SVD

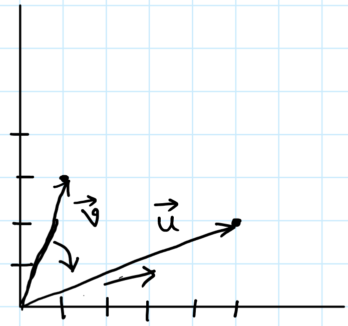
14 August 2021 10:08 AM

$A = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ $\vec{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ $A\vec{v} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \vec{u}$

$A \rightarrow$ rot. + scaling

$\searrow$ SVD $\nearrow$

$\sqrt{10}$ $\sqrt{\frac{29}{10}}$ $\sqrt{29}$



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

$U \rightarrow$ rotation
$\Sigma \rightarrow$ scaling
$V^* \rightarrow$ another rotation

$* \rightarrow$ complex conjugate transpose / Hermitian transpose

$V \rightarrow V^T \; (i \rightarrow -i) \rightarrow V^*$

$U, V \rightarrow$ orthogonal (property of rot. matrices) real matrices, same as V^T.

- $\boxed{U^T \cdot U = U \cdot U^T = I}$

$\Sigma \rightarrow$ diagonal $\begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_n \end{bmatrix}$

$\sigma_j \rightarrow$ singular values

$U \rightarrow \vec{u}_1, \vec{u}_2 \ldots$
$V \rightarrow \vec{v}_1, \vec{v}_2$

$U = \begin{bmatrix} | & | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \cdots & \vec{u}_n \\ | & | & | & & | \end{bmatrix}$ $V = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$

$A \rightarrow$ any dimension $\rightarrow \underline{m \times n}$ SVD ALWAYS WORKS!

Mathematically, $U \rightarrow (m \times n)$, $\Sigma \rightarrow (n \times n)$, $V^T \rightarrow (n \times n)$ if $m > n$
$\underline{U} \rightarrow (m \times m)$, $\Sigma \rightarrow (m \times m)$, $\underline{V^T} \rightarrow (m \times n)$ if $m < n$

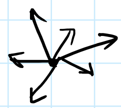
$m > n$

In $U \rightarrow$ add $(m-n)$ columns $\Rightarrow U \rightarrow (m \times m)$

n columns
m rows

$\begin{bmatrix} 1 & 2 & 3 & 4 & \ldots & n \\ 2 & & & & & \\ 3 & & & & & \\ 4 & & & & & \\ 5 & & & & & \\ \vdots & & & & & \\ m & & & & & \end{bmatrix}$

$\mathbb{R}^m$ m m-dim. orthogonal vec.



n m-dim vectors $\vec{u}_j = m$-dim.

$U \rightarrow (m \times m)$

under-determined, $\rightarrow$ add $(m-n)$ more orth. vectors. to make the m vectors cover $\mathbb{R}^m$.

$\Sigma \rightarrow$ add $(m-n)$ rows of zeros. $\Rightarrow \Sigma \rightarrow (m \times n)$ [same as A].

$V^T \rightarrow (n \times n)$

STD.

$A\,(\underline{m \times n}) = U\,(\underline{m \times m}) \times \Sigma\,(\underline{m \times n}) \times V^T\,(\underline{n \times n})$ ~~for m > n.~~

$A\,(m \times n) = \underline{U\,(m \times m)} \times \underline{\Sigma\,(m \times n)} \times \underline{V^T\,(n \times n)}$ ~~for m < n~~

Calculate ?

$\Sigma \rightarrow$ diagonal $\Sigma^T = \Sigma$

$A = U \Sigma V^T$ $A^T = (U \Sigma V^T)^T = V \Sigma^T U^T = \underline{V \Sigma U^T}$

$AA^T = U\Sigma \underbrace{V^T \cdot V}_{I}\Sigma U^T = U\Sigma \cdot \Sigma U^T = U\Sigma^2 U^T$ → multiply both sides with U.

$AA^T \cdot U = U\Sigma^2 \underbrace{U^T U}_{I} \Rightarrow (AA^T) U = U\Sigma^2 \Rightarrow (AA^T)\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \dots \vec{u}_m \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \dots \vec{u}_m \end{bmatrix}\Sigma^2$

$$\Sigma^2 = \begin{bmatrix} \sigma_1^2 & & 0 \\ 0 & \sigma_2^2 & \\ 0 & & \ddots \\ 0 & & \sigma_n^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\boxed{(AA^T)\vec{u}_1 = \sigma_1^2 \vec{u}_1}$ → EIGENVALUE PROBLEM!

number

$\boxed{(AA^T)\vec{u}_j = \sigma_j^2 \vec{u}_j}$ → I KNOW HOW TO SOLVE!

find all eigenvalues of AA^T and their corresponding eigenvectors.

↳ KNOWN

eigenvalues $\lambda_j = \sigma_j^2 \Rightarrow \sigma_j = \sqrt{\lambda_j}$ & have all $\sigma_j \Rightarrow \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \ \sigma_n \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

eigenvectors $\vec{u}_j \Rightarrow U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \dots & \vec{u}_m \end{bmatrix}$

Used AA^T for Σ and U. $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \sigma_n$ how to write Σ.

Now, $A^T A = (U\Sigma V^T)^T \cdot U\Sigma V^T = V\Sigma \underbrace{U^T \cdot U}_{I}\Sigma V^T = V\Sigma^2 V^T$ → multiply by V both sides

$\boxed{A^T A V = V\Sigma^2}$ → same as before, find Σ and V ✓

↳ obtained again

NOW, $\boxed{A = U\Sigma V^T}$ → serve as basis for ML.

- → image compression
- → categorize faces and recognize (face recognition)
- → classify handwritten digits.
- → PCA

$\Sigma \to (m \times n)$.

$\sigma_1 \ \sigma_2 \dots \cancel{\sigma_n}$

first k.

$\sigma_1 \to \sigma_k$ now Σ is $k \times k$.

from U $(m \times m)$ take first k columns ⇒ U $(m \times k)$

from V $(n \times n)$ take first k columns ⇒ V $(n \times k)$.

$A = U\Sigma V^T$ $\quad \underset{(m\times k)}{U} \ \underset{(k\times k)}{\Sigma} \ \underset{(k\times n)}{V^T}$ → reconstruct A $(m \times n)$

$(m \times n)$

BUT WITH LESS DATA! → basis for image compression.