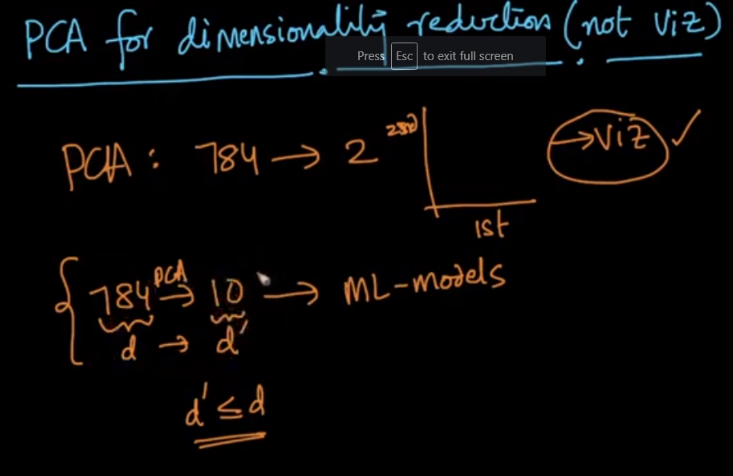
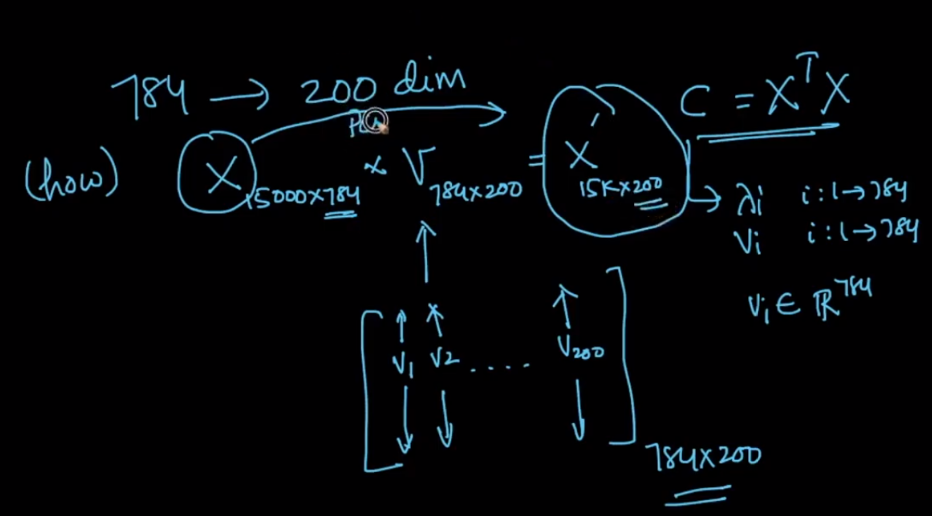
This can be says as a summary of PCA, how we do it for dimensionality reduction.

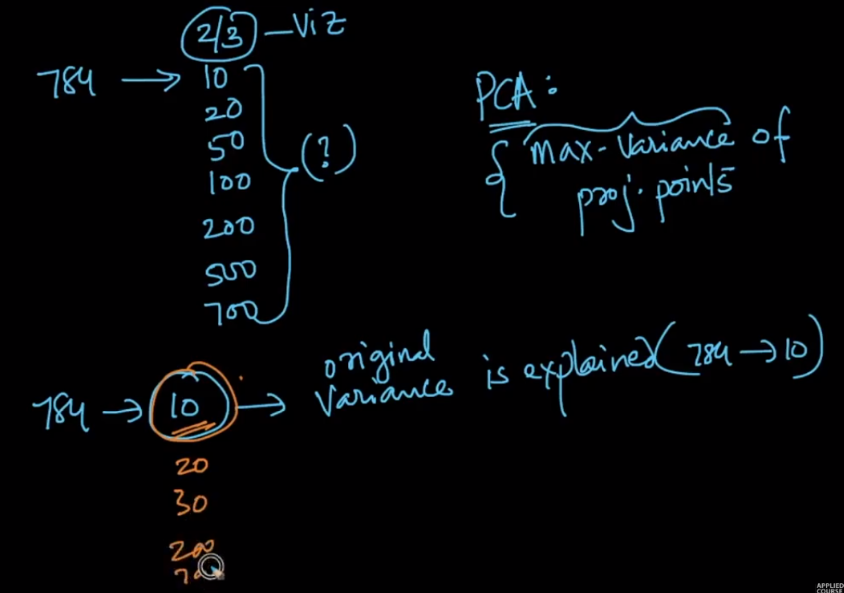
Our ultimate aim is to obtain features **d’** such that **d’ <= d**, where d is original no. of features.

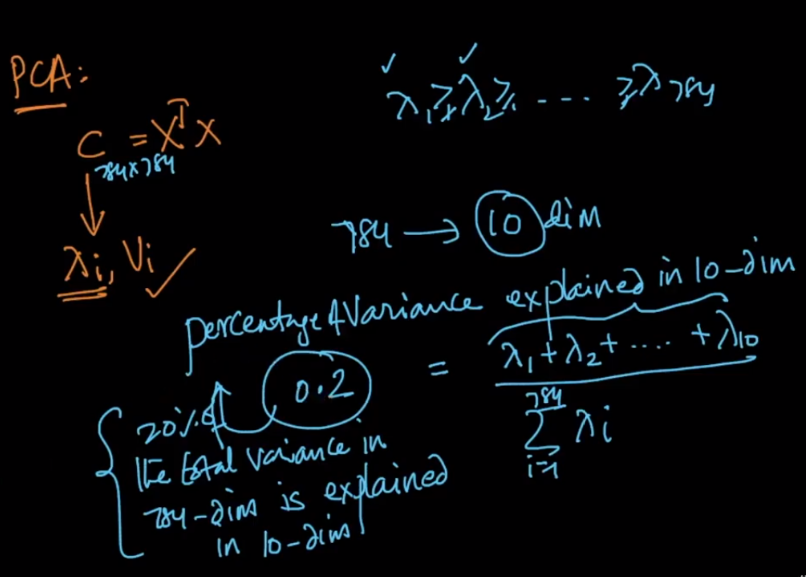


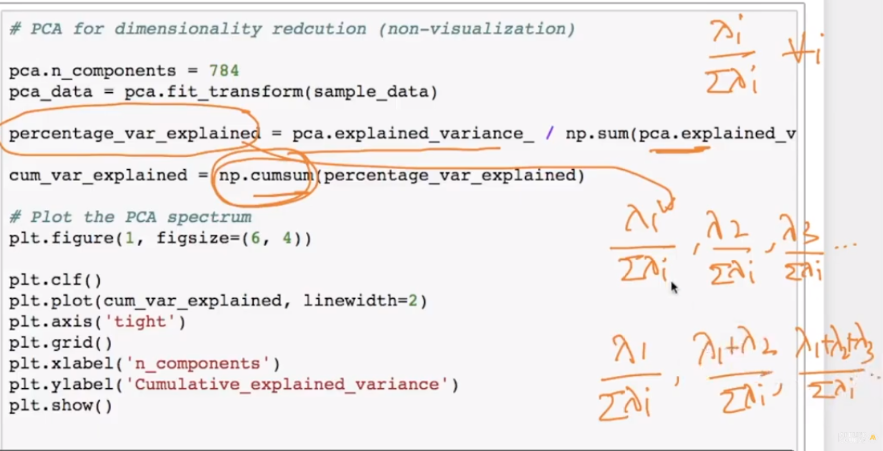
Let’s say for mnist where we are obtaining 200 features from 784 features, here we have X of dimension 15000 \* 784 and the eigen vectors matrix of dimension 784 \* 200, which results the final matrix X’ of dimension 15k \* 200

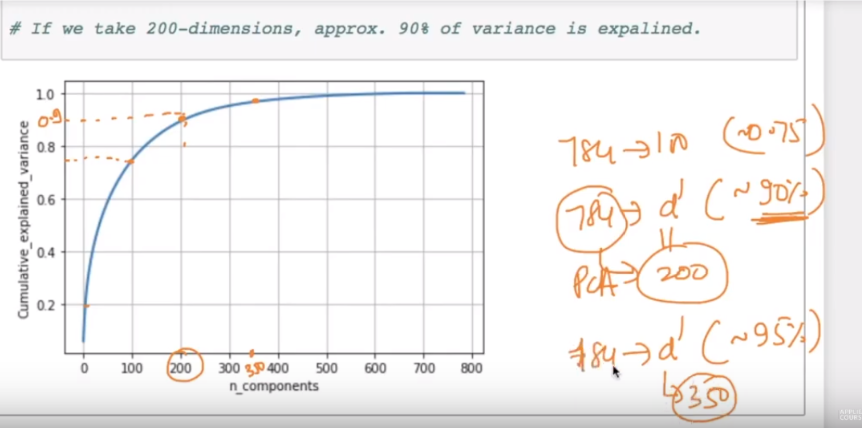


So our ultimate aim is to find the no of features for which the information retain percentage Is what we want which may b 75%, 99% and how it’s calculated is given in below image.









As we can see for 100 components we have 75%, for 200 we have 90% and for 300 components we have 95%.