Lecture 5

- SVD $\to$ geometry
- low rank approximation

SVD Review

$$A = U \Sigma V^T \xleftarrow{\text{SVD}} \longrightarrow A^T A$$

$\lambda_1 \geq \; \dots \geq \lambda_r$:= non-zero e-vals of $A^T A$

$\lambda_{r+1}, \dots, \lambda_n = 0$

$\vec{v}_1, \; \dots, \vec{v}_n$:= e-vects corr λ_i

$$V = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix}$$

Define: $\sigma_i = \sqrt{\lambda_i}$

$\vec{u}_i$ defined as : $A\vec{v}_i = \sigma_i \vec{u}_i$ $\quad i \leq r$ (i.e, non-zero e-vals)

Claim: $\vec{u}_i$ vectors are orthonormal

attempt

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

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

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

$$= V \Sigma^t \Sigma V^T$$

$\vec{u}_i, \vec{u}_j$ } orthonorm $\to$ dp should $= 0$

$$\vec{u}_i^T \vec{u}_j = \frac{(A\vec{v}_i)}{\sigma_i}^T \frac{(A\vec{v}_j)}{\sigma_j}$$

$$= \frac{1}{\sigma_i \sigma_j} \vec{v}_i^T A^T A \vec{v}_j$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j} \vec{v}_i^T \vec{v}_j \qquad \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times m}{V^T}$$

↳ to find remaing $\vec{u}_x$'s use gram-schmidt

$$V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_r & \vec{v}_{r+1} & \cdots & \vec{v}_n \end{bmatrix}$$

$$AV_r = \underline{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & & \sigma_r \end{bmatrix}}$$

$$AV = U\Sigma$$

↳ multiply both sides by $V^{-1} = V^T$

$$A\underbrace{VV^T}_{I} = U\Sigma V^T$$

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

Geometry of SVD

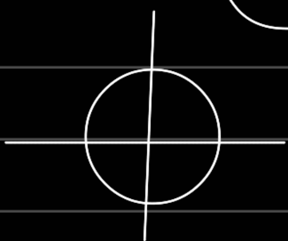
$$A = U\Sigma V^T$$

$$A\vec{x} = U\Sigma V^T\vec{x}$$

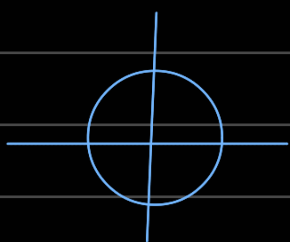
V^T ↳ orthonormal ⇒ rotation/reflection of $\vec{x}$

Σ ↳ scaling

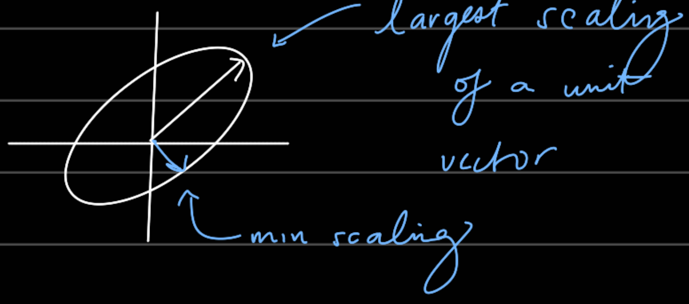
U ↳ "I" ⇒ rot/ref



⇒ orthonormal



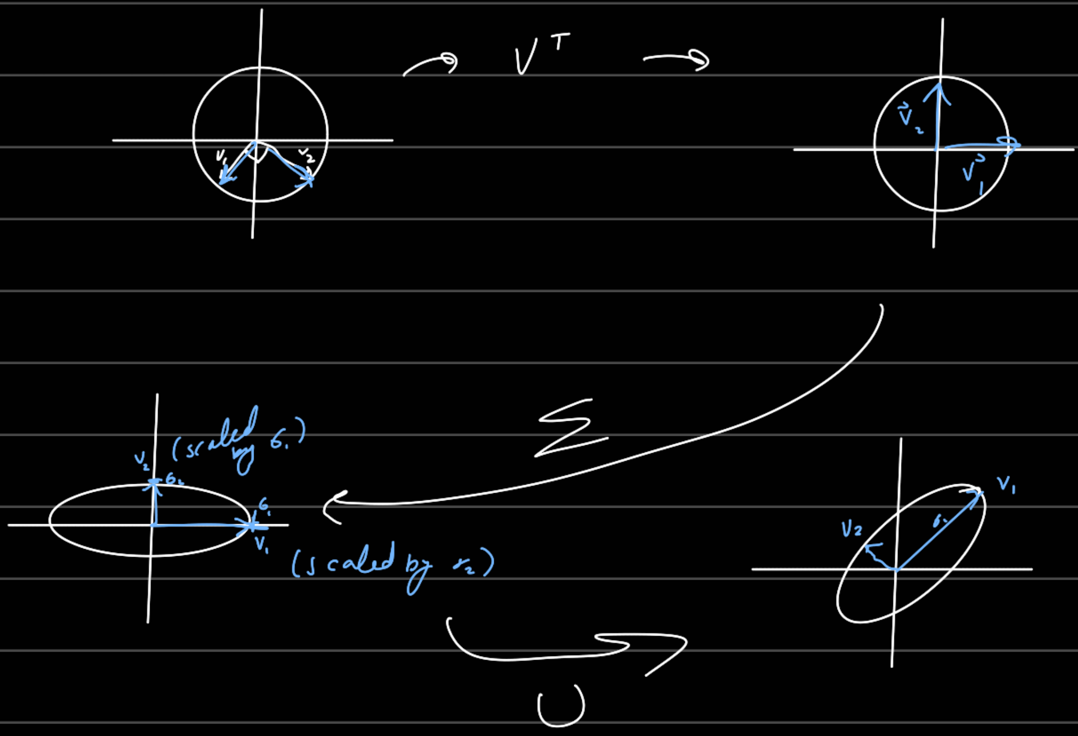
↳ general



$\vec{v}_1, \vec{v}_2 \to$ eigenvectors of A^TA

$\vec{v}_1 \perp \vec{v}_2$

$A\vec{v}_1 \perp A\vec{v}_2$



$\Rightarrow \vec{x} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2$

Low Rank Approximation

↳ about approximating matrices

Matrix A

↳ operators

↳ block of data

Matrix Norms

① Frobenius norm $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{\text{trace}(A^TA)}$

↳ trace of the matrix A^TA

↳ invariant to orthonormal transformations

$$\|UA\|_F = \|AU\|_F = \|A\|_F$$

∘ Spectral Norm (ℓ_2-norm)

$$\underbrace{\|A\|_2}_{\text{matrix}} = \max_{\|\vec{x}\|_2=1} \underbrace{\|Ax\|_2}_{\text{vector norm}}$$

Ralleigh coeff

$$= \max_{\|\vec{x}\|_2=1} x^T A^T A \vec{x} = \sqrt{\lambda_{max}(A^TA)}$$

$$= \sigma_{max} A$$

Eckart-Young-Mirsky Thm

$A \in \mathbb{R}^{n \times n}$ $A = U\Sigma V^T$ $A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^T$

$\sigma_1 \geq \sigma_2 \cdots \geq \sigma_3$ $A = \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T$

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

② $\underset{\substack{B \in \mathbb{R}^{m \times n} \\ \text{rank}(B) = k}}{\text{argmin}} \|A - B\|_2 = A_k$

Proof of Eckert Young ②

Consider $\|A - A_k\|_2 = \|\sum_{i=k+1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T\|$

$= \sigma_{k+1}(A)$ ↳ following from $\|A\|_2 = \sigma_{max}(A)$

Now, to show for all rank(B) = k matrices that:

$\|A - B\|_2 \geq \sigma_{k+1}$

$\forall \vec{\omega}$:

$\|A - B\|_2 \geq \|(A-B)\vec{\omega}\|_2$ s.t. $\|\vec{\omega}\|_2 = 1$

↳ want to choose a $\vec{\omega}$ that gives you a nice

Choose $\vec{\omega} \in N(B)$

↓

$\|(A-B)\vec{\omega}\|_2 = \|A\vec{\omega}\|_2$

↳ want this to look like σ_{k+1}

Consider $V_{k+1} = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_{k+1} \end{bmatrix}$ (where $A = U \Sigma V^T$)

$\text{Rank}(V_{k+1}) = k+1$ (bc $\vec{v}_i$ // ⊥ //)

$\dim(N(B)) = n - k$

$(n-k) + (k+1) = n + 1 > n$

↳ V_{k+1} & B can't be completely disjoint

$\Rightarrow$ there must be at least 1 dimension of overlap in Range (V_{k+1}) & Null(B)

$\vec{\omega} \in N(B)$
$\in R(V_{k+1})$

$$\vec{\omega} = V\vec{\alpha} = \begin{bmatrix} V_{k+1} & V_{rest} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \alpha_1 \vec{v}_1 + \cdots + \vec{\alpha}_{k+1} \vec{v}_{k+1}$$

want $\|\vec{\omega}\| = 1 \Rightarrow$ choose $\sum_{i=1}^{k+1} \alpha_i^2 = 1$

$$\begin{aligned}
\|(A - B)\vec{\omega}\|_2^2 &= \|A\vec{\omega}\|_2^2 \\
&= \|U \Sigma V^T \cdot V\vec{\alpha}\|_2^2 \\
&= \|U \Sigma \vec{\alpha}\|_2^2 \\
&= \|\Sigma \vec{\alpha}\|_2^2 \\
&= \alpha_1^2 \sigma_1^2 + \cdots + \alpha_{k+1}^2 \sigma_{k+1}^2 \\
&\geq \sigma_{k+1}^2 \left(\sum \alpha_i^2\right) \\
&= \sigma_{k+1}^2
\end{aligned}$$