11/22/2021

Last time:

→ Symmetric positive (semi)definite matrices.

Thm: A symmetric matrix is ~~[illegible]~~ positive (semi) definite $\Leftrightarrow$ all of its eigenvalues are (nonnegative) positive.

→ Intro to SVD.

SVD: $\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^T$, $\underline{\underline{A}} \in \mathbb{R}^{m \times n}$.

* columns of $\underline{\underline{U}}$ are eigenvectors of $\underline{\underline{A}}\,\underline{\underline{A}}^T$.

* columns of $\underline{\underline{V}}$ are eigenvectors of $\underline{\underline{A}}^T\underline{\underline{A}}$.

* $\underline{\underline{\Sigma}}$ contains square roots of nonzero eigenvalues of $\underline{\underline{A}}^T\underline{\underline{A}}$ or $\underline{\underline{A}}\,\underline{\underline{A}}^T$.

$\left\{ \text{We saw that if } (\underline{x}, \lambda) \text{ is an eigenpair of } \underline{\underline{A}}\,\underline{\underline{A}}^T \text{ with } \lambda \neq 0, \text{ then } (\underline{\underline{A}}^T\underline{x}, \lambda) \text{ is an eigenpair of } \underline{\underline{A}}^T\underline{\underline{A}}. \right.$

Construction of SVD

Take $(\underline{u}_i, \sigma_i^2)$, $\sigma_i \neq 0$, an eigenpair of $\underline{\underline{A}}\,\underline{\underline{A}}^T$.

Then, $(\underline{v}_i = \sigma_i^{-1} \underline{\underline{A}}^T \underline{u}_i, \sigma_i^2)$ is an eigenpair of $\underline{\underline{A}}^T \underline{\underline{A}}$.

Note:

$$\underline{\underline{A}}\,\underline{v}_i = \underline{\underline{A}}\,(\sigma_i^{-1} \underline{\underline{A}}^T \underline{u}_i) = \sigma_i^{-1} \underline{\underline{A}}\,\underline{\underline{A}}^T \underline{u}_i$$

$$= \sigma_i^{-1} \sigma_i^2 \underline{u}_i = \sigma_i \underline{u}_i.$$

$$\rightsquigarrow \quad \underline{\underline{A}}\,\underline{v}_i = \sigma_i \underline{u}_i, \quad i = 1, \dots, r$$

where r = # of nonzero singular values (repetitions allowed)

Written in matrix form, we have:

$$\underbrace{\overset{\overset{n}{\longleftrightarrow}}{\underline{\underline{A}}}}_{m} \underbrace{\begin{pmatrix} | & & | \\ \underline{v}_1 & \cdots & \underline{v}_r \\ | & & | \end{pmatrix}}_{n \times r} = \underbrace{\begin{pmatrix} | & & | \\ \underline{u}_1 & \cdots & \underline{u}_r \\ | & & | \end{pmatrix}}_{m \times r} \underbrace{\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}}_{r \times r}.$$

If $r < \min(m,n)$, then we can add vectors

$$\underline{v}_{r+1}, \ldots, \underline{v}_n \quad \text{and}$$
$$\underline{u}_{r+1}, \ldots, \underline{u}_m$$

which correspond to zero eigenvalues of $\underline{\underline{A}}^T\underline{\underline{A}}$ and $\underline{\underline{A}}\,\underline{\underline{A}}^T$ respectively.

$$\underline{\underline{A}} \underbrace{\begin{pmatrix} | & & | & | & & | \\ \underline{v}_1 & \cdots & \underline{v}_r & \underline{v}_{r+1} & \cdots & \underline{v}_n \\ | & & | & | & & | \end{pmatrix}}_{\substack{n \times n \\ = \underline{\underline{V}}}} = \underbrace{\begin{pmatrix} | & & | & | & & | \\ \underline{u}_1 & \cdots & \underline{u}_r & \underline{u}_{r+1} & \cdots & \underline{u}_m \\ | & & | & | & & | \end{pmatrix}}_{\substack{m \times m \\ = \underline{\underline{U}}.}} \underbrace{\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots \ \sigma_r & \\ 0 & & 0 \end{pmatrix}}_{\substack{m \times n \\ = \underline{\underline{\Sigma}}.}}$$

Note that $\underline{\underline{V}}\,\underline{\underline{V}}^T = \underline{\underline{I}}$, so we have

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^T.$$

$$= \begin{pmatrix} | & & | \\ \underline{u}_1 & \cdots & \underline{u}_m \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots \ \sigma_r & \\ 0 & & 0 \end{pmatrix} \begin{pmatrix} — \underline{v}_1 — \\ — \underline{v}_2 — \\ \vdots \\ — \underline{v}_n — \end{pmatrix}.$$

$\underline{\text{Ex}}$: $\quad \underline{\underline{A}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}$.

We need to diagonalize $\underline{\underline{A}}\,\underline{\underline{A}}^T$ and $\underline{\underline{A}}^T\underline{\underline{A}}$:

$$\underline{\underline{A}}^T\underline{\underline{A}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

$$\underline{\underline{A}}\,\underline{\underline{A}}^T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}.$$

$\underline{\text{Note}}$: $\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, 6 \right)$ eigenpair of $\underline{\underline{A}}^T\underline{\underline{A}}$.

$\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, 6 \right)$ eigenpair of $\underline{\underline{A}}\,\underline{\underline{A}}^T$.

We can write:

$$\underline{\underline{A}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \sqrt{6}\, \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}}_{\substack{\text{Normalized} \\ \text{eigenvector} \\ \text{of } \underline{\underline{A}}\,\underline{\underline{A}}^T}} \underbrace{\left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}} \right)^T}_{\substack{\text{Normalized} \\ \text{eigenvector} \\ \text{of } \underline{\underline{A}}^T\underline{\underline{A}}.}}$$

We also see that:

$\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, 0 \right)$, $\left(\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, 0 \right)$ are the remaining eigenpairs of $\underline{A}^T \underline{A}$.

and $\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, 0 \right)$ is the remaining eigenpair for $\underline{A}\,\underline{A}^T$.

So, we can these eigenpairs to get the SVD:

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{=\underline{\underline{A}}} = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}}_{=\underline{\underline{U}}} \underbrace{\begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\underline{\underline{\Sigma}}} \underbrace{\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{pmatrix}}_{\underline{\underline{V}}^T}.$$

<u>Observe:</u> rank $\underline{A}$ = dimension $R(\underline{A}) = \underline{1}$ in the above example, and also the # of nonzero singular values $= \underline{1}$.

In general: rank $\underline{A} = \begin{pmatrix} \text{\# of nonzero} \\ \text{singular values} \end{pmatrix}$

i.e. # of nonzero eigenvalues of $\underline{A}^T\underline{A}$ or $\underline{A}\,\underline{A}^T$.

Step by Step procedure for computing SVD

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

①. Compute $\lambda_1 \geq \dots \geq \lambda_m$, the eigenvalues of $\underline{A}\,\underline{A}^T$, and corresponding eigenvectors $\{\underline{u}_1, \dots, \underline{u}_m\}$.

let r = largest index such that $\lambda_r > 0$.

set $\underline{\underline{U}} = \begin{pmatrix} | & & | \\ \underline{u}_1 & \cdots & \underline{u}_m \\ | & & | \end{pmatrix}$.

② compute $\sigma_j = \sqrt{\lambda_j}$, $j = 1, \dots, r$.

③ Compute $\underline{v}_j = \sigma_j^{-1} \underline{\underline{A}}^T \underline{u}_j$, $j = 1, \dots, r$.

compute $\{\underline{v}_{r+1}, \dots, \underline{v}_n\}$ as a orthonormal basis for $N(\underline{A}) = N(\underline{A}^T\underline{A})$. set $\underline{V} = \begin{pmatrix} | & & | \\ \underline{v}_1 & \cdots & \underline{v}_n \\ | & & | \end{pmatrix}$.

④ $\underline{\underline{A}} = \underline{\underline{U}}\, \underline{\underline{\Sigma}}\, \underline{\underline{V}}^T$.