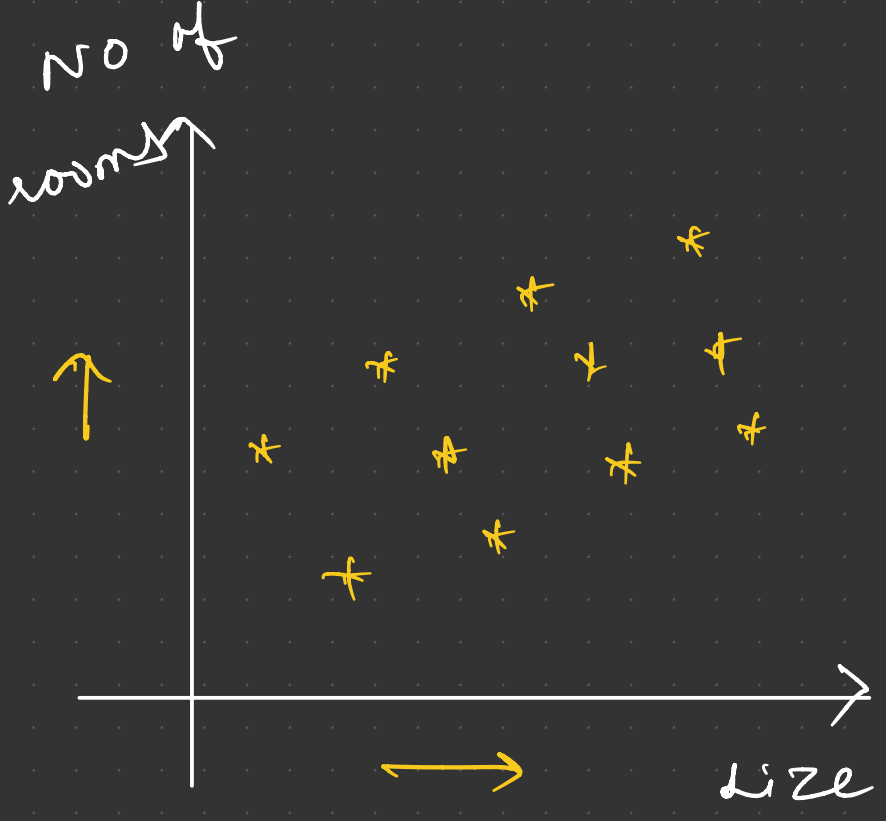
PCA Geometric Intution

Housing Dataset

house size	No. of Rooms	Price

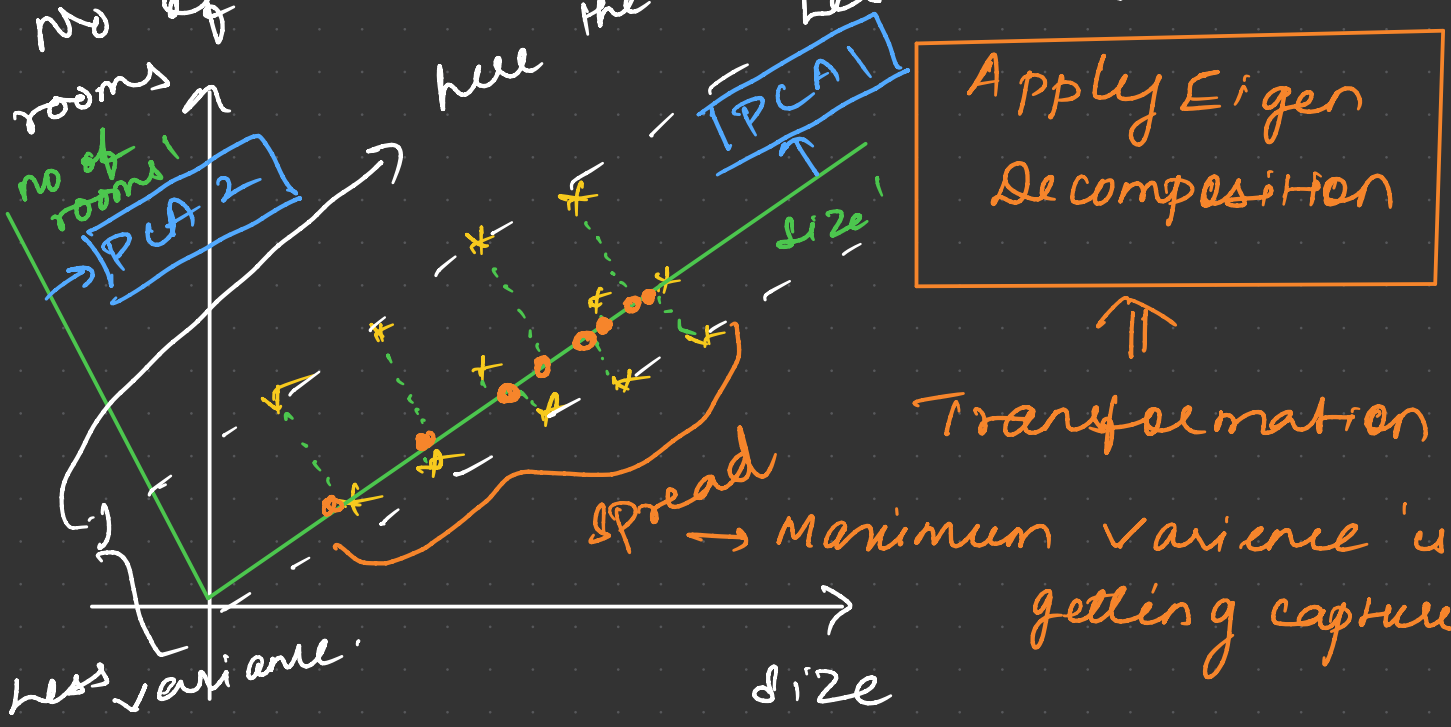


with PCA → I want to reduce 2 dimension to 1 dimension

If we drop any of the information/ feature, then we will end up with loss of information. in our case size of the house and no. of rooms are two essential feature to predict the price of the house. If we remove any of these feature, then our model may not perform really well, thus the accuracy would be inconsistent or would also result in least accurate.

Whenever we work on feature extraction

→ we should ensure that loss of information is very minimal or zero. The more the variance is less, loss of information is minimal.



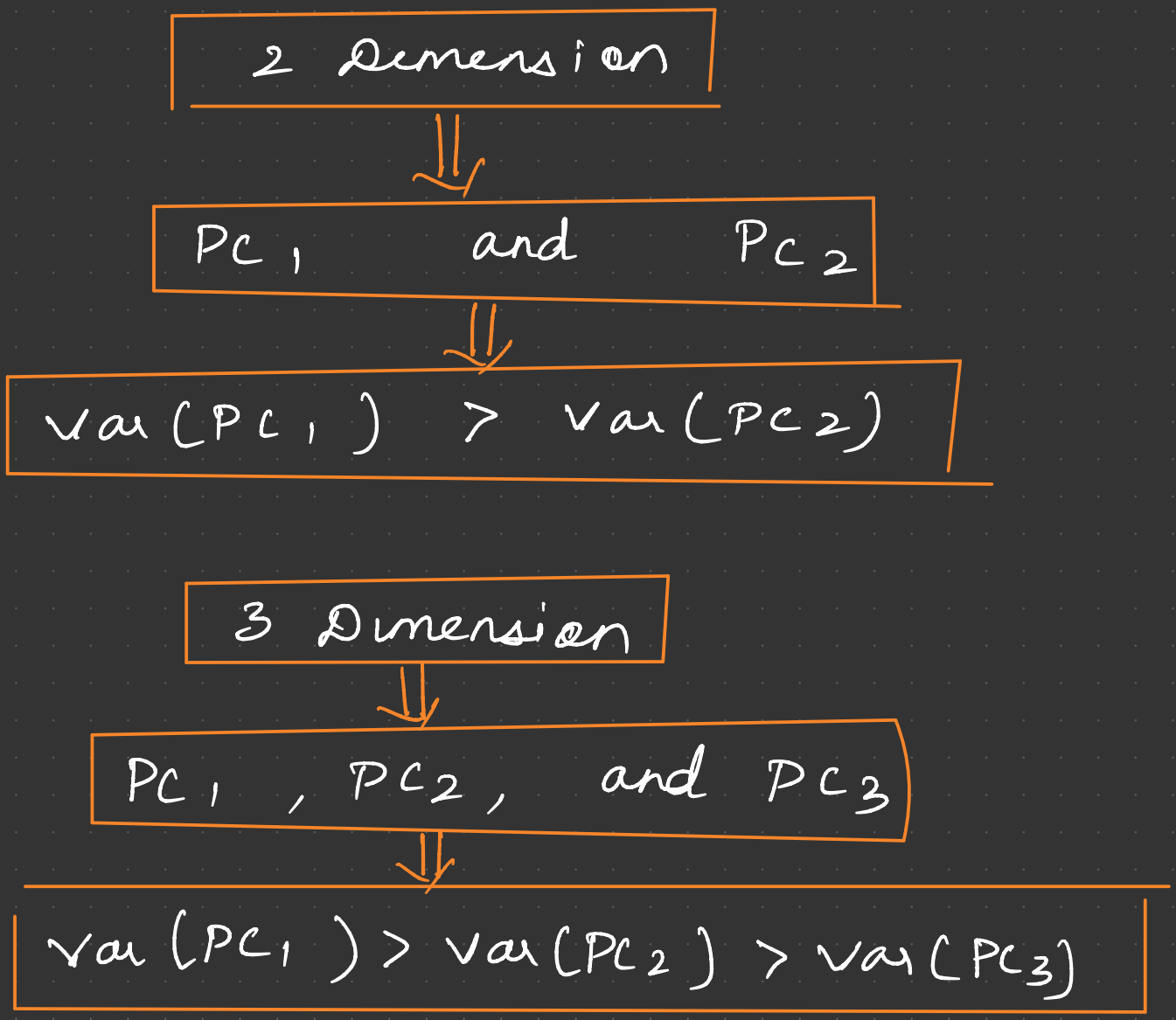
here we had converted

2D → 1D

ensuring much information are not Lost.

Always, PCA 1, will capture the maximum amount of variance, and PCA 2 will have the next maximum amount of variance and so on..

the main of PCA is to capture the maximum variance or spread.



we say, its is best PC, only when, the spread has the maximum variance.

To get the best principle component which captures maximum variance.