EECS 245, Winter 2026

LEC 25 Principal Component Analysis, Continued

Agenda

last day of class!

- Recap: finding the "best direction"
- The Rayleigh quotient
- Principal components analysis

  → just the formal name for what we've already been doing!

Announcements

- HW 11 due tomorrow
- Final Exam on

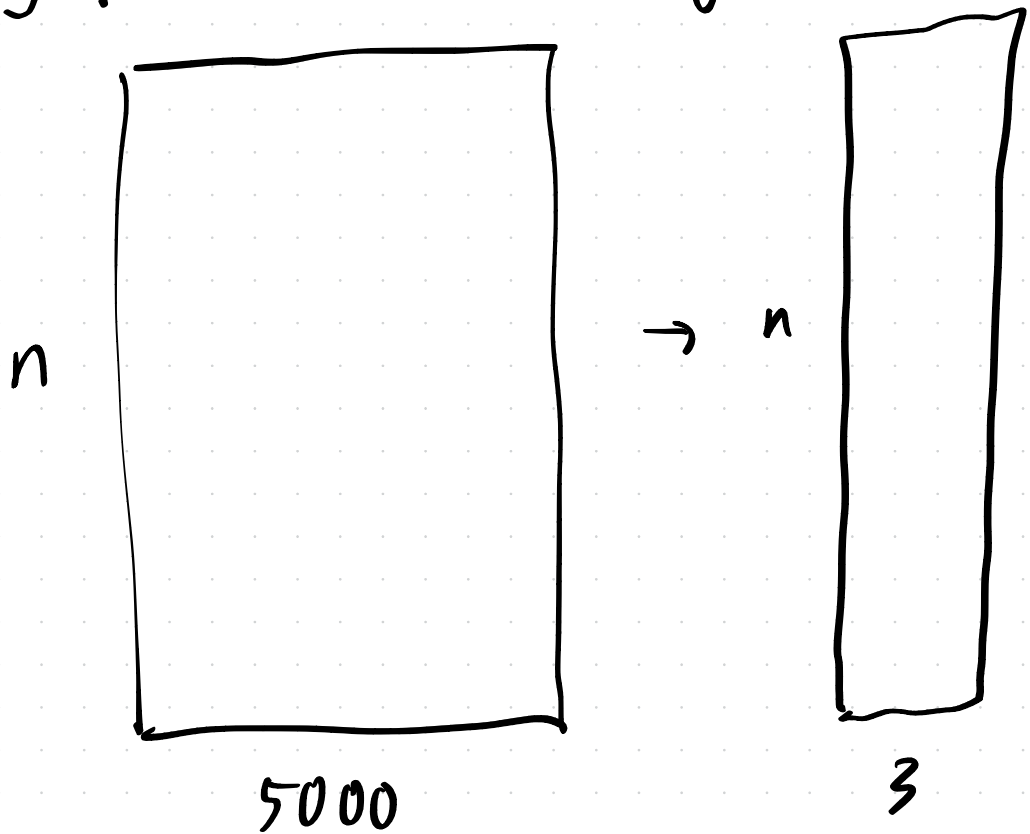
  Monday, April 27

  1:30-3:30PM

  1013 Dow

  see logistics on Ed
- If 80% of the class fills out both the End-of-Sem survey and official Evals, everyone gets 1% overall EC: due Wednesday!

Big picture: dimensionality reduction



How do we create these 3 features?

each new feature is a

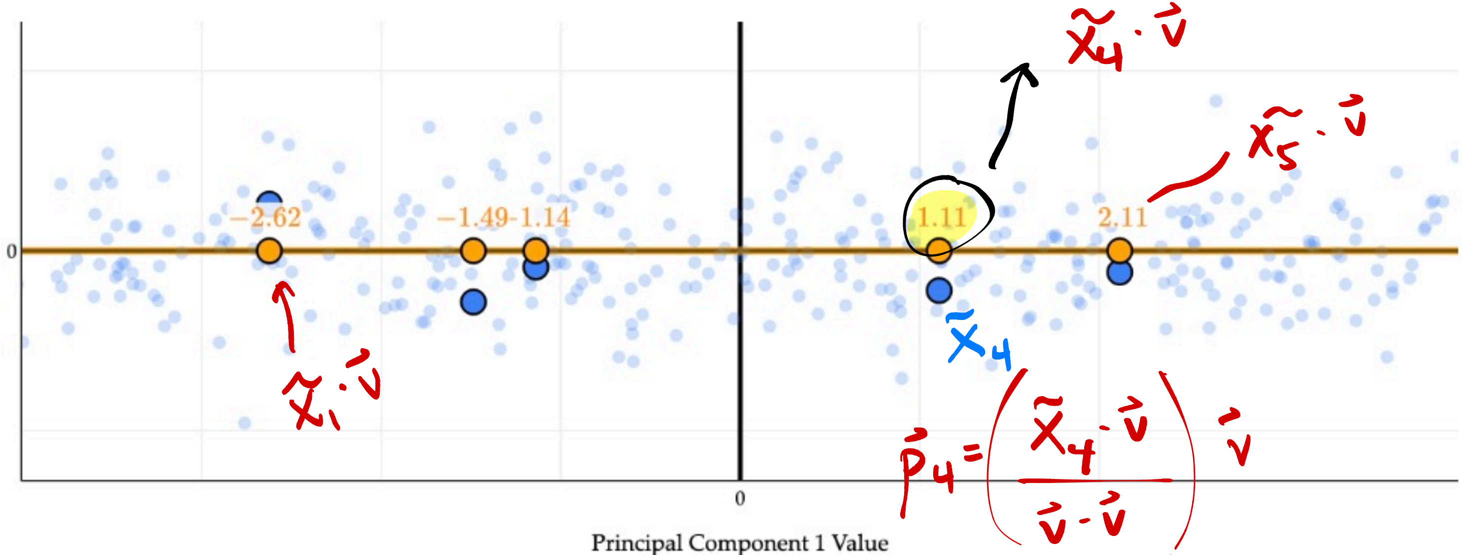
linear combination of

existing features!

⇒ "principal component" = "new feature"

⇒ PC 1 ⇒ first new feature

Once I've picked a unit vector $\vec{v}$, each orange point is described by just one number



$$\vec{p}_4 = \left(\frac{\tilde{x}_4 \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\right) \vec{v}$$

$$= (\tilde{x}_4 \cdot \vec{v})\, \vec{v}$$

Goal: Minimize mean squared orthogonal error

$\vec{v}$ unit vector

$$J(\vec{v}) = \frac{1}{n}\sum_{i=1}^{n} \|\tilde{x}_i - \vec{p}_i\|^2$$

row i of $\tilde{X}$, whose columns are mean centered

$$= \frac{1}{n}\sum_{i=1}^{n} \|\tilde{x}_i - (\tilde{x}_i \cdot \vec{v})\vec{v}\|^2$$

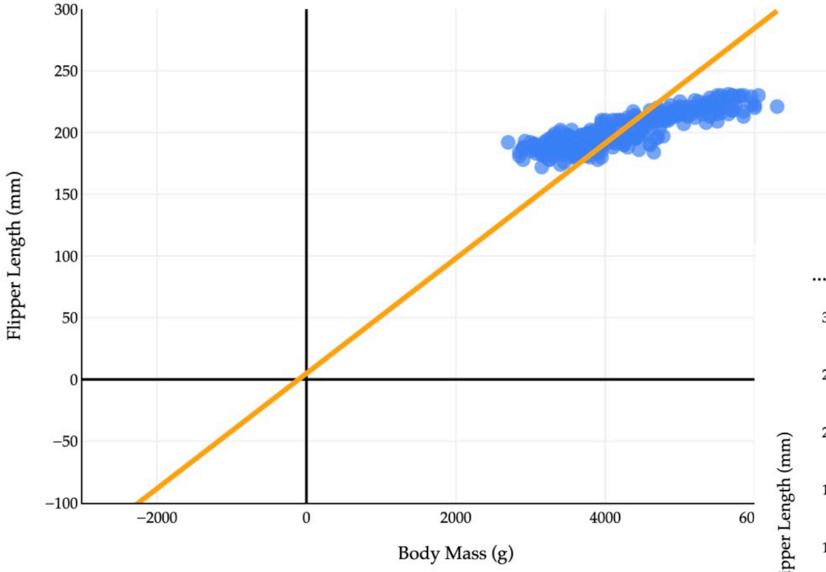
⋮

see next slide

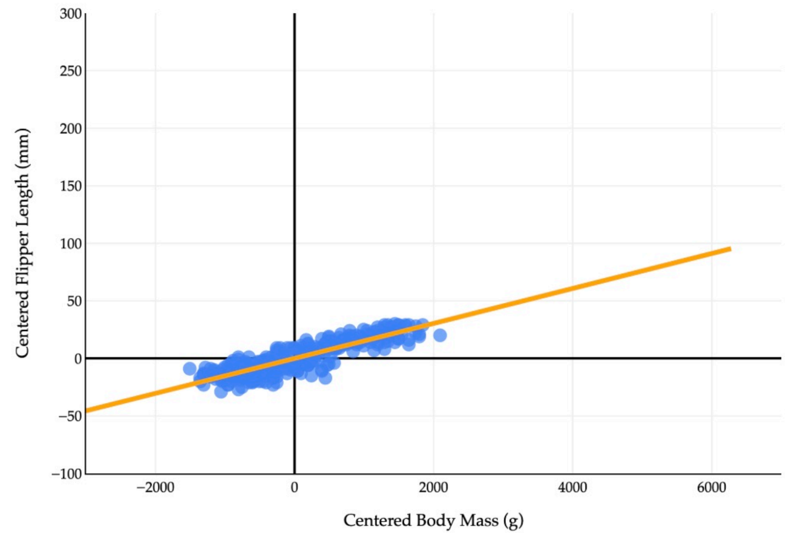
$$= \underbrace{\frac{1}{n}\sum_{i=1}^{n} \|\tilde{x}_i\|^2}_{\text{constant!}} - \frac{1}{n}\sum_{i=1}^{n} (\tilde{x}_i \cdot \vec{v})^2$$

$$
\begin{aligned}
J(\vec{v}) &= \frac{1}{n}\sum_{i=1}^{n}\|\tilde{x}_i - (\tilde{x}_i \cdot \vec{v})\vec{v}\|^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\tilde{x}_i\|^2 - 2\tilde{x}_i \cdot (\tilde{x}_i \cdot \vec{v})\vec{v} + ((\tilde{x}_i \cdot \vec{v})\vec{v}) \cdot ((\tilde{x}_i \cdot \vec{v})\vec{v})\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\tilde{x}_i\|^2 - 2\tilde{x}_i \cdot (\tilde{x}_i \cdot \vec{v})\vec{v} + (\tilde{x}_i \cdot \vec{v})^2\underbrace{\vec{v} \cdot \vec{v}}_{=1}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\tilde{x}_i\|^2 - 2(\tilde{x}_i \cdot \vec{v})^2 + (\tilde{x}_i \cdot \vec{v})^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\|\tilde{x}_i\|^2 - (\tilde{x}_i \cdot \vec{v})^2\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\|\tilde{x}_i\|^2 - \frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i \cdot \vec{v})^2
\end{aligned}
$$

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

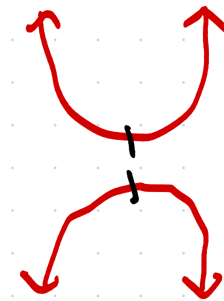


...which is why we center the data first! This doesn't change its shape.



Discovery: minimizing

$$J(\vec{v}) = \frac{1}{n}\sum_{i=1}^{n} \|\tilde{x}_i - (\tilde{x}_i \cdot \vec{v})\vec{v}\|^2$$



is the same as minimizing

$$J(\vec{v}) = \frac{1}{n}\sum_{i=1}^{n} \|\tilde{x}_i\|^2 - \underline{\frac{1}{n}\sum_{i=1}^{n} (\tilde{x}_i \cdot \vec{v})^2}$$

is the same as maximizing

$$\underline{PV}(\vec{v}) = \frac{1}{n}\sum_{i=1}^{n} (\tilde{x}_i \cdot \vec{v})^2 = \frac{1}{n}\|\tilde{X}\vec{v}\|^2$$

projected variance : variance of the new coordinates

Why is $\frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i \cdot \vec{v})^2 = \frac{1}{n}\sum_{i=1}^{n}(\tilde{x}_i \cdot \vec{v} - 0)^2$

equal to the variance of $\underbrace{\tilde{x}_1 \cdot \vec{v}, \tilde{x}_2 \cdot \vec{v}}$, ..., $\underbrace{\tilde{x}_n \cdot \vec{v}}$?

⇒ why do they have a mean of 0?

→ hint: write $\tilde{x}_1 \cdot \vec{v}, \tilde{x}_2 \cdot \vec{v}, \ldots, \tilde{x}_n \cdot \vec{v}$ involving $\tilde{X}$ and $\vec{v}$

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

sum of these is 0, so avg is 0 too!

$$(\tilde{X}\vec{v})^T \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \vec{v}^T \tilde{X}^T \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$= \vec{v}^T \begin{bmatrix} \text{sum of } \tilde{X}\text{'s first col} \\ \text{sum of } \tilde{X}\text{'s second col} \\ \vdots \end{bmatrix}$$

$$= \vec{v}^T \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

because $\tilde{X}$ is mean-centered

Goal: Find $\vec{v}$ maximizes

$$\cancel{\frac{1}{n}} \|\tilde{X}\vec{v}\|^2 \quad \text{subject to} \quad \|\vec{v}\| = 1$$

Equivalent goal: Find $\vec{v}$ that maximizes

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

if $\vec{v}$ isn't unit length,

its length "cancels" in the numerator

denominator

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

see HW 9!

$$\nabla f(\vec{v}) = \frac{2}{\|\vec{v}\|^2}\left(\underbrace{\tilde{X}^T\tilde{X}}_{\text{"A" in hw question}} - f(\vec{v})\right)\vec{v}$$

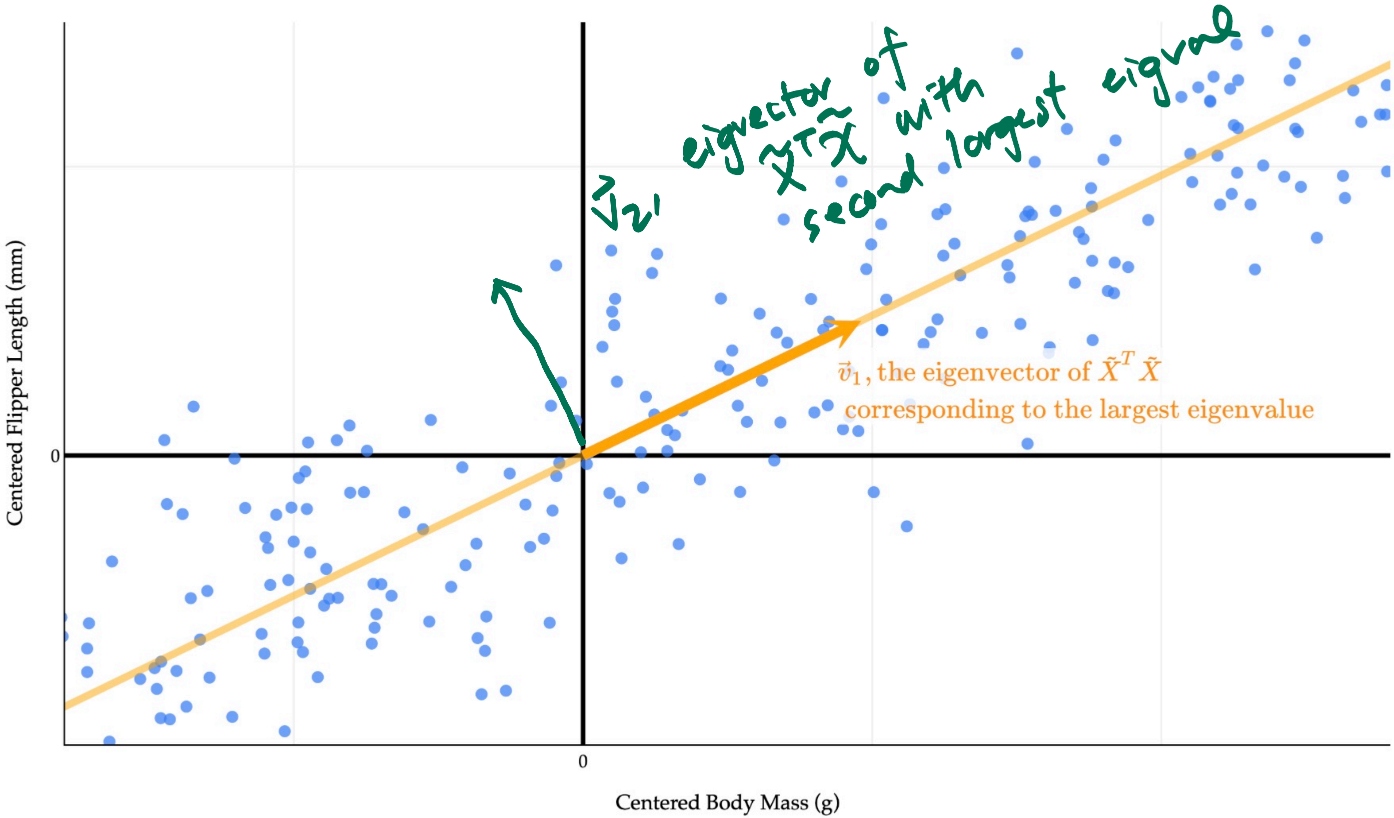
set to 0: $\tilde{X}^T\tilde{X}\vec{v} - f(\vec{v})\,\vec{v} = \vec{0}$

$$\tilde{X}^T\tilde{X}\vec{v} = f(\vec{v})\,\vec{v}$$

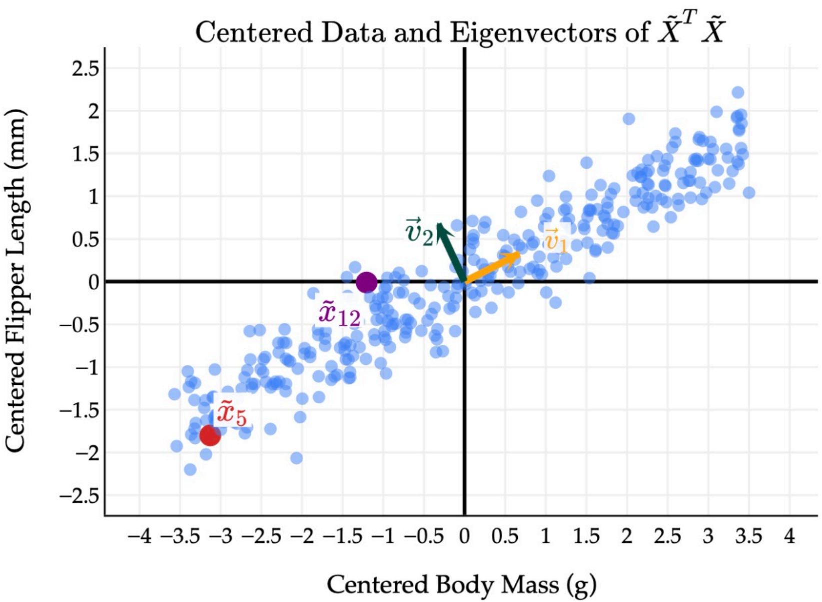
$$\tilde{X}^T \tilde{X} \vec{v} = f(\vec{v}) \vec{v}$$

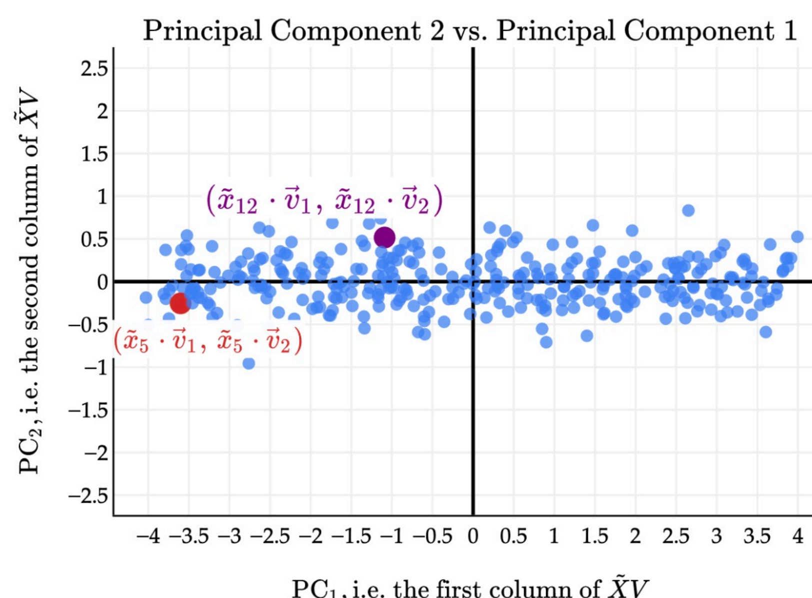
If $\vec{v}$ maximizes $f(\vec{v})$,
it must be an eigenvector
of $\tilde{X}^T \tilde{X}$!!!!

→ pick eigenvector with largest
eigenvalue


$\vec{v}_2$, eigvector of $\tilde{X}^T\tilde{X}$ with second largest eigval
$\vec{v}_1$, the eigenvector of $\tilde{X}^T \tilde{X}$
corresponding to the largest eigenvalue
Centered Flipper Length (mm)
0
0
Centered Body Mass (g)

rotated!





principal components are uncorrelated! (r=0)

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

SVD of $\tilde{X}$

cols of V
(rows of V^T)
describe the best directions

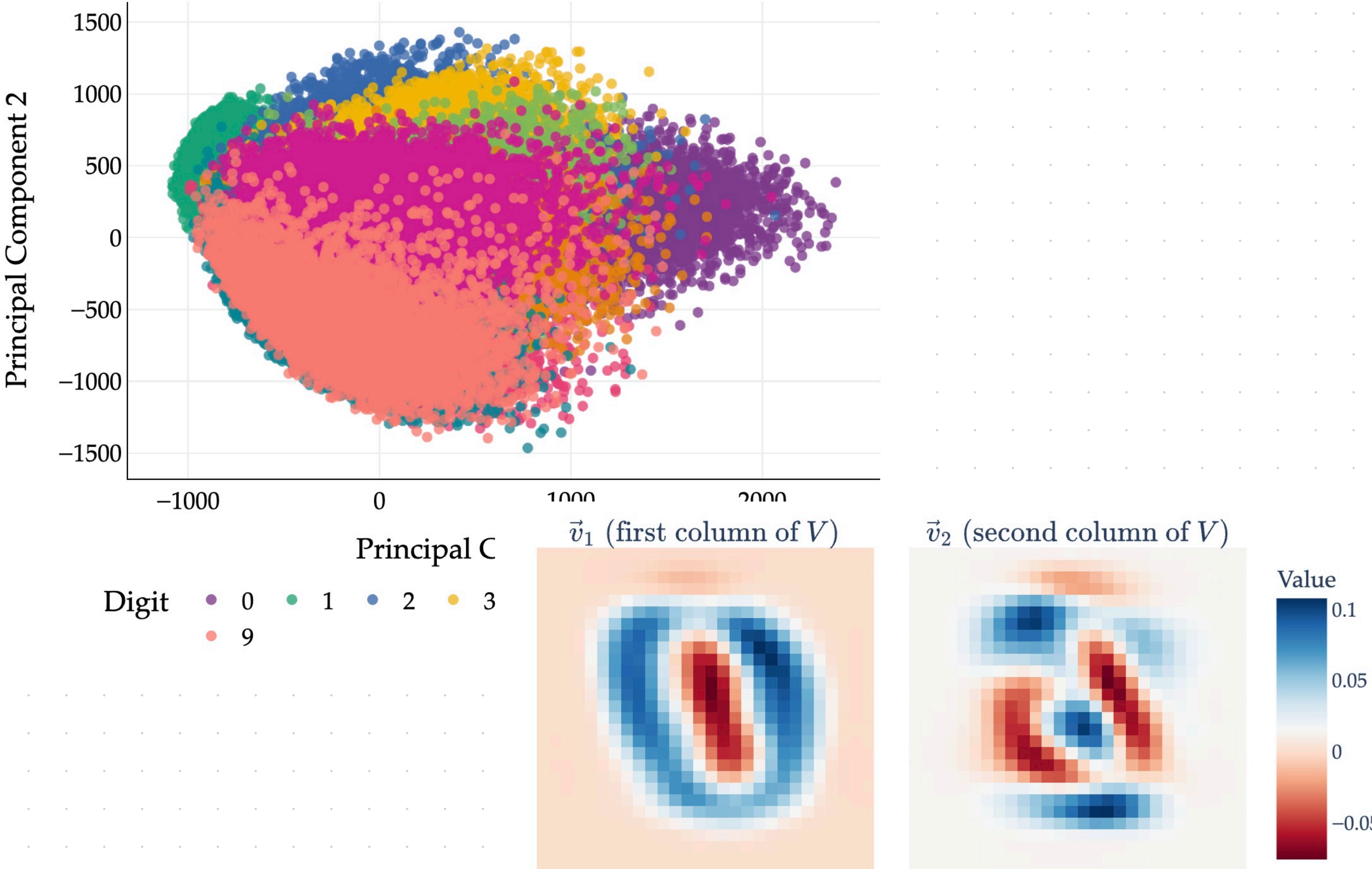
1. Variance of a principal component?

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

$$= \frac{1}{n} \| \sigma_j \vec{u}_j \|^2$$

j^{th} column of U in $\tilde{X} = U \Sigma V^T$

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



Principal Component 2
1500
1000
500
0
−500
−1000
−1500
−1000
0
1000
2000
Principal C
Digit
0
1
2
3
9
$\vec{v}_1$ (first column of V)
$\vec{v}_2$ (second column of V)
Value
0.1
0.05
0
−0.05