SIT787: Mathematics for Artificial Intelligence
Preliminaries and Calculus
Topic 1

Asef Nazari

School of Information Technology, Deakin University

AI, DS, BA, ...

Data → table
matrix

features

price

x_1	x_2	x_3	...	y
3	600	10	..	

case

"house"

\# of rooms

area

age

in CBD

machine learning

- supervised learning ←
- unsupervised learning

$y = f(x)$

$$y = f(x_1, x_2, x_3, \dots, x_n)$$

price

functions

multi-variate

$$\boxed{y = f(x)}$$

output

input variable

variables x → set

x could take any value - independent variable

$y = f(x)$

variable

$y = 3x + 2$

dependent variable

indep. var

Sets

Sets

- A **set** is a collection of elements
  - $A = \{1, 5, -2, 7\} = \{1, 1, 5, -2, -2, 7, 7, 7\}$
  - $B = \{\text{monkey}, \text{money}, \text{key}\}$
- **Membership**
  - $1 \in A$ and $100 \notin A$
  - money $\in B$ and pool $\notin B$
- **Subsets**
  - $A \subset C$ if every element of A is in C
    Consider $A = \{1, 5, -2, 7\}$ and $C = \{1, 5, 8, -2, 7, 12\}$. We say $A \subset C$ but $C \not\subset A$.
- **Cardinality** of a set is the number of elements in that set
  - $|A| = 4$, $|B| = 3$, $|C| = 6$
- **Empty set**, $\emptyset$, has no member. $|\emptyset| = 0$.

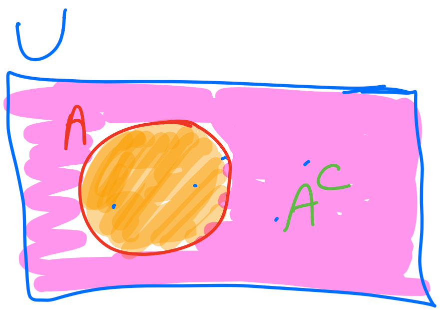
$\emptyset = \{ \quad \}$

$\emptyset \subset A$

body temperature (0, 100)

- **Universal set U:** the set that contains all elements under consideration.
- **Set compliment:** The complement of A is the set of all elements in the universal set U, but not in A, and is denoted by A^c

$$A^c = \{x | x \in U \text{ and } x \notin A\} = \bar{A}$$



$A \subset U$

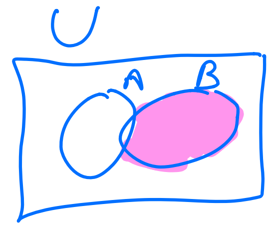
$A^c \subset U$

$A \quad A^c$

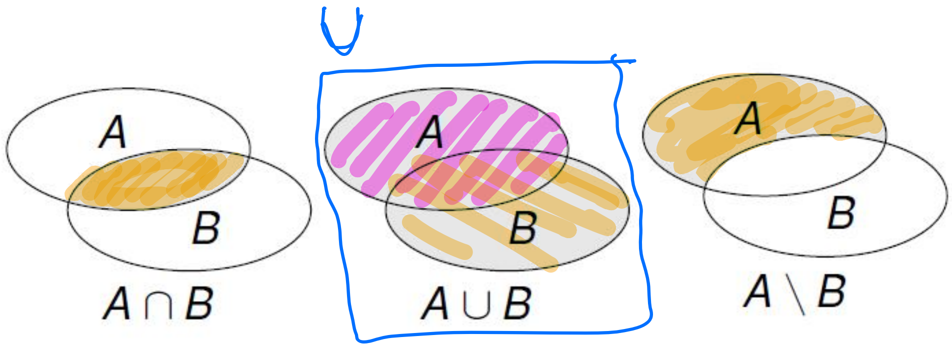
Operations on Sets

Search

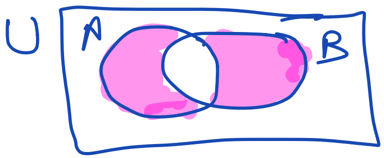
- **Union** of A and B is $A \cup B$
  - $A \cup B = \{x | x \in A \text{ OR } x \in B\}$
- **Intersection** of A and B is $A \cap B$
  - $A \cap B = \{x | x \in A \text{ AND } x \in B\}$
- **Set difference** of A and B is $A \backslash B$ or $A - B$
  - $A \backslash B = \{x | x \in A \text{ AND } x \notin B\} = A \cap B^c = A - B$



$B - A \neq A - B$

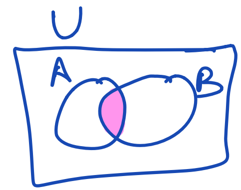


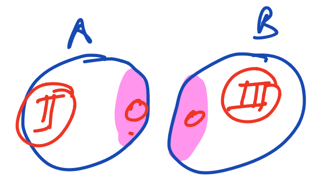
- You search and learn **symmetric difference** between two sets



$A \Delta B = B \Delta A$

counting A $|A|$ cardinality of A





- **Inclusion-exclusion Theorem**

$$\boxed{\underset{\mathrm{I}}{|A\cup B|} = \underset{\mathrm{II}}{|A|} + \underset{\mathrm{III}}{|B|} - \underset{\mathrm{IV}}{|A\cap B|}}$$

- Example: consider $A = \{1, 5, -7, 3\}$ and $B = \{1, -5, -7, 4\}$
  - Find $A \cup B$ and $A \cap B$
  - Find $|A|, |B|, |A \cup B|$ and $|A \cap B|$
  - Verify the Inclusion-exclusion Theorem.

$|A| = 4$ $|B| = 4$ $A \cap B = \{1, -7\}$ $|A \cap B| = 2$

$A \cup B = \{1, 5, -7, 3, 1, -5, -7, 4\}$

$\to A \cup B = \{1, 5, -7, 3, -5, 4\}$

$|A \cup B| = 6$

$$|A \cup B| = |A| + |B| - |A \cap B| = 4 + 4 - 2 = 6$$

Scalers = numerical values

$\mathbb{N} \quad \mathbb{N}$

$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$

- Scalers are just numbers, like 1, $-\frac{1}{3}$, $\frac{\pi}{3}$
- Frequent sets of scalers
  - Natural Numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$ or the counting numbers
  - Integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ or

I

$|\mathbb{N}|$ $\boxed{0 \notin \mathbb{N}}$

$$\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

$|\mathbb{Z}|$

  - Rational Numbers $\mathbb{Q}$ numbers of the form $\frac{m}{n}$ where $n \neq 0$ and $m, n \in \mathbb{Z}$

$\frac{2}{10} = \boxed{0.2}$

$n=1$ $\frac{5}{1} = 5$ $\frac{-3}{1} = -3$ $\frac{\bigcirc}{\bigcirc \neq 0}$

    - number can be written as a Ratio of two integers
    - fractional, decimal $\frac{3}{2} = 1.5$ or $7 = \frac{7}{1}$
    - what about $\frac{1}{3} = \boxed{0.333333\ldots}$?
  - The decimal representation of a rational number always either terminates after a finite number of digits or begins to repeat the same digits over and over.
  - any repeating or terminating decimal represents a rational number.

$\pi = 3.1415 \ldots$ irrational

Scalers

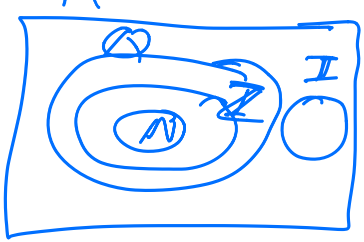
$\mathbb{Q} \cap \mathbb{I} = \phi$

- Irrational Numbers $\mathbb{I}$:
  - numbers cannot be written as a ratio of two integers
  - $\pi = 3.14159265358979323846264338327950...$
  - $e = 2.71828182845904523536028747135266...$
  - Surds: $\sqrt{2}$ $\quad \sqrt{3}$
  - numbers non-terminating decimal with no pattern
- Real Numbers $\mathbb{R}$: all numbers mentioned above all together

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

$$\mathbb{I} \subseteq \mathbb{R}$$

$\mathbb{R} = U$

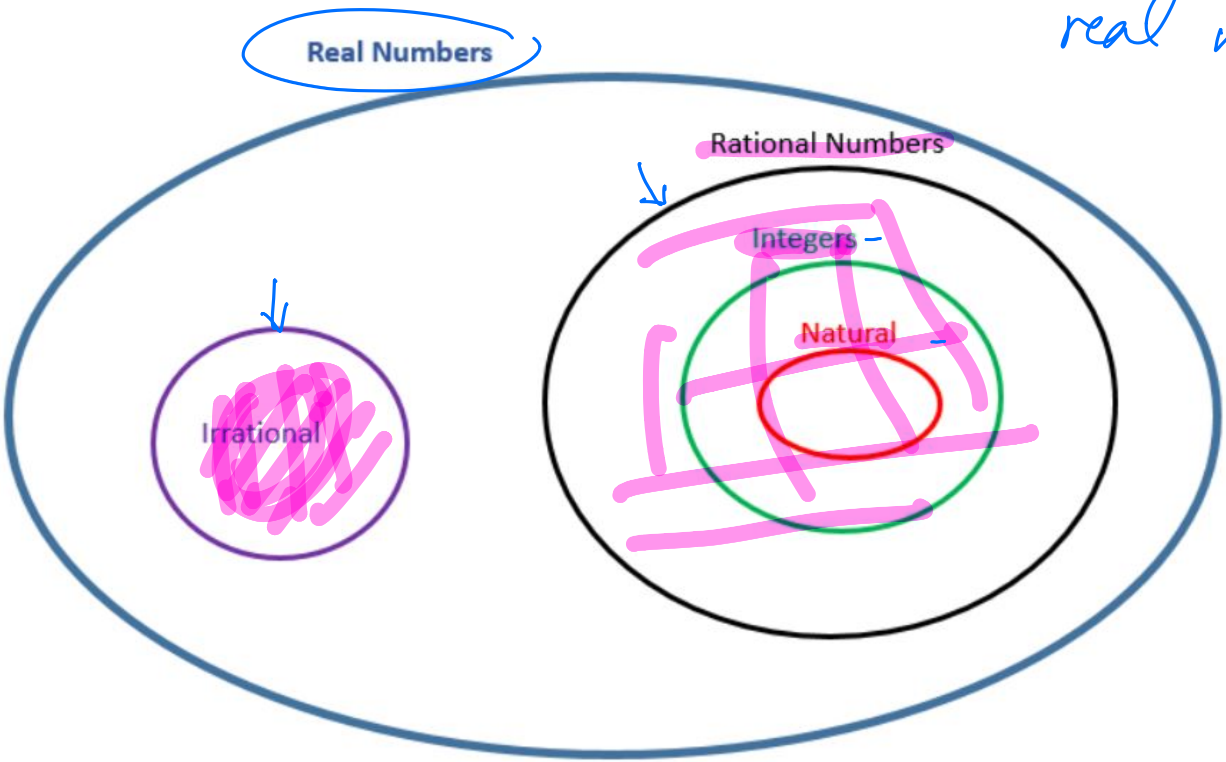


- $\mathbb{Q}^c = \mathbb{I}$ and $\mathbb{I}^c = \mathbb{Q}$
- $\mathbb{Q} \cup \mathbb{I} = \mathbb{R}$

Scalers

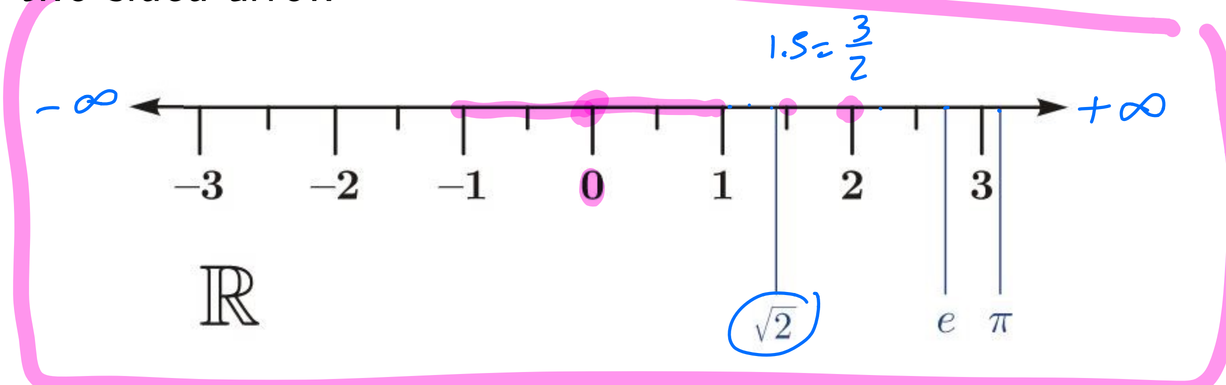
reference set

Scalars
real numbers



$x \in \mathbb{R}$

- Real numbers can be shown on a line (two-sided arrow)
  - reference point as origin
  - Positive and negative direction
  - coordinate line or real line
  - The real numbers are ordered
- Concept of ∞ and $-\infty$
- If you have a real variable x, it could be everywheree on this two-sided arrow



Inequalities

$x \in \mathbb{R}$

$x, y \in \mathbb{R}$

- $x = y$, x is equal to y
- $x < y$, x is less than y
- $y < x$, x is greater than y
- $x \leq y$, x is less than or equal to y
- $x \geq y$, y is greater than or equal to x
- $x \ll y$, x is much, much less than y
- important considerations
  - If $x = y$, then for any real number a, $ax = ay$
  - $\boxed{\text{if } x \leq y}$
    - if a is a positive number, $ax \leq ay$
    - if a is a negative number, $ax \geq ay$

$5 \leq 7$

$3 \leq 3$

less than or equal

$4 \cancel{\leq} 2$

$(-3)5 = (-3)5$

$-15 = -15$

$(2)\, 3 \leq (2) 4 \rightarrow 6 \leq 8$

$\underbrace{(-2)3}_{-6} \geq \underbrace{(-2)4}_{-8}$

Intervals

$\mathbb{R}$

Subsets of

- **Closed** intervals $[a, b] = \{x | a \leq x \leq b\}$
- **Open** intervals $(a, b) = \{x | a < x < b\}$
- **Half-open** intervals ,
  - $(a, b] = \{x | a < x \leq b\}$
  - $[a, b) = \{x | a \leq x < b\}$
- **Rays**
  - $[a, \infty) = \{x | x \geq a\}$
  - $(-\infty, a) = \{x | x < a\}$
  - similarly, (a, ∞), , $(-\infty, a]$

$\mathbb{R} = (-\infty, +\infty)$

Summation and Product notations

$\text{Data} = \{x_1, x_2, \ldots, x_n\}$

$x_1 + x_2 + \cdots + x_n = \sum x_i$

$x_1 x_2 \cdots x_n = \prod x_i$

Sigma Notation for Sums

- $\sum_{i=2}^{6} i = 2 + 3 + 4 + 5 + 6$ i∈{2, 3, 4, 5, 6}
- i $\sum_{i=2}^{6} i^2 = (2)^2 + (3)^2 + (4)^2 + (5)^2 + (6)^2 = 4 + 9 + 16 + 25 + 36$

  i∈{2, 3, 4, 5, 6}
- k $\sum_{k=2}^{6} (2k+1) =$ k∈{2, 3, 4, 5, 6}

  $(2(2)+1) + (2(3)+1) + (2(4)+1) + (2(5)+1) + (2(6)+1)$ = - - -
- → $\sum_{j=0}^{3} \left(\frac{j}{k}\right) = \left(\frac{0}{k}\right) + \left(\frac{1}{k}\right) + \left(\frac{2}{k}\right) + \left(\frac{3}{k}\right)$ j∈{0, 1, 2, 3} ✓

  j=0 j=1 j=2 j=3
- $\sum_{j=0}^{3} 5 = 5 + 5 + 5 + 5 = (3 - 0 + 1)5$ = 20

  j
- $\sum_{j=m}^{n} a = (n - m + 1)a$
- Check that $\sum_{j=0}^{3} (j^2 + 2j) = \left(\sum_{j=0}^{3} j^2\right) + \left(2\sum_{j=0}^{3} j\right)$

Sigma Notation for Sums

- $\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$
- Suppose set $X = \{x_1, x_2, \ldots, x_n\}$ is given.
  - The average of the values of X

$|X| = n$

$$\bar{x} = \frac{\text{add all values}}{\text{divide by } |X|} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Notation for products

- $\prod_{i=1}^{n} x_i = x_1 x_2 \dots x_n$
- $\prod_{i=2}^{5} x_i = x_2 x_3 x_4 x_5$
- For a nonegative integer n, the factorial is defined as

$$n! = n(n-1)(n-2)\dots(3)(2)(1)$$

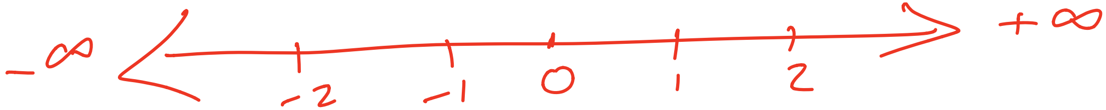
$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120, \text{ also } 1! = 0! = 1$$

- We can represent factorial as

$$n! = \prod_{i=1}^{n} i$$

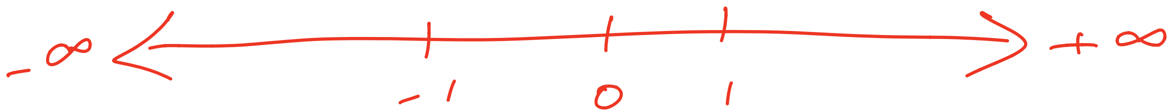
$n! = (1)(2)(3) \dots (n-1)(n)$

$x \in \mathbb{R}$



Coordinate Systems and Distance

$y \in \mathbb{R}$

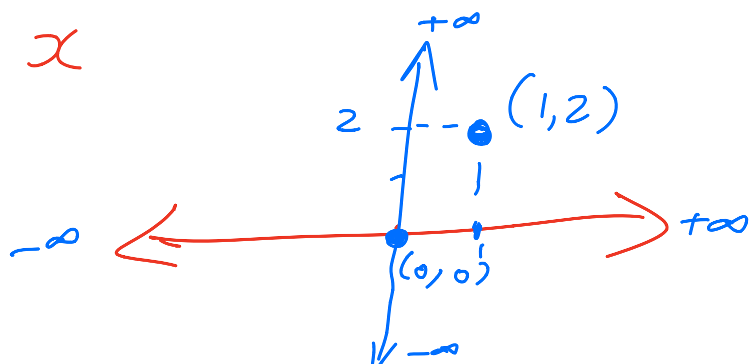


y is dependt on x

(x, y)

(0,0)

(1,2)



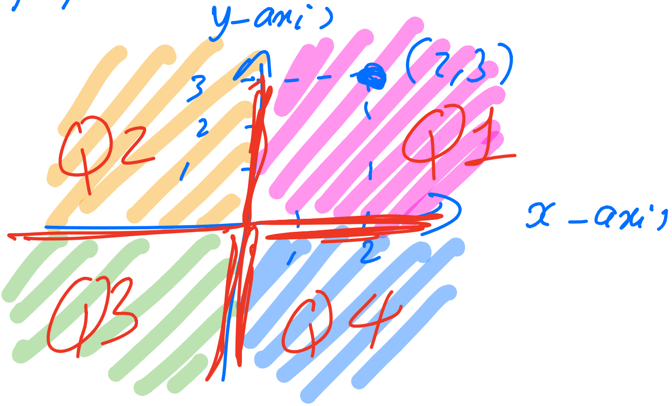
Coordinates of $\mathbb{R}^2$ (Cartesian Plane)

- Points on a line can be identified with real numbers
- Points on a plane can be identified by with ordered pairs of numbers

(x, y) ordered pair

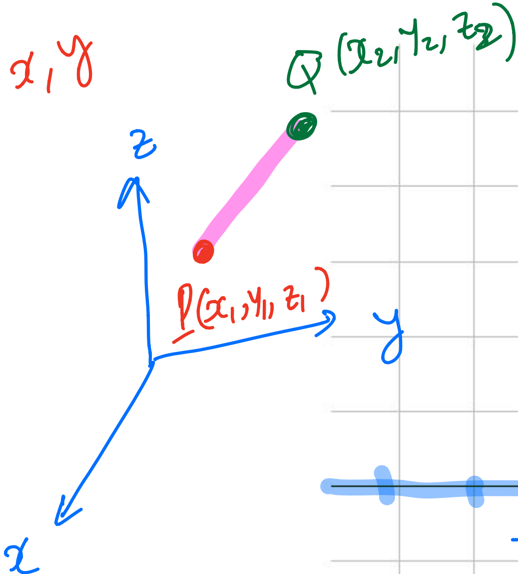
- axis and quadrants

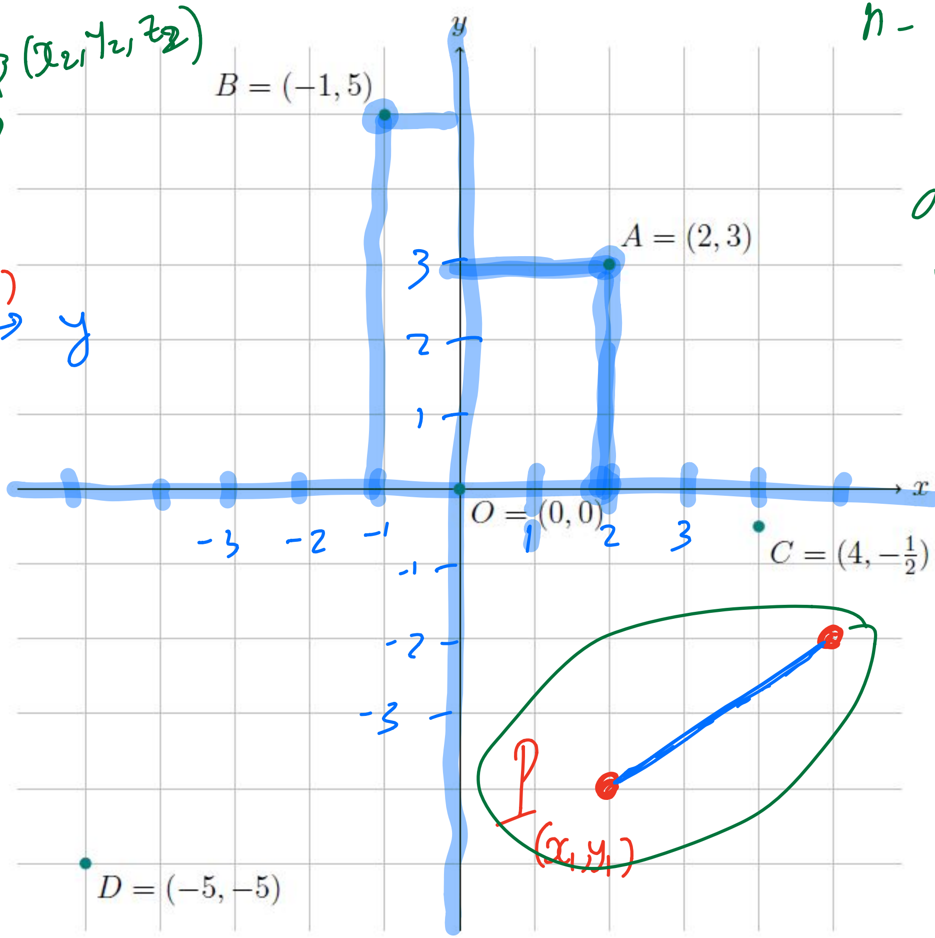
(2, 3)



Cartesian plane

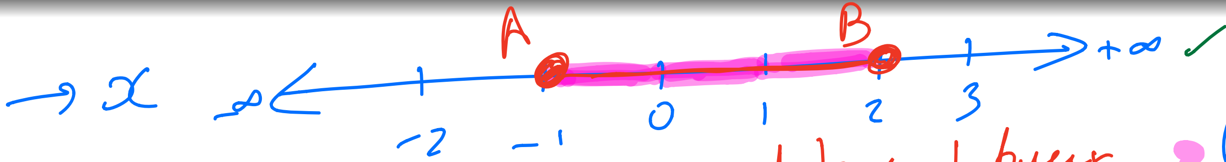
x, y



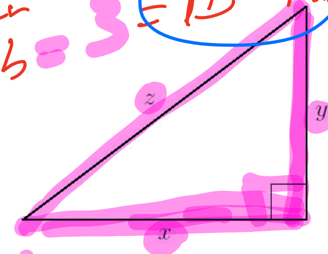


n-dimensional space

distance betwe two points



distance between A, and b = 3 = |B − A|

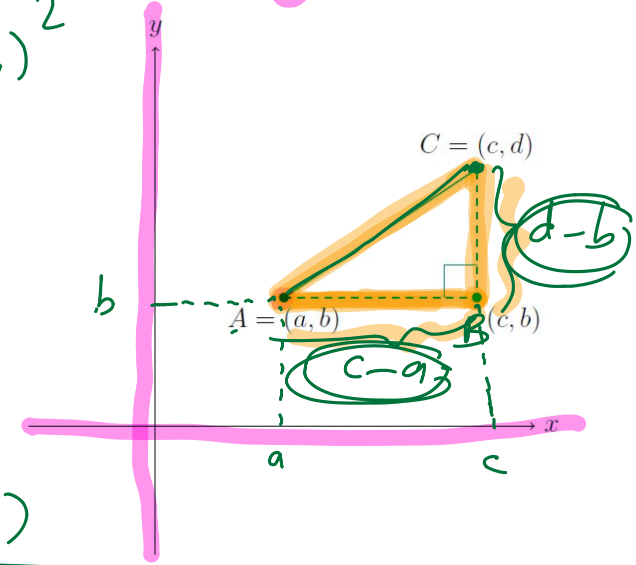


- Pythagoras' Theorem
- $z^2 = x^2 + y^2$

$$(\text{length A to C})^2 = (c-a)^2 + (d-b)^2$$

- Distance between two pints

$$\text{dist}(A, C) = \sqrt{(c-a)^2 + (b-d)^2}$$



$A(x_1, y_1, z_1)$ $C(x_2, y_2, z_2)$

$$\text{dist}(A, C) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

in n-dimensional space

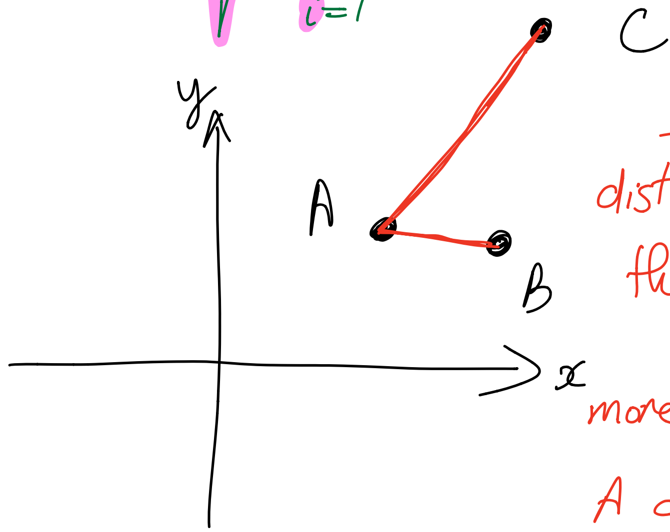
$(x_1, x_2, \ldots, x_n)$

$A(a_1, a_2, \ldots, a_n)$

$B(b_1, b_2, \ldots, b_n)$

$$\text{dist}(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} \boxed{(b_i - a_i)^2}}$$



similarity measure

dist(A,B) < dis(A,C)

then

A and B are more similar than A and C.

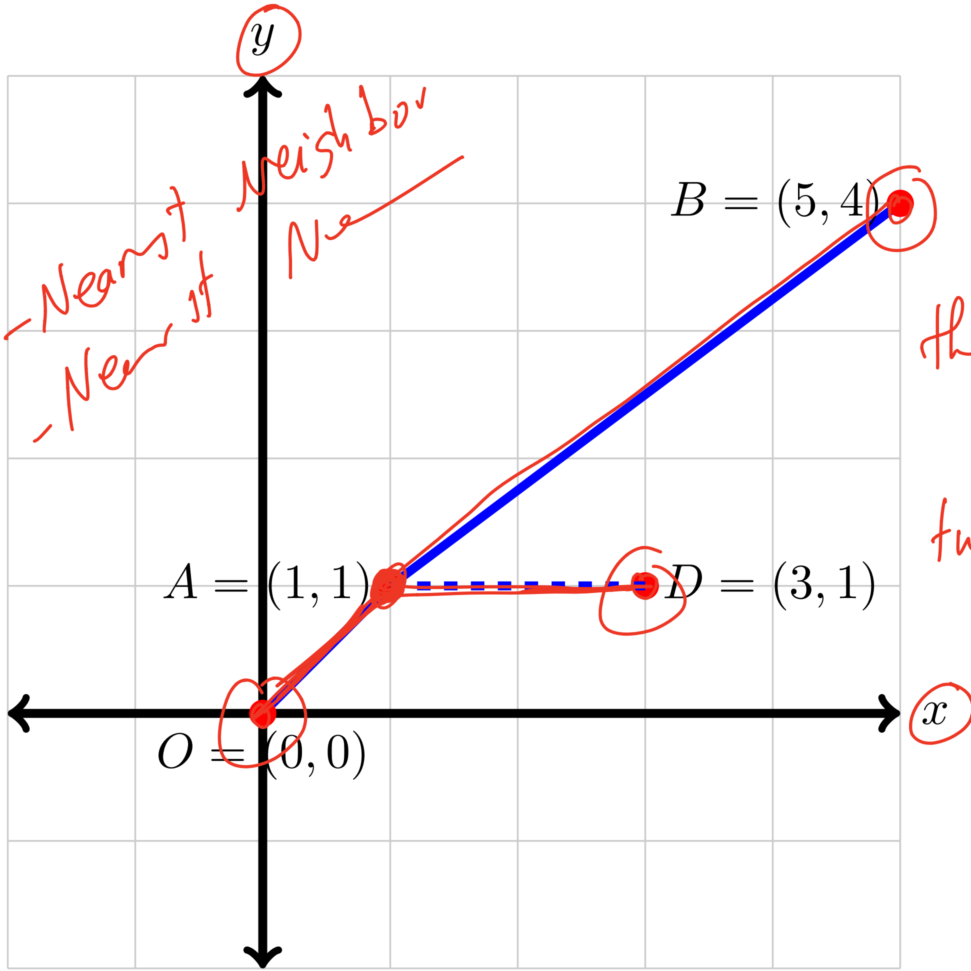
Distance applications: k Nearest Neighbours

- dist(A, B), dist(A, O), dist(A, D)

2-Dimensional space

1-Nearest Neighbor

2-Nearest Ne

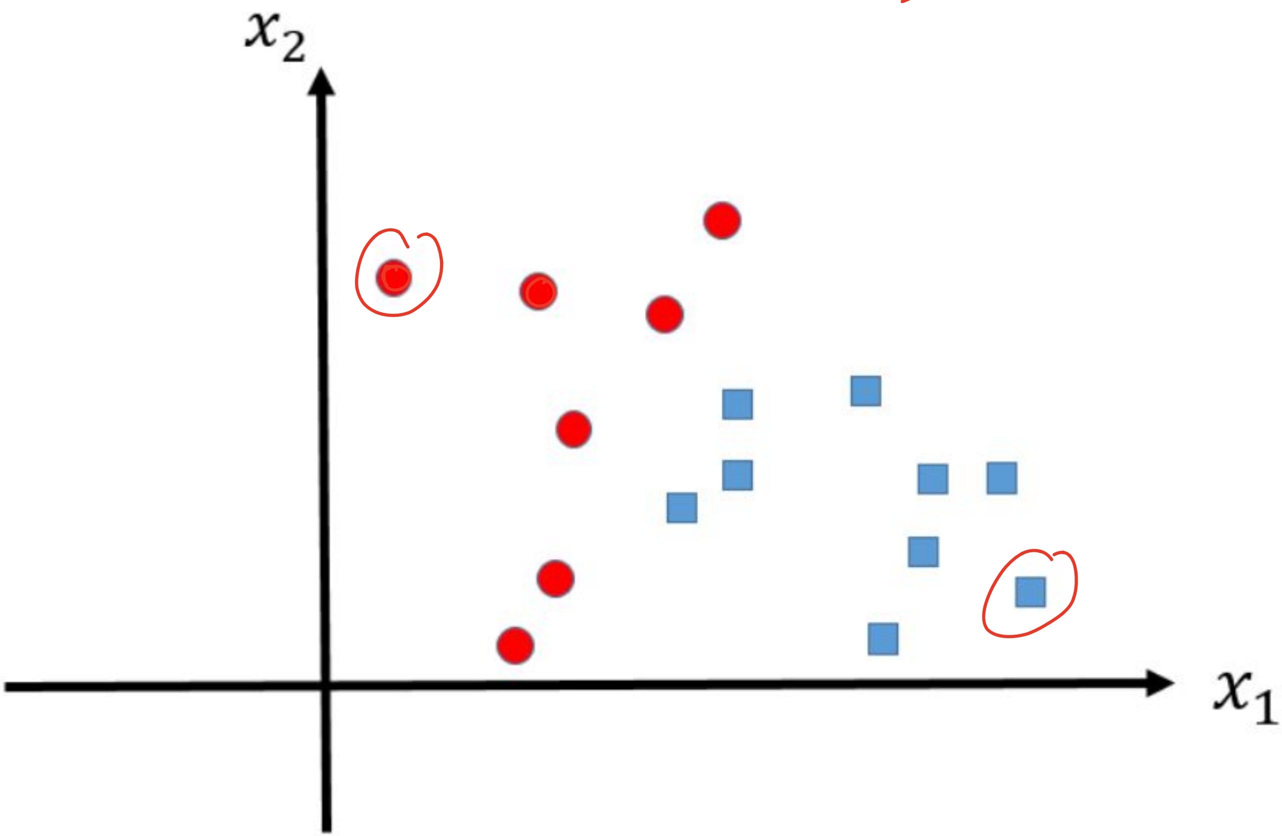


the most similar = {O}

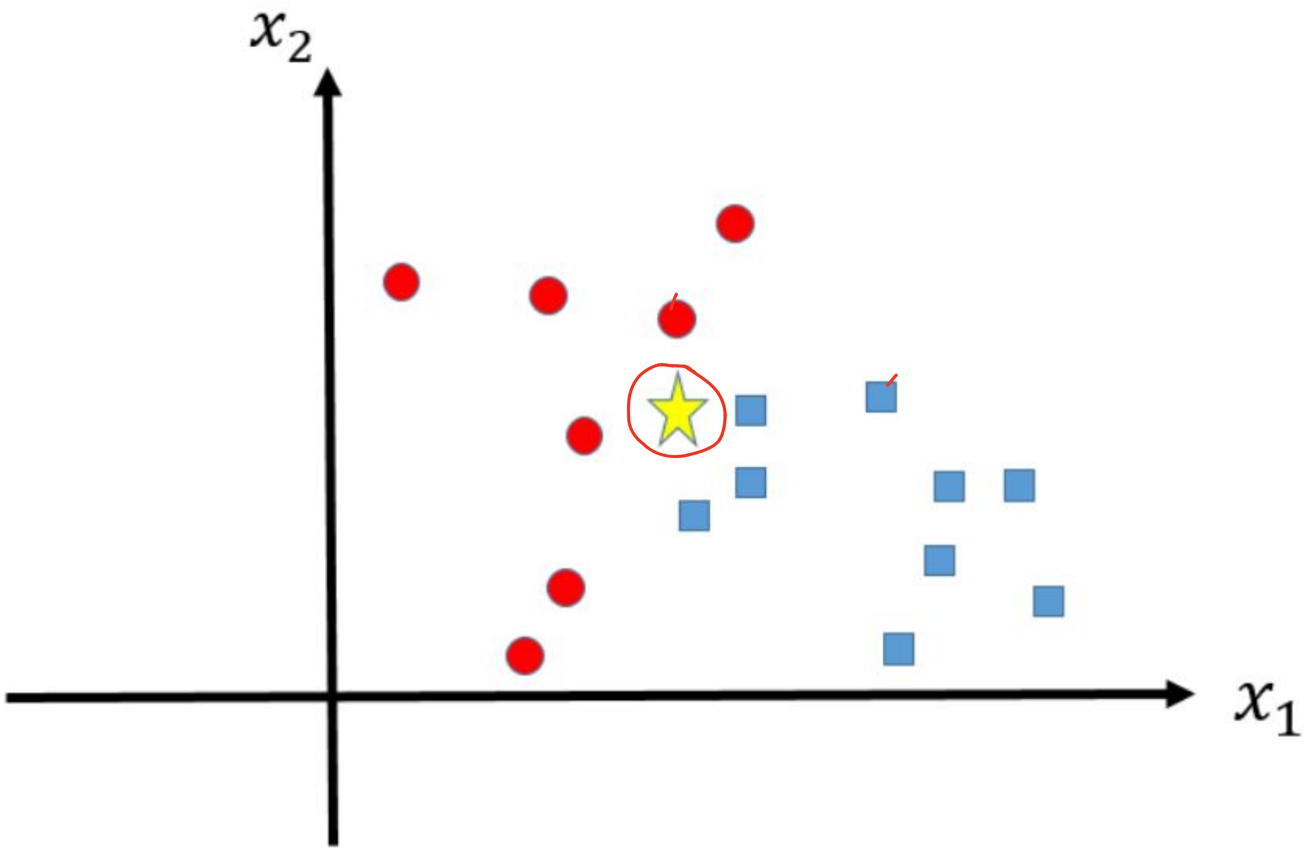
two most similars = {O, D}

Distance applications: k Nearest Neighbours

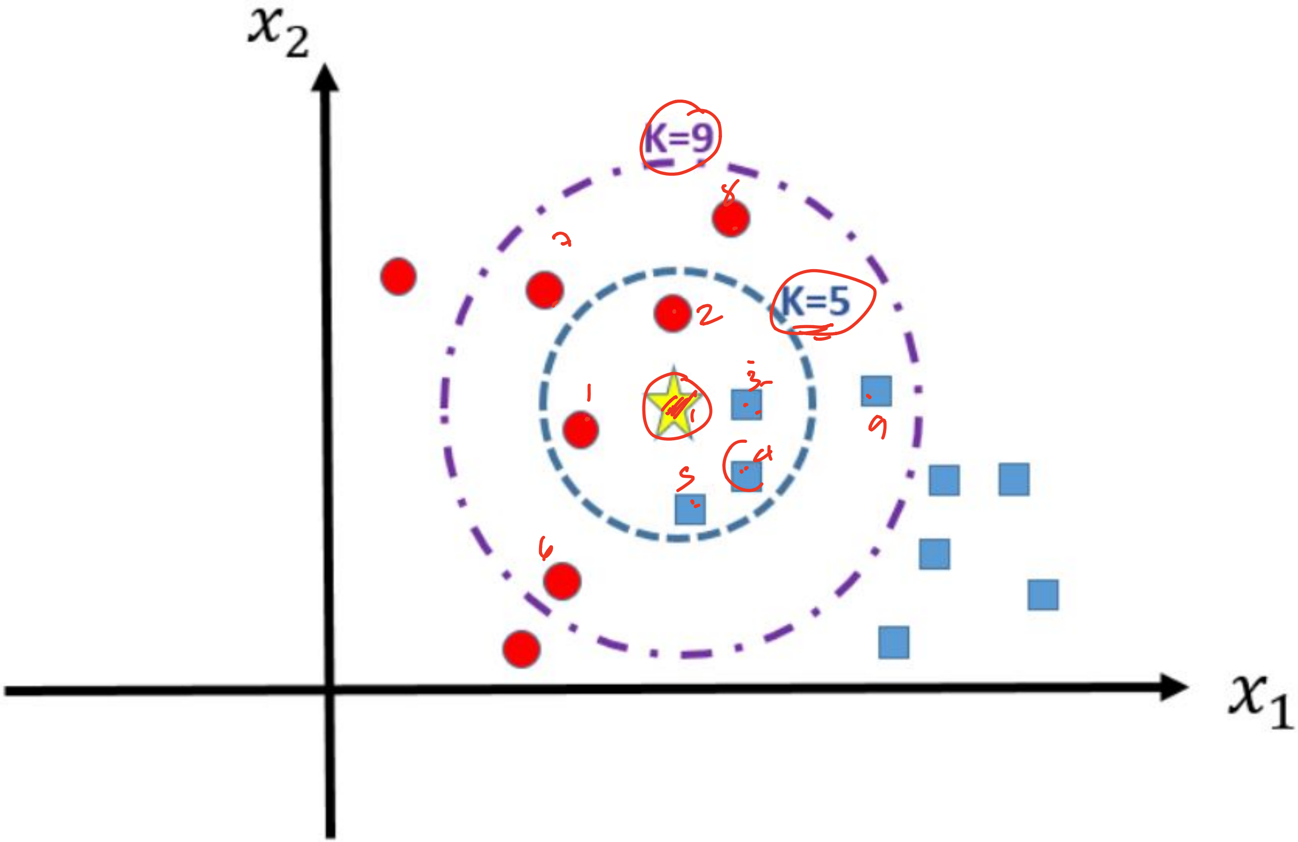
training data



Distance applications: k Nearest Neighbours



Distance applications: k Nearest Neighbours

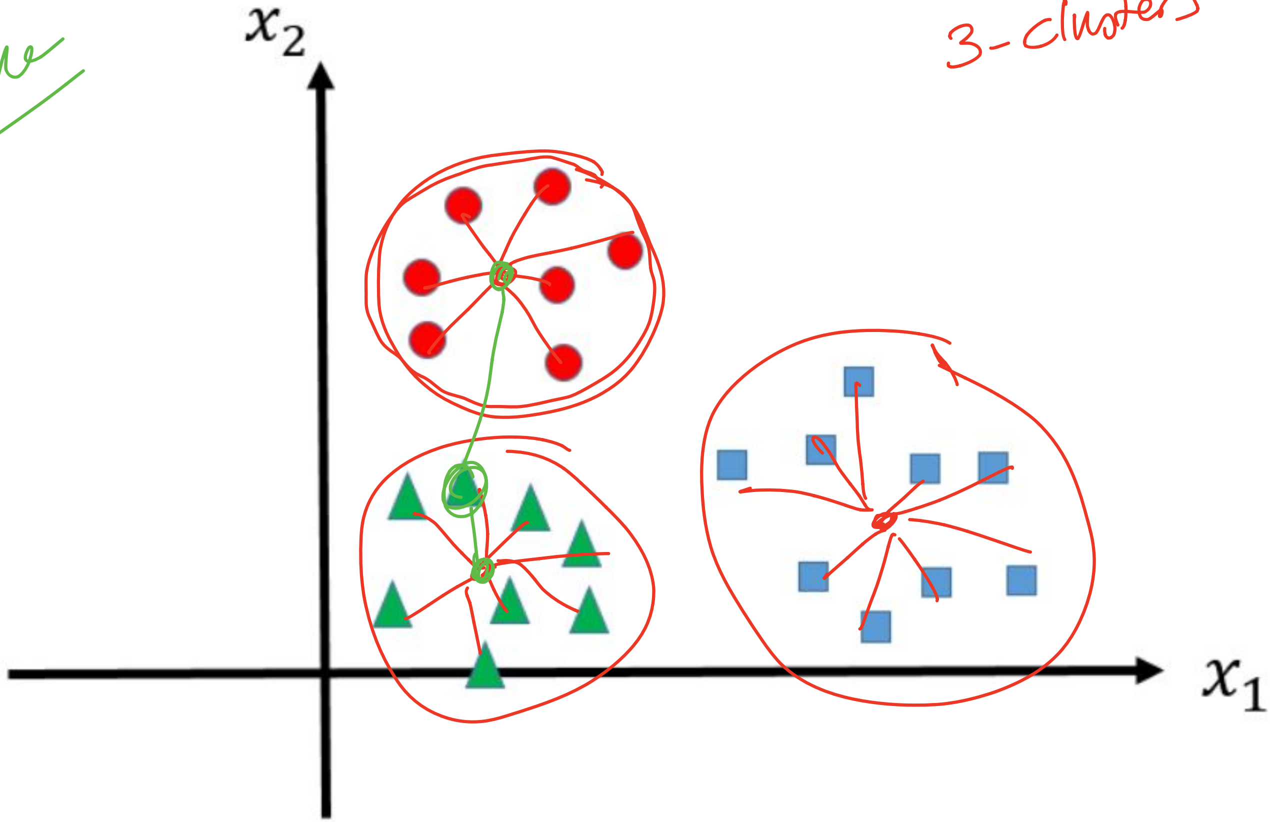


Distance applications: Clustering

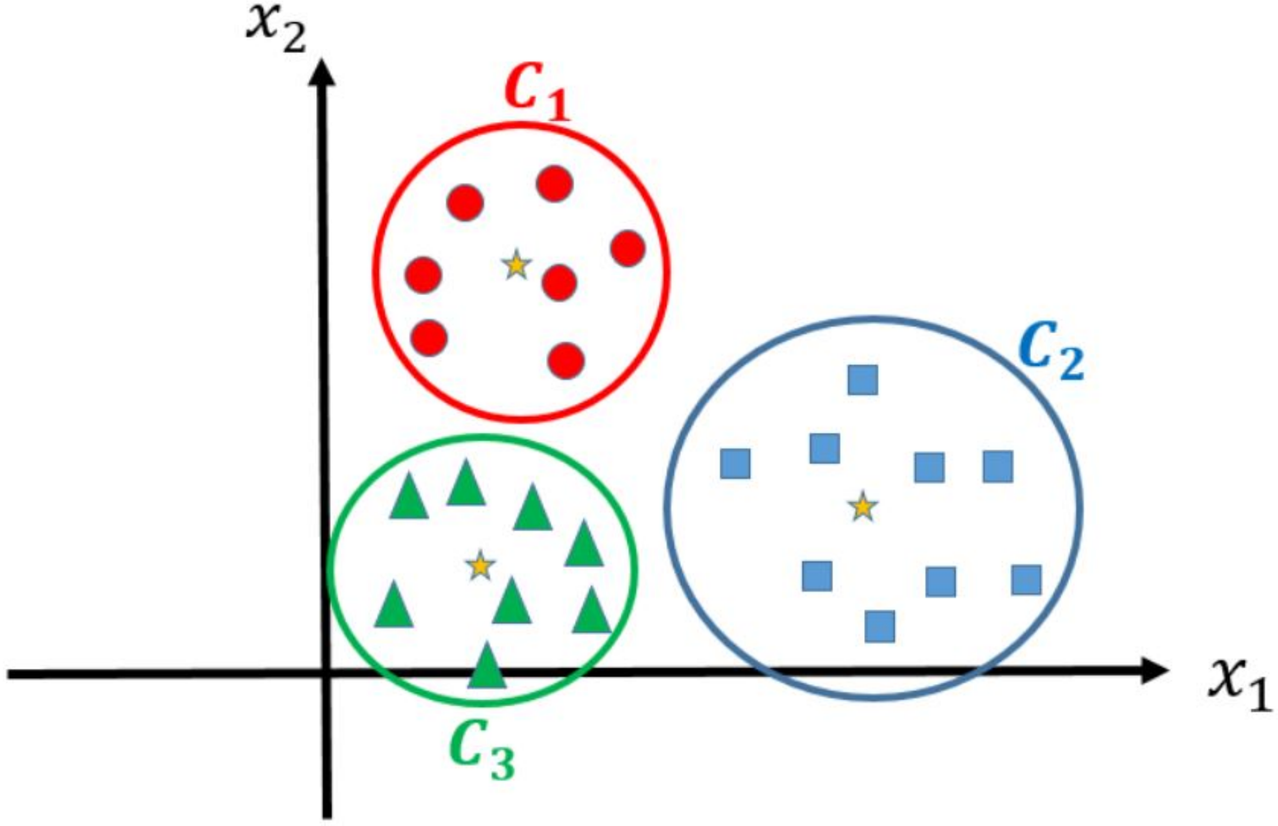
K-means

3-clusters

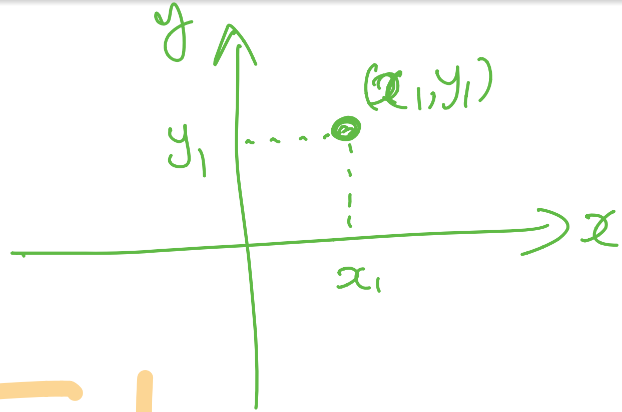
distance



Distance applications: Clustering



ordered pair (x,y)



Functions

$y = f(x)$

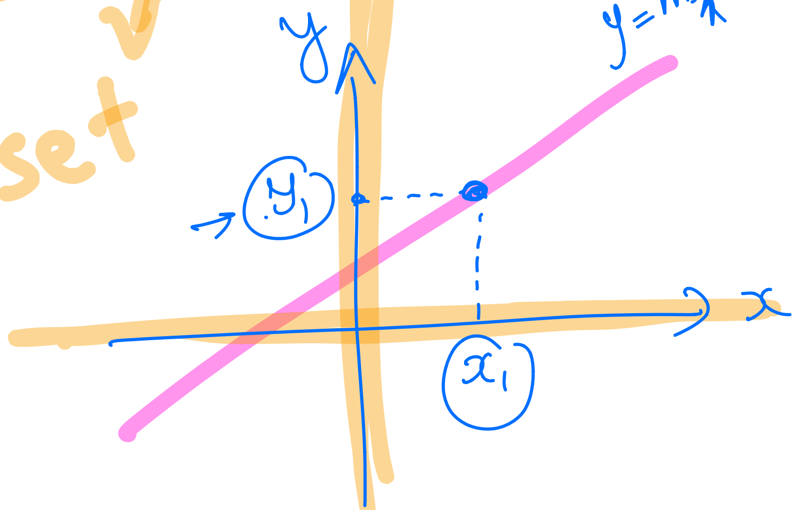
indep

dep

set

set

$y = mx + b$



$x \quad y$

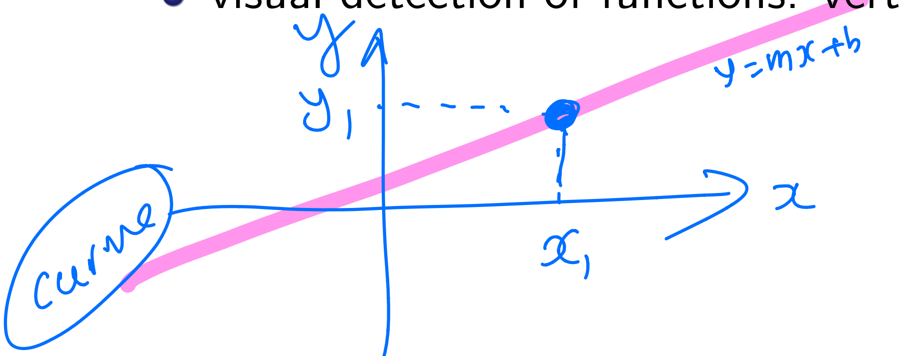
$f : A \to B$

f

$y = f(x)$

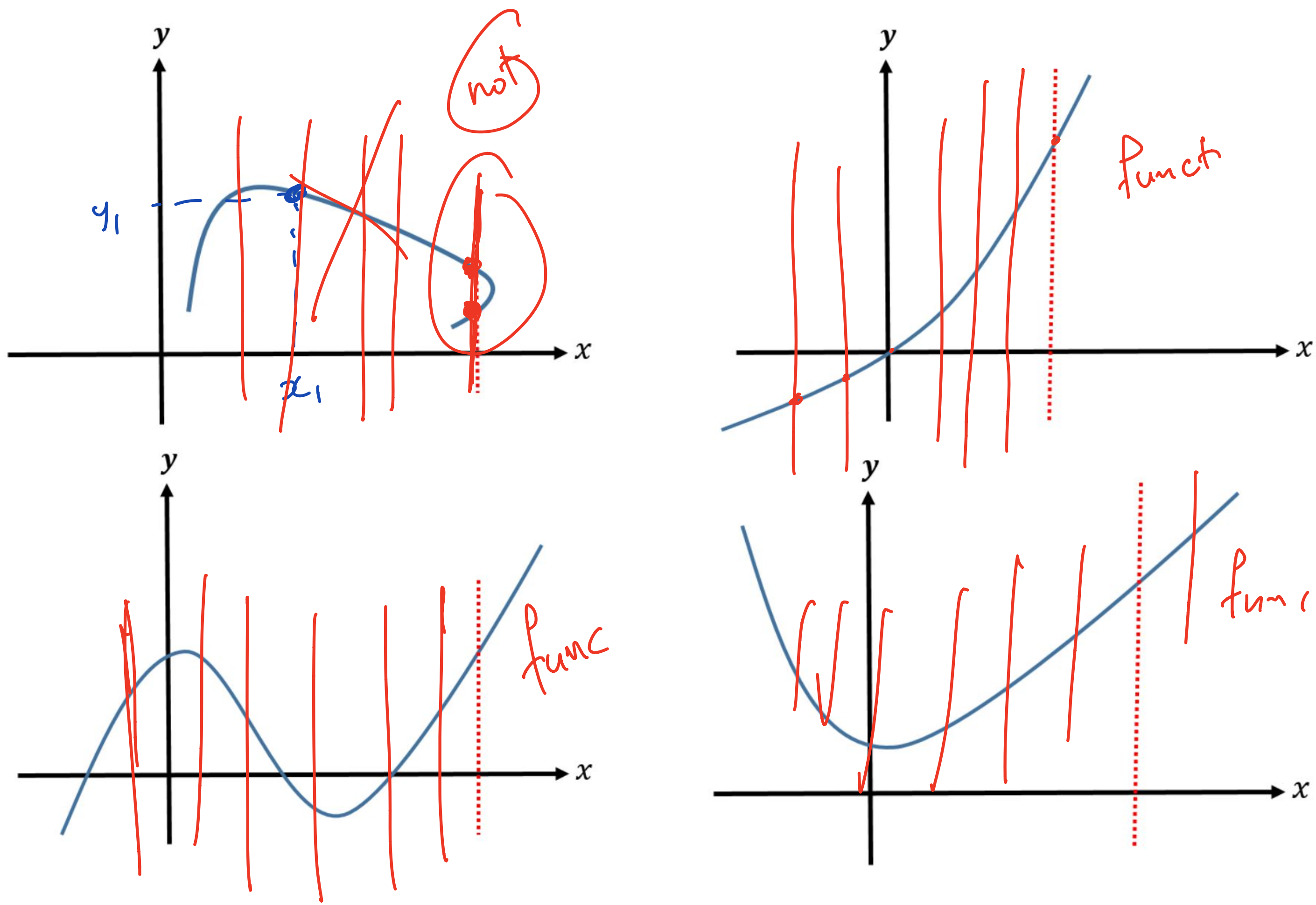
- Ordered pairs (x, y)
- Independent variable (shown generally by x) and dependent variable (shown generally by y)
- A function $f : A \to B$ transforms each elemnt $x \in A$ into $f(x) \in B$
- a particular ordered pair $(x, f(x))$
- visual detection of functions: vertical line test

$(x, y) = (x, f(x))$



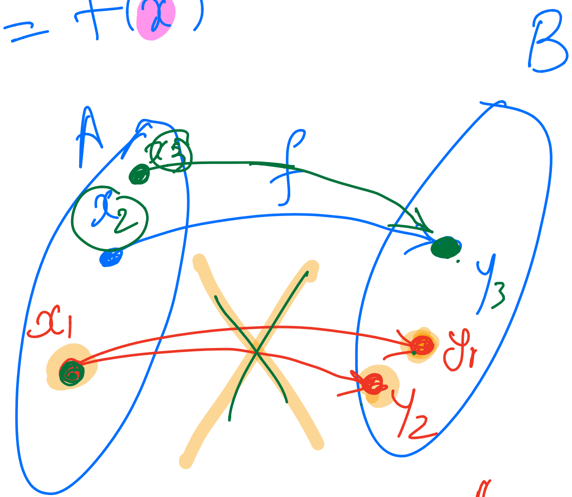
$(x_1, y_1) = (x_1, mx_1 + b)$

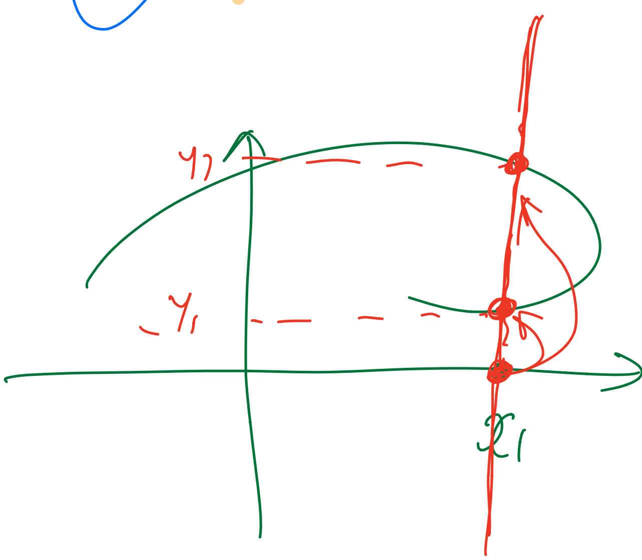
Vertical line tests



$f: A \to B$

$y = f(x)$





not funct

Functions

$f : \mathbb{R} \to \mathbb{R}$

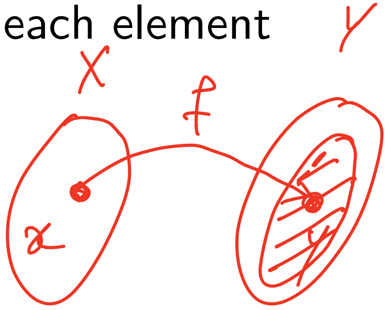
domain

Co-domain

range

$y = f(x)$ → all possible values

- Let X and Y be sets.
- A function f from X to Y assigns (images) each element $x \in X$ to exactly one element in $y \in Y$.



$y = 2x + 1$ $\quad x \in \mathbb{R}$

$\mathrm{dom}(f) = \mathbb{R}$

$$f : X \to Y$$

$$y = f(x)$$

- Functions are also called mappings or transformations.
- X is called domain of f
- Y is called co-domain of f
- The range of f is the set of all images of elements in X

range $\subseteq$ Co-domain

$y = \frac{1}{x}$ $\quad \mathrm{Dom}(f) = \mathbb{R} - \{0\} = \mathbb{R} \setminus \{0\}$

Finding domain of functions

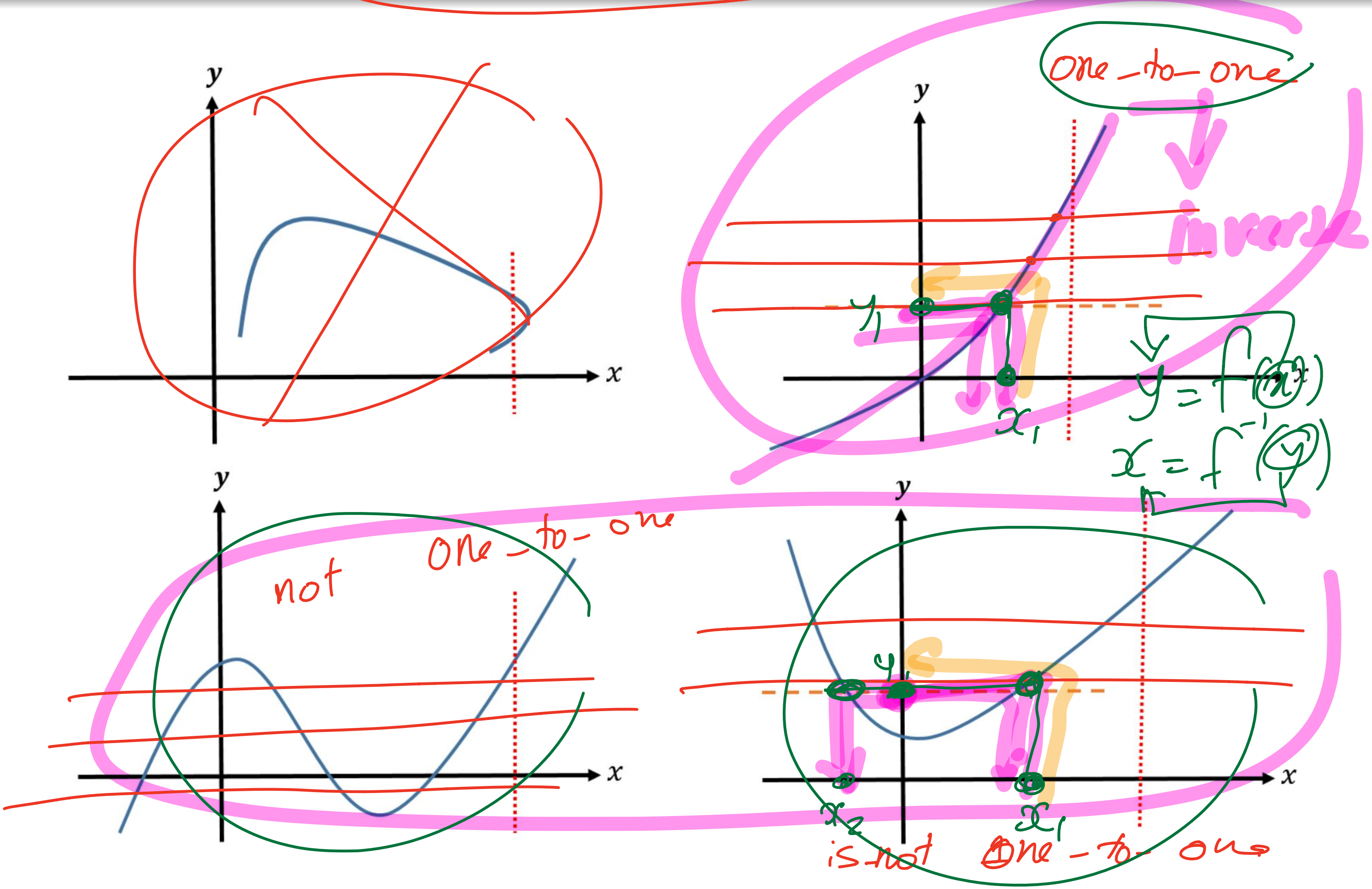
$y=f(x)$ domain, range

$f(x)= \frac{\bigcirc}{\bigcirc} \neq 0$

- Be carefull
  - Never devide by zero! There should not be zero in the denominator. $\mathbb{R}\setminus\{\ldots\}$
  - No negative number under the square root sign. $\sqrt{\boxed{}} \geq 0$
- Generally finding the domain of a function is easier than finding the range of that function.
- Injective or one-to-one
  For all $a, b \in X$, if $f(a) = f(b)$ then $a = b$

assis 1 Q1

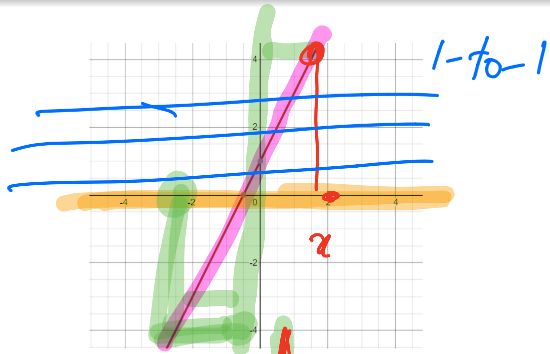
Vertical and horizontal line tests



Some Functions and their graphs

domain co-domain

- A linear function: $f : \mathbb{R} \to \mathbb{R}$

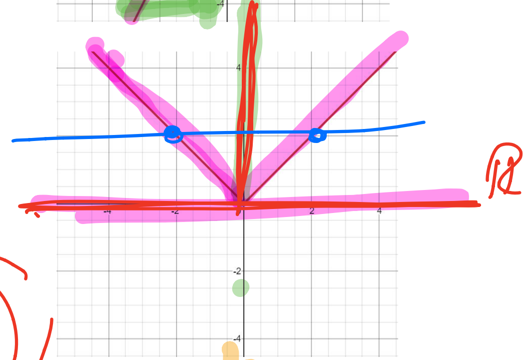


$$y = f(x) = mx + b$$

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

domain co-domain

- Absolute value function: $f : \mathbb{R} \to \mathbb{R}$



$x = 3 \quad y = |3| = 3$

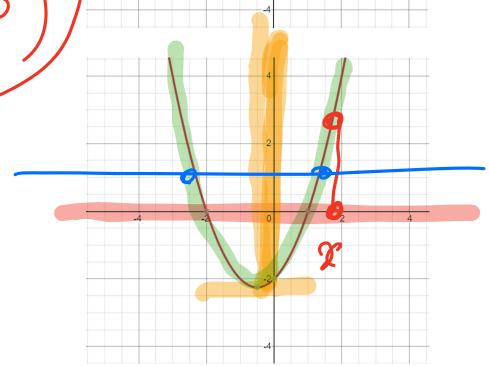
$x = -4 \quad y = |-4| = 4$

$$y = f(x) = |x|$$

range

range of $f = [0, \infty)$

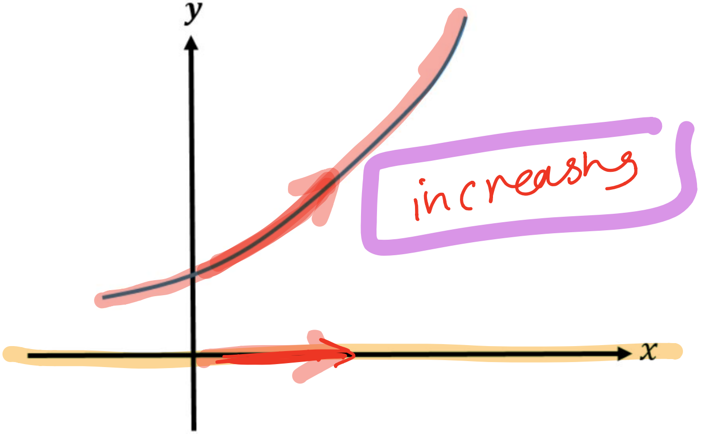
- Qusdratic functions $f : \mathbb{R} \to \mathbb{R}$

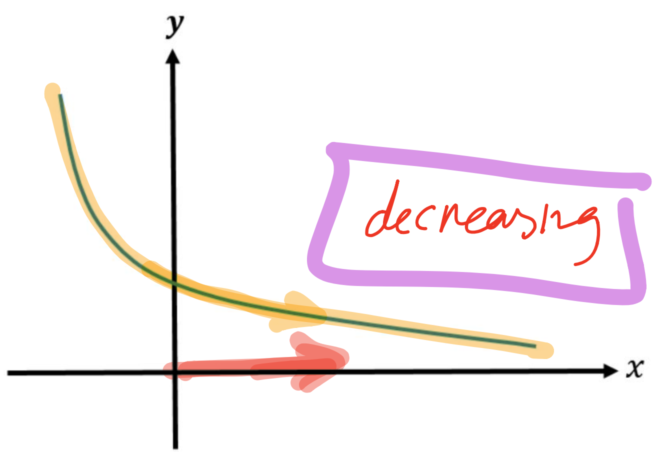


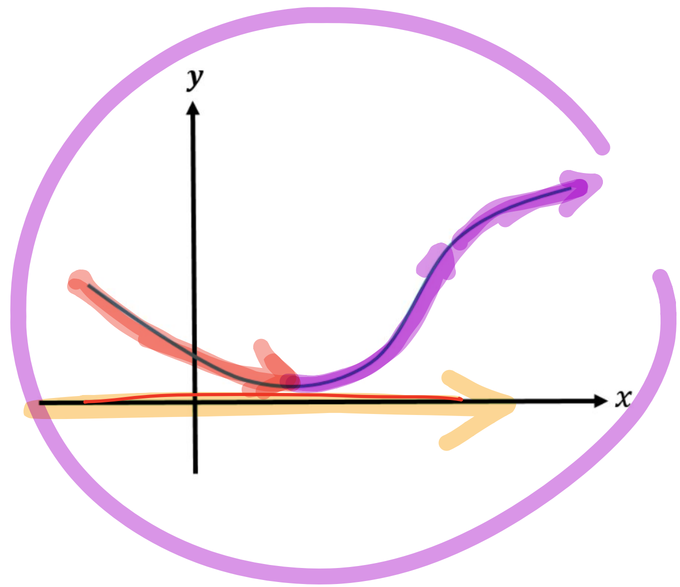
$$y = f(x) = ax^2 + bx + c$$

graph → using vertical line test

→ using horizontal line





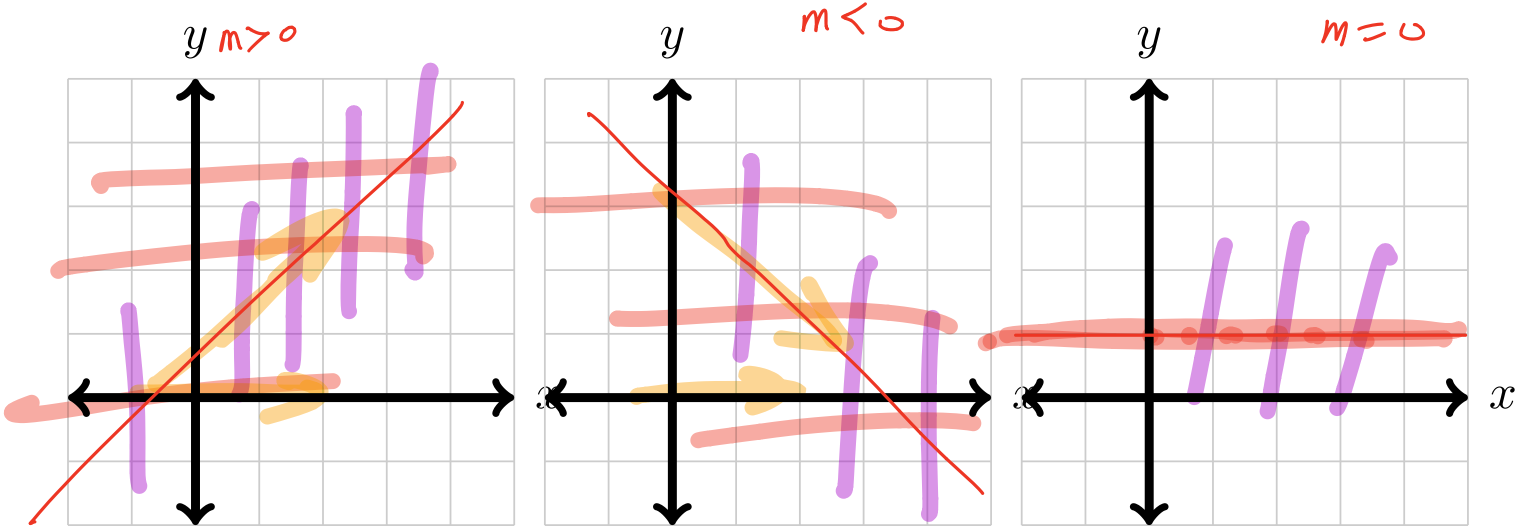


Frequent Classes of functions: Linear Models

- $y = f(x) = mx + b$
  - m slope
  - b y-intercept

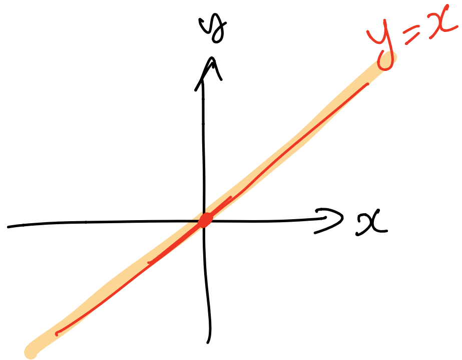
$m \begin{cases} m>0 & \text{increasing} \\ m<0 & \text{decreasing} \\ m=0 & \text{flat} \end{cases}$

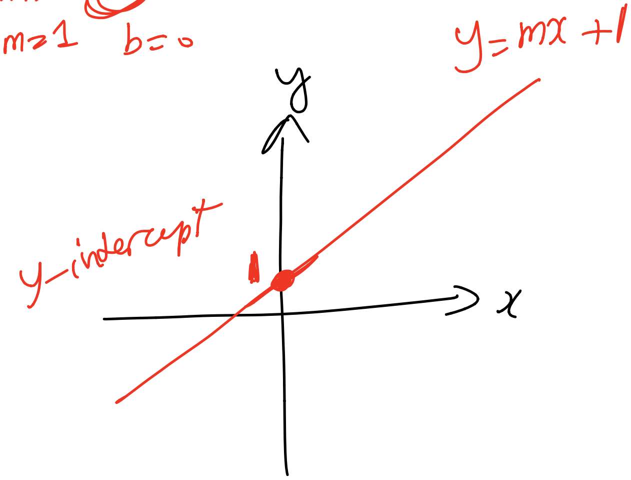
- different types of lines based on different values of m and b



$y = mx + b$

$m=1 \quad b=0$





- $y = f(x) = P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$
  - n nonnegative ineteger
  - $a_n, a_{n-1}, \ldots, a_2, a_1, a_0$ coefficients
  - if $a_n \neq 0$, the degree of polynomial is n
  - domain is $(-\infty, +\infty) = \mathbb{R}$
- Example: $P(x) = 2x^5 - x^3 + \frac{3}{7}x^2 + \sqrt{3}$
- Linear function: polynomial of degree 1, $y = ax + b$
- Quadratic function: polynomial of degree 2. $a \neq 0$

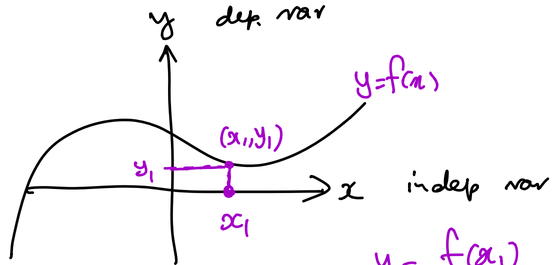
$$P(x) = ax^2 + bx + c, \quad a \neq 0$$

- Cubic function $P(x) = ax^3 + bx^2 + cx + d$ given that $a \neq 0$

$y = f(x)$

↳ independent variable → possible value : domain of f

↳ dependent variable → possible values for y is the range of f



$y_1 = f(x_1)$

$y = f(x) = mx + b$ linear func.s

$y = f(x) = ax^2 + bx + c \quad a \neq 0$ quadratic

$y = f(x) = ax^3 + bx^2 + cx + d \quad a \neq 0$ cubic func.s

$\vdots$

Polynomial $\quad y = f(x) = P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$

$a_n \neq 0$

$\boxed{y = f(x)}$

Frequent Classes of functions: Power and Rational Functions

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

$y = x^{1/3} = \sqrt[3]{x}$

$y = x^{\frac{m}{n}} = \sqrt[n]{x^m} = \left(\sqrt[n]{x}\right)^m$

- Power functions
  - $f(x) = x^a$, where a is a constant
    - $f(x) = x^n$, n is a positive integer
    - $f(x) = x^{\frac{1}{n}} = \sqrt[n]{x}$, n is a positive integer, root function
    - $f(x) = x^{-1} = \frac{1}{x}$ reciprocal function

$x \neq 0$

- Rational functions
  - $f(x) = \frac{P(x)}{Q(x)}$ where P and Q are polynomials

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

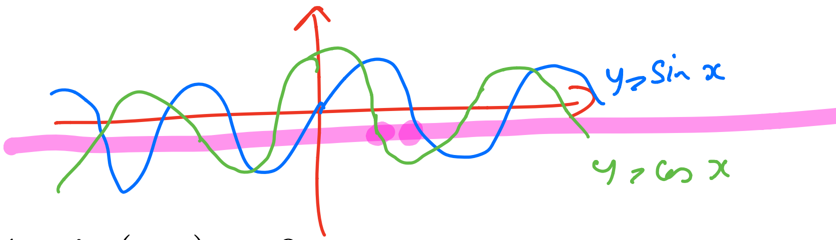
$f(x) = \frac{x^2+1}{x^3-x-1}$

$y = x^{1/2} = \sqrt{x}$ $\quad x \geq 0$

$y = x^{1/4}$ $\quad$ 4 even $\quad y = \sqrt[4]{x}$ $\quad x \geq 0$

$y = x^{1/3} = \sqrt[3]{x}$ $\quad x \in \mathbb{R}$

$y = x^{\frac{1}{2k}} = \sqrt[2k]{x}$ $\quad x \geq 0$



- $y = \sin(x)$, $-1 \leq \sin(x) \leq 1$, $\sin(n\pi) = 0$
- $y = \cos(x)$, $-1 \leq \cos(x) \leq 1$
- $y = \tan(x) = \dfrac{\sin(x)}{\cos(x)}$
- $y = \cot(x) = \dfrac{\cos(x)}{\sin(x)}$
- $y = \sec(x) = \dfrac{1}{\cos(x)}$
- $y = \csc(x) = \dfrac{1}{\sin(x)}$

power functn $y = f(x) = x^n$

base

- $y = f(x) = b^x$ where the base b is a positive constant $y = 2^x$
  - $b > 1$ growth function
  - $b < 1$ decay function $y = f(x) = (\frac{1}{2})^x$
- Index laws
  - $(a^n)(a^m) = a^{n+m}$
  - $\frac{a^n}{a^m} = a^{n-m}$
  - $(ab)^n = (a^n)(b^n)$
  - $(\frac{a}{b})^n = \frac{a^n}{b^n}$
  - $(a^n)^m = a^{mn}$
  - $\boxed{\text{for } a \neq 0,\ a^0 = 1}$ and $a^1 = a$
  - $a^{-n} = \frac{1}{a^n}$
- $y = e^x$ where $e \approx 2.718$ is called natural exponential function

$\frac{2^3}{2^5} = 2^{3-5} = 2^{-2} = \frac{1}{2^2}$

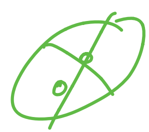
$(a^n)^m = \underbrace{a^n \times a^n \times \cdots \times a^n}_{m\text{-times}}$

$= a^{n+n+\cdots+n} = a^{nm}$

$y = b^x$

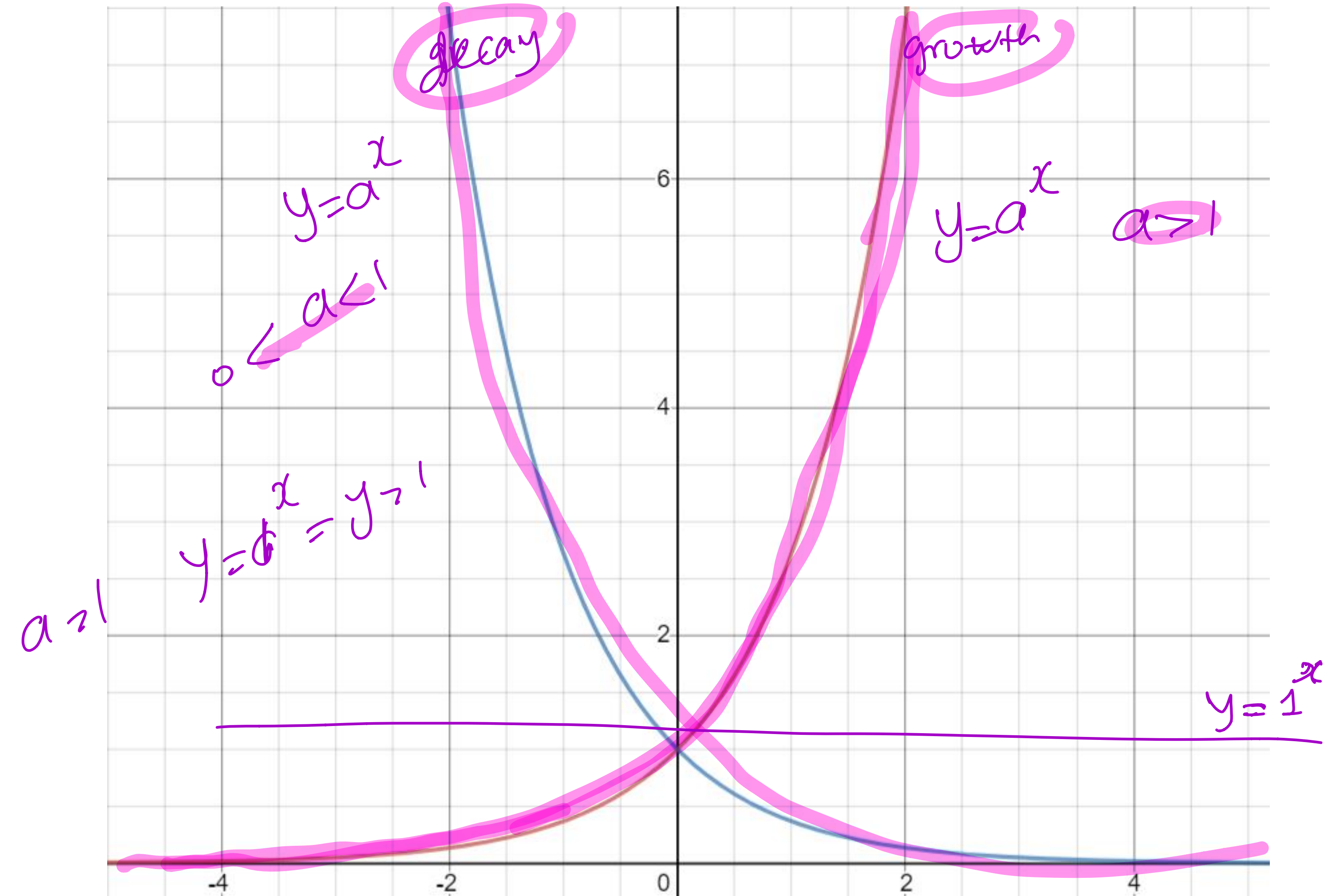
$y = e^x$

$b = \underbrace{2.718}_{e}$



$\boxed{a^0 = 1 \quad \boxed{a \neq 0}}$

Growth and Decay



Frequent Classes of functions: Logarithmic

- $y = f(x) = \log_b(x)$ where the base b is a positive constant
- $\boxed{y = \log_b(x)}$ same as $x = b^y$
- Logrtaithimic properties
  - $\log_b(xy) = \log_b(x) + \log_b(y)$
  - $\log_b(\frac{x}{y}) = \log_b(x) - \log_b(y)$
  - $\log_b(x^r) = r\log_b(x)$
- if $b = e$, $\boxed{\log_e(x)} = \boxed{\ln(x)}$ natural logrith
  - $\ln(x) = y \iff e^y = x$
  - $\ln(e^x) = x$ for all $x \in \mathbb{R}$
  - $e^{\ln(x)} = x$ for all $x > 0$
  - $\ln(e) = 1$
- $\log(x_1x_2x_3) = \log(x_1) + \log(x_2) + \log(x_3)$
- $\log\left(\prod_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} \log(x_i)$ and

$$\log\left(\prod_{i=1}^{n} f(x_i)\right) = \sum_{i=1}^{n} \log(f(x_i))$$

exponential $y = b^x$ $b > 0$

logarithmic $y = \log_b(x)$

$\log_a(AB) = \log_a(A) + \log_c(B)$

$\log_a\left(\frac{A}{B}\right) = \log_a(A) - \log_a(B)$

$\ln x$

e^x

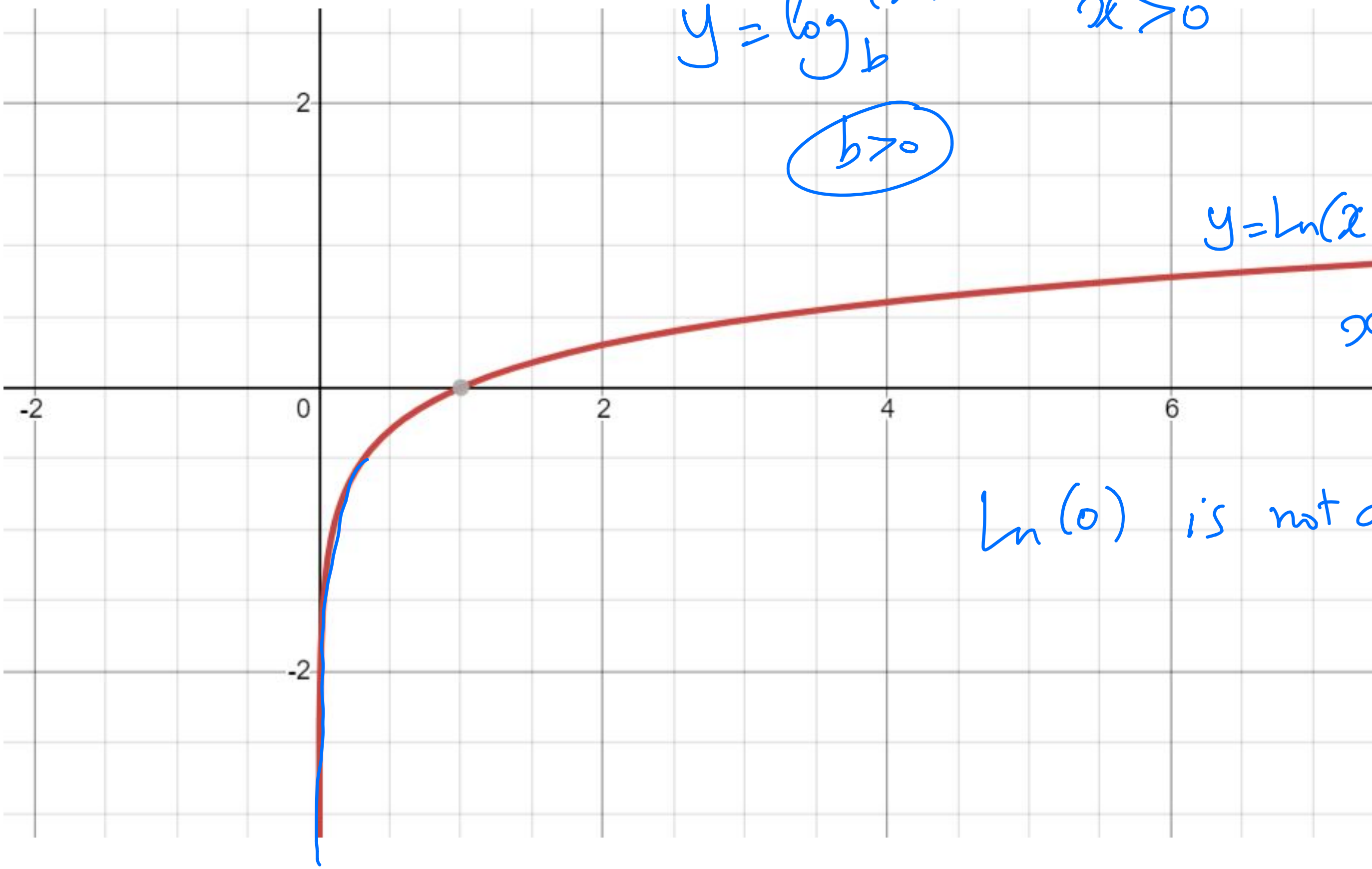
$\log_e e = 1$

$\ln(e) = 1$

Logarithmic function

$y = \log_b(x)$ $\quad x > 0$

$b > 0$



$y = \ln(x)$

$x > 0$

$\ln(0)$ is not defined

Composition of functions

$y = f(x)$

$y = g(x)$

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

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

- $(f \circ g)(x) = f(g(x))$
- Example
  - $f(x) = x^2$, and $g(x) = x - 2$, then
    $(f \circ g)(x) = f(g(x)) = f(x-2) = (x-2)^2$
  - $(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 - 2$
- In general $(f \circ g)(x) \neq (g \circ f)(x)$
- $(f \circ g \circ h)(x) = f(g(h(x)))$

Derivatives

$y = f(x)$

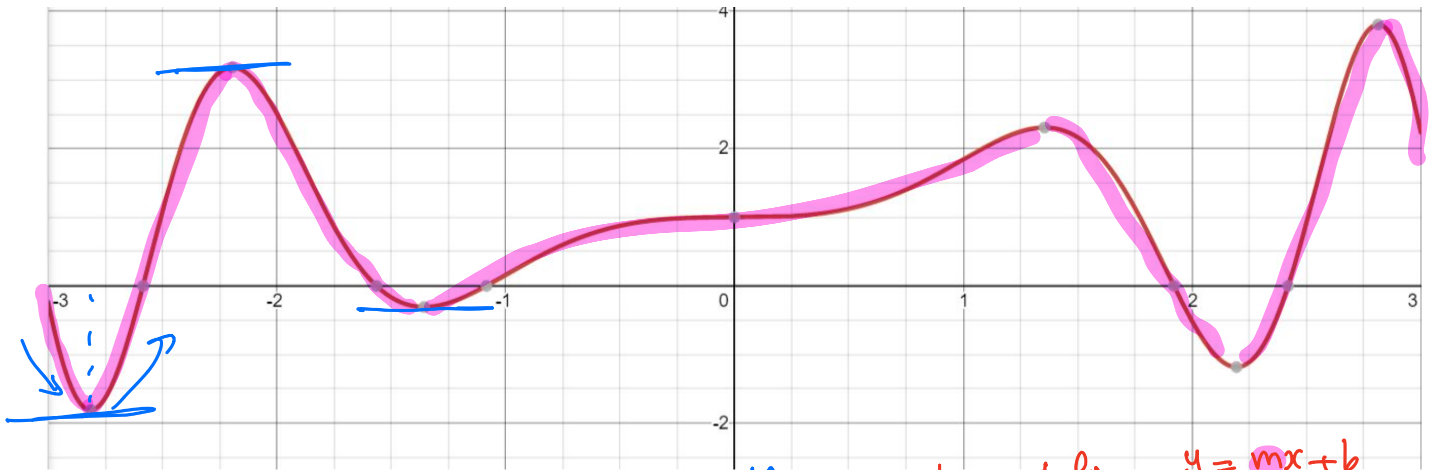
$x \to$ indep. variable

$y \to$ dep. var

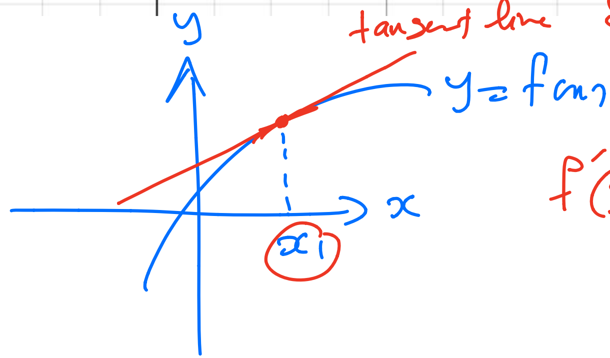
the rate of change of y, with respect to the rate of change of x.

Why do we need derivative?

$$\boxed{y = f(x)}$$



tangent line $y = mx + b$



$f'(x_1) = m$

Derivative: the rate of change

$y = f(x)$

$f'(a)$

- slope of $y = f(x)$ at $x = a$

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

- sign and magnitude of slope
  - Slope is positive at $x = a$: increasing
  - Slope is negative at $x = a$: decreasing
  - $f'(a) \leq f'(b)$
  - Slope is zero at $x = a$: stationary point
- for $y = f(x)$, $\boxed{f'(x) = y' = \frac{dy}{dx} = \frac{df}{dx}}$ are the same

$\dot{y}$

Derivative Table

- $f(x) = c$, then $f'(x) = 0$
- $y = x$, then $y' = 1$
- $y = x^n$, then, $y' = nx^{n-1}$
- $(cf(x))' = cf'(x)$
- $y = af_1(x) + bf_2(x)$, then $y' = af_1'(x) + bf_2'(x)$
- $y = f_1(x)f_2(x)$, then $y' = f_1'(x)f_2(x) + f_1(x)f_2'(x)$
- $y = \frac{f(x)}{g(x)}$, then

linear combination of two functions

product rule

quotient rule

$$y' = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

$$\cos(x)\cos(x) = (\cos(x))^2 = \cos^2(x)$$

- $y = \sin(x)$, then $y' = \cos(x)$
- $y = \cos(x)$, then $y' = -\sin(x)$
- find the derivative of $y = \tan(x)$. (ans. $\sec^2(x)$)
- find the derivative of $y = \cot(x)$. (ans: $-\csc^2(x)$)
- find the derivative of $y = \sec(x)$. (ans: $\sec(x)\tan(x)$)
- find the derivative of $y = \csc(x)$. (ans: $-\csc(x)\cot(x)$)

$$y = \tan(x) = \frac{\sin(x)}{\cos(x)}$$

$$y' = \frac{(\sin(x))'\cos(x) - (\sin(x))(\cos(x))'}{[\cos(x)]^2}$$

$$= \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{[\cos(x)]^2}$$

$$= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \left[\frac{1}{\cos(x)}\right]^2 = \sec^2(x)$$

More derivatives

$y = \ln(\sin(x))$

$y' = \frac{\cos(x)}{\sin(x)} = \cot an(x)$

- $\frac{d}{dx}(\ln(x)) = \frac{1}{x}$
- $\frac{d}{dx}[\ln(g(x))] = \frac{g'(x)}{g(x)}$
- $\frac{d}{dx}(\sin(kx)) = k\cos(kx)$
- $\frac{d}{dx}(\cos(kx)) = -k\sin(kx)$
- $\frac{d}{dx}(e^{kx}) = ke^{kx}$

The Chain Rule

$y = \sin(x^2+1)$

$f(x) = \sin(x) \to f'(x) = \cos(x)$

$g(x) = x^2+1 \quad g'(x) = 2x$

- When $y = f \circ g(x)$ or $\boxed{y = f(g(x))}$
- $y' = g'(x)f'(g(x))$
- Another interpretation:
  - if $y = f(u)$, we can talk about the rate of change of y when independent variable u is changing, which is $\frac{dy}{du}$.
  - If $u = g(x)$, then we can talk about the rate of change of u, when independent variable x is changing, which is $\frac{du}{dx}$.
  - $y = f(u)$ and $u = g(x)$, then $y = f(g(x))$, and y is a function of x. The rate of change of y with respect to x is

$f(g(x)) = \sin(g(x)) = \sin(x^2+1)$

$$\boxed{\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}}$$

$y = f(g(x))$

$y = f(u) \qquad u = g(x)$

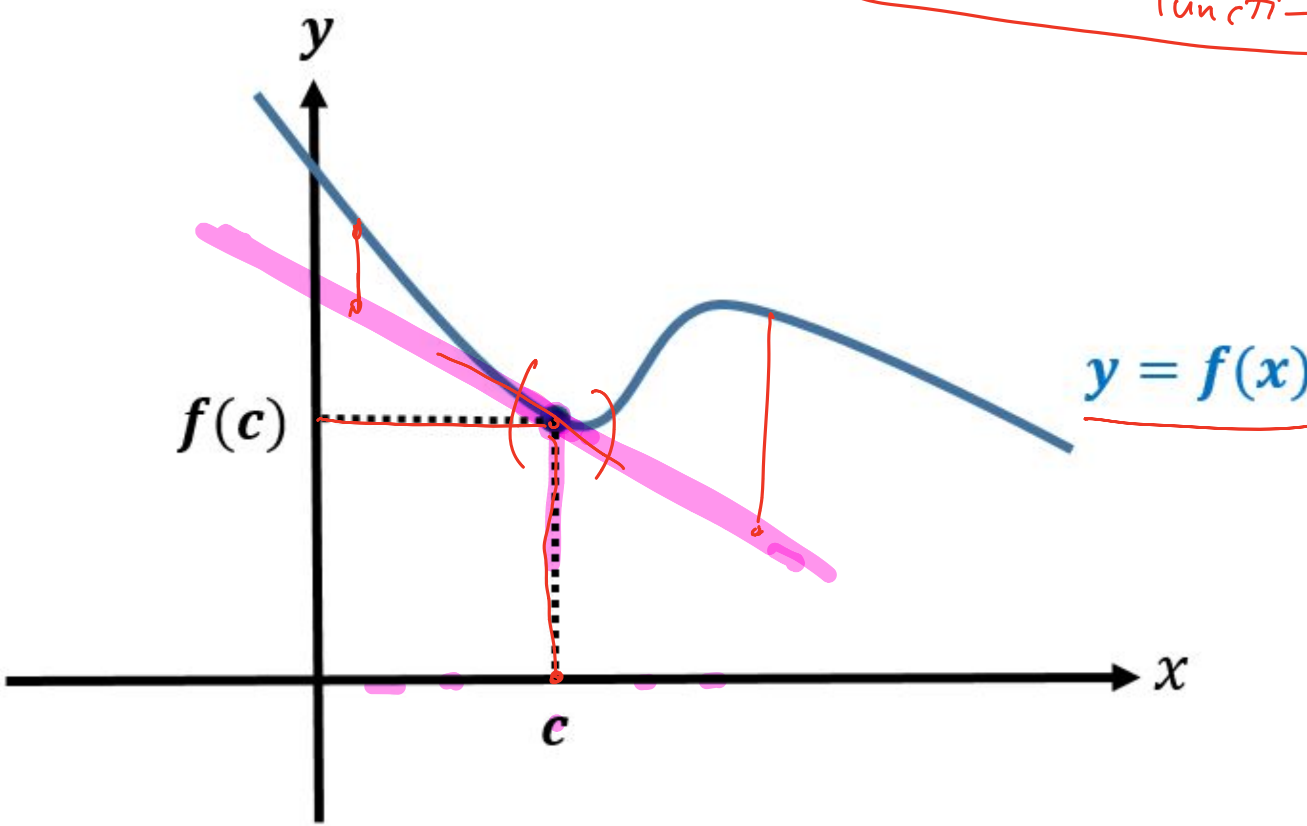
$\frac{dy}{dx} = \boxed{\frac{dy}{du}} \cdot \frac{du}{dx}$

$y' = (2x)\cos(x^2+1)$

$y' = 2x\cos(x^2+1)$

Linear Approximation

linear funcs. most simple functions



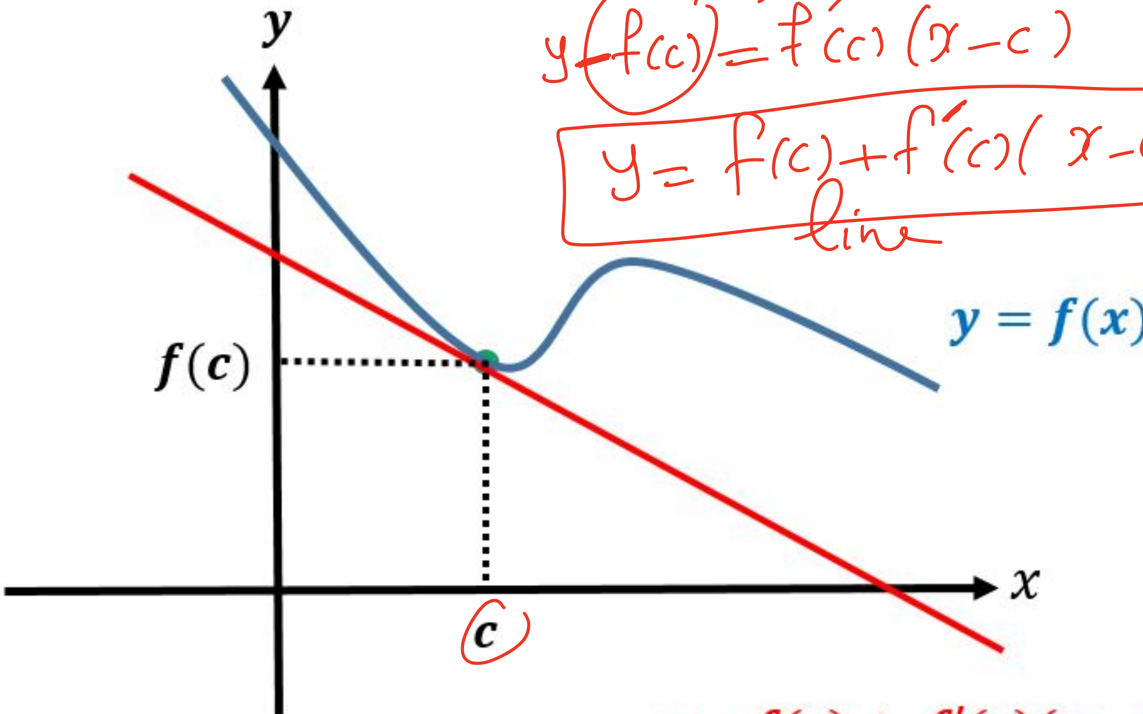
Linear Approximation

$y = f(x)$ at $x = c$

$y - f(c) = f'(c)(x - c)$

$\boxed{y = f(c) + f'(c)(x - c)}$

line



$$y = f(c) + f'(c)(x - c)$$

Finding maxima and minima

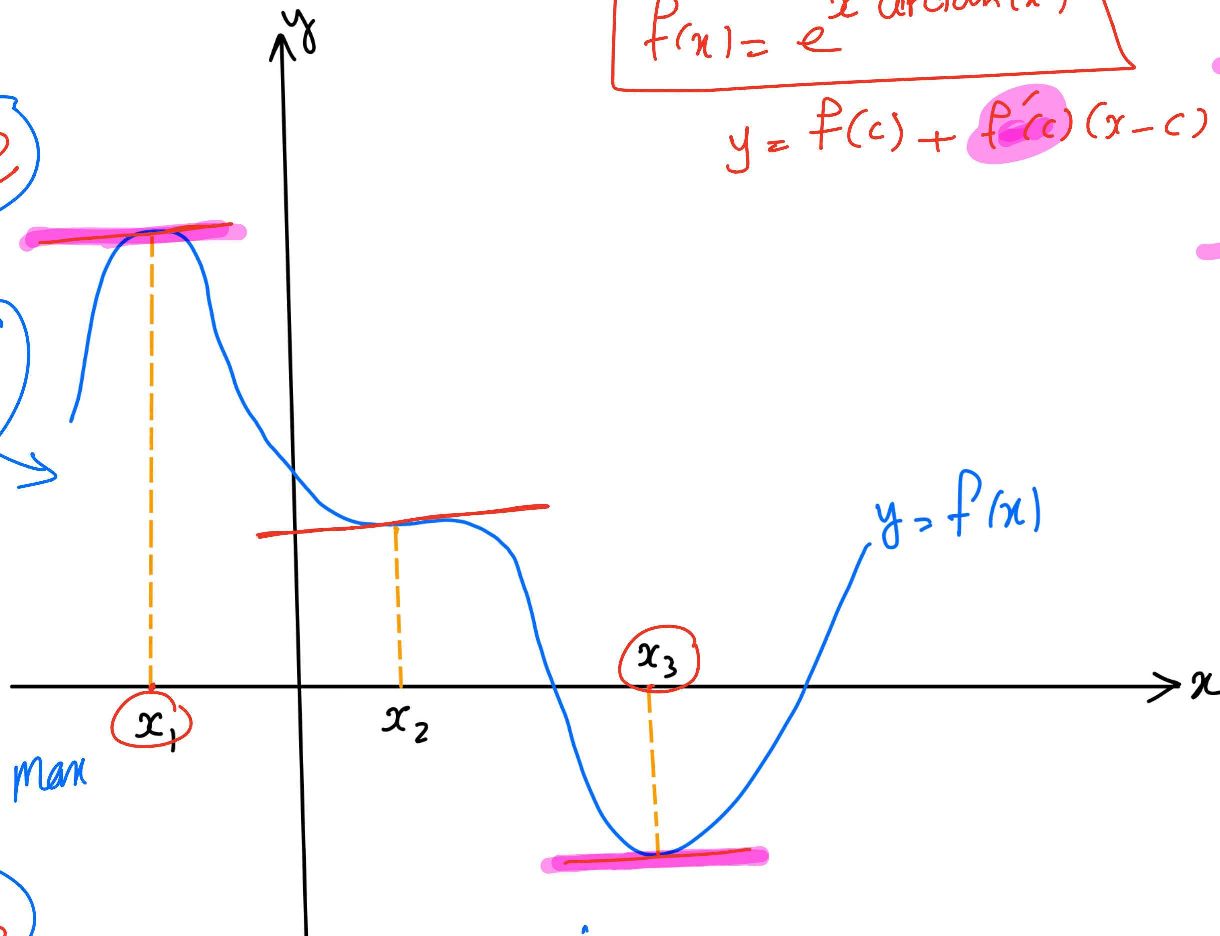
$y = f(x)$

$f'(x) = 0$

stationary points

$$\boxed{f(x) = e^{x^2 \arctan(x)}}$$

$y = f(c) + f'(c)(x - c)$

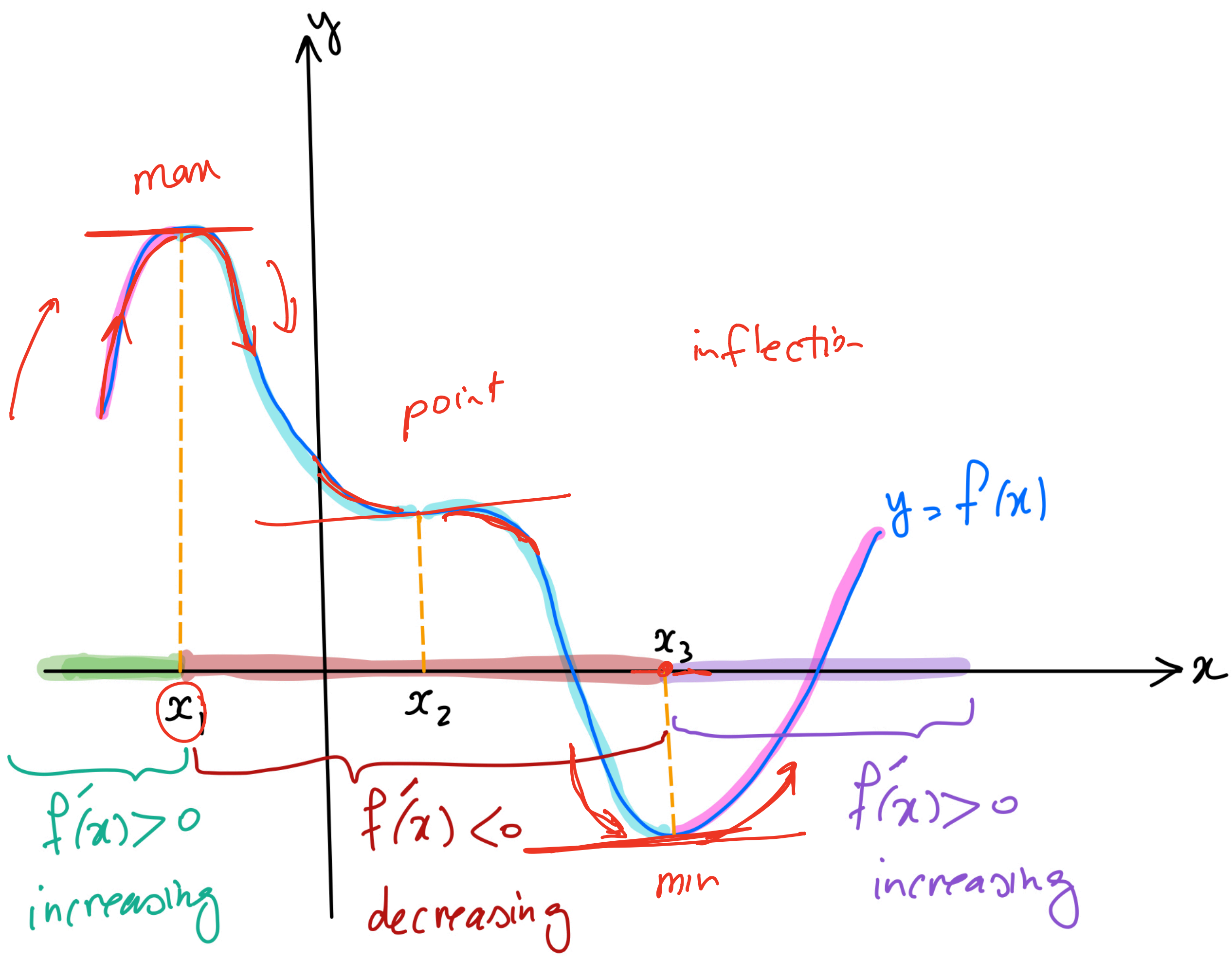


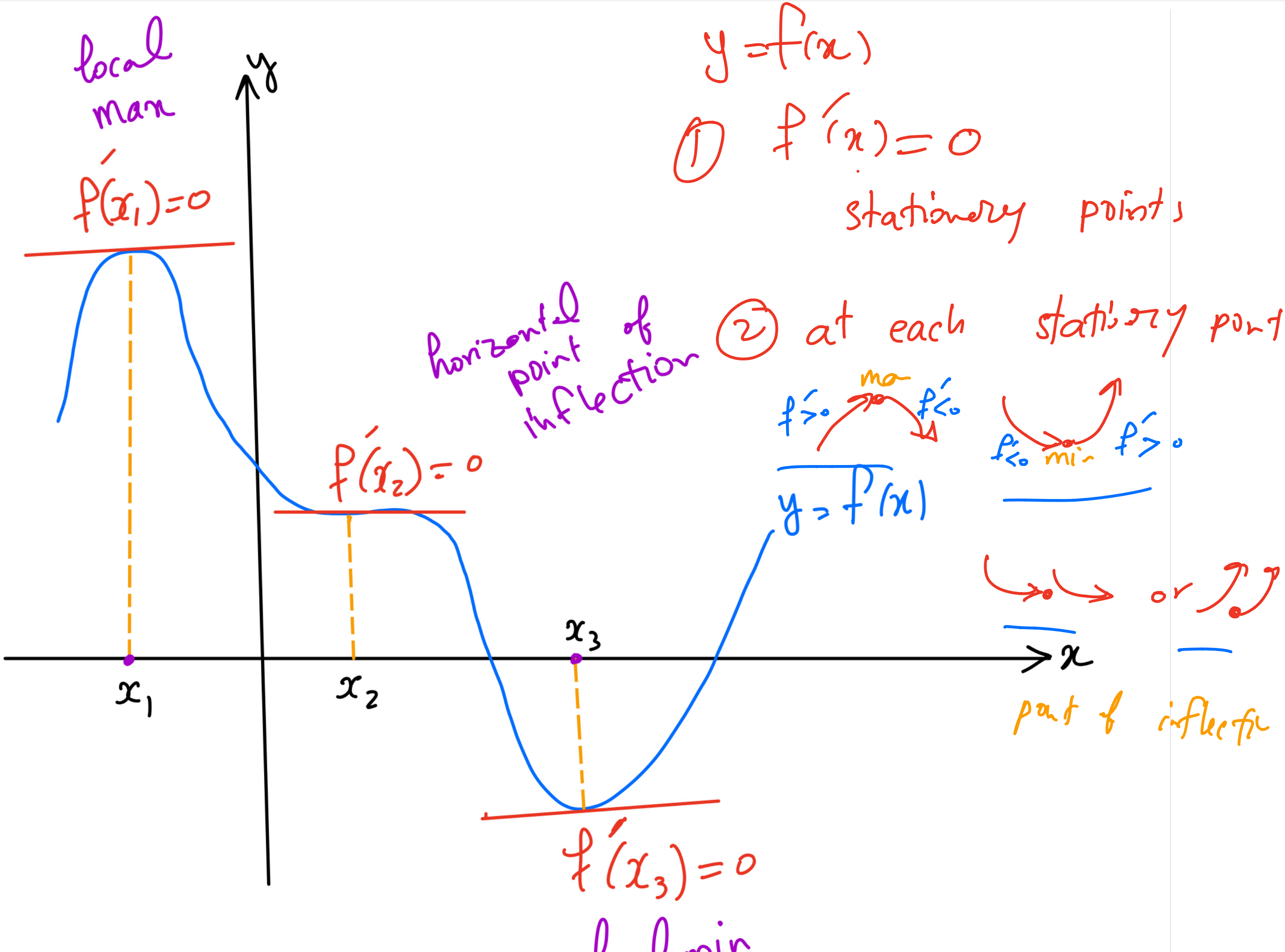
$f'(x_1) = 0$ max

$f'(x_2) = 0$

$f'(x_3) = 0$ min

Finding maxima and minima





$y=f(x)$

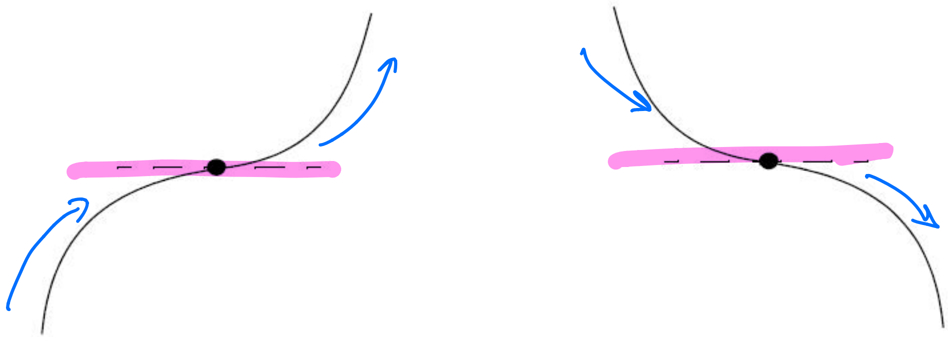
① $f'(x)=0$

stationary points

② at each stationary point

Finding Maxima and minima

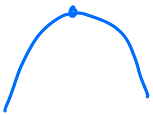
- The points on a curve $y = f(x)$ at which the slope (gradient) is 0 are called **stationary points** (or **critical points**).
- For $y = f(x)$ if $f'(c) = 0$, then c is a stationary point
- A stationary point could be
  - a local maximum,
  - a local minimum, or
  - a horizontal point of inflection



Finding Maxima and minima of $y = f(x)$

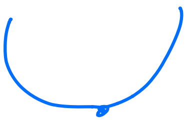
- solve $\boxed{f'(c) = 0}$ and find c values
  - The stationary point $x = c$ is a local maximum point if

$$\begin{cases} x < c & f'(x) > 0 \\ x > c & f'(x) < 0 \end{cases}$$



  - The stationary point $x = c$ is a local minimum point if

$$\begin{cases} x < c & f'(x) < 0 \\ x > c & f'(x) > 0 \end{cases}$$



  - a horizontal point of inflection otherwise.

Maxima and minima of $f(x)$ using second derivative

$y = f(x) \longrightarrow y' = f'(x) \rightarrow y'' = f''(x) \rightarrow y''' = f'''(x)$

$\rightarrow \cdots \rightarrow y^{(n)} = f^{(n)}(x)$

- Stationary (critical) points: $\boxed{f'(c)} = 0$
  - if $f''(c) > 0$, $x = c$ is a local minima
  - if $f''(c) < 0$, $x = c$ is a local maxima
  - if $f''(c) = 0$, $x = c$ no conclusion
- $f''(x) > 0$ at a point, the curve is concave up (convex) at that point
- $f''(x) < 0$ at a point, the curve is concave down (concave) at that point

$y = x^2 \longrightarrow y' = 2x \longrightarrow \boxed{y'' = 2} \longrightarrow y''' = 0, \quad y^{(4)} = 0$

$y = f(x)$

↳ independent var

↳ dependent variable

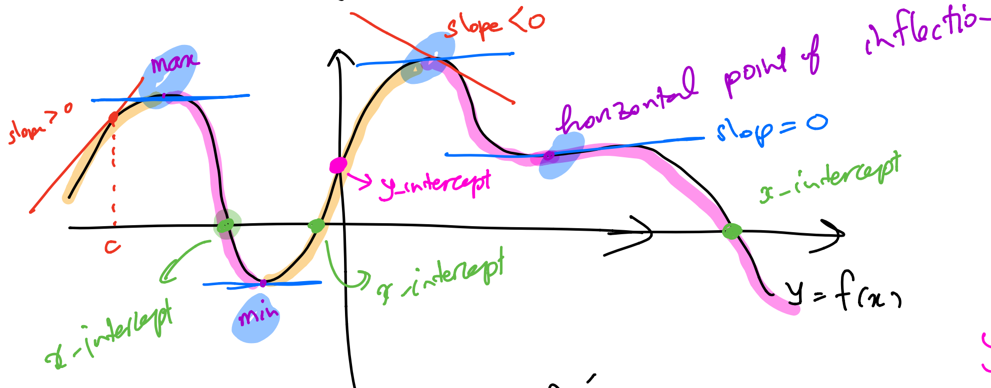
model

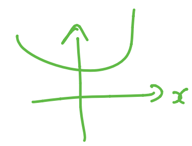
$y = f(x_1, x_2, \dots, x_n)$

↳ dep. var

indep. vars

derivative $\quad y = f(x) \qquad y' = f'(x)$





$f(x) = 0$

$y = f(0)$

at any point x_0 $\begin{cases} f'(x_0) > 0 \rightarrow \text{inc} \\ f'(x_0) < 0 \rightarrow \text{dec} \\ f'(x_0) = 0 \quad \text{stationary point} \end{cases}$

$y = f(x)$

$y' = f'(x)$, $\quad f'(x) = 0 \rightarrow$ find x values of the stationary point

use the second derivative to classify them

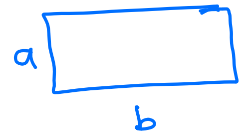
at x_0, that $f'(x_0) = 0$, $\quad f''(x_0) > 0 \rightarrow x_0$ a min point

$f''(x_0) < 0 \rightarrow x_0$ a max point

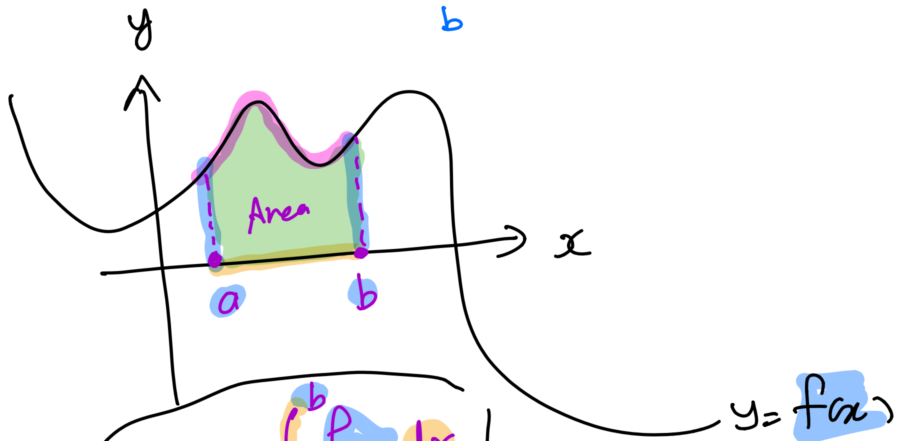
$f''(x_0) = 0$ no information

Integrals

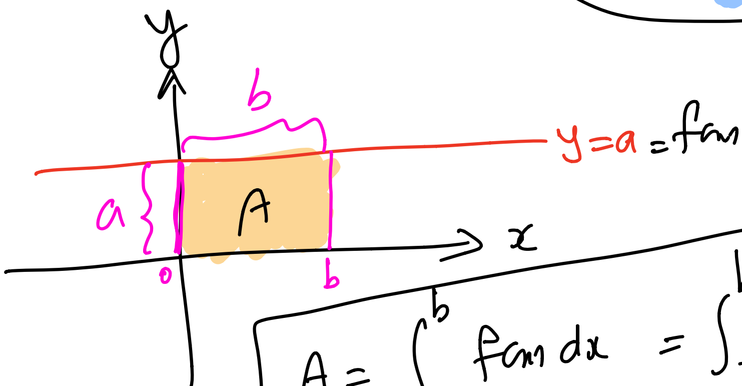
areas



Area = ab



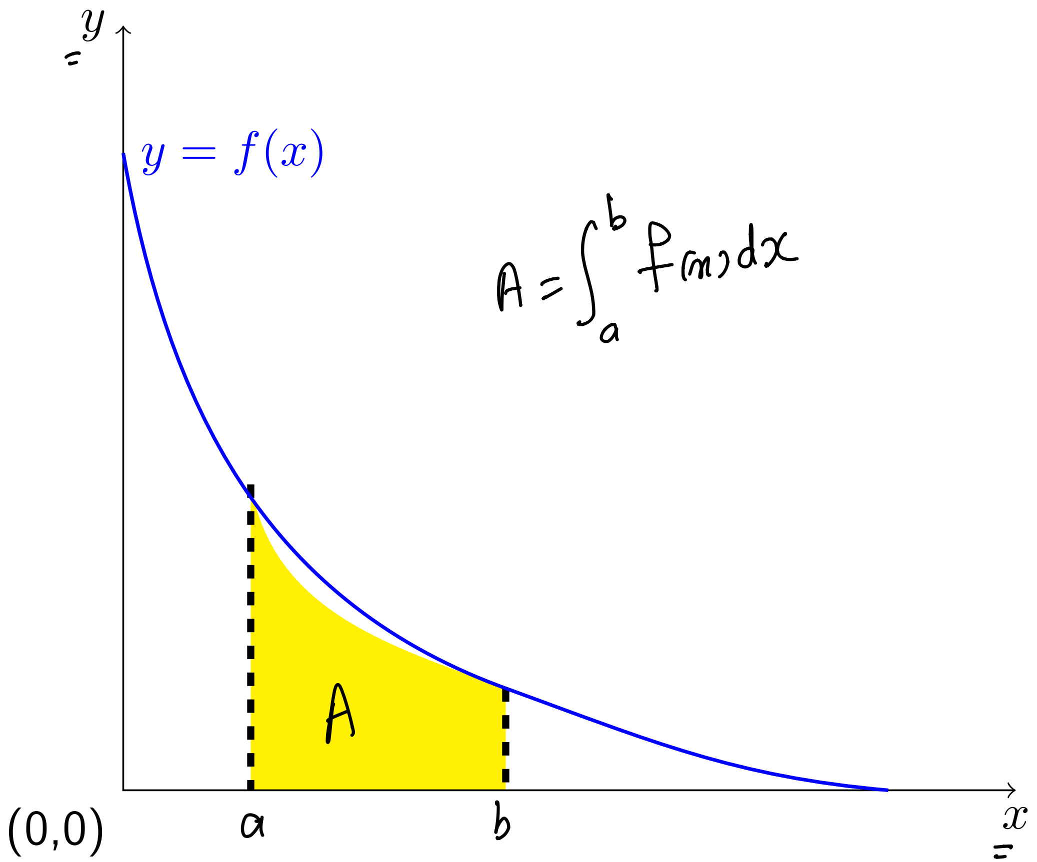
$$\text{Area} = \int_a^b f(x)\,dx$$



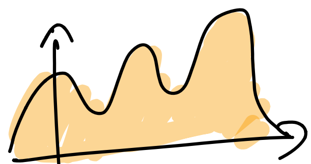
$$A = \int_0^b f(x)\,dx = \int_0^b a\,dx$$

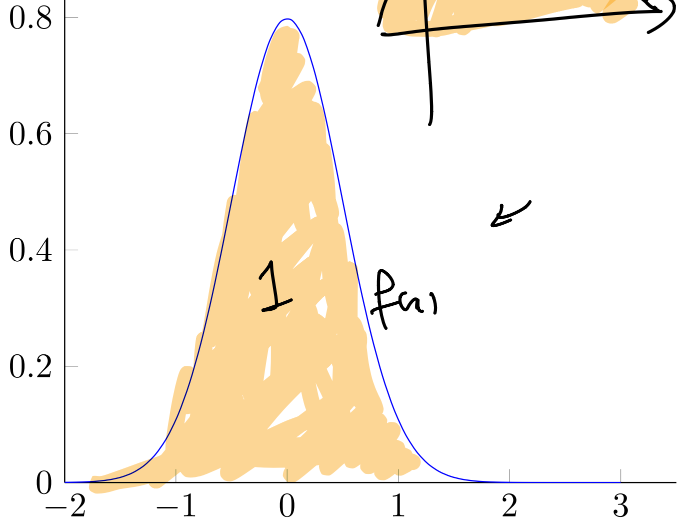
why?

Integrals

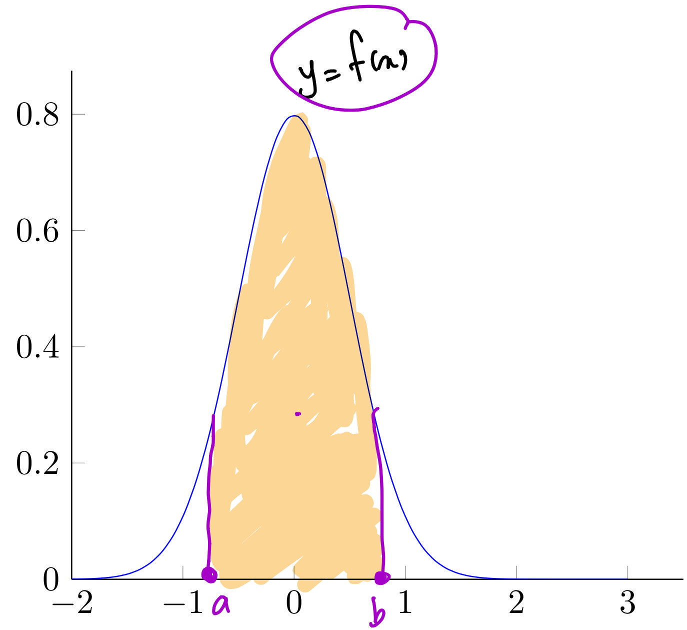


probability distributions





$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$



$$P(a \le X \le b)$$

$$= \int_a^b f(x)\,dx$$

$y = f(x) \xrightarrow{\text{derivative}} y' = f'(x)$

integration

$y = x^2 + 2 \xrightarrow{\text{derivative}} y' = 2x$

integration

- Indefinite Integral: If $F'(x) = f(x)$, then

integrand

antiderivative

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

$F'(x) = f(x)$

$y = x^2 + 2 \rightarrow y' = 2x$

$y = x^2 + 5 \rightarrow y' = 2x$

$y = x^2 + c \rightarrow y' = 2x$

$F(x)$ is called an antiderivetaive of f.

- Definite Integral: $\underbrace{\int_a^b f(x)dx}_{\text{number}} = F(b) - F(a)$

number

Area

Integrals

$\circledast$

$\int 3\,dx = 3x + C$

- $\int f(x)dx = F(x)+C$ means $F'(x) = f(x)$
- $\int cf(x)dx = c\int f(x)dx$
- $\int [f(x)+g(x)]dx = \int f(x)dx + \int g(x)dx$
- $\int k dx = kx + C$
- $\int x^n dx = \frac{x^{n+1}}{n+1} + C, \quad (n \neq -1)$
- $\int \sin(x) dx = -\cos(x) + C$
- $\int \cos(x) dx = \sin(x) + C$

$(-\cos(x)+C)' = \sin x$

Integrals

- $\int \sec^2(x)dx = \tan(x) + C$
- $\int \csc^2(x)dx = -\cot(x) + C$
- $\int \sec(x)\tan(x)dx = \sec(x) + C$
- $\int \csc(x)\cot(x)dx = -\csc(x) + C$
- $\int e^{kx}dx = \frac{1}{k}e^{kx} + C$
- $\int \frac{1}{x}dx = \ln|x| + C$

$x \neq 0$

$\sec(x) = \frac{1}{\cos(x)}$

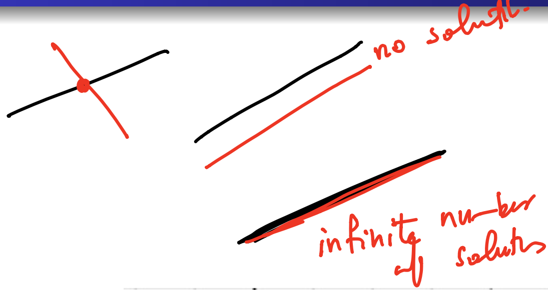
$(\tan(x) + C)' = \sec^2(x)$

$= \frac{1}{\cos^2(x)}$

$\csc(x) = \frac{1}{\sin(x)}$

$y = \ln(x) \quad x > 0$

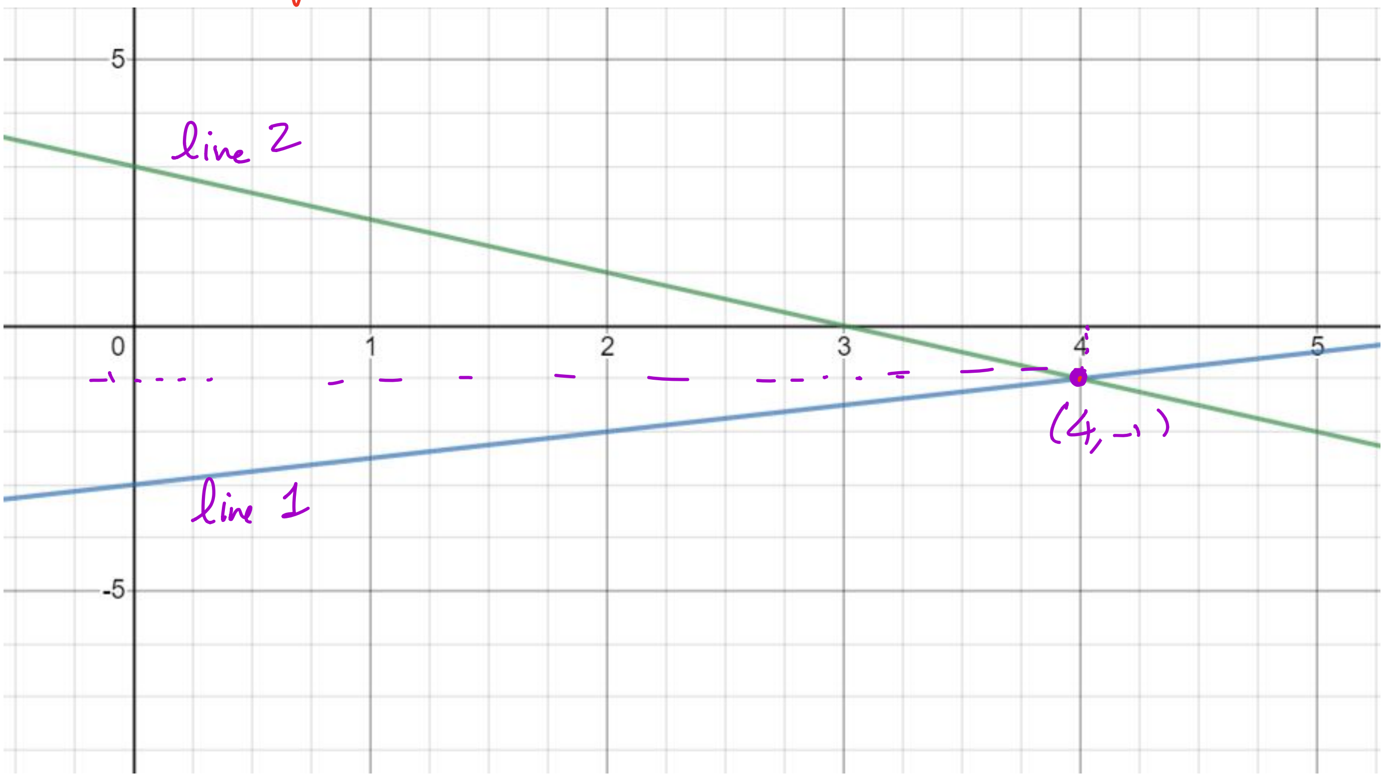
System of Simultaneous Equations



$$\begin{cases} x-2y=6, & \text{line 1} \\ x+y=3, & \text{line 2} \end{cases}$$

unique sol.
no sol
many sols.

$x+y=3$
$y=-x+3$



- Substitution
  1. Make one variable the subject of one of the equations
  2. Substitute for this variable in the other equation
  3. Solve this equation to find the solution for one variable
  4. Substitute the answer found in 3 into the equation obtained in 1 to find the solution for the remaining variable
- solve

$$\begin{cases} x - 2y = 6, & \text{line 1} \\ x + y = 3, & \text{line 2} \end{cases}$$

$x - 2y = 6$

$-2y = 6 - x$

$y = \dfrac{6-x}{-2}$

$x - 2y = 6$

$\boxed{x = 2y + 6}$

$x + y = 3$

$\downarrow$

$\rightarrow 2y + 6 + y = 3$

$3y + 6 = 3$

$3y = -3$

$\boxed{y = -1}$

$x = 2y + 6$

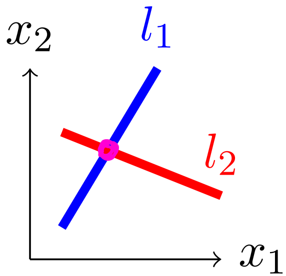
$x = 2(-1) + 6$

$\boxed{x = 4}$

$\boxed{(4, -1)}$

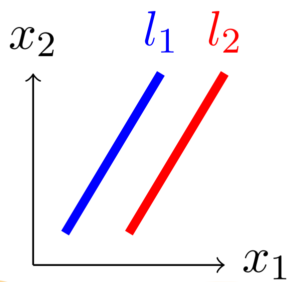
Linear System of two equations

$$\begin{cases} a_1x_1 + b_1x_2 = c_1 \dashrightarrow (l_1) \\ a_2x_1 + b_2x_2 = c_2 \dashrightarrow (l_2) \end{cases}$$



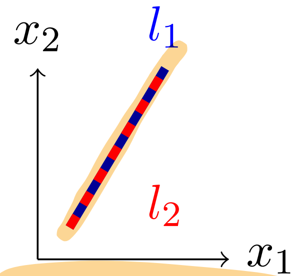
Definite system

single solution



Inconsistent system

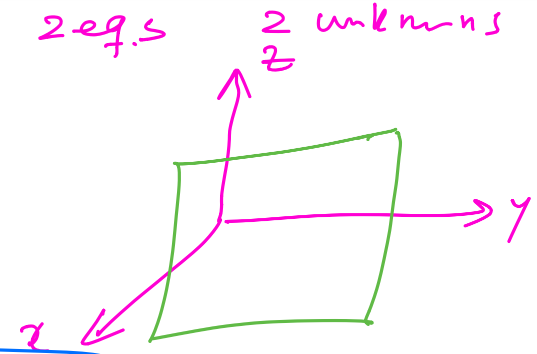
no solution



Indefinite system

many solutions

system 2 eq.s 2 unknowns



- Solve by substitution

$$\begin{cases} x + y + z = 6, & \text{plane 1} \\ x + 2y = 5, & \text{plane 2} \\ 2x + z = 5, & \text{plane 3} \end{cases}$$

(1) $x + y + z = 6 \longrightarrow x = \underline{6 - y - z}$

plug x in the 2^{nd} and 3^{rd} equation

$$\begin{cases} 6 - y - z + 2y = 5 \\ 2(6 - y - z) + z = 5 \end{cases}$$

substitution

$$\begin{cases} y - z = -1 \\ -2y - z = -7 \end{cases}$$

Which system is easier to solve?

- System 1

$$\begin{cases} x + y + z = 6, & \text{plane 1} \\ x + 2y + 0z = 5, & \text{plane 2} \\ 2x + 0y + z = 5, & \text{plane 3} \end{cases}$$

- System 2

$$\begin{cases} x + 0y + 2z = 7, & \text{plane 1} \\ 0x + y - z = -1, & \text{plane 2} \\ 0x + 0y + z = 3, & \text{plane 3} \end{cases}$$

- System 3

$$\begin{cases} x + 0y + 0z = 1, & \text{plane 1} \\ 0x + y + 0z = 2, & \text{plane 2} \\ 0x + 0y + z = 3, & \text{plane 3} \end{cases}$$

$$\begin{cases} x = 1 \\ y = 2 \\ z = 3 \end{cases}$$

0

Useful but harmless operations

row elementary operation

$$\begin{cases} 2x - 4y = 12 \\ 3x + 3y = 9 \end{cases}$$

$$\begin{cases} x - 2y = 6, \\ x + y = 3, \end{cases}$$

×2 ×2 ×3

$$\begin{cases} x + y = 3 \\ x - 2y = 6 \end{cases}$$

- Interchange two equations
- Multiply each element in an equation by a non-zero number
- Multiply an equation by a non-zero number and add the result to another equation.

$$\begin{array}{r} 2x - 4y = 12 \\ x + y = 3 \\ \hline 3x - 3y = 15 \end{array}$$

$$\begin{cases} x - 2y = 6 \\ 3x - 3y = 15 \end{cases}$$

$$\begin{cases} y = f(x) \\ y' = f'(x) \\ \int f(x)\,dx \end{cases}$$

algebra

prior knowledge

search

Assignment 1