Data Science & Machine Learning

Eman Raslan

Course Schedule

Week #	Day1	Day2
1	24-Nov-2023	
2	1-Dec-2023	2-Dec-2023
3	8-Dec-2023	9-Dec-2023
4	15-Dec-2023	16-Dec-2023
5	22-Dec-2023	23-Dec-2023
6	29-Dec-2023	30-Dec-2023
7	5-Jan-2023	6-Jan-2023
8	12-Jan-2023	13-Jan-2023
	19-Jan-2023	
Project		

Course Agenda

- Statistics, Probability & Linear Algebra Basics
- Python for Data Science
- Data Exploration & Preparation
- Machine Learning

Linear Algebra Basics

Machine Learning = Mathematics

- Behind every ML success there is **Mathematics**.
- All ML models are constructed using solutions and ideas from math.
- The **purpose** of ML is to create **models** for understanding **thinking**.

Examples of Linear Algebra in Machine Learning

- Data Representation
- One-Hot Encoding
- Linear Regression
- Regularization
- Principal Component Analysis
- Recommender Systems
- Deep Learning

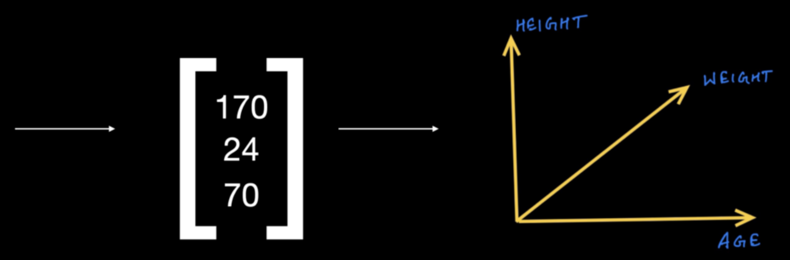
Data Representations

How can we represent data
(images, text, user preferences, etc.)
in a way that computers can understand?

Datasets

- Each dataset resembles a table-like structure consisting of **rows** and **columns**.
- This dataset is handled as a **Matrix**, which is a key data structure in Linear Algebra.
- Further, when this dataset is divided into input and output for the supervised learning model, it represents a **Matrix(X)** and **Vector(y)**, where the vector is also an important concept of linear algebra.

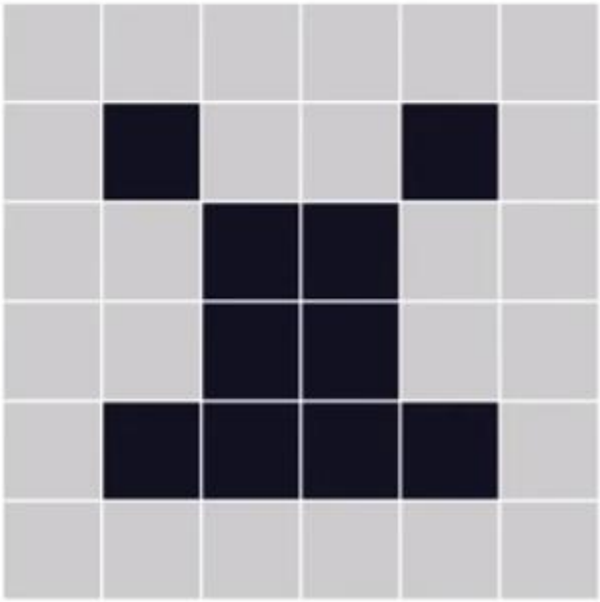
Height	Age	Weight
170	24	70
165	45	65
190	28	102
180	34	83
182	30	79
178	67	85



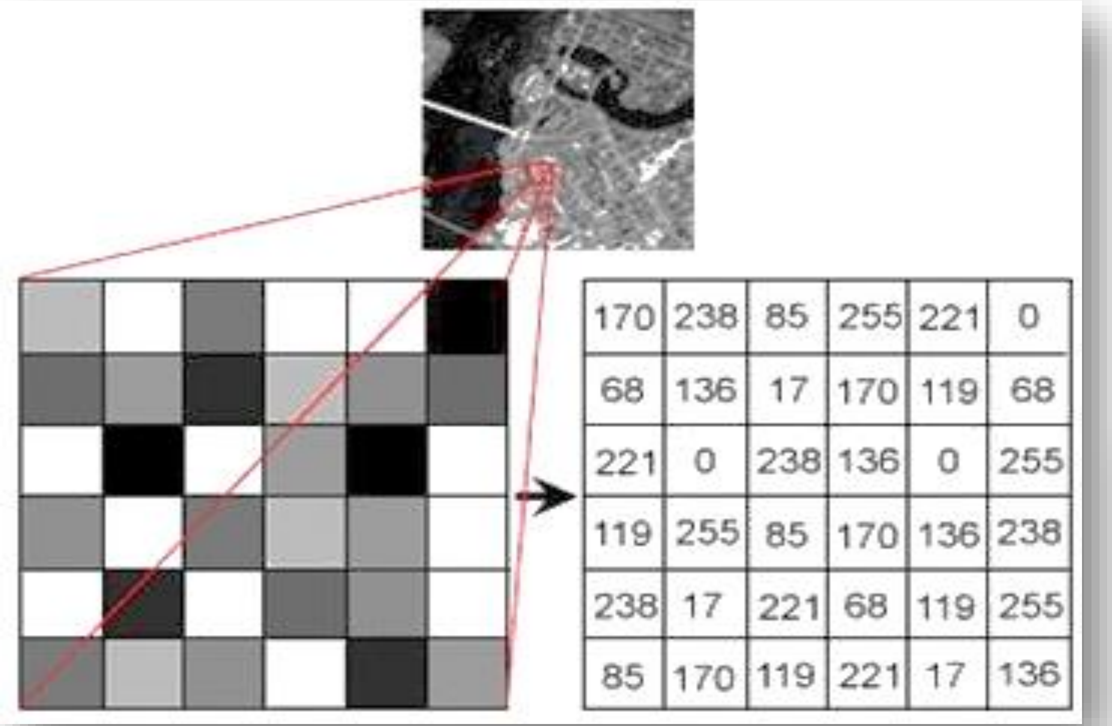
Examples of data representations

Images

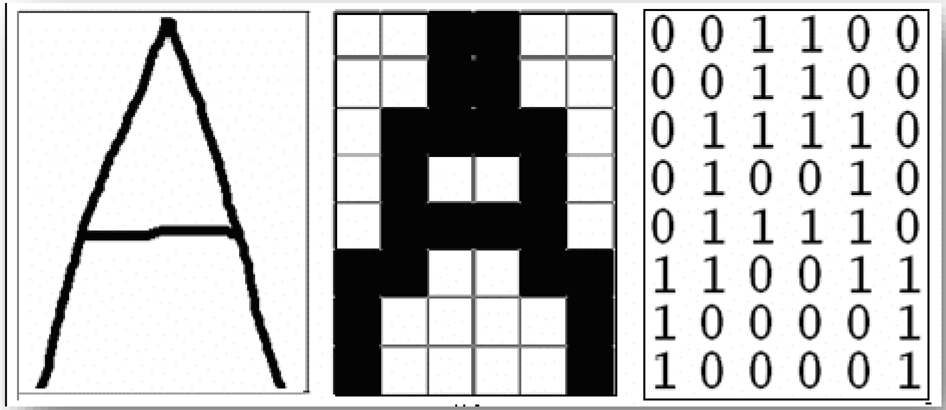
In black and white images, **black and white pixels** correspond to 0s and 1s. Grayscale pixels are numbers between 0 and 255. Both assemble into a 1-dimensional array of numbers.



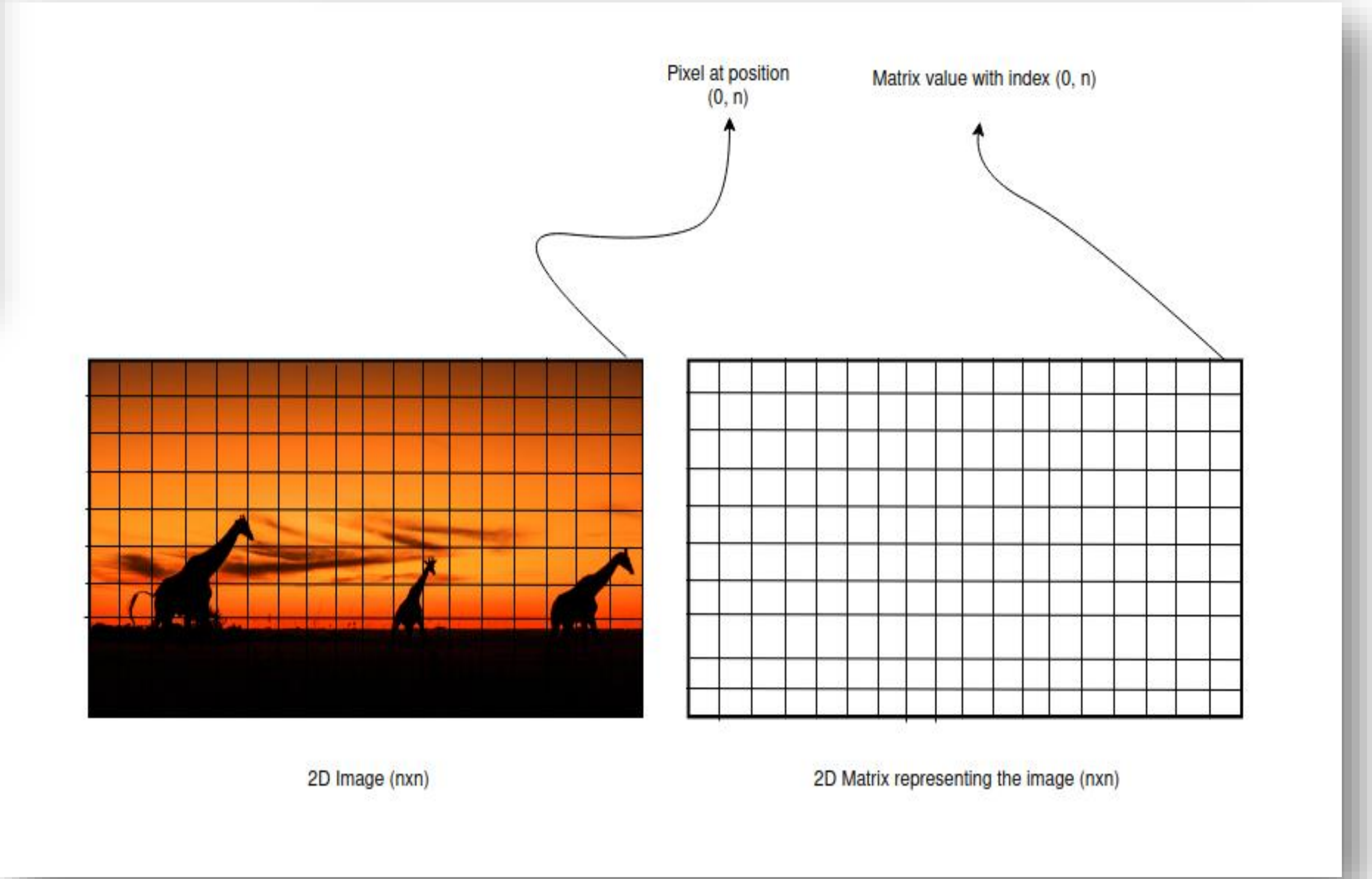
$$\rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$



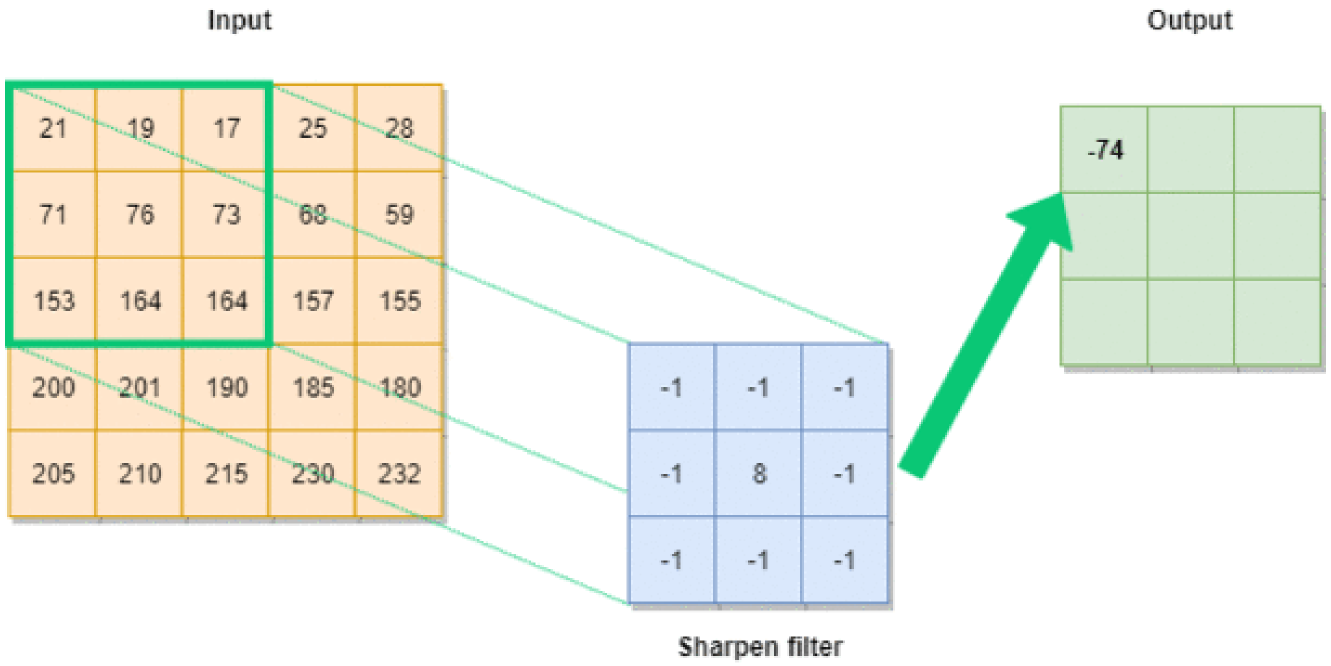
170	238	85	255	221	0
68	136	17	170	119	68
221	0	238	136	0	255
119	255	85	170	136	238
238	17	221	68	119	255
85	170	119	221	17	136



0	0	1	1	0	0
0	0	1	1	0	0
0	1	1	1	1	0
0	1	0	0	1	0
0	1	1	1	1	0
1	1	0	0	1	1
1	0	0	0	0	1
1	0	0	0	0	1



Operation	Kernel	Image result
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	



Examples of data representations

Words and Documents

Given a collection of documents (e.g. Wikipedia articles), assign to every word a vector whose i^{th} entry is the number of times the word appears in the i^{th} document.

$$\textbf{dog} = \begin{array}{cl} \begin{bmatrix} 0 \\ 7 \\ 0 \\ 0 \\ 51 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{l} \text{Wiki \#1} \\ \text{Wiki \#2} \\ \text{Wiki \#3} \\ \text{Wiki \#4} \\ \text{Wiki \#5} \\ \vdots \\ \text{Wiki \#54,000,000} \end{array} \end{array}$$

Examples of data representations

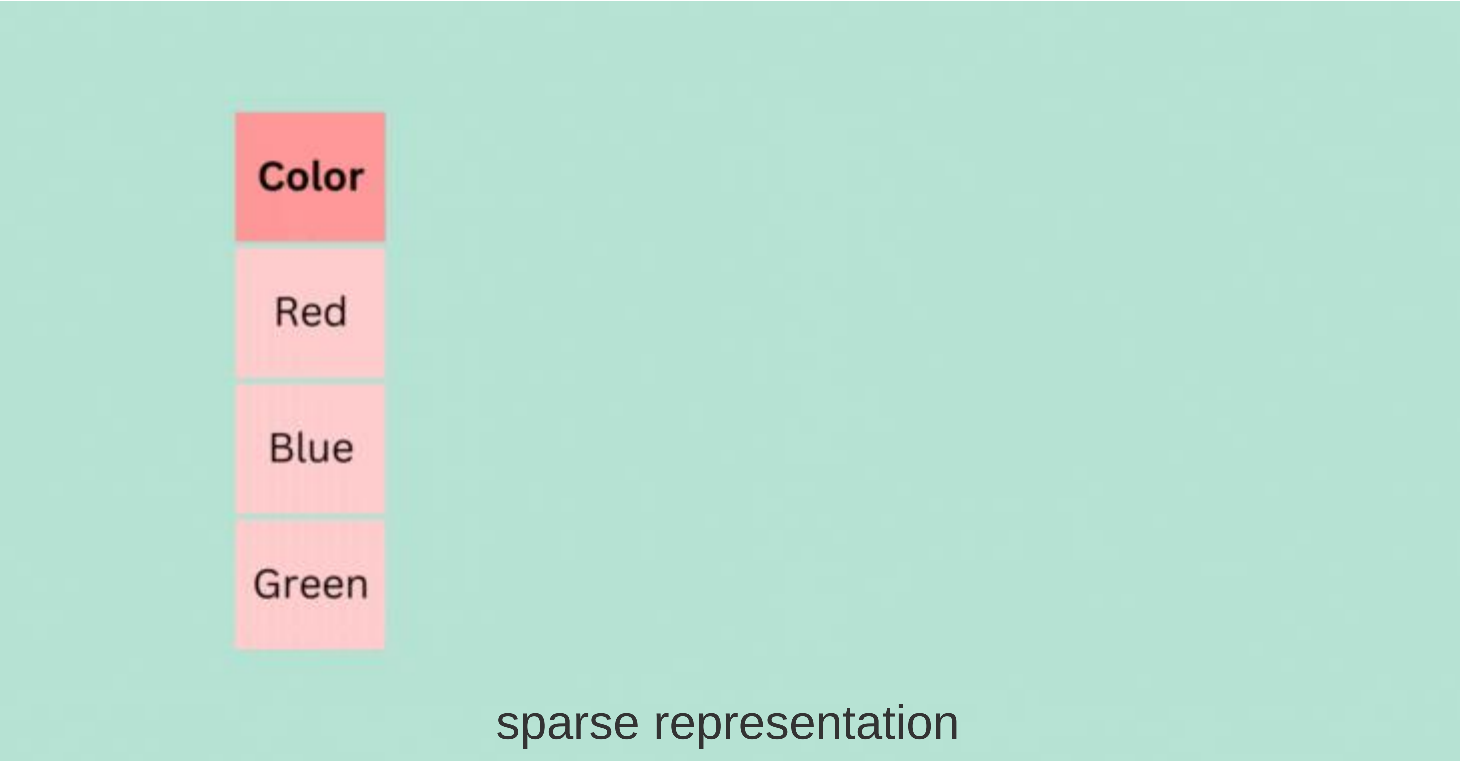
“One-Hot Encodings”

Assign to each word a vector with one 1 and 0s elsewhere. This is called a one-hot encoding (or a “standard basis vector”). For example, suppose our language only has four words:

$$\text{apple} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{cat} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{house} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{tiger} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Now imagine having *tens of thousands* of words....

One Hot Encoding



Examples of data representations
Yes/No or Ratings

Given users and items (e.g. movies), vectors can indicate if a user has interacted with the item (1 = yes, 0 = no) or the user's ratings, say a number between 0 and 5.

$$\text{User 1} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix} \begin{array}{c} \text{No} \\ \text{Yes} \\ \text{No} \\ \text{No} \\ \vdots \\ \text{Yes} \\ \text{No} \end{array} \quad \text{or} \quad \text{User 4} = \begin{bmatrix} 0 \\ 5 \\ 0 \\ 3 \\ \vdots \\ 0 \\ 2 \end{bmatrix} \begin{array}{c} ? \\ \text{Love} \\ ? \\ \text{Like} \\ \vdots \\ ? \\ \text{Dislike} \end{array}$$

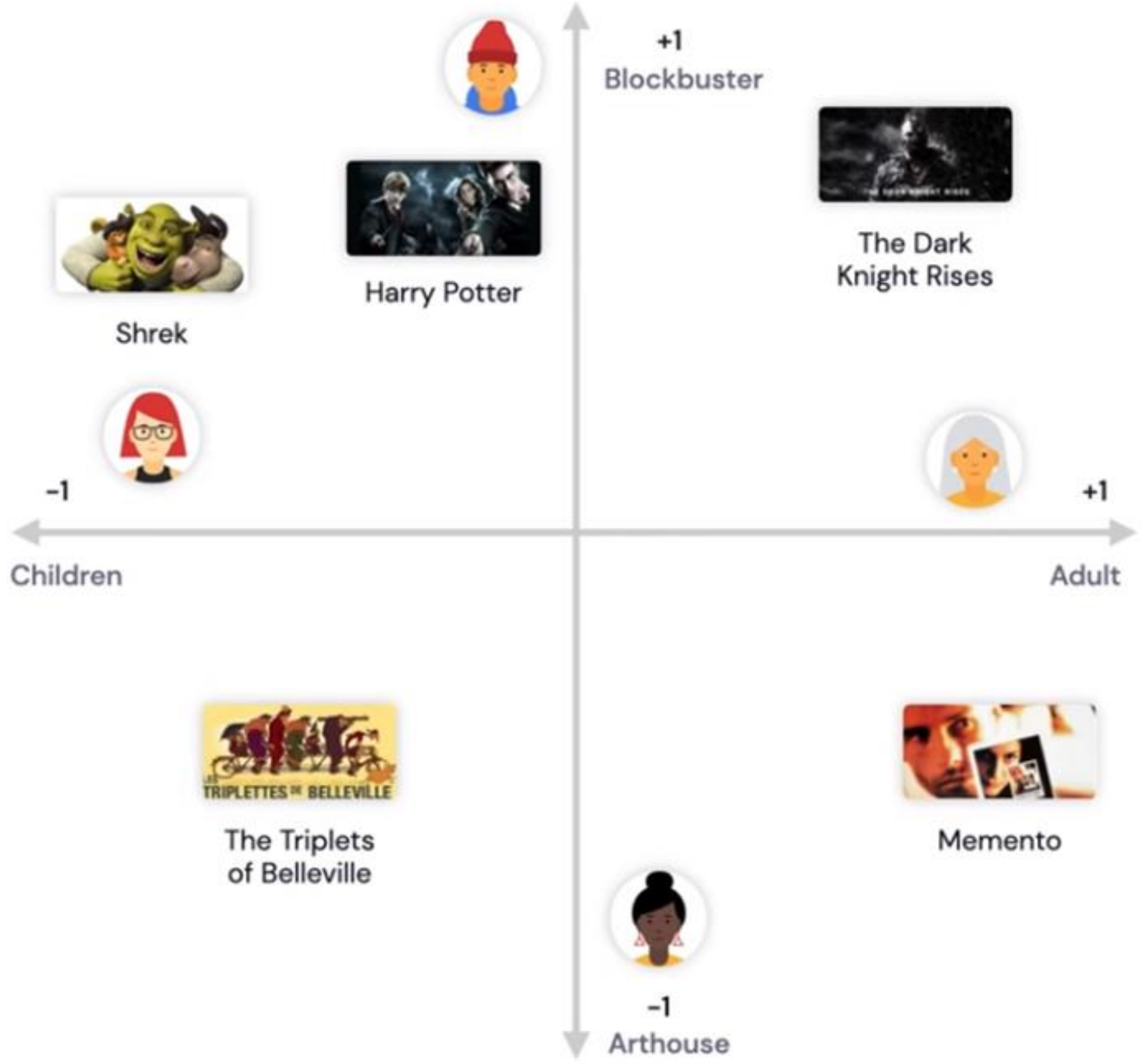
Imagine a **user-by-movie matrix** obtained by stacking users' data vectors side by side. (Checkmarks are 1s and empty cells are 0s or negatives.)

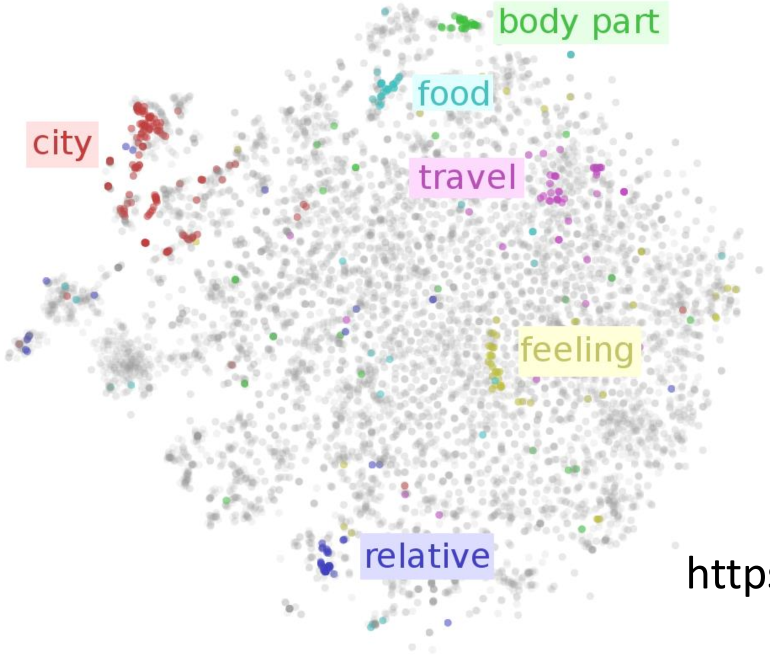
Images from Google Developers ML Crash Course on Collaborative Filtering.

For example, user #3 and *Shrek* correspond to these vectors:

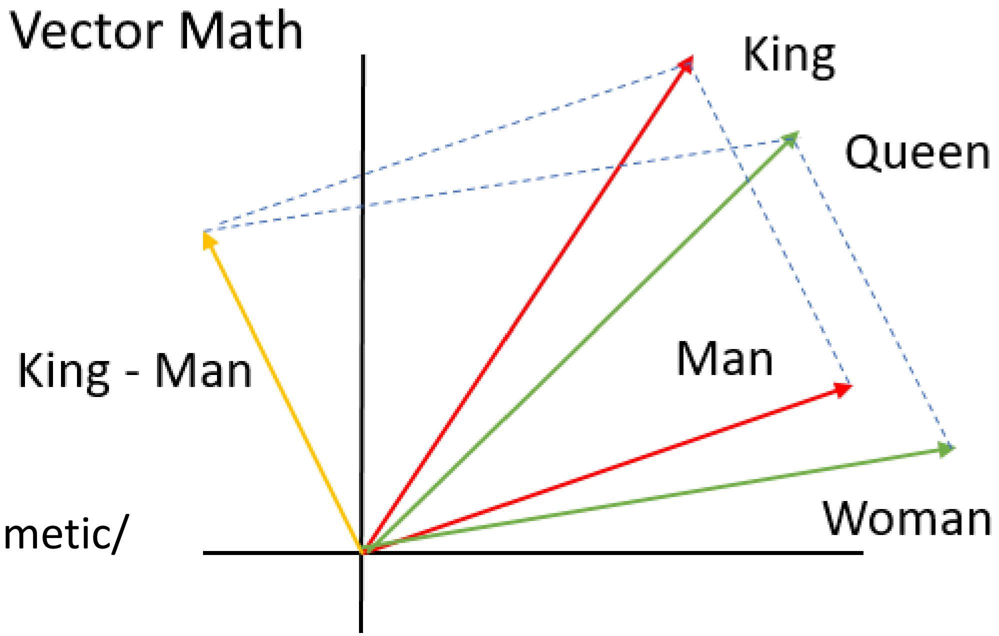
$$\text{User 3} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Images from Google Developers ML Crash Course on Collaborative Filtering.


+1
Blockbuster
The Dark
Knight Rises
Harry Potter
Shrek
-1
+1
Children
Adult
The Triplets
of Belleville
Memento
-1
Arthouse



https://dash.gallery/dash-word-arithmetic/



Vocabulary:

Man, woman, boy, girl, prince, princess, queen, king, monarch

	1	2	3	4	5	6	7	8	9
man	1	0	0	0	0	0	0	0	0
woman	0	1	0	0	0	0	0	0	0
boy	0	0	1	0	0	0	0	0	0
girl	0	0	0	1	0	0	0	0	0
prince	0	0	0	0	1	0	0	0	0
princess	0	0	0	0	0	1	0	0	0
queen	0	0	0	0	0	0	1	0	0
king	0	0	0	0	0	0	0	1	0
monarch	0	0	0	0	0	0	0	0	1

Each word gets a 1x9 vector representation

Machine Learning experts cannot live without Linear Algebra

Scalar

1

Vector(s)

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

Matrix

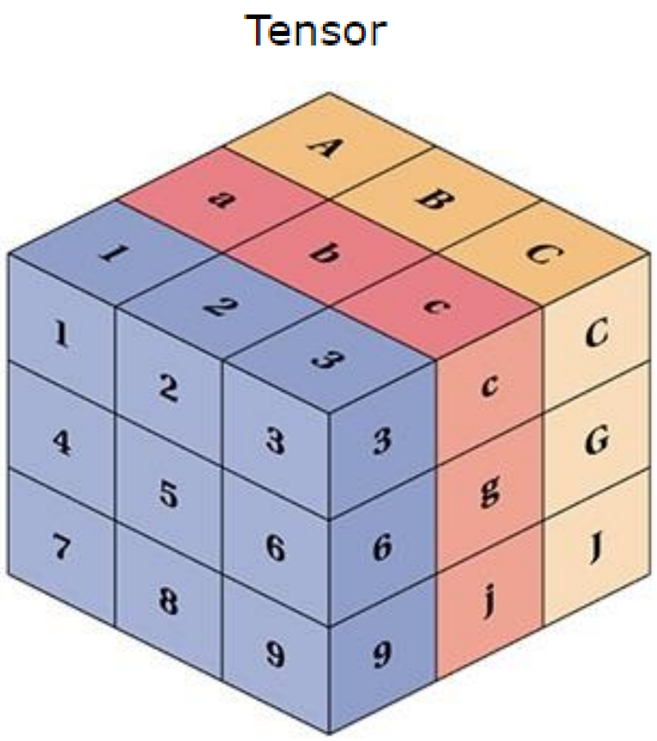
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Tensor

$$\begin{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} & \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} \end{bmatrix}$$

Linear Algebra

- A **Tensor** is a **N-dimensional Matrix:** A **Tensor** is a generalization of **Vectors** and **Matrices** to higher dimensions.
- A **Scalar** is a 0-dimensional tensor
- A **Vector** is a 1-dimensional tensor
- A **Matrix** is a 2-dimensional tensor



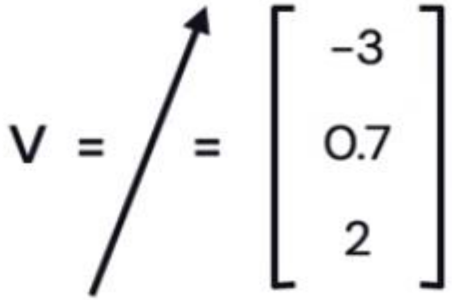
What are vectors?

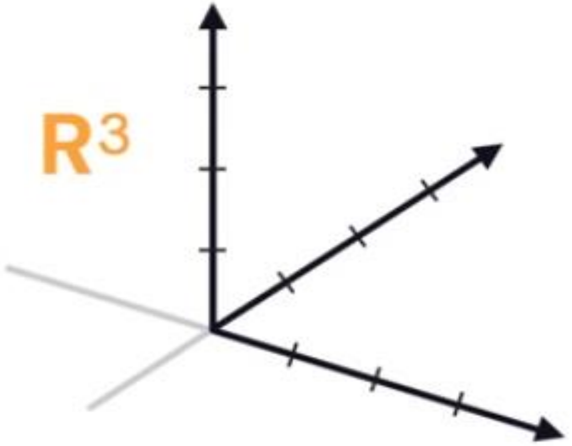
A **vector** is a 1-dimensional array of numbers.
It has both a *magnitude* (length) and a *direction*.

What are vectors?

In the context of machine learning...

A feature vector is a vector whose entries represent the “features” of some object.





“3-dimensional space” consists of all vectors with 3 entries:

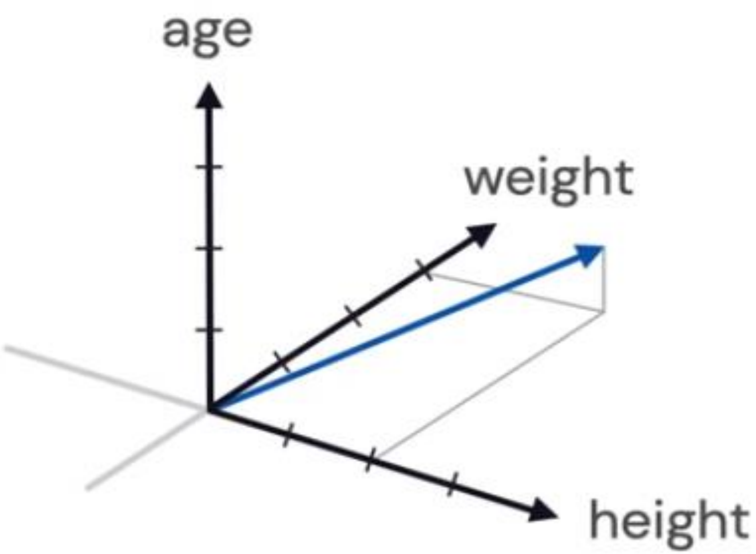
$$\begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

In the context of machine learning...

A feature vector is a vector whose entries represent the “features” of some object.

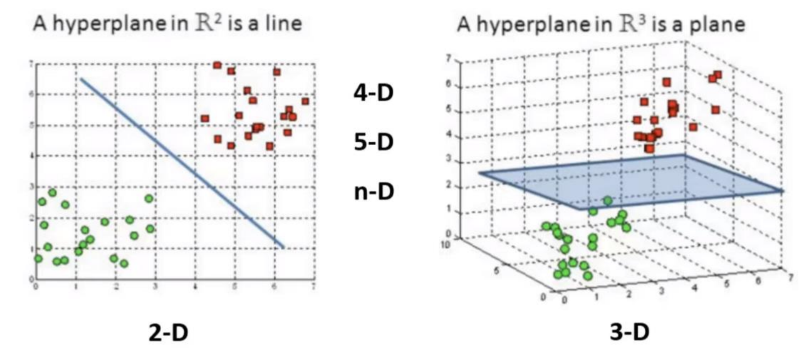
$$p = \begin{bmatrix} 64 \\ 131 \\ 23 \end{bmatrix} \begin{array}{l} \text{height} \\ \text{weight} \\ \text{age} \end{array}$$

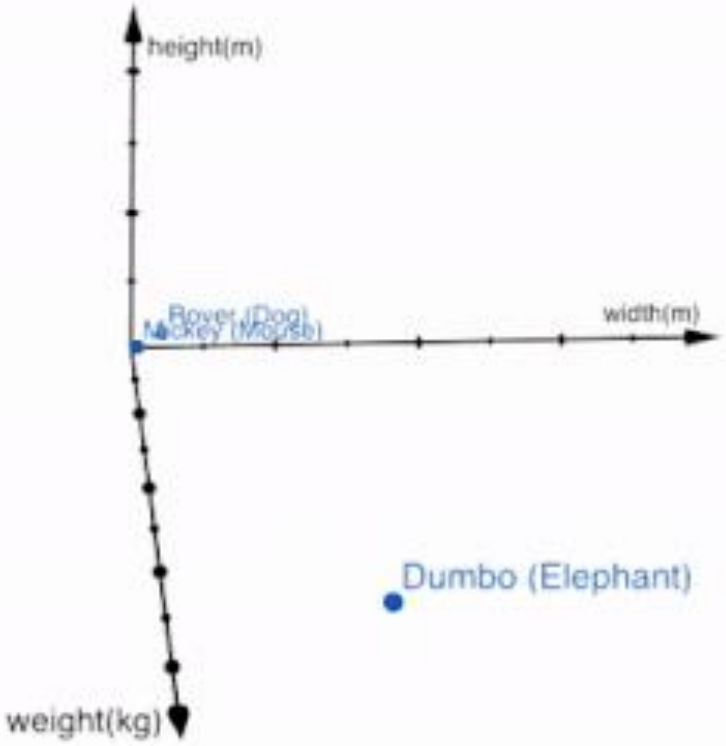
“p” for “patient”

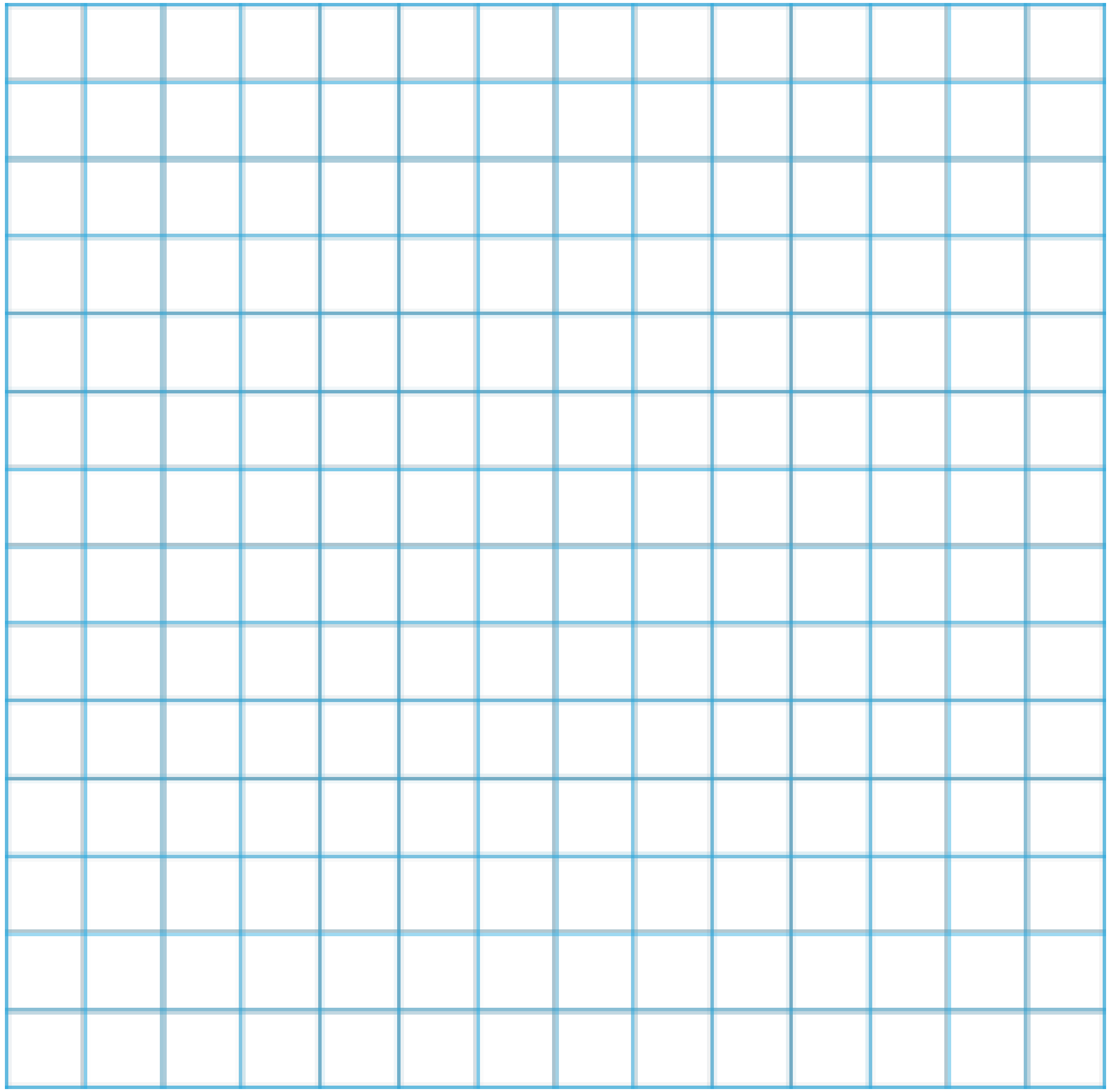


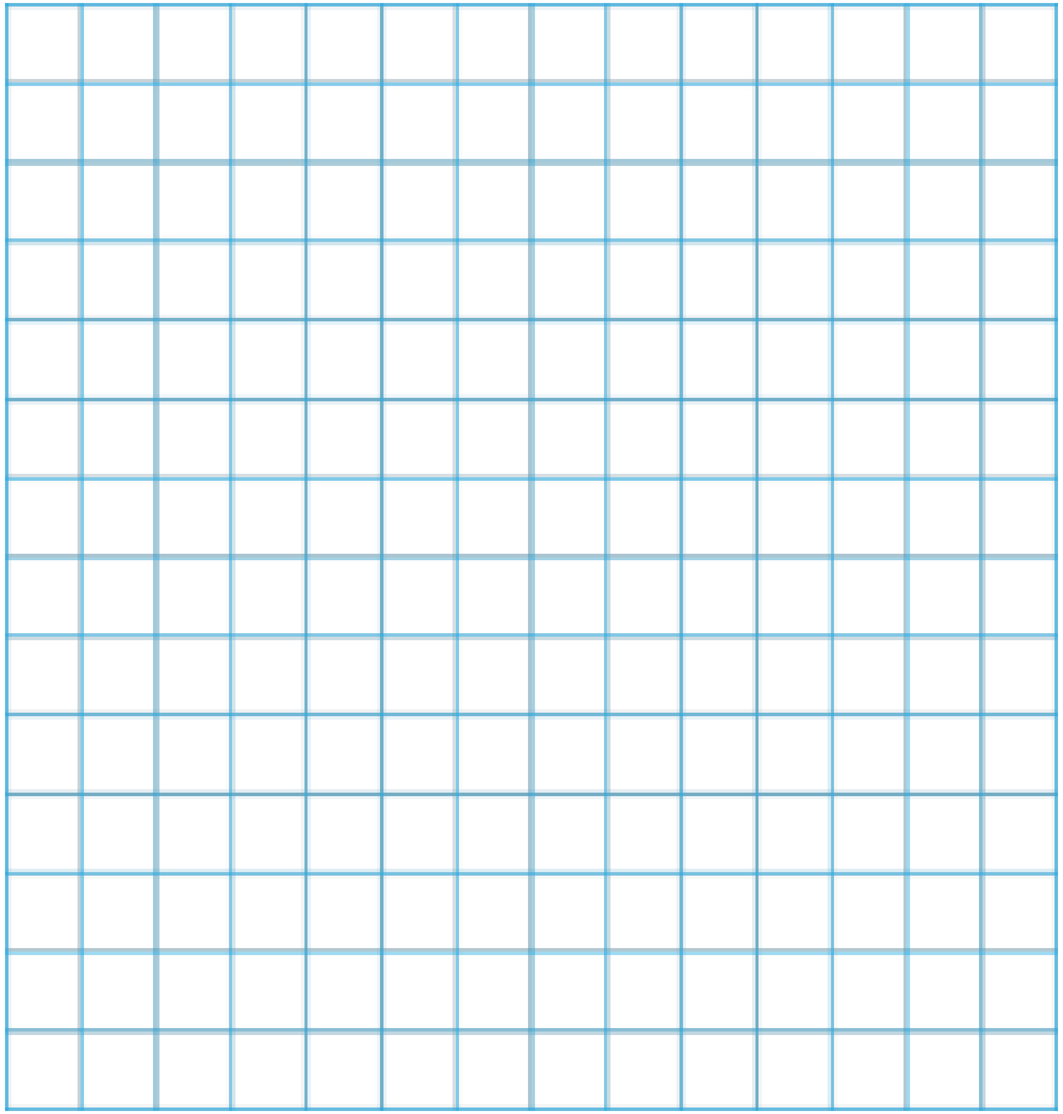
The vector space containing them is called **feature space**.

Feature Space



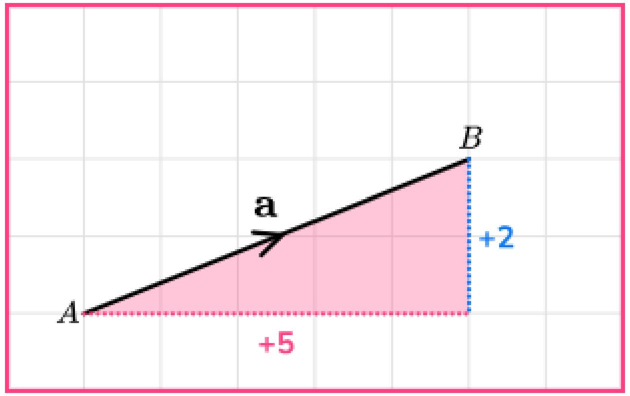


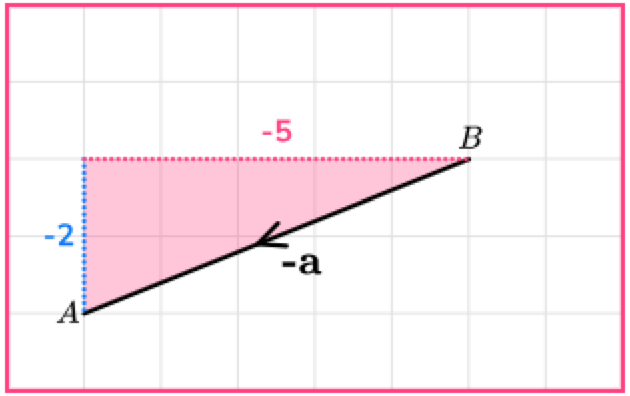




Vector notation

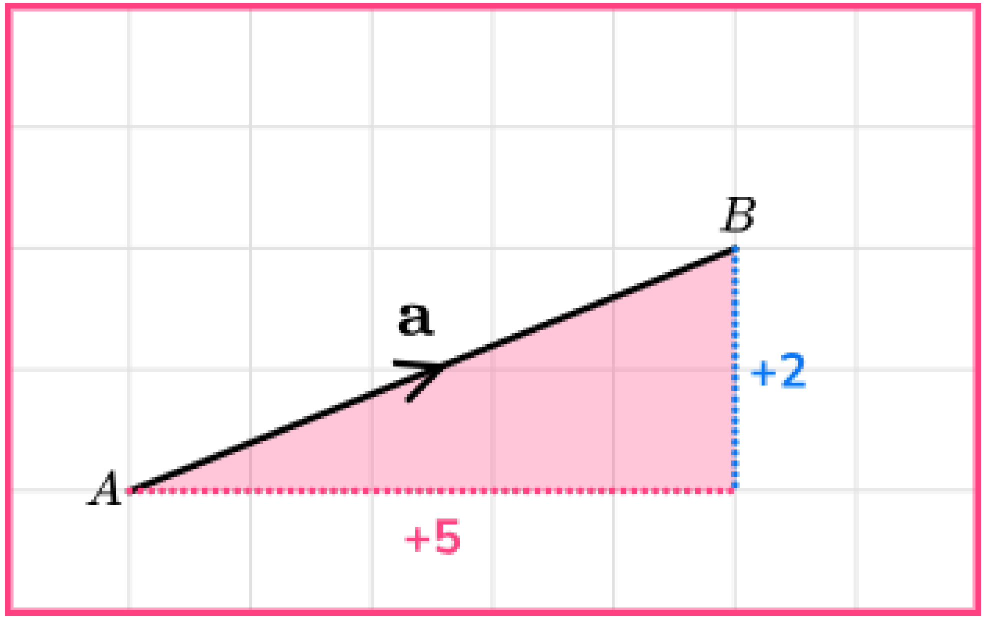
- Vectors can be represented by a straight line segment with an arrow to show the direction of the vector.





$$\boldsymbol{a} = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \qquad -\boldsymbol{a} = \begin{pmatrix} -5 \\ -2 \end{pmatrix}$$

Magnitude of a vector



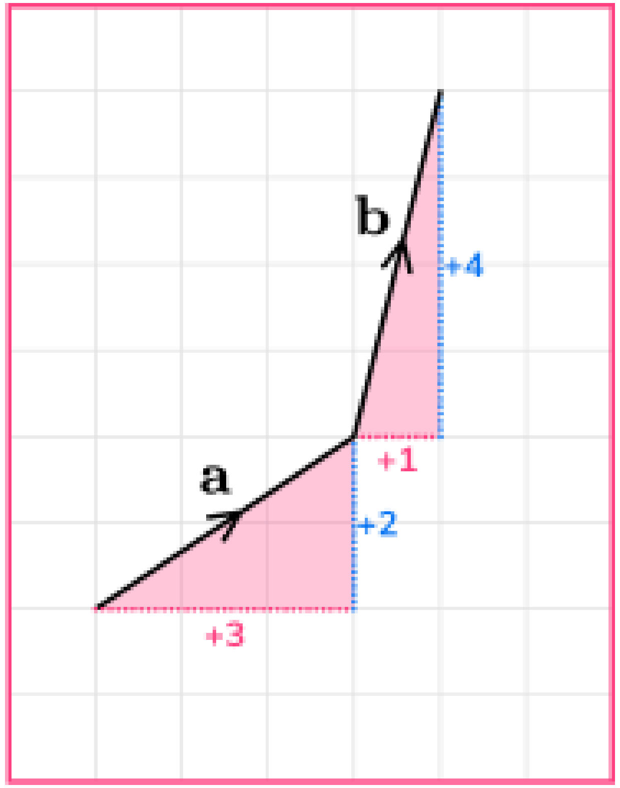
$$|\mathbf{a}| = \sqrt{x^2 + y^2} = \sqrt{5^2 + 2^2} = \sqrt{29}$$

If the magnitude is equal to 1, then the vector is known as a **unit vector.**
If the magnitude is equal to 0 , then the vector is known as a **zero vector**.

Vector addition

$$\mathbf{a} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

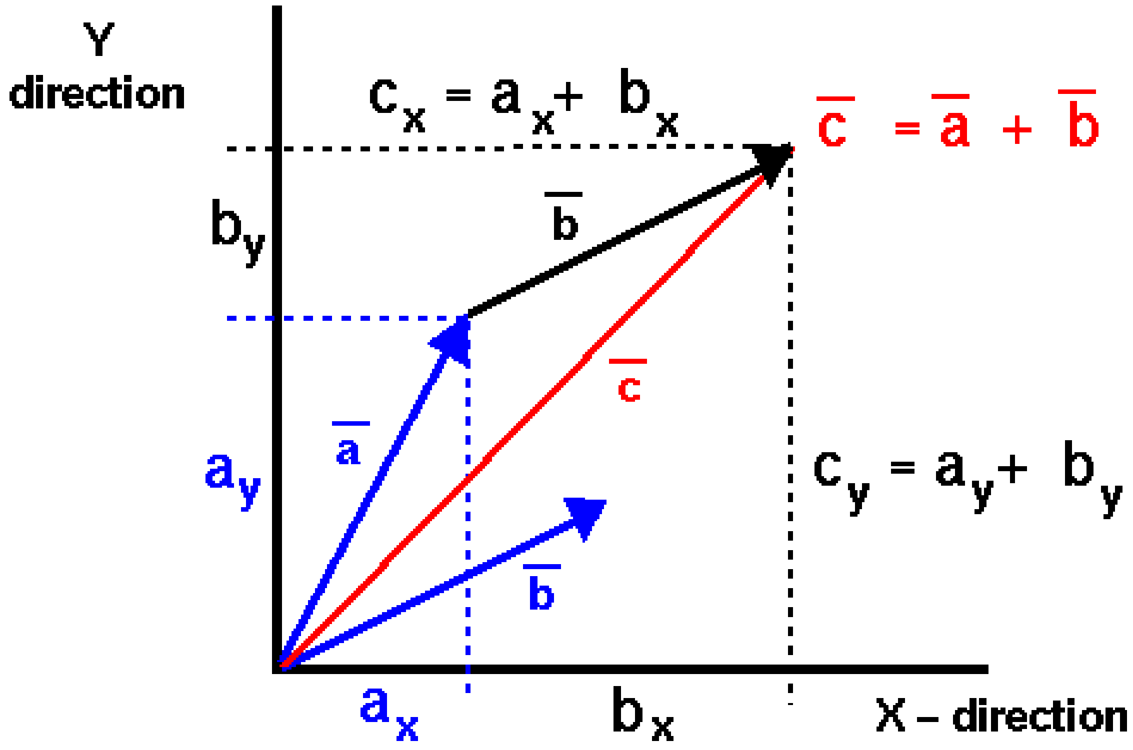
$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$



Vector addition

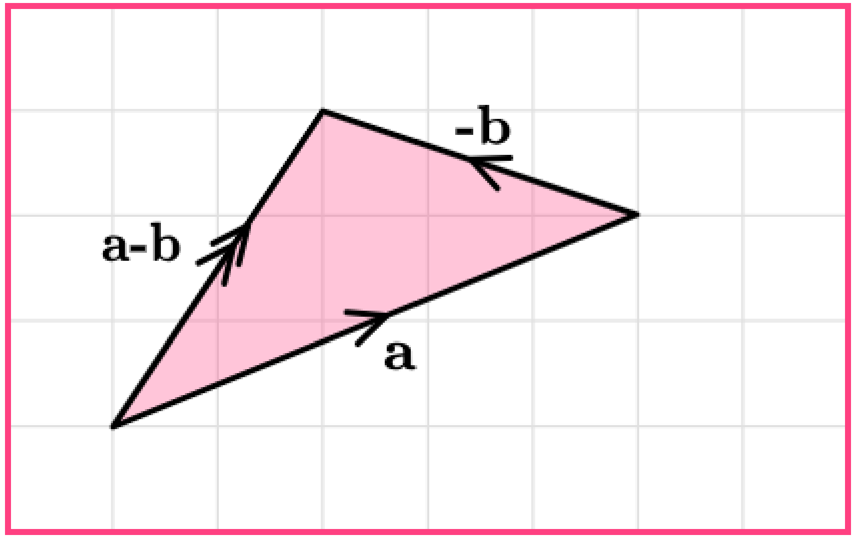
A vector quantity has both magnitude and direction.

Add the vector components.



Vector subtraction

When subtracting vectors, the order is important.



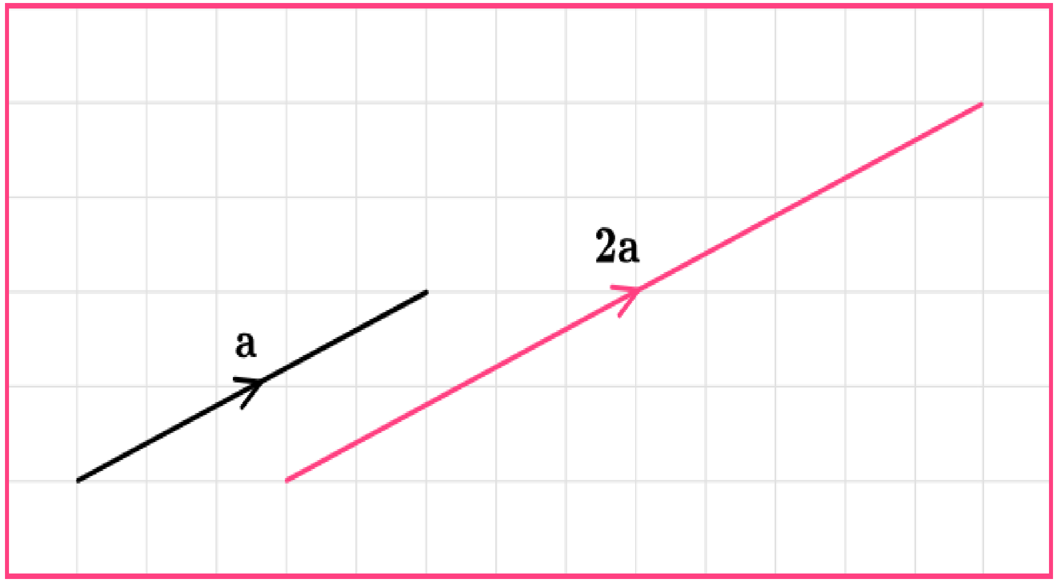
$$\boldsymbol{a} - \boldsymbol{b} = \begin{pmatrix} 5 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

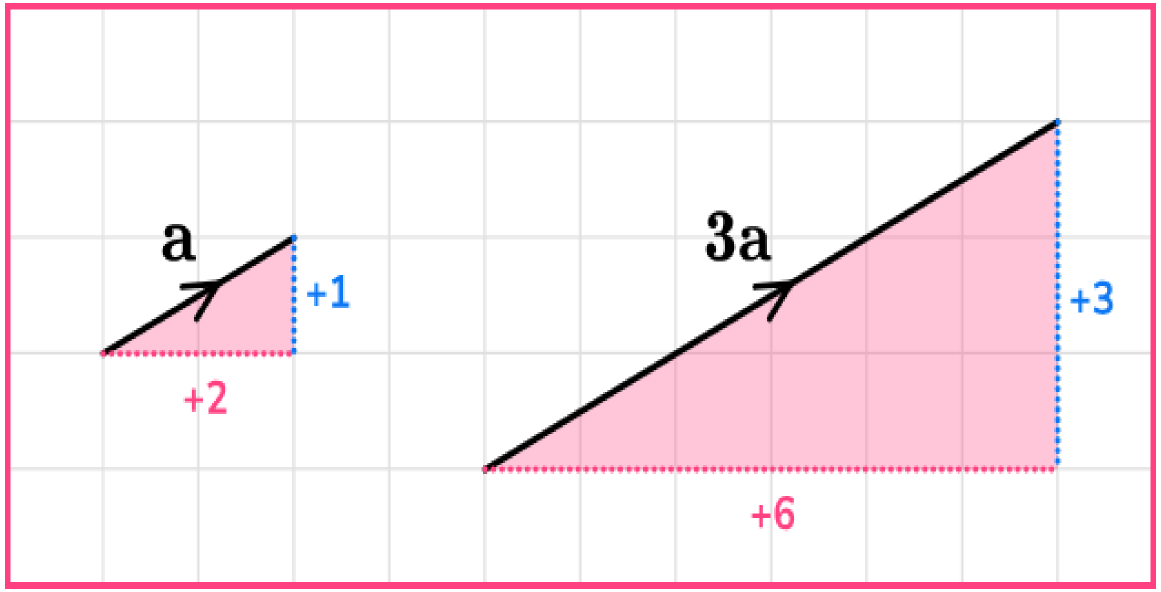
Scalar multiplication

$$\boldsymbol{a} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

$$2\boldsymbol{a} = 2\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \times 5 \\ 2 \times 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}$$

$$3\mathbf{a} = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \times 2 \\ 3 \times 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$





Scalar multiplication

The first key operation is **scalar multiplication**, multiplying a scalar and a vector. If k is a scalar and $\vec{v}$ is a vector, their product $k\vec{v}$ is defined as follows:

- If $k > 0$, then $k\vec{v}$ is the vector pointing in the same direction as $\vec{v}$ that's k times as long as $\vec{v}$.
- If $k = 0$, then $k\vec{v}$ is $\vec{0}$.
- If $k < 0$, then $k\vec{v}$ is the vector pointing in the opposite direction from $\vec{v}$ that's $|k|$ times as long as $\vec{v}$.

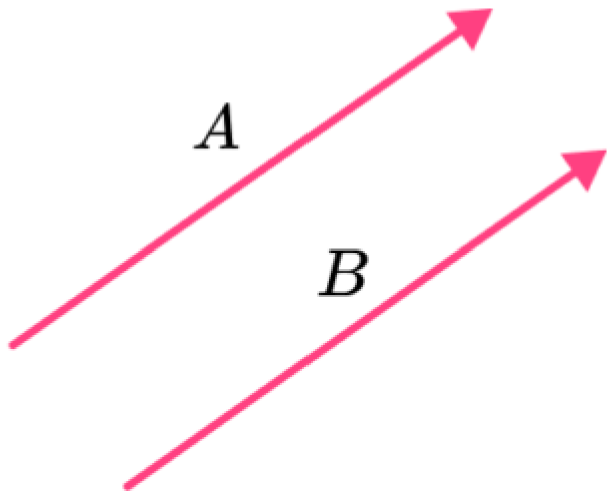
$\vec{v}$

$2\vec{v}$: pointing in the same direction as $\vec{v}$ and twice as long as $\vec{v}$

$-0.6\vec{v}$: pointing in the opposite direction from $\vec{v}$ and 0.6 times as long as $\vec{v}$

Equal Vectors

Vectors are **equal** if they have the same **magnitude** and **direction** regardless of **where they are**.



Vectors A and B are equal. They are **travelling** in the same **direction** and have the same **magnitude** (length).

Algebraic Properties of Vectors

For any vectors $\vec{u} = (u_1, u_2)$, $\vec{v} = (v_1, v_2)$, and $\vec{\omega} = (\omega_1, \omega_2)$ and scalars n and m, consider the following.

- Properties of vector addition:

$\vec{u} + \vec{v} = \vec{v} + \vec{u}$	(commutative property)
$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$	(associative property)
$\vec{u} + \vec{0} = \vec{u}$	(additive identity property)
$\vec{u} + (-\vec{u}) = \vec{0}$	(additive inverse property)
If $\vec{u} + \vec{v} = \vec{u} + \vec{w}$, then $\vec{v} = \vec{w}$	(elimination property)

- Properties of scalar multiplication of vectors:

$n(\vec{u} + \vec{v}) = n\vec{u} + n\vec{v}$	(distributive property)
$(n + m)\vec{u} = n\vec{u} + m\vec{u}$	(distributive property)
$1\vec{u} = \vec{u}$	(multiplicative identity property)
$(nm)\vec{u} = n(m\vec{u})$	(associative property)
If $n\vec{u} = n\vec{v}$, then $\vec{u} = \vec{v}$	(elimination property)

The dot product

Just like we can multiply numbers, we can also multiply vectors.
This is called the **dot product**.

The product of numbers is another number.
The dot product of vectors is *not* another vector!

$2 \times 5 = 10$ **vs** $\begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} = 7$

Numbers Vectors A number

The dot product

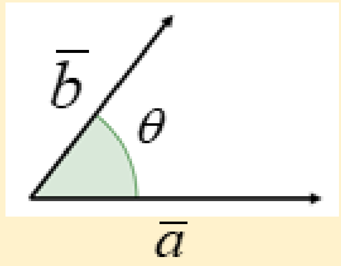
$$\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 2 \\ -1 \end{bmatrix} = (1)(7) + (0)(2) + (3)(-1) = 4$$

Dot Product

If $\bar{a} = < a_1, a_2, a_3 >$ and $\bar{b} = < b_1, b_2, b_3 >$

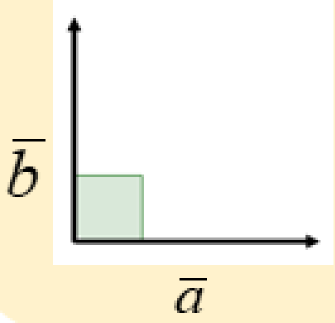
then the dot product is

$$\bar{a} \bullet \bar{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$



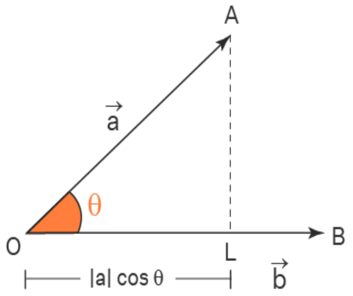
If θ is the angle between $\bar{a}$ and $\bar{b}$ then

$$\bar{a} \bullet \bar{b} = |\bar{a}||\bar{b}|\cos\theta$$



$\bar{a} \bullet \bar{b}$ are orthogonal (perpendicular)

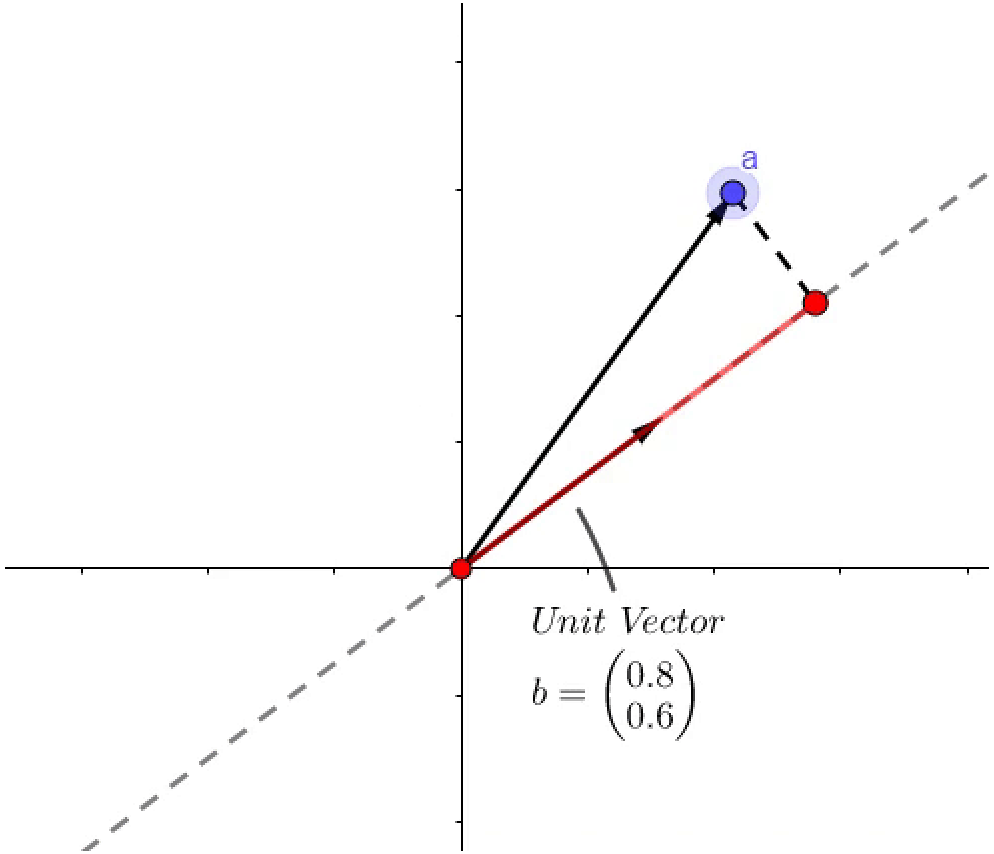
if and only if $\bar{a} \bullet \bar{b} = 0$



Vector projection

https://www.geogebra.org/m/segQU7mb

$$\text{Projection of Vector a on Vector b} = \frac{a.b}{|b|}$$



Here's something to think about:

Dot products between words are zero, no matter how "similar" they are.

$$\text{apple} \cdot \text{cat} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \text{tiger} \cdot \text{cat}$$

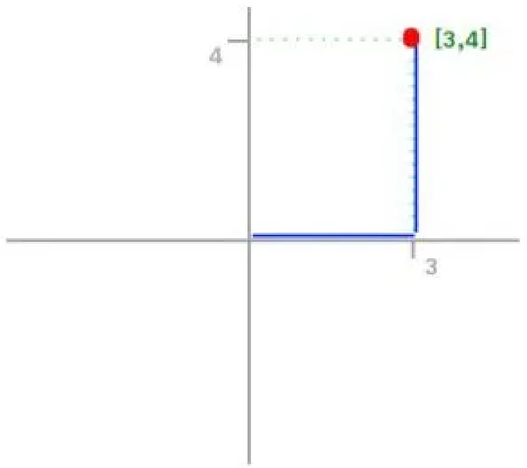
Vector norm

- The length of a vector is a **nonnegative number** that describes the **extent of** the vector in space, and is sometimes referred to as the vector’s **magnitude** or the **norm**.
- Common vector norm calculations used in machine learning.
  - Vector Norm
  - Vector L1 Norm
  - Vector L2 Norm
  - Vector Max Norm

L1 Norm

$$\underbrace{\|x\|_1}_{L1\ \ Norm} = (\sum_i^n |x_i|) = (|x_1| + |x_2| + \cdots + |x_n|)$$

Having, for example, the vector X = [3,4]:



The L1 norm is calculated by

$||X||_1 = |3| + |4| = 7$

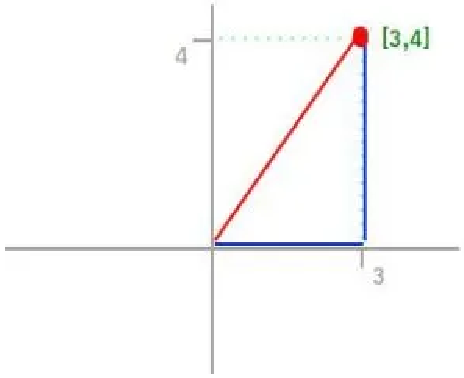
L2 Norm

$$\underbrace{\|x\|_2}_{L-2\ \ Norm} = (\sum_{i}^{n} |x_i|^2)^{\frac{1}{2}} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_n \end{bmatrix}$$

$$||u||_2 = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Is the most popular norm, also known as the Euclidean norm. It is the shortest distance to go from one point to another.



Using the same example, the L2 norm is calculated by

$$||X||2 = \sqrt{(|3|^2 + |4|^2)} = \sqrt{9+16} = \sqrt{25} = 5$$

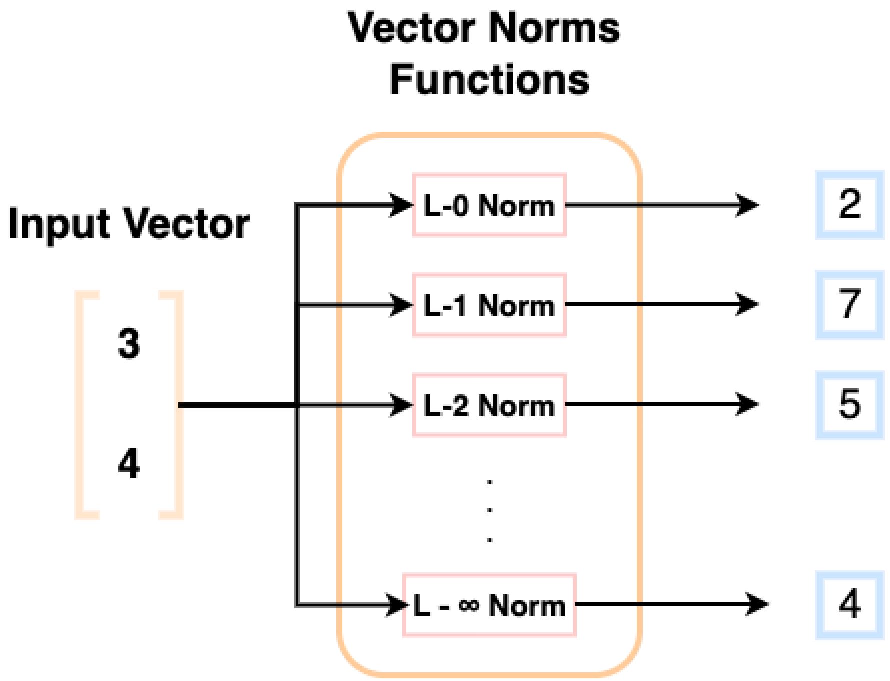
As you can see in the graphic, L2 norm is the most direct route.

L-infinity norm

- Gives the **largest absolute magnitude** among each element of a vector.
- Having the vector X= [-6, 4, 2], the L-infinity norm is 6.

$$\underbrace{\|x\|_\infty}_{Max\ Norm} = \max_i |x_i|$$

Vector norm



- Vector **Norms** are defined as a set of functions that **take** a vector as an **input** and **output** a **positive value** against it.
- This is called the ***magnitude*** of a vector.
- We can obtain **different lengths** for the **same vector** depending on the type of **function** we use to calculate the magnitude.

Norm in ML

Manhattan distance

$$\sum_{i=1}^{n} |x_i - y_i|$$

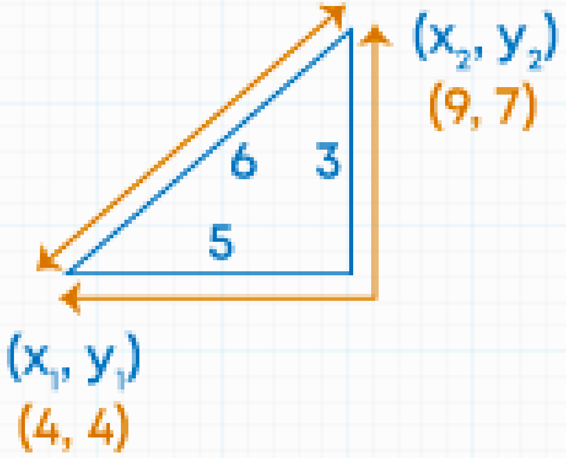
L1 Norm

Squared Euclidean distance

$$\sum_{i=1}^{n} (x_i - y_i)^2$$

L2 Norm

Example:



Euclidean distance

$= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

$= \sqrt{(9 - 4)^2 + (7 - 4)^2}$

$= \sqrt{5^2 + 3^2}$

$= \sqrt{25 + 9}$

$= \sqrt{34}$

$= 5.83$

Manhattan distance

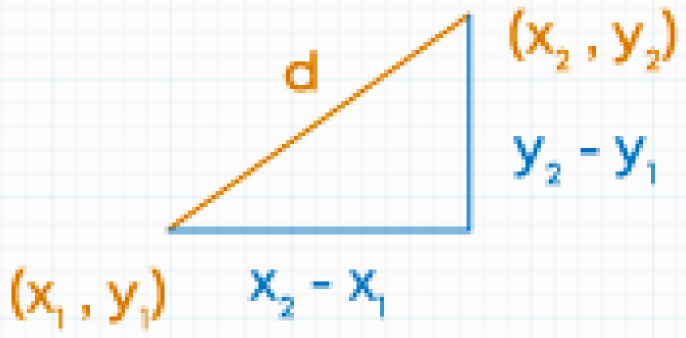
$= |x_2 - x_1| + |y_2 - y_1|$

$= |9 - 4| + |7 - 4|$

$= 5 + 3$

$= 8$

Euclidean distance (d) = $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$



In 3 dimensional space for points (x_1, y_1, z_1) and (x_2, y_2, z_2)

Euclidean distance (d) = $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$

Norm in ML

- **Vector norm** is a function that takes a vector as an input and outputs a positive value.
- The **L1 norm** is also referred to as the ***Mean Absolute Error***.
- The **L2 Norm** is also referred to as the ***Root Mean Squared Error***.
- The **Squared L2 Norm** is also referred to as the ***Mean Squared Error***.

Matrix

$$
\begin{array}{rc}
 & \overbrace{\begin{array}{cccc} 1 & 2 & \cdots\cdots & n \end{array}}^{\text{Columns}} \\
\text{Rows}\left\{\begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \\ m \end{array}\right. &
\begin{bmatrix}
a_{11} & a_{12} & \cdots\cdots & a_{1n} \\
a_{21} & a_{22} & \cdots\cdots & a_{2n} \\
a_{31} & a_{32} & \cdots\cdots & a_{3n} \\
\vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots\cdots & a_{mn}
\end{bmatrix} = A_{m \times n}
\end{array}
$$

Matrix addition

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix}4 & 8\\ 3 & 7\end{bmatrix}+\begin{bmatrix}1 & 0\\ 5 & 2\end{bmatrix}$$

$$=\begin{bmatrix}4+1 & 8+0\\ 3+5 & 7+2\end{bmatrix}$$

$$=\begin{bmatrix}5 & 8\\ 8 & 9\end{bmatrix}$$

Matrix subtraction

$$\mathbf{C} - \mathbf{D} = \begin{bmatrix} 2 & 8 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} 5 & 6 \\ 11 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 2-5 & 8-6 \\ 0-11 & 9-3 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & 2 \\ -11 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 4 \\ 5 & 3 \end{bmatrix}$$

❌ We cannot add them

Or Subtract them

Types of Matrices

Row Matrix

$$\begin{pmatrix} a & b & c \end{pmatrix}$$

Column Matrix
Vector Matrix

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Zero Matrix
Null Matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Diagonal Matrix

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Scalar Matrix

$$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$$

Unit Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Upper Triangular Matrix

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

Lower Triangular Matrix

$$\begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}$$

Properties of matrix addition

Property	Example
Commutative property of addition	$A + B = B + A$
Associative property of addition	$A + (B + C) = (A + B) + C$
Additive identity property	For any matrix A, there is a unique matrix O such that $A + O = A$.
Additive inverse property	For each A, there is a unique matrix $-A$ such that $A + (-A) = O$.
Closure property of addition	$A + B$ is a matrix of the same dimensions as A and B.

Multiplying matrices by scalars

$$\boxed{\text{num}} \qquad \boxed{X} \qquad \boxed{\text{num * X}}$$

$$100 \quad * \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1*100 & 2*100 & 3*100 \\ 4*100 & 5*100 & 6*100 \\ 7*100 & 8*100 & 9*100 \end{bmatrix}$$

$$\boxed{\text{num * X}} = \begin{bmatrix} 100 & 200 & 300 \\ 400 & 500 & 600 \\ 700 & 800 & 900 \end{bmatrix}$$

Properties of matrix addition & scalar multiplication

Property	Example
Associative property of multiplication	$(cd)A = c(dA)$
Distributive properties	$c(A+B) = cA + cB$
	$(c+d)A = cA + dA$
Multiplicative identity property	$1A = A$
Multiplicative properties of zero	$0 \cdot A = O$
	$c \cdot O = O$
Closure property of multiplication	cA is a matrix of the same dimensions as A.

Matrix Transpose

transpose of $m \times n$ matrix A, denoted A^T or A', is $n \times m$ matrix with

$$(A^T)_{ij} = A_{ji}$$

rows and columns of A are transposed in A^T

example: $\begin{bmatrix} 0 & 4 \\ 7 & 0 \\ 3 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 7 & 3 \\ 4 & 0 & 1 \end{bmatrix}$.

- transpose converts row vectors to column vectors, vice versa
- $(A^T)^T = A$

Matrix Multiplication

$$c_{11}=a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31}+a_{14}b_{41}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

2 x 4 4 x 3 2 x 3

$$(m \times n) \cdot (n \times k) = (m \times k)$$

product is defined

$$c_{22}=a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32}+a_{24}b_{42}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

Matrix Multiplication

In this table, A, B, and C are $n \times n$ matrices, I is the $n \times n$ identity matrix, and O is the $n \times n$ zero matrix

Property	Example
The commutative property of multiplication does not hold!	$AB \neq BA$
Associative property of multiplication	$(AB)C = A(BC)$
Distributive properties	$A(B+C) = AB + AC$
	$(B+C)A = BA + CA$
Multiplicative identity property	$IA = A$ and $AI = A$
Multiplicative property of zero	$OA = O$ and $AO = O$
Dimension property	The product of an $m \times n$ matrix and an $n \times k$ matrix is an $m \times k$ matrix.

Matrix Determinant

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$|A| = ad - bc$$

Remember by this -

$$\begin{matrix} a & & b \\ & \times & \\ c & & d \end{matrix}$$

$$A = \begin{bmatrix} 2 & 1 \\ -6 & 4 \end{bmatrix}$$

$$\det(A) = [(2)(4)] - [(1)(-6)]$$

$$= 8 - (-6)$$

$$= 14$$

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

$$\det(A) = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}$$

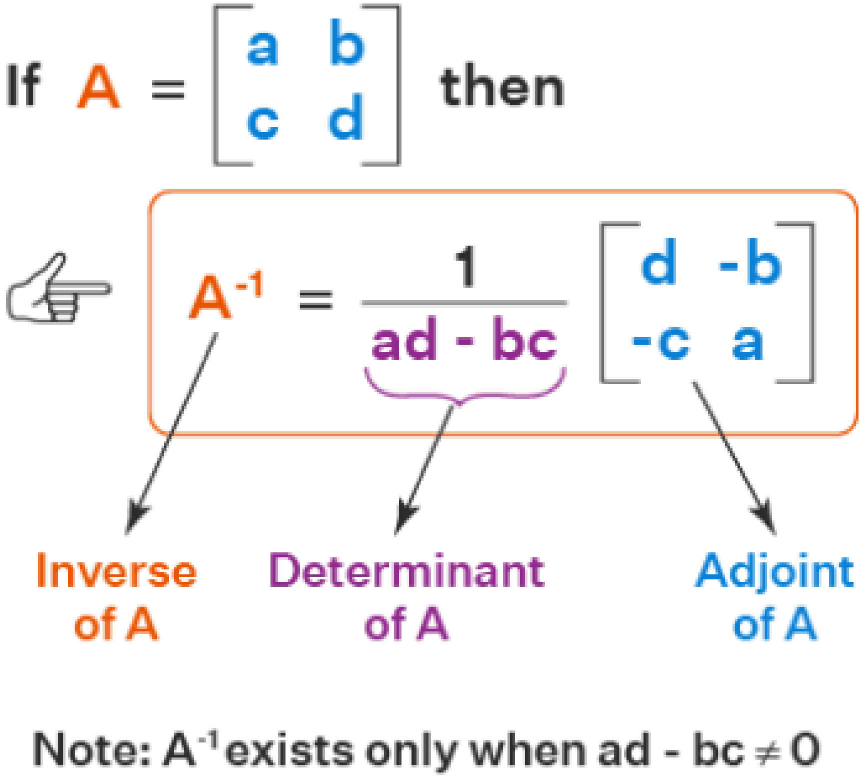
$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

$$\det(A) = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

$$= a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)$$

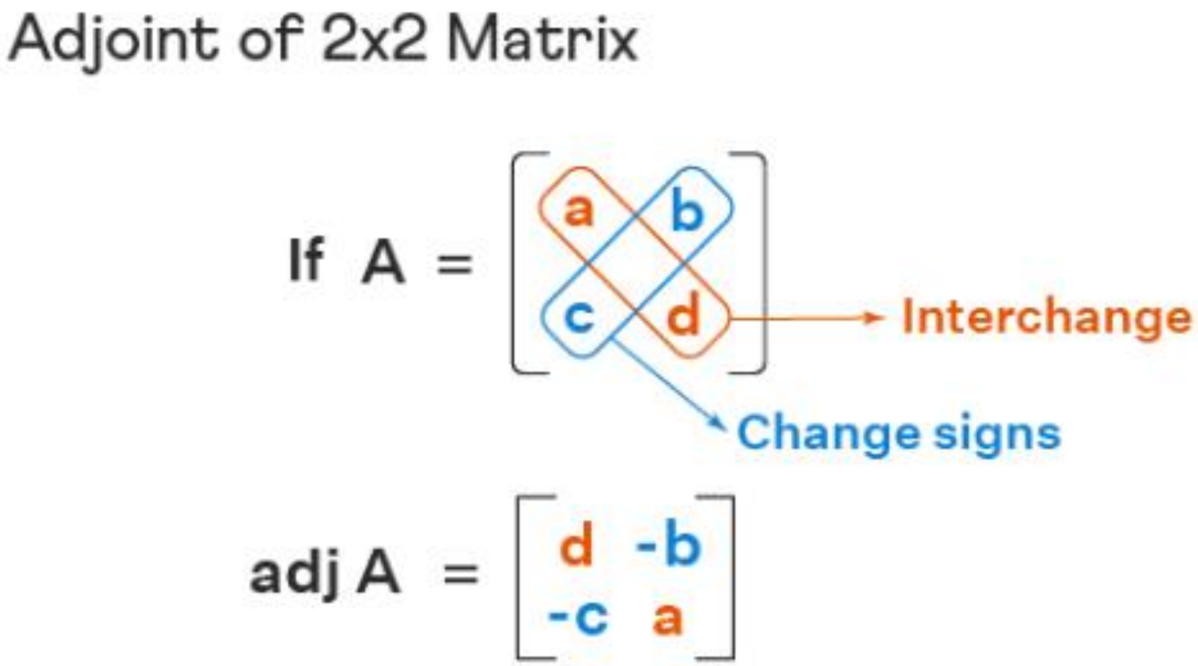
Matrix Inverse

$I = A^{-1} A$



If ad - bc = 0, then A^{-1} does NOT exist and in this case, we call A to be a singular matrix.

Matrix Adjoint

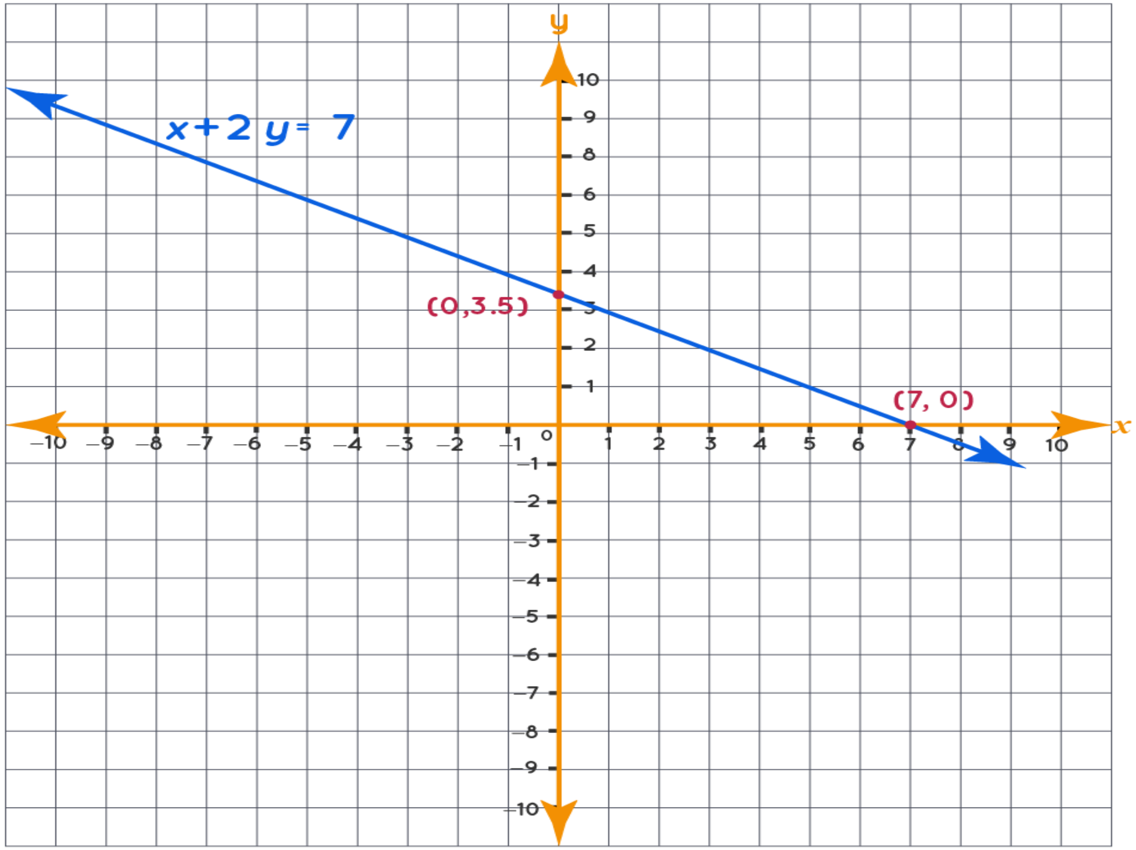


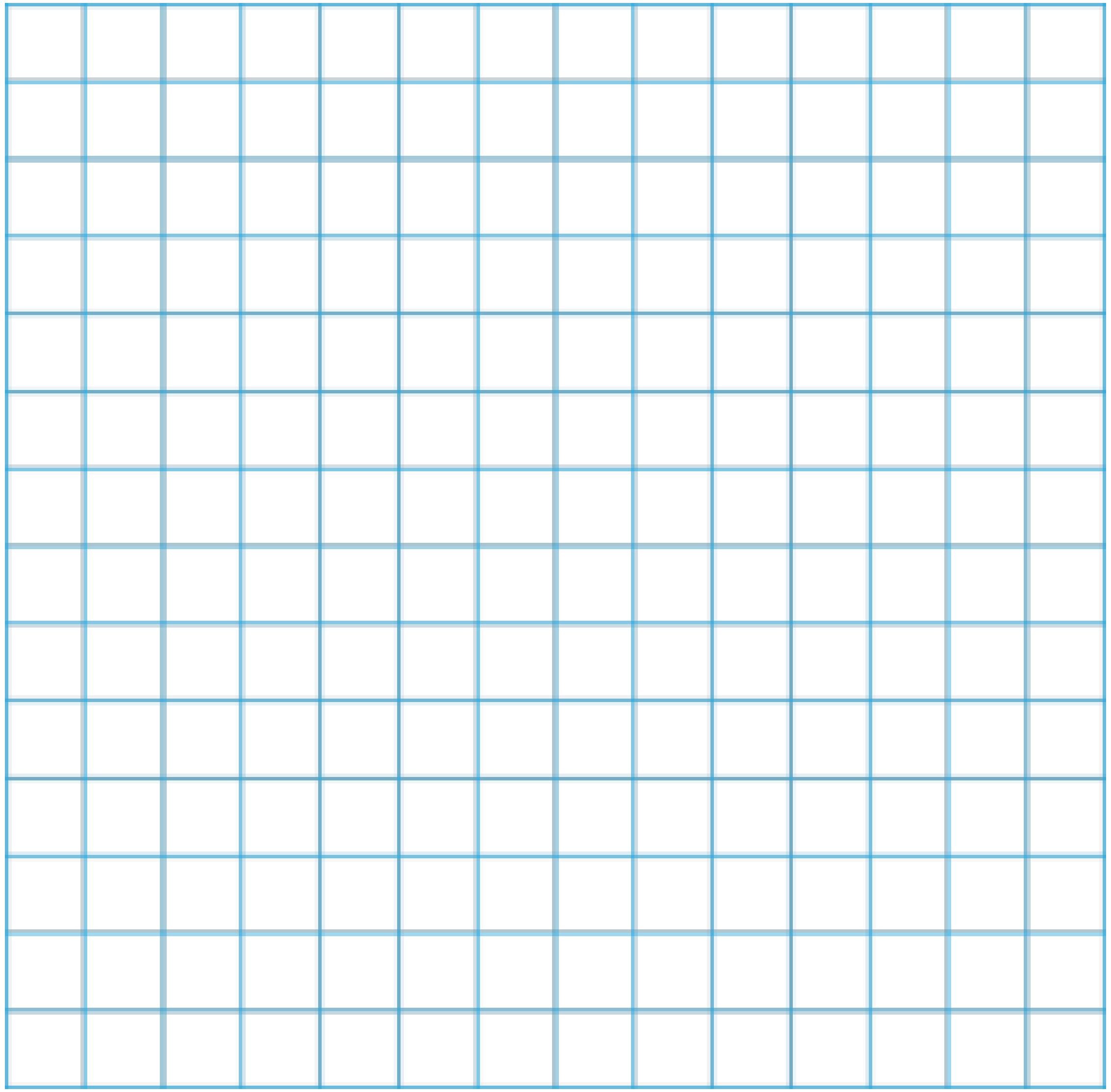
Linear Equations

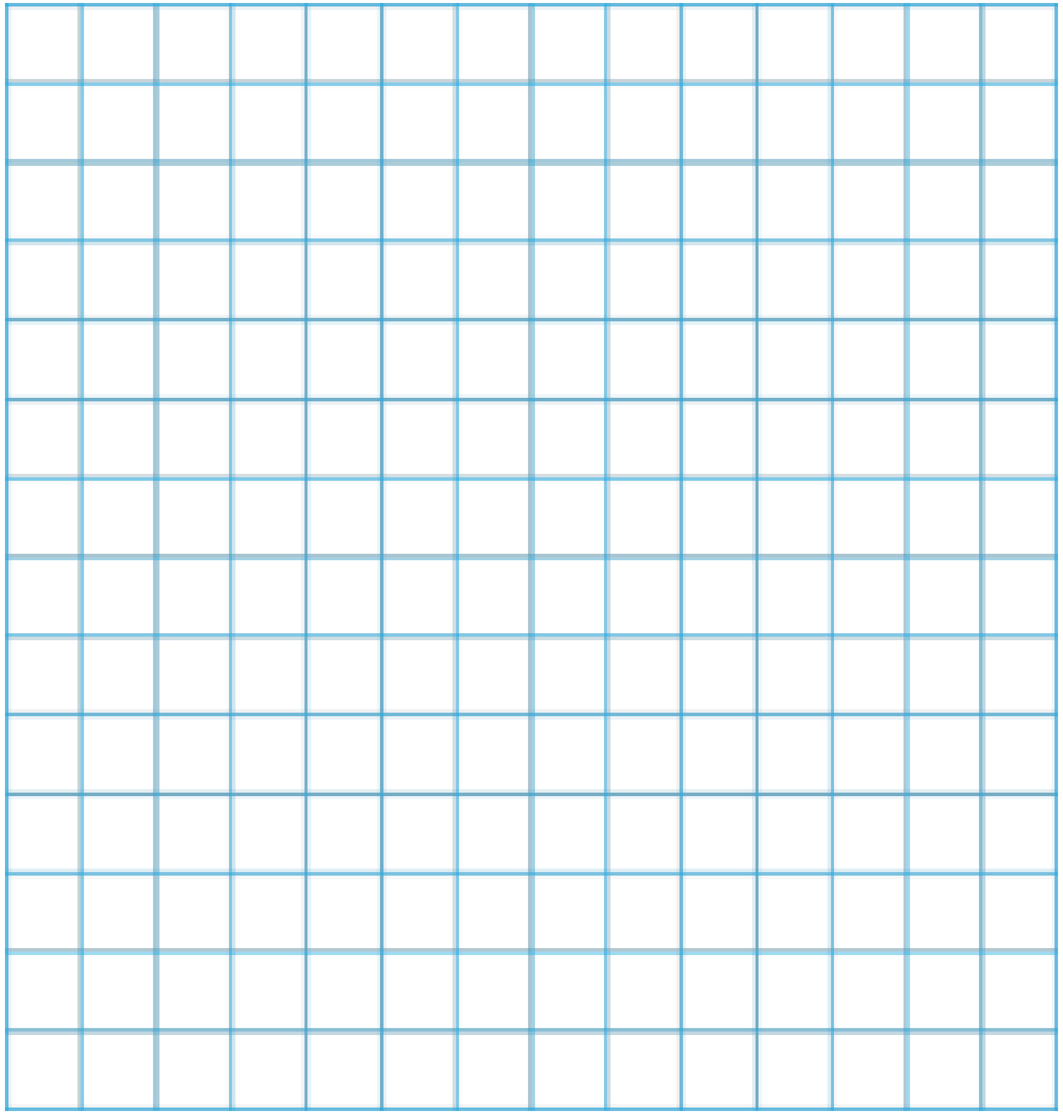
Equations	Linear or Non-Linear
$y = 8x - 9$	Linear
$y = x^2 - 7$	Non-Linear, the power of the variable x is 2
$\sqrt{y} + x = 6$	Non-Linear, the power of the variable y is 1/2
$y + 3x - 1 = 0$	Linear
$y^2 - x = 9$	Non-Linear, the power of the variable y is 2

Graphing Linear Equations

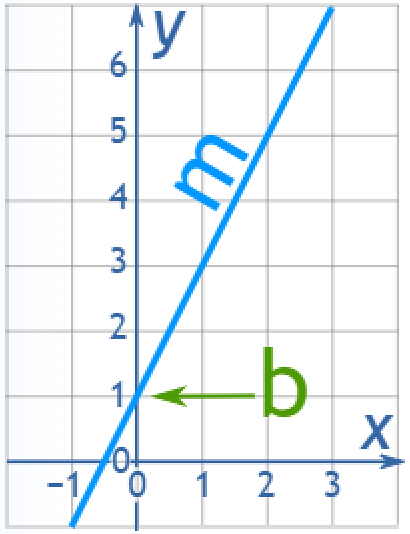
x	7	5	3
y	0	1	2







Equation of a Straight Line



$$y = mx + b$$

Slope or Gradient

y value when **x=0** *(see Y Intercept)*

y = how far up

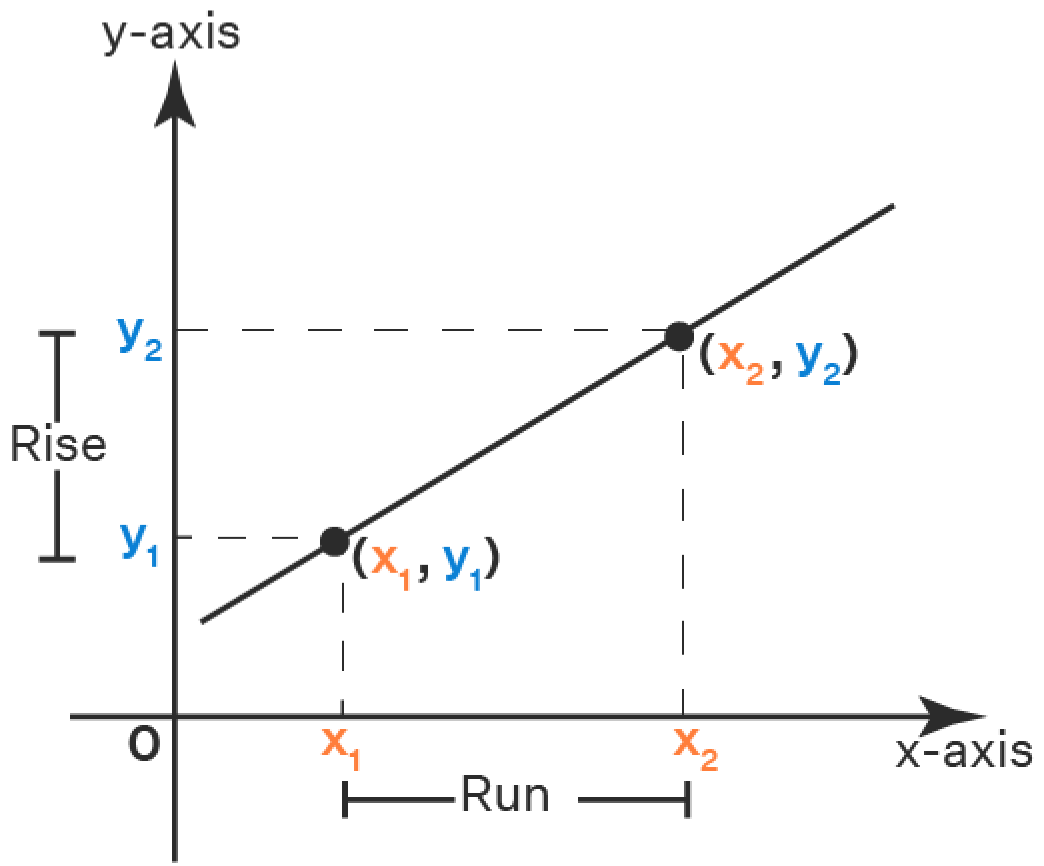
x = how far along

m = Slope or Gradient (how steep the line is)

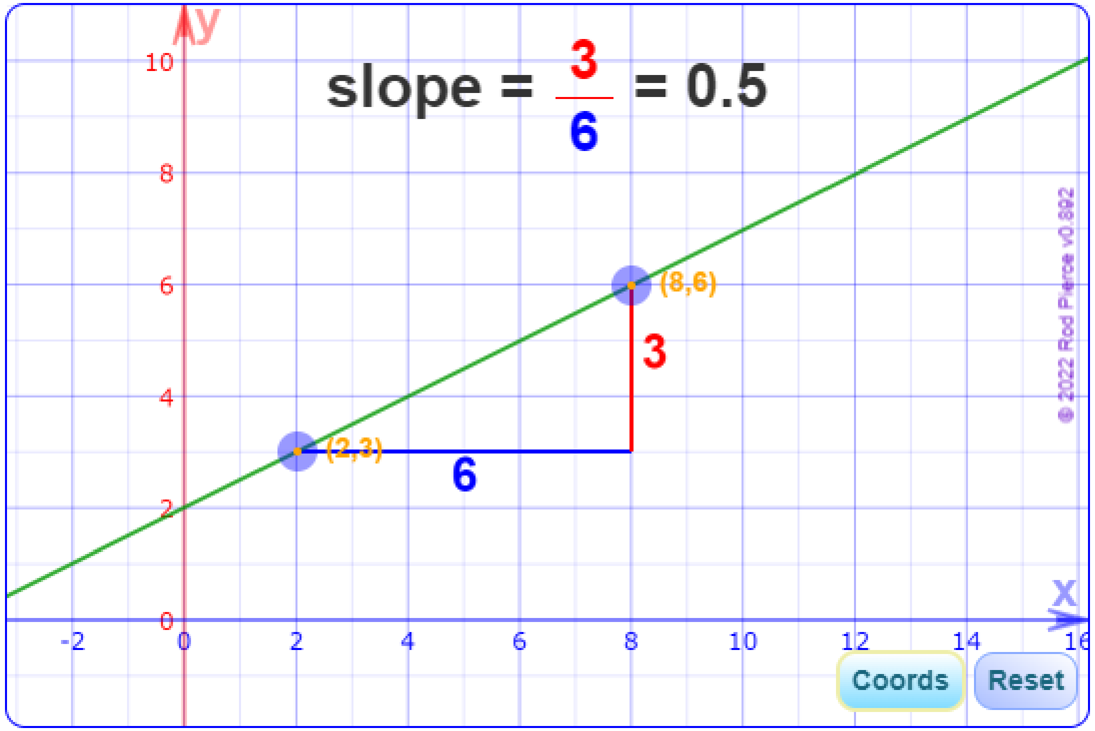
b = value of **y** when **x=0**

Rise Over Run

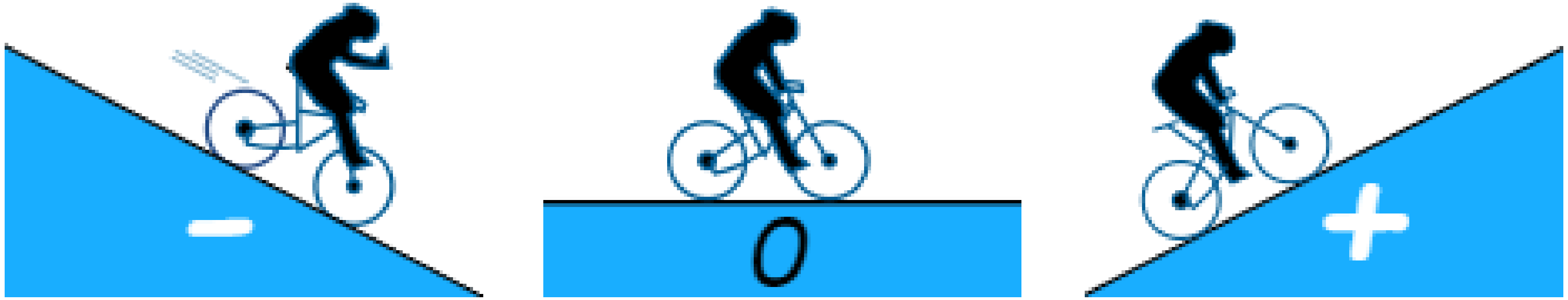
$$\text{Slope} = \frac{\text{Rise}}{\text{Run}} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$



Rise Over Run



Rise Over Run



Functions In Algebra

Functions in algebra are used to describe the operation being applied to an input in order to get an output.

Example The rule “multiply by 5, then add 3”

Function machine

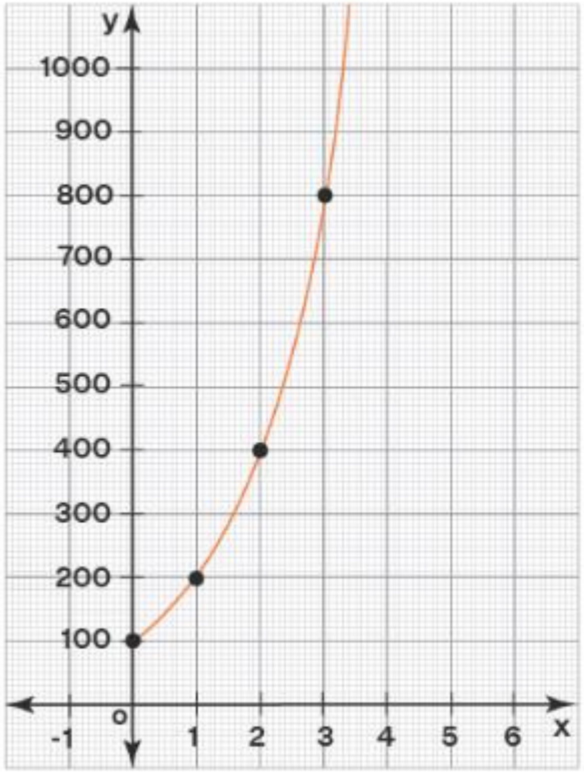
Input $\rightarrow \times 5 \rightarrow -3 \rightarrow$ Output

Function notation

$$f(x) = 5x - 3$$

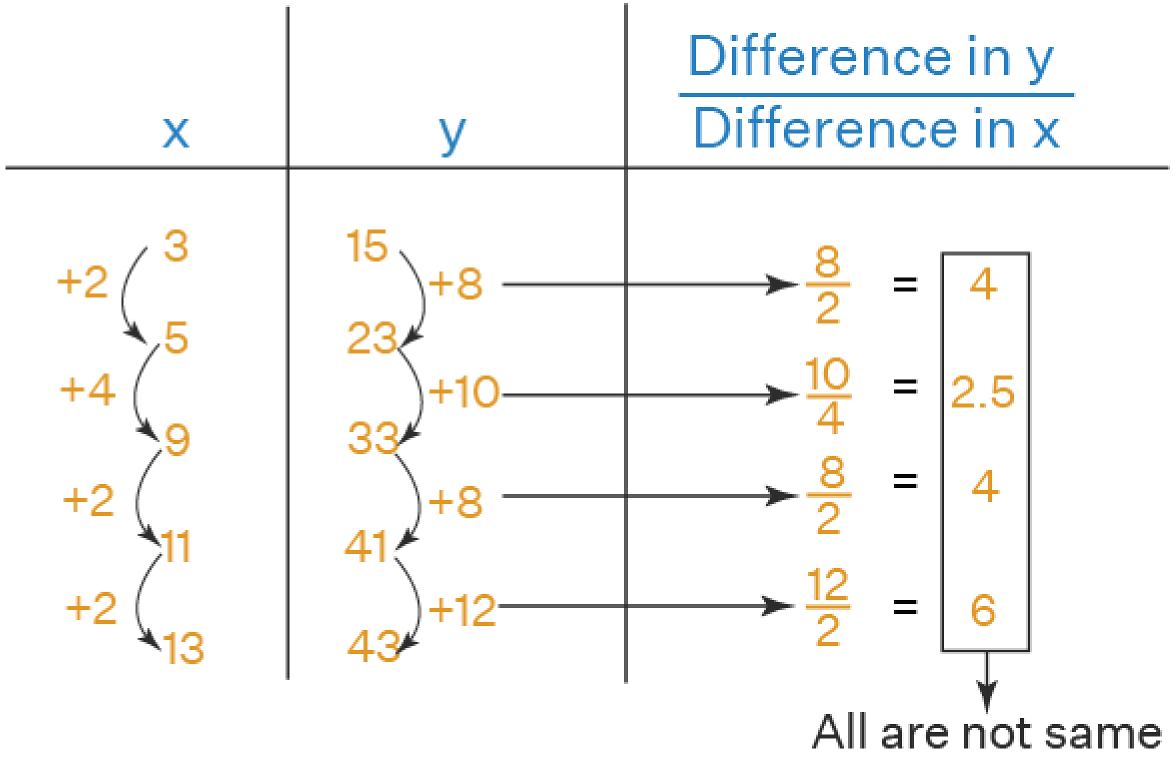
Functions In Algebra

x	y
0	100
1	200
2	400
3	800



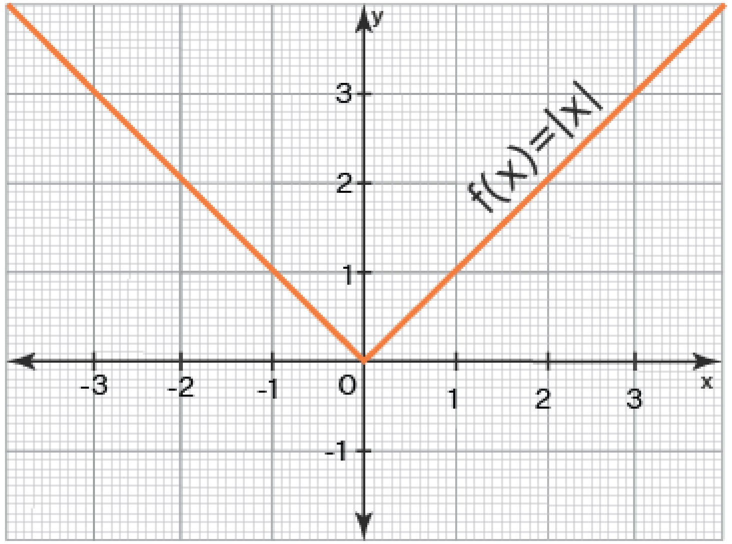
How to identify a Nonlinear Function?

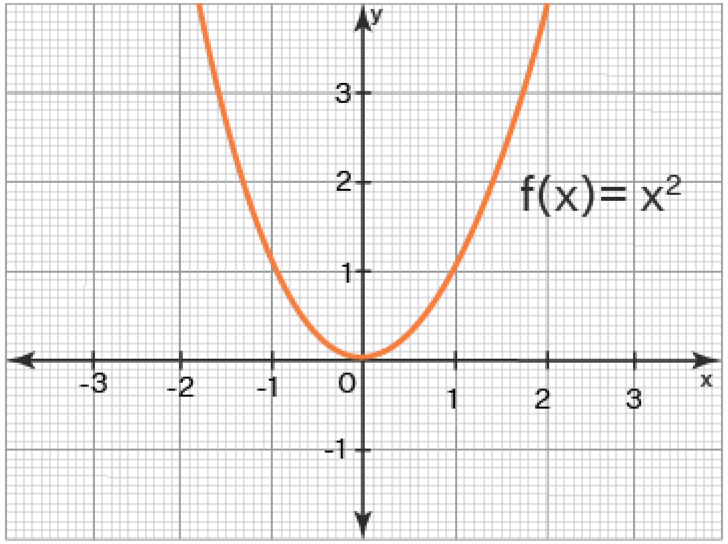
x	y
3	15
5	23
9	33
11	41
13	43

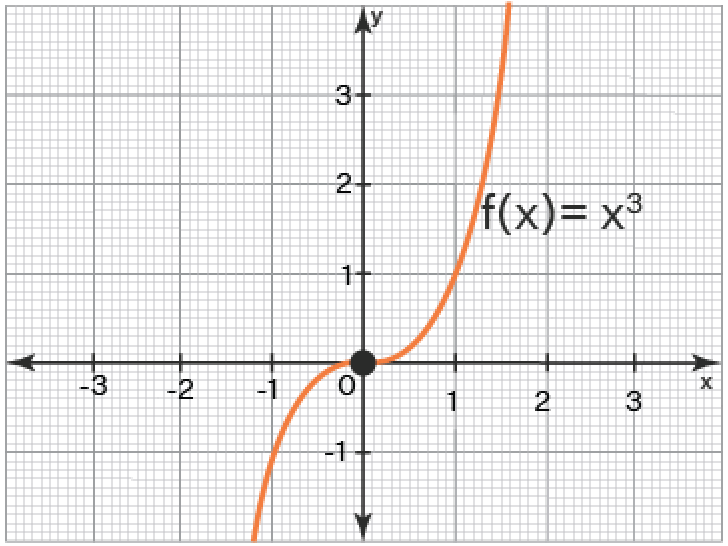


Hence, its Nonlinear

Graphs of Nonlinear Functions







General Form of System of Linear Equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots. + a_{1n}x_n + b_1 = 0$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots. + a_{2n}x_n + b_2 = 0$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots. + a_{3n}x_n + b_3 = 0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots. + a_{mn}x_n + b_m = 0$$

System of Linear Equations - Vector Form

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ . \\ . \\ . \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ . \\ . \\ . \\ a_{m2} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ . \\ . \\ . \\ a_{m3} \end{bmatrix} + \ldots\ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ . \\ . \\ . \\ a_{mn} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ . \\ . \\ . \\ b_m \end{bmatrix} = 0$$

System of Linear Equations - Matrix Form

$$A = \begin{bmatrix} a_{11} + a_{12} + a_{13} + \dots + a_{1n} \\ a_{21} + a_{22} + a_{23} + \dots + a_{2n} \\ a_{31} + a_{32} + a_{33} + \dots + a_{3n} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1} + a_{m2} + a_{m3} + \dots + a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$$

Coefficient matrix ***Variable matrix*** ***Constant matrix***

Representing systems of equations with matrices

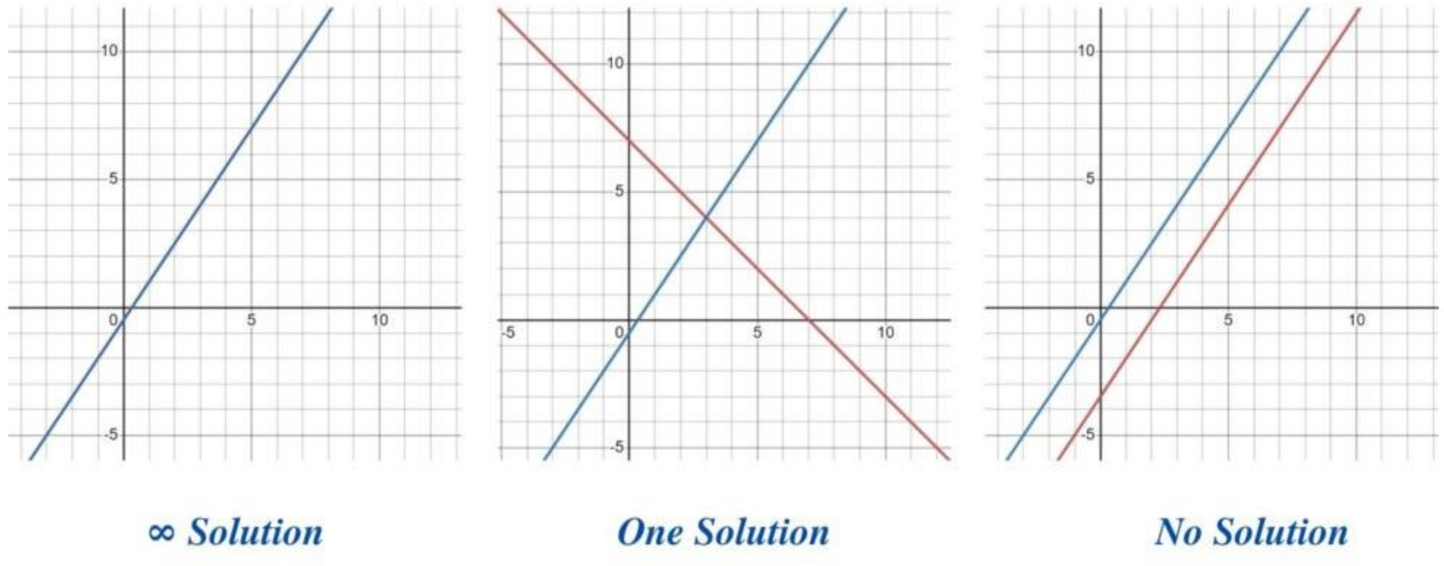
$$2.0x + 4.0y + 6.0z = 18$$
$$4.0x + 5.0y + 6.0z = 24$$
$$3.0x + 1y - 2.0z = 4$$

Matrix representation

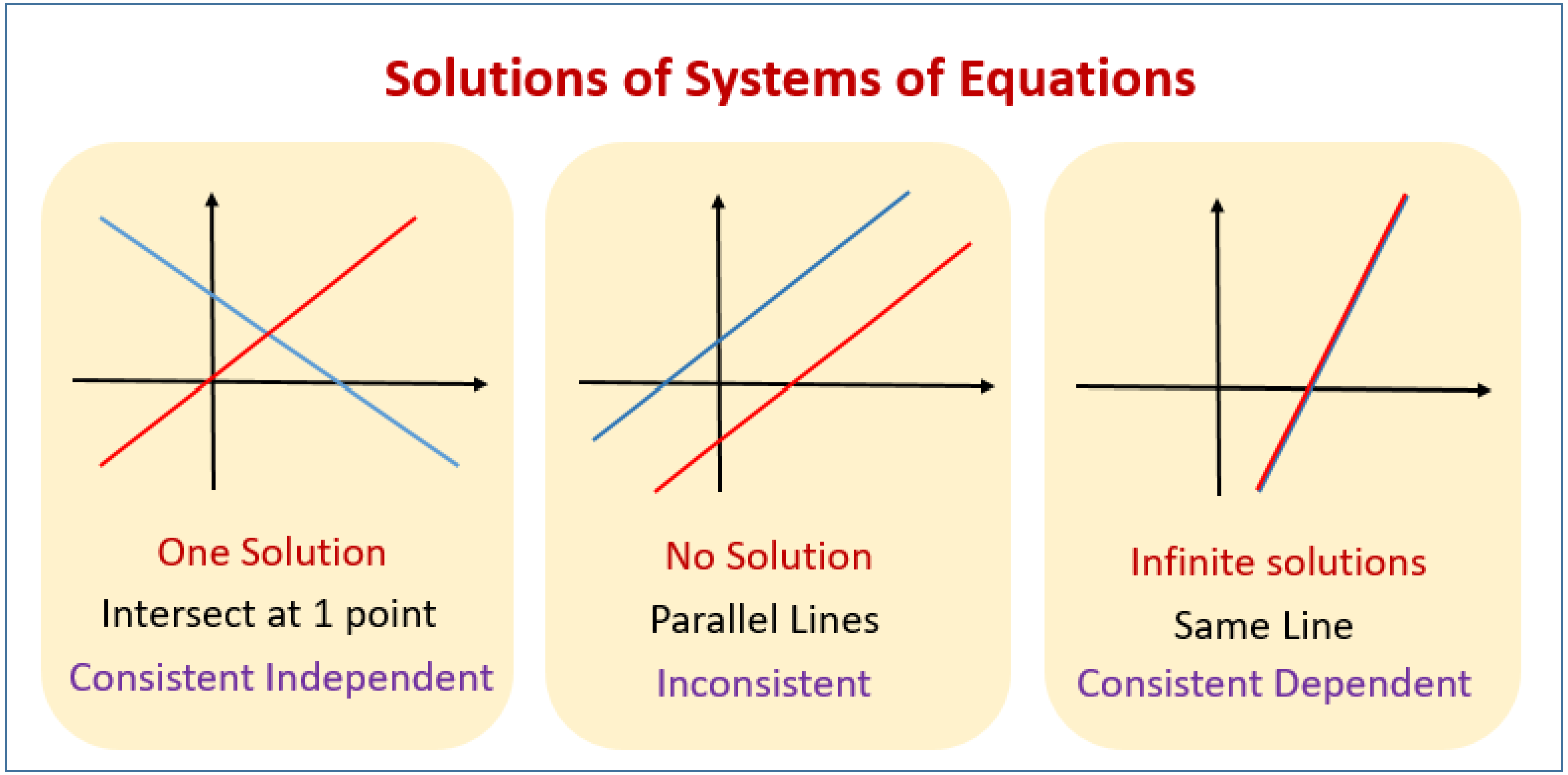
$$A = \begin{bmatrix} 2.0 & 4.0 & 6.0 \\ 4.0 & 5.0 & 6.0 \\ 3.0 & 1.0 & -2.0 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad b = \begin{bmatrix} 18.0 \\ 24.0 \\ 4.0 \end{bmatrix}$$

Types of Solutions of a System of Linear Equations

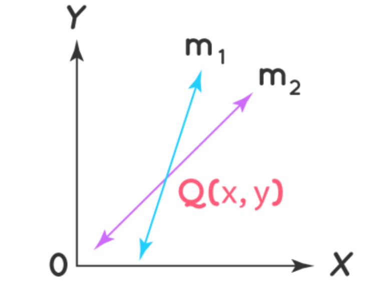
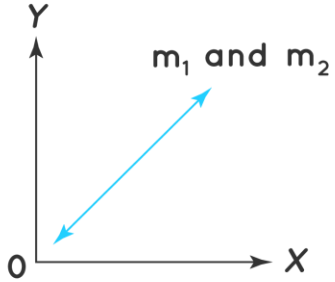
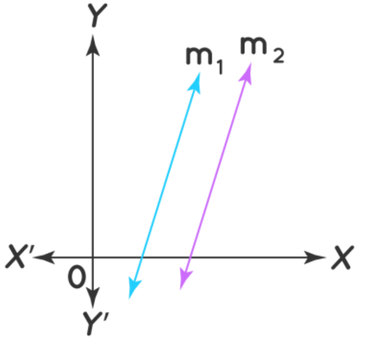
- Infinite solution: When lines are the same
- One solution or unique solution: When lines(graph) cross each other
- No solution: When lines are parallel



Solve a System of Linear Equations Using graphs



Number of Solutions of a Linear System of Equations			
Slopes	**Intercepts**	**Type of Lines**	**Number of Solutions**
Different		Intersecting	1 point
Same	Different	Parallel	No solution
Same	Same	Coincident	Infinitely many solutions

Given Condition	Graphical Representation	Algebraic Interpretation
$\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$	 Intersecting Lines	Exactly one solution (unique)
$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$	 Coincident Lines	Infinitely Many Solutions
$\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}$	 Parallel Lines	No solution

Number of Solutions of a Linear System of Equations			
Slopes	**Intercepts**	**Type of Lines**	**Number of Solutions**
Different		Intersecting	1 point
Same	Different	Parallel	No solution
Same	Same	Coincident	Infinitely many solutions

How to Solve a System of Linear Equations?

- There are many ways to solve a system of linear equations, here we are discussing three of them
- Using elimination
- Using substitution
- Using graphs
- Using matrices

Elimination Method

cuemath THE MATH EXPERT

$$x - 2y = 8 \longleftarrow (1)$$
$$2x + y = 5 \longleftarrow (2)$$

Multiply equation (2) by 2,

$$(2x + y = 5) \times 2$$
$$4x + 2y = 10$$

Add this equation to equation (1), we get,

$$\begin{array}{rl} & x - \cancel{2y} = 8 \\ + & 4x + \cancel{2y} = 10 \\ \hline & 5x = 18 \end{array}$$

$$\boxed{x = \frac{18}{5}}$$

Put this value in equation (2),

$$y = 5 - 2x$$
$$y = 5 - 2 \times \frac{18}{5}$$
$$y = 5 - \frac{36}{5}$$
$$\boxed{y = -\frac{11}{5}}$$

$$\boxed{\therefore\ x = \frac{18}{5} \text{ and } y = -\frac{11}{5}}$$

Elimation method

$$\begin{cases} 3x + y = 10 \\ -4x - 2y = 2 \end{cases}$$

$$2 \times (3x + y = 10)$$
$$-4x - 2y = 2$$

$$6x + 2y = 20$$
$$-4x - 2y = 2$$

$$6x + \cancel{2y} = 20$$
$$-4x - \cancel{2y} = 2$$

$$2x = 22$$

$$x = 11$$

$$3(11) + y = 10$$

$$33 + y = 10$$

$$y = 10 - 33$$

$$y = -23$$

Elimination Method with Three Equations

$$3x - y + 2z = 5 \longleftarrow (1)$$

$$4x + 2y - z = 6 \longleftarrow (2)$$

$$5x - 3y + z = 1 \longleftarrow (3)$$

Add equation (2) & (3) to eliminate z,

$$\begin{array}{rcl} 4x + 2y - \cancel{z} & = & 6 \\ 5x - 3y + \cancel{z} & = & 1 \\ \hline 9x - y & = & 7 \longleftarrow (P) \end{array}$$

Multiply equation (3) by 2,

$$10x - 6y + 2z = 2$$

Subtract equation (1) from this equation,

$$\begin{array}{rcl} 10x - 6y + \cancel{2z} & = & 2 \\ 3x - y + \cancel{2z} & = & 5 \\ - \quad + \quad - & & - \\ \hline 7x - 5y & = & -3 \longleftarrow (Q) \end{array}$$

Solve equation (P) & (Q) using elimination,

$$(9x - y = 7) \times (-5) \longrightarrow \begin{array}{rcl} -45x + \cancel{5y} & = & -35 \\ 7x - \cancel{5y} & = & -3 \\ \hline -38x & = & -38 \end{array}$$

$$\boxed{x = 1}$$

Using substitution

$$x - 2y = 8 \longleftarrow (1)$$

$$x + y = 5 \longleftarrow (2)$$

From equation (2),

$$x = 5 - y$$

Substitute this value in equation (1),

$$x - 2y = 8$$

$$5 - y - 2y = 8$$

$$5 - 3y = 8$$

$$-3y = 8 - 5$$

$$-3y = 3$$

$$y = -1$$

Put this value in equation (2),

$$x - 1 = 5$$

$$x = 5 + 1$$

$$x = 6$$

$$\therefore\ x = 6 \text{ and } y = -1$$

Solve a system of equations using matrices

Here is a visual to show the order for getting the 1's and 0's in the proper position for row-echelon form.

2 × 3 matrix

Step 1 $\left[\begin{array}{cc|c} 1 & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare \end{array}\right]$ **Step 2** $\left[\begin{array}{cc|c} 1 & \blacksquare & \blacksquare \\ 0 & \blacksquare & \blacksquare \end{array}\right]$ **Step 3** $\left[\begin{array}{cc|c} 1 & \blacksquare & \blacksquare \\ 0 & 1 & \blacksquare \end{array}\right]$

3 × 4 matrix

Step 1 $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{array}\right]$ **Step 2** $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ 0 & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{array}\right]$ **Step 3** $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ 0 & \blacksquare & \blacksquare & \blacksquare \\ 0 & \blacksquare & \blacksquare & \blacksquare \end{array}\right]$ **Step 4** $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ 0 & 1 & \blacksquare & \blacksquare \\ 0 & \blacksquare & \blacksquare & \blacksquare \end{array}\right]$

Step 5 $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ 0 & 1 & \blacksquare & \blacksquare \\ 0 & 0 & \blacksquare & \blacksquare \end{array}\right]$ **Step 6** $\left[\begin{array}{ccc|c} 1 & \blacksquare & \blacksquare & \blacksquare \\ 0 & 1 & \blacksquare & \blacksquare \\ 0 & 0 & 1 & \blacksquare \end{array}\right]$

We use the same procedure when the system of equations has three equations.

How to Solve a System of Equations Using a Matrix

Solve the system of equations using a matrix: $\begin{cases} 3x + 4y = 5 \\ x + 2y = 1 \end{cases}$.

Step 1. Write the augmented matrix for the system of equations.		$\begin{cases} 3x + 4y = 5 \\ x + 2y = 1 \end{cases}$ $\left[\begin{array}{cc\|c} 3 & 4 & 5 \\ 1 & 2 & 1 \end{array}\right]$
Step 2. Using row operations get the entry in row 1, column 1 to be 1.	Interchange the rows, so 1 will be in row 1, column 1.	$\begin{array}{c} R_2 \\ R_1 \end{array} \left[\begin{array}{cc\|c} 1 & 2 & 1 \\ 3 & 4 & 5 \end{array}\right]$
Step 3. Using row operations, get zeros in column 1 below the 1.	Multiply row 1 by –3 and add it to row 2.	$\begin{array}{c} \\ -3R_1 + R_2 \end{array} \left[\begin{array}{cc\|c} 1 & 2 & 1 \\ 0 & -2 & 2 \end{array}\right]$
Step 4. Using row operations, get the entry in row 2, column 2 to be 1.	Multiply row 2 by $-\frac{1}{2}$.	$\begin{array}{c} \\ -\frac{1}{2}R_2 \end{array} \left[\begin{array}{cc\|c} 1 & 2 & 1 \\ 0 & 1 & -1 \end{array}\right]$
Step 5. Continue the process until the matrix is in row-echelon form.	The matrix is now in row-echelon form.	$\left[\begin{array}{cc\|c} 1 & 2 & 1 \\ 0 & 1 & -1 \end{array}\right]$
Step 6. Write the corresponding system of equations.	$\begin{array}{ccc} x & y & \\ \end{array}$ $\left[\begin{array}{ccc} 1 & 2 & 1 \\ 0 & 1 & -1 \end{array}\right]$	$\begin{cases} x + 2y = 1 \\ y = -1 \end{cases}$
Step 7. Use substitution to find the remaining variables.	Substitute $y = -1$ into $x + 2y = 1$.	$y = -1$ $x + 2y = 1$ $x + 2(-1) = 1$ $x - 2 = 1$ $x = 3$

Solve the system of equations using a matrix: $\begin{cases} 3x+8y+2z=-5 \\ 2x+5y-3z=0 \\ x+2y-2z=-1 \end{cases}$.

$3x+8y+2z=-5$

Using row operations, get zeros in column 1 below the 1. $\begin{matrix} \\ -2R_1+R_2 \end{matrix} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -1 \\ 0 & 1 & 1 & 2 \end{array}\right]$

Write the augmented matrix for the equations.

Interchange row 1 and 3 to get the entry in row 1, column 1 to be 1.

The entry in row 2, colum

Continue the process unti

is in row-echelon forı

The matrix is now in row-echelon form. $\left[\begin{array}{ccc|c} 1 & 2 & -2 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right]$

Write the corresponding system of equations. $\begin{cases} x+2y-2z=-1 \\ \quad\; y+z=2 \\ \qquad\;\; z=-1 \end{cases}$

Use substitution to find the remaining variables.

$$y+z=2$$
$$y+(-1)=2$$
$$y=3$$

$$x+2y-2z=-1$$
$$x+2(3)-2(-1)=-1$$
$$x+6+2=-1$$

$$x=-9$$

Write the solution as an ordered pair or triple. $(-9, 3, -1)$

Check that the solution makes the original equations true. We leave the check for you.

- If $|A| \neq 0$ then the system is consistent and has a unique solution then $X = A^{-1}B$ will give the solution where A is the coefficient matrix, X is the variable matrix, and B is the constant matrix.

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}$$

$$AX = B$$

$$A^{-1}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix}^{-1} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix}$$

Then multiply A^{-1} by B (we can use the Matrix Calculator again):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}$$

$$= \frac{1}{-21} \begin{bmatrix} -105 \\ -63 \\ 42 \end{bmatrix}$$

$$= \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$$

And we are done! The solution is:

$x = 5$

$y = 3$

$z = -2$

For example, consider the system $2x + y + z = 7$, $3x + 2y + 3z = 16$, and $x + 3y + 2z = 13$, Here we get matrices A, B, and X as follows

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad B = \begin{bmatrix} 7 \\ 16 \\ 13 \end{bmatrix}$$

We know $X = A^{-1}B$, so we need to find A^{-1}

$$A^{-1} = \frac{1}{6} \times \begin{bmatrix} 5 & -1 & -1 \\ 3 & -3 & 3 \\ -7 & 5 & -1 \end{bmatrix}$$

Now we get

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{6} \times \begin{bmatrix} 5 & -1 & -1 \\ 3 & -3 & 3 \\ -7 & 5 & -1 \end{bmatrix} \begin{bmatrix} 7 \\ 16 \\ 13 \end{bmatrix} = \frac{1}{6} \times \begin{bmatrix} 5 \times 7 - 1 \times 16 - 1 \times 13 \\ 3 \times 7 - 3 \times 16 + 3 \times 13 \\ -7 \times 7 + 5 \times 16 - 1 \times 13 \end{bmatrix}$$

Simplify this, then

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

So $x = 1$, $y = 2$ and $z = 3$

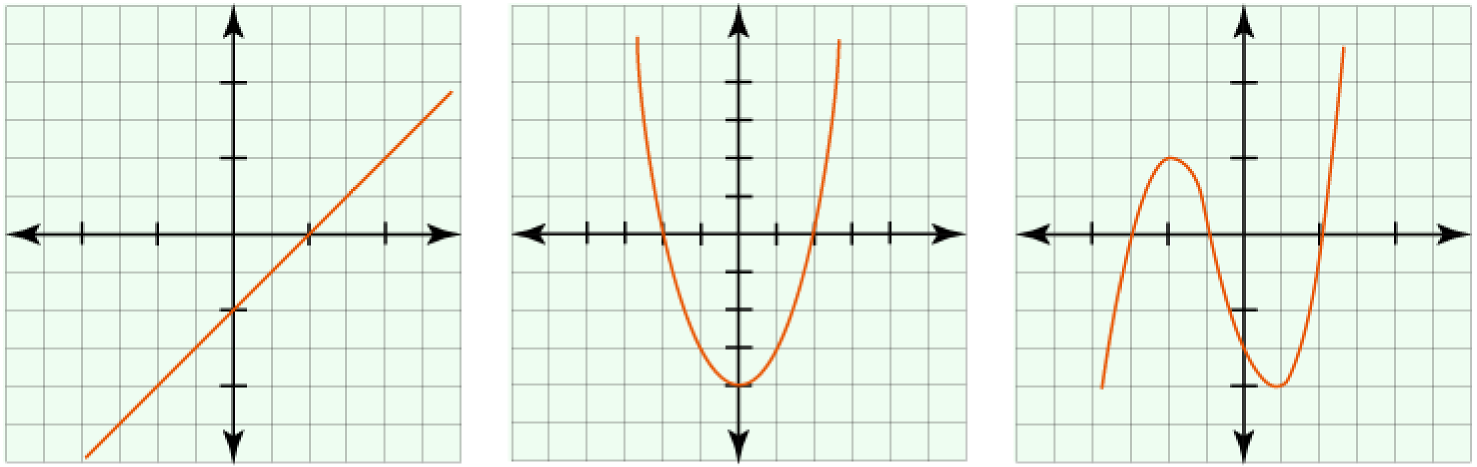
polynomial

Polynomials	Form	Degree	Examples
Linear Polynomial	$p(x): ax+b, a \neq 0$	Polynomial with Degree 1	$x + 8$
Quadratic Polynomial	$p(x): ax^2+b+c,$ $a \neq 0$	Polynomial with Degree 2	$3x^2-4x+7$
Cubic Polynomial	$p(x): ax^3+bx^2+cx,$ $a \neq 0$	Polynomial with Degree 3	$2x^3+3x^2+4x+6$

polynomial

Graphs of Polynomial Functions

cuemath THE MATH EXPERT



$y = ax + b$ $\quad$ $y = ax^2 + bx + c$ $\quad$ $y = ax^3 + bx^2 + cx + d$

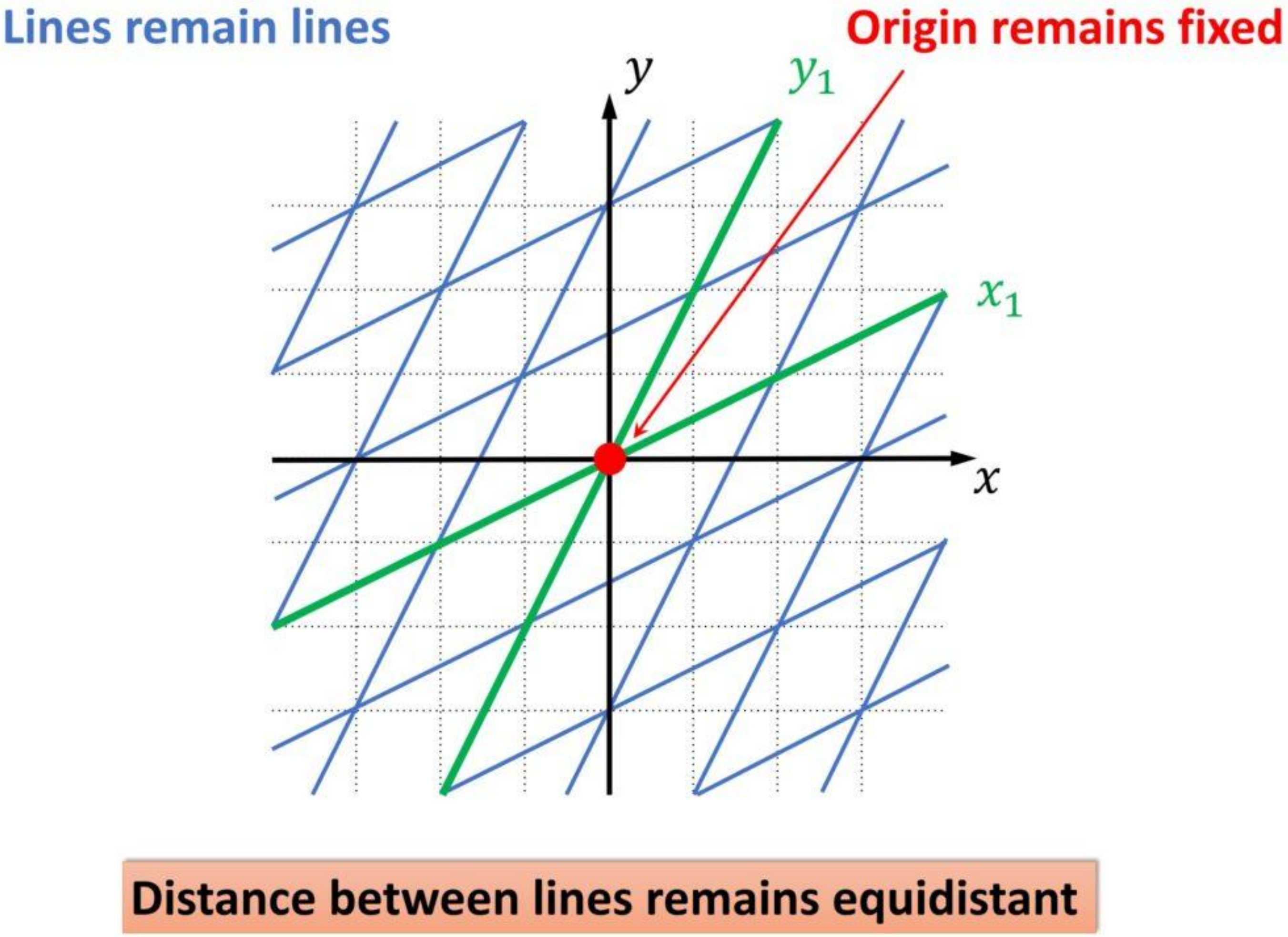
linear transformations

a linear transformation is actually a function that maps an input vector into an output vector

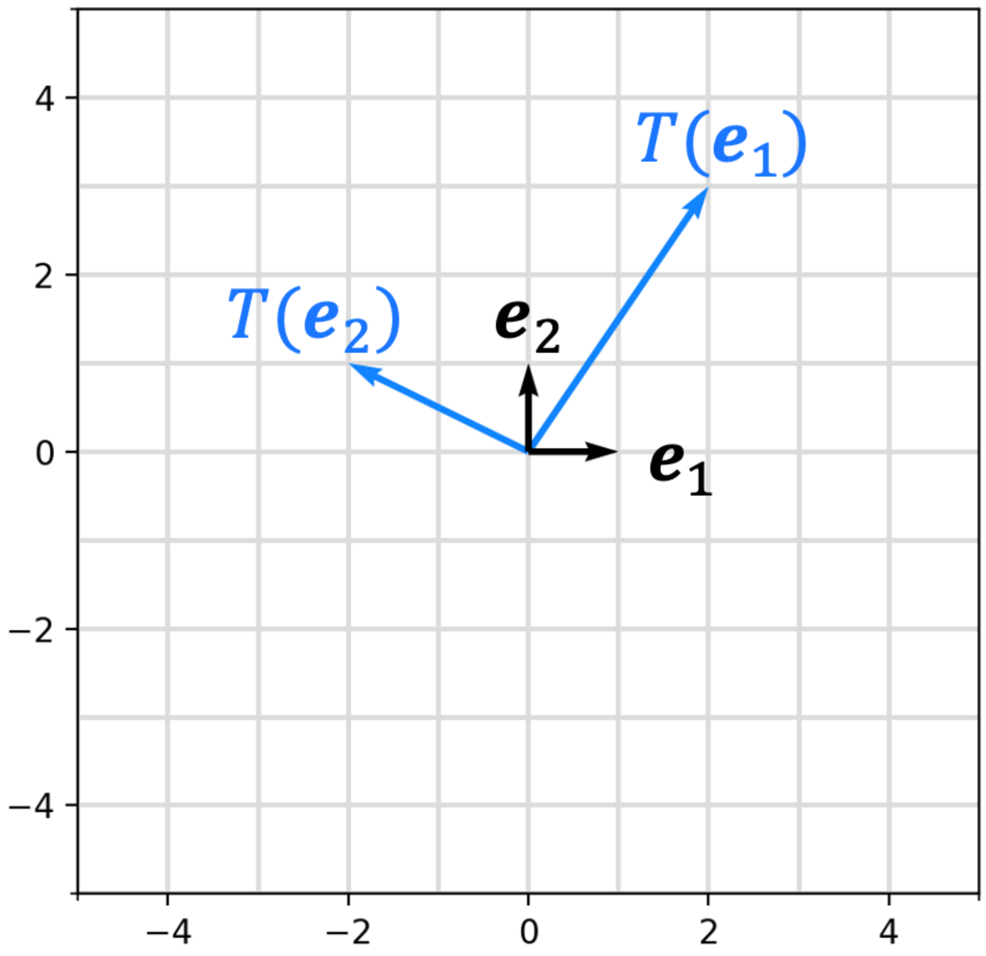
$$\begin{bmatrix}1\\2\\3\end{bmatrix} \qquad f(\vec{v}) \qquad \begin{bmatrix}2\\4\\6\end{bmatrix}$$

Input Vector Output vector

Matrices linear transformations



Matrices linear transformations

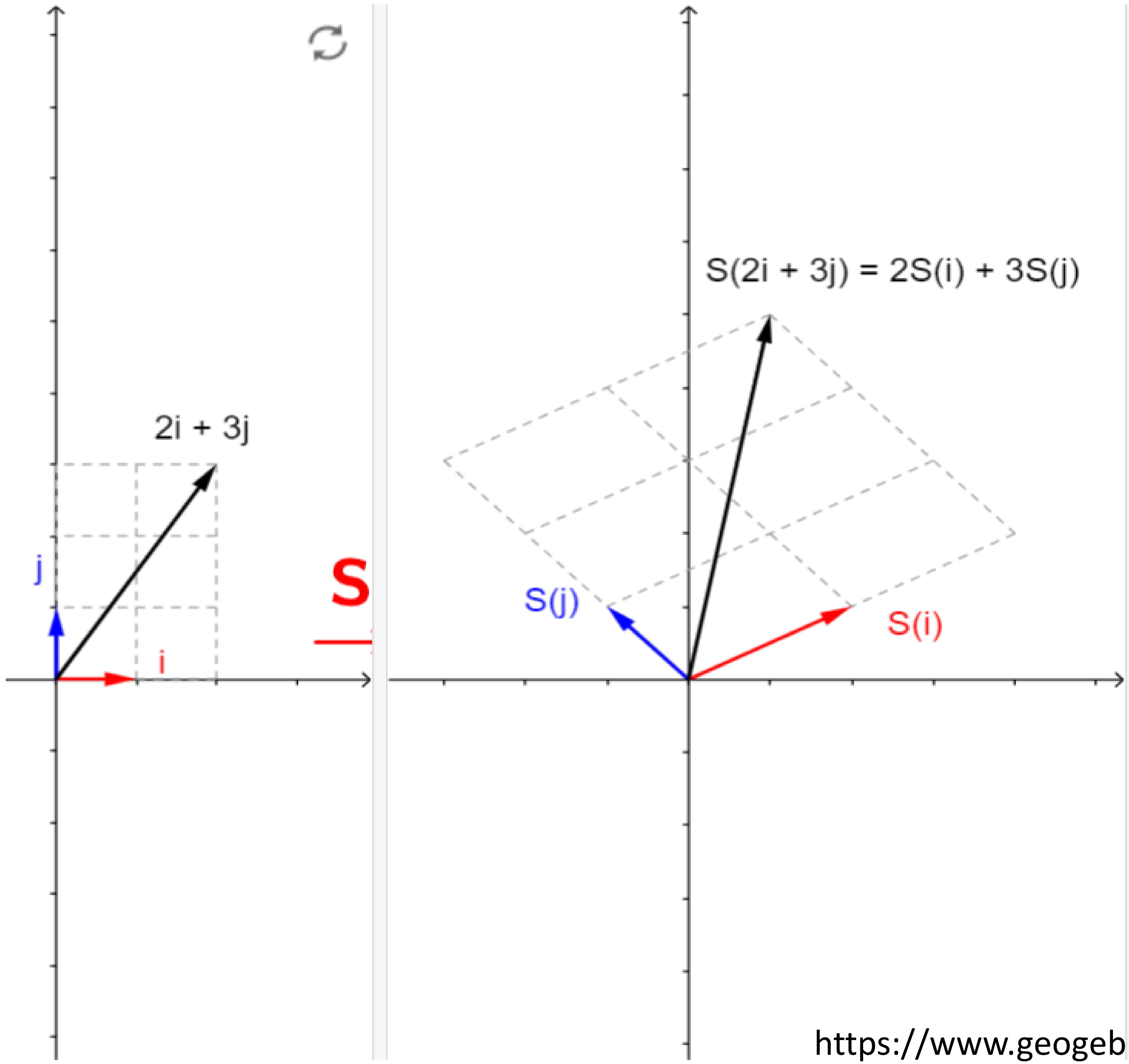


$TA = B$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

T - Transformation Matrix

every linear transformation from vectors to vectors is a matrix multiplication



$$S\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}$$

$$S\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$$

$$\begin{aligned} S\begin{pmatrix}2\\3\end{pmatrix} &= 2\,S\begin{pmatrix}1\\0\end{pmatrix} + 3\,S\begin{pmatrix}0\\1\end{pmatrix} \\ &= 2\begin{pmatrix}2\\1\end{pmatrix} + 3\begin{pmatrix}-1\\1\end{pmatrix} \\ &= \begin{pmatrix}1\\5\end{pmatrix} \end{aligned}$$

In general,

$$\begin{aligned} S\begin{pmatrix}x\\y\end{pmatrix} &= x\,S\begin{pmatrix}1\\0\end{pmatrix} + y\,S\begin{pmatrix}0\\1\end{pmatrix} \\ &= x\begin{pmatrix}2\\1\end{pmatrix} + y\begin{pmatrix}-1\\1\end{pmatrix} \\ &= \begin{pmatrix}2 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} \end{aligned}$$

https://www.geogebra.org/m/kGJZH7AU

linear transformations

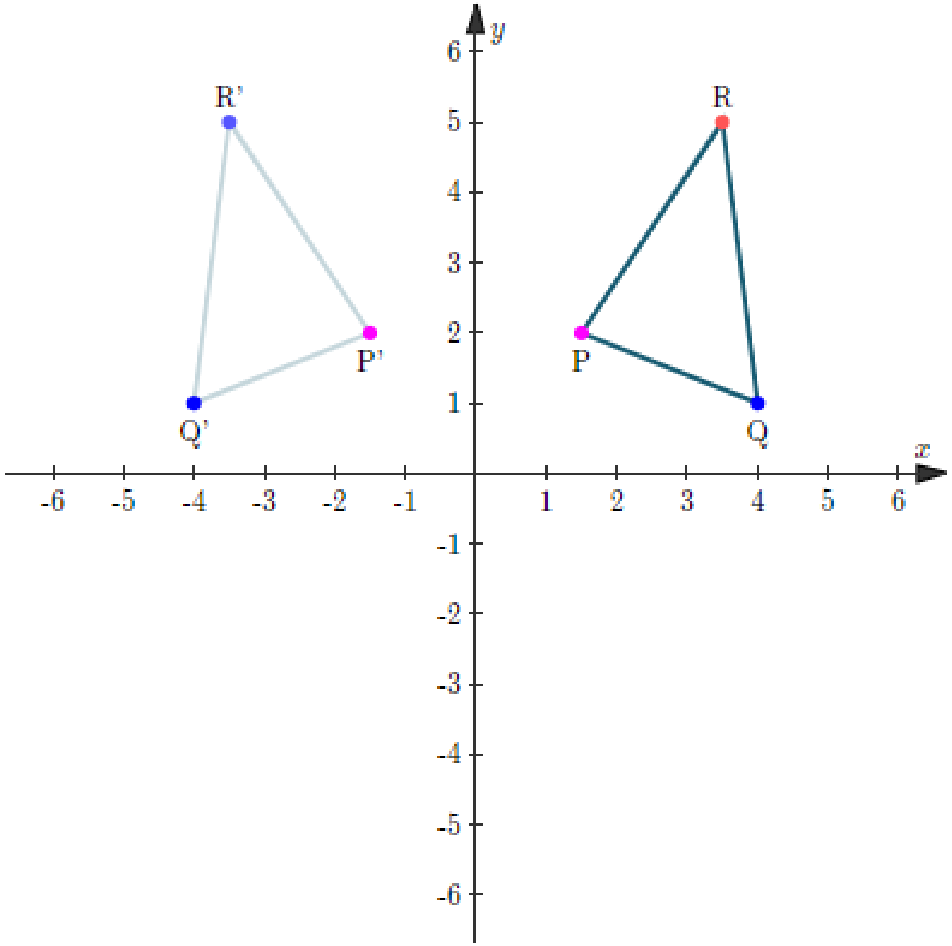
We can represent the point $\mathrm{P}(1.5, 2)$ as a **column vector** $\begin{bmatrix} 1.5 \\ 2 \end{bmatrix}$.

Similarly, we can represent two points $\mathrm{P}(1.5, 2)$ and $\mathrm{Q}(4, 1)$ as a 2×2 vector $\begin{bmatrix} 1.5 & 4 \\ 2 & 1 \end{bmatrix}$.

In most of the examples on this page, we have a triangle PQR, where R is $(3.5, 5)$, and we represent the vertices of the triangle PQR as $\begin{bmatrix} 1.5 & 4 & 3.5 \\ 2 & 1 & 5 \end{bmatrix}$.

$$\mathbf{Av} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1.5 & 4 & 3.5 \\ 2 & 1 & 5 \end{pmatrix}$$

linear transformations



Reflection in the y-axis

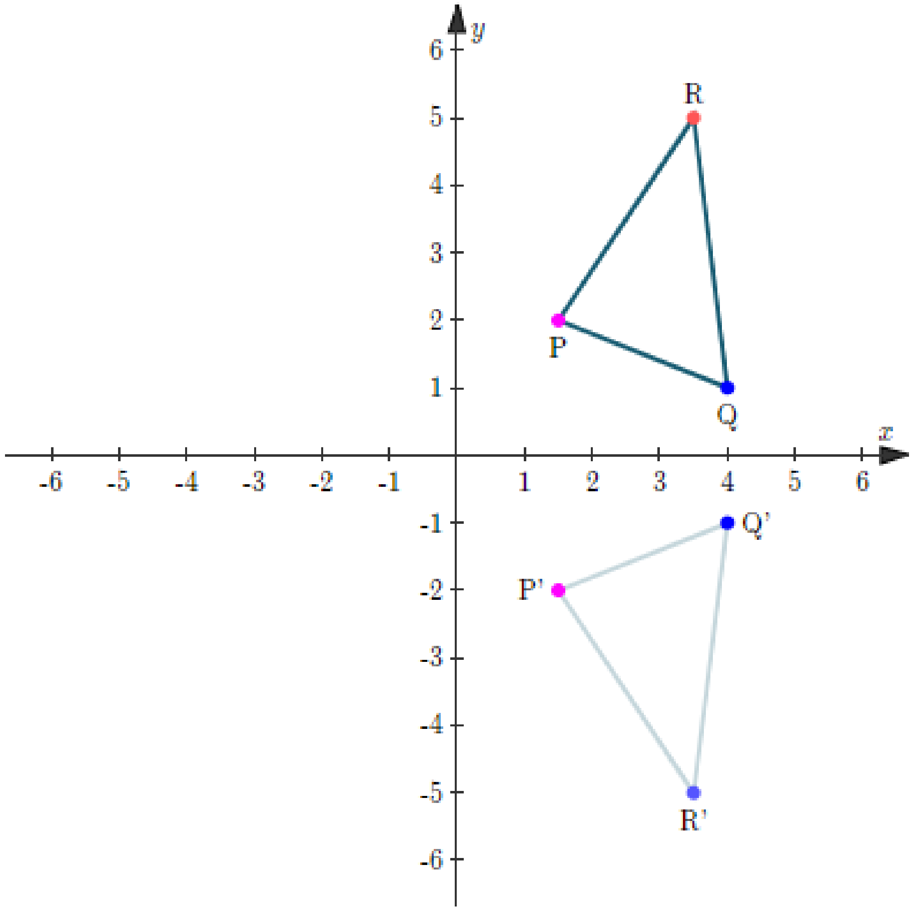
Matrix $\mathbf{A}$ in this example has the effect of producing the negative of each x-value in matrix $\mathbf{v}$, producing the resulting vector $\mathbf{v}_{\mathrm{R'}}$ that represents triangle $\mathrm{P'Q'R'}$.

$$\mathbf{Av} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1.5 & 4 & 3.5 \\ 2 & 1 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} -1.50 & -4.00 & -3.50 \\ 2.00 & 1.00 & 5.00 \end{pmatrix}$$

$$= \mathbf{v}_R$$

linear transformations



Reflection in the x-axis

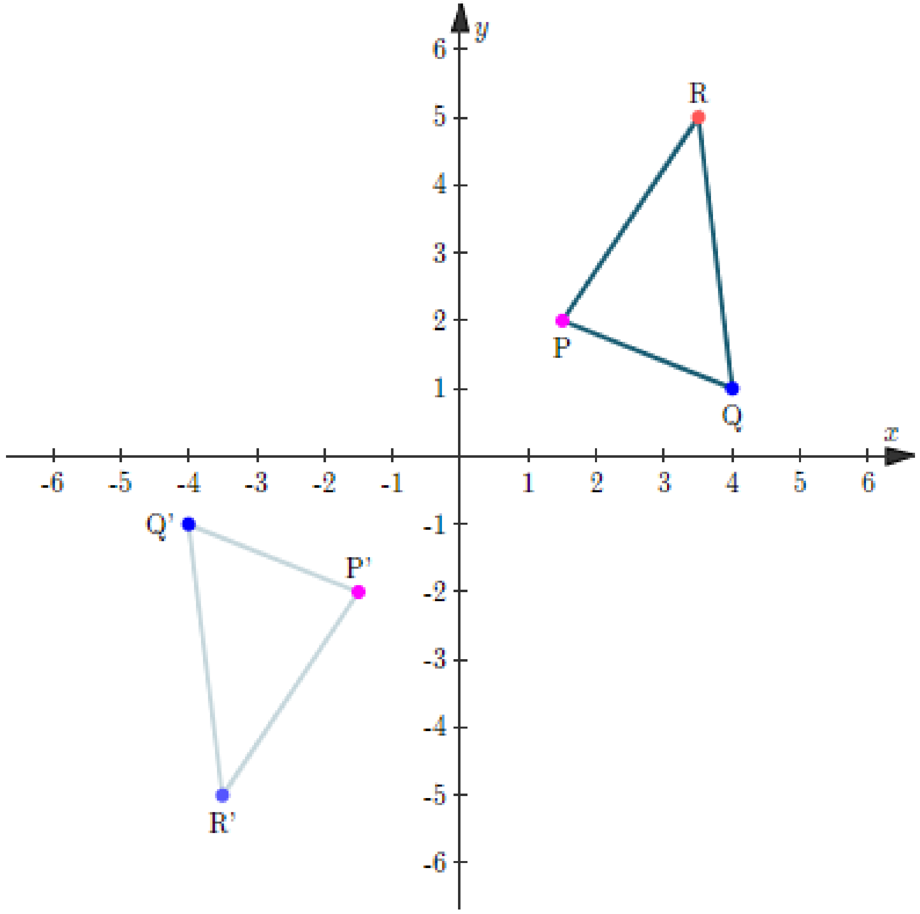
Matrix $\mathbf{A}$ in this example has the effect of producing the negative of each y-value in matrix $\mathbf{v}$, producing $\mathbf{v}_{\mathrm{R}}$.

$$\mathbf{A}\mathbf{v} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1.5 & 4 & 3.5 \\ 2 & 1 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 1.50 & 4.00 & 3.50 \\ -2.00 & -1.00 & -5.00 \end{pmatrix}$$

$$= \mathbf{v}_R$$

Matrices linear transformations



Reflection in the origin

Matrix $\mathbf{A}$ has the effect of producing the negative of each x-value and y-value in matrix $\mathbf{v}$, producing $\mathbf{v}_{\mathrm{R}}$.

$$\mathbf{Av} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1.5 & 4 & 3.5 \\ 2 & 1 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} -1.50 & -4.00 & -3.50 \\ -2.00 & -1.00 & -5.00 \end{pmatrix}$$

$$= \mathbf{v}_R$$

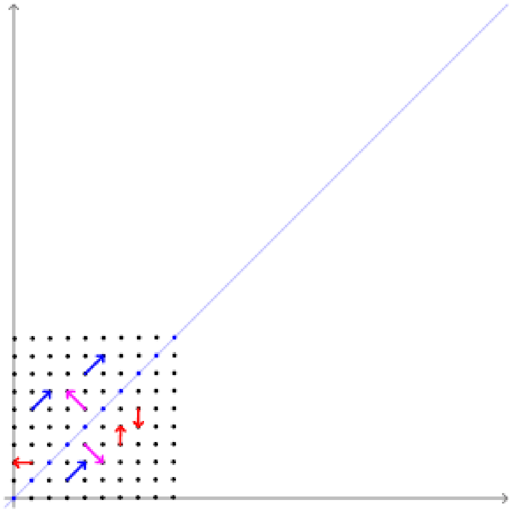
Matrices linear transformations

https://www.intmath.com/matrices-determinants/matrices-linear-transformations.php

https://web.ma.utexas.edu/users/ysulyma/matrix/

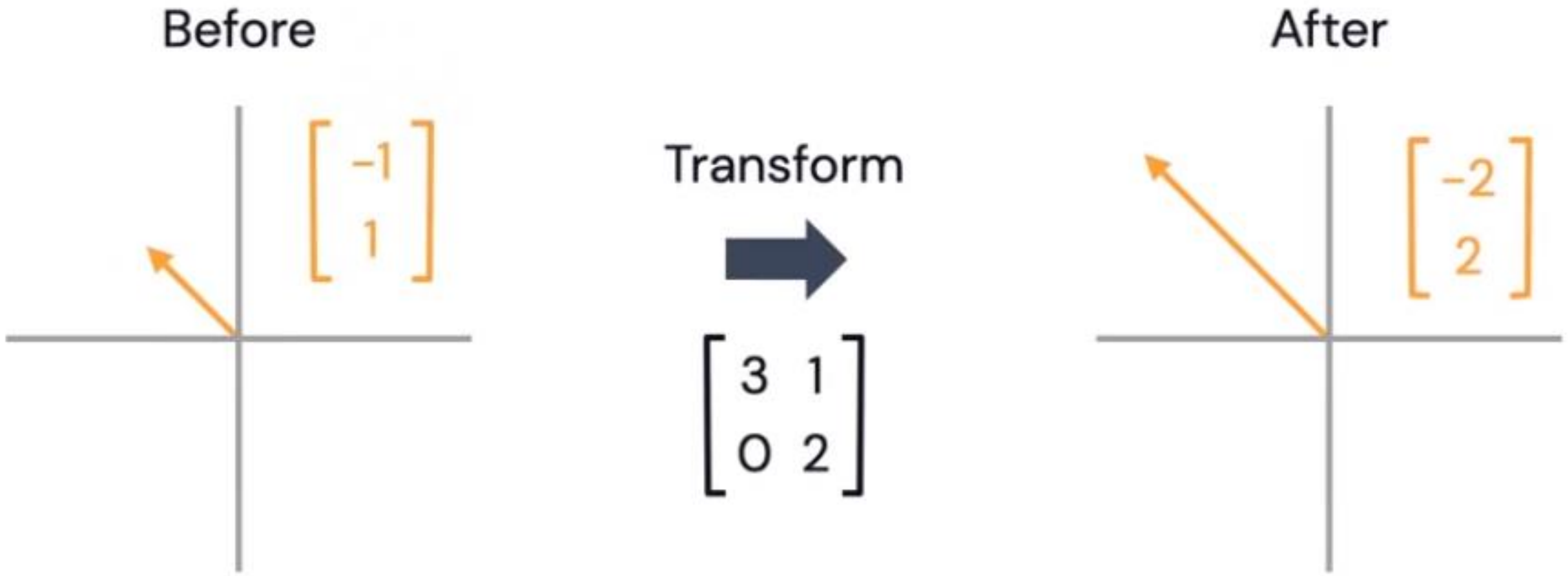
Eigenvalues and eigenvectors

- The **eigenvectors** for a **linear transformation** matrix are the set of vectors that are only stretched, with no rotation or shear.
- The eigenvalue is the **factor** by which an eigenvector is **stretched/scaled**.
- An **eigenvector** is a **nonzero** vector that changes at most by a scalar factor when that linear transformation is applied to it.



Eigenvalues and eigenvectors

Recall that a matrix represents a **transformation** between vector spaces. There are some transformations for which some vectors *never change direction*, but are only scaled.



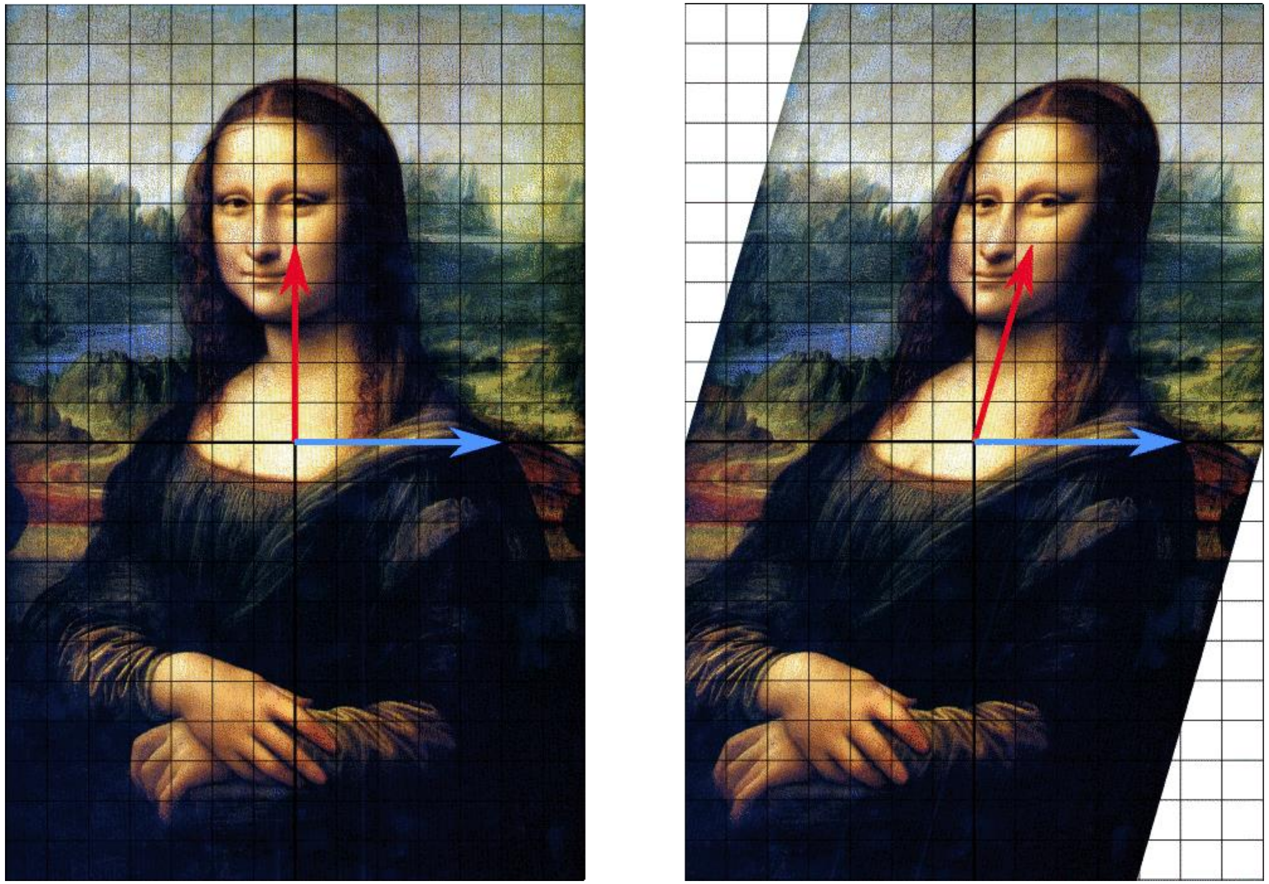
Eigenvalues and eigenvectors

These special vectors are called **eigenvectors**.
The scaling factor is called an **eigenvalue**.

$$\begin{bmatrix}3 & 1\\ 0 & 2\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} = \begin{bmatrix}-2\\ 2\end{bmatrix} = 2\begin{bmatrix}-1\\ 1\end{bmatrix}$$

$$\mathbf{M}\quad \mathbf{v} \quad = \quad \mathbf{2}\quad \mathbf{v}$$

Eigenvalues and Eigenvectors



- In this shear mapping the red arrow changes direction, but the blue arrow does not.
- The blue arrow is an eigenvector of this shear mapping because it does not change direction, and since its length is unchanged, its eigenvalue is 1.

Bag of Words (BOW)

Statement 1: One cat is sleeping, and the other one is running.

Statement 2: One dog is sleeping, and the other one is eating.

	One	Cat	Is	Sleeping	And	The	Other	Dog	Running	Eating
S1	2	1	2	1	1	1	1	0	1	0
S2	2	0	2	1	1	1	1	1	0	1

Term Frequency X Inverse Document Frequency

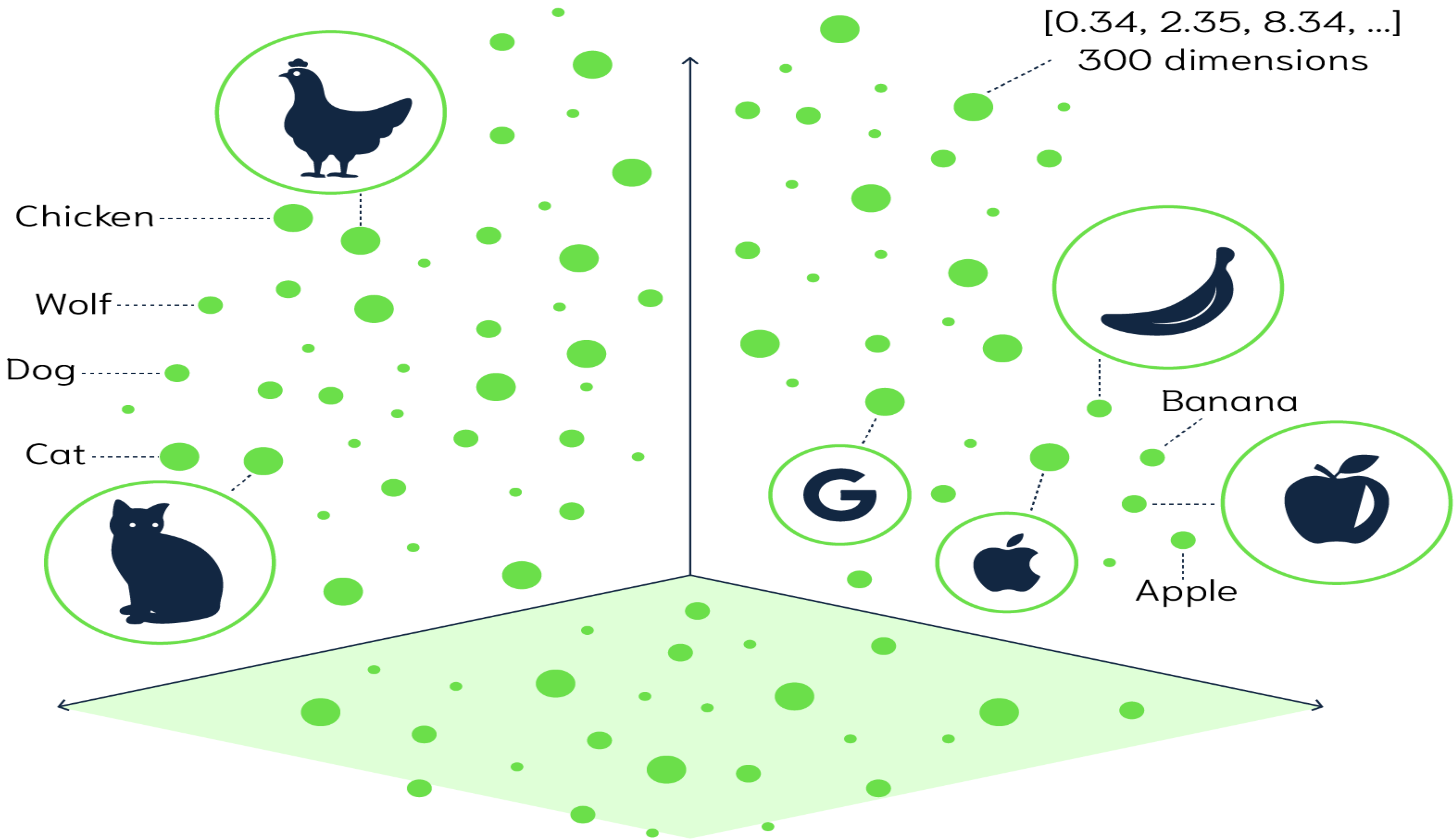
$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ = number of occurrences of i in j
df_i = number of documents containing i
N = total number of documents

$$tf(w,d) = \frac{occurence\ of\ w\ in\ document\ d}{total\ number\ of\ words\ in\ document\ d}$$

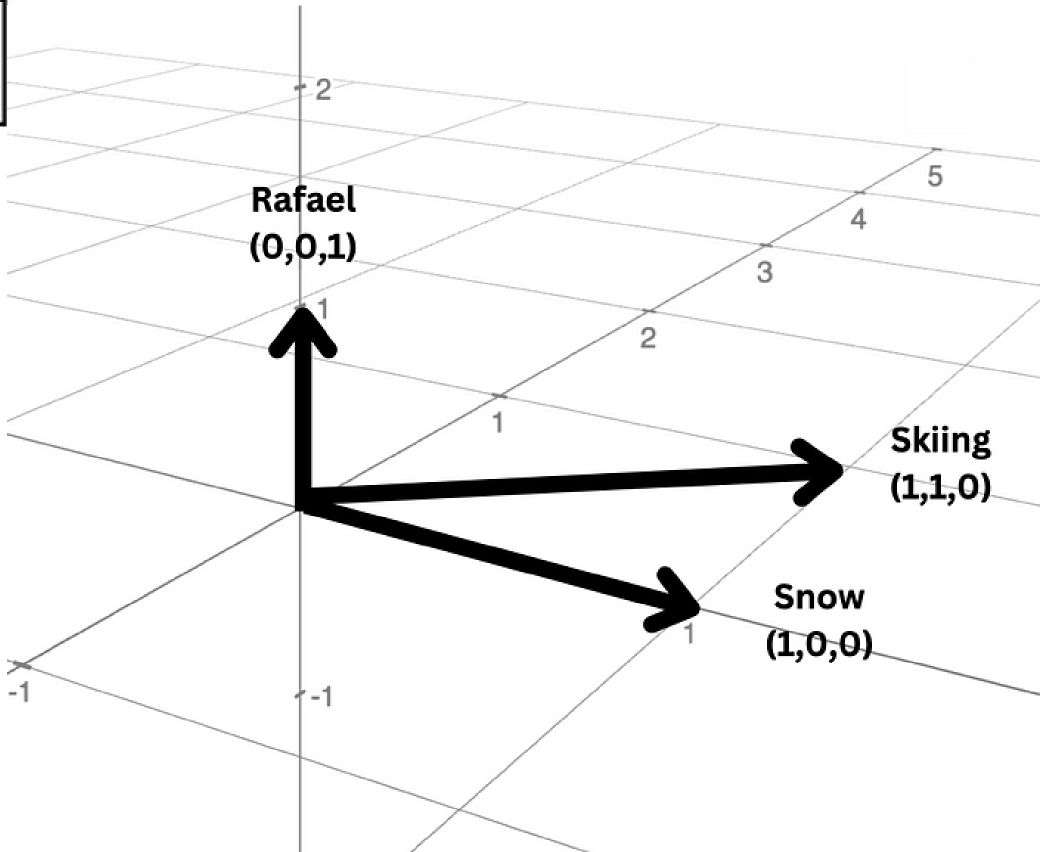
Text1: Basic Linux Commands for Data Science
Text2: Essential DVC Commands for Data Science

	basic	commands	data	dvc	essential	for	linux	science
Text 1	0.5	0.35	0.35	0.0	0.0	0.35	0.5	0.35
Text 2	0.0	0.35	0.35	0.5	0.5	0.35	0.0	0.35

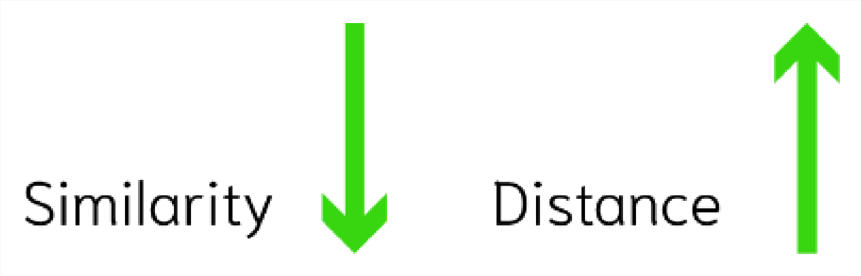

[0.34, 2.35, 8.34, ...]
300 dimensions
Chicken
Wolf
Dog
Cat
Banana
Apple

	animal_name	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed	backbone	breathes	venomous	fins	legs	tail	domestic	catsize	class_ty
0	aardvark	1	0	0	1	0	0	1	1	1	1	0	0	4	0	0	1	
1	antelope	1	0	0	1	0	0	0	1	1	1	0	0	4	1	0	1	
2	bass	0	0	1	0	0	1	1	1	1	0	0	1	0	1	0	0	
3	bear	1	0	0	1	0	0	1	1	1	1	0	0	4	0	0	1	
4	boar	1	0	0	1	0	0	1	1	1	1	0	0	4	1	0	1	
5	buffalo	1	0	0	1	0	0	0	1	1	1	0	0	4	1	0	1	
6	calf	1	0	0	1	0	0	0	1	1	1	0	0	4	1	1	1	
7	carp	0	0	1	0	0	1	0	1	1	0	0	1	0	1	1	0	
8	catfish	0	0	1	0	0	1	1	1	1	0	0	1	0	1	0	0	
9	cavy	1	0	0	1	0	0	0	1	1	1	0	0	4	0	1	0	
10	cheetah	1	0	0	1	0	0	1	1	1	1	0	0	4	1	0	1	
11	chicken	0	1	1	0	1	0	0	0	1	1	0	0	2	1	1	0	
12	chub	0	0	1	0	0	1	1	1	1	0	0	1	0	1	0	0	
13	clam	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	
14	crab	0	0	1	0	0	1	1	0	0	0	0	0	4	0	0	0	

Word	Is this token associated with cold climates?	Is this token a verb?	Is this token the name of a person?
 Snow	1	0	0
 Skiing	1	1	0
 Rafael	0	0	1



Cosine similarity between vectors



The cosine similarity is calculated as:

$$\mathrm{Cos}(A, B) = \frac{A \cdot B}{||A|| * ||B||}$$

A·B is the product (dot) of the vectors A and B

||A|| and ||B|| is the length of the two vectors

||A|| * ||B|| is the cross product of the two vectors

[Cosine Similarity, Clearly Explained!!! (youtube.com)](youtube.com)

Cosine similarity between vectors

Similarity ↓ Distance ↑

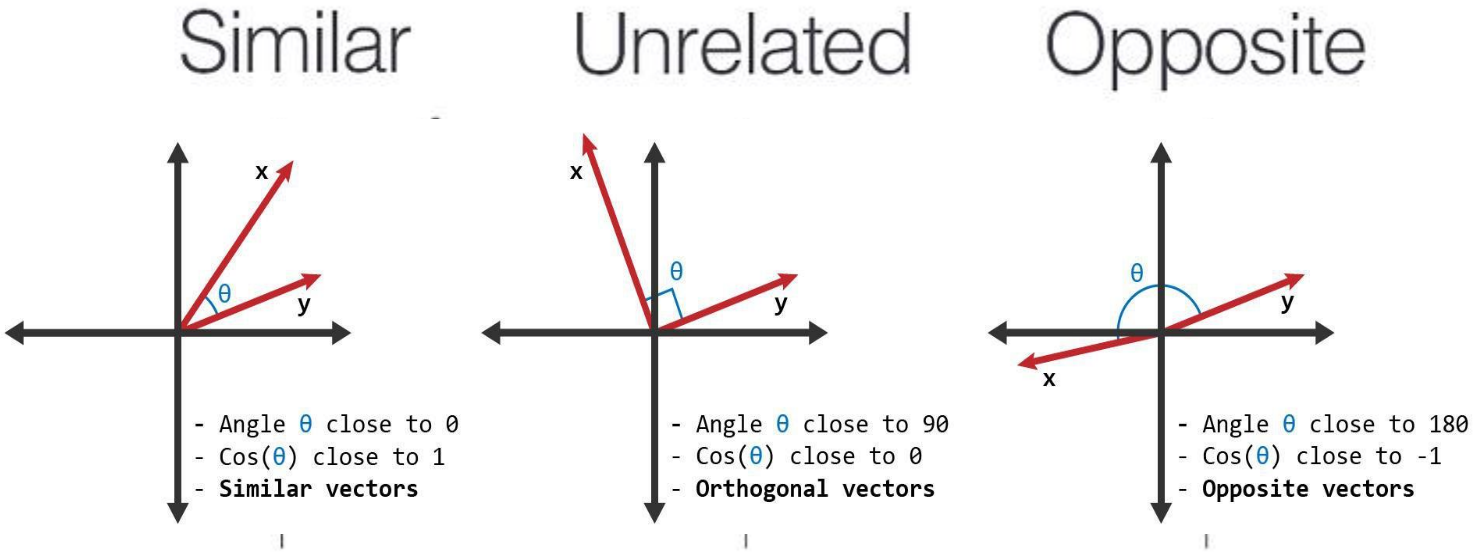
The cosine similarity is calculated as:

$$\mathrm{Cos}(A, B) = \frac{A \cdot B}{||A|| * ||B||}$$

A·B is the product (dot) of the vectors A and B

||A|| and ||B|| is the length of the two vectors

||A|| * ||B|| is the cross product of the two vectors



Cosine similarity

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}},$$

Let's use an example to calculate the **similarity** between two fruits

orange (vector A) and **apple** (vector B).

orange → [4, 0, 1]　　**apple → [3, 0, 1]**

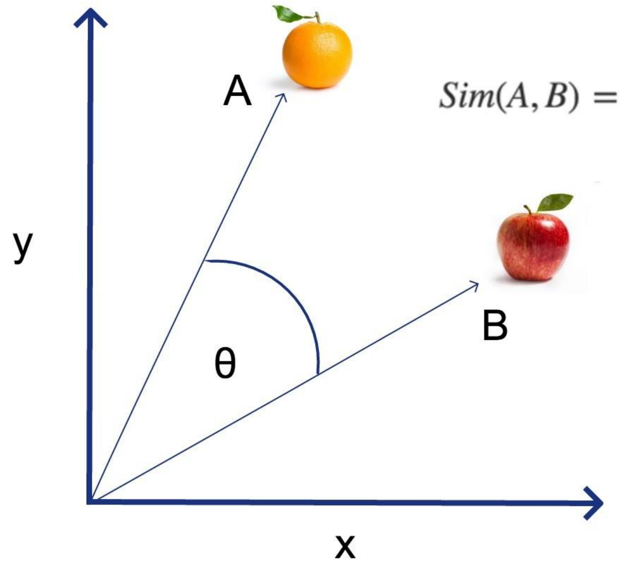
$A \cdot B = 4 * 3 + 0 * 0 + 1 * 1 = 13$

$||A|| = \sqrt{4^2 + 0^2 + 1^2} = 4.12$

$||B|| = \sqrt{3^2 + 0^2 + 1^2} = 3.16$

Cos(A,B) = 13/ (4.12 * 3.16) = 0.998

Cosine Distance = 1 – 0.998 = 0.002



$$Sim(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

Cosine similarity

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}},$$

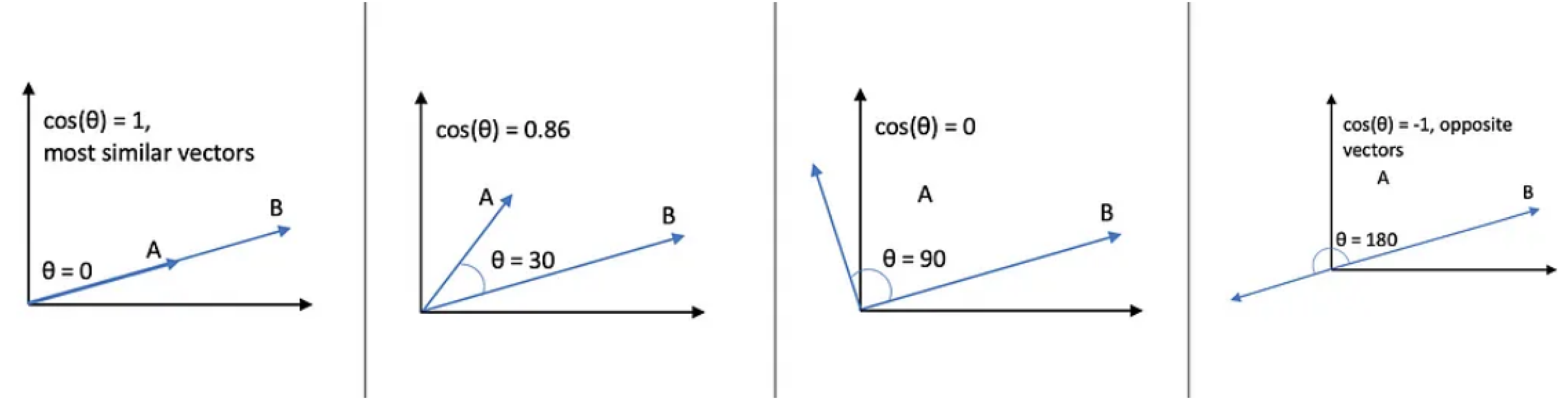
	Allen Solly	Arrow	Peter England	US Polo	Van Heusen	Zodiac
Customer 1	4	5	3	5	2	1
Customer 2	1	2	4	3	3	5

A.B = (4*1) + (5*2) + (3*4) + (5*3) + (2*3) + (1*5) = 52

|| A || = √(4² + 5² + 3² + 5² + 2² + 1² = 8.94

|| B || = √(1² + 2² + 4² + 3² + 3² + 5² = 8

$$\text{Cosine Similarity} = \frac{A.B}{\|A\| * \|B\|} = \frac{52}{8.94*8} = 0.72$$



values of cosine at different angles (Image by author)

Covariance

Covariance is a measure of the relationship between two random variables.

$$COV(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$\begin{bmatrix} Var(x_1) & \cdots\cdots & Cov(x_n,x_1) \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ Cov(x_n,x_1) & \cdots\cdots & Var(x_n) \end{bmatrix}$$

[Covariance, Clearly Explained!!! (youtube.com)]

Covariance Matrix

- Representing Covariance between dimensions as a matrix e.g. for 3 dimensions:

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix}$$

Variances

- Diagonal is the variances of x, y and z
- cov(x,y) = cov(y,x) hence matrix is symmetrical about the diagonal
- N-dimensional data will result in NxN covariance matrix

Covariance Matrix

	Math	Science	History
Math	64.9	33.2	-24.4
Science	33.2	56.4	-24.1
History	-24.4	-24.1	75.6

Example 1: Find the population covariance matrix for the following table.

Score	Age
68	29
60	26
58	30
40	35

Solution: The formula for population variance is $\frac{\sum_1^n (x_i - \mu)^2}{n}$.

$\mu_x = 56.5, n = 4$

var(x) = $[(68 - 56.5)^2 + (60 - 56.5)^2 + (58 - 56.5)^2 + (40 - 56.5)^2] / 4 = 104.75$

$\mu_y = 30, n = 4$

var(y) = $[(29 - 30)^2 + (26 - 30)^2 + (30 - 30)^2 + (35 - 30)^2] / 4 = 10.5$

$\text{cov}(x, y) = \frac{\sum_1^4 (x_i - \mu_x)(y_i - \mu_y)}{4}$

cov(x, y) = -27

The variance covariance matrix is given as follows:

$$\begin{bmatrix} 104.7 & -27 \\ -27 & 10.5 \end{bmatrix}.$$

Example 3: How will you interpret the covariance matrix given below?

$$\begin{array}{c|ccc} & X & Y & Z \\ \hline X & 500 & 320 & -40 \\ Y & 320 & 340 & 0 \\ Z & -40 & 0 & 800 \end{array}$$

1) The diagonal elements **500, 340** and **800** indicate the **variance** in data sets **X, Y** and **Z** respectively. **Y** shows the **lowest** variance whereas **Z** displays the **highest** variance.

2) The **covariance** for **X** and **Y** is **320**. As this is a **positive** number it means that when **X increases** (or decreases) **Y** also **increases** (or decreases)

3) The covariance for **X** and **Z** is **-40**. As it is a **negative** number it implies that when **X increases Z decreases** and vice - versa.

4) The **covariance** for **Y** and **Z** is **0**. This means that there is **no** predictable **relationship** between the two data sets.

Correlation

$$COV(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y}$$

Covarianced normalized by Standard Deviation

Correlation between X and Y

Standard deviation of X

Standard deviation of Y

[Pearson's Correlation, Clearly Explained!!! (youtube.com)]

Correleatiom Matrix