = (b-a)(c-a)[c+a-(b+a)] \\ = (b-a)(c-a)(c-b). \end{matrix}$$

(3) Suppose $A = \begin{pmatrix} A_{11} & 0 & \cdots & \cdots & 0 \\ A_{21} & A_{22} & 0 & & \vdots \\ \vdots & & \ddots & & \vdots \\ A_{n1} & .. & .. & .. & A_{nn} \end{pmatrix}$

is lower Δ^{ar}. Then by repeated exp. using r_1 obtain

$$|A| = A_{11} \begin{vmatrix} A_{22} & & & \\ \vdots & \ddots & & \\ \vdots & & \vdots & \\ A_{n2} & \cdots & \cdots & A_{nn} \end{vmatrix} \begin{matrix} = A_{11}\, A_{22} \cdots A_{nn} \begin{pmatrix} \text{Product of diag.} \\ \text{entries} \end{pmatrix} \\ -\text{ similarly for upper } \Delta^{ar} \\ \text{matrices, and diagonal} \\ \text{matrices.} \end{matrix}$$

(4) In particular

$$|I| = 1.$$

CONNECTION WITH INVERSES AND EQUATIONS

An important property of determinants is

FACT: Let A, B be $n \times n$ matrices. Then

$$|AB| = |A| \cdot |B|.$$

Theorem: *A is non-singular* $\Rightarrow |A| \neq 0$.

Proof: Suppose A has inverse A^{-1} so

$$\begin{aligned} & A \cdot A^{-1} = I \\ \Rightarrow \quad & |A \cdot A^{-1}| = |I| = 1 \\ \therefore \quad & |A| \cdot |A^{-1}| = 1 \quad \Rightarrow \quad |A| \neq 0 \end{aligned}$$

(and $|A^{-1}| = 1/|A|$). Conversely if $|A| \neq 0$ then it can be shown that A is non-singular. Explicitly A^{-1} is given by

$$A_{ij}^{-1} = C_{ji}/|A| = \frac{\text{cofactor of } A_{ji}}{|A|} \text{ – Cramer's Rule.}$$

Corollary: System of equations

$$A\boldsymbol{x} = \mathbf{0}$$

has a solution $\boldsymbol{x} \neq \mathbf{0}$ if and only if $|A| = 0$.

Proof: Have seen that $\boldsymbol{x} = 0$ is unique solution $\Leftrightarrow$ A non-singular $\Leftrightarrow$ $|A| \neq 0$ therefore $\boldsymbol{x} \neq \mathbf{0}$ a solution $\Leftrightarrow$ $|A| = 0$.

NOTES: (1) Cramer's rule is useful for inverses but Gauss–Jordan reduction much more efficient.

(2) Since $|A| = \left|A^T\right|$ above result implies

$$A \text{ non-singular } \Leftrightarrow A^T \text{ non-singular}$$

as noted previously.

Examples: (1) In 2×2 case have

$$|A| = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{22} - A_{12}A_{21}$$

and co-factors

$$C_{11} = A_{22} \quad , \quad C_{12} = -A_{21} \quad , \quad C_{21} = -A_{12} \quad , \quad C_{22} = A_{11}$$

$$\therefore \quad |A| \neq 0 \Rightarrow A^{-1} = \tfrac{1}{|A|}\begin{pmatrix} C_{21} & C_{21} \\ C_{12} & C_{22} \end{pmatrix} = \tfrac{1}{|A|}\begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix}$$

in a agreement with previous result.

(2) Determine whether vectors

$$\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} , \begin{pmatrix} 5 \\ 0 \\ 3 \end{pmatrix} , \begin{pmatrix} 7 \\ 2 \\ 0 \end{pmatrix} \quad \text{are linear independent}$$

Solution: Given vectors form columns of matrix

$$A = \begin{pmatrix} 3 & 5 & 7 \\ 1 & 0 & 2 \\ 0 & 3 & 0 \end{pmatrix}$$

$\therefore$ vectors linear independent $\Leftrightarrow$ A non-singular $\Leftrightarrow$ $|A| \neq 0$. From previous example, $|A| = 3 \neq 0$ $\quad\therefore$ Given vectors are linearly independent.

Vector Products in 3-Space

Definition: For vectors $\boldsymbol{v}$, $\boldsymbol{w}$ in 3-*space* $\mathbb{R}^3$, the *vector* (or *cross*) product of $\boldsymbol{v}$ and $\boldsymbol{w}$ is that vector $\boldsymbol{v} \times \boldsymbol{w}$ s.t.

(i) $\|\boldsymbol{v} \times \boldsymbol{w}\| = \|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\| \sin\theta$, where θ is angle between $\boldsymbol{v}$ and $\boldsymbol{w}$

(ii) $\boldsymbol{v} \times \boldsymbol{w}$ is perp. to $\boldsymbol{v}$ and $\boldsymbol{w}$, in right hand direction from $\boldsymbol{v} \times \boldsymbol{w}$ (right hand rule).

DIAGRAM 96

NOTES: (1) For $\boldsymbol{v}$, $\boldsymbol{w} \neq \boldsymbol{0}$

$\boldsymbol{v} \times \boldsymbol{w} = \boldsymbol{0} \Leftrightarrow \sin\theta = 0 \Leftrightarrow \theta = 0$ or $180°$

$\therefore \quad \boldsymbol{v} \times \boldsymbol{w} = \boldsymbol{0} \Leftrightarrow \boldsymbol{v}$ and $\boldsymbol{w}$ are parallel (or anti-parallel).

(2) $\boldsymbol{v} \times \boldsymbol{w} = -\boldsymbol{w} \times \boldsymbol{v}$ since they have same magnitude but opposite direction – so $\boldsymbol{v} \times \boldsymbol{v} = \boldsymbol{0}$ (as can be seen from (1)).

(3)

If $\boldsymbol{v} \times \boldsymbol{w} \neq \boldsymbol{0}$, then $\boldsymbol{v} \times \boldsymbol{w}$ is perp. to the plane determined by $\boldsymbol{v}$ and $\boldsymbol{w}$ – useful for obtaining equation of a plane in 3-space.

Example:

$$\begin{aligned} \boldsymbol{i} \times \boldsymbol{i} &= \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = \boldsymbol{0} \\ \boldsymbol{i} \times \boldsymbol{j} &= \boldsymbol{k} = -\boldsymbol{j} \times \boldsymbol{i} \quad , \quad \boldsymbol{j} \times \boldsymbol{k} = -\boldsymbol{k} \times \boldsymbol{j} \\ \boldsymbol{k} \times \boldsymbol{i} &= \boldsymbol{j} = -\boldsymbol{i} \times \boldsymbol{k} \end{aligned}$$

DIAGRAM 98

NOTE: Vector products also arise naturally in physics and engineering eg. particle of unit charge moving with velocity $\boldsymbol{v}$ in a magnetic field $\boldsymbol{B}$ experiences a force $\boldsymbol{F} = \boldsymbol{v} \times \boldsymbol{B}$. The *angular momentum* of a particle of mass m moving with velocity $\boldsymbol{v}$ is given by $\boldsymbol{L} = m(\boldsymbol{r} \times \boldsymbol{v})$ where $\boldsymbol{r}$ is the position vector of particle – important for central force problems.

Scalar Triple Product

The *number* $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ is called the *scalar triple product* of $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$. In terms of components we have

$$\begin{aligned}\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w}) &= u_1\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix} - u_2\begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix} + u_3\begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix} \\ &= \begin{vmatrix} v_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \qquad - 3\times 3 \text{ determinant.}\end{aligned}$$

Properties

(1) Have seen that

$$\boldsymbol{v}\times\boldsymbol{w} = -\boldsymbol{w}\times\boldsymbol{v}$$

so commutative law fails. Also we do not have associative law: ie. in general

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

Example: $(\boldsymbol{i}\times\boldsymbol{i})\times\boldsymbol{k} = \boldsymbol{0}\times\boldsymbol{k} = \boldsymbol{0}$, but $\boldsymbol{i}\times(\boldsymbol{i}\times\boldsymbol{k}) = \boldsymbol{i}\times(-\boldsymbol{j}) = -\boldsymbol{k}$.

(2) However it can be shown that

$$\begin{aligned}&\boldsymbol{v}\times(\boldsymbol{w}+\boldsymbol{v}) = \boldsymbol{v}\times\boldsymbol{w}+\boldsymbol{v}\times\boldsymbol{u} \\ &(\boldsymbol{v}+\boldsymbol{w})\times\boldsymbol{u} = \boldsymbol{v}\times\boldsymbol{u}+\boldsymbol{w}\times\boldsymbol{u} \\ &\boldsymbol{v}\times(\alpha\boldsymbol{w}) = (\alpha\boldsymbol{v})\times\boldsymbol{w} = \alpha(\boldsymbol{v}\times\boldsymbol{w}) \quad , \quad \alpha\in\mathbb{R}.\end{aligned}$$

Using above results we have, in terms of components

Theorem: Suppose $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$. Then

$$\begin{aligned}\boldsymbol{v}\times\boldsymbol{w} &= (v_2w_3 - v_3w_2)\boldsymbol{i} + (w_1v_3 - v_1w_3)\boldsymbol{i} + (v_1w_2 - v_2w_1)\boldsymbol{k} \\ \text{ie.}\quad \boldsymbol{v}\times\boldsymbol{w} &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}\boldsymbol{k}\end{aligned}$$

can be evaluated by expanding above determinant along *1st* row.

Proof: Follows directly by expanding

$$\boldsymbol{v}\times\boldsymbol{w} = (v_1\boldsymbol{i} + u_2\boldsymbol{j} + v_3\boldsymbol{k})\times(w_1\boldsymbol{i} + w_2\boldsymbol{j} + w_3\boldsymbol{k}).$$

Application: Find area of ΔABC with vectors $\boldsymbol{v}$, $\boldsymbol{w}$ along edges as shown.

DIAGRAM 99

Solution:

$$\begin{aligned} \text{Area} &= \frac{1}{2}\,\text{base}\ \times\ (\text{perp. ht}\cdot h) \\ &= \frac{1}{2}\,\|\boldsymbol{v}\|\cdot\|\boldsymbol{w}\|\sin\theta = \frac{1}{2}\,\|\boldsymbol{v}\times\boldsymbol{w}\|\,. \end{aligned}$$

Geometrical Interpretation

volume V of parallelepiped determined by vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ along edges as shown is given by

DIAGRAM 100

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

Proof: $V =$ (area of base $ABCD$) $\times\, h$ where

$$\begin{aligned} h &= \text{perp. ht.} \\ &= \text{size of compt. of } \boldsymbol{u} \text{ in direction of } \boldsymbol{v}\times\boldsymbol{w} \\ &= \text{size of} \frac{\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})}{\|\boldsymbol{v}\times\boldsymbol{w}\|^2}(\boldsymbol{v}\times\boldsymbol{w}) = \frac{|\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})|}{\|\boldsymbol{v}\times\boldsymbol{w}\|} \end{aligned}$$

From previous example

DIAGRAM 101

$$\begin{aligned} \text{area } \|gm\ ABCD &= 2\times(\text{area }\Delta\, ABD) \\ &= \|\boldsymbol{v}\times\boldsymbol{w}\| \end{aligned}$$

$$\Rightarrow \quad V = \|\boldsymbol{v}\times\boldsymbol{w}\|\cdot h = |\boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})|\,.$$

NOTE: This gives a useful application of determinants.

Example: Find volume V of parallelepiped determined by vectors

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad , \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad , \quad \boldsymbol{c} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Solution: $V = |\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})|$, where

$$\begin{aligned} \boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) &= \begin{vmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1\cdot\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \quad \text{(exp. along } r_3\text{)} \\ \Rightarrow \qquad V &= |-1| = 1. \end{aligned}$$

NOTE: Volume V of parallepiped determined by vectors $\boldsymbol{u}$. $\boldsymbol{v}$, $\boldsymbol{w}$ vanishes iff vectors are co-planar therefore vectors co-planar iff

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0 \quad - \text{ gives useful method for determining when 3 vectors are co-planar.}$$

Example: Find area of Δ with vertices $A = (0,1,3)$, $B = (1,2,1)$, $C = (r,1,0)$ and obtain a vector perp. to plane of ΔABC.

Solution: $\boldsymbol{w} = \overrightarrow{AB} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$, $\boldsymbol{v} = \overrightarrow{AC} = \begin{pmatrix} 4 \\ 0 \\ -3 \end{pmatrix}$ $\therefore$ Vector perp. to plane of ΔABC is

$$\boldsymbol{v}\times\boldsymbol{w} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 4 & 0 & -3 \\ 1 & 1 & -2 \end{vmatrix} = 3\boldsymbol{i} + 5\boldsymbol{j} + 4\boldsymbol{k}.$$

Area of

$$\begin{aligned} \Delta ABC = \frac{1}{2}\|\boldsymbol{v}\times\boldsymbol{w}| &= \frac{1}{2}\|3\boldsymbol{i} + 5\boldsymbol{j} + 4\boldsymbol{k}\| \\ &= \frac{1}{2}\sqrt{9+25+16} = \frac{1}{2}\sqrt{50}. \end{aligned}$$

LINES and PLANES in 3-space

LINES:

DIAGRAM 102

A line L in $\mathbb{R}^3$ is determined when we know a point $U = (x_0, y_0, z_0)$ on it and a vector $\boldsymbol{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ parallel to L. Then if $X = (x,y,z)$ is any point on L, must have $\overrightarrow{UX} = t\boldsymbol{v}$, some scalar t. Conversely any point X with $\overrightarrow{UX} = t\boldsymbol{v}$, some scalar t must lie on line L. Thus:

X is on L if and only if $\overrightarrow{UX} = t\boldsymbol{v}$, for some scalar t.
In terms of components this gives equation

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = t \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

or

$$\left.\begin{array}{l} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{array}\right\} \quad \begin{array}{l} \textit{parametric equ.} \text{ of st. line} \\ \text{(with parameter } t\text{).} \end{array}$$

Solving each equation for t gives following (they all equal t)

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \quad - \textit{symmetric} \text{ equs. of a line.}$$

NOTE: If any of a, b or c vanish corresp. term is to be replaced by $x = x_0$, $y = y_0$ or $z = z_0$ respectively.

Examples: (1) Find equation of st. line through point $(1, 2, 1)$ and parallel to vector $\boldsymbol{v} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$.

Solution: Here $U = (1, 2, 1)$ so param. equations of st. line are

$$\begin{aligned} x &= 1 + 3t \\ y &= 2 + t \\ z &= 1 + 2t \end{aligned}$$

and symm. equation is

$$\frac{x - 1}{3} = y - 2 = \frac{z - 1}{2}$$

(2) Param. equation of st. line through pt. $(1, 1, 1)$ and parallel to vector $\boldsymbol{v} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$ are

$$\begin{aligned} x &= 1 + 2t \\ y &= 1 \\ z &= 1 - t \end{aligned}$$

and symm. equation is

$$\frac{x-1}{2} = \frac{z-1}{-1} \quad , \quad y = 1 \quad \text{or} \quad \frac{x-1}{2} = 1 - z \quad , \quad y = 1.$$

Planes:

DIAGRAM 103

A plane π is determined by a point $P = (x_0, y_0, z_0)$ on it and a vector $\boldsymbol{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ *normal* (ie. perpendicular) to it. Then point $X = (x, y, z)$ lies on plane π if and only if

$$\boldsymbol{n} \cdot \overrightarrow{PX} = 0$$

or

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0.$$

This gives the equation of plane π:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

or

$$ax + by + cz = d$$

where $d = ax_0 + by_0 + cz_0$.

Example: Equation of plane through point $(0, 1, 0)$ normal to $\boldsymbol{n} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ is

$$\begin{aligned} x + 2(y-1) + z &= 0 \\ \text{or} \qquad x + 2y + z &= 2. \end{aligned}$$

DIAGRAM 104

A plane π is also determined if we know three points (not all on same line) A, B, C on it. Then

$$\boldsymbol{n} = \overrightarrow{AB} \times \overrightarrow{AC}$$

is a normal vector to plane, so equation of plane is

$$\boldsymbol{n} \cdot \overrightarrow{AX} = 0$$

where $X = (x, y, z)$ is any pt. on plane.

Example: Find equation of plane through the points

$$A = (0, 1, 3) \quad , \quad B = (1, 2, 1) \quad , \quad C = (4, 1, 0).$$

Solution: From previous example, a vector normal to plane is

$$\boldsymbol{n} = \overrightarrow{AB} \times \overrightarrow{AC} = 3\boldsymbol{i} + 5\boldsymbol{j} + 4\boldsymbol{k}.$$

Therefore equation of plane is

$$0 = \boldsymbol{n} \cdot \overrightarrow{AX} = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} x - 0 \\ y - 1 \\ z - 3 \end{pmatrix}$$

so

$$3x + 5(y - 1) + 4(z - 3) = 0$$

or

$$3x + 5y + 4z = 17.$$

Eigenvalues and eigenvectors

Let A be an $n \times n$ matrix and consider the equation

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \tag{1}$$

with λ a number (possibly complex). It is clear that $\boldsymbol{x} = \boldsymbol{0}$ is always a solution to (1). A value λ, for which (1) has a non-zero solution is called an *eigenvalue* of A. The corresponding solution vector $\boldsymbol{x} \neq \boldsymbol{0}$ is called an *eigenvector* of A corresponding to eigenvalue λ.

FACT: Let λ be an eigenvalue of A. Then the set of solutions to (1) forms a vector space – called the *eigenspace* of A corresponding to eigenvalue λ.

Proof: Let $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ be solutions to (1). Then

$$\begin{aligned} A(\boldsymbol{x}_1 + \boldsymbol{x}_2) &= A\boldsymbol{x}_1 + A\boldsymbol{x}_2 = \lambda\boldsymbol{x}_1 + \lambda\boldsymbol{x}_2 = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2) \\ A(\alpha\boldsymbol{x}_1) &= \alpha A\boldsymbol{x}_1 = \alpha\lambda\boldsymbol{x}_1 = \lambda(\alpha\boldsymbol{x}_1) \quad , \quad \alpha \in \mathbb{R} \end{aligned}$$

so $\boldsymbol{x}_1 + \boldsymbol{x}_2$, $\alpha\boldsymbol{x}_1$ are also solutions therefore set of solutions is closed under vector addition and scalar mult. and hence forms a vector space.

NOTE: The dimension of the eigenspace corresponding to λ is called the *multiplicity* of λ.

We now show that A has at least 1 and at most n distinct (real or complex) eigenvalues. To see this first note that equation (1) may be written

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0}$$

which has a non-zero solution iff

$$|A - \lambda I| = 0. \tag{2}$$

The determinant

$$D(\lambda) = |A - \lambda I|$$

determines a polynomial of degree n in λ called the *characteristic polynomial* of A and we call equation (2) the *characteristic equation.*

NOTE: It can be shown that A satisfies its own characteristic equation, so

$$D(A) = 0 \qquad \text{– Cayley-Hamilton Theorem.}$$

The eigenvalues of A are thus given by the roots of the characteristic equation (2) so that A has at least 1 and at most n distinct eigenvalues. The corresponding eigenvectors may then be obtained by solving

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0}$$

using Gauss elimination.

NOTE: If λ has multiplicity m, it suffices to find m linearly independent eigenvectors.

Examples: (1) Find the eigenvalues and e-vectors of

$$A = \begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}$$

Solution:

$$|A - \lambda I| = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = (\lambda+5)(\lambda+2) - 4$$
$$= \lambda^2 + 7\lambda + 6 = (\lambda+6)(\lambda+1) \Rightarrow \lambda = -1\,, -6$$

are e-values.

NOTE: A satisfies its own char. equation

$$(A + 6I)(A + I) = A^2 + 7A + 6I = 0 \qquad \text{(check).}$$

□

To find e-vectors need to solve

$$(A-\lambda I)\boldsymbol{x}=\mathbf{0} \quad , \qquad \text{for } \lambda=-1\,,-6.$$

Aug. Matrix $(A-\lambda I\vdots\mathbf{0})$ corresp. to $\lambda=-1,\ -6$ is

$$\lambda=-1: \quad \begin{pmatrix} -4 & 2 & \vdots & 0 \\ 2 & -1 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & \vdots & 0 \\ 2 & -1 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{pmatrix} \quad \therefore \text{ Gen. solution is}$$

$$\boldsymbol{x}=\begin{pmatrix} \alpha \\ 2\alpha \end{pmatrix}=\alpha\begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ an } e\text{-vector corresp. to } e\text{-value } \lambda=-1$$

$$\lambda=-6: \quad \begin{pmatrix} 1 & 2 & \vdots & 0 \\ 2 & 4 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{pmatrix} \Rightarrow \text{Gen. solution } \boldsymbol{x}=\begin{pmatrix} -2\alpha \\ \alpha \end{pmatrix}=\alpha\begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$\therefore \binom{-2}{1}$ an e-vector corresponding to e-value $\lambda=-6$.

(2) Find e-values and e-vectors of $A=\begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$.

Solution:

$$\begin{aligned} |A-\lambda I| &= \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = -(\lambda+2)\begin{vmatrix} 1-\lambda & -6 \\ -2 & -\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & -6 \\ -1 & -\lambda \end{vmatrix} - 3\begin{vmatrix} 2 & 1-\lambda \\ -1 & -2 \end{vmatrix} \\ &= -(\lambda+2)(\lambda^2-\lambda-12)+7(\lambda+3) = -(\lambda+2)(\lambda+3)(\lambda-4)+7(\lambda+3) \\ &= (\lambda+3)[7-(\lambda+2)(\lambda-4) = -(\lambda+3)(\lambda^2-2\lambda-15) = -(\lambda+3)^2(\lambda-5) \end{aligned}$$

$\therefore A$ has eigenvalues $\lambda=-3,\ 5$.

NOTE: In this case A satisfies the Char. equation

$$(A+3I)^2(A-5I)=0.$$

To find e-vectors corresponding to $\lambda=-3,\ 5$ need to solve $(A-\lambda I)\boldsymbol{x}=\mathbf{0}$. The Aux. matrix $(A-\lambda I\vdots\mathbf{0})$ for $\lambda=-3$ is

$$\begin{pmatrix} 1 & 2 & -3 & \vdots & 0 \\ 2 & 4 & -6 & \vdots & 0 \\ -1 & -2 & 3 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -3 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix} \Rightarrow z=\alpha,\ y=\beta$$

arb. and $x=3z-2y=3\alpha-2\beta$.

$\therefore$ General solution $\boldsymbol{x} = \begin{pmatrix} 3\alpha - 2\beta \\ \beta \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \therefore \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ are linearly independent e-vectors with e-value $\lambda = -3$ therefore this e-value has multiplicity 2.

For $\lambda = 5$ the Aux. matrix is

$$\begin{pmatrix} -7 & 2 & -3 & \vdots & 0 \\ 2 & -4 & -6 & \vdots & 0 \\ -1 & -2 & -5 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 5 & \vdots & 0 \\ 2 & -4 & -6 & \vdots & 0 \\ -7 & 2 & -3 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 5 & \vdots & 0 \\ 0 & -8 & -16 & \vdots & 0 \\ 0 & 16 & 32 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 5 & \vdots & 0 \\ 0 & 1 & 2 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

$\therefore z = \alpha$ arb. and back subst. $\Rightarrow y = -2\alpha$, $x = -5z - 2y = -\alpha$

$\therefore$ General solution is $\boldsymbol{x} = \begin{pmatrix} -\alpha \\ -2\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$ is an e-vector with e-value $x = 5$.

(3) Find the eigenvalues and e-vectors of

$$A = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

Solution: $|A - \lambda I| = \begin{vmatrix} -\lambda & 2 \\ -2 & -\lambda \end{vmatrix} = \lambda^2 + 4 = (\lambda - 2i)(\lambda + 2i)$.

$\therefore$ In this case A has two complex e-values $\lambda = \pm 2i$. To obtain e-vectors solve $(A - \lambda I)\boldsymbol{x}$, $\lambda = \pm 2i$, exactly as before using complex numbers. Aux. matrix $(A - \lambda I \vdots \mathbf{0})$ corresponding to $\lambda = 2i$ is

$$\begin{pmatrix} -2i & 2 & \vdots & 0 \\ -2 & -2i & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} -2i & 2 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{pmatrix}$$

$\Rightarrow y = ix \quad \therefore$ Gen. solution $\boldsymbol{x} = \begin{pmatrix} \alpha \\ i\alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ i \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ i \end{pmatrix}$ an e-vector corresponding to $\lambda = 2i$.

Similarly can show e-vector corresponding to $\lambda = -2i$ is $\begin{pmatrix} 1 \\ -i \end{pmatrix}$.

Applications

E-values and e-vectors have many applications eg. in Quantum Mechanics energy of a system is represented by a matrix whose eigenvalues give the allowed energies – this is why discrete energy levels are observed.

Many other applications in Math., Physics, Chemistry and Biology eg.

Coupled ODEs

DIAGRAM 105

$$m\ddot{y}_1 = \underset{\substack{\uparrow \\ \text{upward force} \\ \text{exerted by} \\ \text{1st. spring}}}{k_1 y_1} + \underset{\substack{\uparrow \\ \text{downward force} \\ \text{exerted by} \\ \text{2nd spring}}}{k_2(y_2 - y_1)}$$

where

$$m\ddot{y}_2 = -k_2(y_2 - y_1)$$

or

$$\begin{aligned} \ddot{y}_1 &= -(k_1+k_2)/m\, y_1 + \frac{k_2}{m} y_2 \\ \ddot{y}_2 &= \frac{k_2}{m}(y_1 - y_2) \end{aligned}$$

$$\Rightarrow \quad \ddot{\boldsymbol{y}} = A\boldsymbol{y}$$

$$A = \frac{1}{m}\begin{pmatrix} -(k_1+k_2) & k_2 \\ k_2 & -k_2 \end{pmatrix}$$ – solved using e-values & e-vectors of A (see MT152).

Rotations in $\mathbb{R}^n$ and Orthogonal Matrices

An $n \times n$ matrix A is called *orthogonal* if

$$A^T A = AA^T = I$$

– ie. an orthogonal matrix is one for which $A^{-1} = A^T$.

Since $|A| = |A^T|$,

$$\begin{aligned} A \text{ orthog.} \quad \Rightarrow 1 = |I| &= |A^T A| \\ &|A^T| \cdot |A| = |A|^2 \\ \Rightarrow |A| = \pm 1. \end{aligned}$$

An orthogonal matrix A s.t. $|A| = 1$ is called a *proper orthogonal* matrix – of interest since they determine rotations in $\mathbb{R}^n$.

Rotations in $\mathbb{R}^2$

Consider $A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Then it is easily seen that

$$A_\theta^T \cdot A_\theta = A_\theta A_\theta^T = I$$

so A_θ is orthogonal. Moreover

$$|A_\theta| = \cos^2\theta + \sin^2\theta = 1 \Rightarrow A_\theta \text{ is proper orthogonal.}$$

Now consider action of A_θ on coordinate vectors $\boldsymbol{e}_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix}0\\1\end{pmatrix}$

$$\begin{aligned} A_\theta \boldsymbol{e}_1 &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} &= \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} \\ A_\theta \boldsymbol{e}_2 &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} &= \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix} \end{aligned}$$

DIAGRAM 106

$\therefore$ A_θ rotates $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ about origin through angle θ. Since $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ span $\mathbb{R}^2$, A_θ determines a rotation through an angle θ anti-clock wise about origin.

Example: Determine vector obtained by rotating $\begin{pmatrix}1\\1\end{pmatrix}$ through an angle of 60°.

Solution:

$$A_{\pi/3} = \begin{pmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$$

DIAGRAM 107

$$\Rightarrow A_{\frac{\pi}{3}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1-\sqrt{3}\\ 1+\sqrt{3}\end{pmatrix}.$$

Eigenvalues: Determined by

$$|A_\theta - \lambda I| = \begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix} = (\lambda - \cos\theta)^2 + \sin^2\theta$$

$\therefore$ e-values λ given by

$$\begin{aligned} &(\lambda - \cos\theta)^2 = -\sin^2\theta \Rightarrow \lambda - \cos\theta = \pm i \sin\theta\,, \quad \text{so} \\ &\lambda = \cos\theta \pm i\sin\theta = e^{\pm i\theta} \end{aligned}$$

are the eigenvalues of A_θ.

In general e-values of an orthogonal matrix are of the form $e^{\pm i\theta}$; ie. complex of absolute value 1 – includes ± 1 possible as e-values.

Rotations in $\mathbb{R}^3$

Consider

$$A = \begin{pmatrix} A_\theta & & \vdots & 0 \\ & & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with A_θ the matrix considered above. Using previous example it is easily seen that A is proper orthogonal (check!). Also from previous example, A rotates x, y components through an angle θ but leaves z-component of a vector unchanged. Therefore A represents a rotation about z-axis through an angle θ.

To obtain e-values of A we have

$$|A-\lambda| = \left|\begin{array}{cc:c} A_\theta - \lambda I & & 0 \\ & & 0 \\ \hdashline 0 & 0 & 1-\lambda \end{array}\right| = (1-\lambda)\,|A_\theta - \lambda I| \quad \text{(exp.along row 3)}$$

$\Rightarrow \quad \lambda = 1,\ e^{\pm i\theta}$ are the e-values.

In this case eigenvector corresponding to $\lambda = 1$ is $\boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ – gives the axis of rotation.

DIAGRAM 108

In general a proper orthogonal 3×3 matrix A has e-values $\lambda = 1,\ e^{\pm i\theta}$. If $\boldsymbol{v}$ is e-vector corresponding to $\lambda = 1$, so $A\boldsymbol{v} = \boldsymbol{v}$, then A determines a rotation about axis $\boldsymbol{v}$ through an angle θ. Thus eigenvalues of A and e-vector corresponding to $\lambda = 1$ give the angle and axis of rotation.

Example: Show that

$$A = \frac{1}{5}\begin{pmatrix} -3 & 0 & 4 \\ 0 & -5 & 0 \\ 4 & 0 & 3 \end{pmatrix}$$

determines a rotation in $\mathbb{R}^3$ and find the angle and axis of rotation.

Solution: It is easily checked (Exercize) that

$$A^T A = A A^T = I \quad , \quad |A| = 1$$

so A is a proper orthogonal matrix and thus determines a rotation in $\mathbb{R}^3$. For e-values of

A we have

$$|A-\lambda I| = \begin{vmatrix} -\frac{3}{5}-\lambda & 0 & \frac{4}{5} \\ 0 & -1-\lambda & 0 \\ \frac{4}{5} & 0 & \frac{3}{5}-\lambda \end{vmatrix}$$

$$= -(\lambda+1)\begin{vmatrix} -\frac{3}{5}\lambda & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5}-\lambda \end{vmatrix} \qquad \text{(exp. along } r_2\text{)}$$

$$= -(\lambda+1)(\lambda^2 - \frac{9}{25} - \tfrac{16}{25}) = -(\lambda+1)(\lambda^2-1)$$

Therefore e-values are $\lambda = 1,\ -1 = e^{\pm i\pi}$ – so determines a rotation through 180°.

For axis of rotation need e-vector corresponds to $\lambda = 1$ – given by solution to $(A - I)\boldsymbol{x} = \mathbf{0}$. Corresponding Aug. Matrix $(A - I \vdots \mathbf{0})$ is

$$\begin{pmatrix} -\frac{8}{5} & 0 & \frac{4}{5} & \vdots & 0 \\ 0 & -1 & 0 & \vdots & 0 \\ \frac{4}{5} & 0 & -\frac{2}{5} & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 0 & 1 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 2 & 0 & -1 & \vdots & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 & 1 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

$\Rightarrow \quad y = 0,\ z = 2x$ therefore gen. solution $\boldsymbol{x} = \alpha \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ an e-vector corresponding to $\lambda = 1$.

Therefore A represents a rotation about axis $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$

DIAGRAM 109 through an angle of 180°