

Q How to choose optimum no of principle components?

Ans. We know that each Principle components ~~have~~ are eigen vectors and each of them have its respective eigen value.

• These each eigen values tells how much variance of data is explained by their respective eigen vector

Ex → 784 dimensions → Applied PCA → 784 Principle components & 784 eigen values.

EV = $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_{784}$

• let's say $\lambda_1 = \text{any value}$.

• Convert λ_1 value into percentage, i.e. express how much variance is explained by λ_1 in percentage

Var Explained by $\lambda_1 \rightarrow = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_{784}} = \text{let's say} = 30\%$

• for $\lambda_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_{784}} = 25\%$

• Similarly, you have to find percentage for other λ 's until you minimums

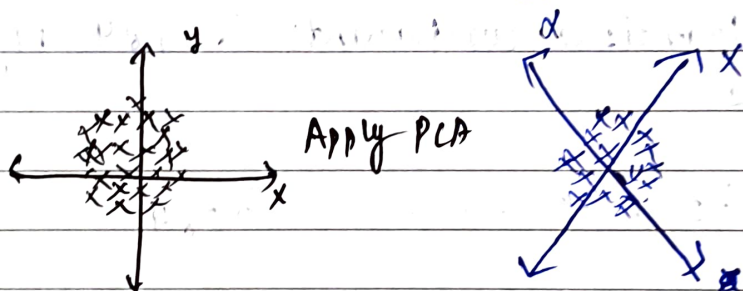
explain 90% variance.

Ex That means, you will select that many principal components for which the total variance explained is minimum 90%.

Q When PCA doesn't work

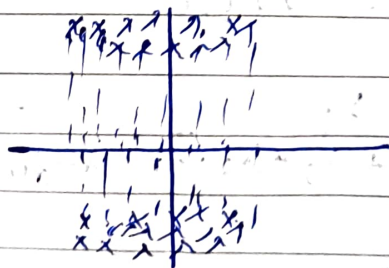
- There will be times, when data won't allow PCA to work

Ex1 : Data is circle



The variance on x axis & y axis, is still same even after applying PCA

Ex2:



Ex3 : • When there is a pattern in higher dimension
• Ex $\rightarrow y = x^2$
• In higher dimension, you may lose this information.

