EECS 245, Spring 2026

LEC 14 PCA, Review

→ No new readings, other than Ch. 10

Agenda

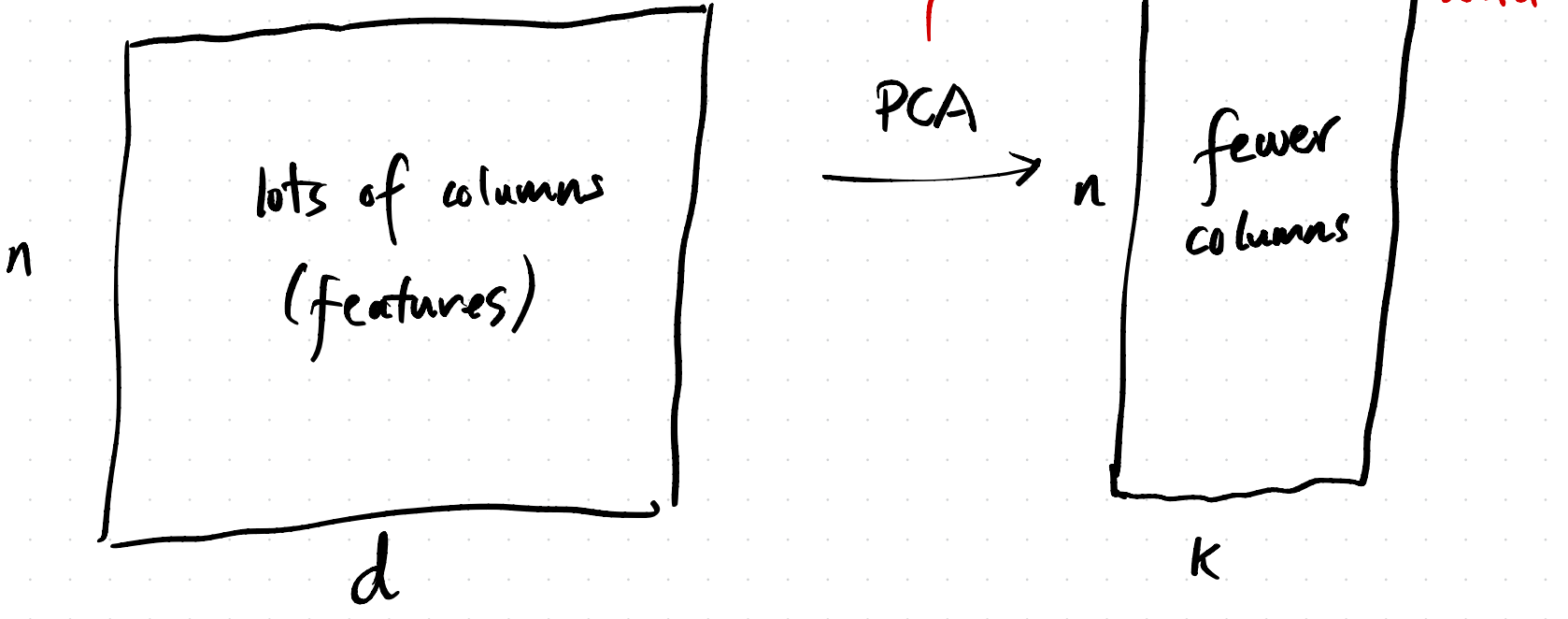
- Recap: PCA
- Take up selected problems from WN26 Final Exam
- Open office hours

Announcements

- Exam tomorrow: 8-10AM, 1018 DOW
- All assignment solutions posted
- Fill out the End-of-Semester Survey + Evals for 1% EC by tonight!
- Get some rest!!!

Recap: what is PCA?

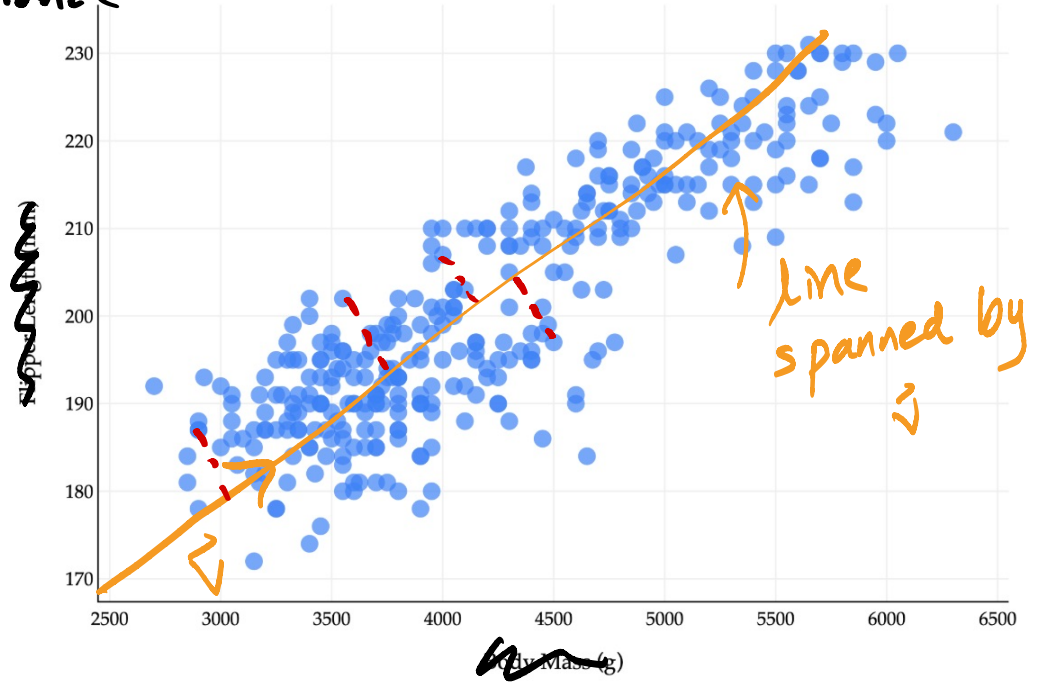
big idea: dimensionality reduction



"Best direction vector" $\vec{v}$

- $\vec{v}$ minimizes mean squared orthogonal error
- $\vec{v}$ maximizes variance

↑ equivalent!



$\tilde{X}$: mean-centered version of X

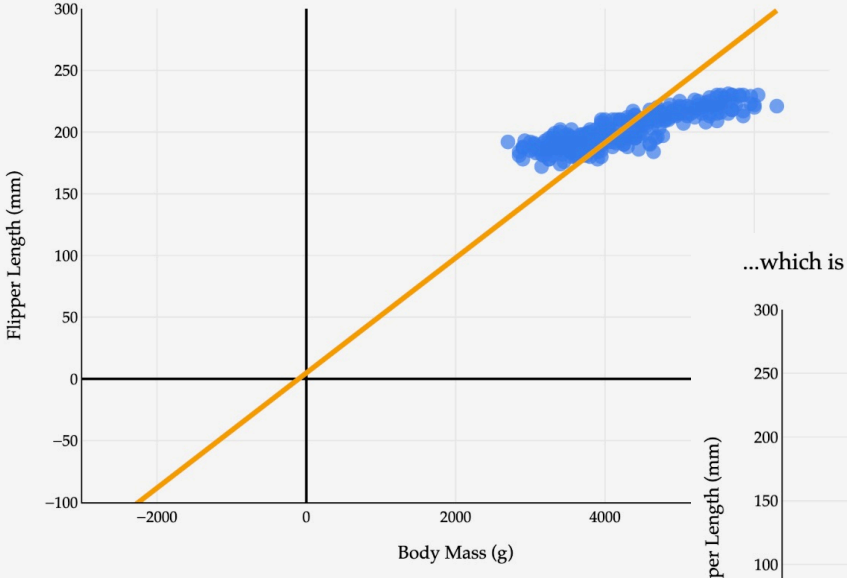
e.g.

$\tilde{X}$: 1000×70

$\tilde{X}\vec{v} \in \mathbb{R}^{1000}$ → one new feature

↑ among all vectors $\vec{v}$, we choose the one that makes $X\vec{v}$'s variance as large as possible

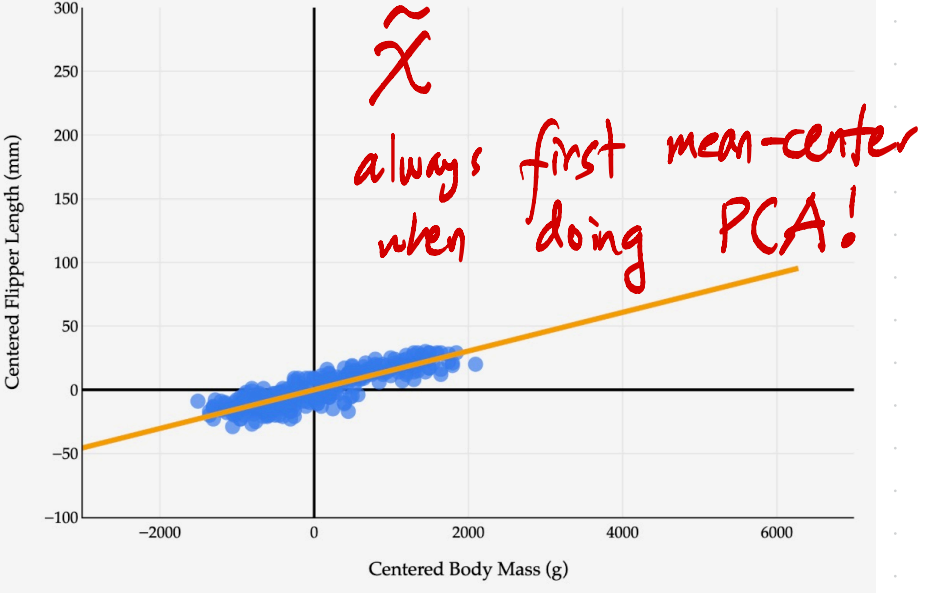
Our data isn't usually located near the origin...



X

$\downarrow$

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



$\tilde{X}$

always first mean-center when doing PCA!

Where does the best direction vector $\vec{v}$ come from?

key idea: find the singular value decomposition of $\tilde{X}$

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

⇒ the columns of V contain the "best" direction vectors!

$$V = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_d \\ | & | & & | \end{bmatrix}$$

$\vec{v}_1$: the best (max variance)

$\vec{v}_2$: the best, while being orthogonal to $\vec{v}_1$

$\vec{v}_7$: the direction that captures the most variance, while being orthogonal to $\vec{v}_1, \dots, \vec{v}_6$

"new feature" = "principal component"

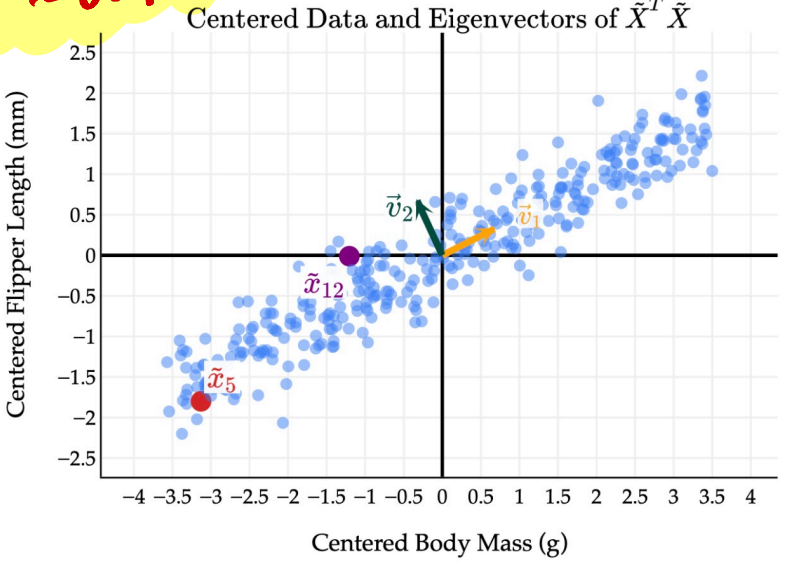
$$PC\ j = \tilde{X}\vec{v}_j = \sigma_j \vec{u}_j$$

$\vec{u}_j$ ~ col j of U

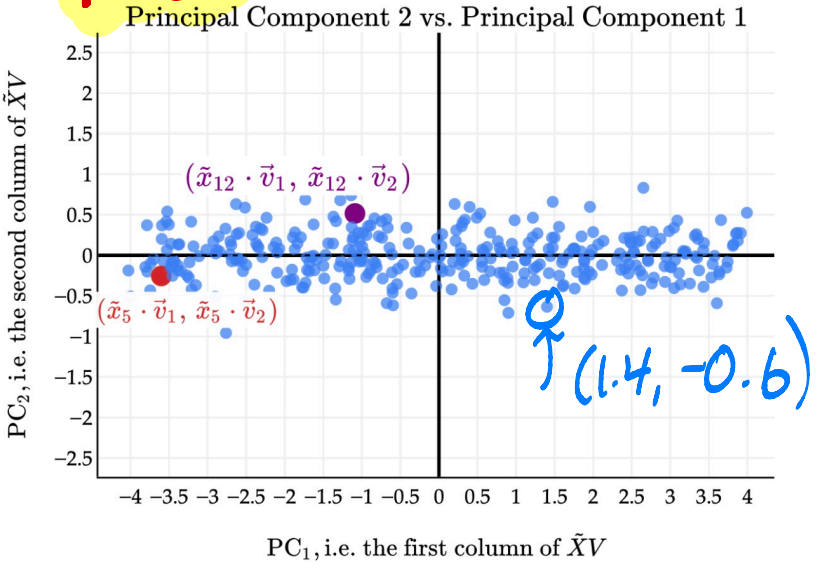
$\vec{v}_j$ col j of V in $\tilde{X} = U\Sigma V^T$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}$$

$r \approx 0.7$



$r = 0$



$$PC\ 1 = \tilde{X}\vec{v}_1 = \begin{bmatrix} \tilde{x}_1 \cdot \vec{v}_1 \\ \tilde{x}_2 \cdot \vec{v}_1 \\ \vdots \\ \tilde{x}_n \cdot \vec{v}_1 \end{bmatrix}$$

PC 1 for row i:

$$\tilde{x}_i \cdot \vec{v}_1$$

PC 2 for row i:

$$\tilde{x}_i \cdot \vec{v}_2$$

variance of PC $j = \frac{\sigma_j^2}{n}$

PC $j = \tilde{X}\vec{v}_j$

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

σ_j: constant

$\|\vec{u}_j\| = 1$, because U in $U\Sigma V^T$ is orthogonal

$$= \frac{\sigma_j^2}{n}$$

"projected variance" = "variance"

proportion of variance captured by first k PCs

$$= \frac{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} + \cdots + \frac{\sigma_k^2}{n}}{\text{sum of variances of original cols}}$$

$\sim$ total variance

$$= \frac{\frac{\sigma_1^2}{n} + \frac{\sigma^2}{n} + \cdots + \frac{\sigma_k^2}{n}}{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} + \cdots + \frac{\sigma_r^2}{n}}$$

r = rank of data

---

prop captured by first k PCs $= \dfrac{\sum_{j=1}^{k} \sigma_j^2}{\sum_{j=1}^{r} \sigma_j^2}$

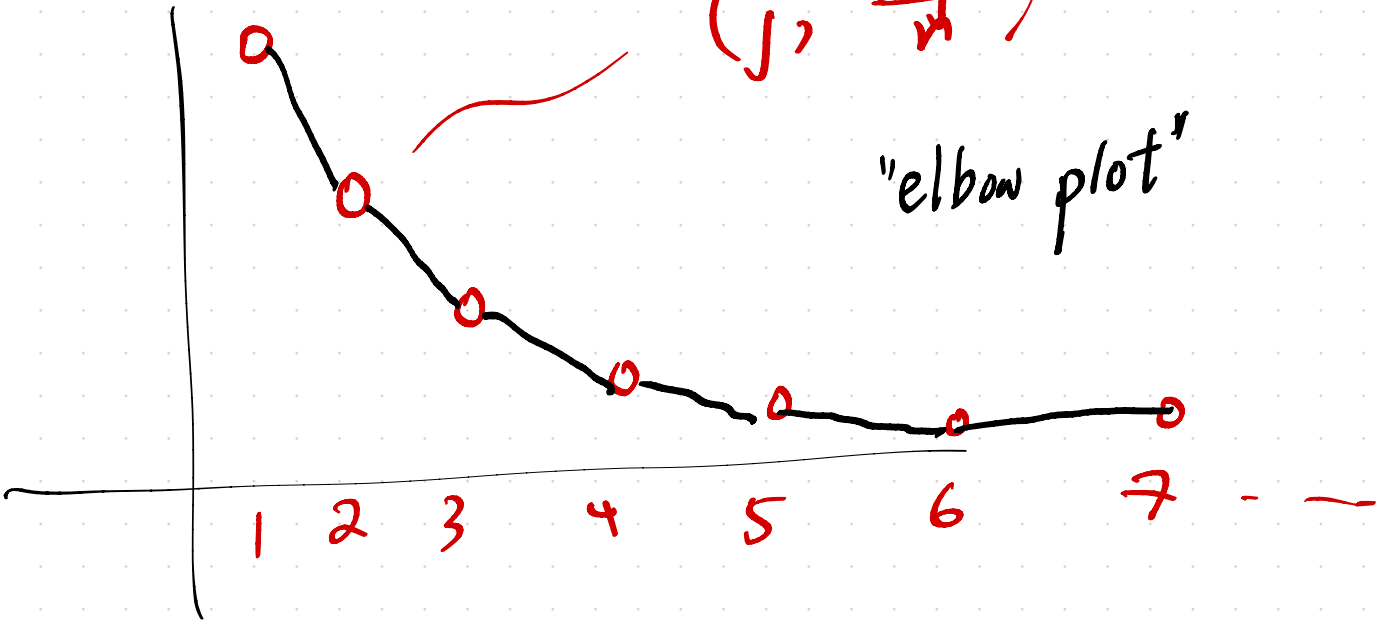
rank of $\tilde{X}$

you already practiced with this in Lab 12!

see notes for visuals

$d = 100$ features $\longrightarrow$ $k = 10$ PCs

$\left(j, \frac{\sigma_j^2}{n}\right)$



"elbow plot"

Problem 2 (13 pts) **Counts towards Midterm 1 redemption score**

Suppose a dataset of n points, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, has the following properties:

mean of y-values $= \bar{y} = 11$, standard deviation of x-values $= \sigma_x = 2$, $\sigma_y = 6$

The simple linear regression line that minimizes mean squared error for predicting y_i from x_i is

$$h(x_i) = 15 - x_i$$

$w_1^* = -1$ $w_0^* = 15 = \bar{y} - w_1^* \bar{x}$

a) (3 pts) What is $\bar{x}$, the mean of the x-values? Give your answer as a number with no variables.

$h(\bar{x}) = \bar{y}$ $h(\bar{x}) = 15 - \bar{x} = 11$

$\bar{x} =$ 4

Now, consider a new dataset, $(t_1, z_1), (t_2, z_2), \ldots, (t_n, z_n)$, defined by $t_i = 5 - x_i$ and $z_i = 2y_i - 1$.

Let $g(t_i) = \beta_0^* + \beta_1^* t_i$ be the best simple linear regression line for predicting z_i from t_i.

$\sigma_z = |2| \sigma_y$

b) (6 pts) Find β_0^*, the intercept of the best simple linear regression line for predicting z_i from t_i. Show your work, and write your final answer in the box provided. Your answer should be a number with no variables.

$\beta_0^* = \bar{z} - \beta_1^* \bar{t} = 21 - 2 \cdot 1 = 19$

$\bar{z} = 2\bar{y} - 1 = 2 \cdot 11 - 1 = 21$

$\bar{t} = 5 - \bar{x} = 5 - 4 = 1$

$\beta_1^* = r_{tz} \frac{\sigma_z}{\sigma_t} = r_{tz} \frac{2\sigma_y}{\sigma_x} = -r_{xy} \frac{2\sigma_y}{\sigma_x} = -2 r_{xy} \frac{\sigma_y}{\sigma_x}$ (w_1^*)

$= -2(-1) = 2$

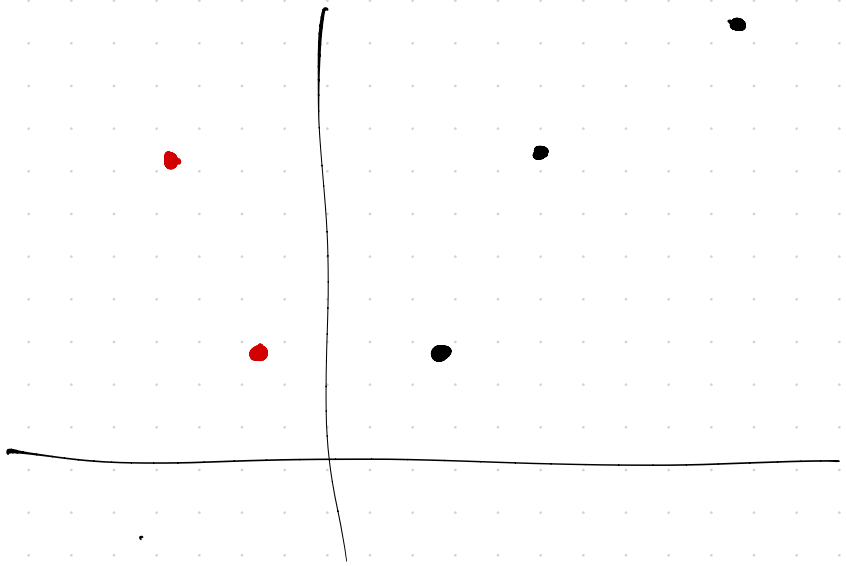
$\beta_0^* =$ 19

c) (4 pts) Let M be the mean squared error of the model $h(x_i) = 15 - x_i$'s predictions on the dataset $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, and M' be the mean squared error of the model $g(t_i) = \beta_0^* + \beta_1^* t_i$'s predictions on the dataset $(t_1, z_1), (t_2, z_2), \ldots, (t_n, z_n)$.

What is the value of the fraction $\frac{M}{M'}$? *If it's not clear, M' is on the denominator.*

◯ 1/5 ◯ 1/4 ◯ 1/2 ◯ 1 ◯ 2 ◯ 4 ◯ 5 ◯ Impossible to tell

$(1,1)$
$(2,3)$
$(4,7)$



$(-1,1)$
$(-2,3)$
$(-4,7)$

$$r_{tz} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{t_i-\bar{t}}{\sigma_t}\right)\left(\frac{z_i-\bar{z}}{\sigma_z}\right)$$

$$= \frac{1}{n}\sum\left(\frac{5-x_i-(5-\bar{x})}{\sigma_x}\right)\left(\frac{2y_i-1-(2\bar{y}-1)}{2\sigma_y}\right)$$

$$= \frac{1}{n}\sum\left(-\frac{(x_i-\bar{x})}{\sigma_x}\right)\left(\frac{y_i-\bar{y}}{\sigma_y}\right)$$

$$= -r_{xy}$$

Problem 7 (8 pts) **Counts towards Midterm 2 redemption score**

Suppose we'd like to fit a multiple linear regression model **without an intercept term** to predict an apartment's monthly rent (in hundreds of dollars) using various features.

For apartment i, the corresponding feature vector is $\vec{x}_i = \begin{bmatrix}\text{bedrooms}_i & K_i & C_i & N_i\end{bmatrix}^T$, where bedrooms_i is the number of bedrooms in apartment i, and K_i, C_i, and N_i are one hot encoded features for the Kerrytown, Central Campus, and North Campus neighborhoods, respectively.

The model is fit by minimizing mean squared error. **All rows of the dataset are shown to the right.** The model's predictions, $h(x_i)$, are shown, along with the true rents, y_i. Several values are missing.

bedrooms_i	neighborhood_i	y_i	$h(x_i)$
4	K	17	(i)
1	C	(ii)	9
3	C	15	13
2	C	10	11
1	N	9	(iii)
4	N	13	(iv)

For instance, the first row of the design matrix is $\begin{bmatrix}4 & 1 & 0 & 0\end{bmatrix}$.

Find all four missing values in the table. Show your work, and write your final answers in the boxes provided. Your answers should be integers with no variables. *Hint: Think about orthogonality.*

$$X = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 4 & 0 & 0 & 1 \end{bmatrix}_{6\times 4}$$

$\vec{e}$ (error) orthogonal to every col of X!

$$\vec{e} = \begin{bmatrix} 0 \\ -1 \\ 2 \\ -1 \\ 9-c \\ 13-d \end{bmatrix}$$

$\vec{x}^{(4)} \cdot \vec{e} = 0 \Rightarrow 9-c+13-d=0 \Rightarrow c+d=22$

$\vec{x}^{(1)} \cdot \vec{e} = 0 \Rightarrow$ another eq'n

$(i) =$ $(ii) =$ $(iii) =$ $(iv) =$

Problem 9 (12 pts)

Consider the matrix $A = \begin{bmatrix} 2 & 3 \\ -4 & k \end{bmatrix}$ where $k \in \mathbb{R}$ is some unknown constant.

a) (3 pts) Suppose $\lambda_1 = 0$ is an eigenvalue of A. Find the value of k. Give your answer as a number with no variables.

$k =$ $\boxed{-6}$

b) (4 pts) Suppose $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A. Find the value of k. Give your answer as a number with no variables.

$$\begin{bmatrix} 2 & 3 \\ -4 & k \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -4+k \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$-4+k=5$$

$k =$ $\boxed{9}$

c) (5 pts) Suppose $\lambda_1 = 3$ is an eigenvalue of A. Find λ_2, the **other eigenvalue** of A. Show your work, and write your final answer in the box provided. Give your answer as a number with no variables.

$$A = \begin{bmatrix} 2 & 3 \\ -4 & k \end{bmatrix} \qquad \lambda_1 = 3, \ \lambda_2 = ?$$

trace: $2+k = 3+\lambda_2 \Rightarrow k = 1+\lambda_2$

det: $2k+12 = 3\lambda_2$

$$2(1+\lambda_2)+12 = 3\lambda_2$$

$$2+2\lambda_2+12 = 3\lambda_2$$

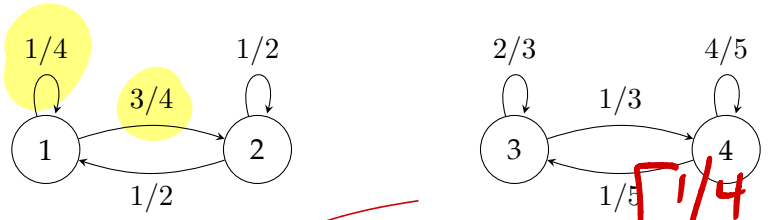
$$14 = \lambda_2$$

$\lambda_2 =$ $\boxed{14}$

uniqname: _____

Problem 10 (14 pts)

The state diagram below describes a Markov chain with four states.



a) (4 pts) Find the adjacency matrix A for this Markov chain.

$$A = \begin{bmatrix} 1/4 & 1/2 & 0 & 0 \\ 3/4 & 1/2 & 0 & 0 \\ 0 & 0 & 2/3 & 1/5 \\ 0 & 0 & 1/3 & 4/5 \end{bmatrix}$$

$\rightarrow 1$ $\rightarrow 2$ $\rightarrow 3$ $\rightarrow 4$

$1 \rightarrow$ $2 \rightarrow$ $3 \rightarrow$ $4 \rightarrow$

$$\begin{bmatrix} 1/4 \\ 3/4 \end{bmatrix} \quad \begin{bmatrix} 1/2 & \\ 1/2 & \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

$\frac{1}{4}a + \frac{1}{2}b = a$

$\frac{1}{2}b = \frac{3}{4}a$

$b = \frac{3}{2}a$

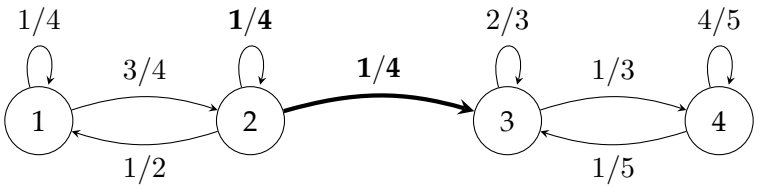
$a = 2, \; b = 3$

b) (6 pts) Suppose the chain starts in **state 1**. Fill each box with the **long-run fraction** of time spent in each state. Your answers should be numbers with no variables, and should sum to 1.

State 1: 2/5 State 2: 3/5 State 3: 0 State 4: 0

$$\begin{bmatrix} 2/5 \\ 3/5 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 3/8 \\ 5/8 \end{bmatrix}$$

Now, consider a **modified** version of the Markov chain. Changes have been emphasized in **bold**.



c) (4 pts) Consider the statement: "'If we start in ____, the long-run fraction of time spent in each state is the same as in the original chain."

Which of the following could be placed in the blank to make the statement true? **Select all** that apply.

☐ state 1 ☐ state 2 ☑ state 3 ☑ state 4 ☐ none of these are valid

Problem 11 (10 pts)

Let S be a 3×3 **symmetric** matrix with eigenvectors $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ corresponding to eigenvalues 5, 2, and -1, respectively. Assume that each $\vec{v}_i$ is a unit vector.

Suppose $\vec{x} \in \mathbb{R}^3$ and that

$$\vec{x} = 3\vec{v}_1 - 4\vec{v}_2 + \vec{v}_3 \Rightarrow$$

$$S\vec{x} = S(3\vec{v}_1 - 4\vec{v}_2 + \vec{v}_3) = 3S\vec{v}_1 - 4S\vec{v}_2 + S\vec{v}_3$$

a) (6 pts) Write $S^2\vec{x}$ as a linear combination of $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$. Fill in each box with a number with no variables.

$$S^2\vec{x} = \boxed{75}\vec{v}_1 + \boxed{-16}\vec{v}_2 + \boxed{1}\vec{v}_3$$

$$= 3(5\vec{v}_1) - 4(2\vec{v}_2) + (-\vec{v}_3)$$

$$= 15\vec{v}_1 - 8\vec{v}_2 - \vec{v}_3$$

$$S^2\vec{x} = 15(5\vec{v}_1) - 8(2\vec{v}_2) - (-\vec{v}_3)$$

$$= 75\vec{v}_1 - 16\vec{v}_2 + \vec{v}_3$$

b) (4 pts) What is the value of $\|S\vec{x}\|^2$?

○ 24 ○ 26 ○ 218 ● 290 ○ 5882 ○ Not enough information

$$S\vec{v}_1 = 5\vec{v}_1$$
$$S\vec{v}_2 = 2\vec{v}_2$$
$$S\vec{v}_3 = -\vec{v}_3$$

$$\|S\vec{x}\|^2 = (S\vec{x})^T(S\vec{x}) = \vec{x}^T S^T S\vec{x} = \vec{x}^T S^2 \vec{x}$$

$$S = V\Lambda V^T$$
$$S^2 = V\Lambda V^T V\Lambda V^T = V\Lambda^2 V^T$$

$$= \vec{x}^T V\Lambda^2 V^{-1}\vec{x}$$
$$= \vec{x}^T V\Lambda^2 \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$
$$= \vec{x}^T V \begin{bmatrix} 25 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$$\vec{x} = 3\vec{v}_1 - 4\vec{v}_2 + \vec{v}_3 = \underbrace{V}_{\text{matrix}} \underbrace{\begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}}_{\text{vector}}$$

$$V = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$$

$$\vec{x} = V \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$$V^{-1}\vec{x} = V^{-1} V \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$$V^{-1}\vec{x} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$V^{-1}\vec{x}$?

$$\Lambda^2 = \begin{bmatrix} 5^2 & & \\ & 2^2 & \\ & & (-1)^2 \end{bmatrix} = \begin{bmatrix} 25 & & \\ & 4 & \\ & & 1 \end{bmatrix}$$

$$\|S\vec{x}\|^2 = (S\vec{x})^T(S\vec{x}) = \vec{x}^T S^T S \vec{x}$$

$$= \vec{x}^T S^2 \vec{x}$$

$$= \vec{x}^T V \Lambda^2 V^T \vec{x}$$

$$= \vec{x}^T V \Lambda^2 \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$$= \vec{x}^T V \begin{bmatrix} 25 & & \\ & 4 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$$

$$= \vec{x}^T V \begin{bmatrix} 75 \\ -16 \\ 1 \end{bmatrix}$$

$$= (V^T\vec{x})^T \begin{bmatrix} 75 \\ -16 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}^T \begin{bmatrix} 75 \\ -16 \\ 1 \end{bmatrix} = 75 \cdot 3 + 64 + 1$$

$$= 225 + 64 + 1 = 290$$

12, 11, 10, 7, 2

uniqname: _____

Problem 12 (12 pts)

Suppose $\tilde{X}$ is an $n \times 2$ matrix whose columns are mean-centered (i.e. have a mean of 0). Furthermore, suppose

$$\tilde{X}^T\tilde{X} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$$

~ cols of V in $\tilde{X} = U\Sigma V^T$ are eigenvectors of $\tilde{X}^T\tilde{X}$

Note that $\tilde{X}^T\tilde{X}$ has eigenvalues of 7 and 2. Let $\tilde{X} = U\Sigma V^T$ be the singular value decomposition of $\tilde{X}$, and let $\vec{v}_1$ be the first column of V (not V^T).

a) (4 pts) What is $\vec{v}_1$? Give your answer as a vector with no variables. If there are multiple correct answers, you only need to provide one.

$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$

$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 7a \\ 7b \end{bmatrix}$ $\quad \begin{bmatrix} 1 \\ 2 \end{bmatrix} \to \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$

$3a + 2b = 7a$

$2b = 4a$

$b = 2a$

b) (3 pts) Suppose the variance of the **second** principal component is $1/15$. What is n, the number of rows in $\tilde{X}$? Give your answer as a number with no variables.

$n =$ 30

$\frac{\sigma_2^2}{n} = \frac{1}{15} \Rightarrow \frac{(\sqrt{2})^2}{n} = \frac{1}{15}$ $\quad \sigma_2 = \sqrt{2}$ $\quad n = 30$

c) (5 pts) Suppose that $\vec{u}_2$ is the second column of U, corresponding to the singular value σ_2, in the singular value decomposition of $\tilde{X}$.
Prove that $\tilde{X}\vec{v}_1$ and $\sigma_2\vec{u}_2$ are orthogonal. You do not need to re-prove any facts about the singular value decomposition, but you should state any facts you use.

$$(\tilde{X}\vec{v}_1) \cdot (\sigma_2\vec{u}_2)$$
$$= (\tilde{X}\vec{v}_1)^T(\sigma_2\vec{u}_2)$$
$$= \vec{v}_1^T\tilde{X}^T\sigma_2\vec{u}_2$$
$$= \sigma_2\vec{v}_1^T\tilde{X}^T\vec{u}_2$$

$\vec{u}_2^T\tilde{X}\vec{v}_2 = \vec{u}_2\sigma_2\vec{u}_2$

$(\vec{u}_2^T\tilde{X}\vec{v}_2)^T = \sigma_2$

$\vec{v}_2^T\tilde{X}^T\vec{u}_2 = \sigma_2$

$$(\tilde{\chi}\vec{v}_1) \cdot (\sigma_2 \vec{u}_2)$$

$$= (\sigma_1 \vec{u}_1) \cdot (\sigma_2 \vec{u}_2)$$

$$= \sigma_1 \sigma_2 (\vec{u}_1 \cdot \vec{u}_2) \quad 0$$

$$= 0$$