Lecture 6

- Eckart - Young Mirsky (continued)
  - ↳ Frobenius norm proof
- A wrong proof

Frobenius Norm Property: invariant to orthonormal transformations

$$\| A \overset{U \text{ orthonormal}}{U} \|_F = \| UA \|_F = \| A \|_F$$

- want to connect Frobenius norm to singular values:

Ayan attempt:

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

$$\| AU \|_F = \| U \Sigma V^T U \| =$$

Recall: • $\| A \|_F = \sqrt{\text{trace}(A^T A)}$

• $\text{trace}(AB) = \text{trace}(BA)$

$$\| AU \|_F = \sqrt{\text{trace}((U \Sigma V^T U)^T (U \Sigma V^T U)}$$

$$= \sqrt{\text{trc}(U^T V \Sigma^T U^T U \Sigma V^T U}$$

$$= \sqrt{\text{trc}(U^T V \Sigma^2 V^T U)}$$

$$= \sqrt{\text{trc}(U U^T V \Sigma^2 V^T)}$$

$$= \sqrt{\text{trc}(V \Sigma^2 V^T)}$$

$$= \sqrt{\text{trc}(V^T V \Sigma^2)}$$

$$= \sqrt{\text{trc}(\Sigma^2)}$$

Ranade's Approach:

$$\| AU \|_F = \sqrt{\text{trace}((AU)^T AU)}$$

$$= \sqrt{\text{tr}(U^T A^T A U)}$$

$$= \sqrt{\text{tr}(U U^T A^T \text{-})}$$

$$= \sqrt{\text{tr}(A^T A)}$$

$$= \| A \|_F$$

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

$= \|\Sigma\|_F$

$= \sqrt{\sum_{i=1}^{n} \sigma_i^2}$

Thm: $A \in \mathbb{R}^{m \times n}$ $\quad A = U\Sigma V^T$ $\quad A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^T$ $\quad \sigma_1 \geq \dots \geq \sigma_k \geq \sigma$

$m > n$

then we have that

① $A_k = \underset{\substack{B \in \mathbb{R}^{m \times n} \\ B = \text{rank}(k)}}{\text{argmin}} \|A - B\|_2$

② $A_k = \underset{\substack{B \in \mathbb{R}^{m \times n} \\ B = \text{rank}(k)}}{\text{argmin}} \|A - B\|_F$

→ proving ②, want $\|A - B\|_F \geq \|A - A_k\|_F$

↳ no matter what B I choose, want this inequality to hold (ie to have $\|A - A_k\|_F$ be the minimizer)

→ evaluate $\|A - A_k\|_F$

$$\|A - A_k\| = \left\| \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T - \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^T \right\|$$

$$= \left\| \sum_{i=k+1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T \right\|$$

$$= \sqrt{\sum_{i=k+1}^{n} \sigma_i^2}$$

(proved above)

↳ Frobenius norm of a matrix = the square root of the sum of its singular values

$$= \sqrt{\sum_{i=k+1}^{n} \sigma_i^2(A)}$$

↳ notation – just singular vals of A

→ compare to the singular values : Frob. norm of $A - B$

wts: $\sum_{i=1}^{n} \sigma_i^2(A - B) \geq \sum_{i=k+1}^{n} \sigma_i^2(A)$

$\sigma_i(A - B) \geq \sigma_{k+1}(A)$

* this is the most important thing to establish bc it shows this which implies this

then sum all the sigmas

$\sigma_i^2(A - B) \geq \sigma_{k+1}^2(A)$

$\sum_{i=1}^{n-k} \sigma_i^2(A - B) \geq \sum_{i=1}^{n-k} \sigma_{k+1}^2(A)$

what we know:

- B : rank(K)
- Frobenius norm connected to singular values
- spectral norm of a matrix is the max singular value of the matrix

Ayah attempt

$\sigma_{k+i}(A)$ → largest singular value of A after the first $(k+i-1)$ are removed

↳ $(k+i)^{th}$ largest s.v. of A

$= \|A - A_{k+i-1}\|_2 \qquad A_j = \sum_{i=1}^{j} \sigma_i \vec{u}_i \vec{v}_i^T$

$\|A - A_{k+i-1}\| = \sqrt{tr(A - A_{k+i-1})}$

$= \| \sum_{i=1}^{n} \sigma_i - \sum_{i=k+1}^{n} \sigma_i \|$ ↳ $\sigma_1 = \sigma_{max}$

↳ Denote $A - B = C$

$\sigma_i(A-B) = \sigma_i(C) = \underbrace{\|C - C_{i-1}\|_2}$

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} - \begin{bmatrix} | & | & & | & & \\ a_1 & a_2 & \cdots & a_k & 0 & 0 \cdots \\ | & | & & | & & \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & a_{k+1} & \cdots & a_n \end{bmatrix}$$

Consider top $(i-1)$ SV's. Remove them. The spectral norm of the remaining matrix must be the i^{th} singular value of C.

↳ use the fact that the two norm is a norm ⇒ connect it back to A

$\sigma_i(A - B) = \|C - C_{i-1}\|_2 + \|B - B_k\|_2$

↳ let's think about B (rk(k))

$\sigma_{k+1}(B) = 0$ (by def of rk(k))

$\underbrace{\|B - B_k\|_2 = 0}$

→ triangle ineq.

$C_i = \sum_{\ell=1}^{i} \sigma_\ell(C) \vec{u}_\ell(C) \vec{v}_\ell(C)$

$\sigma_i(A - B) = \|C - C_{i-1}\|_2 + \|B - B_k\|_2$

$$\geq \| \underbrace{C + B}_{A} - C_i - B_k \|_2$$

$$= \| A - C_i - B_k \|_2$$

→ know that $Rk(B_k) = k$ $Rk(C_{i-1}) \leq i-1$ (bc C is a dyad)

Define $D = C_{i-1} + B_k$

↳ $rk(D) \leq rk(B_k) + rk(C_{i-1}) = i-1+k$

$$\sigma_i(A-B) \geq \| A - (C_i + B_k) \|_2$$

$$= \| A - D \|_2$$

→ Consider the optimization problem

$\operatorname{argmin} \| A - D \|_2$, $rk(D) \leq i+k-1$

↳ $\operatorname{argmin} \| A - D \| = A_{k+i-1}$ } minimum value of this is given by $\sigma_{k+i}(A)$ (bc of the spectral norm properties)

↓

$\min \| A - D \| = \sigma_{k+i}(A)$

$$\sigma_i(A-B) \geq \sigma_{k+i}(A)$$

A wrong proof :

$$\min_{rk(B)=k} \| A - B \|_F = \min_{rk(B)=k} \| U \Sigma V^T - B \|_F$$

$$= \min \| \Sigma - U^T B V \|_F$$

$$= \min \| \Sigma - Z \|_F$$

$Z : rk(k)$, diag

↳ choose

Proof by lots of talking