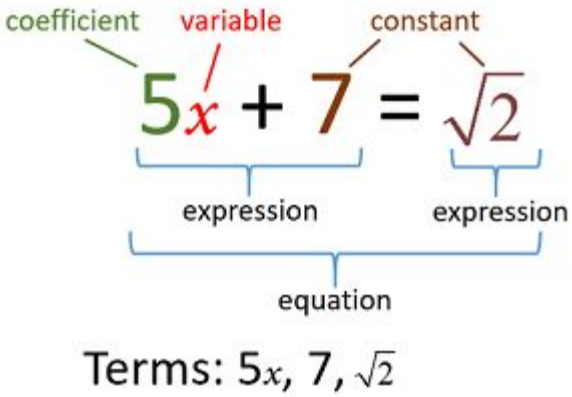
Math for MLCC

by Pratik Parmar

Algebra

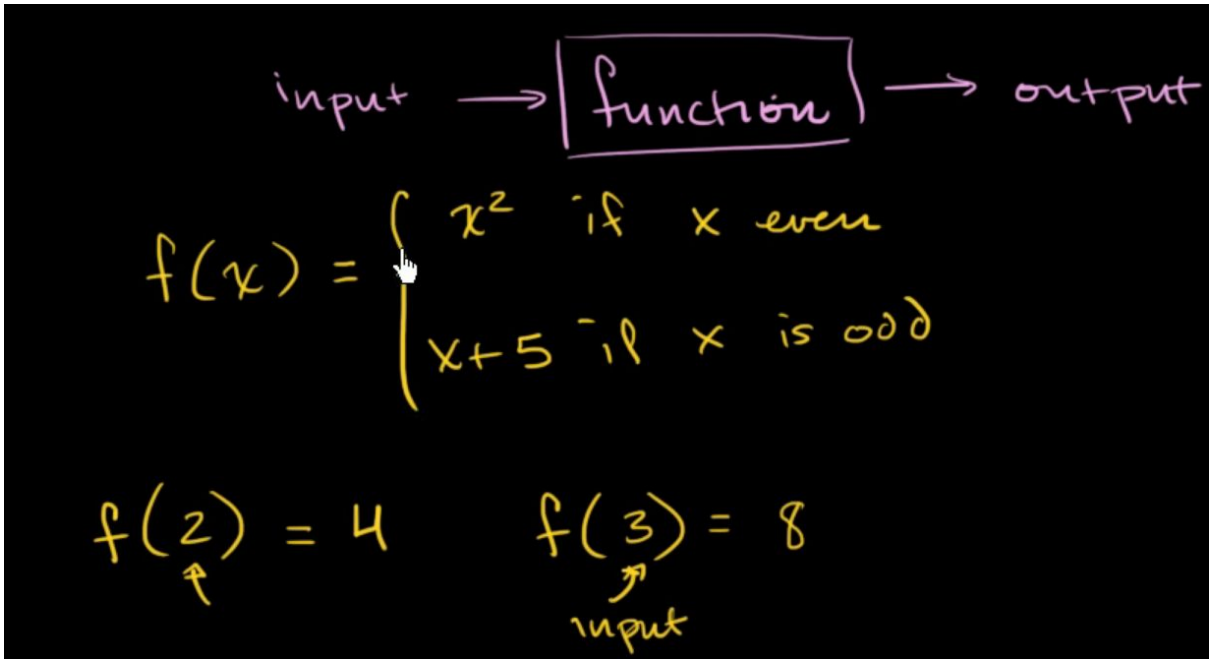
- **Variables**: A variable is a quantity that may change within the context of a mathematical problem or experiment. (restaurant tip example)
- **Coefficients**: A number used to multiply a variable.



Source: [Online Math Learning]

Algebra

- **Functions**: Functions is something that takes input and generates output based on that input.



Source: [Khan Academy]

Algebra

- **Linear Equations**: A **linear** equation is an equation for a straight **line**

  Example: $y = b + w_1x_1 + w_2x_2$

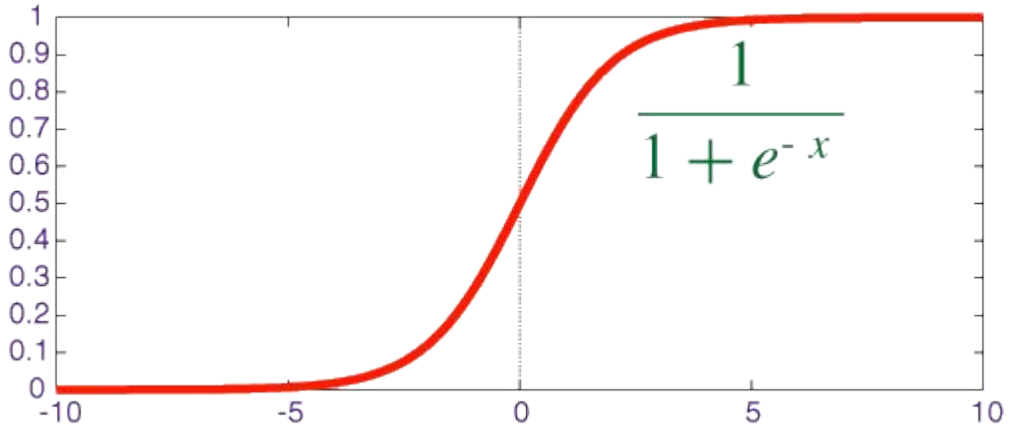
- **Logarithms**: In its simplest form, a logarithm answers the question:
  - How many of *one number* do we multiply to get *another number?*

    Example: How many 2s do we multiply to get 8?
    Answer: **2 × 2 × 2 = 8**, so we had to multiply 3 of the **2**s to get **8**
    So the logarithm is 3

Algebra

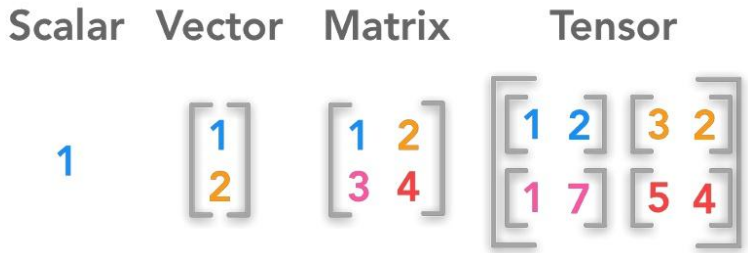
- **Sigmoid Function:** A sigmoid function is a mathematical function having a characteristic "S"-shaped curve or sigmoid curve. It takes any number as an input and squashes it into a range between 0 & 1.

  Often, *sigmoid function* refers to the special case of the logistic function.



Linear Algebra

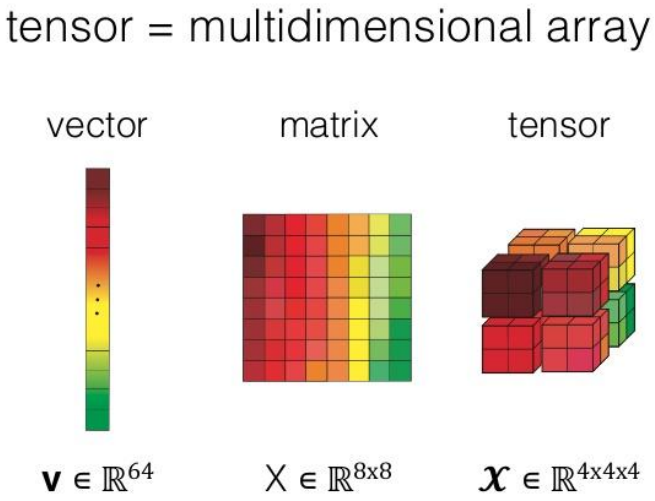
- **Tensor and Tensor rank**: A **tensor** is often thought of as a generalized matrix. That is, it could be a 1-D matrix (a vector is actually such a tensor), a 3-D matrix (something like a cube of numbers), even a 0-D matrix (a single number), or a higher dimensional structure that is harder to visualize. The dimension of the tensor is called its *rank*.



Source: hadrienj

Linear Algebra

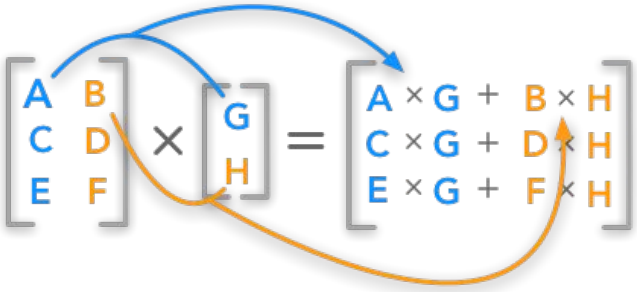
- **Vector vs Matrix vs Tensor**



Source: SlideShare

Linear Algebra

- **Matrix Multiplication**
  To know more: https://www.mathsisfun.com/algebra/matrix-multiplying.html



$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

A B C

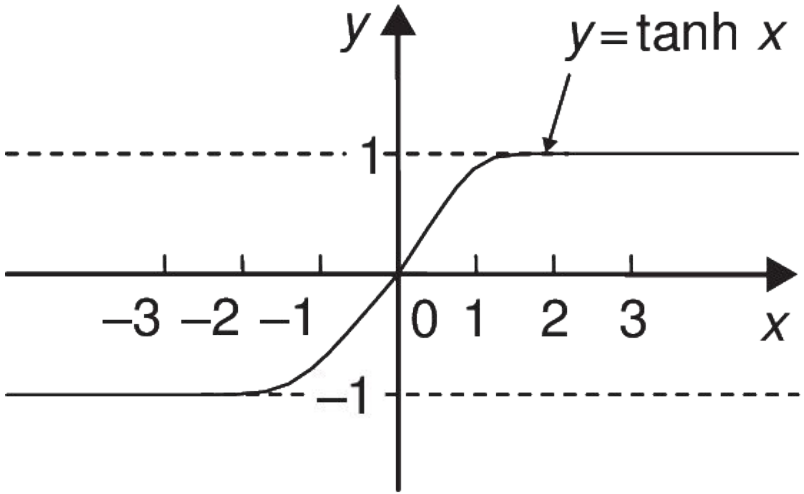
A, B and C are square metrices of size N x N
a, b, c and d are submatrices of A, of size N/2 x N/2
e, f, g and h are submatrices of B, of size N/2 x N/2

Source: GeeksforGeeks, hadrienj

Linear Algebra

- **Hyperbolic function**
  - **Tanh(z)**: gives the hyperbolic tangent of z.

    To know more: https://www.mathsisfun.com/sets/function-hyperbolic.html



Statistics

- **Mean:** Mean is said to be the average value of all the numbers. It can sometimes be used as a representation of the whole data.

  For instance, if you have the marks of students from a class, and you asked about how good is the class performing. It would be irrelevant to say the marks of every single student, instead, you can find the mean of the class, which will be a representative for class performance.

  To find the mean, sum all the numbers and then divide by the number of items in the set.

$$\begin{aligned}\text{Mean} &= \frac{\text{Sum of all data values}}{\text{Number of data values}} \\ &= \frac{15+13+18+16+14+17+12}{7} \\ &= \frac{105}{7} \\ &= 15\end{aligned}$$

Source: [Mathsteacher]

Statistics

- **Median:** The Median is the *"middle"* of a sorted list of numbers.
  When the total numbers in the set are even, the median will be the average of the two middle values.
  Median is used to measure the central tendency.

  To calculate the median for a set of numbers, follow the below steps:
  - Arrange the numbers in ascending or descending order
  - Find the middle value, which will be n/2 (where n is the numbers in the set)

1, 3, 3, **6**, 7, 8, 9

Median = 6

1, 2, 3, **4**, **5**, 6, 8, 9

Median = (4 + 5) ÷ 2

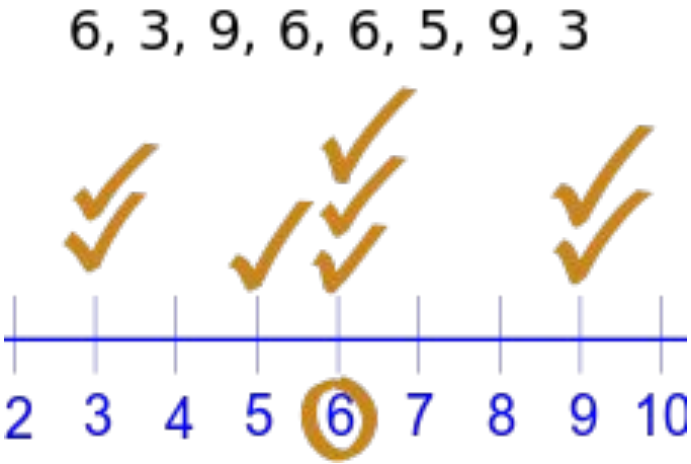
= 4.5

Source: Wikipedia

Statistics

- **Mode:** The number which appears most often in a set of numbers/data.

  Example: in {6, 3, 9, 6, 6, 5, 9, 3} the Mode is 6 (it occurs most often).



- Learn more about Mean, Mode and Median: https://youtu.be/k3aKKasOmIw

Source: Math is Fun

Statistics

- **Outliers:** "Outliers" are values that "lie outside" the other values.
    - Extra: calculate the median and mode when you have outliers.

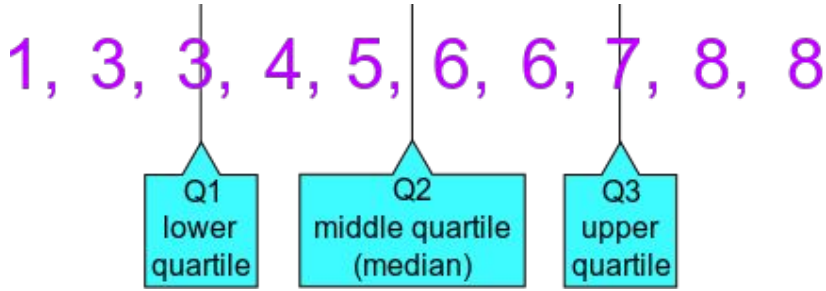
- Learn more : https://www.mathsisfun.com/data/outliers.html
  https://youtu.be/FRlTh5HQORA

Statistics

- **Quartiles:** Quartiles are the values that divide a list of numbers into quarters.

  To find quartiles:
  - Put the list of numbers **in order**
  - Then cut the list into **four equal parts**
  - The Quartiles are at the "cuts"

  Example: 1, 3, 3, 4, 5, 6, 6, 7, 8, 8 →



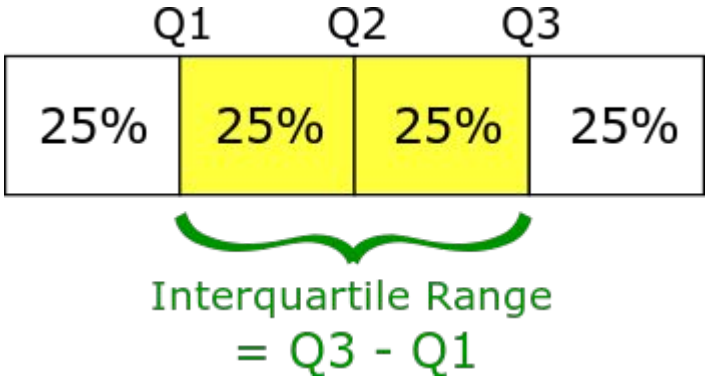
  - The numbers are already in order
  - Cut the list into quarters

  - In this case Quartile 2 is halfway between 5 and 6:
    Q2 = (5+6)/2 = 5.5

  - And the result is:
    - Quartile 1 (Q1) = **3**
    - Quartile 2 (Q2) = **5.5**
    - Quartile 3 (Q3) = **7**

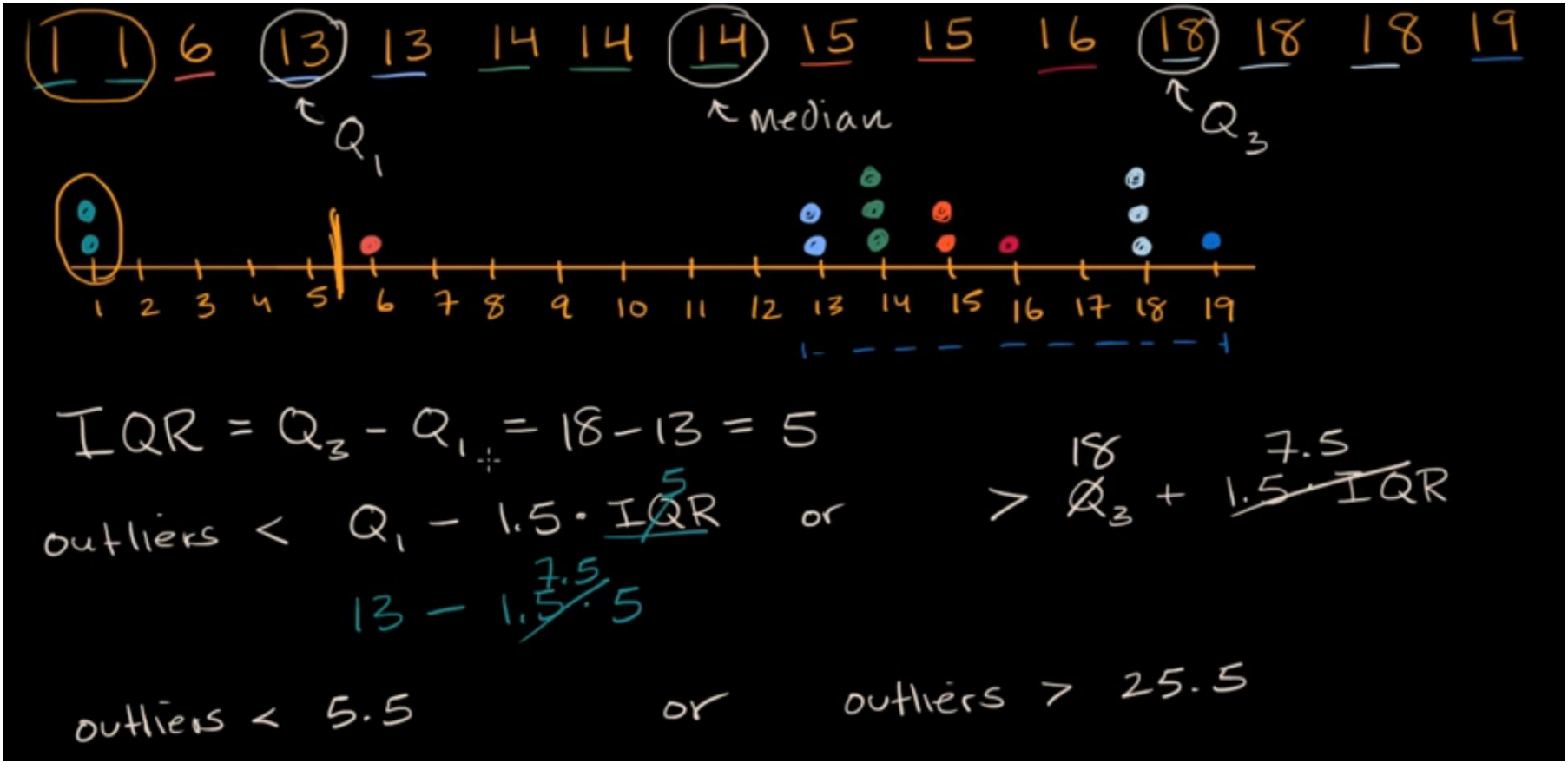
Statistics

- **Inter Quartile Range (IQR) :** The "Interquartile Range" is from Q1 to Q3:



- To calculate it just **subtract Quartile 1 from Quartile 3**, continuing our last example:
  Q3 - Q1 = 7 - 3 = 4

- Outlier example:



Source: [Khan Academy]

Statistics

- **Standard Deviation :**
    - Deviation just means how far from the normal and The Standard Deviation is a measure of how spread out numbers are.

      The formula is easy: it is the **square root** of the **Variance.**

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

But what is the **Variance**?

Statistics

- **Standard Deviation :**
    - Before you ask “What is the variance?”
      The Variance is defined as the average of the **squared** differences from the Mean.
    - To calculate the variance follow these steps:
      1.) Work out the Mean (the simple average of the numbers)
      2.) Then for each number: subtract the Mean and square the result (the *squared difference*).
      3.) Then work out the average of those squared differences. (Why Square?)

- Learn more : https://www.mathsisfun.com/data/standard-deviation.htm
- https://www.youtube.com/watch?time_continue=365&v=E4HAYd0QnRc

Statistics

- **Example of Standard Deviation : Find standard deviation of 1 2 3 8 7**

$$\mu(mean) = \frac{1+2+3+7+8}{5} = 4.2$$

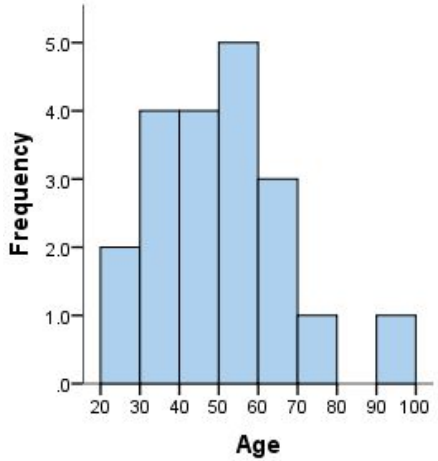
$$\sigma^2(variance) = \frac{(1-4.2)^2 + (2-4.2)^2 + (3-4.2)^2 + (7-4.2)^2 + (8-4.2)^2}{5} = \frac{38.5}{5} = 7.76$$

$$\sigma(standard\,deviation) = \sqrt{\sigma^2} = \sqrt{7.76} = 2.79$$

Statistics

- **Histogram**: A graphical display where the data is grouped into ranges (such as "100 to 149", "150 to 199", etc), and then plotted as bars.

  Similar to a Bar Graph, but in a Histogram each bar is for a range of data.



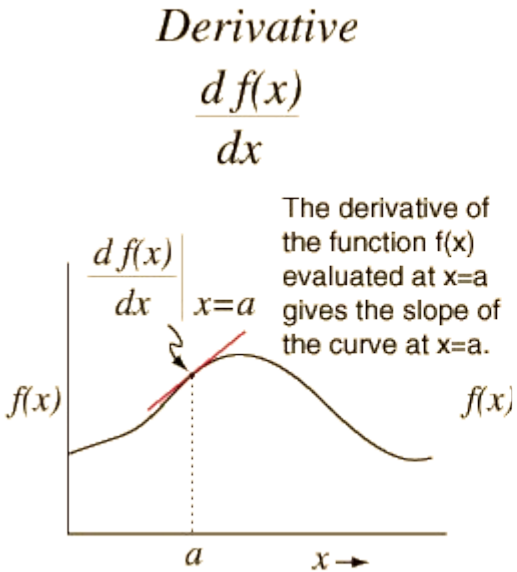
- Learn more : https://www.mathsisfun.com/data/histograms.html

Source: Laerd Statistics

Calculus (optional, for advanced topics)

- **Derivative**: The Derivative tells us the slope of a function at any point.

  NOTE: You only have to be familiar with concept of Derivative, you won't have to actually calculate derivatives.

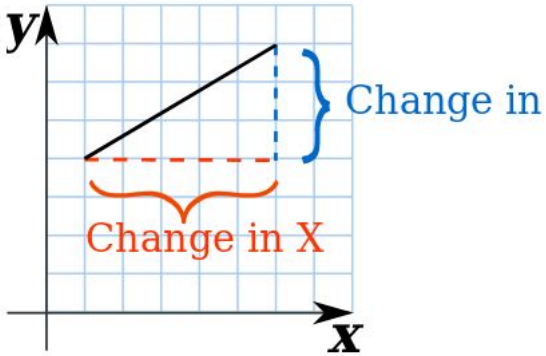


Source: Math is Fun

Calculus

- **Gradient**: The Gradient (also called Slope) of a straight line shows **how steep** a straight line is.

$$\text{Gradient} = \frac{\text{Change in Y}}{\text{Change in X}}$$



Source: Math is Fun

Calculus

- **Partial Derivatives**: Derivatives where we treat other variables as constants.

$$f(x) \quad \Rightarrow \quad f'(x) = \frac{df}{dx}$$

$$f(x,y) \quad \Rightarrow \quad f_x(x,y) = \frac{\partial f}{\partial x} \;\; \& \;\; f_y(x,y) = \frac{\partial f}{\partial y}$$

- Learn more : https://www.mathsisfun.com/data/outliers.htm

  https://youtu.be/FRlTh5HQORA

Calculus

- **Chain rule (for a full understanding of backpropagation algorithm for training neural network):**
    - The chain rule states that the derivative of f(g(x)) is f'(g(x))·g'(x). In other words, it helps us differentiate **composite functions**.
    - Let's simply this:

      Chain Rule using Leibniz's notation, we have:

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(g(x))] = \frac{\mathrm{d}}{\mathrm{d}g(x)}[f(g(x))] \cdot \frac{\mathrm{d}}{\mathrm{d}x}[g(x)]$$

      By letting $u=g(x)$ and $y=f(u)$, we can condense this to be:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \cdot \frac{\mathrm{d}u}{\mathrm{d}x}$$

- Know more: https://youtu.be/XIQ-KnsAsbg
- Visual demo of back propogation: https://google-developers.appspot.com/machine-learning/crash-course/backprop-scroll/

Found this guide helpful?

Then please give a star on github and contribute to make it more informative.

Keep rocking…

- Pratik