Principal Component Analysis. PCA.

given a dataset $X_{(N \times D)}$

N - rows - datpoints.

D - columns - dimensions.

$x_i \in R^D$

we want to transform

$X_{(N \times D)} \rightarrow Z_{(N \times d)}$

such that $d \ll D$

while performing transformation, lets preserve variance in the dataset.

→ perform mean centering on the dataset

Say there is a unit vector $\hat{u}_{(D \times 1)}$ such that the variance across $\hat{u}$ is maximum.

we want to compute $\hat{u}$

Variance ⇒ projecting x across $\hat{u}$

$$= \frac{1}{N} \sum_{i=1}^{N} (x_i^T \hat{u})^2$$

(mean centered x)

$$= \hat{u}^T \left(\frac{1}{N} \sum x_i x_i^T \right) \hat{u}$$

$$\frac{1}{N} \sum_{i=1}^{N} x_i x_i^T = S =$$ covariance matrix of dataset X.
mean centered.

we want to maximize $\hat{u}^T S \hat{u}$ such that
$\hat{u}^T u = 1$ (unit vector).

→ adding this as lagrangian constraint

$$\hat{u}^T S \hat{u} - \lambda (u^T u - 1). \Rightarrow \text{maximize.}$$

∴ differentiating wrt $\hat{u}$ ⇒.

$$2 S \hat{u} - 2\lambda \hat{u} = 0.$$

$$S \hat{u} = \lambda \hat{u}.$$

this is of form eigen transformation analysis.

∴ $\hat{u}$ is eigen vectors of S
eigen values $= \lambda$

when $S\hat{u} = \lambda \hat{u}$.

for $\hat{u}$ that maximizes variance,

$$\text{variance} = \hat{u}^T S \hat{u} = \hat{u}^T \lambda \hat{u} = \lambda \hat{u}^T \hat{u}$$

$= \lambda =$ max variance across $\hat{u}$ eigen vector