Dis 4B Notes

Tuesday, July 14, 2020 11:06 AM

SVD II

* Recap of SVD
* Geometric Interpretation
  → Worksheet
* A little bit about PCA?

I. Recap of SVD

a) $A = U \Sigma V^* = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^*$

A: $m \times n$, U: $m \times m$, Σ: $m \times n$, V^*: $n \times n$, $\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^*$: Compact SVD

$U^* U = I$

$V^* V = I$

U, V are square unitary (col's form orthonormal bases)

Σ is rectangular diagonal matrix

Σ is rectangular diagonal matrix

$$m > n \quad \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ \hline & 0 & \end{bmatrix} \quad \text{or} \quad \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_m & \end{array}\right]$$

$m < n$

$$\sigma_i = \Sigma_{ii} > 0 \quad \text{for } 1 \le i \le r$$
$$= 0 \quad \text{for } i > r$$

$(r = \text{rank } A)$

b) Constructing the SVD

(1) Picking the smaller of A^*A, AA^*

(2) Find eigenvalues and eigenvecs of the matrix from (1):

Order them $\lambda_1 \ge \lambda_2 \dots \ge \lambda_r > 0$

$\lambda_{r+1} = \lambda_{r+2} = \dots \; 0$

a) if A^*A: right singular vectors

$$(A^*A)\vec{v_i} = \lambda_i \vec{v_i}$$

b) if AA^*: left singular vectors

$$(AA^*)\vec{u_i} = \lambda_i \vec{u_i}$$

3 $\sigma_i = \sqrt{\lambda_i}$

4 Complete the U, V bases w/ Gram-Schmidt (or some other orthonormalization process)

c) Relation to Matrix Subspaces

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

$$\Longrightarrow \boxed{A\vec{v_i} = \sigma_i \vec{u_i}}$$

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

$$\Longrightarrow A^*\vec{u_i} = \sigma_i^* \vec{v_i}$$

$$\sigma_i \in \mathbb{R}$$

$$\Rightarrow \boxed{A^* \vec{u_i} = \sigma_i v_i}$$

$$\text{eigenvectors}$$

$$AA^*: \left.\underbrace{\vec{u_1} \cdots \vec{u_r}}_{\text{Span } Col(A)}, \underbrace{\vec{u_{r+1}} \cdots \vec{u_m}}_{\text{Span } Null(A^*)}\right] \text{span } \mathbb{C}^m$$

$$A^*A: \left.\underbrace{\vec{v_1} \cdots \vec{v_r}}_{\text{Span } Col(A^*)}, \underbrace{\vec{v_{r+1}} \cdots \vec{v_n}}_{\text{Span } Null(A)}\right] \text{Span } \mathbb{C}^n$$

II. Geometric Interpretation of the SVD

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

unitary

$$\langle U\vec{v}, U\vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle$$

$$\Rightarrow \text{length-preserving}$$

$U, V \longrightarrow$ rotations

$\Sigma \longrightarrow$ scales axes

How does this affect a circle, i.e, image of a circle under a linear transformation A

$\Longrightarrow$ Worksheet!

Dis 4B Worksheet

Tuesday, July 14, 2020 11:07 AM

1 Geometric interpretation of the SVD

In this exercise, we explore the geometric interpretation of symmetric matrices and how this connects to the SVD. We consider how a real 2×2 matrix acts on the unit circle, transforming it into an ellipse. It turns out that the principal semiaxes of the resulting ellipse are related to the singular values of the matrix, as well as the vectors in the SVD.

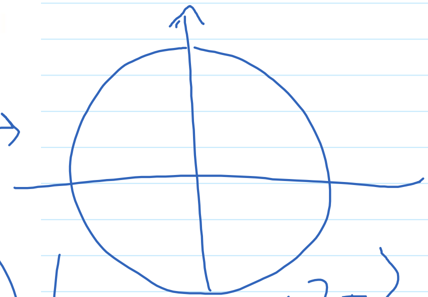
a) Consider the real 2×2 matrix

$$A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}.$$

Now consider the unit circle in $\mathbb{R}^2$,

$$S = \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \,\middle|\, 0 \le \theta < 2\pi \right\}.$$

Plot AS on the $\mathbb{R}^2$ plane.

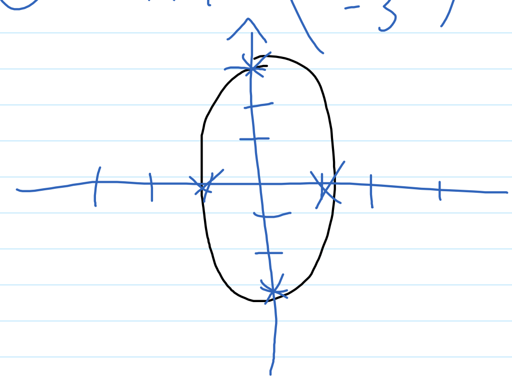


$$AS = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \middle| \, 0 \le \theta < 2\pi \right\}$$

$$= \left\{ \begin{pmatrix} -\sin\theta \\ 3\cos\theta \end{pmatrix} \middle| \, 0 \le \theta < 2\pi \right\}$$

$\theta = 0$: $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$ $\qquad$ $\theta = \frac{\pi}{2}$: $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$

$\theta = \pi$: $\begin{pmatrix} 0 \\ -3 \end{pmatrix}$ $\qquad$ $\theta = \frac{3\pi}{2}$: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$



b) Calculate the SVD of A. Write this as a matrix factorization, i.e. $A = U\Sigma V^*$.

$\lambda_1 = 9, \ \lambda_2 = 1$ $\qquad$ $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\sigma_1 = 3, \ \sigma_2 = 1$

$$A^*A = \begin{pmatrix} 0 & 3 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$$

diagonal: $\lambda_1 = 9, \ \lambda_2 = 1$

diagonal: $\lambda_1 = 9,\ \lambda_2 = 1$

$$\vec{v_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \vec{v_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A\vec{v_i} = \sigma_i \vec{u_i} \qquad \vec{u_1} = \frac{A\vec{v_1}}{\sigma_1} = \frac{\begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}}{3} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Downarrow$$

$$\frac{A\vec{v_i}}{\sigma_i} = \vec{u_i} \qquad \vec{u_2} = \frac{A\vec{v_2}}{\sigma_2} = \frac{\begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}}{1} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

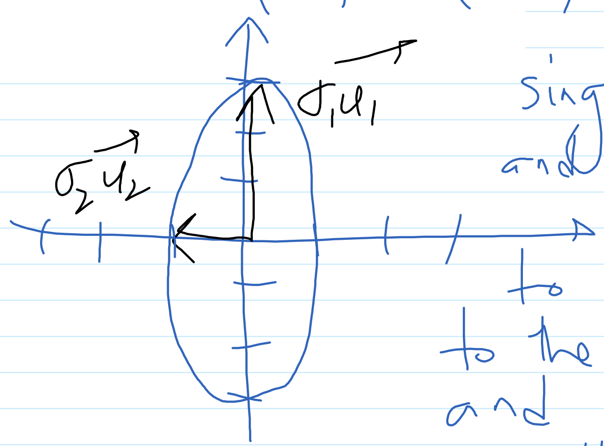
$$\boxed{A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^*}$$

c) Consider the columns of the matrices U, V obtained in the previous part, and treat them as vectors in $\mathbb{R}^2$. Let $U = (\vec{u}_1\ \vec{u}_2)$, $V = (\vec{v}_1\ \vec{v}_2)$. Let σ_1, σ_2 be the singular values of A, where $\sigma_1 \geq \sigma_2$.

Draw in your plot of AS the vectors $\sigma_1\vec{u}_1$ and $\sigma_2\vec{u}_2$, drawn from the origin. What do these vectors correspond to geometrically?

$$\sigma_1\vec{u_1} = 3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

$$\sigma_2\vec{u_2} = 1\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$



singular values and left-singular vecs seem to correspond to the semimajor and seminor axes of the ellipse!

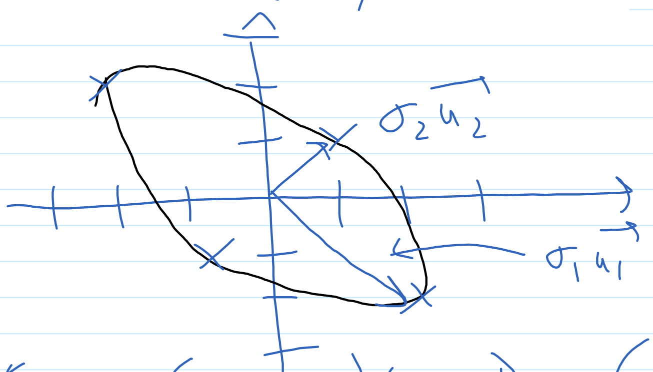
d) **Repeat what you did above for the matrix $A = \begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$.**

$$AS = \left\{ \begin{pmatrix} 2\cos\theta + \sin\theta \end{pmatrix} \middle| \ 0 \leq \theta < 2\pi \right\}$$

$$AS = \left\{ \begin{pmatrix} 2\cos\theta + \sin\theta \\ -2\cos\theta + \sin\theta \end{pmatrix} \;\middle|\; 0 \le \theta < 2\pi \right\}$$

$\theta = 0: \begin{pmatrix} 2 \\ -2 \end{pmatrix}$ $\qquad \theta = \frac{\pi}{2}: \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\theta = \pi: \begin{pmatrix} -2 \\ 2 \end{pmatrix}$ $\qquad \theta = \frac{3\pi}{2}: \begin{pmatrix} -1 \\ -1 \end{pmatrix}$



$$A^*A = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$$

$\lambda_1 = 8, \lambda_2 = 2 \longrightarrow \vec{v_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\sigma_1 = 2\sqrt{2}, \sigma_2 = \sqrt{2}$ $\qquad \vec{v_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\vec{u_1} = \frac{A\vec{v_1}}{\sigma_1} = \frac{\begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}}{2\sqrt{2}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\vec{u_2} = \frac{A\vec{v_2}}{\sigma_2} = \frac{\begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}}{\sqrt{2}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\boxed{A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^*}$$

Using AA^*: $\Rightarrow \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}$

$\lambda_1 = 8$, $\lambda_2 = 2$

$\vec{u_1} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}}$, $\vec{u_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

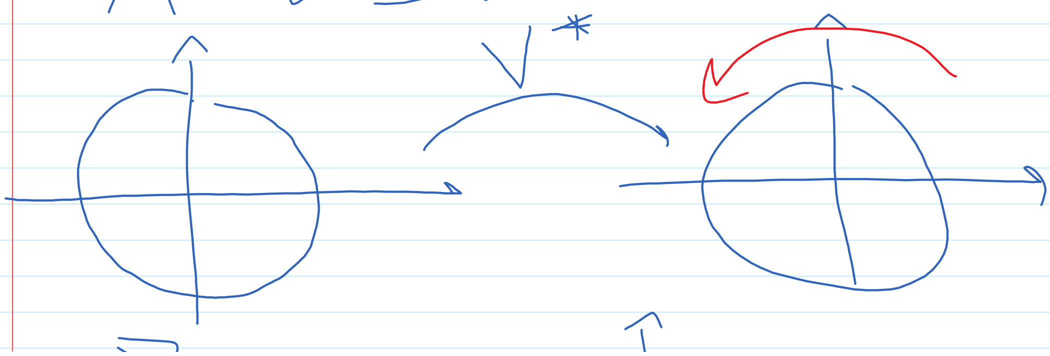
$$\sigma_1 \vec{u_1} = 2\sqrt{2} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

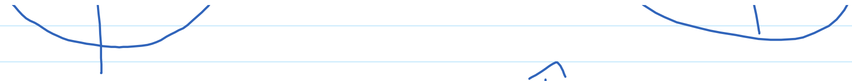
$$\sigma_2 \vec{u_2} = \sqrt{2} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Seems to match intuition that σ_i determine length of principle axes, $\vec{u_i}$ determine the direction of the principle axes

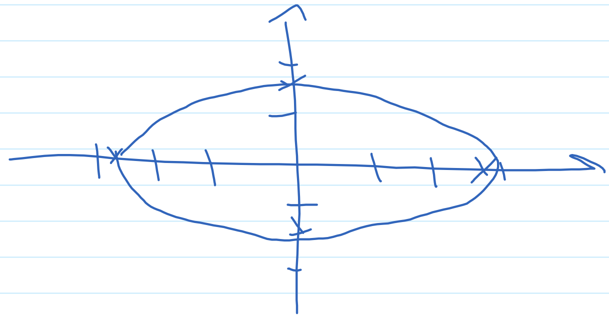
Why just $\sigma_i \vec{u_i}$? No $\vec{v_i}$?
(that determine final size and orientation of the ellipse)

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





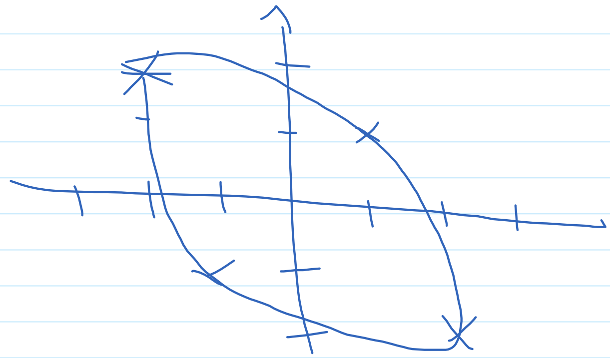
Σ



U

Look closely:

$R(\theta = -45^\circ)$



1) V^* : rotation
→ circle is invariant under rotation

2) Σ : scaling
circle → ellipse

3) U : rotation
ellipse is not invariant under arbitrary rotation

Must be $U\Sigma$ that carries the effect of determining the size and orientation of the ellipse

① $$\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} \cos\theta + 2\sin\theta \\ 2\cos\theta + \sin\theta \\ \cos\theta + \sin\theta \end{pmatrix}$$

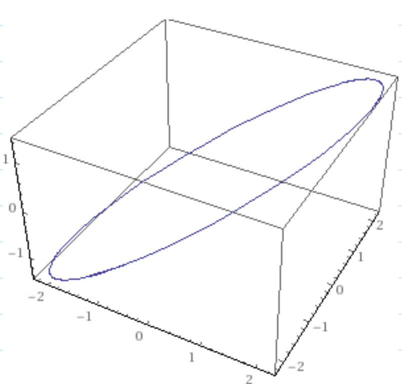
$$\begin{pmatrix} 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} \cos\theta + \sin\theta \end{pmatrix}$$

$$A = \begin{pmatrix} \frac{3}{\sqrt{22}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{11}} \\ \frac{3}{\sqrt{22}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{11}} \\ \sqrt{\frac{2}{11}} & 0 & \frac{3}{\sqrt{11}} \end{pmatrix} \begin{pmatrix} \sqrt{11} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\sigma_1 \vec{u_1} = \begin{pmatrix} \frac{3}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} \\ \sqrt{2} \end{pmatrix} \approx \begin{pmatrix} 2.1 \\ 2.1 \\ 1.4 \end{pmatrix}$$

$$\sigma_2 \vec{u_2} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0.7 \\ -0.7 \\ 0 \end{pmatrix}$$

https://www.wolframalpha.com/input/?i=plot+%28x%3Dcos%28theta%29%2B2sin%28theta%29%2C+y%3D2cos%28theta%29%2Bsin%28theta%29%2C+z%3Dcos%28theta%29%2Bsin%28theta%29%29



https://www.nature.com/articles/nphoton.2015.153.pdf

e) **Consider the case where A is a real $n \times n$ symmetric matrix. What do you observe geometrically in this case?**

SVD(A), where $A = A^T$

$\Rightarrow A^T A \overset{?}{=} A A^T$

$v_i \qquad u_i$

$\Rightarrow V = U$ (the same!)

$\Rightarrow \quad V = U \quad$ (the same!)

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

$\Rightarrow$ Circle gets stretched by σ_i in direction of $\vec{v}_i$ (eigenvecs of A^2)

2 SVD and Induced 2-Norm

a) Show that if U is a unitary matrix then for any $\vec{x}$

$$\|U\vec{x}\| = \|\vec{x}\|.$$

$\|U\vec{x}\| = \|\vec{x}\|$

$\longrightarrow \|U\vec{x}\| = \sqrt{\langle U\vec{x}, U\vec{x}\rangle}$

$= \sqrt{x^* \underbrace{U^* U}_{} \vec{x}}$

$= \sqrt{x^* x}$

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

$= \|\vec{x}\|$

b) Find the maximum

$$\max_{\{\vec{x}:\|\vec{x}\|=1\}} \|A\vec{x}\|$$

in terms of the singular values of A.

$= \max_{\vec{x}:\|\vec{x}\|=1} \|U\Sigma V^* \vec{x}\|$

$= \max_{\vec{x}:\|\vec{x}\|=1} \|\Sigma V^* \vec{x}\|$

force in somehow?

$\vec{x} = V\vec{y}$

$\|\Sigma V^* V \ldots\|$

$\vec{x} = V\vec{y}$

$$= \max_{\vec{y} : \|V\vec{y} = \vec{x}\| = 1} \|\Sigma \cancel{V^*} V \vec{y}\|$$

$$= \max_{\vec{y} : \|V\vec{y}\| = 1} \|\Sigma \vec{y}\|$$

$\Longrightarrow$ Grab max Singular value!

$$\Sigma = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ 0 & & \sigma_n \\ \hline & 0 & \end{pmatrix}$$

$\Longrightarrow$ Pick $\vec{y} = \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

$$\therefore \boxed{\|A\vec{x}\|_{max} = \sigma_1}$$

How do we know $\vec{x} = V\vec{y} = V\vec{e}_1$ has length 1?

$\Longrightarrow$ V is unitary (length preserving)

c) Find the $\vec{x}$ that maximizes the expression above.

$$\vec{x}_{max} = V\vec{e}_1 = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{pmatrix} \vec{e}_1$$

$$= \vec{v}_1$$

$= \vec{v_1}$

picks out direction of max amplification

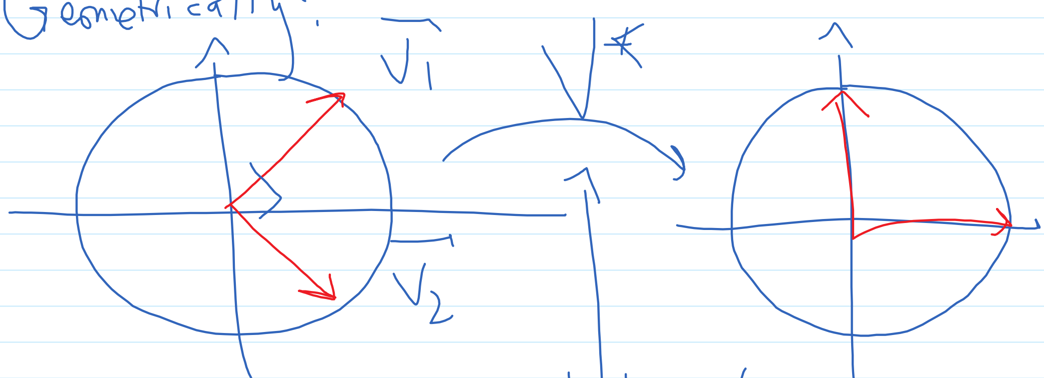
$$A\vec{v_1} = \sigma_1 \vec{u_1}$$

picks out direction to amplify

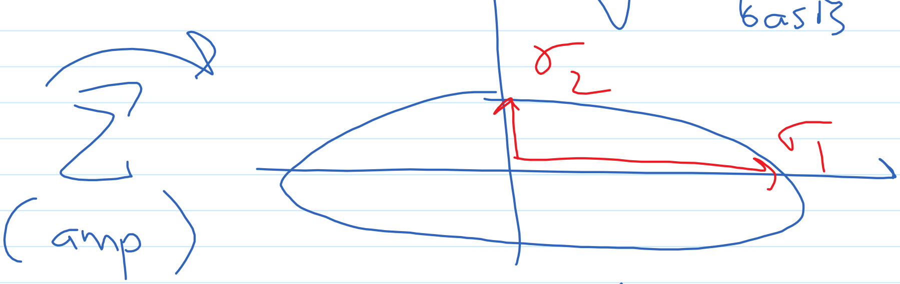
amp. factor

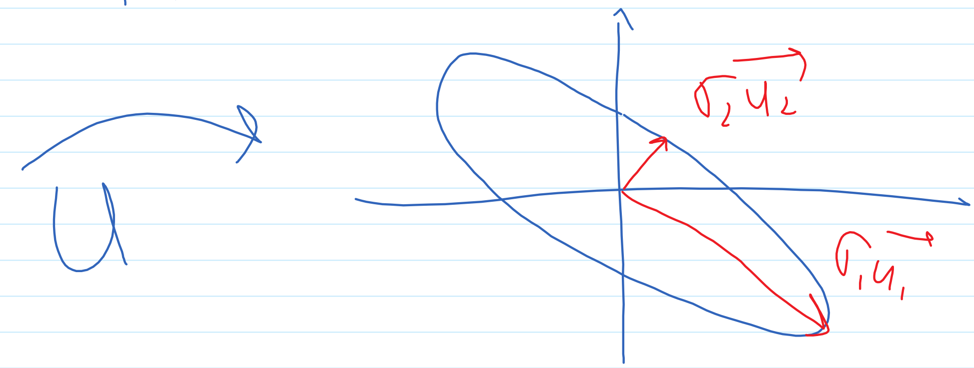
output direction that is amplified

Geometrically?



rotation / change of basis into basis / projection onto basis




U
σ₂u₂
σ₁u₁