LECTURE 31
04/09/12

RECALL THAT FOR A COMPLEX INNER PRODUCT SPACE V,

N NORMAL $\iff$ $N = UDU^{-1}$
$(N^*N = NN^*)$

DIAGONAL MATRIX OF E-VALUES.

ORTHONORMAL E-VECTORS OF N

SPECIAL CASE:

L SELF-ADJOINT $\iff$ $L = UDU^{-1}$
$(L^* = L)$

REAL DIAGONAL MATRIX OF E-VALUES

IF V REAL INNER PRODUCT SPACE,

S SYMMETRIC $\iff$ $S = ODO^{-1}$
$(S^T = S)$

REAL MATRICES.

Q: WHAT ARE PROPERTIES OF U?

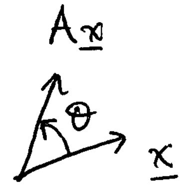
ISOMETRIES (7.4):

DEF. U IS AN ISOMETRY IF IT PRESERVES LENGTH — I.E.,

$$\|U\underline{x}\| = \|\underline{x}\| \quad \text{FOR ALL } x \in V.$$

NOTATION: IF V COMPLEX, U IS CALLED UNITARY.
" " REAL, " " " ORTHOGONAL.

Ex. $V = \mathbb{R}^2$, $\quad A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$



$\underline{x} = (x_1, x_2)^T$

$$\|A\underline{x}\|^2 = \|(cx_1 - sx_2, sx_1 + cx_2)^T\|^2$$
$$= c^2x_1^2 + s^2x_2^2 + s^2x_1^2 + c^2x_2^2$$
$$= x_1^2 + x_2^2$$
$$= \|\underline{x}\|^2, \quad \text{where} \quad c = \cos\theta,\ s = \sin\theta$$

$\Rightarrow$ A ISOMETRY (I.E., A ORTHOGONAL MATRIX).

REMARK: IN FACT, FOR ANY REAL SPACE V, THERE IS AN ORTHONORMAL BASIS $\mathcal{E}$ OF V SUCH THAT FOR AN ORTHOGONAL MATRIX A,

$$[A]_{\mathcal{E}} = \begin{pmatrix} \boxed{B_1} & & & 0 \\ & \boxed{B_2} & & \\ & & \ddots & \\ 0 & & & \boxed{B_k} \end{pmatrix}, \quad \text{WHERE EACH BLOCK } \boxed{B_i}$$

IS EITHER $\underbrace{\boxed{1}}_{\text{UNCHANGED}}$, $\underbrace{\boxed{-1}}_{\text{REFLECTION}}$, OR $\underbrace{\boxed{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}}_{\text{ROTATION IN PLANE}}$ FOR SOME $\theta \in (0, \pi)$.

THM. THESE ARE ALL EQUIVALENT (I.E., (i)–(iv) AND (i′)–(iv′) ALL SAME):

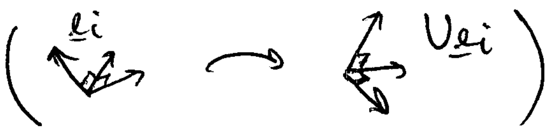
(i) U ISOMETRY	(i′) U^* ISOMETRY
(ii) $\langle U\underline{x} \mid U\underline{y} \rangle = \langle \underline{x} \mid \underline{y} \rangle$ FOR ALL $\underline{x}, \underline{y} \in V$	(ii′) $\langle U^*\underline{x} \mid U^*\underline{y} \rangle = \langle \underline{x} \mid \underline{y} \rangle$ FOR ALL $\underline{x}, \underline{y} \in V$

(iii) $U^* U = I$

(iv) IF $\mathcal{E} = \{\underline{e}_i\}$ ORTHONORMAL

$$[U]_{\mathcal{E}} = (U\underline{e}_1, \ldots, U\underline{e}_n)$$

HAS ORTHONORMAL COLUMNS



(iii') $U U^* = I$

(iv') IF $\mathcal{E} = \{\underline{e}_i\}$ ORTHONORMAL

$$[U^*]_{\mathcal{E}} = (U^*\underline{e}_1, \ldots, U^*\underline{e}_n)$$

HAS ORTHONORMAL COLUMNS.

IN ADDITION:

(a) U IS INVERTIBLE AND $U^{-1} = U^*$.

(PF. $\|U\underline{x}\| = \|\underline{x}\| \Rightarrow U\underline{x} = 0$ IFF $\underline{x} = 0$

$\Rightarrow \mathrm{Ker}(U) = \{0\} \Rightarrow U$ INVERTIBLE.

$U^* U = I \Rightarrow U^{-1} = U^*$.)

NOTE: THIS IMPLIES THAT THE SPECTRAL THM. CAN BE WRITTEN AS

N NORMAL $\iff N = UDU^*$ ← DIAGONAL.

(L SELF-ADJOINT $\iff L = UDU^*$ ← REAL DIAGONAL

S SYMMETRIC $\iff S = ODO^T$ ← REAL.)

(b) ALL E-VALUES OF U HAVE MAGNITUDE $|\lambda| = 1$

(SO IF U ORTHOGONAL, $\lambda = \pm 1$).

(PF. $\|U\underline{x}\| = \|\underline{x}\|$
$\|\lambda\underline{x}\| = |\lambda|\|\underline{x}\|$ $\Rightarrow |\lambda| = 1$.)

(c) $|\det U| = 1$

(PF. $|\det U| = |\lambda_1 \cdots \lambda_n| = 1$.)

PF. OF THM. :

(i) $\Rightarrow$ (ii) : IF V REAL, $\langle \underline{x} | \underline{y} \rangle = \dfrac{\|\underline{x}+\underline{y}\|^2 - \|\underline{x}-\underline{y}\|^2}{4}$ PARALLELOGRAM IDENTITY

$$\Rightarrow \langle U\underline{x} | U\underline{y} \rangle = \frac{\|U(\underline{x}+\underline{y})\|^2 - \|U(\underline{x}-\underline{y})\|^2}{4}$$

$$= \frac{\|\underline{x}+\underline{y}\|^2 - \|\underline{x}-\underline{y}\|^2}{4}$$

$$= \langle \underline{x} | \underline{y} \rangle .$$

SIMILAR IF V COMPLEX..

(ii) $\Rightarrow$ (iii) : $\langle (U^*U - I)\underline{x} | \underline{y} \rangle = \langle U\underline{x} | U\underline{y} \rangle - \langle \underline{x} | \underline{y} \rangle$
$= 0$.

LET $\underline{y} = (U^*U - I)\underline{x} \Rightarrow \|(U^*U - I)\underline{x}\| = 0$ FOR ALL $\underline{x} \in V$

$\Rightarrow U^*U = I$.

(iii) $\Rightarrow$ (iv) : $\langle U\underline{e}_i | U\underline{e}_j \rangle = \langle U^*U\underline{e}_i | \underline{e}_j \rangle = \langle \underline{e}_i | \underline{e}_j \rangle$
$= 0$.

(iv) $\Rightarrow$ (i) : $\underline{x} = \sum_{i=1}^{n} x_i \underline{e}_i$

$\Rightarrow \|U\underline{x}\|^2 = \langle U\underline{x} | U\underline{x} \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle U\underline{e}_i | U\underline{e}_j \rangle$

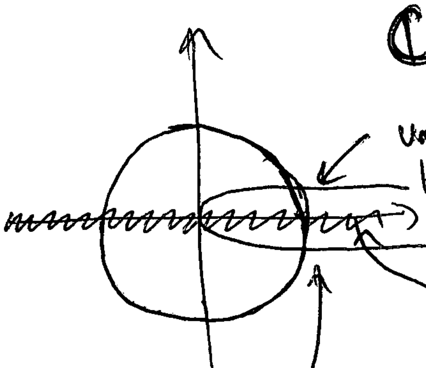
$$= x_1^2 + \cdots + x_n^2 = \|\underline{x}\|^2.$$

(i) $\Leftrightarrow$ (ii): EASY TO SHOW.

TO SUMMARIZE: THERE IS A NICE ANALOGY BETWEEN COMPLEX NUMBERS AND LINEAR OPERATORS!

COMPLEX NUMBERS

$\mathbb{C}$

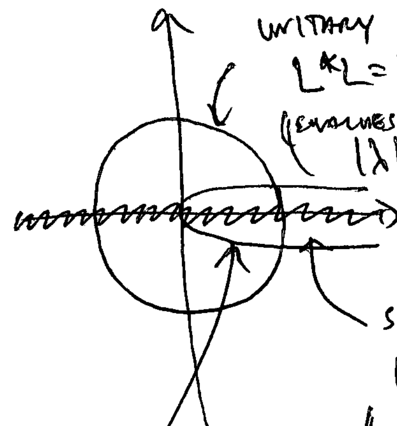


UNIT LENGTH $|z|^2 = \bar{z}z = 1$

REAL NUMBERS $\bar{z} = z$

POSITIVE REALS $\bar{z} = z$ AND $z \geq 0$.

LINEAR OPERATORS



UNITARY $L^*L = I$ (E-VALUES SATISFY $|\lambda|^2 = \bar{\lambda}\lambda = 1$)

SELF-ADJOINT $L^* = L$ (E-VALUES REAL SO $\bar{\lambda} = \lambda$)

POSITIVE OPERATORS $L^* = L$ AND ALL E-VALUES POSITIVE ($\lambda \geq 0$).

POLAR REPRESENTATION: FOR ANY z,

$$z = \frac{z}{|z|}|z| = \underbrace{\left(\frac{z}{|z|}\right)}_{\text{UNIT LENGTH}} \underbrace{\sqrt{\bar{z}z}}_{\text{POSITIVE REAL}}$$

POLAR DECOMPOSITION: FOR ANY L,

$$L = \underset{\text{UNITARY}}{U} \underbrace{\sqrt{L^*L}}_{\text{POSITIVE OPERATOR.}}$$

LECTURE 32
04/11/12

POSITIVE OPERATORS:

DEF. L POSITIVE IF IT IS SELF-ADJOINT AND ALL ITS EIGENVALUES ARE ≥ 0. WE DENOTE THIS BY $L \geq 0$.

(NOTE: WE ALLOW EIGENVALUES TO BE ZERO.)

REMARK: L POSITIVE $\iff$ $L^* = L$ AND $\langle L\underline{x} | \underline{x} \rangle \geq 0$ FOR ALL $\underline{x} \in V$.

THM. IF $L \geq 0$, THERE IS A UNIQUE OPERATOR $B \geq 0$ SUCH THAT $B^2 = L$. IN PARTICULAR,

$$B = U \underbrace{D^{1/2}}_{= \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}} U^*, \quad \text{WHERE} \quad L = U \underbrace{D}_{= \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}} U^*$$

PF. $$B^2 = (UD^{1/2} \underbrace{U^*)(U}_{=I} D^{1/2} U^*) = UD^{1/2}D^{1/2}U^* = UDU^* = L.$$

UNIQUENESS EASY.

EX. SUPPOSE $A \in M_{2,2}(\mathbb{C})$ HAS E-VALUES $\lambda_1 = 9$, $\lambda_2 = 4$ W/ CORRESPONDING E-VECTORS $\underline{\xi}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$, $\underline{\xi}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$.

Q1: WHAT IS A?

A1: SINCE $\langle \underline{\xi}_1 | \underline{\xi}_2 \rangle = (-i, 1)(-i, 1)^T = 0$,

A HAS ORTHOGONAL E-VECTORS AND REAL E-VALUES

$\Rightarrow$ A SELF-ADJOINT.

$$\Rightarrow A = UDU^* = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 13 & 5i \\ -5i & 13 \end{pmatrix}.$$

Q2: WHAT IS $\sqrt{A}$?

A2: SINCE $A \geq 0$, $\sqrt{A}$ EXISTS.

$$\sqrt{A} = U \underbrace{D^{1/2}}_{\begin{pmatrix} \sqrt{9} & 0 \\ 0 & \sqrt{4} \end{pmatrix}} U^* = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix} = \begin{pmatrix} 5 & i \\ -i & 5 \end{pmatrix}.$$

NOTE: $\sqrt{A}$ IS DETERMINED BY E-VALUES OF A, NOT BY INDIVIDUAL ENTRIES a_{ij}!

WE WILL USE POSITIVE OPERATORS TO INTRODUCE ONE OF THE MOST USEFUL MATRIX DECOMPOSITIONS, CALLED SINGULAR VALUE DECOMPOSITION (SVD).

MOTIVATION: SO FAR WE HAVE INTRODUCED SEVERAL DECOMPOSITIONS.

A DIAGONALIZABLE

$$A = PDP^{-1}$$

E-VECTORS

DIAGONAL OF E-VALUES

$\rightsquigarrow$ GENERAL $A \in M_{n,n}(\mathbb{C})$

$$A = \tilde{P}\tilde{D}\tilde{P}^{-1}$$

E-VECTORS AND GENERALIZED E-VECTORS.

BLOCK DIAGONAL

$\boxed{\text{JORDAN DECOMPOSITION}}$

A NORMAL

$$A = UDU^*$$

ORTHONORMAL E-VECTORS

DIAGONAL OF E-VALUES

$\rightsquigarrow$ GENERAL $A \in M_{n,n}(\mathbb{C})$

$$A = UTU^*$$

ORTHONORMAL BASIS

UPPER TRIANGULAR W/ E-VALUES ALONG DIAGONAL

$\boxed{\text{SCHUR DECOMPOSITION}}$

$\rightsquigarrow$ GENERAL $A \in M_{n,n}(\mathbb{C})$

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

UNITARY

DIAGONAL MATRIX OF SINGULAR VALUES.

UNITARY

$\boxed{\text{SINGULAR VALUE DECOMPOSITION}}$

THAT IS,

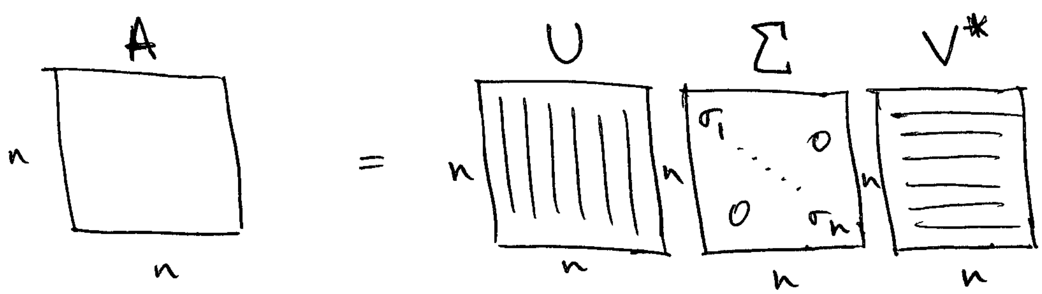
<u>Q</u>: CAN WE FIND TWO ORTHONORMAL BASES $\mathcal{E}$ AND $\mathcal{F}$ SUCH THAT $[A]_{\mathcal{E}\mathcal{F}}$ IS DIAGONAL?

<u>A</u>: YES, FOR <u>ANY</u> A. THEN U WILL HAVE COLUMNS FROM $\mathcal{E}$, V COLUMNS FROM $\mathcal{F}$, AND $\Sigma = [A]_{\mathcal{E}\mathcal{F}}$.

Lecture 33
04/13/12

Singular value decomposition (SVD):

For **any** $A \in M_{n,n}(\mathbb{C})$,



$$\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n), \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

singular values

(e-values of $\sqrt{A^*A}$ or $\sqrt{AA^*}$)

V unitary (orthonormal e-vectors of $\sqrt{A^*A}$)

U unitary (orthonormal e-vectors of $\sqrt{AA^*}$)

Note: A^*A and AA^* are positive operators since they are self-adjoint and have all eigenvalues positive (≥ 0). To see this, note that $(A^*A)^* = A^* A^{**} = A^*A$. And if λ is an e-value w/ e-vector $\underline{v}$ for A^*A, then

$$\lambda \|\underline{v}\|^2 = \lambda \langle \underline{v} | \underline{v} \rangle = \langle \underline{v} | \lambda \underline{v} \rangle = \langle \underline{v} | A^*A\underline{v} \rangle$$

$$= \langle A\underline{v} | A\underline{v} \rangle \quad \text{(by properties of adjoint.)}$$

$$= \|A\underline{v}\|^2 \Rightarrow \lambda \geq 0.$$

Q: WHY SVD?

A: WE WILL SEE LATER THAT SVD IS USEFUL IN APPLICATIONS INVOLVING LARGE, HIGH-DIMENSIONAL DATA SETS. IT IS NOT AS USEFUL FOR DYNAMICAL APPLICATIONS SINCE

$$A^k = (U\Sigma V^*)^k = U\Sigma V^* U\Sigma V^* \cdots U\Sigma V^*$$
$$\neq U\Sigma^k V^*.$$

(UNLESS $U=V$, IN WHICH CASE THIS IS THE SAME AS JORDAN FORM.)

REMARK:

1) A^*A AND AA^* HAVE SAME E-VALUES!

PF. $A^*A\,\underline{v} = \lambda\underline{v} \Rightarrow AA^*(\underbrace{A\underline{v}}_{\text{CALL } \underline{u}}) = \lambda(\underbrace{A\underline{v}}_{\underline{u}})$

$AA^*\underline{u} = \lambda\underline{u} \Rightarrow A^*A(\underbrace{A^*\underline{u}}_{\text{CALL } \underline{v}}) = \lambda(\underbrace{A^*\underline{u}}_{\underline{v}}).$

2) $r \triangleq \text{RANK}(A) = \#$ OF NONZERO σ_i's.

PF. $\text{RANK}(A) = \text{RANK}([A]_{\mathcal{E}\mathcal{F}}) = \text{RANK}(\Sigma)$
$= \#$ NONZERO σ_i's

WHERE $\mathcal{E}, \mathcal{F}$ ARE COLUMNS OF U, V.

PROCEDURE TO FIND SVD:

① FIND NONZERO E-VALUES $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$, $r \leq n$. OF A^*A AND CORRESPONDING E-VECTORS $\{v_i\}_{i=1}^{r}$.

② LET $\sigma_i = \sqrt{\lambda_i}$, $i = 1, .., r$. THESE ARE THE E-VALUES OF $\sqrt{A^*A}$ (OR OF $\sqrt{AA^*}$) — THAT IS, THE <u>SINGULAR VALUES</u> OF A.

③ LET $\underline{u}_i = \frac{1}{\sigma_i} A \underline{v}_i$, $i = 1, .., r$. THESE ARE THE ORTHONORMAL E-VECTORS OF AA^*.

(TO SEE THIS, NOTE THAT $\langle A\underline{v}_i | A\underline{v}_j \rangle = \langle A^*A\underline{v}_i | \underline{v}_j \rangle$
$= \begin{cases} \sigma_i^2 & \text{IF } i=j \\ 0 & \text{ELSE} \end{cases} \Rightarrow \{\underline{u}_i\}$ ORTHONORMAL.)

④ IF NECESSARY: FOR REMAINING COLUMNS

$\underline{v}_{r+1}, \ldots, \underline{v}_n$ (E-VECTORS OF A^*A W/ E-VALUE 0, I.E., IN $\mathrm{Ker}(A^*A)$)

$\underline{u}_{r+1}, \ldots, \underline{u}_n$ (E-VECTORS OF AA^* W/ E-VALUE 0, I.E., IN $\mathrm{Ker}(AA^*)$)

USE GRAM-SCHMIDT. THAT IS, FIND A BASIS FOR $\mathrm{Ker}(A^*A)$ AND ORTHONORMALIZE TO GET $\{v_i\}_{i=r+1}^{n}$, AND SIMILARLY FIND A BASIS FOR $\mathrm{Ker}(AA^*)$ AND ORTHONORMALIZE TO GET $\{\underline{u}_i\}_{i=r+1}^{n}$.

EX. FIND SVD OF $A = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$.

① $A^*A = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \Rightarrow \lambda_1 = 8,\ \lambda_2 = 2$

(TRACE $= 10$, DET $= 16$)

$\Rightarrow \sqrt{\lambda_1} = \sqrt{8},\ \sqrt{\lambda_2} = \sqrt{2}$

② $\Rightarrow \sigma_1 = \sqrt{8},\ \sigma_2 = \sqrt{2}$.

CORRESPONDING EVECTORS OF A^*A ARE $\underline{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \underline{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

ORTHONORMAL.

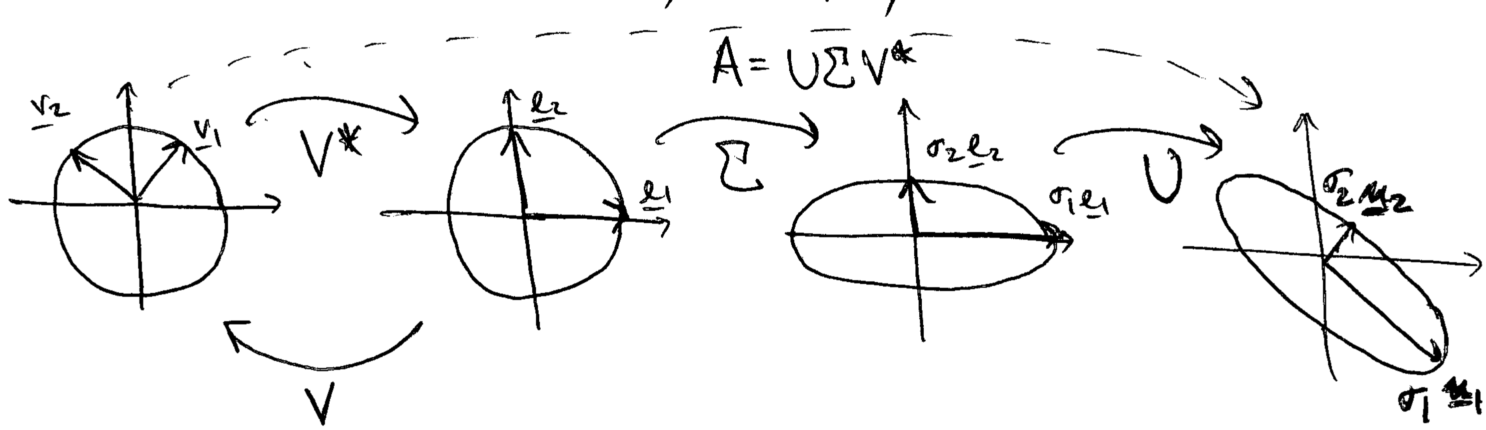
③ $\frac{1}{\sigma_1} A \underline{v}_1 \doteq \underline{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\frac{1}{\sigma_2} A \underline{v}_2 \doteq \underline{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

ORTHONORMAL.

$$\Rightarrow A = \underset{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}{\underset{\shortparallel}{U}}\ \underset{\begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix}}{\underset{\shortparallel}{\Sigma}}\ \underset{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}{\underset{\shortparallel}{V^*}}$$

GEOMETRY OF SVD: EVERY LINEAR TRANSFORMATION IS GIVEN BY A ROTATION, STRETCH, AND ANOTHER ROTATION. FOR EX, IN $\mathbb{R}^2$,



Q: WHY DOES A^*A PLAY A ROLE TO BEGIN WITH?

FROM THE GEOMETRIC PICTURE, WE SEE THAT ANY LINEAR TRANSFORMATION IS A MAPPING FROM THE UNIT SPHERE TO AN ELLIPSE (A HYPERELLIPSE IN HIGHER DIMENSIONS).

WE SEEK $\underline{x}$'s SUCH THAT $\|\underline{x}\|=1$ AND $A\underline{x}$ HAS MAXIMAL (OR MINIMAL) LENGTH.

LET $Q(\underline{x}) = \|A\underline{x}\|^2$. THEN, WE SEEK SOLUTIONS $\underline{x}$ THAT

$$\begin{cases} \text{EXTREMIZE } Q(\underline{x}) = \|A\underline{x}\|^2 = \langle A\underline{x} | A\underline{x}\rangle = \langle \underline{x} | A^*A\underline{x}\rangle \\ \text{w/ CONSTRAINT } \|\underline{x}\| = 1. \end{cases}$$

USING LAGRANGE MULTIPLIERS TO SOLVE THIS CONSTRAINED MAXIMIZATION/MINIMIZATION PROBLEM, WE GET THAT SOLUTIONS SATISFY

$$A^*A\,\underline{x} = \lambda\,\underline{x} \quad \text{FOR SOME } \lambda$$

LAGRANGE MULTIPLIER

$\Rightarrow$ I.E., LOOK AT E-VECTORS OF A^*A!