

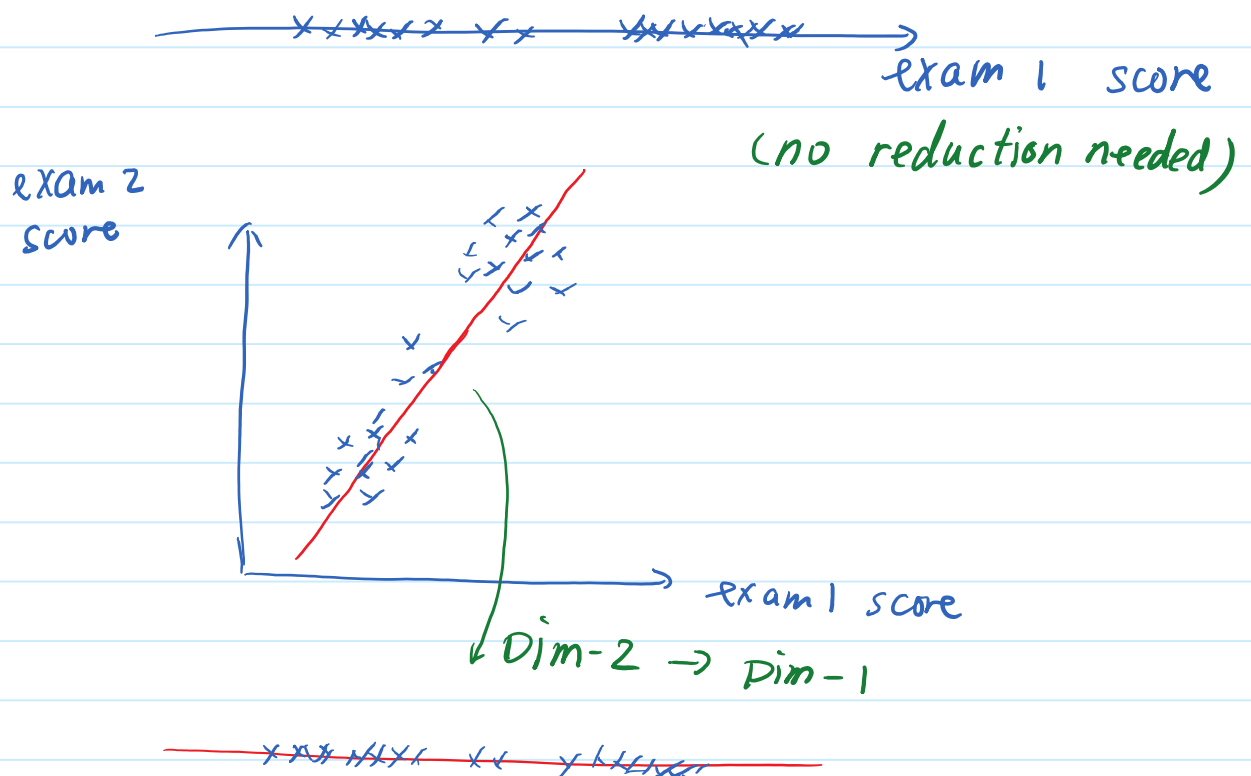
## Lab 14

In this Lab:

Principal Component Analysis (PCA)

---

### Dimensionality Reduction



In ECE 314, we have 14 labs for each student.

	Lab1	Lab2	- - - - -	Lab14
student 1				
student 2				
⋮				
⋮				
⋮				
⋮				

Each student's performance is represented by a point in  $\mathbb{R}^{14}$ .

How can we visualize the students' performance?



# PCA

Big goal: reduce the dimension of the data while preserving the most important features of the data.

random vector  $X = (X_1, \dots, X_{14})^T \rightarrow$  column vector  
 $\swarrow$  Lab 1  $\searrow$  Lab 14.

covariance matrix

$$\text{cov}(X) = \mathbb{E}[(X - \mathbb{E}X)(X - \mathbb{E}X)^T]$$

$$= \begin{bmatrix} \text{cov}(X_1, X_1) & \dots & \text{cov}(X_1, X_{14}) \\ \vdots & & \vdots \\ \text{cov}(X_{14}, X_1) & \dots & \text{cov}(X_{14}, X_{14}) \end{bmatrix}$$

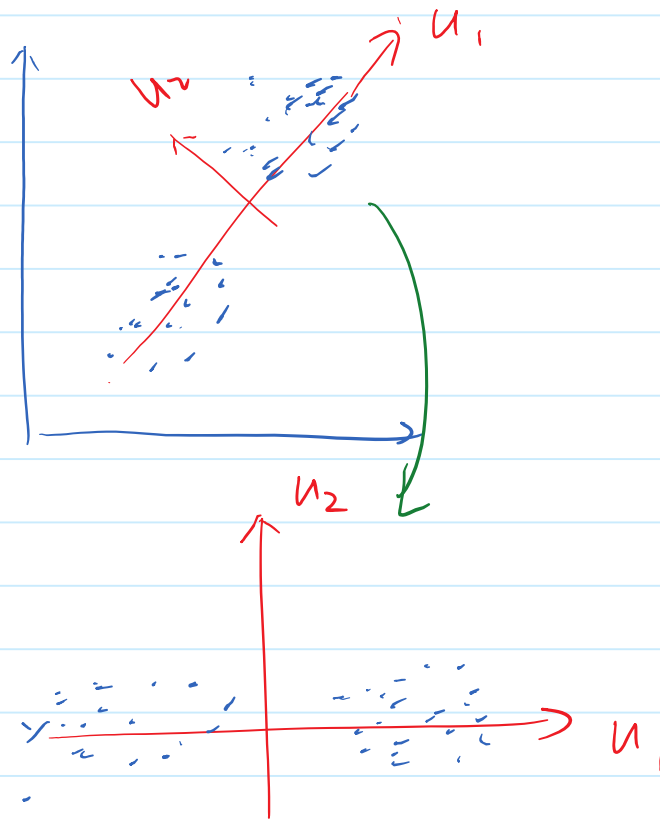
$$= U \Lambda U^T$$

$$= \begin{bmatrix} u_1 & \dots & u_{14} \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{14} \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_{14}^T \end{bmatrix}$$

$\swarrow$  column vector

$$\text{let } y = U^T x \\ = \begin{bmatrix} u_1^T \\ \vdots \\ u_{14}^T \end{bmatrix} x$$

$(u_1, u_2, \dots, u_{14}) \rightarrow$  new coordination system.



Then we throw away the data on  $u_2$  direction.