EECS 245 Fall 2025
Math for ML

Lecture 26: PCA, Continued

Read: Ch 5.4 (more details added)

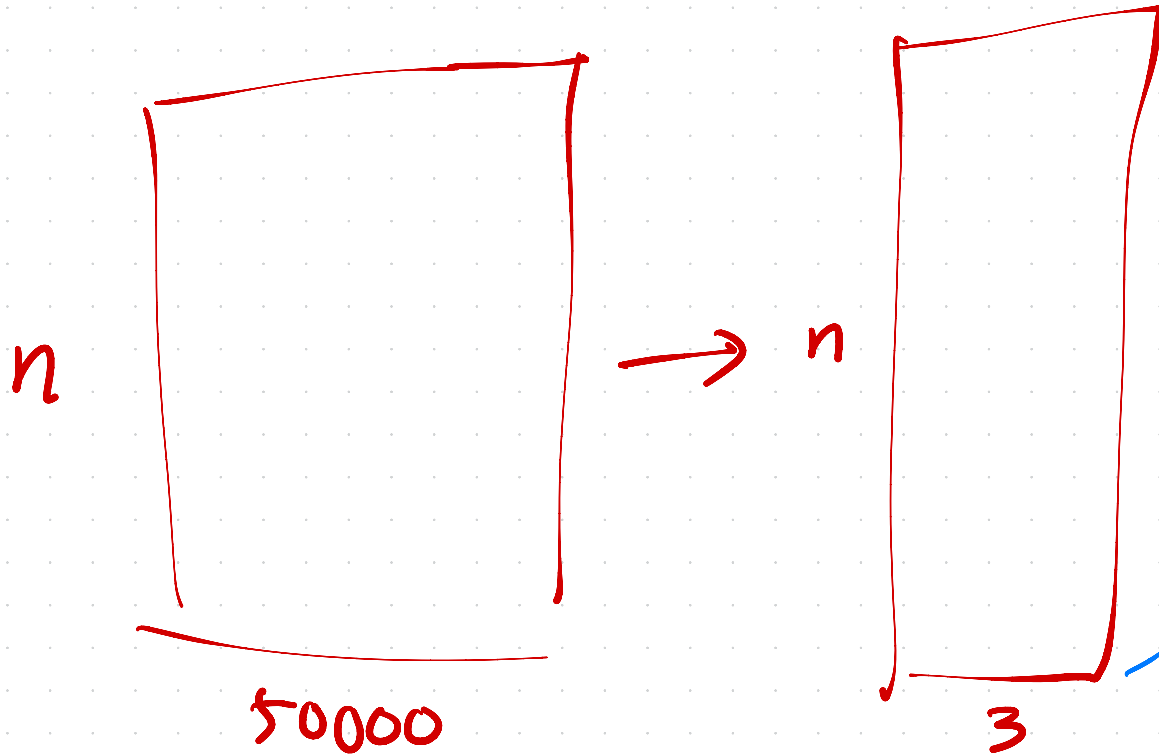
Agenda

More details on PCA!
Will appear on the exam
but will be gentle:
I promise!

Announcements

- HW 11 due Sunday, no slip days
- If 90% of the class fills out BOTH the End-of-Sem survey and evals, everyone gets 2% overall EC
- Practice problems worksheet coming soon!

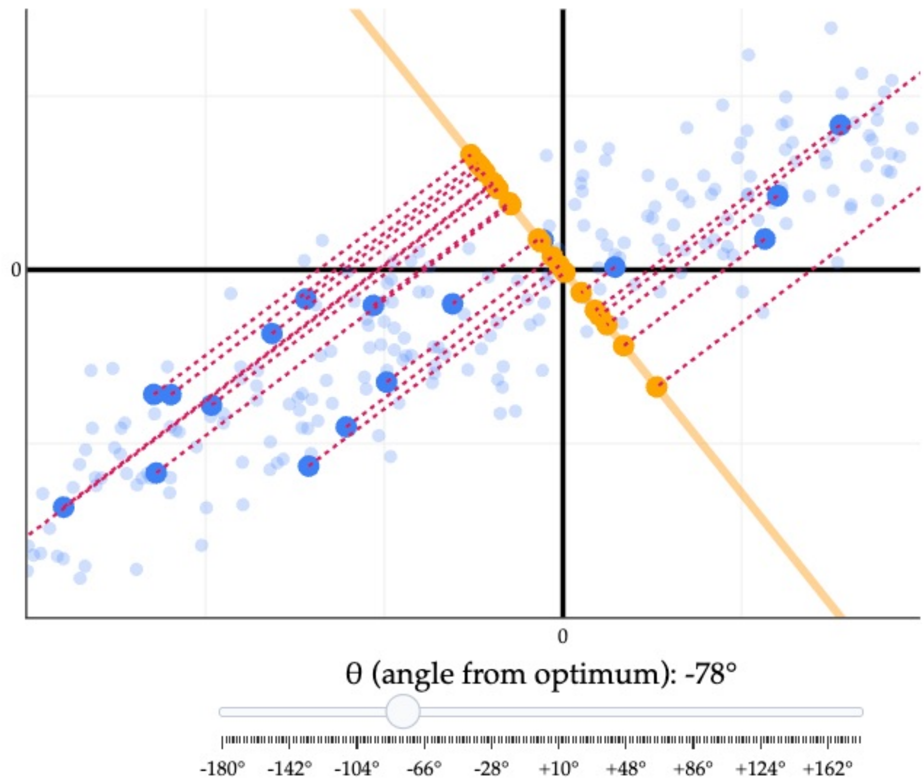
big picture: dimensionality reduction



how do we make those 3 features?

see new interactive animation in Ch. 5.4!

The shorter the orthogonal errors are, the more spread out the orange points are!



$$\tilde{x}_1 \cdot \vec{v} \over \vec{v} \cdot \vec{v}$$ $= \tilde{x}_1 \cdot \vec{v}$, since $\vec{v}$ unit

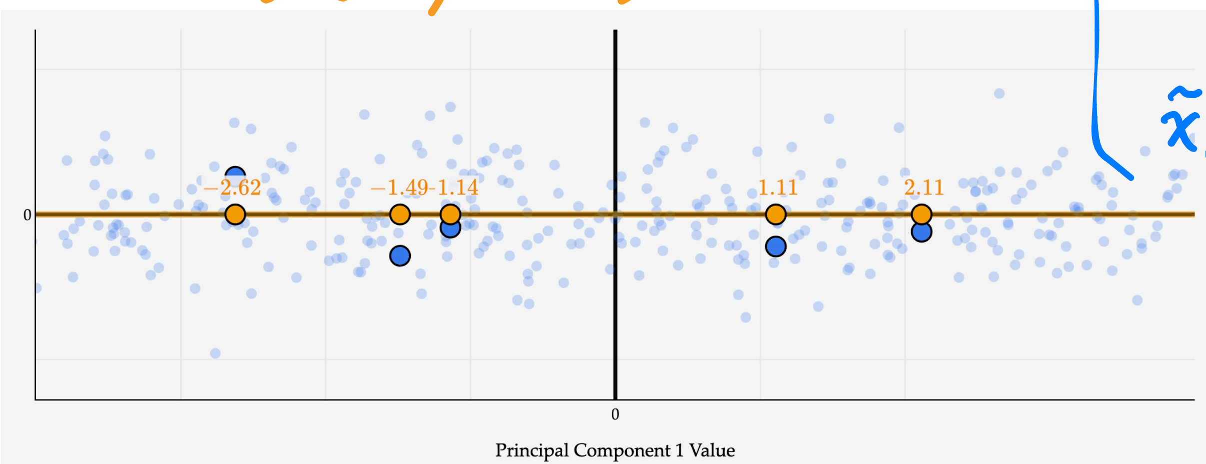
$$PV(\vec{v}) = \frac{1}{n} \| \tilde{X} \vec{v} \|^2$$

projected variance

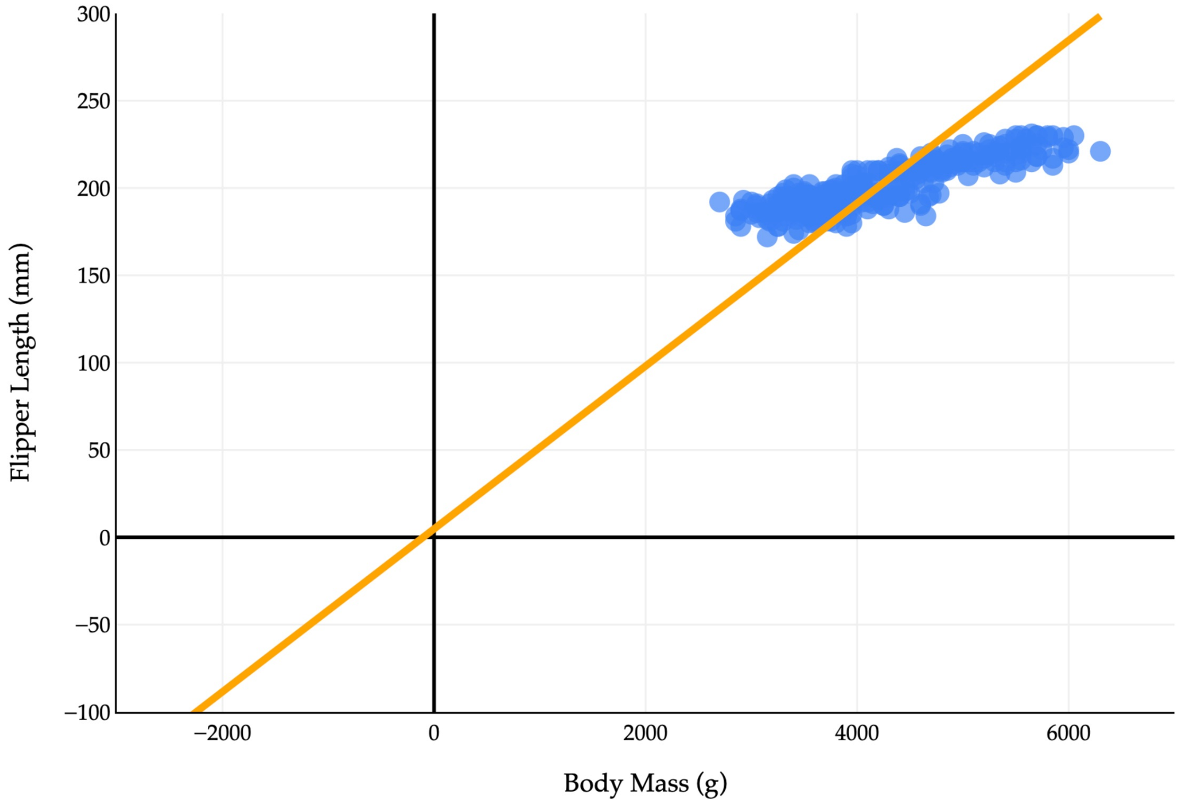
= variance of $-2.62, -1.49, \ldots$

$\tilde{X}$: mean centered

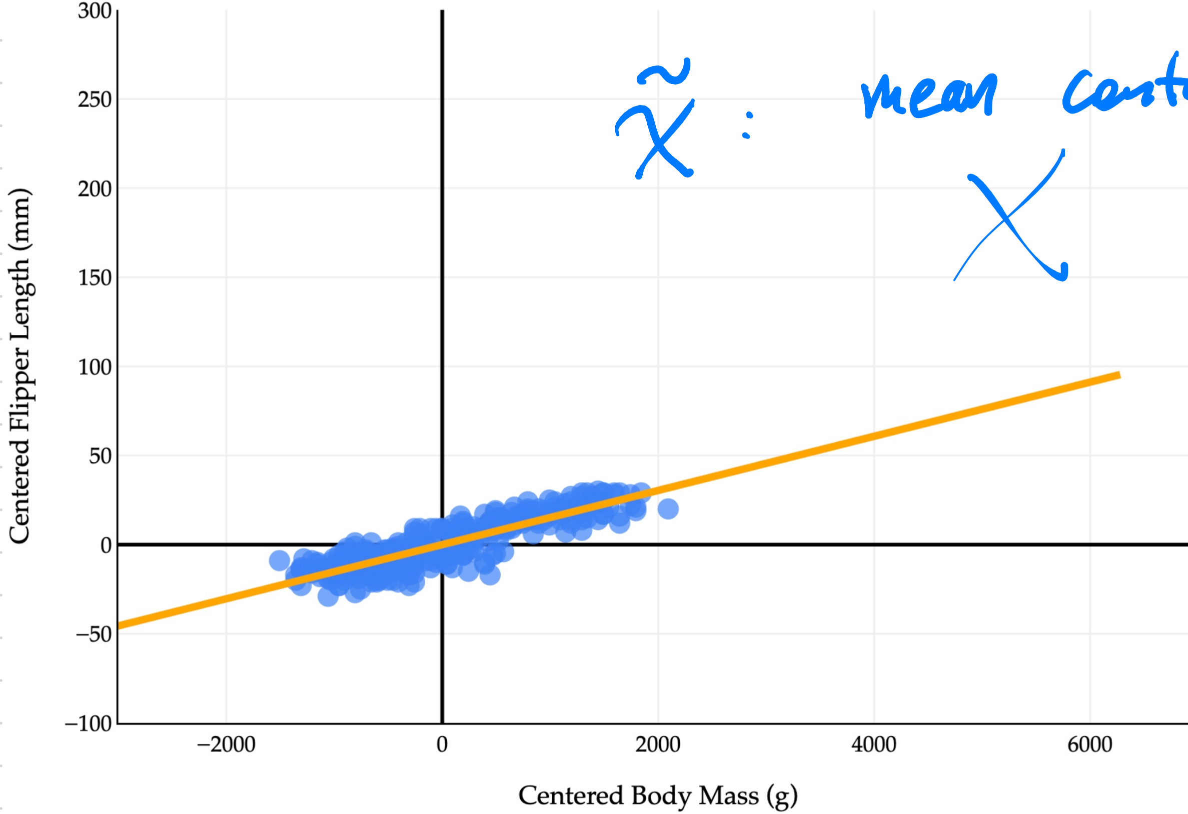
$$\tilde{X} \vec{v} = \begin{bmatrix} \tilde{x}_1 \cdot \vec{v} \\ \tilde{x}_2 \cdot \vec{v} \\ \vdots \\ \tilde{x}_n \cdot \vec{v} \end{bmatrix}$$



Our data isn't usually located near the origin...



...which is why we center the data first! This doesn't change its shape.

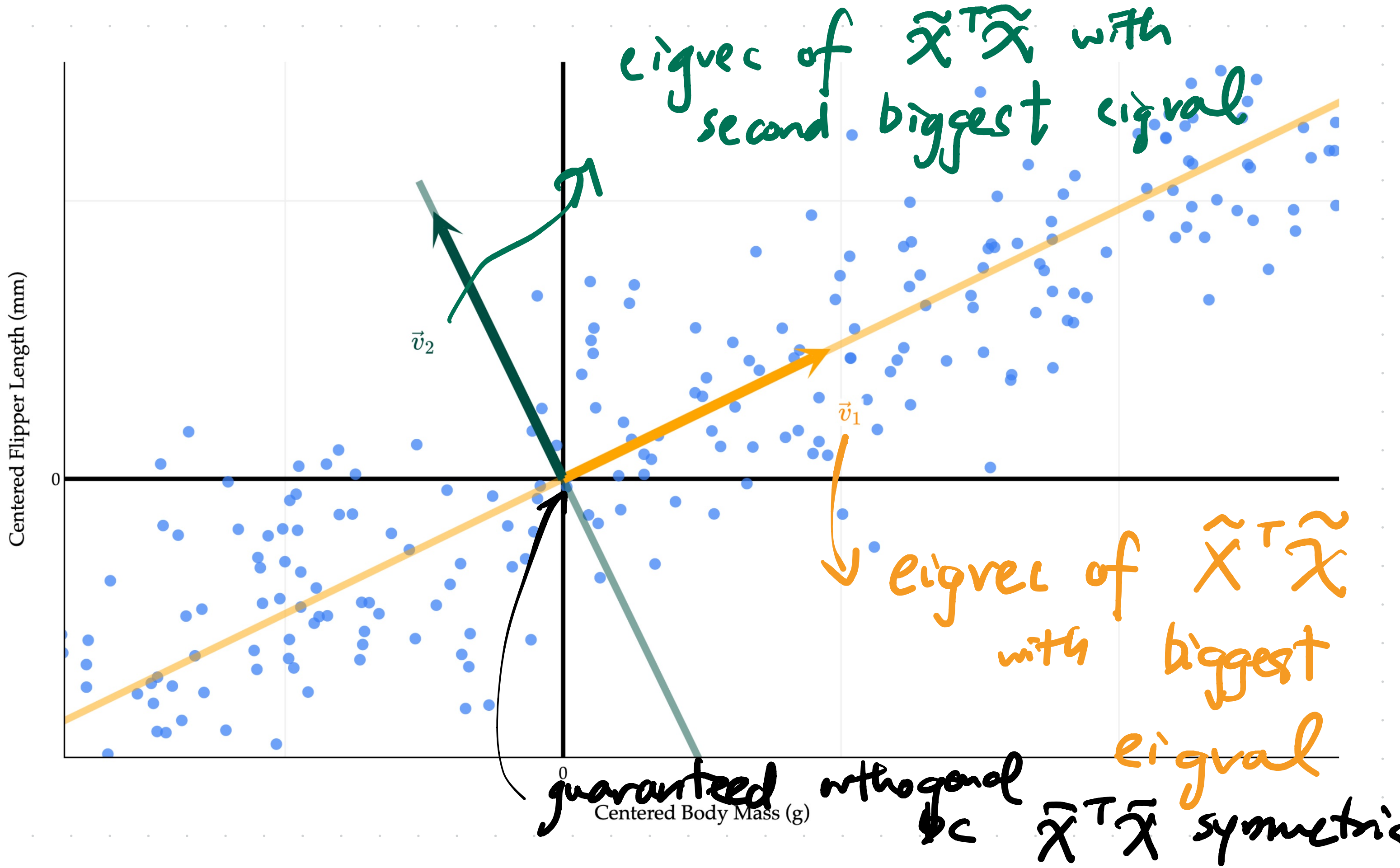


$\tilde{X}$: mean centering X

$$f(\vec{v}) = \frac{\|\tilde{X}\vec{v}\|^2}{\|\vec{v}\|^2}$$

review the Rayleigh quotient

$\Rightarrow$ the $\vec{v}$ that maximizes $f(\vec{v})$ is the eigenvector of $\tilde{X}^T\tilde{X}$ with the biggest eigenvalue!


eigvec of $\tilde{X}^T\tilde{X}$ with second biggest eigval
$\vec{v}_2$
$\vec{v}_1$
eigvec of $\tilde{X}^T\tilde{X}$ with biggest eigval
Centered Flipper Length (mm)
0
0
Centered Body Mass (g)
guaranteed orthogonal bc $\tilde{X}^T\tilde{X}$ symmetric

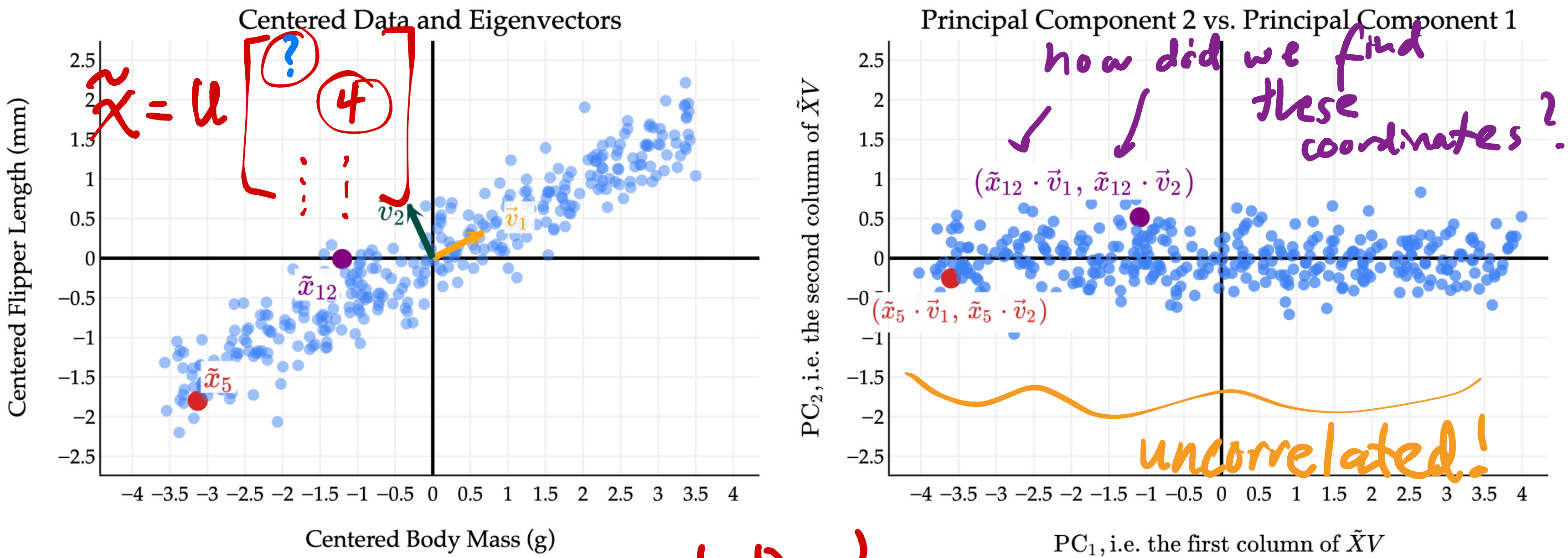
$$\tilde{X} = U \Sigma V^T$$

cols are the eigenvecs of $\tilde{X}\tilde{X}^T$

cols of V are eigenvecs of $\tilde{X}^T\tilde{X}$

big idea: SVD (on the centered matrix) finds the most important directions!

"principal component" = "new feature"



rotation!

computing principal components

$\vec{v}_j$ is the j'th column of V in

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

$$PC_j = \tilde{X} \vec{v}_j$$

if $\tilde{X}$ is $n \times d$,

$\vec{v}_j$ is $\mathbb{R}^d$

$X\vec{v}_j$ is $\mathbb{R}^n$

variance of a PC:

$$PC_j = \tilde{X}\vec{v}_j$$

$$\text{variance of } PC_j = \frac{1}{n}\|\tilde{X}\vec{v}_j\|^2$$

remember, $\vec{v}_j$ is column j in $\tilde{X} = U\Sigma V^T$

$$\tilde{X}\vec{v}_j = \sigma_j \vec{u}_j$$

how does this help?

$$\text{variance} = \frac{1}{n}\|\sigma_j \vec{u}_j\|^2 = \frac{1}{n}\sigma_j^2 \underbrace{\|\vec{u}_j\|^2}_{\text{unit!}} = \frac{\sigma_j^2}{n}$$

!

Suppose X is a 51×5 matrix, whose **first 3 rows** are given by

n d

$$\text{first 3 rows of } X = \begin{bmatrix} 3 & 12 & 5 & 1 & 5 \\ 3 & 4 & 8 & 2 & 1 \\ 1 & 2 & 7 & 2 & 1 \end{bmatrix}$$

not $\tilde{X}$!

$$\text{var } PC_j = \frac{\sigma_j^2}{n}$$

$2 - 1.2$

Consider the following information about the columns of X.

	Column 1	Column 2	Column 3	Column 4	Column 5
Mean	2	3	10	5	1
Variance	0.3	0.3	0.8	0.3	0.3

$$\frac{9^2+4^2}{9^2+4^2+2^2+1^2}$$

Let $\tilde{X}$ be the centered version of X, and let $\tilde{X} = U\Sigma V^T$ be the singular value decomposition of $\tilde{X}$.

Suppose the values along the diagonal of Σ are 9, 4, 2, 1, and 0.

a) What is rank(X)? Give your answer as an integer. 4 (num nonzero singular values)

b) What proportion of the total variance in X is accounted for by the second principal component? Give your answer as a fraction.

$$\frac{4^2}{9^2+4^2+2^2+1^2} = \frac{16}{102} = \frac{8}{51}$$

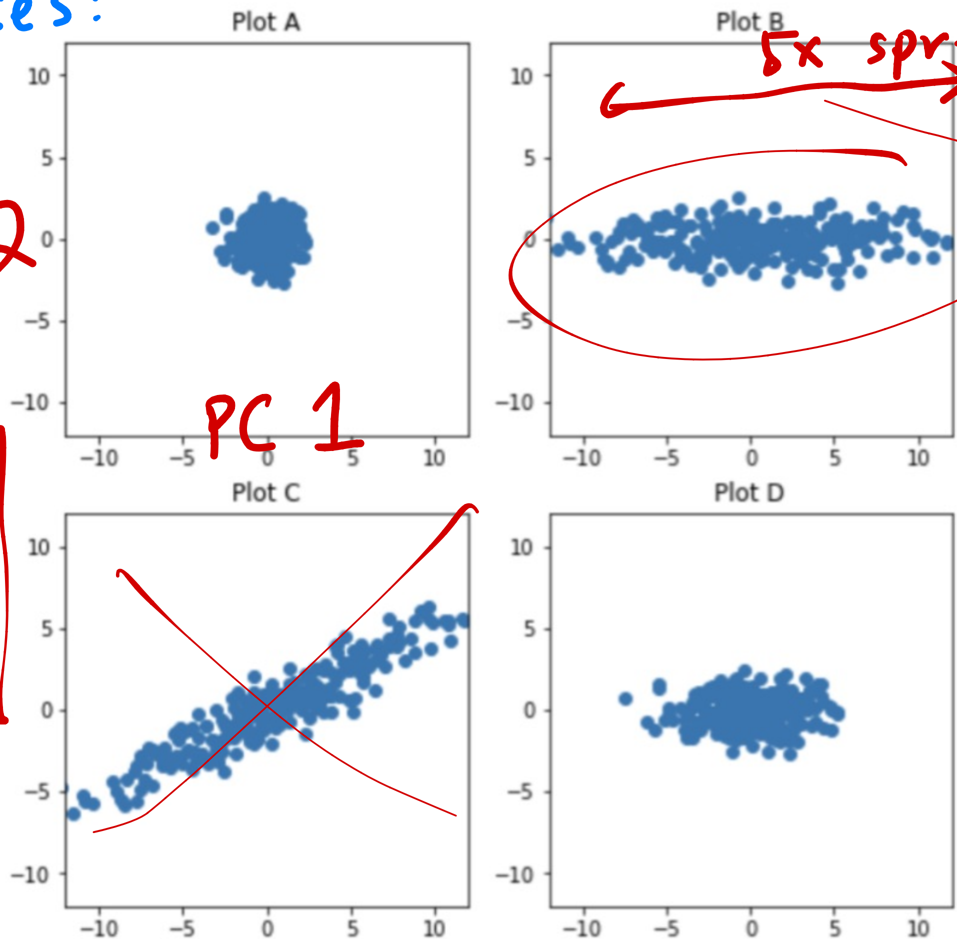
c) We want to choose the first k principal components, such that at least 95% of the variance in X is accounted for. What is the smallest possible value of k that we can choose? 2

d) Notice that the table provided does not include the variance of column 3. Given all the information above, what is the variance of column 3?

ratio of variances: $\frac{9^2}{4^2} = \frac{81}{16} \approx 5$

PC 2

$$\Sigma = \begin{bmatrix} 9 & & & & \\ & 4 & & & \\ & & 2 & & \\ & & & 1 & \\ & & & & 0 \end{bmatrix}$$



$$\text{var } PC_j = \frac{\sigma_j^2}{n}$$

prop of variance accounted for by

$$PC_j = \frac{\frac{\sigma_j^2}{\cancel{n}}}{\frac{\sigma_1^2}{\cancel{n}} + \frac{\sigma_2^2}{\cancel{n}} + \cdots + \frac{\sigma_r^2}{\cancel{n}}}$$

rank

$$= \frac{\sigma_j^2}{\sum_{k=1}^{r} \sigma_k^2}$$