→ LAST TIME

→ PCA (PRINCIPAL COMPONENT ANALYSIS)

1. INTUITIVE IDEAS OF WHAT IT ACHIEVES
2. COVARIANCE MATRICES
3. PCA: EIGENDECOMPOSITION OF CO-VARIANCE MATRICES

→ TODAY

4. PCA: CONNECTION WITH SVDs
5. COMPUTING SVDs VIA PCA/EIGENDECOMPOSITION

— K-MEANS CLUSTERING

1. THE PROBLEM
2. ILLUSTRATION ON A 1-D EXAMPLE
3. THE 1-D ALGORITHM
4. DISTORTION AND ITS MINIMIZATION BY K-MEANS
5. THE K-MEANS ALGORITHM FOR HIGHER-D DATA
6. "FAILURE" OF K-MEANS
7. EXAMPLES AND APPLICATIONS

PCA: SUMMARY: (FROM LAST TIME)

— EIGENDECOMPOSE S (COV. MAT.)

→ $\vec{p}_i$ is the i^{th} "most important" (spread) PRINCIPAL COMPONENT

→ λ_i is the variance along $\vec{p}_i$ of the data

— $\vec{p}_1$ is the direction of <u>greatest visual</u> spread.

→ $\vec{p}_2, \vec{p}_3$, etc. $2^{nd}, 3^{rd}, \ldots$, greatest spread.

---

4. USING THE SVD FOR PCA

$A, \tilde{A} \in \mathbb{R}^{n \times m}$

— SVD OF $\underbrace{\tilde{A}}_{\text{mean centered data}} = \underbrace{U}_{n \times n} \underbrace{\Sigma}_{n \times m} \underbrace{V^T}_{m \times m}$

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

$$\rightarrow S = \frac{1}{n} \tilde{A}^T \tilde{A} = \frac{1}{n} V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T = V \frac{\Sigma^T \Sigma}{n} V^T$$

$$\rightarrow \Sigma^T \Sigma = \begin{array}{c} \xleftrightarrow{\;m\;} \\ \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix} \end{array} \updownarrow m = \begin{bmatrix} \sigma_1 & & & \vdots \\ & \ddots & & \vdots \\ & & \sigma_m & \vdots \end{bmatrix} \underbrace{\begin{array}{c} \xleftrightarrow{\;m\;} \\ \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ \cdots & \cdots & \cdots \end{bmatrix} \end{array}}_{\Sigma} \updownarrow n$$

$$\rightarrow \frac{\Sigma^T \Sigma}{n} = \begin{array}{c} \xleftrightarrow{\;m\;} \\ \begin{bmatrix} \sigma_1^2/n & & \\ & \ddots & \\ & & \frac{\sigma_m^2}{n} \end{bmatrix} \end{array}$$

$$S = V \begin{bmatrix} \sigma_1^2/n & & \\ & \ddots & \\ & & \frac{\sigma_m^2}{n} \end{bmatrix} V^T \; ;$$

→ $VV^T = I \iff V^{-1} = V^T$

→ $\sigma_i > 0$

$S = P \Lambda P^T \quad \rightarrow PP^T = I \checkmark \quad ; \quad \checkmark \lambda_i \geq 0$

$$\boxed{P = V \; ; \quad \lambda_i = \frac{\sigma_i^2}{n}} \longleftarrow \text{PCA FROM SVD OF } \tilde{A}$$

$\longrightarrow$ S NOT NEEDED.

---

5. SVD FROM PCA (EIGENDECOMPOSITION)

$\longrightarrow$ GIVEN SOME A

FIND THESE

$\longrightarrow$ GOAL: FIND ITS SVD: $A = U \Sigma V^T$

NOT MEAN-CENTERED

$\longrightarrow$ FORM $T = A^T A$

NOT DIVIDED BY n

$\longrightarrow$ EIGENDECOMPOSE $T = P \Lambda P^T$

$\rightarrow P^T = P^{-1} \rightarrow$ P is unitary $\checkmark$

$\rightarrow \lambda_i \geq 0 \checkmark$

$\longrightarrow$ CLAIM:

$$\boxed{\begin{aligned} &\rightarrow V = P \\ &\rightarrow \sigma_i = \sqrt{\lambda_i} \quad , \quad i = 1, \ldots, m \end{aligned}}$$

$\longrightarrow$ PROOF: $A = U \Sigma V^T$

$$T = A^T A = V \Sigma^T \underbrace{U^T U}_{} \Sigma V^T = V \underbrace{\Sigma^T \Sigma}_{\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix}} V^T$$

$$T = V \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix} V^T \longleftarrow \text{DON'T KNOW } V,\ \sigma_i \text{ yet}$$

$\rightarrow V V^T = I$; $\rightarrow \sigma_i \geq 0$

$$T = P \Lambda P^T \quad \text{(from EIGENDECOMPOSITION).}$$

$\rightarrow P P^T = I$; $\rightarrow \lambda_i \geq 0$

$$\Longrightarrow \boxed{V = P \ ; \ \sigma_i = \sqrt{\lambda_i}}$$

— DON'T HAVE U yet: HOW TO FIND IT?

$$\rightarrow V = \begin{array}{c} m \\ \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{v}_1 \cdots & \vec{v}_i \cdots & \vec{v}_m \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \end{array} m$$

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

$$V \vec{v}_i = \begin{bmatrix} \leftarrow \vec{v}_1^T \rightarrow \\ \leftarrow \vec{v}_i^T \rightarrow \\ \vdots \\ \leftarrow \vec{v}_m^T \rightarrow \end{bmatrix} \vec{v}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{th}$$

$$\rightarrow A \vec{v}_i = \underbrace{U \Sigma}_{\downarrow} \underbrace{V^T \vec{v}_i}_{\downarrow} = \sigma_i \vec{u}_i$$

$$n \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \sigma_1 \vec{u}_1 & \cdots \sigma_i \vec{u}_i \cdots & \sigma_m \vec{u}_m \\ \downarrow & \downarrow & \downarrow \end{bmatrix}_{m} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_i \vec{u}_i$$

$$\overset{U}{n \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{u}_1 \cdots & \vec{u}_i \cdots & \vec{u}_n \\ \downarrow & \downarrow & \downarrow \end{bmatrix}_{n}} \overset{m}{\begin{bmatrix} \sigma_1 & & \\ & \sigma_i \ddots & \\ & & \sigma_m \\ - & - & - \end{bmatrix}} n = \overset{m}{\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \sigma_1 \vec{u}_1 \cdots & \sigma_i \vec{u}_i \cdots & \sigma_m \vec{u}_m \\ \downarrow & \downarrow & \downarrow \end{bmatrix}} n$$

ASIDE

$$\begin{bmatrix} | & & | \\ \vec{p}_1 & \cdots & \vec{p}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1 \vec{p}_1 + \alpha_2 \vec{p}_2 + \cdots + \alpha_n \vec{p}_n$$

$$\rightarrow A \vec{v}_i = \sigma_i \vec{u}_i, \quad i = 1, \ldots, m, \quad \text{NOT for } i = m+1, \ldots, n$$

→ $\vec{u}_i = \dfrac{A\vec{v}_i}{\sigma_i}$, $i = 1,\dots,m$ AND if $\sigma_i \neq 0$

→ FINDING $\vec{u}_{m+1},\dots,\vec{u}_n$, and $\vec{u}_i$ if $\sigma_i = 0$

→ → $A\vec{v}_i = \underbrace{U\Sigma}\,V^T$

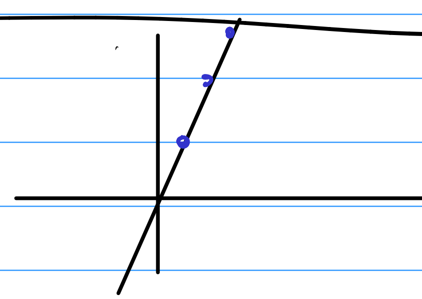
$$\underbrace{\begin{bmatrix} \sigma_1\vec{u}_1 & \cdots & \sigma_i\vec{u}_i & \cdots & \sigma_m\vec{u}_m \end{bmatrix}}_{U\Sigma}$$

→ $\vec{u}_{m+1},\dots,\vec{u}_n$ DON'T MATTER

→ $\sigma_i = 0$, $\vec{u}_i$ DOESN'T MATTER

→ SVD $= \underset{\substack{n\times n\\ \text{orthonormal}}}{U}\Sigma V^T$

→ FIND $\vec{u}_{m+1},\dots,\vec{u}_n$ to COMPLETE $\vec{u}_1,\dots,\vec{u}_m$ as AN ORTHONORMAL BASIS



EXAMPLE

→ $A_1 (= A) = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix}$

$$\rightarrow \; T = A^T A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 6 & 10 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix} = \begin{bmatrix} 35 & 70 \\ 70 & 140 \end{bmatrix}$$

$$\rightarrow T = 35 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

EIGENDEC.

$\downarrow$

$$\rightarrow T = P \Lambda P^T$$

$$= \underbrace{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 175 & \\ & 0 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \frac{1}{\sqrt{5}}}_{P^T}$$

$\downarrow$

CHOSEN ORTHONORMAL

$$\Longrightarrow \left(\text{SVD}: \quad V = P \; ; \; \sigma_i = \sqrt{\lambda_i} \right)$$

$$V = \frac{1}{\sqrt{5}} \begin{array}{c} \begin{array}{cc} \vec{v}_1 & \vec{v}_2 \end{array} \\ \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \end{array} = \quad \Sigma = \begin{bmatrix} \sqrt{175} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{array}{l} \sigma_1 = \sqrt{175} \\ \sigma_2 = 0 \end{array}$$

— Find U

$$\vec{u}_1 = \frac{A\vec{v}_1}{\sigma_1} ; \qquad A\vec{v}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \frac{1}{\sqrt{5}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 \\ 15 \\ 25 \end{bmatrix} = \sqrt{5} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

$$\vec{u}_1 = \frac{\sqrt{5}}{\sqrt{175}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \frac{1}{\sqrt{35}} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \vec{u}_1$$

$$A\vec{v}_2 = \overset{0}{\sigma_2}\vec{u}_2 \Rightarrow A\vec{v}_2 = 0$$

$\rightarrow$ WANT: $\vec{u}_2 \perp \vec{u}_1 \iff \vec{u}_2^T \vec{u}_1 = 0$

$\perp$: "perpendicular/orthogonal to"

$$\frac{1}{\sqrt{35}}\begin{bmatrix}1\\3\\5\end{bmatrix} = \vec{u}_1$$

$\rightarrow$ Say $\vec{u}_2 = \frac{1}{\sqrt{a^2+b^2+c^2}}\begin{bmatrix}a\\b\\c\end{bmatrix}$; $\vec{u}_2^T \vec{u}_1 = \frac{1}{\sqrt{35}}(a+3b+5c) = 0$

$\rightarrow (a+3b+5c) = 0$

$\rightarrow$ PICK ANY 2, GET THE THIRD

$$a=2,\ b=1 \Rightarrow c = \frac{-(a+3b)}{5} = -1$$

$$\vec{u}_2 = \begin{bmatrix}2\\1\\-1\end{bmatrix} \times \frac{1}{\sqrt{6}}$$

$\vec{u}_3 \perp \vec{u}_1$, and $\vec{u}_3 \perp \vec{u}_2$ and $\|\vec{u}_3\| = 1$

$\rightarrow$ Say $\vec{u}_3 = \frac{1}{\sqrt{a^2+b^2+c^2}}\begin{bmatrix}a\\b\\c\end{bmatrix}$; $\left.\begin{array}{l}\vec{u}_3^T\vec{u}_1 \Rightarrow (a+3b+5c) = 0\\ u_3^T\vec{u}_2 \Rightarrow (2a+b-c) = 0\end{array}\right\}$ 2 equns.

$$\vec{u}_3 = \begin{bmatrix}8\\-11\\5\end{bmatrix} \times \frac{1}{\sqrt{210}}$$

$\approx U$

$$U = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \vec{u}_1 & \vec{u}_2 & u_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{35}} & \frac{2}{\sqrt{6}} & \frac{8}{\sqrt{210}} \\ \frac{3}{\sqrt{35}} & \frac{1}{\sqrt{6}} & \frac{-11}{\sqrt{210}} \\ \frac{5}{\sqrt{35}} & \frac{-1}{\sqrt{6}} & \frac{5}{\sqrt{210}} \end{bmatrix}$$

$$V = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \quad ; \quad \Sigma = \begin{bmatrix} \sqrt{175} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$T = A^T A$$

- T is symmetric: $T^T = T = A^T A$

→ Say B is symmetric.

→ 1. B HAS ALL <u>REAL</u> EIGENVALUES

→ $B\vec{p} = \lambda \vec{p}$

→ a: use symmetry: $\vec{p}^T B = \lambda \vec{p}^T$ (1)

→ b: use realness of B: $\overline{B\vec{p}} = \overline{\lambda \vec{p}} \Rightarrow B\overline{\vec{p}} = \overline{\lambda}\, \overline{\vec{p}}$ (2)

$$\vec{p}^T \underbrace{B\overline{\vec{p}}} = \lambda \underbrace{\vec{p}^T\overline{\vec{p}}}_{\|\vec{p}\|^2} = \lambda \|\vec{p}\|^2$$

$$\vec{p}^T \overline{\lambda}\, \overline{\vec{p}} = \overline{\lambda} \|\vec{p}\|^2$$

$$\lambda \cancel{\|\vec{p}\|^2} = \overline{\lambda} \cancel{\|\vec{p}\|^2}$$

$\lambda = \overline{\lambda} \Rightarrow \lambda$ real.