EECS 16B

Spring 2022

Lecture 20

3/31/2022

LECTURE 20 : SVD

Given $A \in \mathbb{R}^{m \times n}$ with rank $= r$, "outer product" form of SVD:

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \dots + \sigma_r \vec{u}_r \vec{v}_r^T = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \qquad (1)$$

where:

- $\vec{u}_1, \dots, \vec{u}_r$ are orthonormal in $\mathbb{R}^m$
- $\vec{v}_1, \dots, \vec{v}_r$ " " in $\mathbb{R}^n$
- $\sigma_1 \geq \dots \geq \sigma_r$ are positive, real numbers.

Equivalently,

$$A = \underbrace{\underbrace{[\vec{u}_1 \cdots \vec{u}_r]}_{U_r}}_{m \times r,\ \text{orthonormal columns}} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \underbrace{\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix}}_{V_r^T} \qquad -- (2)$$

V_r : $n \times r$, orthonormal columns

look at (j,k) element:

$$A(j,k) = \sum_{i=1}^{r} \underbrace{U_r(j,i)}_{(\vec{u}_i)_j} \sigma_i \underbrace{V_r^T(i,k)}_{= V_r(k,i) = (\vec{v}_i)_k}$$

from matrix multiplication

since ith column of U_r is $\vec{u}_i$

$$= \sum_{i=1}^{r} \sigma_i (\vec{u}_i)_j (\vec{v}_i)_k$$

this matches (j,k) entry in (1); therefore (1) and (2) are equivalent

Example 1: $A = \begin{bmatrix} 4 & 4 \\ 3 & 3 \end{bmatrix}$ $r = 1$

$$= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$= \underbrace{5 \cdot \sqrt{2}}_{\sigma_1} \underbrace{\left(\frac{1}{5} \begin{bmatrix} 4 \\ 3 \end{bmatrix} \right)}_{\vec{u}_1} \underbrace{\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \right)}_{\vec{v}_1^T}$$

If we change the signs of $\vec{u}_1$ and $\vec{v}_1$ simultaneously, length and outer product are the same $\Rightarrow$ another SVD $\Rightarrow$ SVD <u>not</u> unique!

How do we find a SVD in general? Will give a procedure to find SVD of $A \in \mathbb{R}^{m \times n}$ from the evalues/evectors of $A^T A \in \mathbb{R}^{n \times n}$.

Some facts about $A^T A$:

<u>Claim 1:</u> $A^T A$ has real evalues/evectors $(\lambda_i, \vec{v}_i)$, $i = 1, \dots n$.

Proof from Spectral Thm because $A^T A$ is symmetric:

$$(A^T A)^T = A^T (A^T)^T = A^T A.$$

<u>Claim 2:</u> Evalues of $A^T A$ are nonnegative.

Proof: $$A^T A \vec{v}_i = \lambda_i \vec{v}_i$$

Multiply both sides from left by $\vec{v}_i^T$:

$$\underbrace{\underbrace{\vec{v}_i^T A^T A \vec{v}_i}_{(A\vec{v}_i)^T (A\vec{v}_i)}}_{\|A\vec{v}_i\|^2} = \lambda_i \underbrace{\vec{v}_i^T \vec{v}_i}_{\|\vec{v}_i\|^2} \Rightarrow \lambda_i = \frac{\|A\vec{v}_i\|^2}{\|\vec{v}_i\|^2} \geq 0$$

Claim 3: If $\text{rank}(A)=r$, then r evalues of A^TA are strictly positive.

Proof: First note that null spaces of A and A^TA are the same: $N(A)=N(A^TA)$ because

(i) $N(A) \subseteq N(A^TA)$

$$A\vec{v}=0 \Rightarrow A^T\underbrace{A\vec{v}}_{=0}=0$$

(ii) $N(A^TA)\subseteq N(A)$

$$A^TA\vec{v}=0 \xRightarrow{\substack{\text{multiply}\\ \text{from left}\\ \text{by } \vec{v}^T}} \underbrace{\vec{v}^TA^T}_{(A\vec{v})^T}A\vec{v}=(A\vec{v})^T(A\vec{v})=\|A\vec{v}\|^2=0$$

$$\Downarrow$$

$$\|A\vec{v}\|=0$$

$$\Downarrow$$

$$A\vec{v}=0$$

From (i) and (ii): $N(A)=N(A^TA)$.

$$\text{Rank}(A)=r \Rightarrow \dim N(A) \underset{\substack{\uparrow\\ \text{1GA,}\\ \text{Rank/Nullity}\\ \text{Thm}}}{=} n-r = \dim N(A^TA)$$

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

A^TA is $n\times n$, has $n-r$ dim. null space.
Elements of $N(A^TA)$ are evectors of A^TA (for eigenvalue $=0$, repeated $n-r$ times).
Claim 2 said n, ≥ 0 evalues. $n-r$ of them are $=0$.
$\Rightarrow$ remaining r evalues of A^TA are strictly positive.

SVD Procedure for $A\in\mathbb{R}^{m\times n}$, rank $=r$ (using A^TA):

Step 1) Find orthogonal matrix V diagonalizing A^TA:
(exists by Spectral Thm)

$$V^T(A^TA)V=\left[\begin{array}{ccc|ccc} \lambda_1 & & & & & 0\\ & \ddots & & & & \\ & & \lambda_r & & & \\ \hline 0 & & 0 & 0 & & \\ & & & & \ddots & \\ & & 0 & & & 0\end{array}\right]\Big\}\,n-r$$

$\lambda_1,\ldots,\lambda_r$ are evalues of A^TA

and make sure to put evalues in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0.$$

Step 2) For $i = 1, \ldots r$, pick ith column $\vec{v}_i$ of V (which is evector of A^TA for evalue λ_i) and let

$$\sigma_i = \sqrt{\lambda_i}, \quad \vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i.$$

singular value of A are square roots of eigenvalues of A^TA

Example 2: $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$ $r = 2$

$$A^TA = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}$$

$$\lambda I - A^TA = \begin{bmatrix} \lambda - 25 & -7 \\ -7 & \lambda - 25 \end{bmatrix} \rightarrow \det = (\lambda - 25)^2 - 7^2 = 0$$

$$\lambda - 25 = \mp 7$$

$$\lambda_{1,2} = 25 \mp 7$$

$\lambda_1 = 32$ $\quad \lambda_1 I - A^TA = \begin{bmatrix} 7 & -7 \\ -7 & 7 \end{bmatrix}$ $\quad \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = 18$ $\quad \lambda_2 I - A^TA = \begin{bmatrix} -7 & -7 \\ -7 & -7 \end{bmatrix}$ $\quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Step 2:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{32} = 4\sqrt{2} \qquad \vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{4\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\sigma_2 = \sqrt{\lambda_2} = \sqrt{18} = 3\sqrt{2}$ $\quad \vec{u}_2 = \frac{1}{\sigma_2} A \vec{v}_2 = \frac{1}{3\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -6 \end{bmatrix}$

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

$$A = \underbrace{4\sqrt{2}}_{\sigma_1} \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{u}_1} \underbrace{\frac{1}{\sqrt{2}}[1 \ 1]}_{\vec{v}_1^T} + \underbrace{3\sqrt{2}}_{\sigma_2} \underbrace{\begin{bmatrix} 0 \\ -1 \end{bmatrix}}_{\vec{u}_2} \underbrace{\frac{1}{\sqrt{2}}[1 \ -1]}_{\vec{v}_2^T}$$

Justification of the Procedure: Do $\vec{v}_i, \vec{u}_i, \sigma_i$ resulting from procedure satisfy the following?

- $\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T = A$
- $\vec{u}_1, \dots \vec{u}_r$ orthonormal
- $\vec{v}_1, \dots \vec{v}_r$ " $\rightarrow$ true by Step 1 of procedure
- $\sigma_1, \dots \sigma_r$ real and positive $\rightarrow$ true by Step 2: $\sigma_i = \sqrt{\lambda_i}$

$$\vec{u}_i^T \vec{u}_j = \left(\frac{1}{\sigma_i} A \vec{v}_i\right)^T \left(\frac{1}{\sigma_j} A \vec{v}_j\right) = \frac{1}{\sigma_i \sigma_j} \vec{v}_i^T \underbrace{A^T A \vec{v}_j}_{= \lambda_j \vec{v}_j}$$

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

$$\frac{\lambda_i}{\sigma_i^2} = \frac{\lambda_i}{\lambda_i} = 1$$

$$\vec{u}_i^T \vec{u}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

<u>First bullet:</u> $\displaystyle\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \overset{?}{=} A$

Note: $\displaystyle \vec{u}_i \vec{v}_i^T = \overbrace{\left(\frac{1}{\sigma_i} A \vec{v}_i\right)}^{\vec{u}_i \text{ by Step 2}} \vec{v}_i^T = \frac{1}{\sigma_i} A \vec{v}_i \vec{v}_i^T$

Therefore,

$$\sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T = \sum_{i=1}^{r} A \vec{v}_i \vec{v}_i^T = A \sum_{i=1}^{r} \vec{v}_i \vec{v}_i^T \quad --(3)$$

Recall V orthogonal $V^T V = \underbrace{V V^T}_{\underbrace{[\vec{v}_1 \cdots \vec{v}_n]\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}}_{\sum_{i=1}^{n} \vec{v}_i \vec{v}_i^T = I}} = I$

$$\Rightarrow \boxed{A \sum_{i=1}^{n} \vec{v}_i \vec{v}_i^T = A \quad --(4)}$$

Note, for $i > r$, $\vec{v}_i$ evector of $A^T A$ corresponding to a zero evalue:

$$\underset{i=r+1,\ldots,n}{A \vec{v}_i = 0} \quad \xRightarrow{\text{multiply from right by } \vec{v}_i^T} \quad \underset{i=r+1,\ldots,n}{A \vec{v}_i \vec{v}_i^T = 0}$$

$$\Rightarrow \sum_{i=r+1}^{n} A \vec{v}_i \vec{v}_i^T = 0 \Rightarrow \boxed{A \sum_{i=r+1}^{n} \vec{v}_i \vec{v}_i^T = 0 \quad -(5)}$$

Subtract (5) from (4): $\quad A \sum_{i=1}^{r} \vec{v}_i \vec{v}_i^T = A \quad --(6)$

and substitute (6) in (3): $\quad \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T = A.$