→ MIDTERM FEEDBACK: ACTIONS

→ HAND-WRITTEN LECTURES (LOW ON SLIDES)

→ MORE EXAMPLES TO ILLUSTRATE CONCEPTS

→ SIMPLER AND MORE CLEAR FLOW

THINGS YOU SHOULD KNOW AT THIS POINT

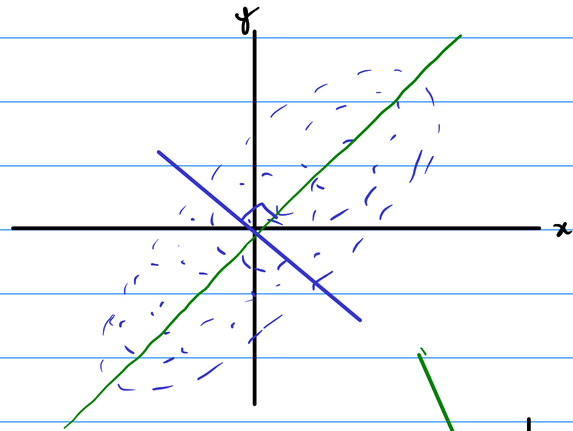
- (EIGENDECOMPOSITION)
- WHAT AN SVD IS (NOT HOW TO COMPUTE IT)
- SOME USES OF SVDs (LOW RANK APPROXIMATION)
  - IMAGE DECOMPOSITION AND COMPRESSION
  - DATA ANALYSIS: ROW/COL FEATURES AND USES FOR CLUSTERING

→ PCA (PRINCIPAL COMPONENT ANALYSIS)

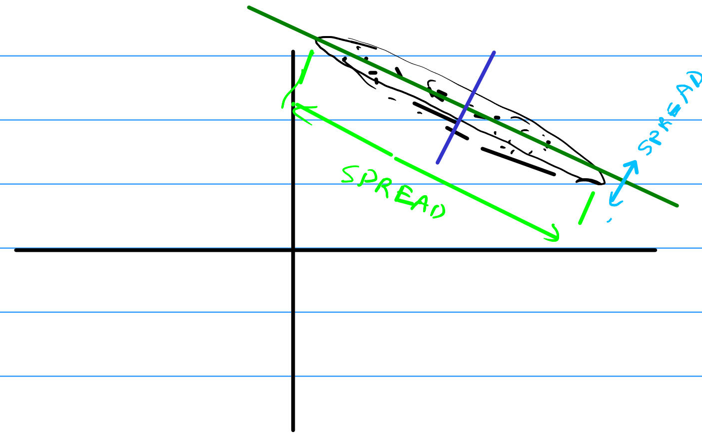
1. INTUITIVE IDEAS OF WHAT IT ACHIEVES
2. COVARIANCE MATRICES
3. PCA: EIGENDECOMPOSITION OF CO-VARIANCE MATRICES
4. PCA: CONNECTION WITH SVDs
5. COMPUTING SVDs VIA PCA/EIGENDECOMPOSITION

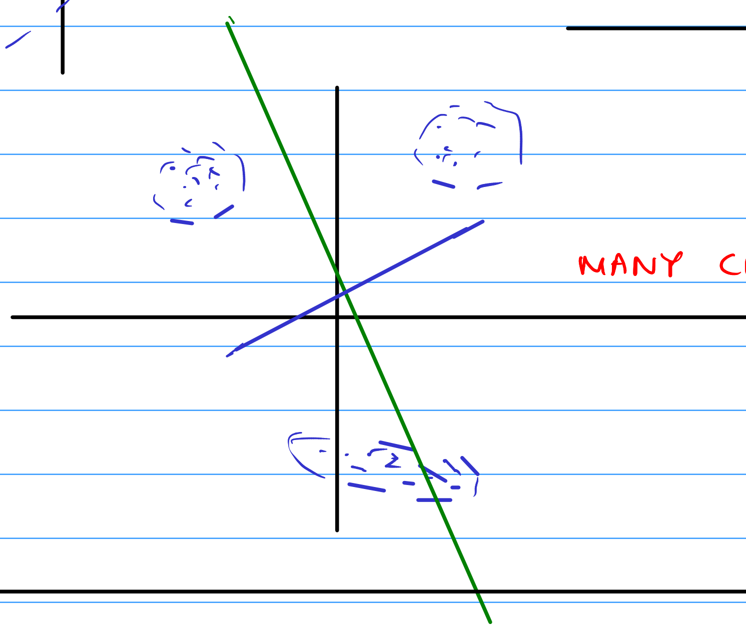
---

1. INTUITION BEHIND PCA



TRENDS IN A SINGLE CLUSTER OF DATA

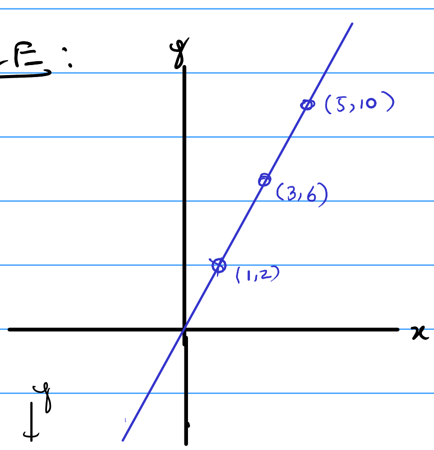


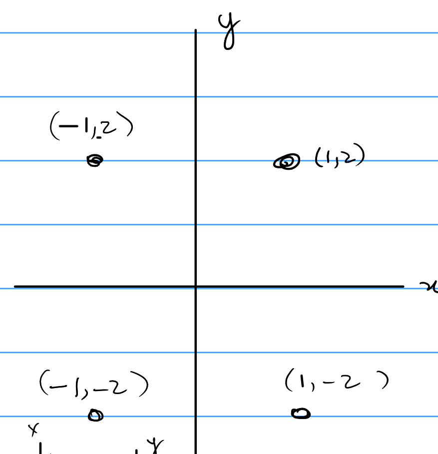


MANY CLUSTERS: PCA LESS EFFECTIVE

2. COVARIANCE MATRICES

— EXAMPLE:





$$A_1 = \begin{array}{c} \begin{array}{cc} x & y \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix} \\ \underleftrightarrow{\quad m \quad} \end{array} \updownarrow n$$

$$A_2 = \begin{array}{c} \begin{array}{cc} x & y \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 1 & 2 \\ 1 & -2 \\ -1 & 2 \\ -1 & -2 \end{bmatrix} \end{array}$$

— STEP 1: SUBTRACT COL. MEANS OUT FROM EACH COL

$$A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 6 \\ 5 & 10 \end{bmatrix} \; ; \text{ mean-centered data } = \tilde{A}_1 = \begin{bmatrix} -2 & -4 \\ 0 & 0 \\ 2 & 4 \end{bmatrix}$$

MEAN $= \frac{1+3+5}{3} = 3$ 6 mean

$$A_2 = \begin{bmatrix} 1 & 2 \\ 1 & -2 \\ -1 & 2 \\ -1 & -2 \end{bmatrix} \Longrightarrow \tilde{A}_2 = A_2$$

0   0

STEP 2: FORM $S = \frac{1}{n} \tilde{A}^T \tilde{A}$ ← COVARIANCE MATRIX

$$S_1 = \frac{1}{3} \underbrace{\begin{bmatrix} -2 & 0 & 2 \\ -4 & 0 & 4 \end{bmatrix}}_{\tilde{A}_1^T} \underbrace{\begin{bmatrix} -2 & -4 \\ 0 & 0 \\ 2 & 4 \end{bmatrix}}_{\tilde{A}_1} = \begin{bmatrix} \frac{8}{3} & \frac{16}{3} \\ \frac{16}{3} & \frac{32}{3} \end{bmatrix} \leftarrow S_1$$

SPREAD OF x

COVARIANCE BETWEEN X AND Y

Variance of x

variance of y = SPREAD OF Y

$$S_2 = \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -2 & 2 & -2 \end{bmatrix} \begin{array}{cc} x_1 & x_2 \\ \left[\begin{matrix} 1 & 2 \\ 1 & -2 \\ -1 & 2 \\ -1 & -2 \end{matrix}\right] \end{array} = \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 0 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \leftarrow S_2$$

COV(x,y)

$\overset{\xleftrightarrow{m}}{A} \in \mathbb{R}^{n \times m}$

IN GENERAL

$$S = \begin{bmatrix} s_1^2 & s_{12} & s_{13} & \cdots & s_{1m} \\ s_{21} & s_2^2 & s_{23} & \cdots & s_{2m} \\ \vdots & \vdots & s_3^2 & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & \cdots & s_m^2 \end{bmatrix} \Bigg\updownarrow m$$

$\xleftrightarrow{\quad m \quad}$

variance of x_1

$var(x_m)$

PROP 1: $s_i^2 \geq 0$

PROP 2: $s_{ij} = s_{ji} \quad \forall\, i \neq j$ (SYMMETRY OF S)

SYMMETRY OF S

$S = \frac{1}{n}\tilde{A}^T\tilde{A}$, $S^T = \frac{1}{n}\tilde{A}^T\tilde{A} = S$

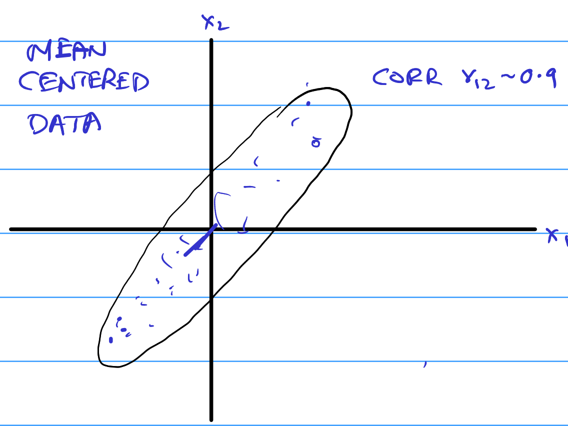
PROP 3: $|s_{ij}| \leq s_i s_j$ $\leftarrow$ CAUCHY INEQUALITY PROVES THIS

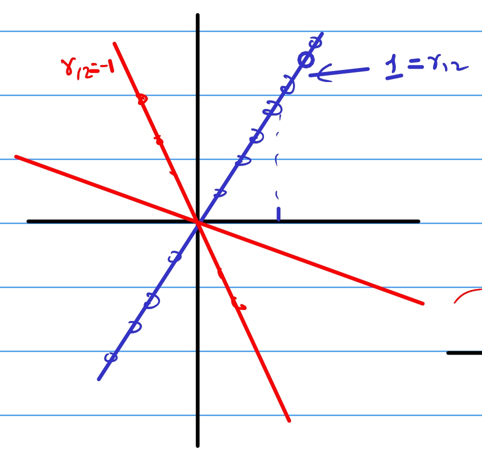
— CORRELATION BETWEEN x_i & x_j (COLS i AND j)

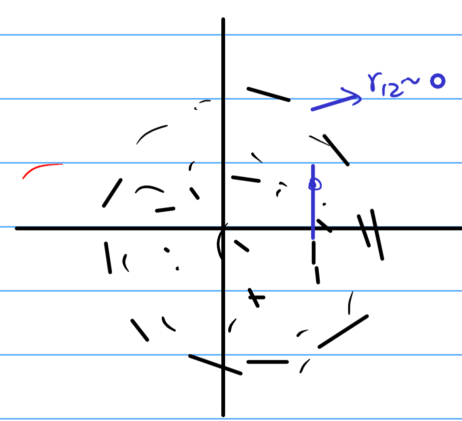
— $\boxed{r_{ij} \triangleq \dfrac{s_{ij}}{s_i\, s_j}}$

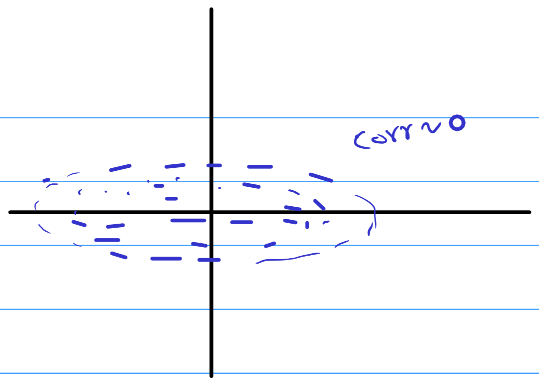
— PROP 4: $|r_{ij}| \leq 1 \implies -1 \leq r_{ij} \leq +1$

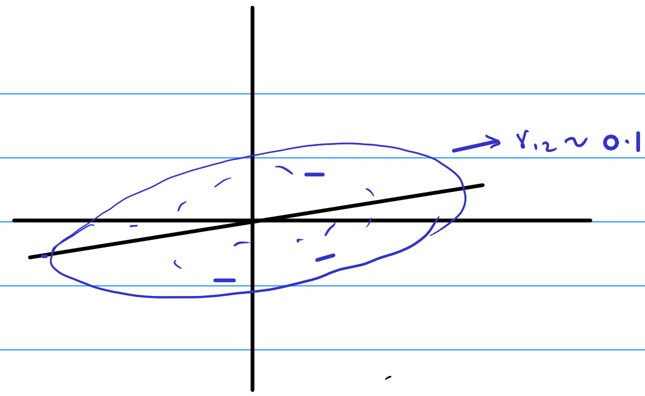
— INTUITION:











---

3. PCA

— EIGENDECOMPOSE S: $\quad S = P \Lambda \underbrace{P^{-1}}_{P^T}$

→ EIGENVECTORS (P) and EIGENVALUES (Λ) HAVE SPECIAL PROPS.:

↳ $\boxed{S = B^T B}$ → symmetric + positive-definite

→ **PROP 1:** P is UNITARY

$$P^T P = P P^T = I \Leftrightarrow P^T = P^{-1}$$

→ **PROP 2:** EIGENVALUES REAL AND ≥ 0

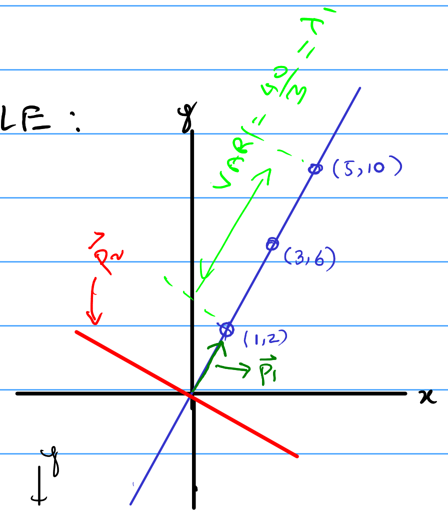
$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix} \qquad \begin{array}{l} \lambda_1, \lambda_2, \dots, \lambda_m \geq 0 \\ \lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_m \geq 0 \end{array}$$

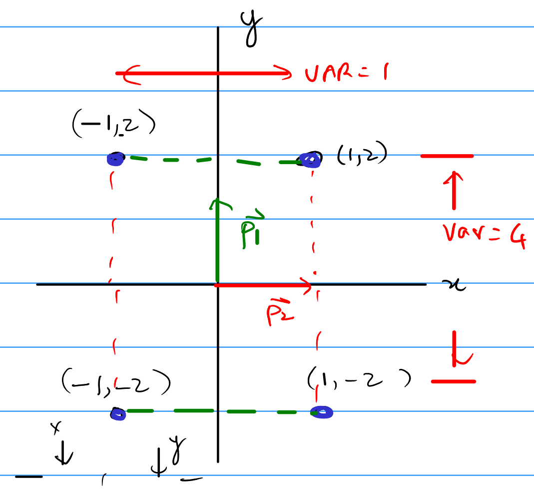
$$S_1 = \begin{bmatrix} \frac{8}{3} & \frac{16}{3} \\ \frac{16}{3} & \frac{32}{3} \end{bmatrix} = \underbrace{\frac{1}{\sqrt{5}} \begin{array}{cc} \downarrow \vec{p}_1 & \downarrow \vec{p}_2 \\ \end{array} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 40/3 & \\ & 0 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \frac{1}{\sqrt{5}}}_{P^T = P^{-1}}$$

$$\begin{array}{ccc} & \vec{p}_1 \quad \vec{p}_2 & \\ S_2 = \begin{bmatrix}1 & 0\\ 0 & 4\end{bmatrix} = & \underbrace{\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}}_{P} & \underbrace{\begin{bmatrix}4 & 0\\ 0 & 1\end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}}_{P^T = P^{-1} (= P)}\end{array}$$

MPLE :

S_1





S_2

$\vec{p}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}$ $\quad$ $\vec{p}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix}-2\\1\end{bmatrix}$ $\qquad$ $\vec{p}_1 = \begin{bmatrix}0\\1\end{bmatrix}, \quad \vec{p}_2 = \begin{bmatrix}1\\0\end{bmatrix}$

PCA: SUMMARY :

— EIGENDECOMPOSE S (COV. MAT.)

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

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

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

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