TOPICS

- COMPLEX #'S VS. MATRICES
- POLAR DECOMPOSITION
- SINGULAR VALUE DECOMPOSITION (SVD)

REVIEW:

$$A = \overbrace{\underline{\tfrac{1}{2}(A+A^T)}}^{S} + \overbrace{\underline{\tfrac{1}{2}(A-A^T)}}^{K}$$

symmetric — skew symmetric

Helmholtz decomp $\Rightarrow$ potential vector field $\Rightarrow$ rotational vector field.

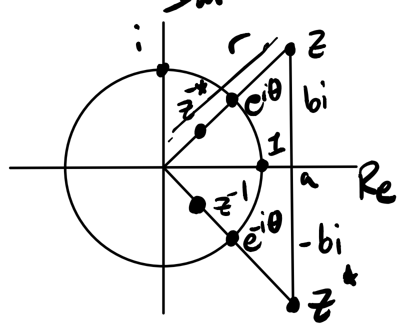
$\dot{x} = \frac{1}{2}(A+A^T)x$ — $\dot{x} = \frac{1}{2}(A-A^T)x$

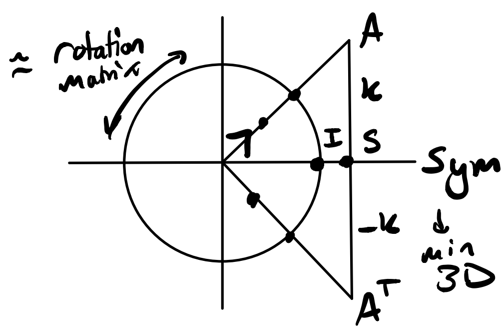
Comparisons

$$z = a + bi$$

$\approx$ skew $\to$ min 1D

Im





$z = a+bi$

$= re^{i\theta}$ $\quad \underline{r>0}$

$r = \sqrt{a^2+b^2} \Longrightarrow$

$= \sqrt{z^* z}$

$A^T = S - K$

$(A^* A)^{1/2}$

"magnitude of a matrix"

$$\underline{A^TA} = \begin{bmatrix} - A_1^T - \\ \vdots \\ - A_n^T - \end{bmatrix} \underbrace{\begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix}}_{} = \begin{bmatrix} A_1^TA_1 & \cdots & A_1^TA_n \\ A_n^TA_1 & \cdots & A_n^TA_n \end{bmatrix}$$

if R rotation

and $A' = [RA_1 \cdots RA_n] \to$ rotated / reoriented ea. col of A. by R.

$A' = R[A_1 \cdots A_n]$

weird

$(A'^T)A' = A^TR^TRA = A^TA \to$ reorienting cols of A doesn't change A^TA

Analogy r in $re^{i\theta}$ stays the same as one changes θ

$$AA^T = \begin{bmatrix} - \bar{a}_1^T - \\ \vdots \\ - \bar{a}_n^T - \end{bmatrix} \begin{bmatrix} \bar{a}_1 \cdots \bar{a}_n \end{bmatrix} = \begin{bmatrix} \bar{a}_1^T\bar{a}_1 & \cdots & \bar{a}_1^T\bar{a}_n \\ \bar{a}_n^T\bar{a}_1 & \cdots & \bar{a}_n^T\bar{a}_n \end{bmatrix}$$

$$A' = \begin{bmatrix} \bar{a}_1^TR \\ \vdots \\ \bar{a}_n^TR \end{bmatrix} = \begin{bmatrix} - \bar{a}_1^T - \\ - \bar{a}_n^T - \end{bmatrix} R$$

$A'A'^T = ARR^TA^T = AA^T$

$\Rightarrow AA^T$ doesn't change if we reorient the rows of A

Both A^TA & AA^T can be thought of as similar to r in some way.

What are some properties of A^TA, AA^T?

r **real** $\qquad A^TA,\ AA^T \to$ sym<u>metric</u>

$r > 0 \qquad \underline{A^TA > 0},\ \underline{AA^T > 0}$

$\to$ positive definite

Positive Definite: $A \in \mathbb{R}^{n \times n}$ $\quad A \succ 0, \; A = A^T$

LATEX: \succ

defn: $x^TAx > 0 \quad \forall x \neq 0 \in \mathbb{R}^n$

SIDE NOTE: usually assume that $A = A^T$

why? x^TAx makes sense even if $A \neq A^T$

$$x^T\left(\tfrac{1}{2}(A+A^T) + \tfrac{1}{2}(A-A^T)\right)x$$

$$x^T\tfrac{1}{2}(A+A^T)x + \tfrac{1}{2}\underbrace{x^T(A-A^T)x}_{=0}$$

$K = -K^T$

$x^TKx = \sum_{ij} K_{ij} x_i x_j$

$= 0$

$$x^TAx = \tfrac{1}{2}x^T(A+A^T)x$$

a skew sym comp of A doesn't affect x^TAx

Equiv Characterization: all eigenvalues are positive $\lambda > 0$

Note: only true if $A = A^T$

Assuming $A = A^T$:

$$A = \underbrace{P}_{\text{rotation}}\underbrace{D}_{\substack{\text{diagonal}\\ \text{real}}}P^{-1} = P\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}P^T$$

$$x^TAx = x^TP\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}\underbrace{P^Tx}_{z} = z^T\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_n\end{bmatrix}z$$

$$= \lambda_1 z_1^2 + \dots + \lambda_n z_n^2$$

if $\lambda_1 \dots \lambda_n > 0 \Rightarrow x^TAx > 0$

if $x^TAx > 0 \;\forall x \Rightarrow \lambda_1 \dots \lambda_n > 0$

select x to be an eigenvector

$\rightarrow x^TAx = x^T\lambda_i x = \lambda_i x^Tx > 0 \Rightarrow \lambda > 0$

here .. choosing x s.t. $P^Tx = z = \begin{bmatrix}0\\ 1\\ 0\end{bmatrix}$

A couple other terms:

Negative definite: $A \prec 0$ $\quad x^T A x < 0 \quad \forall x$

$\Leftrightarrow \lambda_1 \cdots \lambda_n < 0$

positive semi definite: $A \succeq 0$ $\quad x^T A x \geq 0 \quad \forall x \neq 0$

$\Leftrightarrow \lambda_1 \cdots \lambda_n \geq 0$

negative semi definite: $A \preceq 0$ $\quad x^T A x \leq 0 \quad \forall x \neq 0$

$\Leftrightarrow \lambda_1 \cdots \lambda_n \leq 0$

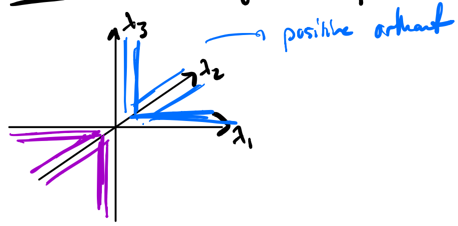
if not pos semi def or neg semi def

$\Rightarrow$ A is <u>indefinite</u>

if $x \neq 0$ $\quad x^T A x = 0 \ \ldots \ x \in N(A)$

A has 0 eigenvalue

<u>Cartoon</u> in the eigenvalue space ...



<u>Positive def matrices</u> are a <u>cone</u>

$A_1 \succ 0 \quad A_2 \succ 0$

$\Rightarrow \alpha A_1 \succ 0 \quad \alpha A_2 \succ 0$ for $\alpha > 0$

$\Rightarrow A_1 + A_2 \succ 0$ $\rightarrow$ positive combs of PD matrices still positive def.

pos def matrices are a convex set.

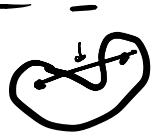
$\Rightarrow$ easily optimize over the set of PD matrices

branch of convex prog.

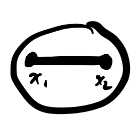
$\Rightarrow$ semidefinite programming (SDP)

<u>Convex</u> <u>set</u>: if two pts are in the set $\rightarrow$ all points between them are

Not convex



Convex



$x_1, x_2 \in \underset{\text{convex}}{\mathcal{X}}$ defn of $\mathcal{X}$ convex $x \in \mathcal{X}$ for $x = (1-\alpha)x_1 + \alpha x_2$ for $0 \le \alpha \le 1$

for PD matrices

$A_1, A_2 \succeq 0$

$\Rightarrow (1-\alpha)A_1 + \alpha A_2 \succeq 0$ for $0 \le \alpha \le 1$

$$x^T\left((1-\alpha)A_1 + \alpha A_2\right)x = (1-\alpha)\underbrace{x^T A_1 x}_{\ge 0} + \alpha \underbrace{x^T A_2 x}_{\ge 0}$$

$$\ge 0$$

Congruent transformation

A is congruent to BAB^T for invertible B

congruent transforms preserve definitness properties

$\underline{x^T A x}\ \forall x \qquad \underline{z^T BAB^T z}\ \underline{\forall z}$ since B is invertible pick $x = B^T z$

A is similar to $\underline{BAB^{-1}}$ for invertible B

similarity transforms preserve eigenvalues

congruent transform = similarity transform when B is a rotation

---

$A^TA,\ AA^T \ldots \Rightarrow A^TA \succeq 0 \quad AA^T \succeq 0$

$$x^T A^T A x = |Ax|^2 \ge 0 \qquad x^T A A^T x = |A^T x|^2 \ge 0$$

even if A is not square ...

tall: $x^T \begin{bmatrix} A^T \end{bmatrix} \begin{bmatrix} A \end{bmatrix} x > 0$ full col rank

fat: $x^T \begin{bmatrix} A^T \end{bmatrix} \begin{bmatrix} A \end{bmatrix} x \ge 0$

still works for non square matrices
and still pos semi def.

$A \in \mathbb{R}^{m \times n}$ $\quad$ $\left.\begin{array}{l} A^T A \in \mathbb{R}^{n \times n} \\ A A^T \in \mathbb{R}^{m \times m} \end{array}\right\} \to$ for non square A not same dim.

$(z^* z)^{1/2} = |z| > 0$

$(A^T A)^{1/2}$: take square root of evals of $A^T A$ ... $\Leftarrow$
(spectral mapping thm)

if $\lambda \in \sigma(A^T A) \Rightarrow \lambda$ real $\lambda \geq 0$

$\Rightarrow \sqrt{\lambda}$ real $\sqrt{\lambda} \geq 0$

$(A A^T)^{1/2}$: if $\lambda \in \sigma(A A^T) \Rightarrow \lambda$ real $\lambda \geq 0$

$\Rightarrow \sqrt{\lambda}$ real $\sqrt{\lambda} \geq 0$

Non zero eigenvalues

$\underline{(A^T A)^{1/2}}, \underline{(A A^T)^{1/2}} = \{\sigma_1 \dots \sigma_k\}$ rank of A

called singular values

Symmetric matrices
$\Rightarrow$ diagonalizable by a rotation, evals are real

much richer than z.

$z = \underline{r} e^{i\theta}$ $\qquad$ A

$r = \sqrt{z^* z} > 0$ $\qquad$ $(A^T A)^{1/2} > 0$ same for $A' = RA$

independ of θ $\qquad$ $(A A^T)^{1/2} > 0$ same for $A' = AR$

mag of z $\qquad$ "mag" of A

$z = re^{i\theta}$

Sym: P, Q

Assume $A \in \mathbb{R}^{n \times n}$ invertible

$A = ?$

needs to be invertible

needs to be invertible

$$A = \underbrace{A(A^TA)^{-1/2}}_{R} \underbrace{(A^TA)^{1/2}}_{P} \iff A = \underbrace{(AA^T)^{1/2}}_{P'} \underbrace{(AA^T)^{-1/2}A}_{R}$$

rotations /reflections | pos def | POLAR DECOMP | pos def | rotation/reflection

$$R^TR = I \qquad RR^T = I$$

$$(A^TA)^{-1/2} A^TA (A^TA)^{-1/2} = I \qquad (AA^T)^{-1/2} AA^T (AA^T)^{-1/2} = I$$

$z = re^{i\theta}$

$A = RP$

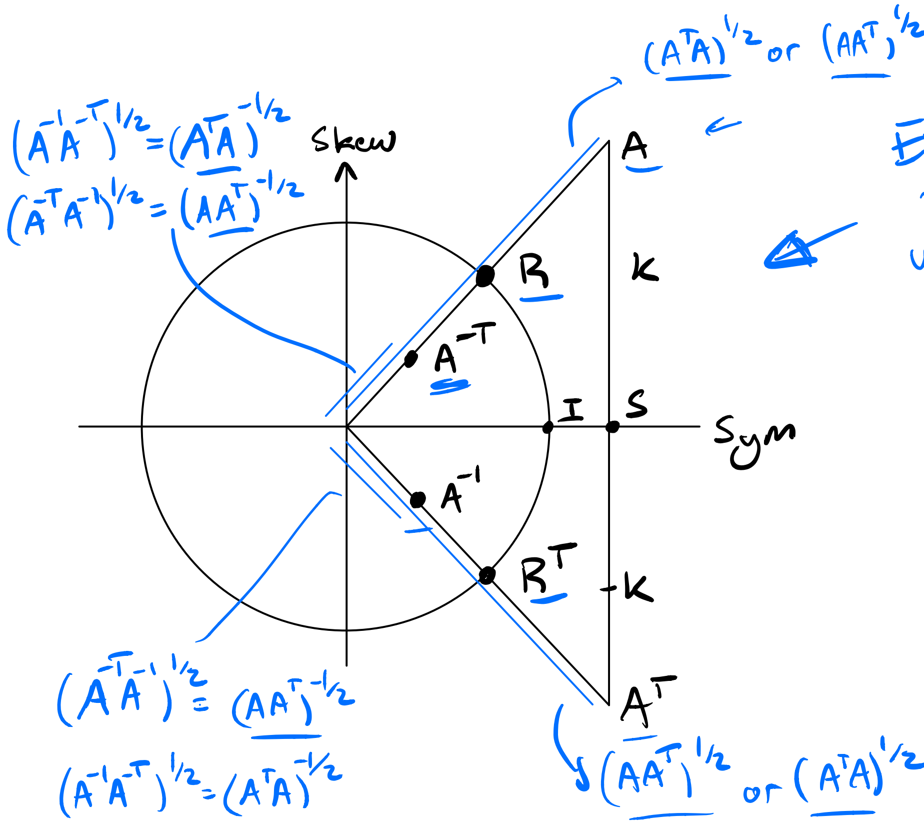
$R \sim e^{i\theta}$

$A = P'R$

Cartesian: $A = \underbrace{\tfrac{1}{2}(A+A^T)}_{S} + \underbrace{\tfrac{1}{2}(A-A^T)}_{K}$

Polar $A = (AA^T)^{1/2}\underline{(AA^T)^{-1/2}A} \qquad A = A(A^TA)^{-1/2}(A^TA)^{1/2}$

$$R = A(A^TA)^{-1/2} = (AA^T)^{-1/2}A$$



Exercise:

write these metrices in terms of u, v, Σ

$$A = PR \Rightarrow A^{-1} = R^{-1}P^{-1} = \underline{R^T}\,\underline{P^{-1}}$$

$$A = RP' \Rightarrow A^{-1} = P'^{-1}R^T$$

SVD : Singular value decomposition:

$(A^*A)^{1/2}$ & $(AA^*)^{1/2}$ $\Rightarrow$ nonzero eigenvalues are called singular values

can be computed whether or not A is square.

$\{\sigma_1 \ldots \sigma_k\}$

how much the matrix stretches the space

$\Rightarrow$ eigenvalues non negative & real.

$\Rightarrow$ eigenvectors are orthogonal

$A \in \mathbb{C}^{m \times n}$

take $A^*A \Rightarrow$ eigenvectors $[v_1 \ldots v_n]$

$$A^*Av_i = \sigma_i^2 v_i \qquad v_i \in \mathbb{C}^n \quad |v_i| = 1.$$

$$v_i^* v_i = 1$$

consider $u_i = \dfrac{Av_i}{\sigma_i} \in \mathbb{C}^n$

direction v_i gets rotated to

$$|u_i| = (u_i^* u_i)^{1/2} = \frac{v_i^* A^*}{\sigma_i}\frac{Av_i}{\sigma_i} = \frac{v_i^* v_i \sigma_i^2}{\sigma_i^2} = 1$$

$$AA^* u_i = AA^* \frac{Av_i}{\sigma_i} = \frac{A\sigma_i^2 v_i}{\sigma_i} = \sigma_i A v_i$$

$$= \sigma_i^2 \frac{Av_i}{\sigma_i}$$

$$AA^* u_i = \sigma_i^2 u_i \qquad = \sigma_i^2 u_i$$

evec u_i

write for all v_i's

to simplify assume $A \in \mathbb{C}^{n \times n}$

$$\left[u_1 \cdots u_n\right] = \left[\frac{Av_1}{\sigma_1} \cdots \frac{Av_n}{\sigma_n}\right] \qquad U = \left[u_1 \cdots u_n\right]$$

$$U = AV\Sigma^{-1} \qquad V = \left[v_1 \cdots v_n\right]$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix}$$

$$A = U \Sigma V^{-1}$$

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

$$V^* V = I$$

Ingeneral:

where $U^* U = I$

$V^* V = I$

$$\underbrace{A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^*}_{\text{singular value decomposition}} \qquad \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{bmatrix}$$

(SVD)

SVD is very general

any $A \in \mathbb{C}^{m\times n}$ has an SVD.

some insight

breaks down the operation of a matrix into 3 steps

$$A = {}_{m}\overset{m}{U}\ \begin{array}{c} \\ k \\ m-k \end{array}\!\!\begin{array}{c} \begin{array}{cc} k & n-k \end{array} \\ \begin{bmatrix} \Sigma_{k\times k} & 0 \\ 0 & 0 \end{bmatrix} \end{array}{}_{n}\overset{n}{V}^{*}$$

another rotation | positive stretching of some axes... | "rotation"

Very good visualization on wikipedia

$$A = \begin{array}{c} \begin{array}{cc} k & m-k \end{array} \\ \begin{bmatrix} \underline{U_1} & U_2 \end{bmatrix} \end{array} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix}\begin{array}{c} k \\ n-k \end{array}$$

U_1 orthonormal basis for $\underline{R(A)}$ $\qquad$ V_1 orthonormal basis for $\underline{R(A^*)}$

U_2 orthonormal basis for $\underline{N(A^*)}$ $\qquad$ $\underline{V_2}$ orthonormal basis for $\underline{N(A)}$

$$A^* = V\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}U^*$$

$$A^{-1} = V\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}U^*$$

$$A^{-*} = U\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}V^*$$

Note: the 0's here are different sizes

$$\underline{(A^*A)^{1/2}} = \left(V\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}U^*U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^*\right)^{1/2} = \underline{V}\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}\underline{V^*}$$

$$\underline{(AA^*)^{1/2}} = \left(U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^*V\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}U^*\right)^{1/2} = \underline{U}\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}\underline{U^*}$$

Moore Penrose Pseudo Inverse

$$A^+ = V^*\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}U$$

$$R = A(A^*A)^{-1/2} = U\begin{bmatrix}\Sigma & 0\\ 0 & 0\end{bmatrix}V^*V\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}V^* = U_1V_1^*$$

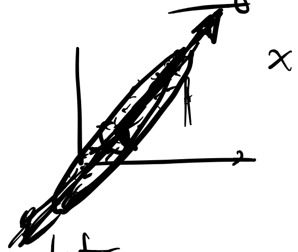
for invertible square A.

$$R = UV^*$$

PCA : Principle component analysis.

$y = Ax$ → least square. → dim reduction technique.

LS: $x = (A^TA)^{-1}A^T y$



$$\text{\# meas}\ \begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

dep var.   data   ind var.

$$A = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

$$y = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ & 0 & \sigma \\ & 0 & \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \begin{matrix} \leftarrow \text{big } \sigma\text{s} \\ \leftarrow \text{small } \sigma\text{s} \end{matrix}$$

$$\begin{bmatrix} \sigma_1 & & \\ & \sigma_{10} & \\ & & \sigma_k \end{bmatrix}$$

largest $\sigma_1 \to 10^5$

smallest $\sigma_k \to 10^2$

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} x$$

ignore → x_2'