

Golden-Thompson Ineq

15-859N
2/22/21

Thm Sym $A \& B$ then (real)

$$\text{Tr}(e^{A+B}) \leq \text{Tr}(e^A e^B)$$

Note $e^{A+B} = e^A e^B$ iff $AB = BA$

In general: $\text{Tr}(e^{A+B+C}) \neq \text{Tr}(e^A e^B e^C)$

The proof uses both eigenvalues
& singular values

Intuition: Think of e^A as transition matrix

Def B similar to A if \exists invertible P
st $B = P^{-1}AP$

Properties preserved:

rank, eigenvalues, trace, minimal polynomial

Not preserved:

angle between eigenvectors, singular values.

Real linear algebra needs complex numbers!

2

$$\mathbb{C} = \{a+ib \mid a, b \in \mathbb{R}\}$$

$$z^* = (a+ib)^* = a - ib.$$

$$\text{Def } (A^*)_{ij} \equiv A^*_{ji} \quad (\text{Hermitian})$$

Def $U \in \mathbb{C}^{n \times n}$ is unitary $U^*U = I$

$$\text{Thus } U^* = U^\dagger \quad \& \quad UU^* = I$$

Thus unitary are closed under Hermitian
& Product.

Note: Def $\langle a, b \rangle = a^* b$.

$$\text{Thus } \langle a, b \rangle = \langle Ua, Ub \rangle.$$

Def A is unitarily similar to B if

$$A = U^\dagger B U \text{ or } A = U^T B U$$

3

Schur Decomp $\forall A \in \mathbb{C}^{n \times n} \exists$ unitary U st

U^*AU is upper triangular &

$$A_{ii} = \lambda_i \quad |\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

P_f is Gram-Schmidt!

A will have n eigenvalues

(roots of polynomial $\det[A - \lambda I] = P(\lambda)$)

but possibly not n eigenvectors.

Thm (Singular Value Decomp) for $A \in \mathbb{R}^{n \times n}$

$\exists U, V \in \mathbb{R}^{n \times r}$ st $U^T U = V^T V = I$

& $A = U \Sigma V^T$

where $\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ 0 & \ddots & 0 \\ & & \sigma_r \end{pmatrix}$ $\sigma_i > 0$.

4

Pf $A^T A$ is SPSD thus $A^T A = \bar{Q} \bar{\Lambda}^2 \bar{Q}^T$

Let Q be \bar{Q} with zero eigenvectors removed.

Λ be $\bar{\Lambda}$ " " eigenvalues " ".

Thus $A^T A = Q \Lambda^2 Q^T$

Note: We have found $V = Q$ & $\Sigma = \Lambda$

We need to find U !

Claim: $U = A Q \Lambda^{-1}$ works i.e.

$$1) U^T U = I$$

$$2) A = U \Lambda Q^T$$

$$1) U^T U = \bar{\Lambda}^{-1} Q^T A^T A Q \bar{\Lambda}^{-1}$$

$$= \bar{\Lambda}^{-1} Q^T Q \bar{\Lambda}^2 Q^T Q \bar{\Lambda}^{-1} = I$$

$$\begin{aligned} \text{Consider } U^T A Q &= \bar{\Lambda}^{-1} Q^T A^T A Q = \bar{\Lambda}^{-1} Q^T Q \bar{\Lambda}^2 \\ &= \bar{\Lambda}^{-1} \bar{\Lambda}^2 = \bar{\Lambda} \end{aligned}$$

$$\text{Thus } A = U \Lambda Q^T$$

Def Λ are singular values of A

The Polar Decomposition

5

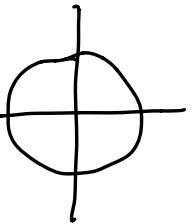
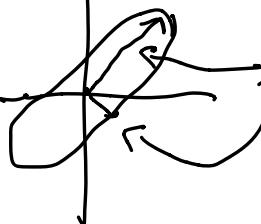
Consider:

$$\begin{aligned} A &= U \Lambda U^T U V^T \\ &= (U \Lambda U^T)(U V^T) = M \cdot P \end{aligned}$$

M spsd & P unitary.

(Polar DecomP) $\exists M, P$ st $A = MP$

eg $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = 1 \quad | \quad \vec{v}_1 \cdot \vec{v}_2 = 1$

$A =$  \Rightarrow  left sing vectors

Weyl: $\Rightarrow \lambda_1 \leq \vec{v}_1$ &

$$\lambda_1 + \lambda_2 \leq \vec{v}_1 + \vec{v}_2$$

$$\text{ie } \lambda_2 \geq \vec{v}_2$$

$$\text{Cor } |\det(A)| = |\prod \lambda_i| = \prod |\vec{v}_i|$$

Weyl's Majorant Thm

6

Thm $A \in \mathbb{R}^{n \times n}$, singular values $\tau_1 \geq \dots \geq \tau_n \geq 0$

with eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$

$$\text{then } \left| \prod_{i=1}^k \lambda_i \right| \leq \prod_{i=1}^k \tau_i \quad \forall k \quad 1 \leq k \leq n$$

$$\text{Cor} \quad \sum_{i=1}^k |\lambda_i|^m \leq \sum_{i=1}^k \tau_i^m \quad m \in \{1, 2, \dots\}$$

pf of Cor, see HW3.

Courant-Fischer for Singular Values

Thm $A \in \mathbb{R}^{n \times n}$

$$\tau_k(A) = \min_{\dim(T)=n-k+1} \max_{x \in T} \frac{\|Ax\|_2}{\|x\|_2} = \frac{(x^T A^T A x)^{1/2}}{(x^T x)^{1/2}}$$

$$\tau_k(A) = \max_{\dim(S)=k} \min_{x \in S} \frac{\|Ax\|_2}{\|x\|_2}$$

Pf The eigenvalue proof works for singular values.

7

Let $a_0 \geq a_1 \geq \dots \geq a_n \quad a_i \in \mathbb{R}$

$b_0 \geq \dots \geq b_n \quad b_i \in \mathbb{R}$

Def $\langle b \rangle$ interlaces with $\langle a \rangle$

if $a_0 \geq b_0 \geq a_1 \geq \dots \geq b_n \geq a_n$

(Cauchy) If A is sym and A_1 is

A with same column and row removed
then $\lambda(A_1)$ interlace with $\lambda(A)$

pf: WLOG assume it is last row & column.

B is A with last row & column removed

$$\lambda_{k+1}(A) = \min_{\dim T=n-k+1} \max_{X \in T} \frac{x^T A x}{x^T x}$$

$$\geq \min_{\dim T=n-k+1} \max_{\substack{X \in T \\ x_n=0}} \frac{x^T A x}{x^T x}$$

$$\geq \min_{\substack{\dim T = n-k \\ X_n = 0}} \max_{\substack{X \in T \\ X_n = 0}} \frac{x^T A x}{x^T x}$$

$$= \min_{\dim(T) = n-k} \max_{X \in T} \frac{x^T B x}{x^T x} = \lambda_k(B) \quad \square$$

the sing values interlace.

8

Thm If A is square ($A \in \mathbb{R}^{n \times n}$) and A upper tri
and B is A with last column & row removed

$$\text{then } \tau_k(A) \geq \tau_k(B) \geq \tau_{k+1}(A)$$

$$\tau_k(A) = \min_{\dim T=n-k+1} \max_{X \in T} \left(\frac{\bar{x}^T A^T A x}{x^T x} \right)^{1/2}$$

$$\geq \min_{\dim T=n-k+1} \max_{\substack{X \in T \\ X_n=0}} \left(\frac{\bar{x}^T A^T A x}{x^T x} \right)^{1/2}$$

$$\geq \min_{\substack{\dim T=n-k \\ X_n=0}} \max_{\substack{X \in T \\ X_n=0}} \left(\frac{\bar{x}^T A^T A x}{x^T x} \right)^{1/2}$$

$$= \min_{\dim(T)=n-k} \max_{X \in T} \frac{\bar{x}^T B x}{x^T x}$$

□

$\text{pf}(\text{Weyl})$

$X^{n \times n}$ wlog X is upper tri

$$\& \text{diag}(X) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad |\lambda_i| \geq |\lambda_{i+1}|$$

Let X_k be $k \times k$ upper principal submatrix.

$$|\lambda_1 \cdots \lambda_k| = |\det(X_k)| = \tau_1(X_k) \cdots \tau_k(X_k) \stackrel{\substack{\text{Used} \\ \text{interlacing}}}{\leq} \tau_1(X) \cdots \tau_k(X) \quad \square$$

$$\text{Def } |X| = (X^T X)^{1/2}$$

$$\text{Note } AA^T = M P P^T M = M^2 = U \Lambda^2 U^T$$

$$\therefore (AA^T)^{1/2} = U \Lambda U^T$$

$$\text{Thus } \lambda(|A|) = \text{sing}(A) \quad (\text{Weyl}) \quad \text{Tr}(X^m) \leq \sum \tau_i(X^m)$$

$$\text{Cor } \text{Tr}(X^m) \leq \text{Tr}(|X|^m)$$

$$\text{pf } \lambda(X) = \lambda_1, \dots, \lambda_n \quad \tau_1, \dots, \tau_n$$

$$\text{Tr}(X^m) = \sum \lambda_i^m \leq \sum \tau_i^m = \text{Tr}(|X|^m) \quad (\text{Weyl})$$

Lie Product Formula

$$e^{A+B} = \lim_{N \rightarrow \infty} \left(e^{A/N} e^{B/N} \right)^N$$

Pf Compare as $N \rightarrow \infty$

$$X_N = e^{(A+B)/N} \quad \& \quad Y_N = e^{A/N} e^{B/N}$$

Tay or Exp

$$X_N = 1 + \frac{A+B}{N} + O(N^{-2})$$

$$Y_N = \left[1 + \frac{A}{N} + O(N^{-2}) \right] \left[1 + \frac{B}{N} + O(N^{-2}) \right]$$

$$= \left[1 + \frac{A}{N} + \frac{B}{N} + O(N^{-2}) \right]$$

$$\text{Thus } X_N - Y_N = O(N^{-2})$$

Let's compare: $\|X_N^N - Y_N^N\|$?

Def $\|\cdot\|$ is a matrix norm if

- 1) $\|A\| \geq 0$
- 2) $\|A+B\| \leq \|A\| + \|B\|$
- 3) $\|AB\| \leq \|A\| \|B\|$

Claim $\|X^N - Y^N\| \leq NM^{N-1} \|X - Y\|$

where $M = \max(\|X\|, \|Y\|)$ & $\|\cdot\|$ norm

Pf

$$\begin{aligned} X^N - Y^N &= (X^N - X^{N-1}Y) + (X^{N-1}Y - X^{N-2}Y^2) \\ &\quad + \dots (XY^{N-1} - Y^N) \\ &= X^{N-1}(X - Y) + X^{N-2}(X - Y)Y \\ &\quad + \dots (X - Y)Y^{N-1} \\ &\leq NM^{N-1} \|X - Y\| \end{aligned}$$

12

$$\text{To Show } \lim_{N \rightarrow \infty} (X_N^N - Y_N^N) = 0$$

$$\text{Note } \|e^A\| = \left\| \left(I + A + \frac{A^2}{2!} + \dots \right) \right\|$$

$$\leq \|I\| + \|A\| + \frac{\|A^2\|}{2} + \dots$$

$$\leq \|I\| + \|A\| + \frac{\|A\|^3}{2} + \dots$$

$$= e^{\|A\|}$$

$$X_N = e^{(A+B)/N}$$

Thus

$$\|X_N\| \leq e^{\|A+B\|/N} \leq e^{(\|A\| + \|B\|)/N}$$

$$\|Y_N\| \leq \|e^{AN} \cdot e^{BN}\| \leq e^{(\|A\| + \|B\|)/N}$$

Therefore

$$M^N = \max(\|X\|, \|Y\|)^N \leq e^{\|A\| + \|B\|}$$

finally

$$\|X_N^N - Y_N^N\| \leq N e^{\|A\| + \|B\|} O(N^{-2})$$

$$= O(1/N)$$

Back to proof of Golden-Thompson

$$\text{Let } X = e^{A/2^N} \quad Y = e^{B/2^N}$$

to show $\text{tr}((XY)^{2^N}) \leq \text{tr}(X^{2^N} \cdot Y^{2^N})$

(unmix the X & Y)

$$\text{Recall: } \text{tr}(e^{A+B}) \stackrel{\text{Lee}}{=} \text{tr}((XY)^{2^N}) \quad N \rightarrow \infty$$

$$\text{tr}(X^{2^N} \cdot Y^{2^N}) = \text{tr}(e^A e^B)$$

$$\text{tr}((XY)^{2^N}) \leq |\text{tr}[(XY)^{2^N}]| \quad 14$$

$$\leq \text{tr}[|XY|^{2^N}] \quad [\text{Cor to Weyl.}]$$

$$= \text{tr}\left[\left(YXXX\right)^{2^{N-1}}\right]$$

$$= \text{tr}\left[YX^2YYX^2YY\cdots X^2Y\right]$$

$$= \text{tr}\left[X^2Y^2\cdots X^2Y^2\right] \quad \begin{bmatrix} \text{tr}(uv) = \\ \text{tr}(vu) \end{bmatrix}$$

$$= \text{tr}\left(\left(X^2Y^2\right)^{2^{N-1}}\right)$$

$$\leq \text{tr}\left(\left(X^4Y^4\right)^{2^{N-2}}\right) \quad (\text{repeating})$$

$$\leq \text{tr}\left(X^{2^N}Y^{2^N}\right)$$

Done !!