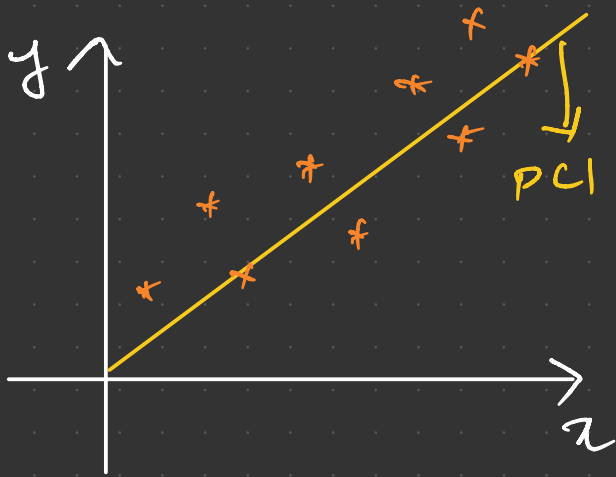


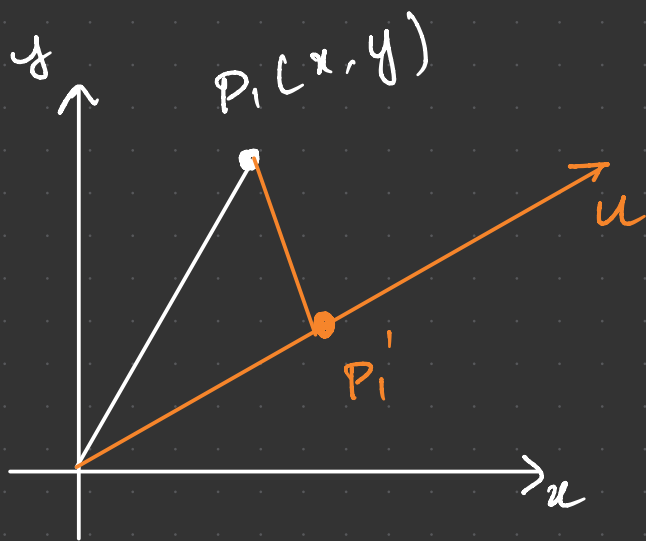
# Math Intuition behind PCA Algorithm



2D  $\rightarrow$  1D

① Projection

② cost function  $\rightarrow$  variance.



$$\text{Projection}_{P_i} u = \frac{P_i \cdot u}{\|u\|}$$

$\|u\| = 1 \rightarrow$  for unit vector,

therefore,

$$\text{Projection}_{P_i} u = P_i \cdot u$$

$\Downarrow$

Scalar value.

$\Downarrow$

variance

$$\text{Projection}_{P_1} u = P_1 \cdot u$$

$\Downarrow$

scalar value of each projection

$$p_0', p_1', p_2', p_3' \dots p_n'$$

$\Downarrow$

$$x_0', x_1', x_2', x_3', x_4' \dots x_n'$$

$$\text{variance} = \frac{\sum_{i=1}^n (x_i' - \bar{x})^2}{n} \Rightarrow \text{cost function}$$

$\Downarrow$

Goal is to find best unit vector which captures maximum variance.

# Eigen Vectors and Eigen Values

- ① Identify covariance between features.
- ② Eigen vectors and Eigen values will found out from this covariance matrix
- ③ Eigen vector where the Eigen value is high

→ Eigen values are the magnitude of the Eigen vector

→ this will capture the maximum variance

Eigen vectors and Eigen values can be found by

$$\boxed{A\vec{v} = \lambda\vec{v}} \rightarrow \text{linear transformation of matrix}$$