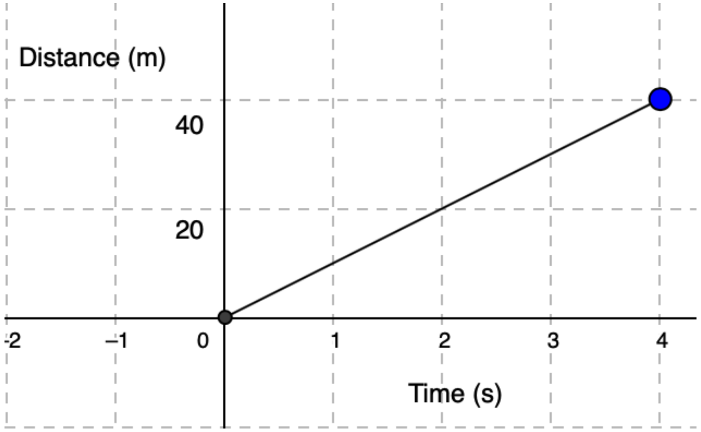
Week 1. Differentiation and some linear algebra

Dr. Shuo Wang

Overview

- Univariate differentiation
- Some rules
- Partial differentiation (more than 1 independent variable)

Rate of Change (Gradient) of a Straight Line

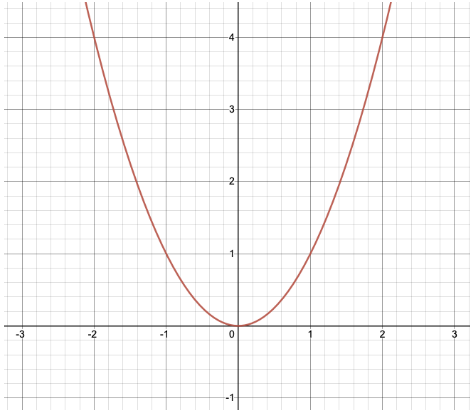


- A car travels 40m over 4s. Speed?
- Gradient (derivative)= speed/slope = rate of distance change, steepness
- $\Delta x, \Delta y$: change of x and y
- Gradient of a straight line: y = 10x Constant gradient, same at every point.

Differentiation

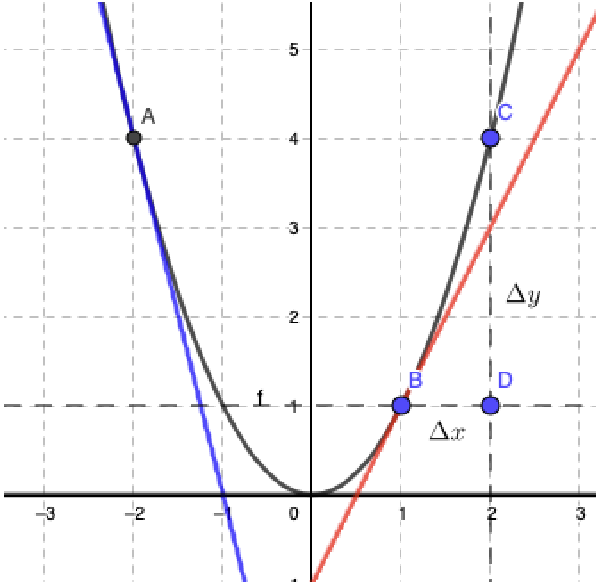
The process of finding the rate at which one variable changes with respect to another (i.e. the gradient).

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



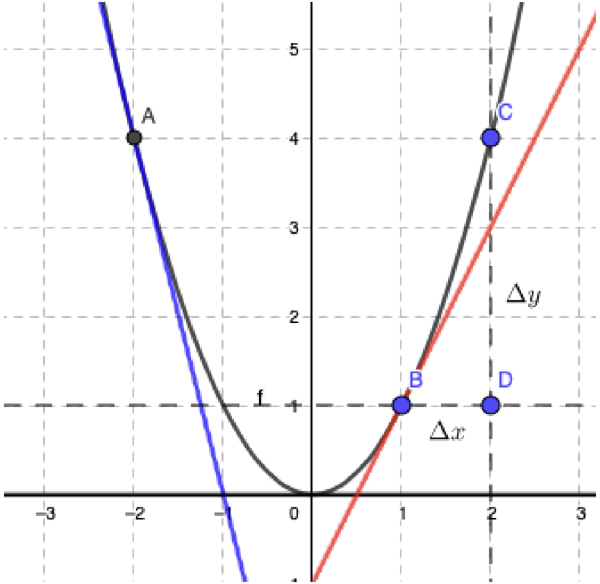
Gradient at a Point, Differentiation from first principles

- The gradient at a point is given by the gradient of the tangent at that point.
- 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.
- It gives you **the direction of the steepest uphill (aka. largest increase).**



Gradient/Derived Function, Derivative

- 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.



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.

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

- Practice: how to obtain the general gradient function of $y = x^2$.

Differentiation of Monomials $y = ax^n$

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

- What Pattern do you notice?
- In general:

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

Differentiation of Multiple Terms - Polynomials

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

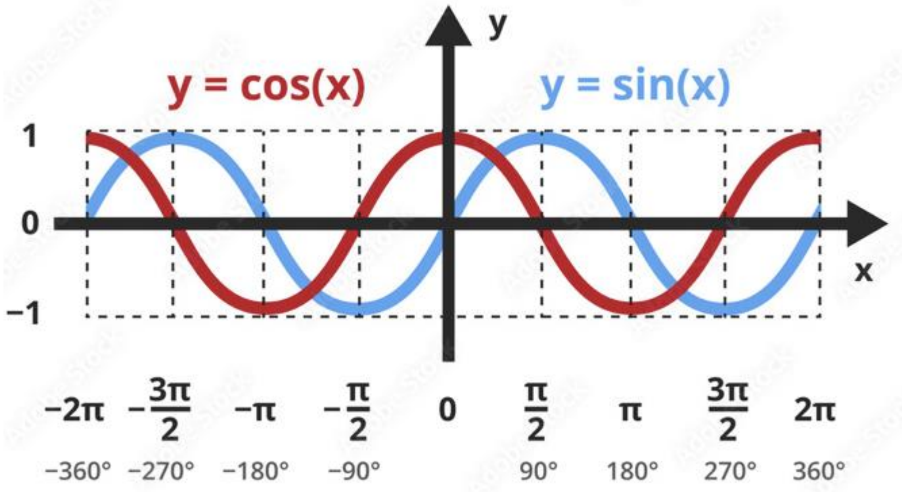
- General rule for sums of functions:

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

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

If $f(x) = e^x$, $f'(x) = e^x$

- Natural logarithm (the inverse of the natural exponential)

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

The Rules – 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 Rules – The Quotient Rule

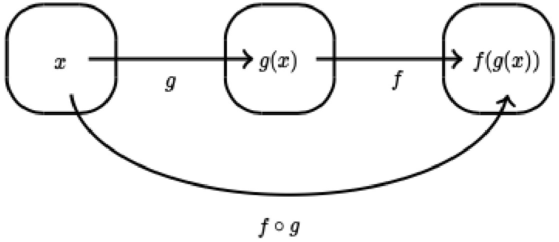
If y $= \frac{f(x)}{g(x)}$, $\frac{dy}{dx} = \frac{g(x)f'(x) - f(x)g\prime(x)}{g(x)^2}$

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

UNIVERSITY OF BIRMINGHAM

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:

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

Outer function differentiated × inner function differentiated

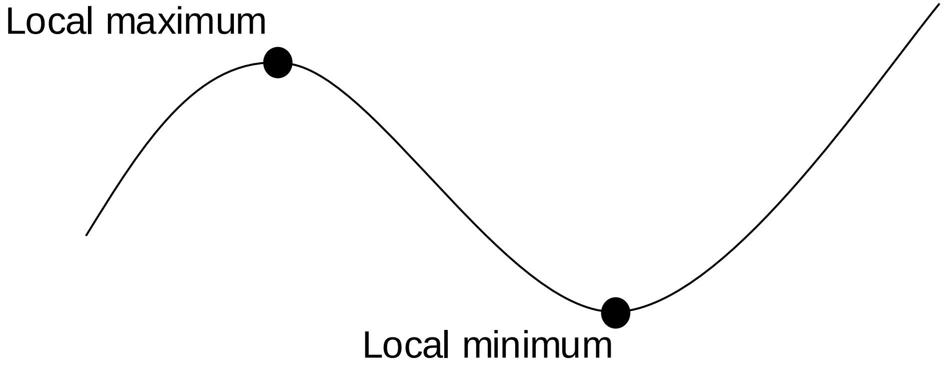
- Example: $y = e^{3x}$

UNIVERSITY OF BIRMINGHAM

Application of Derivatives– Find out Local Max and Min

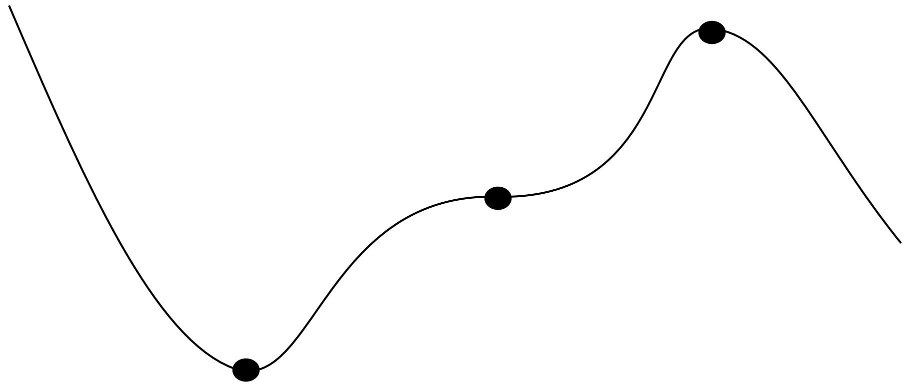
- Stationary point

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



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
?	?	?

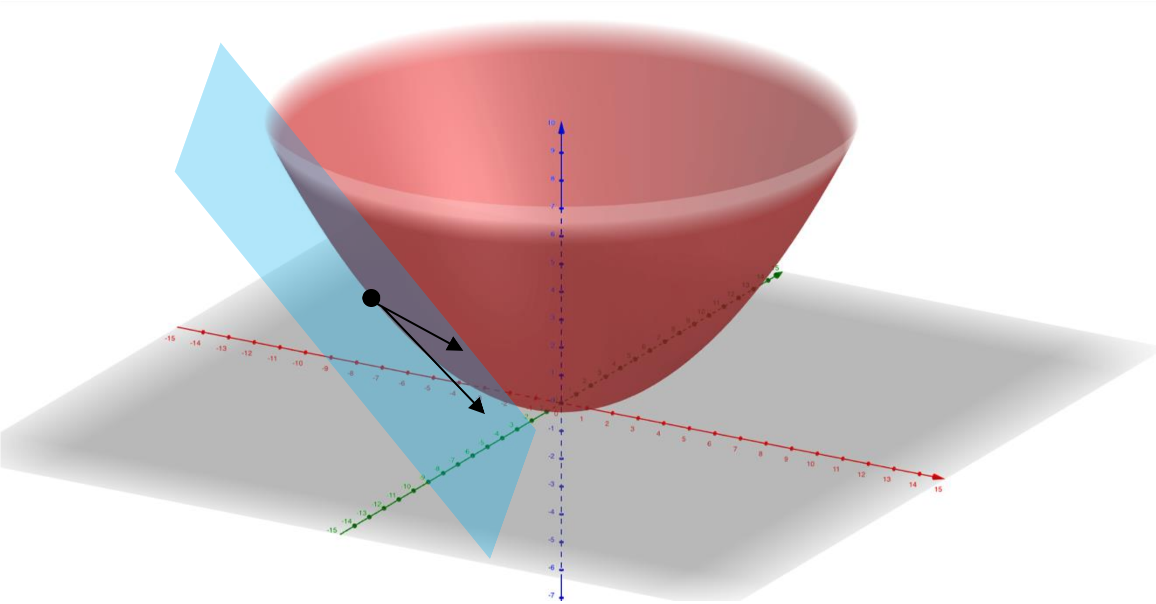
Local Minimum		
Gradient just before	Gradient at min	Gradient just after
?	?	?

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

Multivariate 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

<u>Notations</u>: for $\mathrm{z} = f(x, y)$, 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

<u>Rule</u>:

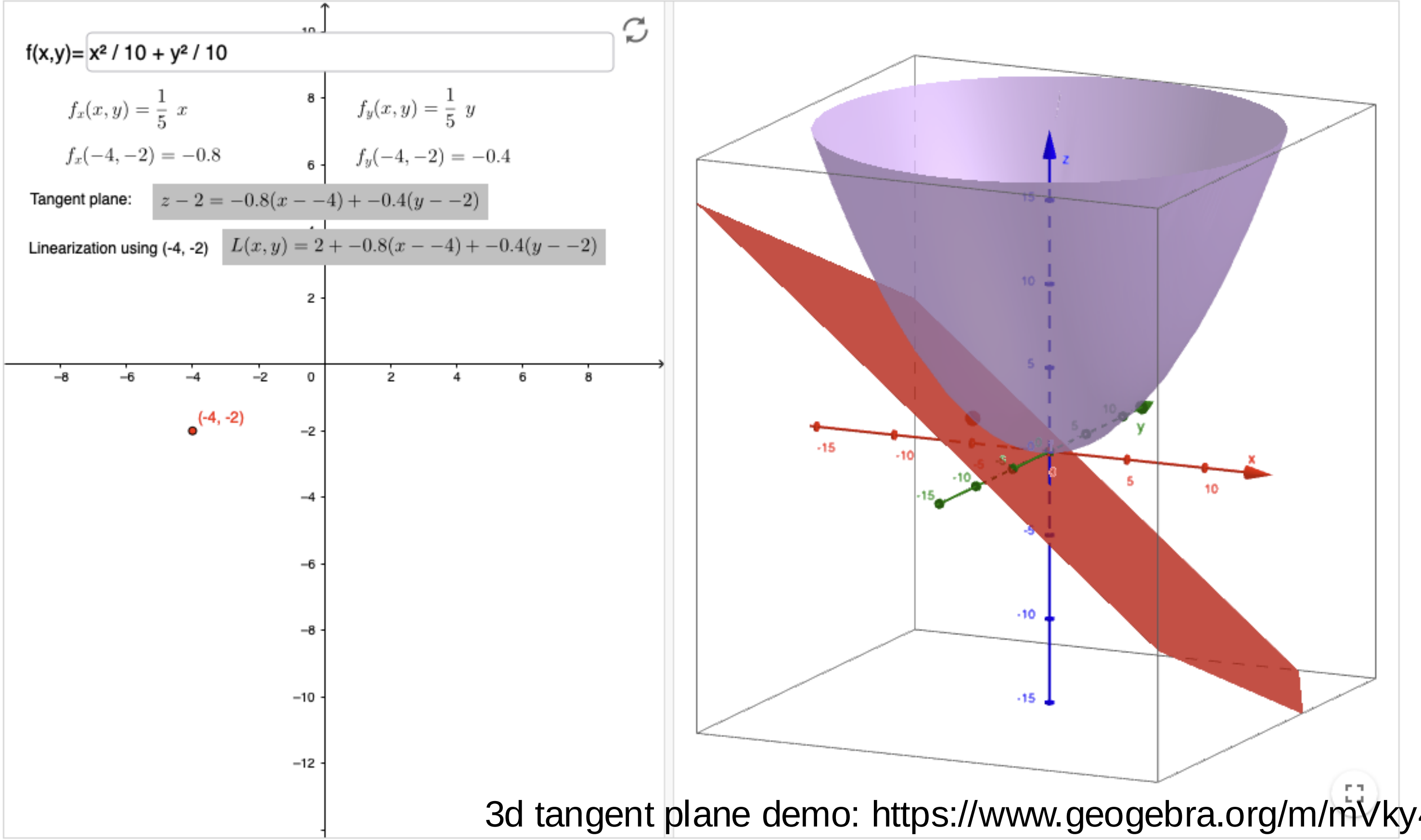
The partial derivative with respect to x is the <u>ordinary</u> derivative of the function of x by treating the other variables as <u>constants</u>.

- 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

Linear functions and the vector notation

- Univariate

$$y = w_0 + w_1 x$$

a line

vector $\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} 1 \\ x \end{pmatrix}$

$$y = \boldsymbol{w}^T \boldsymbol{x}$$

- Multivariate

$$y = w_0 + w_1 x_1 + w_2 x_2 .. + w_n x_n$$

a hyperplane

vector $\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \dots \\ w_n \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$

$$y = \boldsymbol{w}^T \boldsymbol{x}$$

Revise dot product of a row and a column vectors

Polynomial functions and the vector notation

- Contain multiple terms with only variables of varying degrees, coefficients, non-negative integer exponents, and constants.

- Polynomial function with one variable:

$$\boxed{y = w_0 + w_1 x + w_2 x^2 .. + w_n x^n}$$

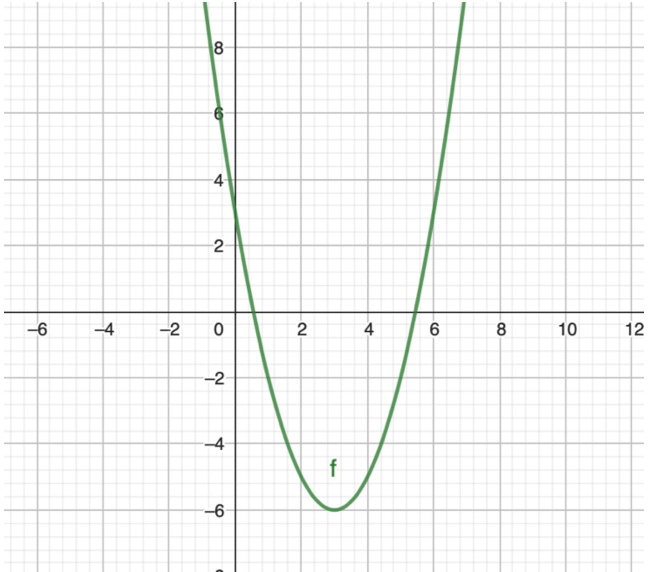
- vector $\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \dots \\ w_n \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} 1 \\ x \\ x^2 \\ \dots \\ x^n \end{pmatrix}$

$$\boxed{y = \boldsymbol{w}^T \boldsymbol{x}}$$

- *y is not a linear function of* $\boldsymbol{x}$*, but a linear function of* $\boldsymbol{w}$

Polynomial function examples

- $y = 3 - 6x + x^2$
- $y = 2 - 3x + 4x^3$
- $y = -1 - 6\sqrt{x} + 4x^2$ (this is not a polynomial function)



Go visualize these functions
https://www.geogebra.org/calculator

Sigmoid (or logistic) function

$$y = \frac{1}{1 + e^{-x}}$$

- Very important function used in machine learning algorithms.
- Homework:
  explore its properties: e.g. shape, output range, symmetry, differentiability.

Q/A

Teams Channel for Week1
Tutorials and Drop-in Sessions from Week2
See Canvas module homepage