Coursera → mathematics for machine learning (Linear Algebra)

→ module 1

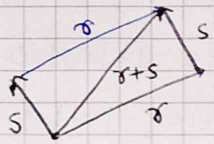
$$2a + 3b = 8$$
$$10a + 8b = 16$$

$$\begin{pmatrix} 2 & 3 \\ 10 & 8 \end{pmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$$

★ In data science, we describe a vector which describes the attributes of an object.

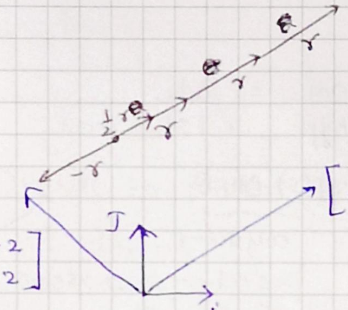
Eg  120 sqm² 1 bathroom 2 bedroom $\begin{bmatrix} 120 \\ 2 \\ 1 \end{bmatrix}$

→ Vector addition



$r + s = s + r$

Vector multiplication with a scalar quantity



ar, $\frac{1}{2}r$, $-r$ → going in opposite direction

$\begin{bmatrix} 3 \\ 2 \end{bmatrix} = r$

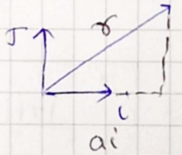
→ Coordinate system

$s = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

$\left.\begin{array}{l}(r+s)+t = \\ r+(s+t)\end{array}\right\}$ associativity.

→ module 2

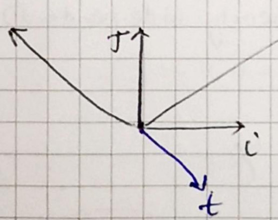
→ Vector length → $|r| = \sqrt{a^2+b^2}$



$r = \begin{bmatrix} a \\ b \end{bmatrix}$

→ vector multiplication → let's say we have 2 vectors namely r and s, the components of which are represented by (r_i, r_j) and (s_i, s_j) resp.

$s = \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} s_i \\ s_j \end{bmatrix}$



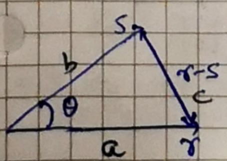
$r = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} r_i \\ r_j \end{bmatrix}$

$\boxed{r \cdot s = r_i s_i + r_j s_j} = s \cdot r$ → commutative

$\boxed{r \cdot (s+t) = r \cdot s + r \cdot t}$ → distributive

$\boxed{r \cdot (as) = a(r \cdot s)}$ → associative over scalar multiplication.

→ Cosine rule



$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

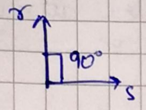
$$|r-s|^2 = |r|^2 + |s|^2 - 2|r||s|\cos\theta \quad \ldots\ldots (i)$$

$$(r-s)(r-s) = r \cdot r - s \cdot r - r \cdot s + s \cdot s$$
$$= |r|^2 - 2\, r \cdot s + |s|^2 \quad \ldots\ldots (ii)$$

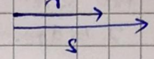
when comparing eq. (i) and (ii) we get

$$-\cancel{2}\, s\cdot r = -\cancel{2}|r||s|\cos\theta$$

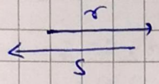
$\boxed{r\cdot s = |r||s|\cos\theta}$ → This shows the direction in which both these vectors are pointing.

Eg. 

$r\cdot s = |r||s|\cos 90°$
$= 0$ (never meet)



$r\cdot s = |r||s|\cos 0$
$= |r||s|$
(same direction)

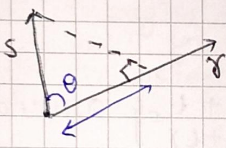


$r\cdot s = |r||s|\cos(180)$

$r\cdot s = -|r||s|$ (opposite direction)

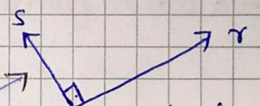
→ Vector projection.

$$\cos\theta = \frac{adj}{hyp} = \frac{adj}{|s|}$$

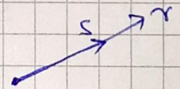


$$r\cdot s = |r|\underbrace{|s|\cos\theta}_{\text{adjacent}}$$

projection or shadow of s on r



no projection ($\theta = 90°$)



full projection ($\theta = 0$)

$\boxed{\frac{r\cdot s}{|r|} = |s|\cos\theta}$ → Scalar projection

$\boxed{r\cdot\frac{r\cdot s}{|r||r|} = \frac{r\cdot s}{r\cdot r}\, r}$ → Vector projection

→ Basis is a set of n vectors that:

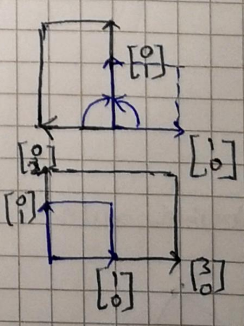
(i) are not linear combinations of each other (linearly independent)
(ii) span the space (that they describe)
→ The space is then n-dimensional.

Matrices.

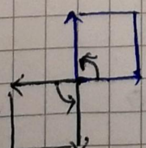
① Identity matrix $\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\y\end{bmatrix}$

Eg. $\begin{bmatrix}-1&0\\0&2\end{bmatrix}$

$\begin{bmatrix}3&0\\0&2\end{bmatrix}$



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



inversion

→ linearly dependent matrices

$$\begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 2 & 3 & 7 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 12 \\ 17 \\ 29 \end{pmatrix} \quad \left\{ \begin{matrix} row_1 + row_2 = row_3 \\ 2 \times col_1 + col_2 = col_3 \end{matrix} \right\}$$

when trying to solve this and do back substitution for calculating the inverse:

$$\underbrace{\begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 12 \\ 5 \\ 0 \end{pmatrix}} \rightarrow \quad 0c = 0$$

i.e. no solution.

det = 0, we cannot solve the system. i.e. no inverse

$\boxed{\text{Singular matrix}}$

$\boxed{\text{determinant} \rightarrow \text{change in area}}$

→ Einstein convention for matrix multiplication. (Summation convention)

$$\begin{array}{ccc} A & B & AB \\ \begin{pmatrix} a_{11} & a_{12} & a_{13} \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} & \begin{pmatrix} b_{11} & b_{12} & b_{13} \cdots & b_{1n} \\ b_{21} & & & \vdots \\ \vdots & & & \\ b_{n1} & \cdots & \cdots & b_{nn} \end{pmatrix} = & \begin{pmatrix} & \end{pmatrix} \end{array}$$

let's say we have to calculate the product of $(ab)_{23}$

$$(ab)_{23} = a_{21} b_{13} + a_{22} b_{23} + \cdots a_{2n} b_{n3}$$

$$\boxed{ab_{ik} = \sum_j a_{ij} b_{jk} = a_{ij} b_{jk}}$$

This helps especially when we are computing. we can go in loops and calculate the elements of the final matrix instead of computing individually.

→ General case of matrix multiplication.

$AB = C$

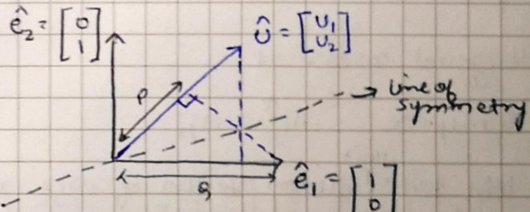
$C_{ik} = A_{ij} B_{jk}$

some no.

$$\text{Eg. } \begin{array}{c} \\ 2 \end{array}\overset{3}{\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}} \overset{4}{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}\begin{array}{c} 3 \end{array} = \begin{array}{c} 2 \end{array}\overset{4}{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}$$

$$\underbrace{\overset{u}{\begin{bmatrix} u_i \end{bmatrix}} \cdot \overset{v}{\begin{bmatrix} v_i \end{bmatrix}}}_{u_i v_i} \sim \begin{bmatrix} u_1 & u_2 & u \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Vector



which implies that dot product is symmetric, the projection is symmetric

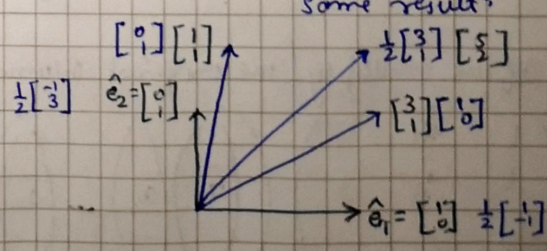
$\boxed{P = Q}$. If we do the same with $\hat{e}_2$, we will get the same result.

→ Change of Basis.

→ let's take an example where we have Pandas.

→ In my world, the Pandas basis vectors are $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

→ my basis vectors are $\hat{e}_1$ and $\hat{e}_2$



→ In Pandas world, its basis vectors are $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$

→ So, Bear's basis vectors are $\begin{bmatrix}3\\1\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$ in my frame.

→ Bear's transformation matrix = $\begin{bmatrix}3 & 1\\1 & 1\end{bmatrix}$

→ Now, let's say we have a vector which we need to transform (depicted in center) which (in bear's world) is $\frac{1}{2}\begin{bmatrix}3\\1\end{bmatrix} \sim \begin{bmatrix}3/2\\1/2\end{bmatrix}$

→ Transforming this into my world, we get

$$\underbrace{\begin{bmatrix}3 & 1\\1 & 1\end{bmatrix}}_{\text{Bear's basis in my coordinates}}\underbrace{\begin{bmatrix}3/2\\1/2\end{bmatrix}}_{\text{Bear's vector}} = \underbrace{\begin{bmatrix}5\\2\end{bmatrix}}_{\text{my vector}} \text{(in my world)}$$

} Here, we have translated the bear's world into my world.
Usually we have to go the other way round.

→ Performing inverse.

→ If we say that the bear's matrix is "B", it's inverse is B^{-1}

$$B = \begin{bmatrix}3 & 1\\1 & 1\end{bmatrix} \qquad B^{-1} = \frac{1}{2}\begin{bmatrix}1 & -1\\-1 & 3\end{bmatrix}$$ } my basis in bear's world.

i.e. my $\begin{bmatrix}1\\0\end{bmatrix}$ ie. $\hat{e}_1$ will be $\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}$
$\begin{bmatrix}0\\1\end{bmatrix}$ i.e. $\hat{e}_2$ will be $\frac{1}{2}\begin{bmatrix}-1\\3\end{bmatrix}$ } in bear's system.

→ If we multiply my basis in bear's world with my vector, we will get the world in bear's basis i.e

$$\frac{1}{2}\begin{bmatrix}1 & -1\\-1 & 3\end{bmatrix}\begin{bmatrix}5\\2\end{bmatrix} = 5\begin{bmatrix}1/2\\-1/2\end{bmatrix} + 2\begin{bmatrix}-1\\3\end{bmatrix} = \underbrace{\frac{1}{2}\begin{bmatrix}3\\2\end{bmatrix}}_{\text{Bear's Vector}}$$

⇒ Orthogonal matrices

→ Transpose $A_{ij}^T = A_{ji}$

→ let's take unit column vectors of size n×n ($A_{n\times n}$) $\left(\begin{pmatrix}\hat{a}_1\end{pmatrix}\begin{pmatrix}\hat{a}_2\end{pmatrix}\cdots\begin{pmatrix}\hat{a}_n\end{pmatrix}\right)$ such that
a_{i1} a_{i2} a_{in}

$$a_1 \cdot a_2 = 0 \sim a_i \cdot a_j = 0 \text{ if } i \neq j$$
$$= 1 \text{ if } i = j$$
} → orthonormal condition.

→ If we multiply the original column matrix with its transpose, we get.

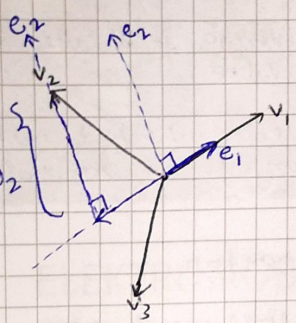
$$\underbrace{\begin{pmatrix}(& \hat{a}_1 &)\\(& \hat{a}_2 &)\\ & \vdots & \\(& \hat{a}_n &)\end{pmatrix}}_{A^T}\underbrace{\left(\begin{pmatrix}\hat{a}_1\end{pmatrix}\begin{pmatrix}\hat{a}_2\end{pmatrix}\cdots\begin{pmatrix}\hat{a}_n\end{pmatrix}\right)}_{A} = \underbrace{\begin{pmatrix}1 & 0 & 0 & \cdots & 0\\0 & 1 & 0 & \cdots & 0\\0 & 0 & 1 & \cdots & 0\\\vdots & \vdots & \vdots & & \vdots\\0 & 0 & 0 & \cdots & 1\end{pmatrix}}_{I}$$

→ If the vectors are orthonormal then $\boxed{A^{-1} = A^T}$. In that case "A" is called an "orthogonal matrix"

→ Also $\boxed{|A| = \pm 1}$

→ Constructing orthonormal basis

Gram-Schmidt → let's say we have vectors " $V = \{v_1, v_2 \dots v_n\}$ " which span the entire space. They are also linearly independent



→ let $\underline{v_1}$ • $e_1 = \frac{v_1}{|v_1|}$ } e_1 is some normalised version of v_1

$\underline{v_2} = \underbrace{(v_2 \cdot e_1)}_{\text{vector projection of } v_2 \text{ on } e_1} \simeq (v_2 \cdot e_1) \frac{e_1}{|e_1|} + u_2$

vector

e_1 is unit length therefore neglecting.

$$u_2 = v_2 - (v_2 \cdot e_1) e_1$$

if we normalise u_2, we will get $\boxed{\frac{u_2}{|u_2|} = e_2}$

Since v_3 is not in the plane of v_1 and v_2, we can project v_3 into their plane so,

$$v_3 - (v_3 \cdot e_1) e_1 - (v_3 \cdot e_2) e_2 = u_3$$ → perpendicular to the plane

if we normalise u_3 $\simeq$ $\boxed{\frac{u_3}{|u_3|} = e_3}$

Eg. let's say we have 3 vectors v_1, v_2 and v_3 such that the two vectors (v_1 and v_2) are in the plane and v_3 is out of plane. i.e.

$$v_1 = \begin{pmatrix}1\\1\\1\end{pmatrix} \quad v_2 = \begin{pmatrix}2\\0\\1\end{pmatrix} \quad v_3 = \begin{pmatrix}3\\1\\-1\end{pmatrix}$$

$$e_1 = \frac{v_1}{|v_1|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$u_2 = v_2 - (v_2 \cdot e_1) e_1 = \begin{bmatrix}2\\0\\1\end{bmatrix} - \left[\begin{bmatrix}2\\0\\1\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} \frac{1}{\sqrt{3}}\right] \frac{1}{\sqrt{3}} \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}2\\0\\1\end{bmatrix} - \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\0\end{bmatrix}$$

$$e_2 = \frac{u_2}{|u_2|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix}$$

$$u_3 = v_3 - (v_3 \cdot e_1) e_1 - (v_3 \cdot e_2) e_2 = \begin{bmatrix}3\\1\\-1\end{bmatrix} - \left[\begin{pmatrix}3\\1\\-1\end{pmatrix} \cdot \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix} \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\right] - \left[\begin{pmatrix}3\\1\\-1\end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}\right]$$

$$= \begin{bmatrix}1\\1\\-2\end{bmatrix}$$

$$e_3 = \frac{u_3}{|u_3|} = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix}$$

If we write our transformation matrix "$\mathcal{E}$" described by the basis vectors,

$$\mathcal{E} = \left(\begin{pmatrix} e_1 \end{pmatrix} \begin{pmatrix} e_2 \end{pmatrix} \begin{pmatrix} e_3 \end{pmatrix} \right)$$

→ Eigen vectors and Eigenvalues.

↳ characteristics.

① If when applying any transformation, the vectors does not change, it is called an eigen vector and the corresponding value is called an eigen value.

② when in 3d, if we rotate a (square, etc) and we find a vector which is not changing, we say that it is an eigenvector. more importantly, we can also say that it is its axis of rotation.

→ creating eigenvalues.

→ If we have an eigenvector (x), then.

$$\underbrace{Ax}_{\text{applying some transformation to "}x\text{"}} = \underbrace{\lambda x}_{\text{applying some scaling factor to "}x\text{"}}$$

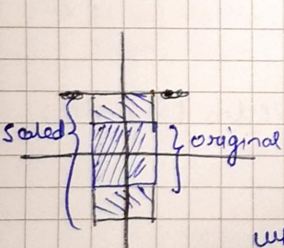
A → n-dimensional transform

x → n-dimensional vector

→ $\underbrace{(A-\lambda I)}_{\text{solving for } 0} x = 0$ $\{$ if $x=0$ → trivial solution $\}$

$$\det(A-\lambda I) = 0 \longrightarrow \det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) = 0$$

$$\rightarrow \lambda^2 - (a+d)\lambda + ad - bc = 0$$



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

solving ↳ $\det\begin{pmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) = 0$

when substituting both values and solving we get.

when $\lambda = 1$ $\begin{pmatrix} 1-1 & 0 \\ 0 & 2-1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \sim \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = 0$

when $\lambda = 2$ $\begin{pmatrix} 1-2 & 0 \\ 0 & 2-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \sim \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 \\ 0 \end{pmatrix} = 0$

from this we can say that

$\boxed{@\lambda = 1 : x = \begin{pmatrix} t \\ 0 \end{pmatrix}}$

at $\lambda = 1$, the eigenvector "x" can be anything along the horizontal axis as long as its vertical axis is 0. we put it to "t"

$\boxed{@\lambda = 2 : x = \begin{pmatrix} 0 \\ t \end{pmatrix}}$