Lesson 22

Chapter 8

Geometry of SVD

Geometric interpretation : $T \quad \mathbb{R}^2 \to \mathbb{R}^2 \quad T(w) = Aw$

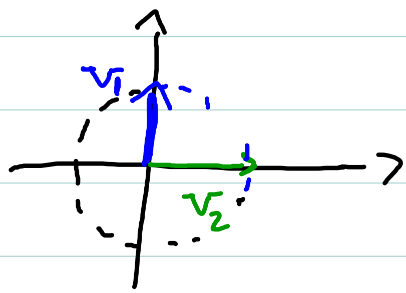
$$A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$$

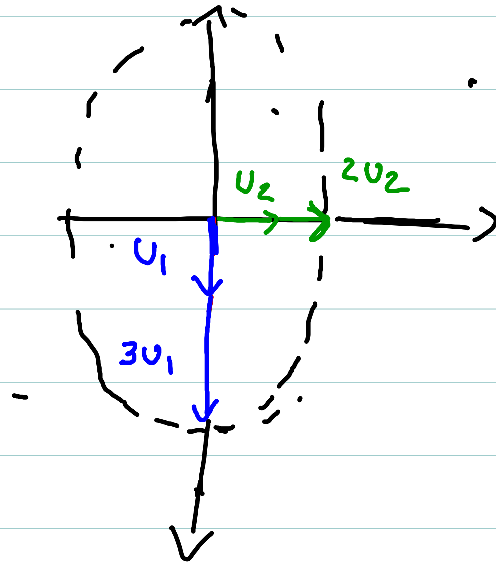
If $v = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$, $\alpha_1^2 + \alpha_2^2 = 1$

$$Av = \begin{bmatrix} 2\alpha_1 \\ -3\alpha_2 \end{bmatrix} \qquad \frac{(2\alpha_1)^2}{4} + \frac{(3\alpha_2)^2}{9} = 1$$

v on circle $x^2 + y^2 = 1$ $\quad Av$ on ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$

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





u_1 along major axis
u_2 along minor axis

$A v_1 = 3 u_1 \qquad A v_2 = 2 u_2$

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

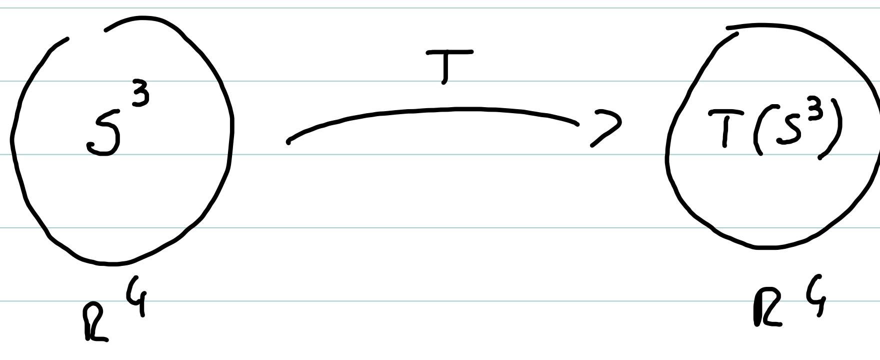
Consider $T \quad R^4 \to R^4 \quad T(v) = Av$

Unit splere $S^3 = \left\{ \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \in R^4 \mid x^2+y^2+z^2+t^2 = 1 \right\}$

SVD Tells vs whet is $T(S^3)$

Fact it is en "hyperellipse"

Left singuler vectors ere along axes of this "hyperellipse"



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

$T \quad R^4 \to R^4$

$T(v) = Av$

$$v = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \text{ on } S^3 \quad \text{iff}$$

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 = 1$$

$$T(v) = Av = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} \alpha_2 \\ 2\alpha_3 \\ 3\alpha_4 \\ 0 \end{bmatrix}$$

$$\alpha_2^2 + \frac{(2\alpha_3)^2}{4} + \frac{(3\alpha_4)^2}{9} = \alpha_2^2 + \alpha_3^2 + \alpha_4^2 \leq$$

$T(v)$ is on or inside hyperellipse

$$x_1^2 + \frac{x_2^2}{4} + \frac{x_3^2}{9} = 1 \qquad x_4 = 0$$

If v is on S^3 $T(v)$

is a vector of the form $\begin{bmatrix} a \\ b \\ c \\ 0 \end{bmatrix}$

with $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ on or inside the

hyperellipse $x^2 + \frac{y^2}{2^2} + \frac{z^2}{3^2} = 1$

Recall from last time:

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

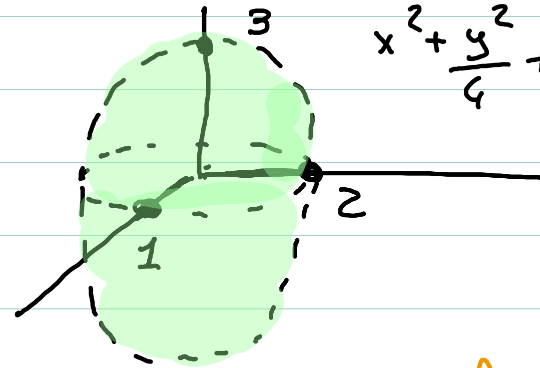
$$S^3 \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha_2 \\ 2\alpha_3 \\ 3\alpha_4 \\ 0 \end{bmatrix}$$ something in R^4 that projects onto

$x^2 + \frac{y^2}{4} + \frac{z^2}{9} = 1$

$\sigma_1 = 3$

$\sigma_2 = 2$

$\sigma_1 = 1$



$v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ $\quad u_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ along major axis z

$Av_1 = 3u_1$

$v_2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ $\quad u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ along y axis

$Av_2 = 2u_2$

$v_3 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ $\quad u_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ along x axis

$Av_3 = 3u_3$

Geometry of SVD

$\underset{m\times n}{A} = \underset{m\times m}{U}\,\Sigma\,\underset{n\times n}{V^T}$ $\quad T: \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad T(w)=Aw$

$\qquad\qquad\qquad\qquad\qquad B_1 \qquad\qquad\qquad B_2$

A has rank r

$B_1 = v_1 \cdots\cdots v_n \qquad B_2 = u_1 \cdots\cdots u_m$

$E = e_1\ e_2 \cdots e_k$ standard basis in $\mathbb{R}^k$

Th: Given $A = U\Sigma V^T \qquad A\in\mathbb{R}^{m\times n}$

$B_1 = v_1 \cdots v_n \qquad B_2 = u_1 \cdots u_m$

and $T \quad \mathbb{R}^n \to \mathbb{R}^m$

$T(w) = Aw$

$$\Sigma\,[w]_{B_1} = [T(w)]_{B_2}$$

which means that if we use B_1 in the domain and B_2 in the codomain, $T \quad \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad T(v)=Av$ is described by $\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$

Proof:

Change of basis stuff: $M_{B_i}^{B_j}$ = matrix of change of basis from B_i to B_j

$B_1 = v_1 \cdots v_n$, $B_2 = u_1 \cdots u_m$

$V = \begin{bmatrix} v_1 \cdots & v_n \\ & \end{bmatrix} = M_{B_1}^{E_n}$ i.e $V[w]_{B_1} = w$

$V^T = V^{-1} = M_{E_n}^{B_1}$ i.e $V^T w = [w]_{B_1}$

$U = \begin{bmatrix} u_1 \cdots & u_m \\ & \end{bmatrix} = M_{B_2}^{E_m}$ i.e $U[z]_{B_2} = z$

$U^T = U^{-1} = M_{E_m}^{B_2}$ i.e. $U^T z = [z]_{B_2}$

$A = U \Sigma V^T$ therefore

$\Sigma = U^T A V$

$\Sigma [w]_{B_1} = U^T A V [w]_{B_1} = U^T A w =$

$U^T \; T(w) = [T(w)]_{B_2}$

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

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "A"] \arrow[d, "V^T"] & \mathbb{R}^m \\
\mathbb{R}^n \arrow[r, "\Sigma"] & \mathbb{R}^m \arrow[u, "U"]
\end{tikzcd}$$

standard basis $E_n: e_1, \dots, e_n$ $\quad x$

standard basis $e_1, \dots, e_m$ $\quad b = Ax$

V basis $B_1: \{v_1, \dots, v_n\}$ $\quad y := V^T x = [x]_{B_1}$

U basis $\{u_1, \dots, u_m\} = B_2$ $\quad \Sigma y = b' = [Ax]_{B_2}$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}}_{m \times n} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sigma_1 y_1 \\ \vdots \\ \sigma_r y_r \\ \vdots \\ 0 \end{bmatrix} = b' = \begin{bmatrix} b_1' \\ \vdots \\ b_r' \\ \vdots \\ 0 \end{bmatrix}$$

if x is a unit vector $\quad y$ is a unit vector too:

if x satisfies $x_1^2 + x_2^2 + \dots + x_n^2 = 1$ so does y

what about b'?

$$\left(\frac{b_1'}{\sigma_1}\right)^2 + \dots \left(\frac{b_r'}{\sigma_r}\right)^2 = y_1^2 + \dots + y_r^2 \leq 1$$

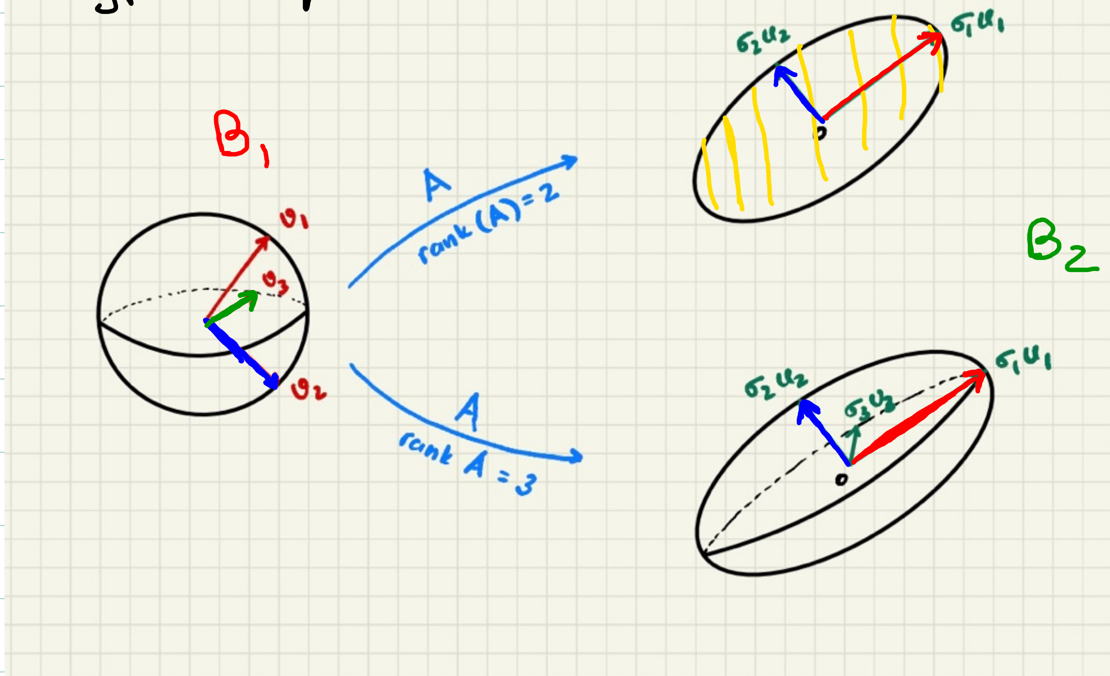
$$b_{r+1} = \dots = b_m = 0$$

$$\frac{b_1'^2}{\sigma_1^2} + \dots + \frac{b_r'^2}{\sigma_r^2} = 1$$

r dimensional hyperellipse in R^m (wrt basis U)

$m \times n$

A sends the unit sphere in $\mathbb{R}^n$ onto an hyperellipse in $\mathbb{R}^m$



The singular values σ_i give the lengths of the semiaxes. The right singular vectors $(v_1, \dots, v_r)$ give the directions of the semiaxes

$A v_1 = \sigma_1 u_1 \; \dots \; A v_r = \sigma_r v_r \qquad A v_{r+1} \cdots A v_n = 0$

Th : If $Q \in R^{n \times n}$ is orthogonal and $v \in R^n$ then $\|Qv\| = \|v\|$

Proof:

$$Qv = \underbrace{\begin{bmatrix} c_1 & \cdots & c_n \\ & & \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{v} = x_1 c_1 + \cdots + x_n c_n$$

$$\|Qv\|^2 = (x_1 c_1 + \cdots + x_n c_n)^T (x_1 c_1 + \cdots + x_n c_n) =$$

$$x_1^2 + x_2^2 + \cdots + x_n^2 = \|v\|^2$$

since $c_i^T c_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{ow} \end{cases}$

Therefore $\|Qv\| = \|v\|$

Example 90° rotation counterclockwise

$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ A is not diagonalizable but A has a SVD:

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

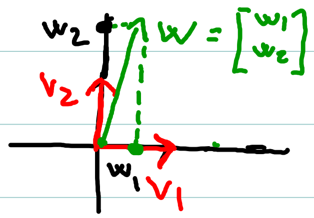
what do you expect σ_1 σ_2 be?

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

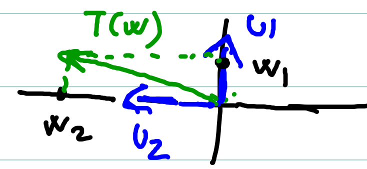
$$A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \sigma_1 = \sigma_2 = 1 \qquad v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

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



$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

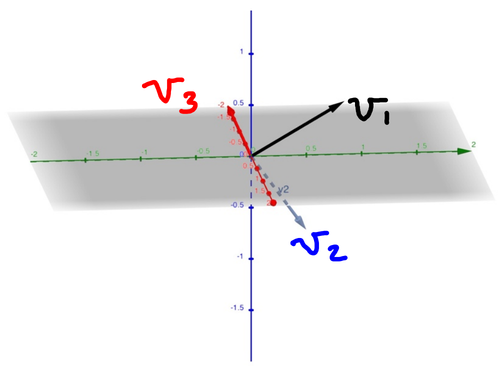


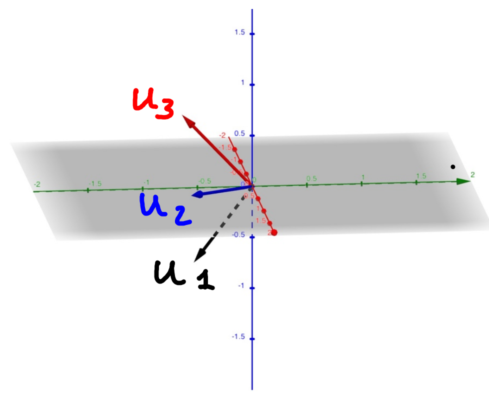
rotation "does nothing" if we use rotated basis in codomain. $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$

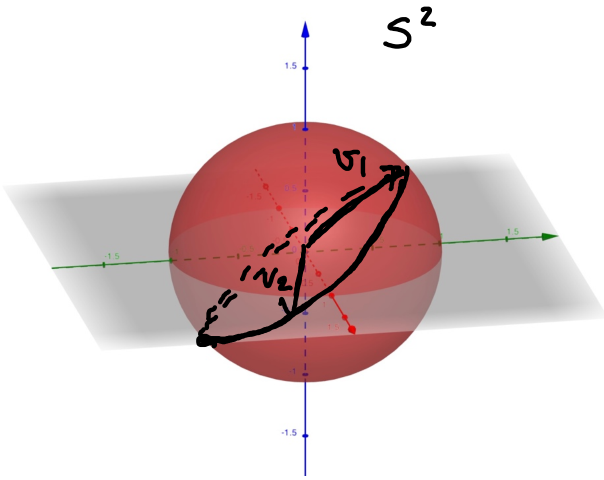
$$T\left(\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}\right) = w_1 u_1 + w_2 u_2$$

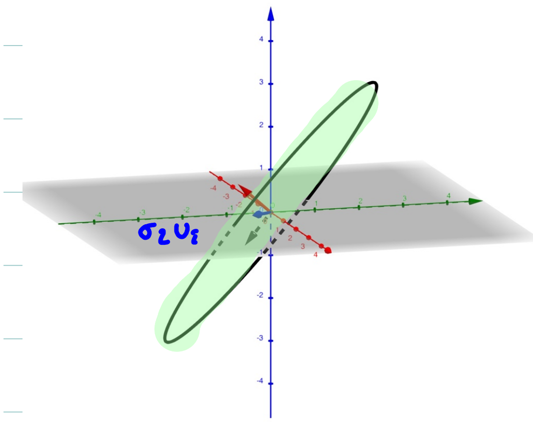
Julia example

$$\begin{bmatrix}1&1&1\\-1&0&2\\0&1&3\end{bmatrix} = \begin{array}{c}\begin{bmatrix}-0.30519&0.75731&-0.57735\\-0.50325&-0.64296&-0.57735\\-0.80844&0.11435&0.57735\end{bmatrix}\\ \begin{array}{ccc}u_1\quad\quad&u_2\quad\quad&u_3\end{array}\end{array}\begin{bmatrix}3.90448&0&0\\0&1.65981&0\\0&0&0\end{bmatrix}\begin{array}{cc}\begin{bmatrix}0.05072&-0.28522&-0.95711\\0.84363&0.52515&-0.11178\\0.53452&-0.80178&0.26726\end{bmatrix}&\begin{array}{c}v_1\\v_2\\v_3\end{array}\end{array}$$

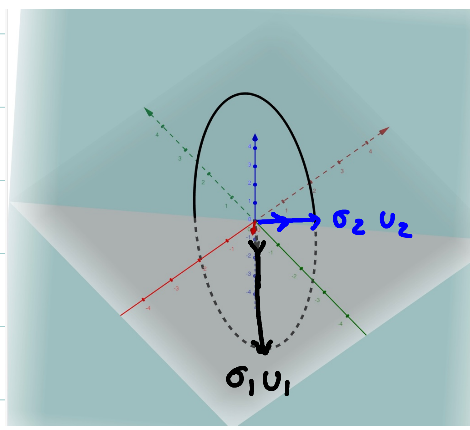








plane of u_1, u_2



Q: For which vectors v in S^2 is $T(v)$ on (not inside) ellipse?

v in span v_1 v_2 i.e v on a