LECTURE 20

Q. What do you think the singular values σ_i of

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 3 & 0 \end{bmatrix}$$

are?

<u>Recall</u> $A \in \mathbb{R}^{m \times n}$

$m \geq n$:

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \Sigma \\ \hline 0 \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

↳ thin SVD:

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} \hat{U} \end{bmatrix} \begin{bmatrix} \Sigma \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

$n \geq m$:

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \Sigma \,|\, 0 \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

↳ thin SVD:

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \Sigma \end{bmatrix} \begin{bmatrix} \hat{V}^T \end{bmatrix}$$

$$\mathbb{R}^n \xrightarrow{V^T} \mathbb{R}^n \xrightarrow{\begin{bmatrix} \Sigma \\ 0 \end{bmatrix}} \mathbb{R}^m \xrightarrow{U} \mathbb{R}^m$$

pretransform

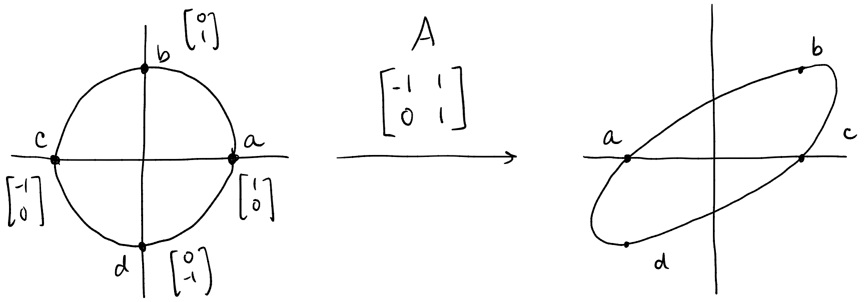
stretch in the $e_1, e_2, \ldots, e_n$ directions, where σ_i is the stretch/shrink factor for e_i

posttransform

* arranged so that e_1 is stretched the most, and e_n the least.

What is SVD doing / why do we need both U and V / why is this decomposition unique if we require $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

eg. to keep in your head



a, b, c, d are marked points on the circle for visualization purposes!

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

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

$A = U\Sigma V^T$ MATLAB: $[U, S, V] = \mathrm{svd}(A) \Rightarrow A = U{*}S{*}V'$

$$U = \begin{array}{cc} \begin{bmatrix} 0.8507 & -0.5257 \\ 0.5257 & 0.8507 \end{bmatrix} \\ \quad u_1 \qquad\qquad u_2 \end{array} \qquad V^T = \begin{array}{cc} \begin{bmatrix} -0.5257 & 0.8507 \\ 0.8507 & 0.5257 \end{bmatrix} & \begin{array}{c} v_1^T \\ v_2^T \end{array} \end{array}$$

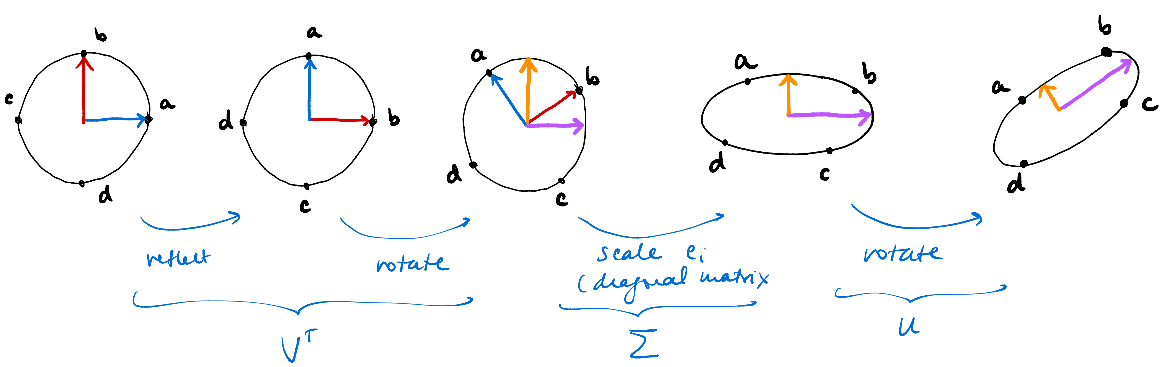
$\theta = \arccos(0.8507\ldots)$
$\approx 31.71°$

MATLAB:
$\theta = \mathrm{acos}(U(1,1))$

$$\Sigma = \begin{bmatrix} \varphi & 0 \\ 0 & \varphi - 1 \end{bmatrix} \qquad \varphi = \sqrt{\frac{3+\sqrt{5}}{2}} = \text{golden ratio}$$

$$U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \text{rotate by } \theta \text{ CCW}$$

$$V^T = \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \text{reflect by } e_1 \leftrightarrow e_2 \text{, then rotate by } \theta \text{ CCW}$$



Motto: If I understand SVD of A, I know everything about it.

Let's see how info about A can be easily extracted from its SVD.

SVD & Matrix norms

Recall $\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sigma_1$!

(Note that since U, V are orthogonal, $\|A\|_2 = \|\Sigma\|_2$.)

and $\|A\|_2 = \sqrt{\max \{ \lambda_i (A^T A) \}}$

Indeed, this is how we compute the singular values σ_i:

(2nd) defn. $\{\sigma_i\} = \{ \sqrt{\lambda_i (A^T A)} \}$ (we could have taken this as the definition)

$$= \{ \sqrt{\lambda_i (A A^T)} \}$$

because:

$$A^T A = (USV^T)^T (USV^T) = V S^T U^T U S V^T = V S^T S V^T$$

$$A A^T = (USV^T)(USV^T)^T = U S V^T V S^T U^T = U S S^T U^T$$

But $\Sigma = \Sigma^T \Rightarrow$

$$\{SS^T, S^T S\} = \left\{ \begin{bmatrix} \Sigma \\ \hline \end{bmatrix} \begin{bmatrix} \Sigma & \vline & \end{bmatrix}, \ \begin{bmatrix} \Sigma & \vline & 0 \end{bmatrix} \begin{bmatrix} \Sigma \\ \hline 0 \end{bmatrix} \right\} = \Bigg\{ \underbrace{\begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix}, \ \begin{bmatrix} \Sigma^2 \end{bmatrix}}_{\text{These have the same eigenvalues.}} \Bigg\}$$

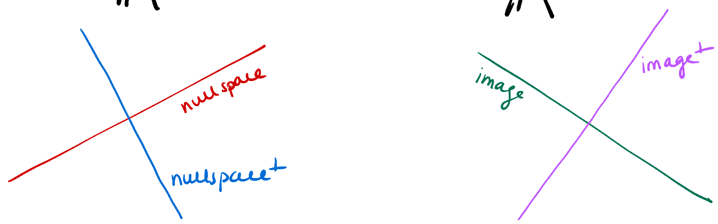
AND eigenvalues don't change under orthogonal transformations!

SVD gives bases for the 4 fundamental subspaces assoc to a matrix:

let $r = rk(A)$

A is allowed to be rank deficient!

$$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$$



$\ker(A) = null(A) = \{x \mid Ax = 0\}$

$= \langle v_{r+1}, v_{r+2}, \ldots, v_n \rangle$

orthogonal complement of $null(A) = (null(A))^\perp$

$= range(A^T)$!

$= \langle v_1, v_2, \ldots, v_r \rangle$

$range(A) = image(A) = \{y \mid y = Ax$ for some x in domain$\}$

$= \langle u_1, u_2, \ldots, u_r \rangle$

orthogonal complement of $range(A)$

$= (range(A))^\perp = null(A^T)$!

$= \langle u_{r+1}, \ldots, u_m \rangle$

<u>Fact to ponder</u> Look at the example we did. U, V are basically giving us the eigenvectors associated to $\sigma_i = \sqrt{\lambda_i(A^TA)} = \sqrt{\lambda_i(AA^T)}$ for A^TA and AA^T!

You'll look at this using MATLAB on HW07.