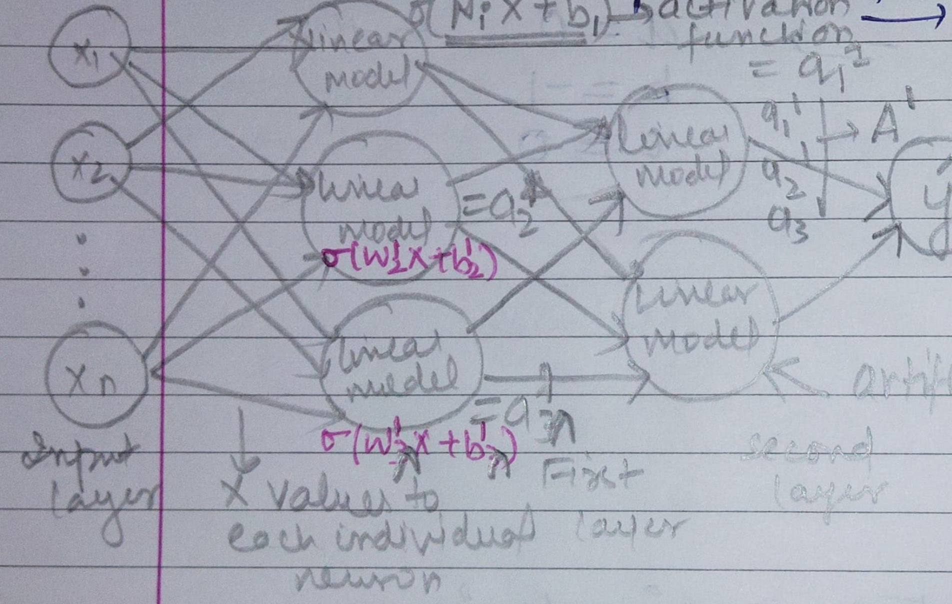
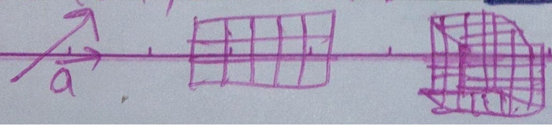
Week 3 (Vectors & Matrices)

Neural Network → one of the powerful ML model for representing non-linear systems

activation function → Under the hood they're really just a large collection of linear models



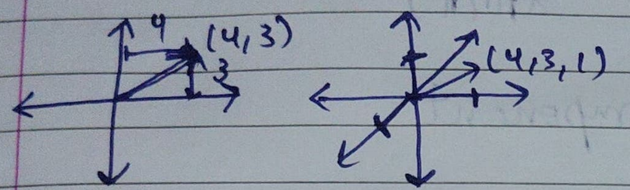
Instead of writing down a zillion little linear eqn to represent a neural network you will instead represent the inputs & outputs of each layer as vectors, matrices & tensors (just like higher dimensional matrices)



you'll apply linear algebra to operate on these vectors matrices & tensors and compute your ML results

① vectors & their properties

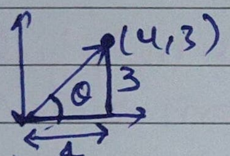
A vector is simply a tuple of numbers. It can be two nos or three no.s representing the no. of dimensions



vectors has magnitude (size) & direction

L_1 norm $= \|(a,b)\|_1 = |a| + |b|$

L_2 norm $= \|(a,b)\|_2 = \sqrt{a^2+b^2}$



$\tan\theta = 3/4 \quad \therefore \theta = 36.87^\circ$

vector notation → Row vector $x = (x_1 \ x_2 \ \ldots \ x_n)$

1, 2 .. n → subscript represents no. of components

→ column vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Other notation

$\vec{x}$ $\quad [x_1, x_2 \ldots x_n]$

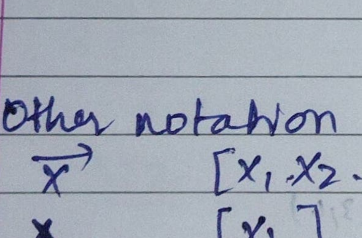
x $\quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$x = (x_1 \ x_2 \ \ldots \ x_n)$

L_1 norm $= \|x\|_1 = |x_1| + |x_2| + \ldots |x_n|$

L_2 norm $= \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots x_n^2}$

② vector operations (added, subtracted)



$u + v = (\cancel{4}\ 4+1, 1+3) = (5,4)$

The sum of vector is precisely diagonal in a $\parallel^{gram}$ formed by the vector u & v.

$u - v = (4-1, 1-3) = (3,-2)$

General definition: Sum & difference

$X = (x_1 \ x_2 \ \dots \ x_n)$ $\quad y = (y_1 \ y_2 \ \dots \ y_n)$

same no. of components

Sum

$x + y = (x_1 + y_1, \ x_2 + y_2 \ \dots \ x_n + y_n)$

Sum component by component.

difference

$x - y = (x_1 - y_1 \ \ x_2 - y_2 \ \dots \ x_n - y_n)$

subtract component by component

The difference between two vectors is helpful to tell how far apart 2 vectors are from eachother

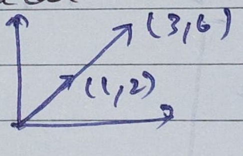
$L_1 = |U - V| = |5| + |-3| = 8$

$L_2 = |U - V|_2 = \sqrt{5^2 + 3^2} = 5.83$

multiplying a vector by a scalar

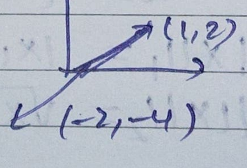
$u = (1,2)$ $\quad \lambda U = (3,6)$

$\lambda = 3$



$u = (1,2)$ $\quad \lambda u = (-2,-4) =$

$\lambda = -2$



multiplication by a scalar

$x = (x_1 \ x_2 \ \dots \ x_n)$

$\lambda x = (\lambda x_1 \ \lambda x_2 \ \dots \ \lambda x_n)$

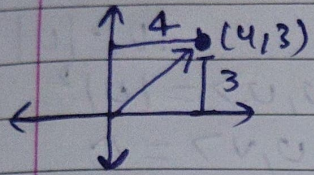
3 Dot Product

$\underset{R}{\boxed{1 \,|\, 2 \,|\, 3}}$ $\qquad$ $\begin{bmatrix}4\\5\\6\end{bmatrix} C$

$u = (1, 2, 3) \quad v = (4, 5, 6)$

$$u \cdot v = (1 \times 4 + 2 \times 5 + 3 \times 6)$$
$$= (4 + 10 + 18) = 32$$

Norm of a vector using dot Product



$\sqrt{4^2 + 3^2} = \sqrt{25} = 5$

$L_2 \text{ norm} = \sqrt{\text{dot product}(v, v)}$

$= \sqrt{\langle v, v \rangle}$

$$\boxed{|v|_2 = \sqrt{\langle v, v \rangle}}$$

Transpose (Rows vector → column vector or vice-versa)

$$\begin{bmatrix}2\\4\\1\end{bmatrix}^T = [2 \ 4 \ 1] \quad \text{or} \quad [2 \ 4 \ 1]^T = \begin{bmatrix}2\\4\\1\end{bmatrix}$$

Matrix → (2×3) (col → Rows)

$$\underset{3\times2}{\begin{bmatrix}2 & 5\\4 & 7\\1 & 3\end{bmatrix}} = \underset{2\times3}{\begin{bmatrix}2 & 4 & 1\\5 & 7 & 3\end{bmatrix}}$$

General definition : dot Product

$x = (x_1 \ x_2 \ \dots \ x_n) \quad y = (y_1 \ y_2 \ \dots \ y_n)$

same no. of components

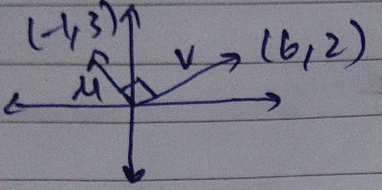
$$x \cdot y = (x_1 \times y_1) + (x_2 \times y_2) + (x_3 \times y_3) + \dots + (x_n \times y_n)$$

$\langle x, y \rangle$ is another notation of dot Product

$$x \cdot y^T = (x_1 \ x_2 \dots x_n) \cdot \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$$

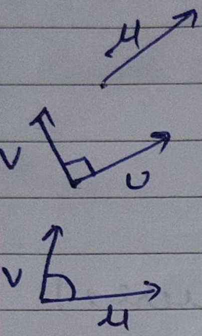
4. Geometric representation of dot Product.

Orthogonal vectors have dot Product 0



$u \cdot v = (-1,3)\cdot(6,2) = -6+6 = 0$

$|u|\cdot|u|$

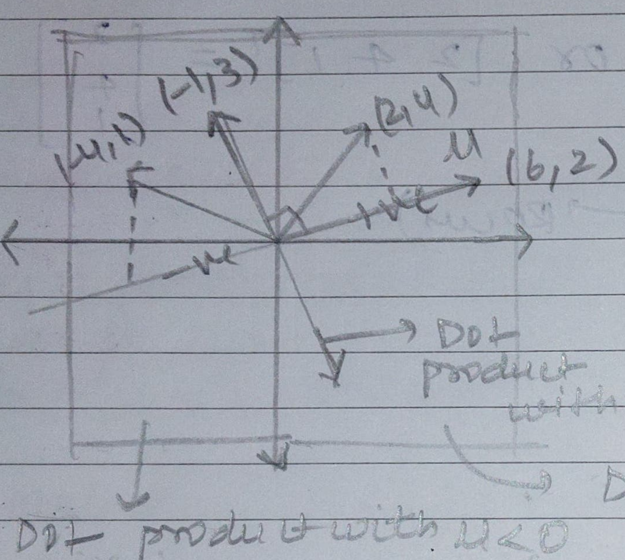


dot product of a vector $\langle u,u\rangle = |u|^2$

dot product of 2 vector $\langle u,v\rangle = 0$

dot product of 2 vector $\langle u,v\rangle = |u||v|\cos\theta$

~~[illegible]~~ ~~[illegible]~~



$(6\ \ 2)\cdot\begin{pmatrix}2\\4\end{pmatrix} = 20$ +ve

$(6\ \ 2)\begin{pmatrix}-1\\3\end{pmatrix} = 0$

$(6\ \ 2)\begin{pmatrix}-4\\1\end{pmatrix} = -22$ -ve

(5) Multiplying a ~~vector~~ matrix by a vector.

equations as dot product

$2a + 4b + c = 28$

$$\begin{bmatrix}2 & 4 & 1\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = 28$$

row vector

column vector

dot product

$a + b + c = 10$ $\qquad$ $a + 2b + c = 15$ $\qquad$ $a + b + 2c = 12$

$[1 \ 1 \ 1] \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 10$ $\qquad$ $[1 \ 2 \ 1] \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 15$ $\qquad$ $[1 \ 1 \ 2] \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 12$

$\Downarrow$

there column vector is same

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 12 \end{bmatrix}$$ Matrix Product

3×3 $\quad$ 3×1 $\quad$ 3 $\qquad$ # columns = length of a vector

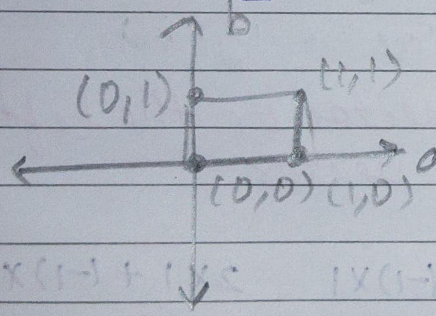
length

Matrix can be rectangular

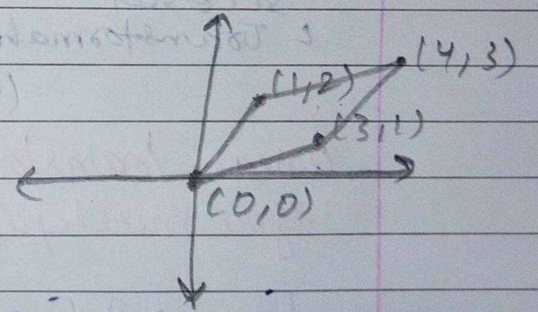
4×3 $\quad$ 3×1 $= 4$

⑥ Matrices as linear transformations

$$\begin{array}{cc} a & b \\ 3 & 1 \\ 1 & 2 \end{array}$$



$(0,0) \to (0,0)$
$(1,0) \to (3,1)$
$(0,1) \to (1,2)$
$(1,1) \to (4,3)$



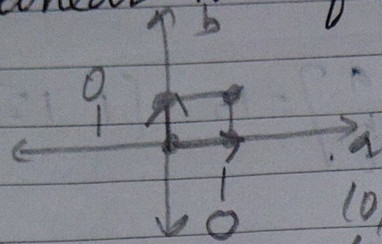
$$\begin{array}{cc} a & b \\ 3 & 1 \\ 1 & 2 \end{array} \begin{array}{c} \\ 0 \\ 0 \end{array} = \begin{array}{c} \\ 0 \\ 0 \end{array} \qquad \begin{array}{cc} 3 & 1 \\ 1 & 2 \end{array} \begin{array}{c} 1 \\ 1 \end{array} = \begin{array}{c} 4 \\ 3 \end{array}$$

$$\begin{array}{ccc} a & b & c \\ 3 & 1 & 1 \\ 1 & 2 & 0 \end{array} = \begin{array}{c} \\ 3 \\ 1 \end{array}$$

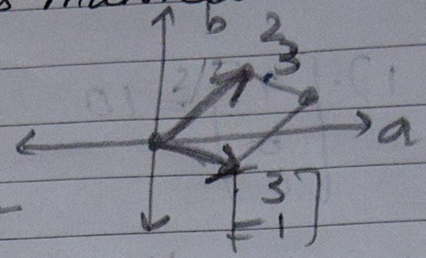
$$\begin{array}{ccc} 3 & 1 & 0 \\ 1 & 2 & 1 \end{array} = \begin{array}{c} 1 \\ 2 \end{array}$$

PAGE NO.:
DATE: / /

(7) linear transformation as matrices



a b
9 9
9 9



$(0,0) \to (0,0)$
$(1,0) \to (3,-1)$
$(0,1) \to (2,3)$
$(1,1) \to (5,2)$

$\begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix} \begin{matrix} 0 \\ 1 \end{matrix} = \begin{matrix} 2 \\ 3 \end{matrix}$

$\begin{matrix} 9 & 9 \\ 9 & 9 \end{matrix} \quad \begin{matrix} 1 \\ 0 \end{matrix} = \begin{matrix} 3 \\ -1 \end{matrix}$

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

fundamental.
you only look two coordinates
(0,1) & (1,0) (bases vector)
goes to
vector with coordinates (1,0) ⟶ (3,−1)
2nd col of matrix
first colum of matrix

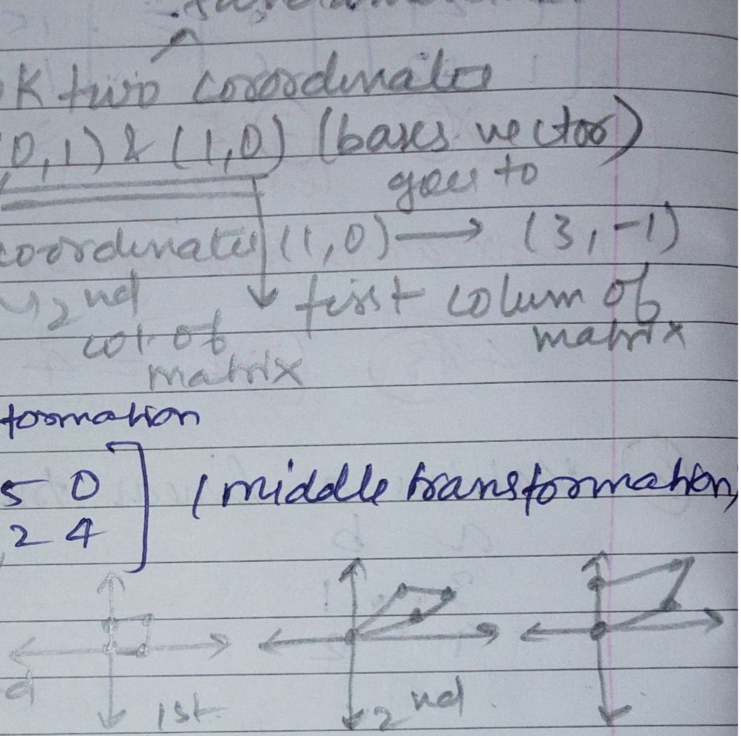
(8) Matrix Multiplication

First L Transformation

(right) $\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 2 & 4 \end{bmatrix}$ (middle transformation)
(left)

second L Transformation

from 1st to 3rd



linear transformation acts on vector on the left so you multiply matrix times vector

$$\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2\times3+(-1)\times1 & 2\times1+(-1)\times2 \\ 0\times3+2\times1 & 0\times1+2\times2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 2 & 4 \end{bmatrix}$$

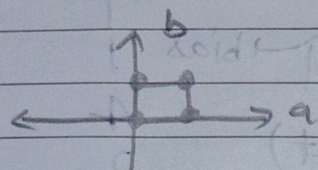
matrix multiplication can be seen it as combining two linear transformations into a 3rd linear transformation.

$2 \times 3 \quad 3 \times 4$
same
answer matrix

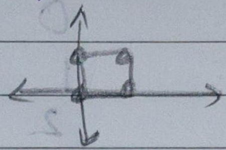
⑨ The identity matrix (ones in the diagonal, zeroes everywhere else) $\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$

matrix multiplied by identity matrix results into same matrix

Linear Transformation also stays same



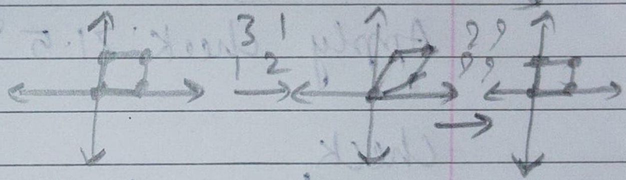
$\begin{matrix}1&0\\0&1\end{matrix} \cdot \begin{matrix}x\\y\end{matrix} = \begin{matrix}x\\y\end{matrix}$



(0,0) → (0,0)
(1,0) → (1,0)
(0,1) → (0,1)
(1,1) → (1,1)

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

⑩ Matrix Inverse



$\begin{matrix}3&1\\1&2\end{matrix} \cdot \begin{matrix}a&b\\c&d\end{matrix} = \begin{matrix}1&0\\0&1\end{matrix}$

↳ inverse

$\begin{bmatrix}3&1\\1&2\end{bmatrix}^{-1} = \begin{matrix}a&b\\c&d\end{matrix} = \begin{matrix}2/5&-1/5\\-1/5&3/5\end{matrix}$

⇒ $3a+c=1$ $a=2/5$
$3b+d=0$ $b=-1/5$
$a+2c=0$ $c=-1/5$
$b+2d=1$ $d=3/5$

⑪ Which matrices have inverse?
non-singular matrix always have inverse
(Invertible matrix)   Det ≠ 0

Singular matrix ~~always~~ does not have inverse
(non-invertible matrix)   Det = 0. → zero

⑫ Neural networks & matrices

Classifiers uses matrix multiplications

Bias	Spam	Lottery	Win	Prod
1	yes	1	1	2
1	yes	2	1	3
1	No	0	0	0
1	yes	0	2	2
1	No	0	1	1
1	No	1	0	1
1	yes	2	2	4
1	yes	2	0	2
1	yes	1	2	3

· model =
1
1
-1.5 → bias
(trained model output)

Apply check: >1.5? (threshold)

Check

1· win + 1· Lottery > 1.5 ← threshold

1· Win + 1· Lottery $\underbrace{-1.5}_{\text{bias}}$ > 0

to add this in matrix multiplication add this in table

The AND Operator

AND	X	Y	model	Dot Product	Check
No	0	0	1	0	No
No	1	0	1	1	No
No	0	1		1	No
Yes	1	1		2	Yes

Check: >1.5

The AND dataset can be modelled as a perceptron of onelayer Neural Network

The perceptron