SINGULAR VALUE DECOMPOSITION

diagonal

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^T$$

orthogonal matrices

STUNNING!

WHAT IS LINEARITY?

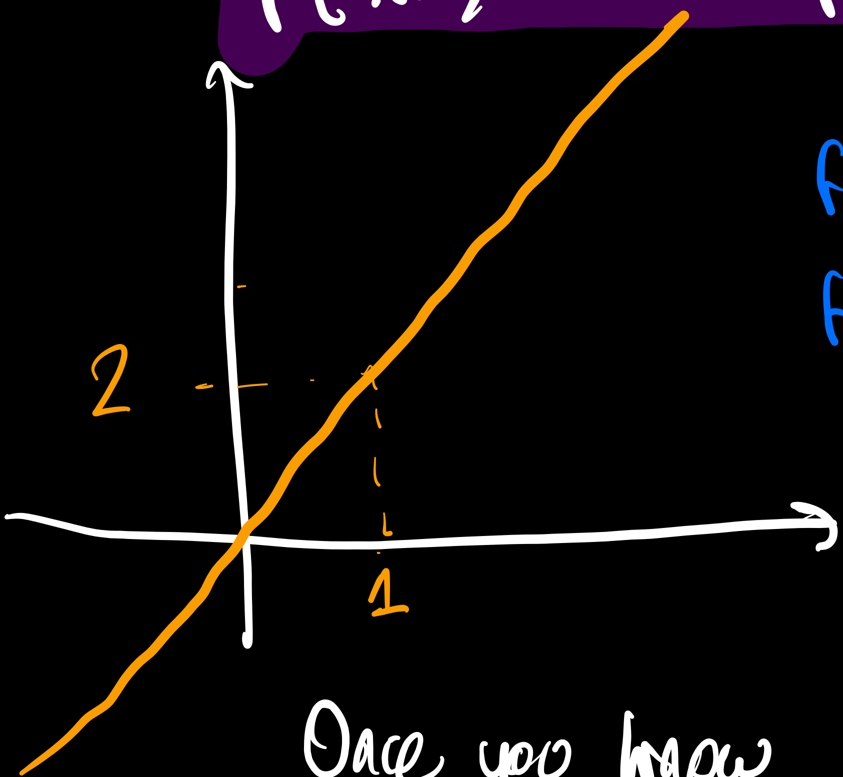
Linear systems

↳ whole = sum of their parts

price = number × price of apple

$$f(x+y) = f(x) + f(y)$$

$$f(\lambda x) = \lambda f(x)$$



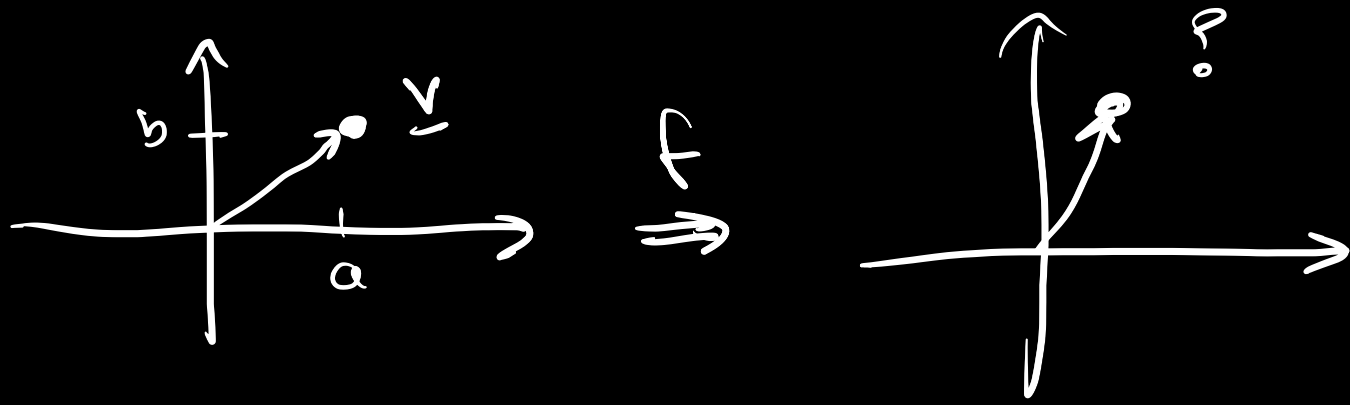
STRAIGHT LINE

$f(2) = 2f(1) = 8$

$f(0) = 0f(1) = 0$

Once you know $f(1)$ you know everything!

MATRICES = LINEAR FUNCTIONS IN ≥ 1 DIMENSIONS

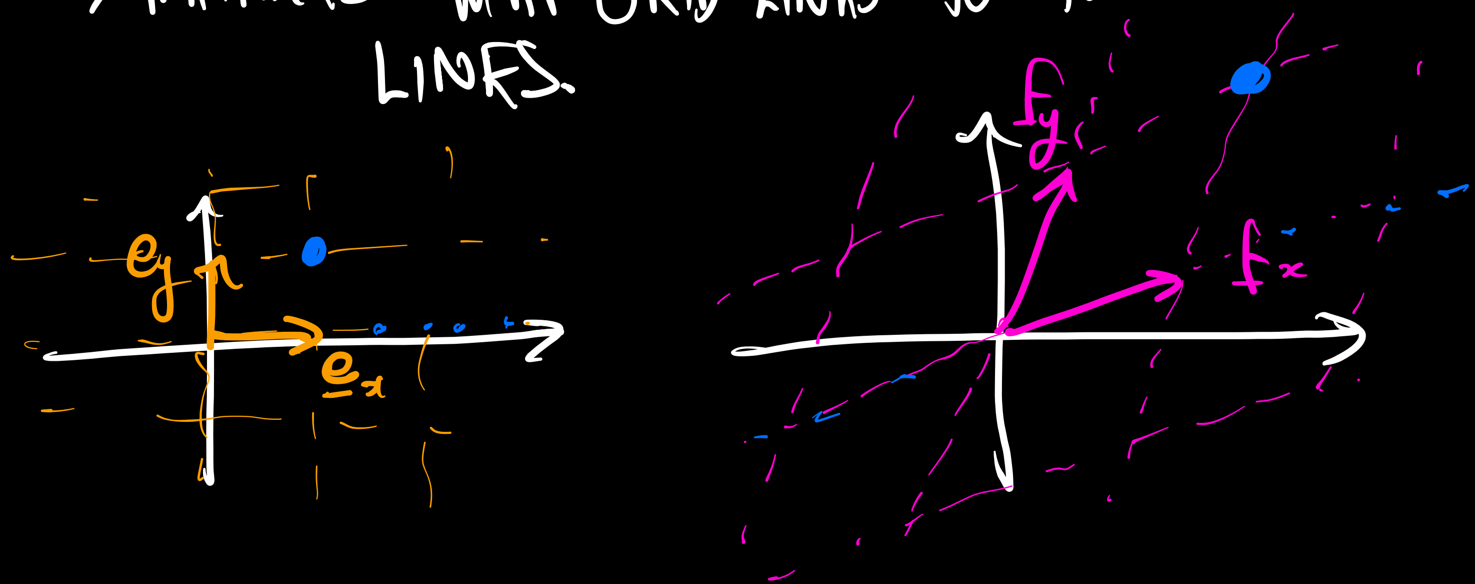


$$f\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = f\left(a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$

$$= a\,\underbrace{f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)}_{\underline{f}_x} + b\,\underbrace{f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)}_{\underline{f}_y}$$

$$\therefore\quad f(\underline{v}) = v_x \underline{f}_x + v_y \underline{f}_y$$

$\Rightarrow$ MATRICES MAP GRID LINES TO TWISTED GRID LINES.



Represent as matrix

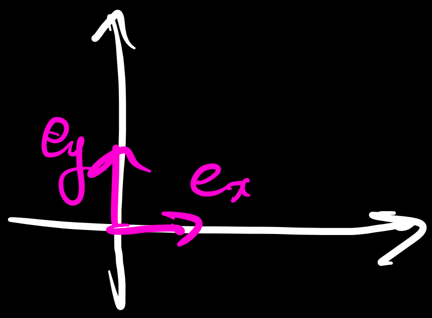
$$F(\underline{v}) = \begin{pmatrix} \underline{f}_x & \underline{f}_y \end{pmatrix}\begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

$\therefore$ columns of matrix = where unit vectors go

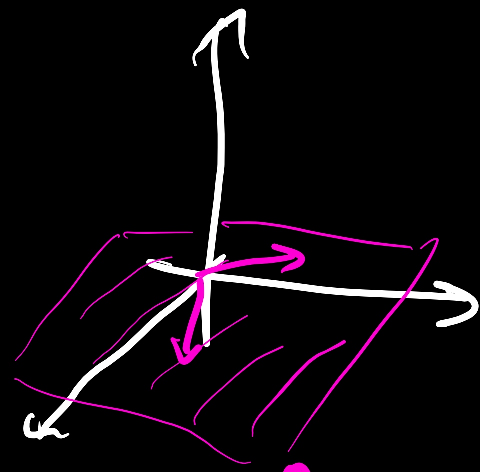
output is linear combination

optimise to deduce

What about 2D ⇒ ND

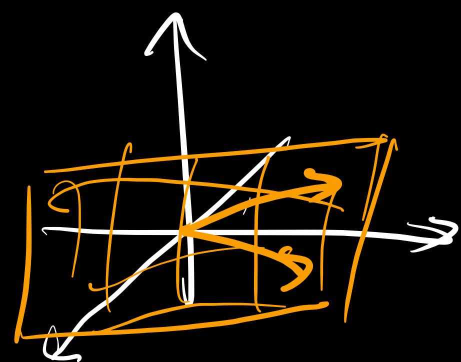


⇒



e.g. $\underline{\underline{M}} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 0 & 0 \end{pmatrix}$ output lives in plane

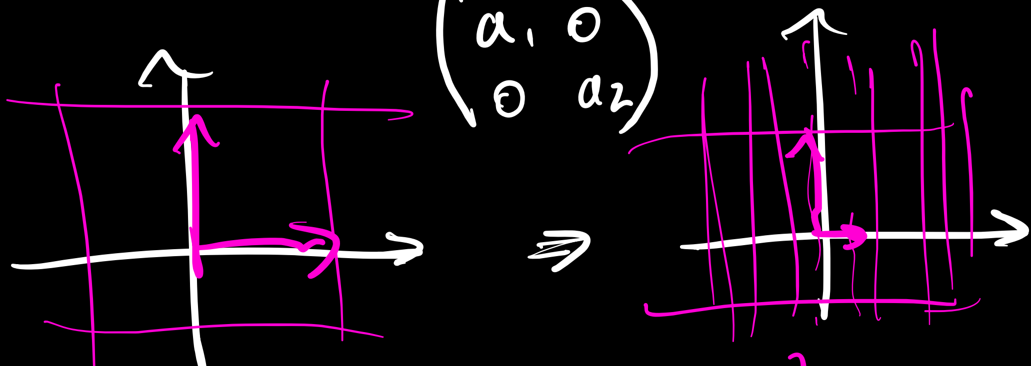
⇒



Some matrices

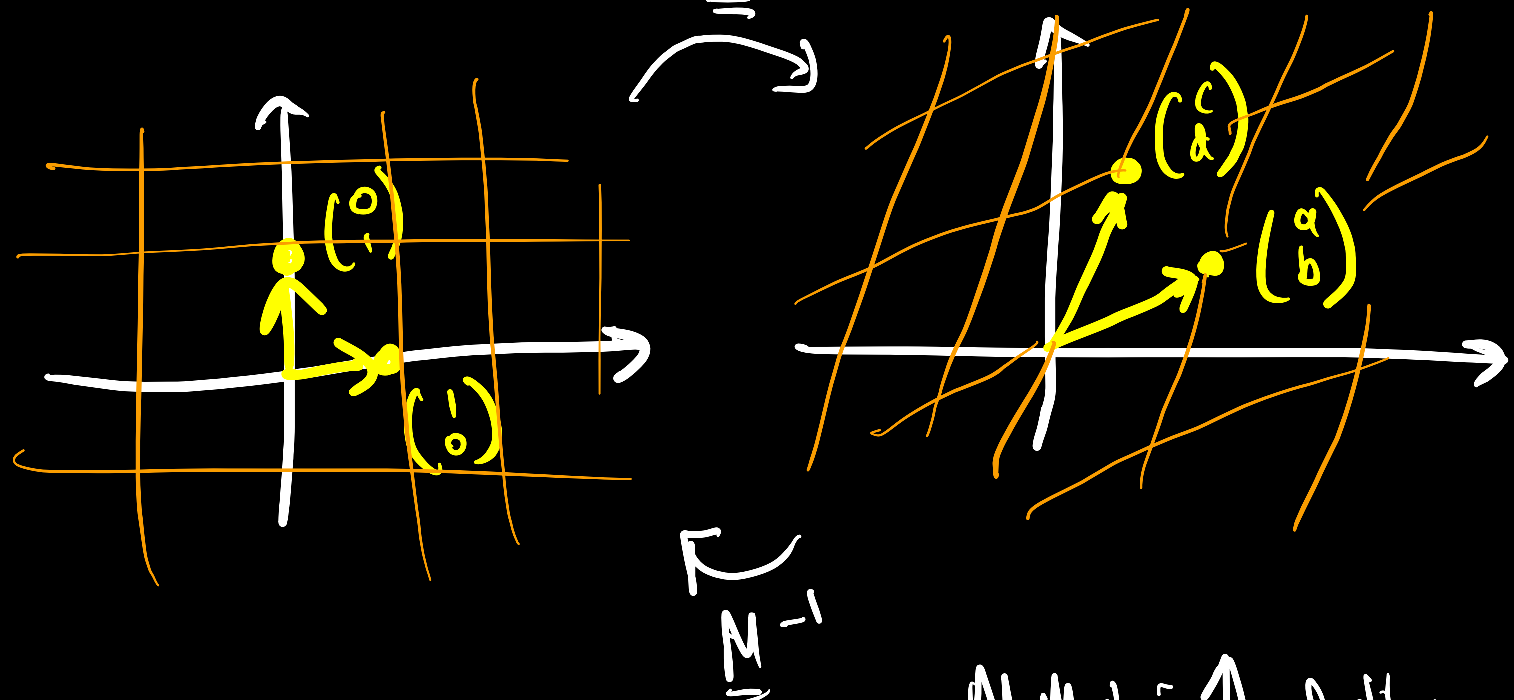
Diagonal:

$$\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$



SQUASH AXES!

Inverse:



$\underline{\underline{M}}\,\underline{\underline{M}}^{-1} = \underline{\underline{1}}$ identity

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

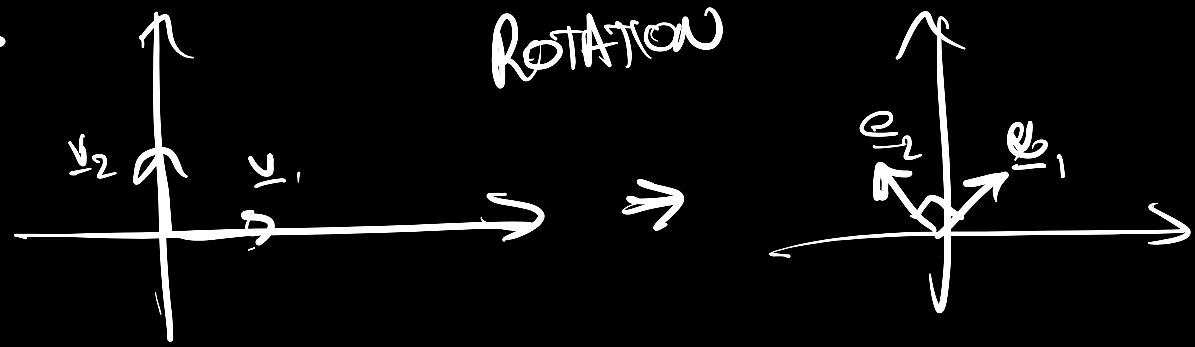
takes you back!

Often doesn't exist!

→ e.g. 2D → 2D, one projection

2D $\Rightarrow$ 3D. DO THESE

Orthogonal



$$\underline{\underline{Q}} = \left(\underline{e}_1 \;\; \underline{e}_2 \right)$$

Check transpose & dot product

$$\underline{\underline{Q}}^T \underline{\underline{Q}} = \begin{pmatrix} - \underline{e}_1 - \\ - \underline{e}_2 - \end{pmatrix} \begin{pmatrix} | & | \\ \underline{e}_1 & \underline{e}_2 \\ | & | \end{pmatrix}$$

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

$\Rightarrow$ transpose is inverse! $\underline{\underline{Q}}^T = \underline{\underline{Q}}^{-1}$

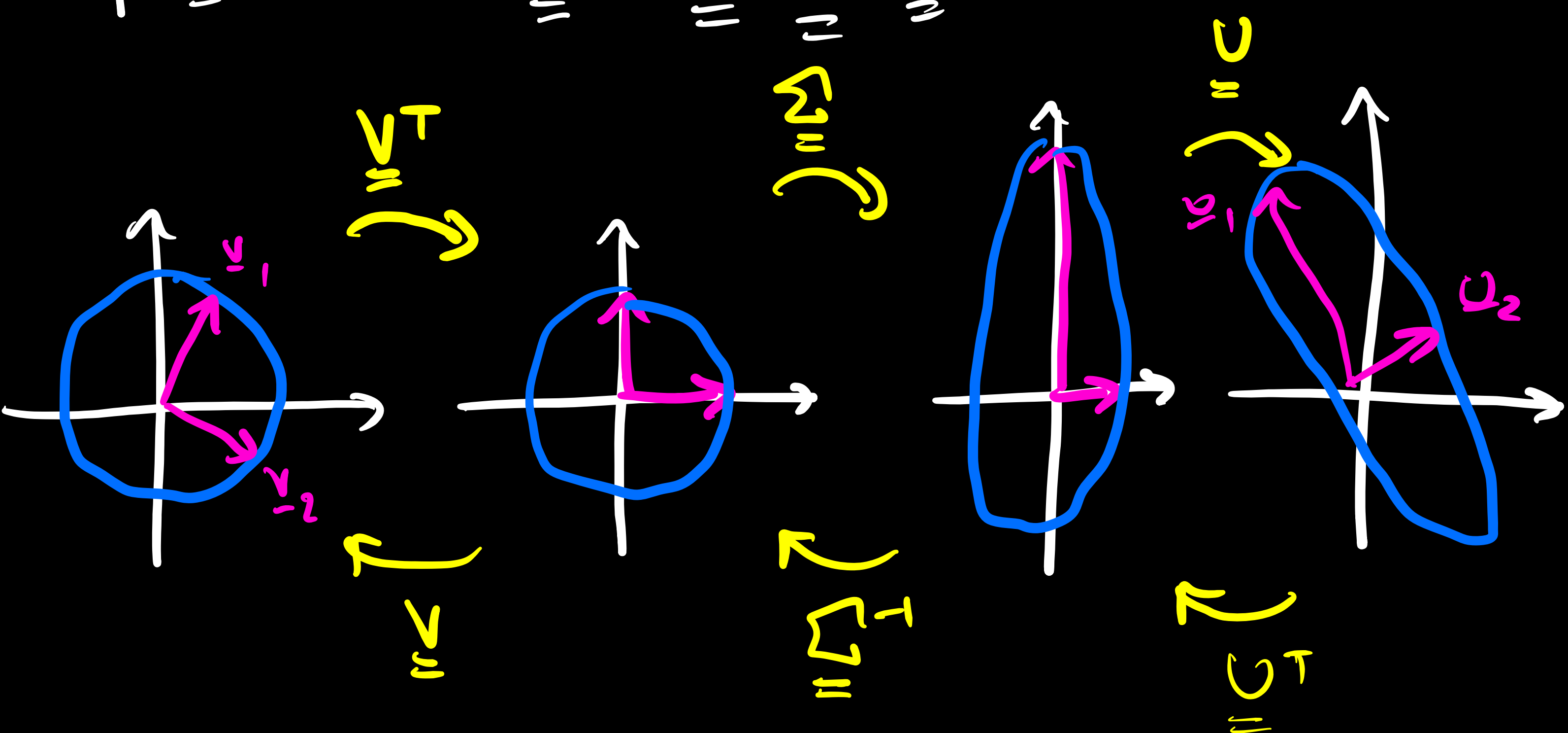
BACK TO SVD.

$$A = U \Sigma V^*$$

eigenvectors of AA^*

eigenvectors of A^*A

If A is real $\quad A = U \Sigma V^T$

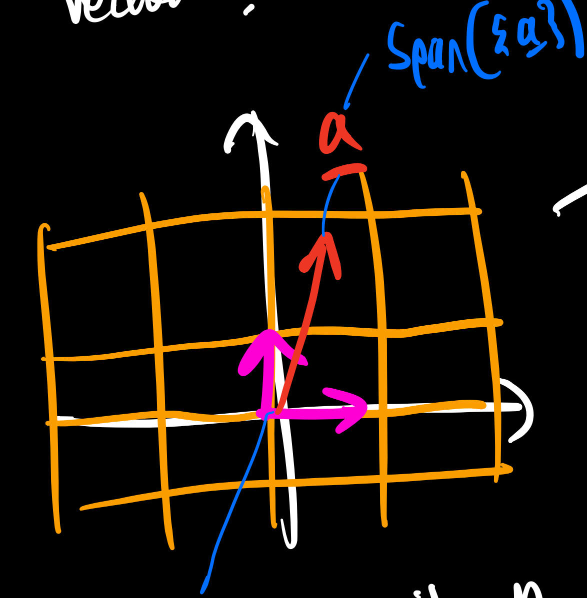


ALL MATRICES ARE JUST DOING THIS!

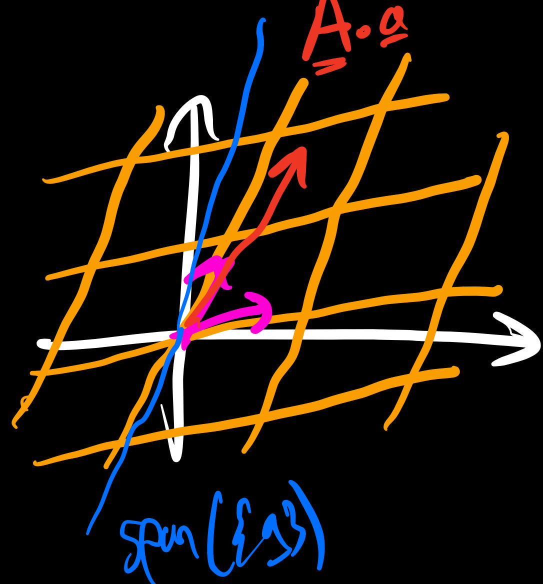
↳ EVEN RECTANGULAR ONES!

Will's Waffle

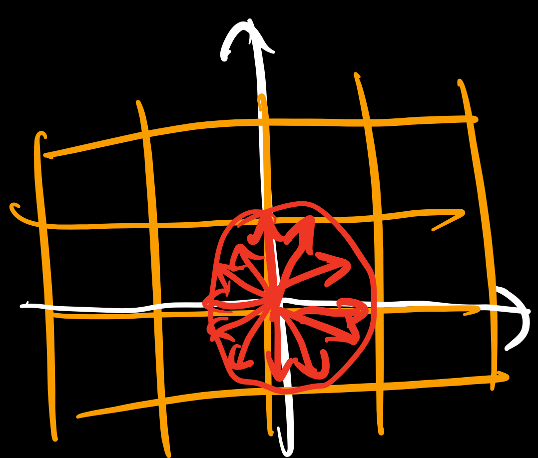
Eigenvector: Take a square matrix $\underline{\underline{A}}$, most vectors are rotated somehow, they lie outside the span of the original vector:



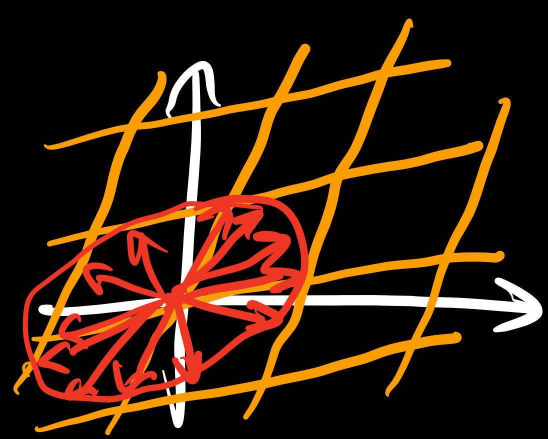
$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$



Some vectors aren't though



$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$



Points being squished towards & away from the axis!

eigenvalue

$$\underline{\underline{A}}\,\underline{x} = \lambda \underline{x}$$

eigenvector

Exercise: why do only square matrices have eigenvectors?

Prerequisites

SPECTRAL THEOREM: $\underline{\underline{H}} \in \mathbb{R}^{N \times N}$ is symmetric.

Then $\exists$ n, not necessarily distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding unit vectors, $\underline{v}_1, \dots, \underline{v}_n$ such that:

$$\underline{\underline{H}}\,\underline{v}_i = \lambda_i \underline{v}_i$$

$$U(\underline{v}_i) = \lambda_i \underline{v}_i$$

And $\underline{v}_i$ are orthonormal.

Def^n: Positive-definite matrix $\underline{x}^T \underline{\underline{A}}\,\underline{x} > 0 \quad \forall \underline{x} \in \mathbb{R}^n$

If $\underline{x}^T \underline{\underline{A}}\,\underline{x} \geq 0$ the semi-definite

RECALL

Theorem: $\underline{x}^T \underline{\underline{A}}\,\underline{x} \geq 0$ for symmetric matrix means $\lambda_j \geq 0$

$$\underline{x} = \sum_i c_i \underline{v}_i$$

$$\underline{x}^T \underline{\underline{A}}\,\underline{x} = \left(\sum_i c_i \underline{v}_i\right)^T \left(\sum_j c_j \underline{\underline{A}}\,\underline{v}_j\right)$$

$$= \left(\sum_i c_i \underline{v}_i\right)^T \left(\sum_j c_j \lambda_j \underline{v}_j\right)$$

$$= \sum_i c_i^2 \lambda_j \geq 0 \quad \therefore \lambda_j \geq 0$$

PROOF: $\underline{\underline{A}} \in \mathbb{R}^{n \times m}$

① $\underline{\underline{A}}^T \underline{\underline{A}}$ is symmetric & +ve semidefinite

- $(\underline{\underline{A}}^T \underline{\underline{A}})^T = \underline{\underline{A}}^T \underline{\underline{A}}$
- $\underline{x}^T \underline{\underline{A}}^T \underline{\underline{A}} \underline{x} = \| \underline{\underline{A}} \underline{x} \| \geq 0$

$\therefore$ Find eigenspace of $\underline{\underline{A}}^T \underline{\underline{A}} \Rightarrow \{ \lambda_i, \underline{v}_i \}_{i=1}^{n}$, ONB.

Choose $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$

② Define $\sigma_j = \| \underline{\underline{A}} \underline{v}_j \|_2$

$$\sigma_j^2 = \| \underline{\underline{A}} \underline{v}_j \|^2 = \underline{v}_j^T \underline{\underline{A}}^T \underline{\underline{A}} \underline{v}_j = \lambda_j \underline{v}_j^T \underline{v}_j = \lambda_j$$

Since $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$

③ Lets find our output vectors

If $\sigma_j \neq 0$ $\quad \underline{u}_j = \frac{1}{\sigma_j} \underline{\underline{A}} \underline{v}_j \quad (\| \underline{u}_j \| = 1)$

Then extend to make an ONB, any vector in orth to first ones

If $\sigma_j \neq 0, \sigma_k \neq 0$

$$\langle \underline{u}_j, \underline{u}_k \rangle = \frac{1}{\sigma_j \sigma_k} \langle \underline{\underline{A}} \underline{v}_j, \underline{\underline{A}} \underline{v}_k \rangle$$

$$= \frac{1}{\sigma_j \sigma_k} \langle \underline{v}_j, \underline{\underline{A}}^T \underline{\underline{A}} \underline{v}_k \rangle = \frac{\lambda_k}{\sigma_j \sigma_k} \langle \underline{v}_j, \underline{v}_k \rangle$$

$$= \frac{\lambda_j}{\sigma_j^2} = 1 \, !$$

$$\underline{\underline{A}} \, \underline{v}_j = \sigma_j \underline{u}_j$$

$$\underline{\underline{A}} \begin{pmatrix} | & & | \\ \underline{v}_1 & \cdots & \underline{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \sigma_1 \underline{u}_1 & \cdots & \sigma_n \underline{u}_n \\ | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ \underline{u}_1 & \cdots & \underline{u}_n \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}$$

$$\underline{\underline{A}} \, \underline{\underline{V}} = \hat{\underline{\underline{U}}} \hat{\underline{\underline{\Sigma}}}$$

Since $\underline{\underline{V}}$ is an ONB $\underline{\underline{V}}^{-1} = \underline{\underline{V}}^T$ $\therefore$ right multiply by $\underline{\underline{V}}^T$

$$\Rightarrow \quad \underline{\underline{A}} = \hat{\underline{\underline{U}}} \hat{\underline{\underline{\Sigma}}} \underline{\underline{V}}^T$$

REDUCED SINGULAR VALUE DECOMPOSITION

Uses: Regression

Linear system $\underline{\underline{A}} \cdot \underline{x} = \underline{b} = USV^T x$

$\therefore \quad SV^T x = U^T b$

In general can't invert and find x for all $\underline{b}$
(recall 3D case)

But can we get close? Perhaps the x that gets closest to $\underline{b}$?

$$\underline{x}^* = \arg\min_x \| \underline{\underline{A}}\underline{x} - \underline{b} \|_2^2 \quad ?$$

Claim if $A = USV^T$, $\quad S = \begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \sigma_3 \end{pmatrix}$

make $A^+ = VS^+U^T$, $\quad S^+ = \begin{pmatrix} 1/\sigma_1 & & 0 \\ & 1/\sigma_2 & \\ 0 & & 1/\sigma_3 \\ & 0 & \end{pmatrix}$

then $x^* = A^+ b$!!

Proof: If $\|v\|^2 = \underline{v}^T\underline{v}$

$$Ax - b = USV^Tx - b = U(SV^Tx - U^Tb)$$

$$\|Ax-b\|^2 = (Ax-b)^T(Ax-b)$$

$$= (SV^Tx - U^Tb)^T U^TU (SV^Tx - U^Tb)$$

$$= \|SV^Tx - U^Tb\|$$

minimise over all $\underline{x}$, let $\underline{y} = \underline{V}^T\underline{x}$, so y also goes everywhere.

minimise $\|\Sigma y - U^Tb\|$

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \\ & 0 & \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \left.\begin{pmatrix} \sigma_1 y_1 \\ \sigma_2 y_2 \\ \sigma_3 y_3 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right\} r = \text{rank}$$

Error is minimised when first r entries agree.

This is true if $\sigma_1 y_1 = (U^Tb)_1$

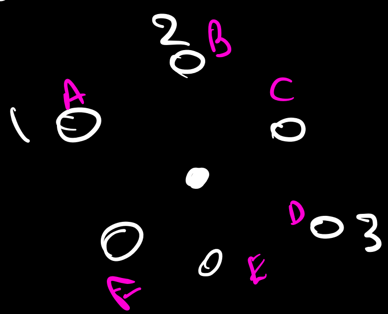
$$y_1 = \frac{1}{\sigma_1}(U^Tb)_1$$

i.e. $\underline{y} = \underline{\Sigma}^+\underline{U}^T\underline{b}$ $\qquad \underline{x} = \underline{V}\,\underline{\Sigma}^+\underline{U}^T\underline{b}$

Principal Angles - SVD in Neuroscience

neurons is large, might be embedded in 1 coding subspace

e.g. Xie et al. 2021



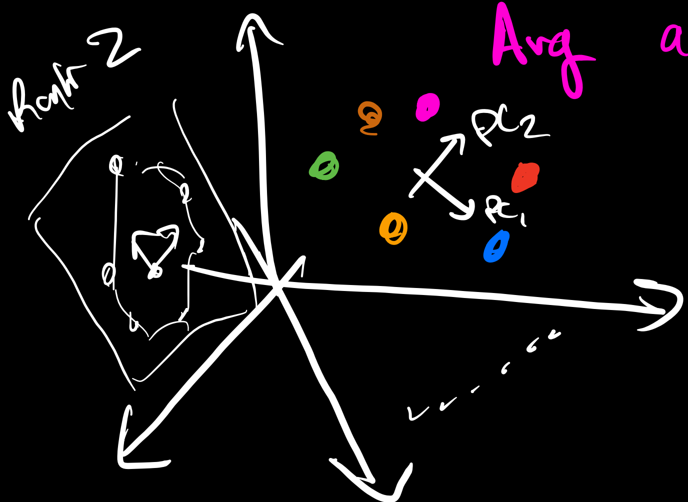
DELAY

Look at neural activity

PERFORM

Find encoding subspaces of position 1, 2, 3

Avg activity if 1 = A



Question how orthogonal are α_1, α_2 from β_1, β_2?

Principal angles

→ Set up bases and talk about their complement

→ Then say you have PCs and do SVD on

$\underset{=}{U^T}\;\underset{=}{V}$.