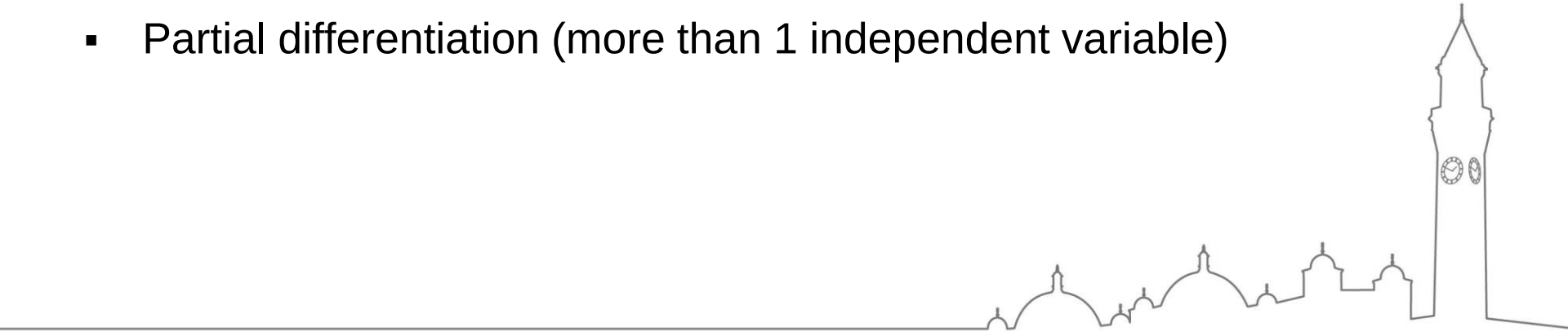
UNIVERSITY OF BIRMINGHAM

Week 1. Differentiation

Dr. Shuo Wang

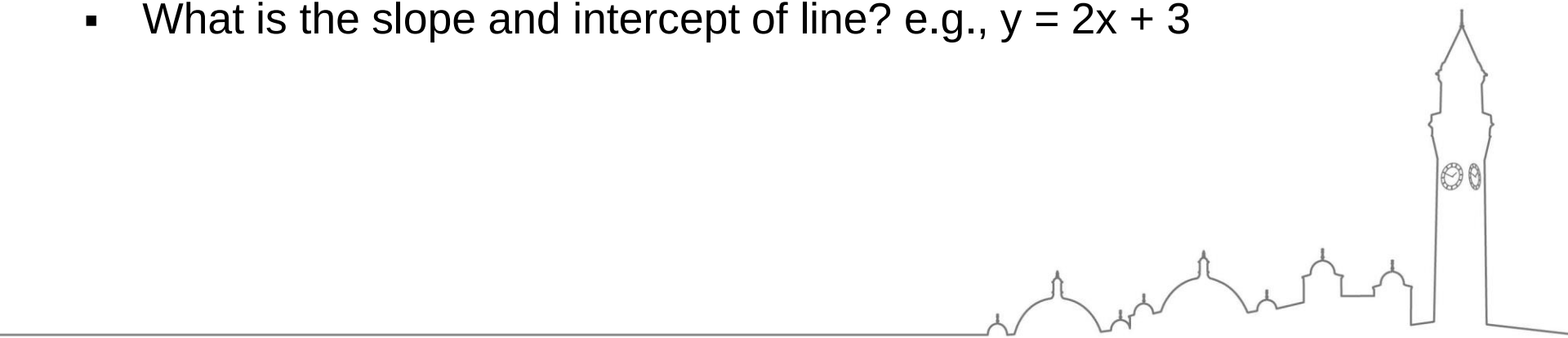
Overview

- Differentiation plays an important role in machine learning/AI
- An important method for machines to learn a function
- Univariate differentiation
- Some rules
- Partial differentiation (more than 1 independent variable)

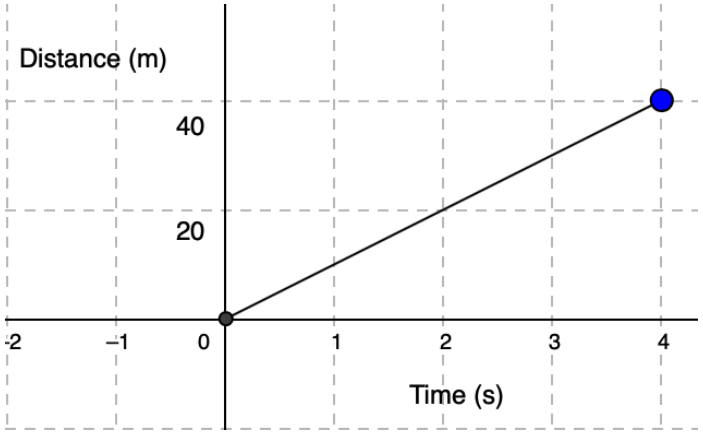


Some Basic Mathematics

- In $y=f(x)$, what are the dependent and independent variables?
- In $z=f(x, y)$, what are the dependent and independent variables?
- What is the equation of a line?
- What is the slope and intercept of line? e.g., y = 2x + 3



Rate of Change (Gradient) of a Straight Line



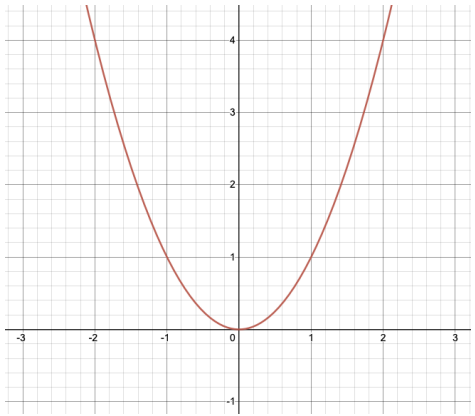
- Gradient $= \frac{\Delta y}{\Delta x}$
- A car travels 40m over 4s. Speed?
- Gradient = speed/slope of the line = rate of change of distance
- $\Delta x, \Delta y$: change of x and y

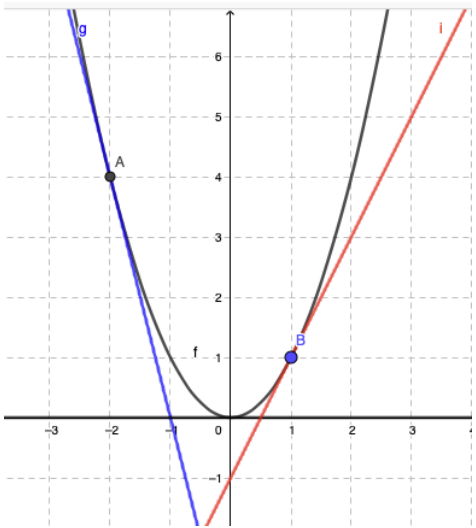
Gradient of the straight line: y = 10x
For a straight line, gradient is constant: same at every point.

Differentiation

The process of finding the rate at which one variable changes with respect to another (i.e. the gradient).

- $\Delta x, \Delta y$ represent a change in the value of x and y
- What if our function is a curve, instead of a straight line?
- What is the slope of a curve?

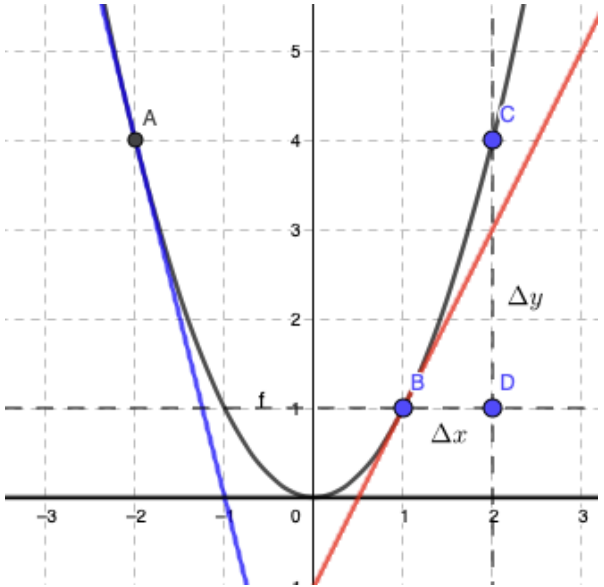




UNIVERSITY OF BIRMINGHAM

Gradient at a Point, Gradient/Derived Function, Derivative

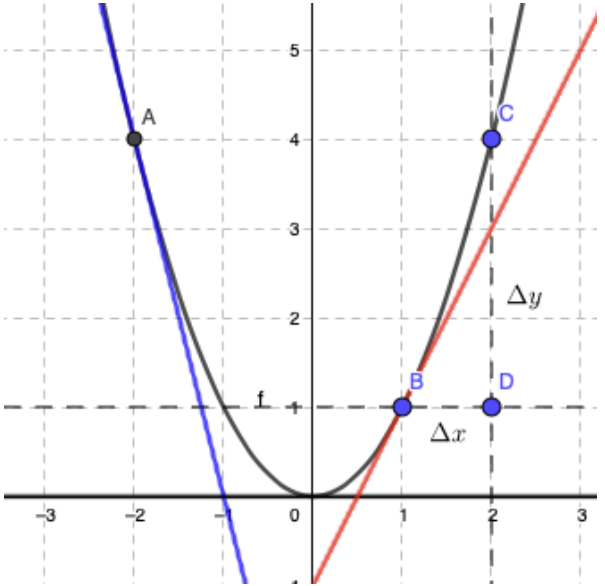
- The gradient at a point is given by the gradient of the tangent* at that point.
- The gradient of the tangent to a curve (non-linear) function $y = f(x)$ varies with variable x. Therefore, it is also a function of x.
- It is called gradient function or derived function.
- Let's see how to obtain the general gradient function of $y = x^2$.



* The tangent to a curve at a given point is a straight line which "just touches" the curve at that point

Differentiation from first principles

- As point C moves closer to B, the gradient of the line BC gets closer to the gradient at B.
- Consider the limit as Δx tends to 0.
- This process called differentiation from first principles.



UNIVERSITY OF BIRMINGHAM

Gradient/Derived Function, Derivative

- Both $f'(x)$ and $\frac{dy}{dx}$ mean the gradient function.
- Also known as the derivative of y with respect to x.

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

Rules of Differentiation

- Differentiation of Monomials $y = ax^n$
- In general:

For $f(x) = x^n$, $f'(x) = nx^{n-1}$
For $f(x) = ax^n$, $f'(x) = anx^{n-1}$

$f(x)$	c	x	x^2	x^3	x^4	x^5
$f'(x)$	0	1	$2x$	$3x^2$	$4x^3$	$5x^4$

Differentiation of Multiple Terms - Polynomials

- A polynomial function: $y = x^3 + 6x^2 - 3x + 1$
- How to differentiate this function with respect to x?

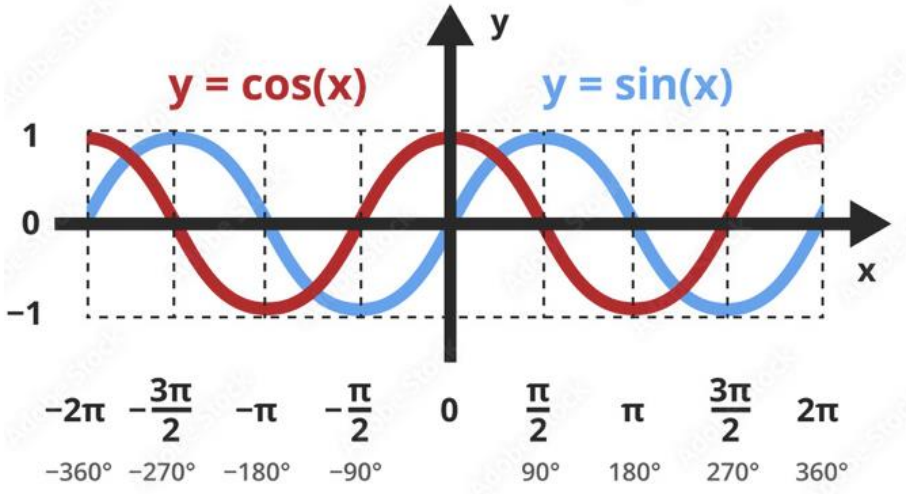
- General rule for sums of functions:

If $y = f(x) \pm g(x)$, $\frac{dy}{dx} = f'(x) \pm g'(x)$

Other Derivatives

- Trigonometric functions: sine and cosine

If $f(x) = \sin x$, $f'(x) = \cos x$
If $f(x) = \cos x$, $f'(x) = -\sin x$



UNIVERSITY OF BIRMINGHAM

Other Derivatives

- Natural exponential

$$\text{If } f(x) = e^x,\ f'(x) = e^x$$

- Natural logarithm (the inverse of the natural exponential)

$$\text{If } f(x) = \ln x \ \ (x > 0),\ f'(x) = \frac{1}{x}$$

The Product Rule

If $y = f(x)g(x)$, $\frac{dy}{dx} = f(x)g'(x) + g(x)f'(x)$

- Example: $y = x^2 \cos x$

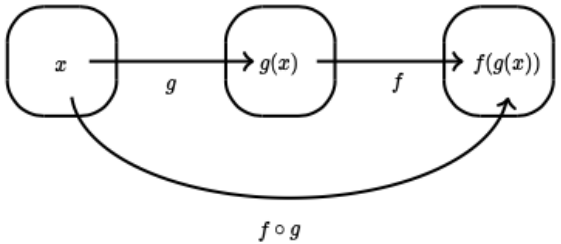
The Quotient Rule

$$\text{If } y = \frac{f(x)}{g(x)}, \frac{dy}{dx} = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2}$$

- Example: $y = \frac{2x+1}{x^2+2x+1}$

The Rules – The Chain Rule

- Allows us to differentiate a composite function, i.e. a function within a function.
- Composite function:



- How to differentiate it:

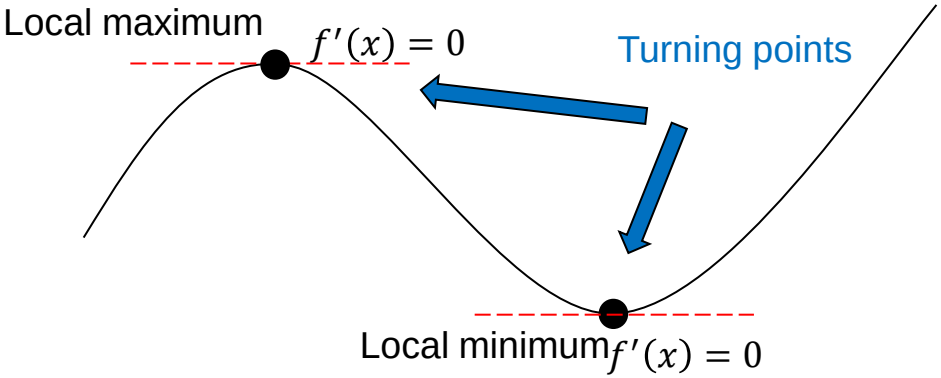
$$\text{If } y = f(g(x)) \text{ , } \frac{dy}{dx} = f'\big(g(x)\big)g'(x)$$

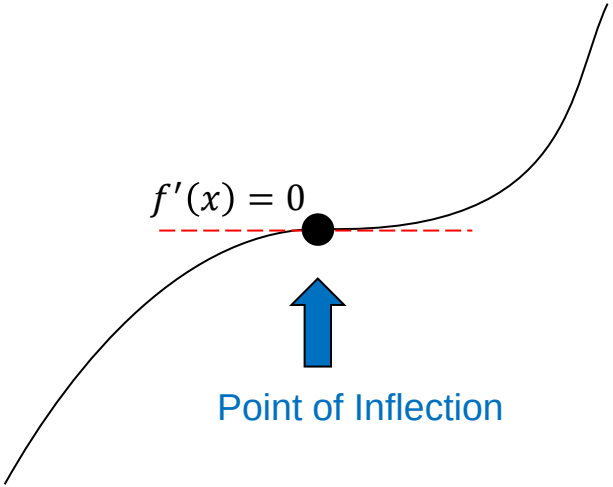
  Outer function differentiated × inner function differentiated
- Example: $y = e^{3x}$

Application of Derivatives– Find out Local Max and Min

A stationary point is where the gradient is 0, i.e. $f'(x_0) = 0$

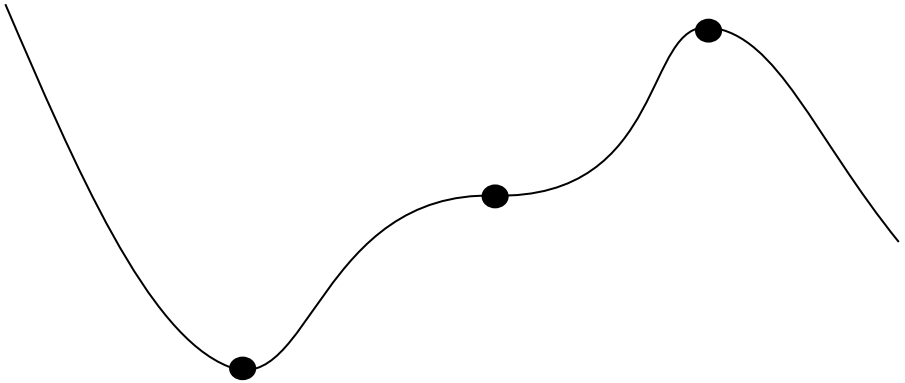
- Stationary point and its 3 types:





How to determine type of stationary point?

- Look at the gradient just before and after the point



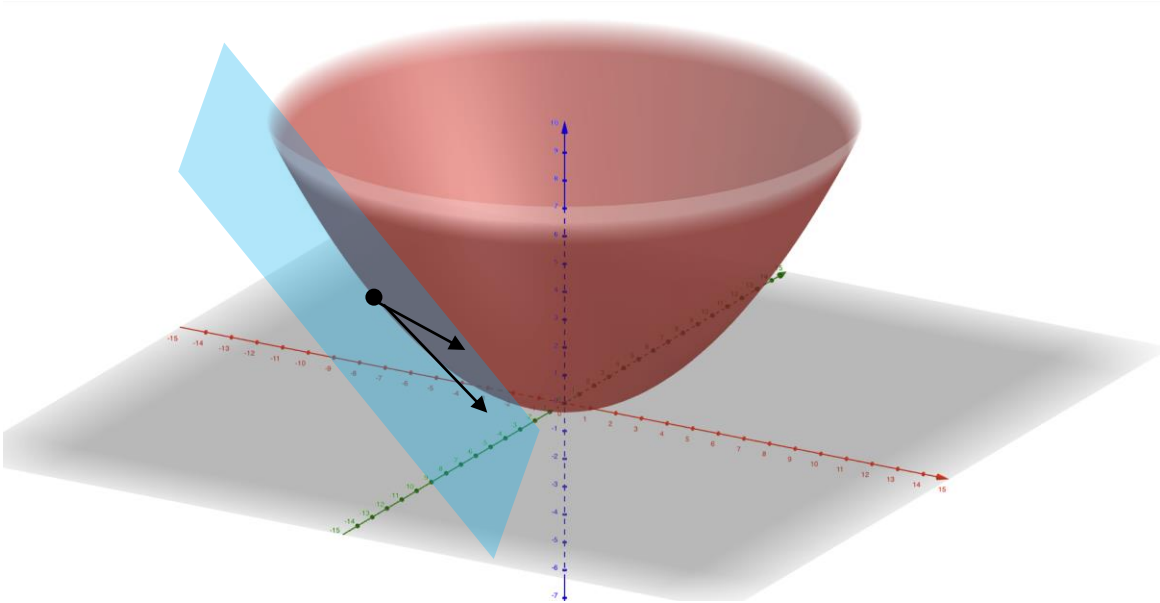
Local Maximum		
Gradient just before	Gradient at max	Gradient just after
?	?	?

Local Minimum		
Gradient just before	Gradient at min	Gradient just after
?	?	?

Point of Inflection		
Gradient just before	Gradient p.o.i	Gradient just after
?	?	?

Multivariable Function and Partial Differentiation

- When a function has more than one independent variable
  e.g. $z = x^2/10 + y^2/10$, or $f(x, y) = x^2/10 + y^2/10$
  3 dimensions, x and y are independent variables and z is the dependent variable.
- In 3D, a tangent line becomes a tangent plane.



- Directional derivative
- Partial derivative

UNIVERSITY OF BIRMINGHAM

Partial Differentiation

Notations: for $z = f(x, y)$, we write:

- $f_x(x, y) = \frac{\partial f}{\partial x} = \frac{\partial z}{\partial x}$, the partial derivative of f with respect to x
- $f_y(x, y) = \frac{\partial f}{\partial y} = \frac{\partial z}{\partial y}$, the partial derivative of f with respect to y

Rule:

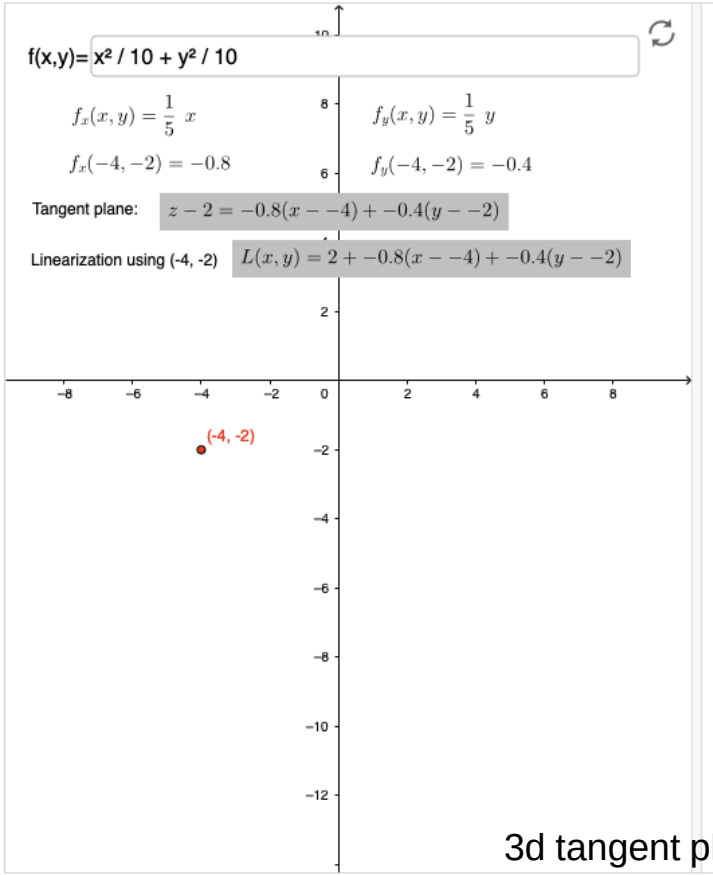
The partial derivative with respect to x is the ordinary derivative of the function of x by treating the other variables as constants.

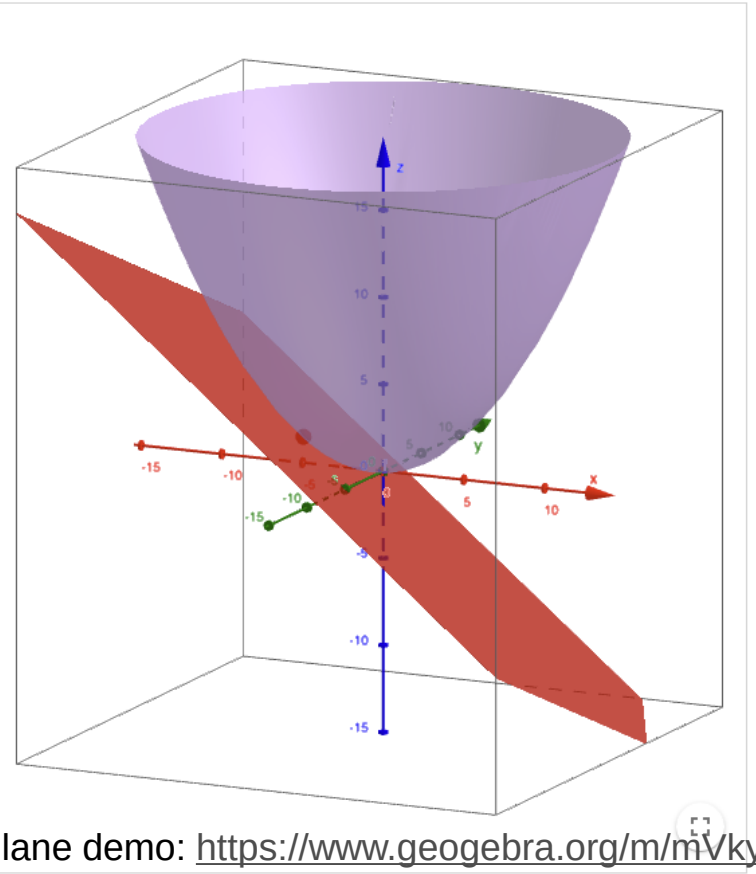
- To find f_x, treat y as a constant and differentiate $f(x, y)$ with respect to x.
- To find f_y, treat x as a constant and differentiate $f(x, y)$ with respect to y.

UNIVERSITY OF BIRMINGHAM

Example: $f(x, y) = x^2y + 2x$

Try yourself: Tangent plane of $f(x, y) = x^2/10 + y^2/10$ at (-4, -2)





3d tangent plane demo: https://www.geogebra.org/m/mVky4Jkw

Summary

- We have studied: slope of a line: y=mx+c
- Tangent to a curve
- The derivative function of differentiation $f'(x)$
- Some rules of differentiation
- We can visualise the functions/derivatives graphs
- Differentiation plays an important role in machine learning