

- Eigenvalue Decomposition:

If $A \in \mathbb{R}^{n \times n}$ is diagonalizable, then

$$A = V D V^{-1} = V \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{bmatrix} V^{-1}$$

$$\Leftrightarrow A V = V D$$

$$\Leftrightarrow A [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] = [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$$

$$\Leftrightarrow [A\vec{v}_1 \ A\vec{v}_2 \ A\vec{v}_3] = [d_1\vec{v}_1 \ d_2\vec{v}_2 \ d_3\vec{v}_3]$$

$$\Leftrightarrow A\vec{v}_i = d_i\vec{v}_i$$

$$\Leftrightarrow \vec{v}_i \text{ are independent eigenvectors with eigen-values } d_i$$

Remark: d_i & $\vec{v}_i$ might be complex

- The following are always diagonalizable with orthonormal eigenvectors:

① complex normal matrix ($AA^* = A^*A, A \in \mathbb{C}^{n \times n}$) $A^* = \bar{A}^T$

② real normal matrix ($AA^T = A^T A, A \in \mathbb{R}^{n \times n}$)

③ real symmetric ($A^T = A$):

- { 1) all eigenvalues are real
- { 2) thus eigenvectors are real

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

orthonormal cols in $V \Rightarrow V^T V = I$

④ real skew-symmetric ($A^T = -A$):

- { 1) all eigenvalues are purely imaginary
- { 2) thus eigenvectors are complex

$$\Rightarrow A = V D V^{-1} = V D V^*$$

complex orthonormal cols in $V \Rightarrow V^* V = I$

In general, A is normal $\Rightarrow A = V D V^*$.

• Singular Value Decomposition (SVD)

For any $A \in \mathbb{R}^{n \times n}$, there are $U, \Sigma, V \in \mathbb{R}^{n \times n}$ st.

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

① $\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}$, $\sigma_i \geq 0$ are singular values

$$\sigma(A) = \sqrt{\lambda(A^T A)} = \sqrt{\lambda(A A^T)}$$

$\hookrightarrow$ eigenvalue

② Cols of U are orthonormal: left singular vectors

$$U^T U = U U^T = I$$

③ Cols of V are orthonormal: right singular vectors

$$V^T V = V V^T = I$$

$$\textcircled{4} \quad A = U \Sigma V^T$$

$$\Leftrightarrow A = U \Sigma V^{-1}$$

$$\Leftrightarrow A V = U \Sigma$$

$$\Leftrightarrow A [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3] \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}$$

$$\Leftrightarrow [A \vec{v}_1 \ A \vec{v}_2 \ A \vec{v}_3] = [\sigma_1 \vec{u}_1 \ \sigma_2 \vec{u}_2 \ \sigma_3 \vec{u}_3]$$

$$\Leftrightarrow A \vec{v}_i = \sigma_i \vec{u}_i \quad \text{The order must match}$$

$$⑤ \quad A = U \Sigma V^T \Rightarrow A A^T = U \Sigma V^T V \Sigma U^T = U \Sigma^2 U^T$$

Review: A is real positive definite if ① $A = A^T$ ② $\vec{x}^T A \vec{x} > 0, \forall \vec{x} \in \mathbb{R}^n, \vec{x} \neq \vec{0}$
 $= \square$

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

$$\Rightarrow \begin{cases} \vec{u}_i \text{ are orthonormal eigenvectors of } A A^T. \\ \sigma_i^2 \text{ are eigenvalues of } A A^T. \end{cases}$$

$$A = U \Sigma V^T \Rightarrow A^T A = V \Sigma U^T U \Sigma V^T = V \Sigma^2 V^T$$

$$\Rightarrow \begin{cases} \vec{v}_i \text{ are orthonormal eigenvectors of } A^T A. \\ \sigma_i^2 \text{ are eigenvalues of } A^T A. \end{cases}$$

And we have pick right order and signs of $\vec{u}_i, \vec{v}_i$ st.

$$A \vec{v}_i = \sigma_i \vec{u}_i$$

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

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

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

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

② Corresponding eigenvectors of $A^T A$ are

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

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

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

$$V = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} & 0 \\ 1/\sqrt{5} & 2/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{matrix} A\vec{v}_i = \sigma_i \vec{u}_i \\ \Leftrightarrow \vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}, \sigma_i \neq 0 \end{matrix}$$

$$\textcircled{2} \quad A A^T = \begin{pmatrix} 4 & 4 & 0 \\ 4 & 5 & -2 \\ 0 & -2 & 4 \end{pmatrix}$$

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

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

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

The sign is to be determined by $A\vec{v}_i = \sigma_i \vec{u}_i$

$$A\vec{v}_2 = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix} \frac{2}{\sqrt{5}} \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} = \underset{\frac{1}{2}}{\sigma_2} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix}$$

$$A\vec{v}_3 = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \underset{0}{\sigma_3} \cdot (\pm) \frac{1}{3} \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$$

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

$$\Rightarrow A = U \begin{pmatrix} 3 & & \\ & 2 & \\ & & 0 \end{pmatrix} V^T \Leftrightarrow A\vec{v}_i = \sigma_i \vec{u}_i$$

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

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

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

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

- $A = U \Sigma V^T \Rightarrow A^T A = (U \Sigma V^T)^T U \Sigma V^T$
 $= V \Sigma U^T U \Sigma V^T$
 $= V \Sigma^2 V^T$
 $= V \begin{pmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2 \end{pmatrix} V^T$

$$\text{tr}(A^T A) = \text{tr}(V \Sigma^2 V^T) = \text{tr}(\Sigma^2 V^T V)$$

$$\text{tr}(ABC) = \text{tr}(BCA)$$

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

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

- Theorem: The number of nonzero singular values is the rank of A

Proof: $A = U \Sigma V^T = U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T, \sigma_1, \sigma_2 > 0.$

$$A = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3] \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T$$

$$= [\sigma_1 \vec{u}_1 \ \sigma_2 \vec{u}_2 \ \vec{0}] V^T$$

$$= [\sigma_1 \vec{u}_1 \ \sigma_2 \vec{u}_2 \ \vec{0}] \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$\Rightarrow$ cols of A are linear combinations of $\sigma_1 \vec{u}_1$ and $\sigma_2 \vec{u}_2$

$$\Rightarrow \text{Rank}(A) = 2.$$

- Low-Rank Approximation Theorem

For $A \in \mathbb{R}^{n \times n}$, let $A = U \Sigma V^T$ be its SVD

with singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_n \geq 0$

For any $r = 1, 2, \dots, n$, let $\tilde{\Sigma} = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}.$

Then $\|A - U\tilde{\Sigma}V^T\|_F \leq \|A - B\|_F$ for any rank- r matrix B .

This means $U\tilde{\Sigma}V$ is the best rank- r approximation to A .

• Example: $r=1$, $U\tilde{\Sigma}V^T = [\vec{u}_1 \ \vec{u}_2 \ \vec{u}_3] \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T$

$$= [\sigma_1 \vec{u}_1 \ \vec{0} \ \vec{0}] \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vec{v}_3^T \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^T$$



• Principal Component Analysis (PCA):

If $A = U\Sigma V^T$, then $\sigma_i \vec{u}_i$ are principal components.

• Low-rank compression:

if we replace A by $\tilde{A} = U\tilde{\Sigma}V^T$

then the error is $\|A - \tilde{A}\|_F = \|U(\Sigma - \tilde{\Sigma})V^T\|_F$

$$= \left\| U \begin{pmatrix} 0 & \dots & 0 & \sigma_{r+1} & \dots & \sigma_n \end{pmatrix} V^T \right\|_F$$

$$= \sqrt{\sigma_{r+1}^2 + \dots + \sigma_n^2}$$

For $\tilde{A}$, we only need to store

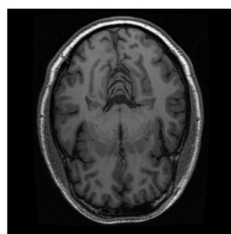
① r scalars: $\sigma_1, \dots, \sigma_r$

② $2r$ vectors: $\vec{u}_1, \dots, \vec{u}_r$
 $\vec{v}_1, \dots, \vec{v}_r$

$$r \times n: U \tilde{\Sigma} V^T = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \begin{array}{|c|} \hline \tilde{\Sigma} \\ \hline \end{array} \begin{array}{|c|c|c|} \hline \text{ } & \text{ } & \text{ } \\ \hline \end{array}$$

Example: $r=1$, $U \tilde{\Sigma} V^T = \sigma \vec{u} \vec{v}^T = \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \begin{array}{|c|} \hline \text{ } \\ \hline \end{array} \begin{array}{|c|c|c|} \hline \text{ } & \text{ } & \text{ } \\ \hline \end{array}$

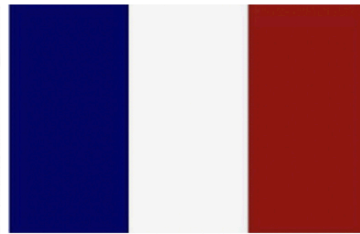
$$\begin{bmatrix} 2 & 2 & 3 & 3 & 4 & 4 \\ 2 & 2 & 3 & 3 & 4 & 4 \\ 2 & 2 & 3 & 3 & 4 & 4 \\ 2 & 2 & 3 & 3 & 4 & 4 \end{bmatrix}$$



(a) MRI Image



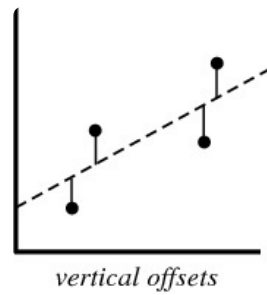
(b) Finger Print



(c) France Flag

- Pictures are naturally stored as matrices: a 1024×1024 black and white picture is stored as a 1024×1024 matrix.
- For many pictures, if we store it as a matrix, it is a waste: MRI image has information only inside the curve boundary; finger prints have repeated/similar structure.
- Compression of data: we do not need to store the whole $n \times n$ pixels for a $n \times n$ picture.
- There are many ways to compress data: low-rank approximation is one of them.
- Example: France Flag is a rank-1 matrix. SVD is $A = U \Sigma V^* = \mathbf{u}_1 \sigma_1 \mathbf{v}_1^*$. So we only need to store two vectors $\mathbf{u}_1$ and $\mathbf{v}_1$ and σ_1 .

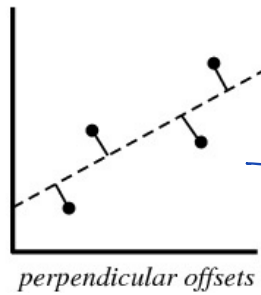
Line Fitting



vertical offsets



Computed by least square



perpendicular offsets

→ Computed by PCA/SVD of a nonsquare matrix

• SVD for $A \in \mathbb{R}^{m \times n}$, $m > n$.

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

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}$$

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$$

The eigen-decomposition: $\boxed{A^T A} = \boxed{V} \boxed{\begin{matrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{matrix}} \boxed{V^T}$

$$\boxed{A A^T} = \boxed{U} \boxed{\begin{matrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 & 0 \\ & & & \ddots & 0 \end{matrix}} \boxed{U^T}$$

$$= \boxed{U} \boxed{\begin{matrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{matrix}} \boxed{U^T}$$

- Compact SVD : for $A \in \mathbb{R}^{n \times n}$, if $\text{rank}(A) = k < n$,

then SVD is $U \Sigma V^T$

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

$$= \boxed{\tilde{U}} \boxed{\tilde{\Sigma}} \boxed{\tilde{V}^T}$$

$= \tilde{U} \tilde{\Sigma} \tilde{V}^T$ is called compact SVD.

- Compact SVD, low-rank approximation and PCA are used for almost all matrix applications.

- σ_i contain lots of information:

even for a nonsquare $A \in \mathbb{R}^{m \times n}$, $m > n$

$$\textcircled{1} \sqrt{\sigma_1^2 + \dots + \sigma_n^2} = \|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

$$\textcircled{2} \text{rank}(A) = \# \text{ of nonzero } \sigma_i$$

$\textcircled{2}$ If $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$, then

$$\sigma_1 = \max_{\vec{x} \in \mathbb{R}^n} \frac{\|A\vec{x}\|}{\|\vec{x}\|}$$

↪ largest change of length from $\vec{x}$ to $A\vec{x}$.