M2L2

1. Standardize the Cov matrix
(Normalize the data)

2. Computing the Cov matrix
relation b/w 2 features (totally diff)
on the standardized data. fit transform fn.

For how many features we want to consider

3. E.values & vectors — how much var is explained along a vector.
calculate using n (nxn) matrix
How well are the datasets are related along the line.
find PC's (like S) multiply with Y ; to effectively reduce dim
how to take a diff transform such that it leads to different subspace.
we need to find a transformation
regression: squares are min
Classification:

Direction along highest & lowest variance.
1st E.vector → Dirn of highest variance.

Compute Cov matrix for all features

Cumulative var: Ex: 30 vectors / components
If we take all 30 PC's ⇒ Original - max var: 100 %.

If we truncate some components
⇒ Ex 15 vectors that can explain the original at 99 %
Implying we don't need to consider other 15 vectors

5. Reducing the dimensions

Y → Eigen vectors

Take all the feature labels

In 3D :

all the vectors get multiplied get decomposed along a line.

dot product along x direction

Ex: Take 100 % VAV

we have a vector along [1,2] to be decomposed into 2 vectors.

~~[1/2/3]~~ x [1,0] + x [0,1]

Basis vectors for ~~original subspace.~~ transformation is same as PC components multiplied with matrix.

defined subspace with basis vectors