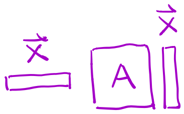
• <u>Def</u> $A \in \mathbb{R}^{n\times n}$ is called positive definite if

① $A = A^T$ (thus all eigenvalues are real)



② $\vec{x}^T A \vec{x} > 0$ for any nonzero $\vec{x} \in \mathbb{R}^n$.

Example: $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is positive definite because

$$(x \ y)\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (x \ y)\begin{pmatrix} 2x-y \\ -x+2y \end{pmatrix}$$

$$= x(2x-y) + y(-x+2y)$$

$$= 2x^2 - 2xy + 2y^2$$

$$= x^2 + (x^2 - 2xy + y^2) + y^2$$

$$= x^2 + (x-y)^2 + y^2 > 0 \text{ if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

• <u>Def</u> $A \in \mathbb{R}^{n\times n}$ is called positive semi-definite if

① $A = A^T$ (thus all eigenvalues are real)

② $\vec{x}^T A \vec{x} \geq 0$ for any nonzero $\vec{x} \in \mathbb{R}^n$.

• If $A = VDV^{-1}$ where D is diagonal, then

$A = VDV^{-1}$ is also called <u>eigen-decomposition</u>.

because $\begin{cases} \text{① diagonal entries in } D \text{ are eigenvalues} \\ \text{② cols of } V \text{ are eigen-vectors.} \end{cases}$

• Theorem A real symmetric $A \in \mathbb{R}^{n\times n}$ is positive definite if and only if all eigenvalues of A are positive.

Proof: $A = A^T \Rightarrow \begin{cases} \text{eigenvalues } \lambda_i \text{ are real.} \\ A \text{ has } n \text{ orthonormal eigenvectors} \end{cases}$

(Apply Gram-Schmit to eigenvectors for each eigenspace)

Let $V \in \mathbb{R}^{n \times n}$ consist of n orthonormal eigenvectors

Then $A = VDV^{-1}$ and $V^{-1} = V^T$

$= VDV^T$ $\quad V^T V = I$

For any $\vec{x} \in \mathbb{R}^n$,

$$\vec{x}^T A \vec{x} = \vec{x}^T V D V^{-1} \vec{x}$$

$(AB)^T = B^T A^T$

$$= \vec{x}^T V D V^T \vec{x}$$

$\vec{y} = \boxed{V^T}\boxed{\vec{x}}$

$(V^T \vec{x})^T = \vec{x}^T V$

$$= (V^T \vec{x})^T D (V^T \vec{x})$$

Change of variable $\vec{y} = V^T \vec{x}$

$$= \vec{y}^T D \vec{y}$$

$$= \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2$$

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 > 0, \text{ for any } y_1, y_2, y_3 \Leftrightarrow \begin{cases} \lambda_1 > 0 \\ \lambda_2 > 0 \\ \lambda_3 > 0 \end{cases}$$

Example: $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

Find 3 orthonormal eigenvectors of A.

Sol: $|A-\lambda I| = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix}$

$$= -\lambda(\lambda^2-1) - (-\lambda-1) + (1+\lambda)$$

$$= -\lambda(\lambda+1)(\lambda-1) + (\lambda+1) + (\lambda+1)$$

$$= (\lambda+1)[-\lambda(\lambda-1)+1+1]$$

$$= (\lambda+1)[-\lambda^2+\lambda+2]$$

$$= -(\lambda+1)^2(\lambda-2)$$

$\lambda_1 = -1$, $\lambda_2 = 2$

① Plug in $\lambda_1 = -1$ into $(A-\lambda I)\vec{v} = \vec{0}$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left.\begin{array}{l} v_2 = s \\ v_3 = t \end{array}\right\} \Rightarrow v_1 = -s-t \Rightarrow \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ t \end{pmatrix}$$

$$= s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$\Rightarrow$ Eigen-Space is Span $\left\{ \overset{\vec{u}_1}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}, \overset{\vec{u}_2}{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}} \right\}$

But two basis eigen-vectors are not orthogonal.

Apply Gram-Schmidt Procedure:

$$\vec{v}_1 = \vec{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{v}_2 = \vec{u}_2 - \frac{\langle \vec{v}_1, \vec{u}_2 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1$$

$$= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Verify $A\vec{v}_2 = (-1)\vec{v}_2$

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \end{bmatrix}$$

② Plug in $\lambda = 2$ into $(A - \lambda I)\vec{v} = \vec{0}$

to find $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

So we get orthogonal eigenvectors

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and orthonormal eigenvectors:

$$\vec{w}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}$$

$$\vec{w}_2 = \frac{1}{\sqrt{\frac{3}{2}}}\begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}\sqrt{\frac{2}{3}} \\ -\frac{1}{2}\sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix}$$

$$\vec{w}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{1}{3}} \\ \sqrt{\frac{1}{3}} \\ \sqrt{\frac{1}{3}} \end{pmatrix}$$

Use $V = [\vec{w}_1 \;\; \vec{w}_2 \;\; \vec{w}_3]$ $\quad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Then $A = VDV^{-1}$

and $V^{-1} = V^T$

$V \; V^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$\Uparrow$

$V^TV = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow V^{-1} = V^T \Rightarrow VV^T = I$

Chapter 7 Singular Value Decomposition (SVD)

- SVD is defined for any matrix, but we only focus on square ones.
- $A \in \mathbb{R}^{n \times n}$ may not have a diagonalization like $A = VDV^{-1}$, but A always has Singular Value Decomposition $A = U\Sigma V^T$

where $\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}$ is diagonal with $\sigma_i \geq 0$

and $\begin{cases} U \text{ has orthonormal cols thus } U^T = U^{-1} \\ V \text{ has orthonormal cols thus } V^T = V^{-1} \end{cases}$

- σ_i are called singular values of A

  Cols of U are left singular vectors of A

  Cols of V are right singular vectors of A.

Remark: ① eigenvalues of A can be complex but singular values of A are always real non-negative.

② Left singular vectors are always orthonormal.

③ Right singular vectors are always orthonormal.

- SVD is defined / computed as the following:

  ① A^TA is real symmetric and positive semi-definite

  $(A^TA)^T = A^T(A^T)^T = A^TA$ $\quad \vec{x}^TA^TA\vec{x} = (A\vec{x})^T(A\vec{x}) = \|A\vec{x}\|^2 \geq 0$

So its eigenvalues $\lambda_i(A^TA) \geq 0$.

The singular value of A, denoted as $\sigma_i(A)$, is computed/defined by $\sigma_i(A) = \sqrt{\lambda_i(A^TA)}$.

It can also be computed by $\sqrt{\lambda(AA^T)}$, which is always the same even if $AA^T \neq A^TA$.

② Cols of U are orthonormal eigenvectors of AA^T.

③ Cols of V are orthonormal eigenvectors of A^TA.

④ Match order : $A\vec{v}_i = \sigma_i \vec{u}_i \Leftrightarrow AV = U\Sigma \Leftrightarrow A = U\Sigma V^T$.

Example: $A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix}$. Find its SVD.

① $A^TA = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\lambda(A^TA) = 9, 4, 0$

$\Rightarrow \sigma(A) = 3, 2, 0$

② Corresponding eigenvectors of A^TA are

$\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ already orthogonal.

If one eigen-space is 2-dim, need Gram-Schmit.

Orthonormal eigenvectors $\left(\frac{\vec{v}_i}{\|\vec{v}_i\|}\right)$:

$$\frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\\0\end{pmatrix},\quad \frac{2}{\sqrt{5}}\begin{pmatrix}-\frac{1}{2}\\1\\0\end{pmatrix},\quad \begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$V=\begin{pmatrix}2/\sqrt{5} & -1/\sqrt{5} & 0\\ 1/\sqrt{5} & 2/\sqrt{5} & 0\\ 0 & 0 & 1\end{pmatrix}$$

③ $$AA^T=\begin{pmatrix}4 & 4 & 0\\ 4 & 5 & -2\\ 0 & -2 & 4\end{pmatrix}$$

$$\lambda(AA^T)=9,\ 4,\ 0$$

eigenvectors: $$\begin{pmatrix}-4\\-5\\2\end{pmatrix},\quad \begin{pmatrix}1\\0\\2\end{pmatrix},\quad \begin{pmatrix}-2\\2\\1\end{pmatrix}$$

orthonormal eigen-vectors $$\frac{1}{3\sqrt{5}}\begin{pmatrix}4\\5\\-2\end{pmatrix},\quad \frac{1}{\sqrt{5}}\begin{pmatrix}-1\\0\\-2\end{pmatrix},\quad \frac{1}{3}\begin{pmatrix}-2\\2\\1\end{pmatrix}$$

$$U=\begin{pmatrix}\frac{4}{3\sqrt{5}} & \frac{-1}{\sqrt{5}} & -\frac{2}{3}\\ \frac{5}{3\sqrt{5}} & 0 & \frac{2}{3}\\ \frac{-2}{3\sqrt{5}} & \frac{-2}{\sqrt{5}} & \frac{1}{3}\end{pmatrix}$$

$$A=\underset{\substack{\downarrow\\ AA^T}}{U}\begin{pmatrix}3 & & \\ & 2 & \\ & & 0\end{pmatrix}\underset{\substack{\downarrow\\ A^TA}}{V^T}$$

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

$$\Rightarrow A^TA=(U\Sigma V^T)^T\,U\Sigma V^T$$

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

$$=V\Sigma^2V^T$$

It is a convention to order σ_i : $\sigma_1\geq\sigma_2\geq\sigma_3\geq\dots$

- We will not discuss why A is equal to $U\Sigma V^T$
- Instead, assume $A = U\Sigma V^T$, then

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

$V^T V = I$

$$= U\Sigma\Sigma U^T$$
$$= U\begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} U^T$$
$$= U\begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} U^{-1}$$

This is the eigen-decomposition of AA^T

$$\Rightarrow \begin{cases} \text{eigenvalues of } AA^T \text{ are } \sigma_i^2 \\ \text{eigenvectors of } AA^T \text{ are cols of } U \end{cases}$$

Similary,
$$A^TA = (U\Sigma V^T)^T U\Sigma V^T$$
$$= V\Sigma U^T U \Sigma V^T$$
$$= V\begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{pmatrix} V^{-1}$$

$$\Rightarrow \begin{cases} \text{eigenvalues of } A^TA \text{ are } \sigma_i^2 \\ \text{eigenvectors of } A^TA \text{ are cols of } V. \end{cases}$$

Ex (True or false):

$A \in \mathbb{R}^{n\times n}$ is real symmetric $\iff$ there are V and diagonal D s.t. $A = VDV^T$

True: "$\Rightarrow$" Let V consist of real orthonormal eigenvectors

$$\Rightarrow A = VDV^{-1} = VDV^T$$

"$\Leftarrow$" $A = VDV^T \Rightarrow A^T = A$.

$(VDV^T)^T = (V^T)^T D^T V^T = VDV^T$

Remark: If A is positive semi-definite, then there are V and $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ with $\lambda_i \geq 0$ s.t.

$$A = VDV^{-1} = VDV^T,$$

which is also the SVD of A.

$$\begin{aligned} A^TA &= (VDV^T)^T VDV^T \\ &= VDV^TVDV^T \\ &= V\begin{pmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{pmatrix} V^T \end{aligned}$$

$\Rightarrow \begin{cases} \sigma_i(A) = \sqrt{\lambda_i^2} = \lambda_i \\ \text{left singular vectors of } A \text{ are cols of } V. \end{cases}$

$$AA^T = V\begin{pmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{pmatrix} V^T$$

$\Rightarrow \begin{cases} \sigma_i(A) = \sqrt{\lambda_i^2} = \lambda_i \\ \text{right singular vectors of } A \text{ are cols of } V. \end{cases}$

- $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}$

- $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, the Frobenius norm of A is defined as

  $$\|A\|_F = \sqrt{\sum_{j=1}^{3} \sum_{i=1}^{3} a_{ij}^2}$$

- Useful formula: $\|A\|_F = \sqrt{tr(A^TA)} = \sqrt{tr(AA^T)}$

  $$tr\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9$$

- $A = U\Sigma V^T \Rightarrow A^TA = (U\Sigma V^T)^T U\Sigma V^T$

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

  $$= V\Sigma^2 V^T$$

  $$= V\begin{pmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2 \end{pmatrix} V^T$$

  $$tr(A^TA) = tr(V\Sigma^2V^T) = tr(\Sigma^2 V^TV)$$

  $$tr(ABC) = tr(BCA)$$

  $$= tr(\Sigma^2) = \sigma_1^2 + \sigma_2^2 + \sigma_3^2$$

  $$\Rightarrow \|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$$