Introduction to Linear Algebra

Mathematical Foundations

Agenda

- Linear Algebra Overview
- Applications of Linear Algebra
- Vector properties and operations
- Linear Independence
- Change of basis (Vector)
- Matrix Introduction
- Types of Matrix
- Implementation in Python

What is Linear Algebra?

Linear Algebra is a branch of Mathematics that deals with symbols and rules that solve linear equations.

For example, Look at the following linear equations:

$$2a + 3b + 4c = 11$$

$$3a + 5b + 5c = 15$$

$$5a + 6b + 3c = 19$$

‘a’, ‘b’ and ‘c’ are the symbols and the equations are the rules.

Linear Algebra Representation

Linear Algebra = Vectors & Matrices and operations on them.

Vectors and Matrices are objects in space with various dimensions.

The equations in the previous slide can be represented using matrix and vectors as:

$$\begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 5 & 6 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 11 \\ 15 \\ 19 \end{pmatrix}$$

Application of Linear system

- Suppose an economy consists of the Coal, Electric (power), and Steel sectors, and the output of each sector is distributed among the various sectors as shown in Table 1, where the entries in a column represent the fractional parts of a sector's total output.
- The second column of Table 1, for instance, says that the total output of the Electric sector is divided as follows: 40% to Coal, 50% to Steel, and the remaining 10% to Electric. (Electric treats this 10% as an expense it incurs in order to operate its business.)
- Since all output must be taken into account, the decimal fractions in each column must sum to 1. Denote the prices (i.e., dollar values) of the total annual outputs of the Coal, Electric, and Steel sectors by $p_C, p_E, and\ p_S$, respectively. If possible, find equilibrium prices that make each sector's income match its expenditures.

Application of Linear system

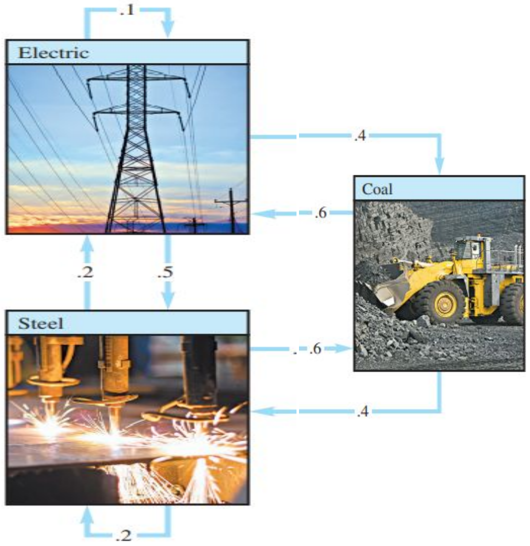


TABLE 1 A Simple Economy

Distribution of Output from:			
Coal	**Electric**	**Steel**	**Purchased by:**
.0	.4	.6	Coal
.6	.1	.2	Electric
.4	.5	.2	Steel

Application of Linear system

- A sector looks down a column to see where its output goes, and it looks across a row to see what it needs as inputs. For instance, the first row of Table 1 says that Coal receives (and pays for) 40% of the Electric output and 60% of the Steel output.
- Since the respective values of the total outputs are p_E and p_S,
- Coal must spend $.4p_E$ dollars for its share of Electric's output and $.6p_S$ for its share of Steel's output.
- Thus Coal's total expenses are $.4p_E + .6p_S$.
- To make Coal's income, p_C, equal to its expenses, we want $pC = .4pE + .6pS$

Application of Linear system

- $p_C = .4p_E + .6p_S$
- $p_E = .6p_C + .1p_E + .2p_S$
- $p_S = .4p_C + .5p_E + .2p_S$

$$\begin{aligned} p_C - .4p_E - .6p_S &= 0 \\ -.6p_C + .9p_E - .2p_S &= 0 \\ -.4p_C - .5p_E + .8p_S &= 0 \end{aligned}$$

Basic Terms

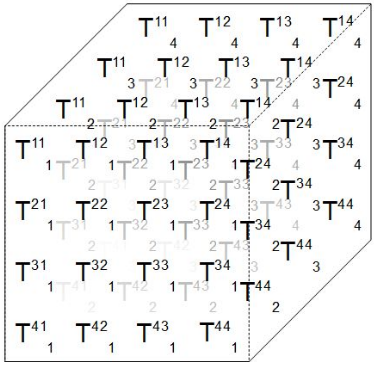
- Scalar - Only Magnitude

- Vector - Direction & Magnitude - one dimensional array of numbers

- Matrices - Two dimensional array of numbers- can represent transforms/ multiple vectors

- Tensors- N Dimensional array of numbers

$$(T^{ij}{}_{k})_{1 \le i, j, k \le 4} =$$



https://commons.wikimedia.org/wiki/File:Coordinates_of_a_tensor_of_type_(2,1)_in_dimension_4.png

Example Scalar and Vectors

- 1.A box being pushed at a force of 125 newtons ANSWER: **scalar**
- 2. wind blowing at 20 knots ANSWER: **scalar**
- 3. a deer running 15 meters per second due west ANSWER: **vector**
- 4. a baseball thrown with a speed of 85 miles per hour ANSWER: **scalar**

Example of Vectors

- **Color**: (R,G,B)
- **Quantities of n different commodities** (or resources), e.g., bill of materials
- **Portfolio**: entries give shares (or $ value or fraction) held in each of n assets, with negative meaning short positions
- **Cash flow**: xi is payment in period i to us
- **Audio**: xi is the acoustic pressure at sample time i (sample times are spaced 1/44100 seconds apart)
- **Features**: xi is the value of ith feature or attribute of an entity
- **Customer purchase**: xi is the total $ purchase of product i by a customer over some period
- **Word count**: xi is the number of times word i appears in a document

Word count vectors

- A short document: Word count vectors are used in computer based document analysis. Each entry of the word count vector is the number of times the associated dictionary word appears in the document.
- Consider The Following Examples
- Here We from set S of words excluding common words also called stop words like the etc etc S = {cat, dog, mat, sat, scratched} in dictionary order

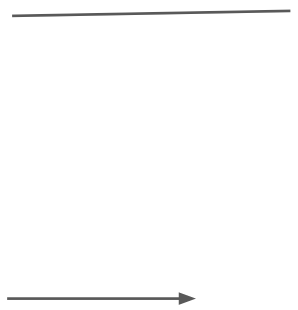
(a)“The dog sat on the mat” is summarised by the vector w = (0 , 1 , 1 , 1 , 0).
(b) “The cat scratched the dog” is summarised by the vector w = (1, 1, 0, 0, 1).
(c) “The cat and dog sat on the mat” is summarised by the vector w = (1, 1, 1, 1, 0).

Some Applications of Linear Algebra

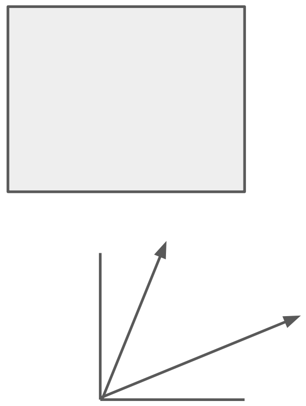
- Formulation of Machine Learning problem
- Image Processing
- Dimensionality Reduction
- Natural Language Processing
- Recommendation System
- Computer Vision

Vector Representations (Geometrically)

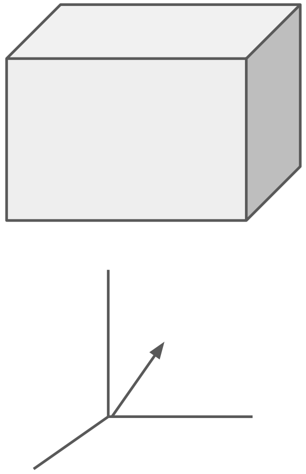
1 D



2 D



3 D



Vectors (Data science Example)

Can be of very higher dimension (To the extent of number of features)

Iris Dataset:

Each flower has 4 features, which can be considered as 4 Dimensions and every row is a Vector

$$\begin{pmatrix} \text{Sepal\_Length} \\ \text{Sepal\_Width} \\ \text{Petal\_Length} \\ \text{Petal Width} \end{pmatrix}$$

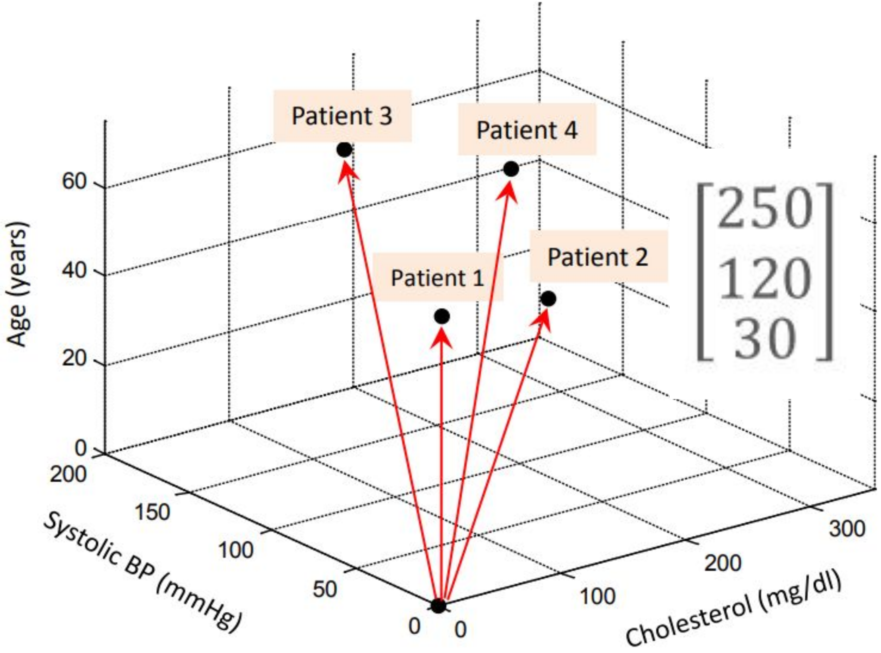
Flower	sepal_length	sepal_width	petal_length	petal_width
Flower1	5.1	3.5	1.4	0.2
Flower2	4.9	3.0	1.4	0.2
Flower3	4.7	3.2	1.3	0.2
Flower 1000	4.9	3.8	1.7	0.7

Vector Representation

- Consider the following Data of Patient , Then measurement of Cholesterol , Systolic BP and Age in each row can be considered as a vector

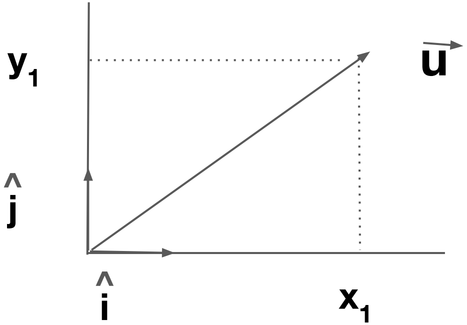
Patient id	Cholesterol (mg/dl)	Systolic BP (mmHg)	Age (years)	Tail of the vector	Arrow-head of the vector
1	150	110	35	(0,0,0)	(150, 110, 35)
2	250	120	30	(0,0,0)	(250, 120, 30)
3	140	160	65	(0,0,0)	(140, 160, 65)
4	300	180	45	(0,0,0)	(300, 180, 45)

Vector Representation



Coordinate System- Vectors

Symbol for vector : $\longrightarrow$ **Symbol for unit vector :** ^

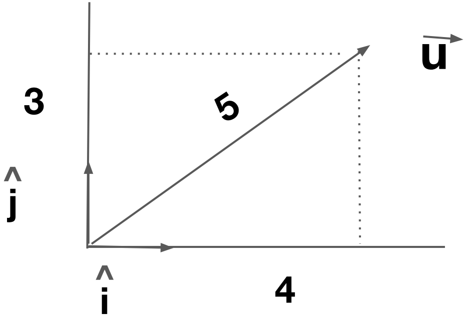


(x_1 , y_1)

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

$$\|u\| = \sqrt{(x_1{}^2 + y_1{}^2}$$

$$\hat{u} = \frac{\vec{u}}{\|u\|} = \left(\frac{x_1}{\|u\|} , \frac{y_1}{\|u\|}\right)$$



$(4 , 3)$

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

$$\|u\| = \sqrt{(4^2 + 3^2}$$

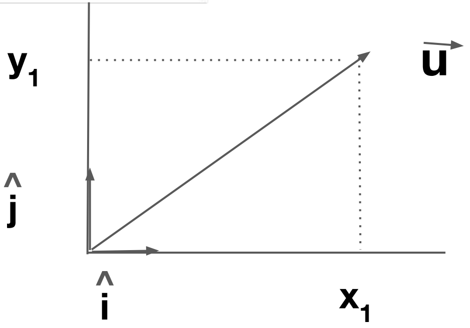
$$\hat{u} = \frac{\vec{u}}{\|u\|} = \left(\frac{4}{5} , \frac{3}{5}\right)$$

Magnitude of a vector - Vector norm

PES UNIVERSITY

The length of vector is called the magnitude / norm of the vector.

There are Three norms - L1 and L2 and L_∞ Norms



$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

L1 Norm : $\|\mathbf{u}\| = |\mathbf{x}_1| + |\mathbf{y}_1|$

L2 Norm : $\|u\| = \sqrt{(x_1^2 + y_1^2}$

L_∞ Norm Chebyshev distance or Maximum Metric

- Given Vector $\mathrm{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, then its L_∞ Norm is defined as follows
- $||u||_\infty = Max\{x_1, y_1\}$
- It can be also thought of as minimum number moves on chess boards for a king to move from one place to other

$\mathrm{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ then $||u||_\infty = Max\{2{,}3\} =$ 3

Minkowski distance

The Minkowski distance of order p (where p is an integer) between two points

$$X = (x_1, x_2, \ldots, x_n) \text{ and } Y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$$

is defined as:

$$D(X,Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}} = ||\boldsymbol{x} - \boldsymbol{y}||_p$$

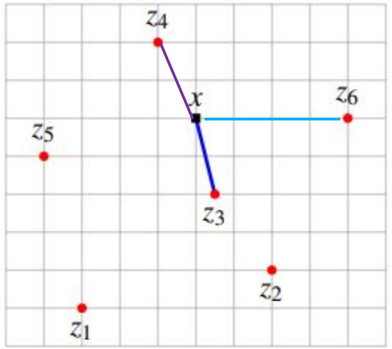
The General Norm is called **Minkowski distance** it is defined as for p =1 it becomes Manhattan for p= 2 it becomes Euclidean and for p = infinity it becomes Chebyshev distance

Application of distance: Nearest Neighbors

- Feature distance and nearest neighbors in the below diagram 3 nearest neighbor of x are z3 , z6, z4
  Based on Euclidean distances.

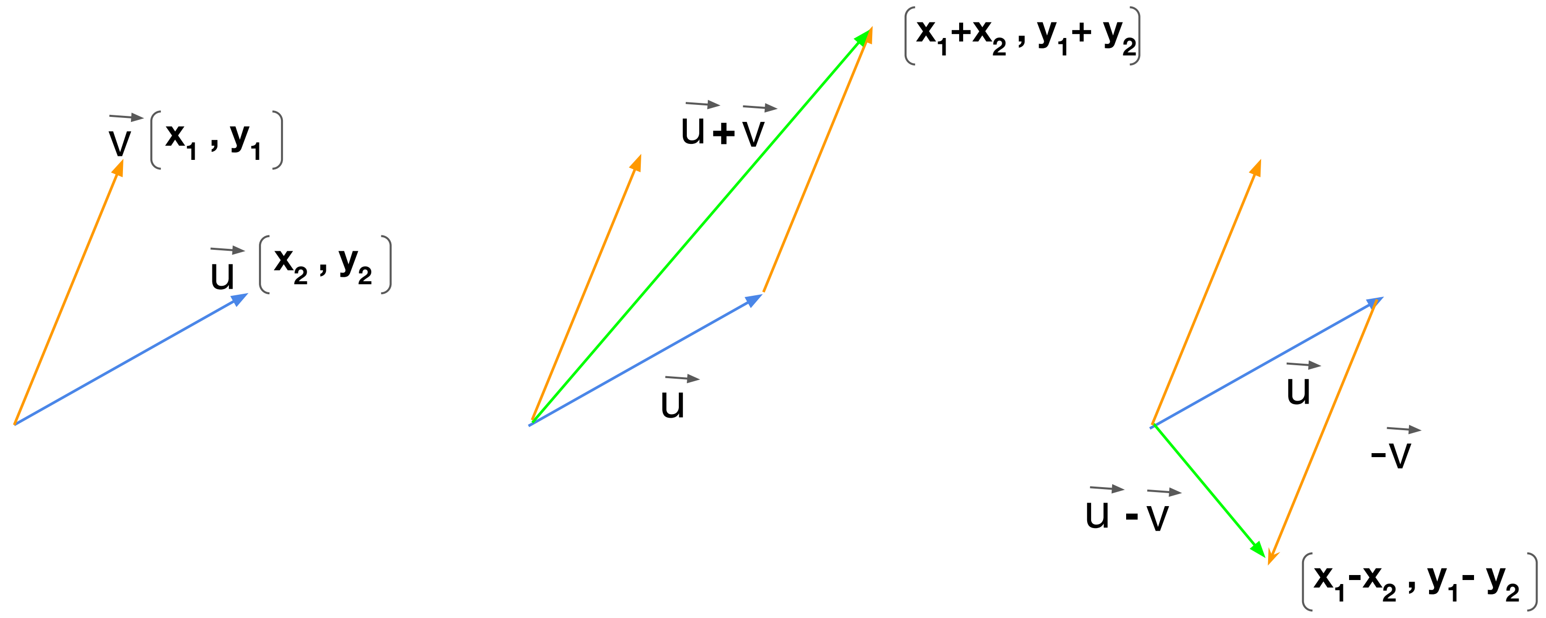
- if x and y are feature vectors for two entities, $\|x - y\|$ is the *feature distance*
- if $z_1, \ldots, z_m$ is a list of vectors, z_j is the *nearest neighbor* of x if

$$\|x - z_j\| \leq \|x - z_i\|, \quad i = 1, \ldots, m$$

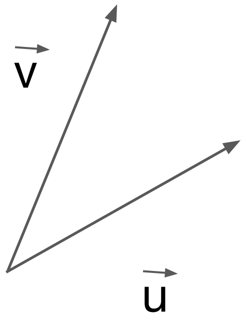


- these simple ideas are very widely used

Vector Addition and subtraction



Vector Properties

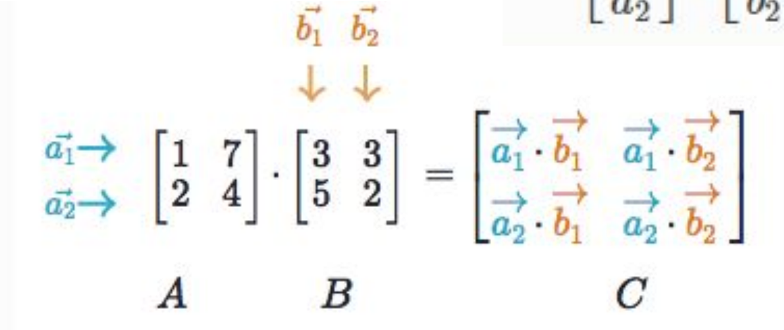


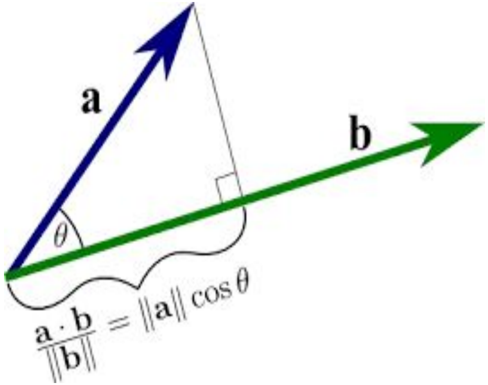
- Associativity of addition $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
- Commutativity of addition $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- Distributivity of scalar multiplication with respect to vector addition $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$
- Multiplicative Identity is 1

Vector Operations- Continued

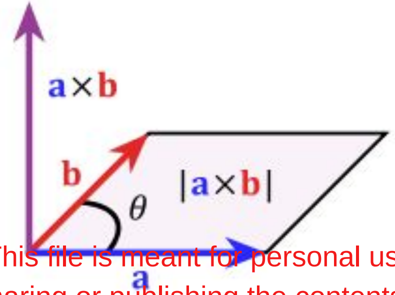
Dot Product -

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1 b_1 + a_2 b_2$$





Cross Product



The magnitude (length) of the cross product equals the area of a parallelogram with vectors **a** and **b** for sides.

Scalar transformations

Dot Product Application

- Consider a firm operating two plants in two different locations. They both produce the same output (say, 10 units) using the same type of inputs. Although the amounts of inputs vary between the plants the output level is the same.
- The firm management suspects that the production cost in Plant 2 is higher than in Plant 1.
- The following information was collected from the managers of these plants.

PLANT 1

Input	Price	Amount used
Input 1	3	9
Input 2	5	10
Input 3	7	8

PLANT 2

Input	Price	Amount used
Input 1	4	8
Input 2	7	12
Input 3	3	9

Dot Product Application

- Question 1. Does this information confirm the suspicion of the firm management?
- Plant 1: Price Vector $p_1 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$ and Quantity Vector $q_1 = \begin{bmatrix} 9 \\ 10 \\ 8 \end{bmatrix}$
- Plant 2 : Price Vector $p_2 = \begin{bmatrix} 4 \\ 7 \\ 3 \end{bmatrix}$ and Quantity Vector $q_2 = \begin{bmatrix} 8 \\ 12 \\ 9 \end{bmatrix}$
- In Plant 1 the total cost is $\mathbf{c_1} = \mathbf{p_1}.\mathbf{q_1} = 133$, which implies that $\boldsymbol{unit\ cost\ is\ 13.3}$.
- In Plant 2, cost of production is $\boldsymbol{c_2} = \boldsymbol{p_2}.\boldsymbol{q_2} = 143$, which gives $\boldsymbol{unit\ cost\ as\ 14.3}$ which is higher than the first plant. That is, the suspicion is reasonable.

Dot Product Application

- Question 2. The manager of the Plant 2 claims that the reason of the cost differences is the higher input prices in her region than in the other.
- Is the available information supports her claim?
- $p_2 = \lambda p_1$
- Plant I enjoys lower prices for inputs 2 and 3,
- whereas Plant 2 enjoys lower price for input 3.
- For a rough guess, one can still compare the lengths (norm)which are
- $||p_1|| = \sqrt{p_1 . p_1} = \sqrt{3^2 + 5^2 + 7^2} = \sqrt{83} = 9.11$ Similarly $||p_2|| = 8.60$

PLANT 1

Input	Price	Amount used
Input 1	3	9
Input 2	5	10
Input 3	7	8

PLANT 2

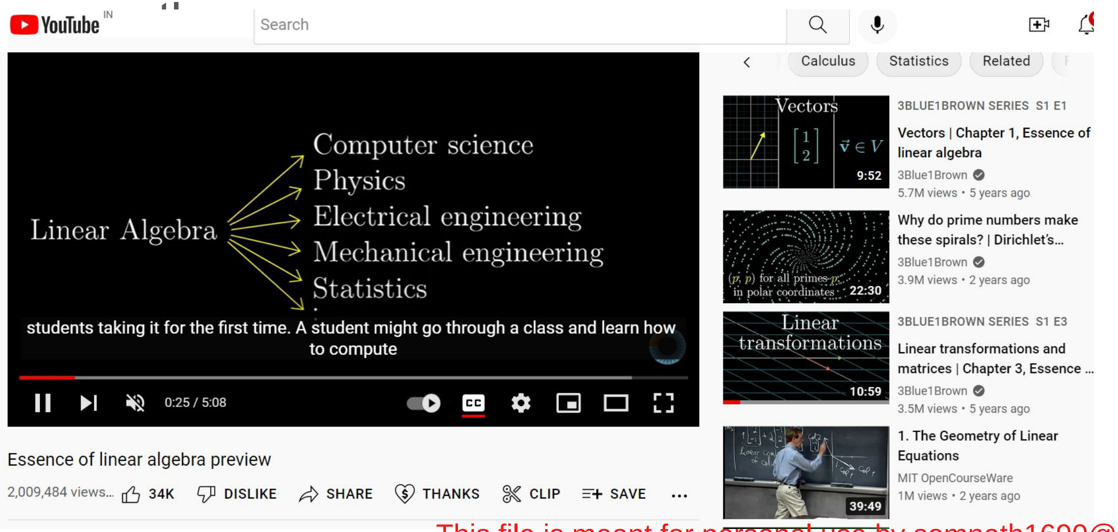
Input	Price	Amount used
Input 1	4	8
Input 2	7	12
Input 3	3	9

Dot Product Application

- which indicates that price data does not support the claim of the manager of the Plant 2.
- When examined more closely, one can see that the Plant 2 uses the most expensive input (input 2) intensely.
- In contrast, Plant 2 managed to save from using the most expensive input (in this case input 3).
- Therefore, the manager needs to explain the reasons behind the choice mixture of inputs in her plant.

Dot Product Application : You tube video recommendation

Consider a user u and watching a video v on you tube , then user u can also see list of recommended videos (or very similar videos) in a window alongside. How does you tube recommend similar videos to the video user is watching



Dot Product Application:

- Lets assume that we n users and m videos , in reality n, m can be in millions or billons. Consider the following arrangement of binary numbers o and 1 in which every row is a user and every column is a video , we define each entry as follows If user $u_i$watched video v_j then we fill that entry with 1 otherwise we fill it as 0 Let's assume we have 5 users and 9 videos

Ssimilarity between video vi and vj is calculated as vi.vj

Dot Product Application

	v1	v2	v3	v4	v5	v6	v7	v8	v9
u1	1	1	0	1	0	1	0	0	0
u2	1	1	1	0	1	0	1	0	0
u3	1	0	0	0	1	0	0	0	1
u4	1	1	1	1	0	1	0	1	0
u5	1	0	0	1	0	1	0	1	0

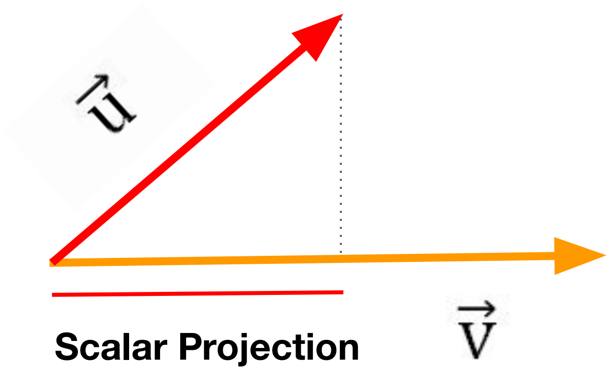
Dot Product Application

- Similarity between video v1 and v3 is , look at v1 column entries and v3 column entries and take a dot product

- $v1.v3 = [\,1\ \ 1\ \ 1\ \ 1\ \ 1\,].\begin{bmatrix}0\\1\\0\\1\\0\end{bmatrix} = 2$

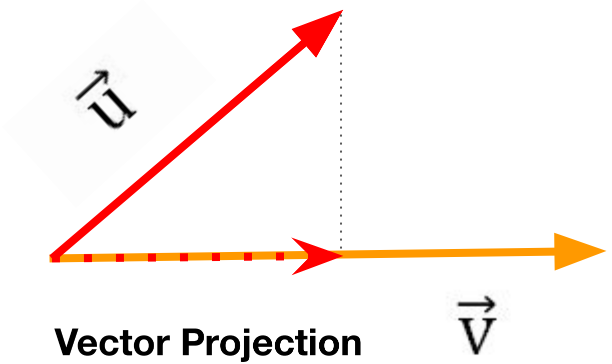
Scalar Projections

- A scalar projection of $\vec{u}$ onto $\vec{v}$ is the length of the shadow that $\vec{u}$ casts on $\vec{v}$:



Vector Projections

- A vector projection of $\vec{\mathbf{u}}$ onto $\vec{\mathbf{v}}$ is the vector in the same direction as $\vec{\mathbf{v}}$, whose length is the scalar projection of $\vec{\mathbf{u}}$ on $\vec{\mathbf{v}}$.



- The scalar projection is the magnitude of the vector projection.

Scalar and Vector Projections- Example

Find the scalar projection, the vector projection, and the orthogonal component of:

(6*i* + 7*j*) onto **(5*i* - 12*j*)**

The scalar projection is: $= 6 \cdot 5 + 7 \cdot (-12)$

$= -54/13$

The vector projection is: $= -54(5i - 12j)/13^2$

$= -270i/169 + 648j/169$

The orthogonal component is: $= (6i + 7j) - (-270i/169 + 648j/169)$

$= 1284i/169 - 535j/169$

Linear Combination of vectors

- Consider the set of n vectors $S = \{v_1, v_2, \dots, v_n\}$ and scalars $c_1, c_2, \dots, c_n$ then

$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$ is called linear combination of $v_1, v_2, \dots, v_n$

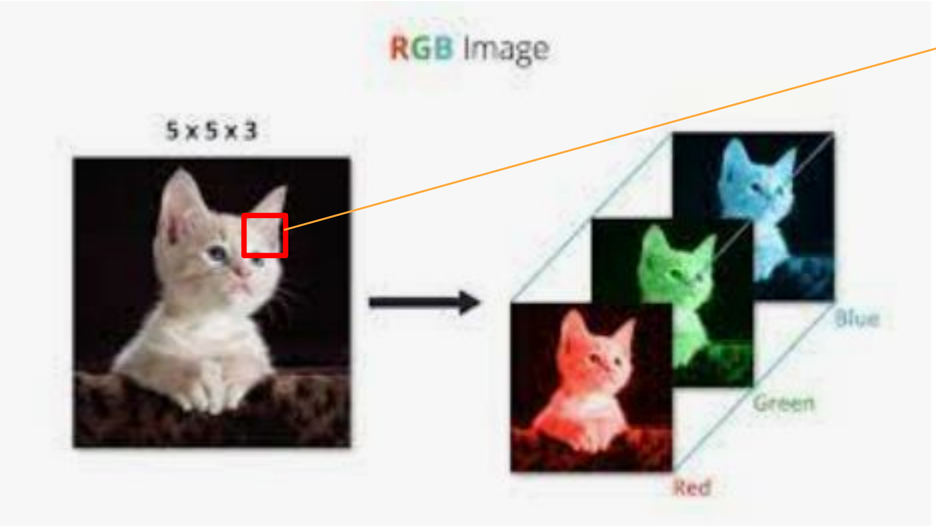
Eexamples :- suppose v1=$\begin{bmatrix}1\\2\end{bmatrix}$ and v2=$\begin{bmatrix}3\\4\end{bmatrix}$ c1 = 2, c2 = - 1.5

Then 2*v1 + (-1.5)*v2 = 2* $\begin{bmatrix}1\\2\end{bmatrix}$+(-1.5)* $\begin{bmatrix}3\\4\end{bmatrix}$ is linear combination of v1,v2

Example of linear combination

Consider color image also called RGB image , it has three component RED Green Blue As we can see red rectangle part of image is linear combination of RED GREEN and BLUE

C1*Red+c2*Green+c3*Blue



Linear Independence

- A Set of vectors is Linearly dependent, if one of the vectors can be defined as a combination of other vectors.
- If no vector of the set can be represented as combination of other vectors in the set, the set of vectors is linearly independent
- Hence, for vectors to be linearly independent, the linear combination of the vectors results in zero vector , if and only if all the constants are zero.

$$c_1\overrightarrow{V_1} + c_2\overrightarrow{V_2} + \cdots + c_n\overrightarrow{V_n} = 0 \qquad c_1 = c_2 = \cdots = c_n = 0$$

Span of a Set

- Consider the set of n vectors $S = \{v_1, v_2, \ldots, v_n\}$ and scalars $c_1, c_2, \ldots, c_n$
- Then $L(S) = \{c_1 v_1 + c_2 v_2 + \cdots + c_n v_n : c_1, c_2, \ldots, c_n$ scalar$\}$is called a span of S

Let S = $\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$ Then $L(S) = \{x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix} : x, y\ scalars\}$

$L(S) = \begin{bmatrix}x\\y\end{bmatrix}$ for all real numbers x and y

Basis Vectors

- If you can write every vector in a given space as a linear combination of some vectors and these vectors are independent of each other then we call them as basis vectors for that given space.
- **Properties of basis vector:**
- **Basis vectors must be linearly independent of each other.**
- **Basis vectors must span the whole space**
- **Basis vectors are not unique**
- **Basis vectors need not be unit vectors and orthogonal**

Linear Independence (Example)

$$\begin{array}{c} R1 \\ R2 \\ R3 \end{array} \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 5 & 6 & 3 \end{pmatrix}$$

$$\begin{array}{c} R1 \\ R2 \\ R3 \end{array} \begin{pmatrix} 2 & 3 & 4 \\ 4 & 6 & 8 \\ 5 & 6 & 3 \end{pmatrix}$$

R2= 2 R1
2R1-R2+0R3=0

$$\begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 5 & 7 & 9 \end{pmatrix} \begin{array}{c} R1 \\ R2 \\ R3 \end{array}$$

R1 + R2 = R3
1R1+1R2+(-1)R3=0

Only 2 vectors are sufficient, 3rd vector does not add any additional information

Importance of Linear Independence

- Direction of two linearly dependent vectors is the same. This shows that there can be only as many solutions as the no. of independent vectors.
- Determines if the vectors form a basis
- Determines the no. of independent equations
- Determine the no. of independent vectors

Basis

- Let S = $\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$ Then $L(S) = \{x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix} : x, y\ scalars\}$

$L(S)$ = $\begin{bmatrix}x\\y\end{bmatrix}$ for all real numbers x and y

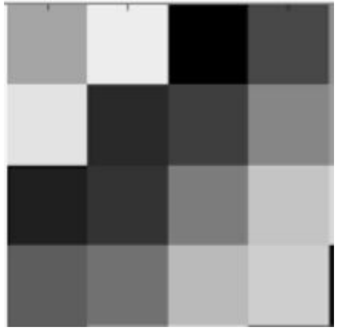
Here aby vector in 2d plane can be written as linear combination of S.

S is Linearly independent as $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$ are not multiples of each other

Therefore S is a basis for 2d plane

Basis Vector Application

- Consider The following Vectors v1 =  $= \begin{bmatrix} 255 \\ 0 \\ 255 \\ 0 \end{bmatrix}$
- Vector v2 =  $= \begin{bmatrix} 0 \\ 255 \\ 0 \\ 255 \end{bmatrix}$
- Then the image  is linear combination of Vectors v1 and v2
-  $= \frac{200}{255}$  $+ \frac{75}{255}$  $= \frac{200}{255} v_1 + \frac{75}{255} v_2$
- So v1 and v2 are basis vectors for any grey scale 4 by 4 image



Change of basis

Change of basis is a technique applied to finite-dimensional **vector spaces** in order to rewrite **vectors** in terms of a different set of **basis** elements.

Given two **bases** of a **vector space**, there is a way to translate **vectors** in terms of one **basis** into terms of the other; this is known as **change of basis**.

Basis is a set of N vectors

i) that are linearly independent

ii) spans the space with the N dimensions

Change of Basis

- Lets consider $\mathbb{R}^2$ we know the standard basis for $\mathbb{R}^2$ is $\{e_1 = \begin{bmatrix}1\\0\end{bmatrix}, e_2 = \begin{bmatrix}0\\1\end{bmatrix}\}$
- So $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ is a standard basis matrix
- We can write any $\begin{bmatrix}x\\y\end{bmatrix} \in \mathbb{R}^2$ as $\begin{bmatrix}x\\y\end{bmatrix} = x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix}$
- let's take another basis vector $\{u_1 = \begin{bmatrix}2\\1\end{bmatrix}, u_2 = \begin{bmatrix}-1\\1\end{bmatrix}\}$ which has matrix form
- $\begin{bmatrix}2 & -1\\1 & 1\end{bmatrix}$
- consider the vector $v = \begin{bmatrix}3\\2\end{bmatrix}$ then we can write v in two ways as follows
- $\begin{bmatrix}3\\2\end{bmatrix} = 3\begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\end{bmatrix}$ Or $\begin{bmatrix}3\\2\end{bmatrix} = \alpha\begin{bmatrix}2\\1\end{bmatrix} + \beta\begin{bmatrix}-1\\1\end{bmatrix}$ ....(1)

Change of Basis

- $\therefore 3\begin{bmatrix}1\\0\end{bmatrix}+2\begin{bmatrix}0\\1\end{bmatrix}=\alpha\begin{bmatrix}2\\1\end{bmatrix}+\beta\begin{bmatrix}-1\\1\end{bmatrix}$ using (1)

$\Rightarrow 3e_1+2e_2=\alpha u_1+\beta u_2$

$\Rightarrow 3e_1+2e_2=\alpha(2e_1+1.e_2)_1+\beta(-1e_1+1.e_2)$

$\Rightarrow 3e_1+2e_2=(2\alpha e_1-\beta e_1)+(\alpha e_2+\beta e_2)$

$\Rightarrow 3e_1+2e_2=(2\alpha-\beta)e_1+(\alpha+\beta)\ e_2$

Equating

$\Rightarrow (2\alpha-\beta)$=3 and $(\alpha+\beta)$=2

Which can be rewritten as

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

Change of Basis

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

$\Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ as $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

$\Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \end{bmatrix}$ This means that the vector $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ has coordinates 3, 2 in standard basis

But the same vector $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ has coordinates $\frac{5}{3}, \frac{1}{3}$ with u_1, u_2 basis i.e

Change of Basis

- $\begin{bmatrix}3\\2\end{bmatrix} = \mathbf{3}\begin{bmatrix}1\\0\end{bmatrix} + \mathbf{2}\begin{bmatrix}0\\1\end{bmatrix}$
- $\begin{bmatrix}3\\2\end{bmatrix} = \frac{5}{3}\begin{bmatrix}2\\1\end{bmatrix} + \frac{1}{3}\begin{bmatrix}-1\\1\end{bmatrix}$
- so if $\begin{bmatrix}x\\y\end{bmatrix}$ is any vector with respect to standard basis then its coordinates with respect to any given basis $u_1 = \begin{bmatrix}a\\b\end{bmatrix}$ and $u_2 = \begin{bmatrix}c\\d\end{bmatrix}$ is given by
- $\begin{bmatrix}a & c\\b & d\end{bmatrix}^{-1}\begin{bmatrix}x\\y\end{bmatrix}$ and conversely given any vector $\begin{bmatrix}p\\q\end{bmatrix}$ with repect to $\{u_1, u_2\}$ then we can represent it in standard basis as $\begin{bmatrix}a & c\\b & d\end{bmatrix}\begin{bmatrix}p\\q\end{bmatrix}$

Orthonormal Basis

- Two vectors are orthogonal , when their dot product is zero.
- When the basis is orthogonal and normalized, The basis is Orthonormal

$\vec{u}$ $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\qquad \|\vec{u}\| = \sqrt{1^2 + 1^2}$ $\qquad \hat{u} = \dfrac{\vec{u}}{\|\vec{u}\|}$ $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

$\vec{v}$ $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ $\qquad \|\vec{v}\| = \sqrt{(-1)^2 + 1^2}$ $\qquad \hat{v} = \dfrac{\vec{v}}{\|\vec{v}\|}$ $\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

Introduction of Matrix (as Linear transformation)

- A matrix is a 2 dimensional array of numbers arranged in rows and columns. Matrices provide a method of organizing, storing, and working with mathematical information.
- Matrices provide a useful tool for working with models based on systems of linear equations.
- If two matrices have the same size, they can be added or subtracted.
- If a matrix is multiplied by a scalar, each entry is multiplied by that scalar. We can consider scalar multiplication as multiplying a number and a matrix to obtain a new matrix as the product.
- To perform **multiplication of two matrices**, we should make sure that the number of columns in the 1st matrix is equal to the rows in the 2nd matrix.

Introduction of Matrix (as Linear transformation)

- Matrix transform Space/ Vector
- Change of basis can be effected through a matrix transform

$$\begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 5 & 6 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 11 \\ 15 \\ 19 \end{pmatrix}$$

Matrix Properties: Associative

A + B = B + A

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \\ a_{31}+b_{31} & a_{32}+b_{32} & a_{33}+b_{33} \end{pmatrix}$$

Transpose

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

If $A=A^T$, Then A is Symmetric Matrix

$$\begin{pmatrix} 1 & 3 & 4 \\ 3 & 2 & 5 \\ 4 & 5 & 8 \end{pmatrix}$$

Identity Matrix

Identity matrix have no of rows same as no of columns

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

Other Types of Matrices

Lower Triangular Matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 5 & 0 \\ 7 & 8 & 9 \end{bmatrix}$$

Upper Triangular Matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{bmatrix}$$

Diagonal Matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

Orthogonal Matrix

When the product of a square matrix and its transpose gives an identity matrix, then the square matrix is known as orthogonal matrix.

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

If $AA^T = A^TA = I$, then matrix A is orthogonal.

Sparse Matrix

- Sparse matrix has its maximum entries zero

Dense Matrix

1	2	31	2	9	7	34	22	11	5
11	92	4	3	2	2	3	3	2	1
3	9	13	8	21	17	4	2	1	4
8	32	1	2	34	18	7	78	10	7
9	22	3	9	8	71	12	22	17	3
13	21	21	9	2	47	1	81	21	9
21	12	53	12	91	24	81	8	91	2
61	8	33	82	19	87	16	3	1	55
54	4	78	24	18	11	4	2	99	5
13	22	32	42	9	15	9	22	1	21

Sparse Matrix

1	.	3	.	9	.	3	.	.	.
11	.	4	.	.	.	.	.	2	1
.	.	1	.	.	.	4	.	1	.
8	.	.	.	3	1	.	.	.	.
.	.	.	9	.	.	1	.	17	.
13	21	.	9	2	47	1	81	21	9
.	.	.	.	.	.	.	.	.	.
.	.	.	.	19	8	16	.	.	55
54	4	.	.	.	11	.	.	.	.
.	.	2	.	.	.	.	22	.	21

Applications of Matrices

User-rating matrix , it is sparse matrix why ?

John	5	1	3	5
Tom	?	?	?	2
Alice	4	?	3	?

Appendix

- **numpy.linalg.norm(u)** # *Euclidean Distance; u is a array*
- **numpy.linalg.norm(u, ord=1)** # *Manhattan Distance; u is a array*
- ***Note:*** *linalg is a module in numpy to specifically work with linear algebra*

- Dot product: **numpy.dot(a, b)** *#Where a, b are two array/Matrix of same size*

- Scalar Projection: **sc_proj = np.dot(u, v) / np.linalg.norm(v)** *#Refer to the formula mentioned in the deck above*
- Vector Projection: **vec_proj = (v * sc_proj)/ np.linalg.norm(v)**

- **To inverse a matrix A:** V = inv(A)

Thank You!