Linear Algebra

11 April 2023 07:29

Linear algebra is a branch of mathematics that deals with the study of linear systems, which are sets of equations involving linear functions of variables. It is a foundational subject in mathematics and has applications in many areas, including computer science, engineering, physics, economics, and more.

advance ← vectors →

Scalars: 2, 5, -6

Vectors: [22 36 45]

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

Tensors

3D → 1D, 2D → n-D

ML? — ② reasons

1) generalizing concepts in higher dim

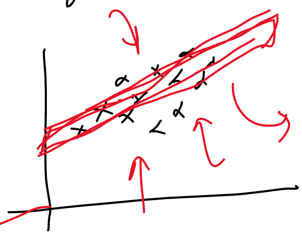
2D, 3D

data → 5000 dims

| educ | exp | salary |

given

solve higher



hyperplane

2) Data representation

→ tabular
→ text
→ image
→ video

→ represent linear algebra → ml algorithm

CPU / GPU

prepare GPU

What are Vectors

11 April 2023 07:30

- maths
- physics
- computer → ML

dimension (2) (3) (n)

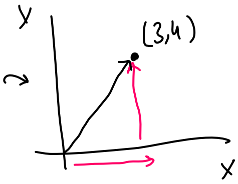
$a = [x_1, x_2, x_3 \ldots x_n]$

n dim vector → $\boxed{\text{n\_dim}}$

$a = [x_1\ x_2\ x_3 \ldots x_n]$

2D coordinate

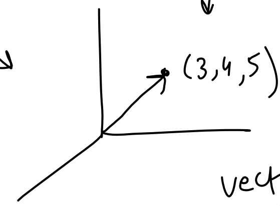
point



components

$a = [3,4]$

x comp, y comp



2D, 3D nD

feature vectors

$a = [3,4,5]$

in 3D space

vectors → iris

input feature

target

feature vector

$[5.2, 2.1, 7.2, 9.1]$ → model → setosa

4 dim

$[3.1, 4.2, 7, 6.5]$

numbers

(150)

150 flower species

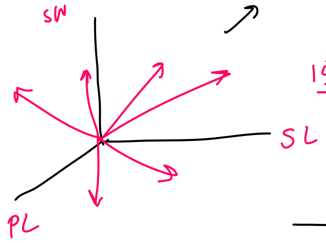
SL	SW	PL	PW
5.2	2.1	7.2	9.1
3.1	4.2	7	6.5

datapoint

species: setosa, versicolor

$\boxed{4d}$

150 vectors (feature)



titanic

Q C S
0 1 2

age	fare	sex	Pclass	embarked
17	52	$\boxed{M}$	1	S

encode
M → 0
F → 1

$[17, 52, 0, 1, 2]$

a feature vector → 5 dim

Vector example in Machine Learning

11 April 2023 07:30

movies → Matrix

name	summary
m1	s1
m2	s2
m3	s3

→ rows → vectors

m1	Hi how are you
m2	my name is campusx
m3	this is 2023

→ feature vector → BOW

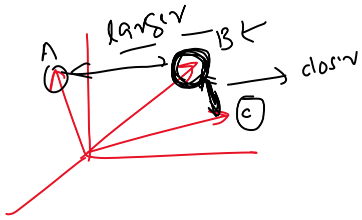
5000 movies

50000 words

feature vectors

Hi	how	are	you	my	name	is	campusx	this	2023
1	1	1	1	0	0	0	0	0	0
0	0	0	0	1	1	1	1	0	0
0	0	0	0	0	0	1	0	1	0

3 feature vectors → 10 dim



Row and Column Vector

11 April 2023 07:30

$$a = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}_{1 \times n}$$

row vector

rows cols

a is a row vector

$$b = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$$

rows cols

SL	SW	PL	PW	Species
–	–	–	–	–
–	–	–	–	–
–	–	–	–	–
⋮				
–	–	–	–	–

row vector

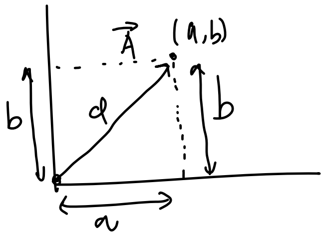
$[1 \times n]$

col vector

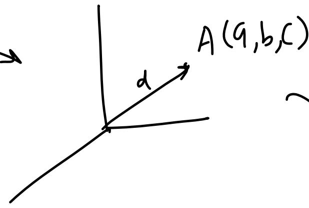
150×1

Distance From Origin ||A||

11 April 2023 07:31



$$\boxed{d = \sqrt{a^2+b^2}}$$



$$d = \sqrt{a^2+b^2+c^2}$$

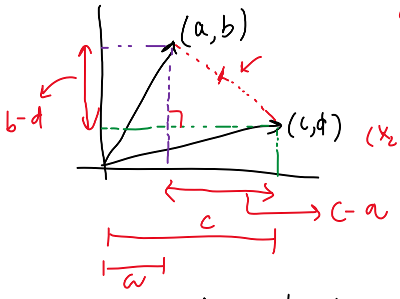
$\boxed{n\text{-}D}$ $A = [x_1\ x_2 \dots x_n]$

$$d = \sqrt{x_1^2 + x_2^2 + x_3^2 + \dots x_n^2}$$

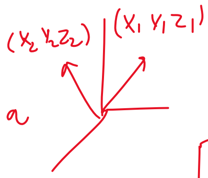
$\boxed{\text{magnitude}}$ $\boxed{||A||}$

Euclidean Distance

11 April 2023 07:31



$$dist = \sqrt{(c-a)^2 + (d-b)^2} \rightarrow 2\text{-D}$$



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

$\boxed{\text{n-dim space}}$

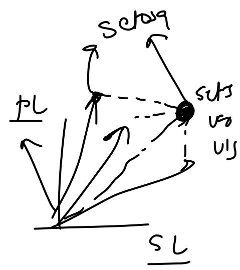
ML algos

K-nearest neighbors - classification

iris

SL | PL | species

→ 2D

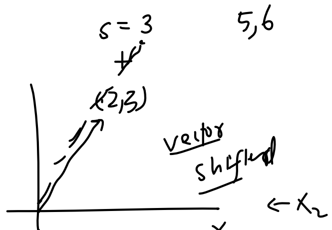


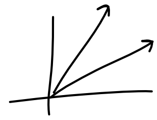
Scalar Addition/Subtraction (Shifting)

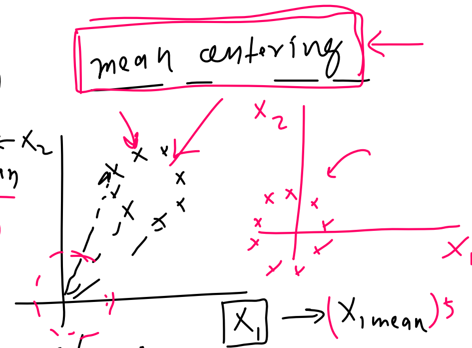
11 April 2023 07:31

Scalar addition/substraction

Scalar multiplication/division







$V = [v_1, v_2, v_3 \dots v_n] \leftarrow n$

$V = [v_1+n, v_2+n, v_3+n, \dots v_n+n]$ ← shifting

$V = [v_1-n, v_2-n, v_3-n \dots v_n-n]$

50 2D vectors / points

vector ← $x_1 - x_{1mean}$, $x_2 - x_{2mean}$ → scalar is

	x_1	x_2
v_1	3	4
v_2	5	6
v_3	7	8

mean → $\frac{15}{3} = 5$; 6

$x_{1\,shifted}$	$x_{2\,shifted}$
-2	-2
0	0
2	2
=0	=0

Mean centering is a useful pre-processing technique in various machine learning applications. It can improve the performance, convergence, and interpretability of the model. Some practical examples where mean centering is applied include:

1. Principal Component Analysis (PCA): PCA is a dimensionality reduction technique that transforms the data into a new coordinate system by identifying the directions (principal components) with the highest variance. Before applying PCA, it is essential to mean center the data to ensure that the first principal component represents the direction with the highest variance in the dataset, rather than being influenced by the location of the data in the coordinate system.
2. Linear regression: In linear regression, mean centering can help improve the interpretability of the model coefficients by making them directly comparable. When the features are mean-centered, the intercept term represents the expected value of the dependent variable when all independent variables are at their mean values. Additionally, mean centering can help with multicollinearity issues, especially when there are interaction terms in the model.
3. Gradient-based optimization algorithms: Some machine learning algorithms, such as gradient descent, can converge faster when the input features are mean-centered. This is because mean centering can lead to better conditioning of the optimization problem, allowing the gradient descent algorithm to take larger, more consistent steps towards the optimal solution.
4. Clustering algorithms: Mean centering can help improve the performance of clustering algorithms like k-means by ensuring that the initial cluster centroids are not heavily influenced by the location of the data in the coordinate system. This can lead to faster convergence and better clustering results.
5. Regularization: In machine learning models that use regularization techniques, such as ridge regression or LASSO, mean centering can help ensure that the regularization term has a consistent effect across all features. By mean centering the features, the model is less likely to penalize the intercept term, which can lead to better generalization.

Scalar Multiplication/Division [Scaling]

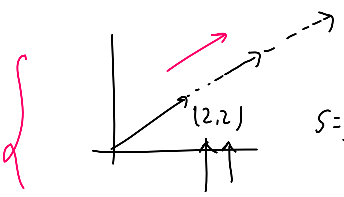
11 April 2023 08:10

$\boxed{15}$

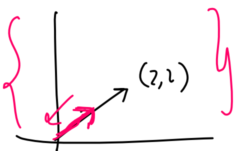
$v = [v_1, v_2, v_3 \dots v_n]$

$v \times s = [v_1 * s, v_2 * s, v_3 * s, \dots v_n * s]$

$\frac{v}{s} = [v_1/s, v_2/s, v_3/s \dots v_n/s]$



s=3



s = 2

direction

magnitude

scale → scalar

vector

$$\frac{X - X_{min}}{X_{max} - X_{min}}$$

scaling

Normalization → $\boxed{0-1}$

age	salary	purchase
32	100000	→ 1
45	250000	→ 0

division → scale

[32, 100000]

32
45
55

[0-1]

max

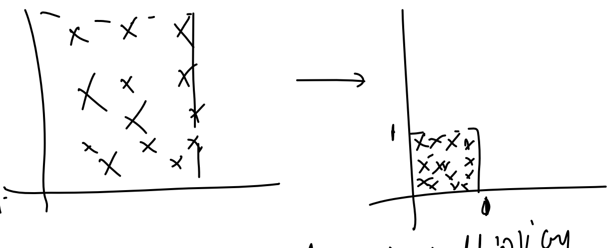
min

0 – 1

0 – 1

0 – 1

1000 people



scalar multiplication

gradient descent

$\boxed{1000}$

$\boxed{w_1}$ age	$\boxed{w_2}$ iq	$\boxed{w_3}$ 10th marks	$\boxed{w_4}$ 12th marks	placement
18	120	90	95	→ 10

$\boxed{1000}$ $\boxed{\text{linear reg}}$ → gradient

$W_{new} = W_{old} - \eta \boxed{\frac{\Delta L}{\Delta W}}$

$\boxed{0.01}$ =

scale your value

$[w_1 \; w_2 \; w_3 \; w_4]$ ⇒ $\boxed{\frac{\partial L}{\partial w_1}}$ $\boxed{\frac{\partial L}{\partial w_2}}$ $\boxed{\frac{\partial L}{\partial w_3}}$ $\boxed{\frac{\partial L}{\partial w_4}}$

Vector Addition/Subtraction

11 April 2023 08:08

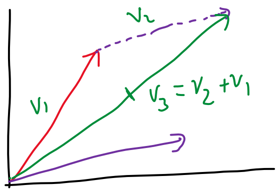
$\boxed{V_1 \quad V_2 \; \ldots \; V_n}$

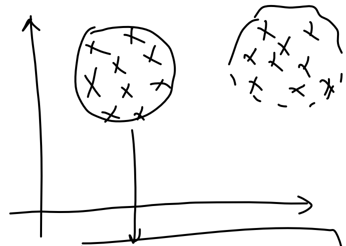
$V_1 = [a_1, a_2, a_3 \ldots a_n]$

$V_2 = [b_1, b_2, b_3 \ldots b_n]$

$V_1 + V_2 = [a_1+b_1, a_2+b_2, a_3+b_3, \ldots a_n+b_n]$

resultant





$\dfrac{V_1 + V_2 + V_3 + \ldots}{n}$ → centroid

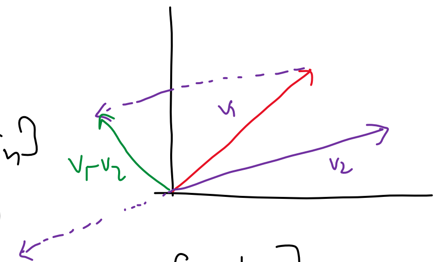
gradient descent

weight $= [w_1, w_2, w_3 \ldots w_n]$

dev $= \left[\dfrac{\partial L}{\partial w_1}, \dfrac{\partial L}{\partial w_2}, \dfrac{\partial L}{\partial w_3} \ldots \dfrac{\partial L}{\partial w_n}\right]$

η

$w_n = w_o - \eta \, \text{dev}$



$v_1 = [a_1, a_2, a_3 \ldots a_n]$

$v_2 = [b_1, b_2 \ldots b_n]$

$v_1 - v_2 = [a_1 - b_1, a_2 - b_2, \ldots a_n - b_n]$

$V_1 + V_2 = V_2 + V_1$

Rules

→ greater 2 vector $\quad V_1 \; V_2 \; V_3 = V_1 + V_2 + V_3$

→ dimension should be same $\quad$ 2D vector ✗ 3D vector

→ $A + B = B + A$

→ $(A + B) + C = A + (B + C)$

Dot Product

11 April 2023 07:32

$A = [a_1, a_2 \ldots a_n]$
$B = [b_1, b_2 \ldots b_n]$

$\rightarrow$ $\boxed{a_1b_1 + a_2b_2 + a_3b_3 + \ldots + a_nb_n}$ =

2 vectors

geometric

matrix multiplication

$1 \times n$ $n \times 1$ $\rightarrow$ $\boxed{1 \times 1}$

Dot product $\rightarrow$ scalar $\leftarrow$ useful

Cross product $\rightarrow$ matrix

vector

$[a_1 \; a_2 \; \ldots \; a_n]_{1 \times n}$ row

$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}_{n \times 1}$ col

$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$ $A^T B$

vectors
$\hookrightarrow$ col vectors

$$\boxed{A \cdot B = A^T B}$$

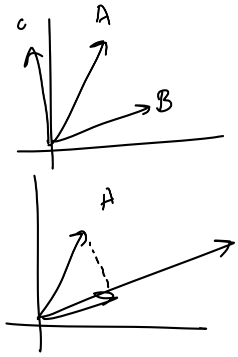
Rules

1) Commutative

$\boxed{A \cdot B = B \cdot A}$ $\leftarrow$ $\boxed{n\text{-}D}$ $\leftarrow$

2) Distributive

$\boxed{A \cdot (B+C) = A \cdot B + A \cdot C}$ $\leftarrow$



proj$_B A$

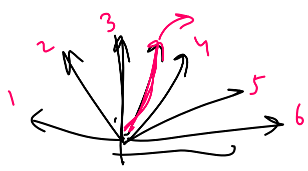
Use

dot product

- $\rightarrow$ Compute similarity between 2 vectors
- $\rightarrow$ projections $\leftarrow$
- $\rightarrow$ perform matrix multiplication

Machine

$\rightarrow$ movies | summary

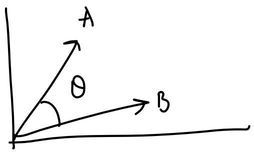


recommend $\rightarrow$ $\boxed{\text{cosine similarity}}$

$\rightarrow$ Deep learning $\rightarrow$ matrix $\rightarrow$ dot product

Angle between 2 vectors

11 April 2023 07:32



$\boxed{A.B = \|A\| \|B\| \cos\theta}$

dist of A from origin

dist of B from origin

A, B both non-zero

1) $A.B = 0 \rightarrow$

$\|A\| \|B\| \cos\theta = 0$

$\boxed{\theta = 90^\circ}$ → A and B dissimilar

SVM

$A \perp B$ → perpendicular → orthogonal

$\underline{A}$ $\underline{B}$ → $A = [a_1\ a_2\ a_3 \dots a_n]$

$B = [b_1\ b_2\ b_3 \dots b_n]$

$\underline{\text{angle}}$

$$\cos\theta = \frac{A \cdot B}{\|A\| \|B\|} \Rightarrow \theta = \cos^{-1}\left\{\frac{A \cdot B}{\|A\| \|B\|}\right\}$$

Cosine similarity ←

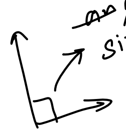
$$\boxed{\cos\theta = \frac{A \cdot B}{\|A\| \|B\|}}$$

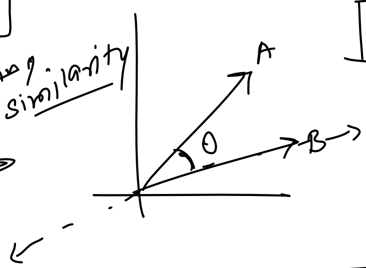
← cos of angle betw A and B

ML → similarity measure

$\boxed{-1 \text{ to } 1}$ $\boxed{-1}$

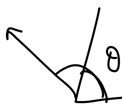
θ -1 . -1 0





$\underline{\theta = 0}$ same direction → $\boxed{\text{similar}}$

$\theta = 180^\circ$ polar opposite

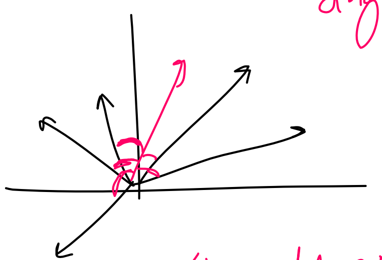


$\boxed{\theta = 90 \quad 0}$

$\underline{\text{movie recommend}}$

$\underline{m}$ | $\underline{s}$ → vectors

{ recommender ML } → Cosine similarity $\underline{\text{dot product}}$



angle small

$\theta = 0$

$\boxed{0\text{-}30\text{-}60\text{-}90}$

Unit Vector

11 April 2023 14:31

Projection of a vector

11 April 2023 07:31

Equation of line in n-D

11 April 2023 07:33

Vector Norms

11 April 2023 07:33