Singular value Decomposition (SVD)

Applications :-

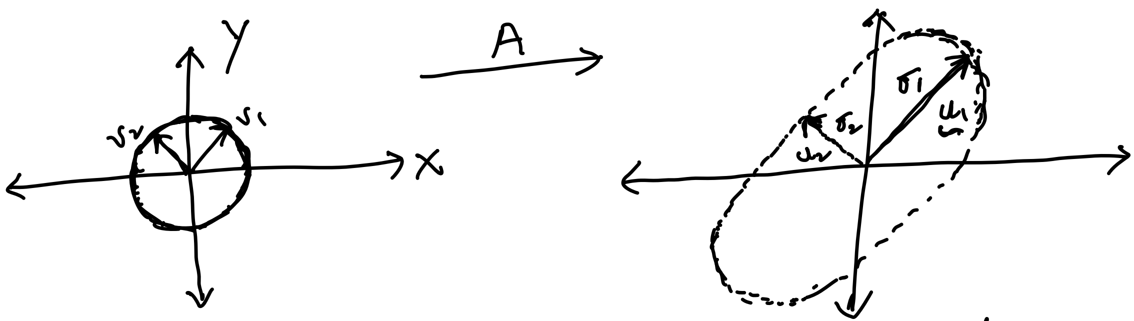
- Searching closest related images
- Image compression (Reducing image size)
- Image recovery
- Principal component analysis (Find most representative/dominant features)
- Solution of Least squares problems
- Find null-space, column-space
- Computing Matrix norm

Geometric Intution:-

Let us consider for simplicity we are working in $\mathbb{R}^2$.

→ Action of any matrix $\underset{\sim}{A} \in \mathbb{R}^{2\times 2}$ on a unit circle → Gives an ellipse

$\underset{\sim}{A}$ is a non-singular and hence full rank



The image of unit circle under the action of matrix $\underset{\sim}{A}$ is ellipse

$$\underset{\sim}{A}\,\underset{\sim}{S} \longrightarrow \text{ellipse}$$

$\underset{\sim}{u_1}, \underset{\sim}{u_2}$ are the principal semi-axes of my ellipse with lengths σ_1, σ_2

$\underset{\sim}{v_1}, \underset{\sim}{v_2}$ are pre-image vectors generating images $\underset{\sim}{u_1}, \underset{\sim}{u_2}$ the axes of the ellipse.

$$\underset{\sim}{A}\,\underset{\sim}{v_1} = \sigma_1 \underset{\sim}{u_1}$$

$$\underset{\sim}{A}\,\underset{\sim}{v_2} = \sigma_2 \underset{\sim}{u_2}$$

[There exists vectors v_1, v_2 which are orthogonal whose images $\underset{\sim}{u_1}, \underset{\sim}{u_2}$ are along the directions of ellipse axes]

$$\underset{\sim}{A}\underbrace{\begin{bmatrix}\underset{\sim}{v_1} & \underset{\sim}{v_2}\end{bmatrix}}_{V} = \underbrace{\begin{bmatrix}\underset{\sim}{u_1} & \underset{\sim}{u_2}\end{bmatrix}}\underbrace{\begin{bmatrix}\sigma_1 & 0 \\ 0 & \sigma_2\end{bmatrix}}$$

$$\underset{2\times 2}{\underset{\sim}{A}}\;\underset{2\times 2}{\underset{\sim}{V}} = \underset{2\times 2}{\underset{\sim}{U}}\;\underset{2\times 2}{\underset{\sim}{\Sigma}}$$

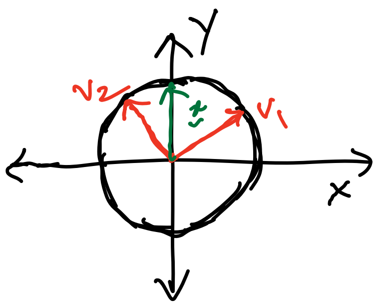
$$\underset{\sim}{A}\,\underset{\sim}{V}\,\underset{\sim}{V}^T = \underset{\sim}{U}\,\underset{\sim}{\Sigma}\,\underset{\sim}{V}^T \implies \underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{\Sigma}\,\underset{\sim}{V}^T$$

Consider action of any matrix on a

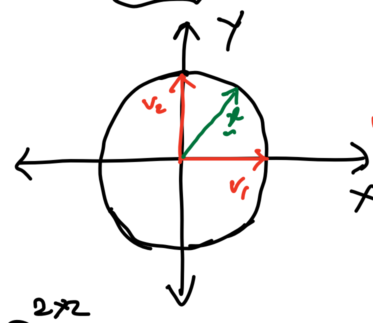
unit circle S -via- sequence of steps

$$\underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{\Sigma}\,\underset{\sim}{V}^T$$

$$\underset{\sim}{A}\,\underset{\sim}{S} = \boxed{\underset{\sim}{U}\,\boxed{\underset{\sim}{\Sigma}\,\boxed{\underset{\sim}{V}^T S}}}$$



Rotation V^T $\longrightarrow$



Action of any matrix $\underset{\sim}{A} \in \mathbb{R}^{2\times 2}$ on a unit circle S

Rotation

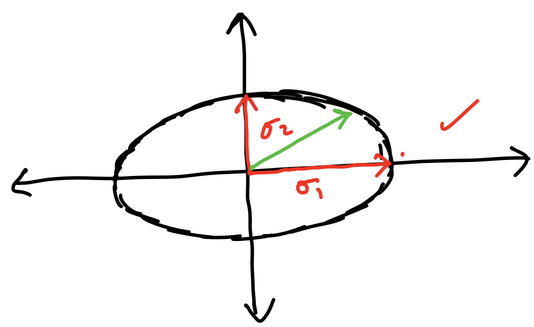
$\downarrow$

Stretching to Ellipse

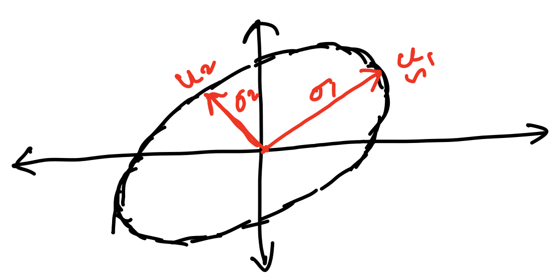
$\downarrow$

Rotation of Ellipse

$\downarrow$ Σ Stretching $\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$ $\sigma_1 > \sigma_2$



$\downarrow$ Rotation U



This process can be generalized to higher dimension ≥ 2

Let $\underset{\sim}{S}$ be unit hypersphere in $\mathbb{R}^n$.
The action of $\underset{\sim}{A} \in \mathbb{R}^{m \times n}$ $(m \geq n)$ and full rank on $\underset{\sim}{S}$ is a hyperellipse in $\mathbb{R}^m$ with the following properties :-

(i) $\underset{\sim}{A}\,\underset{\sim}{v}_j = \sigma_j\,\underset{\sim}{u}_j$

$\sigma_1, \sigma_2, \sigma_3 \ldots \sigma_n$ are lengths of the principal semi-axes of $\underset{\sim}{A}\,\underset{\sim}{S}$ (hyperellipse in $\mathbb{R}^m$) and these σ_j's are called singular values of $\underset{\sim}{A}$. By convention they are ordered in descending order

$$\sigma_1 \geq \sigma_2 \cdots \geq \sigma_n$$

(ii) The set of unit vectors $\{u_1, u_2, \ldots u_n\}$ are directions of principal semi axes of our hyperellipse and these are called left singular vectors and are orthonormal vectors.

$\therefore$ $\sigma_j \underset{\sim}{u}_j$ is the j^{th} largest principal semiaxes of our hyperellipse $\underset{\sim}{A}S$

(iii) The set of unit vectors $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots v_n\}$ which also form an orthonormal set are the pre-images of $\underset{\sim}{u}_i$ and are called right singular vectors $\underset{\sim}{A}$

$$\underset{\sim}{A}\,\underset{\sim}{v}_j = \sigma_j \underset{\sim}{u}_j \quad \text{— (1)}$$

$$[\underset{\sim}{A}]_{m\times n} \underbrace{\begin{bmatrix} | & | & | & & | \\ \underset{\sim}{v}_1 & \underset{\sim}{v}_2 & \underset{\sim}{v}_3 & \cdots & \underset{\sim}{v}_n \\ | & | & | & & | \end{bmatrix}}_{V \;\; n\times n}$$

$$= \underbrace{\begin{bmatrix} | & | & | & & | \\ \underset{\sim}{u}_1 & \underset{\sim}{u}_2 & \underset{\sim}{u}_3 & \cdots & \underset{\sim}{u}_n \\ | & | & | & & | \end{bmatrix}_{m\times n}}_{\hat{U}} \underbrace{\begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{bmatrix}_{n\times n}}_{\hat{\Sigma}}$$

$$\underset{\sim}{A}\,\underset{\sim}{V} = \hat{\underset{\sim}{U}}\,\underset{\sim}{\Sigma}$$

$\underset{\sim}{A} \in \mathbb{R}^{m\times n}$ $\quad \hat{\underset{\sim}{U}} \in \mathbb{R}^{m\times n}$

$\underset{\sim}{V} \in \mathbb{R}^{n\times n}$ $\quad \hat{\Sigma} \in \mathbb{R}^{n\times n}$

Since $\underset{\sim}{V}$ is orthogonal matrix

$$A V V^T = \hat{U} \hat{\Sigma} V^T$$

$$\boxed{A = \hat{U} \hat{\Sigma} V^T}$$

$\downarrow$

Reduced Singular Value Decompositi