

2.) Let X be a $n \times d$ matrix (where n is samples & d is the no. of features)

Assuming, mean is subtracted from X .

$$\text{Covariance matrix} = C = X^T X / n.$$

Now, it is a positive definite matrix and hence diagonalisable.

$C = V L V^T$ where V is matrix of eigenvectors & L is diagonal matrix of eigenvalues.

Here, XV is the new compressed data.

$$\text{Now, using SVD, } X = U D V^T, \quad C = V D U^T U D V^T / n \\ = \frac{V D^2 V^T}{n}$$

Now, comparing it with $V L V^T$, we can intuitively find that V is the matrix of principal components / ~~eigenvalue~~ vectors.

$$\& L = \frac{D^2}{n} \quad \text{and compressed data is } XV = VD.$$

So, this is how we can use the SVD of the data matrix to do dimensionality reduction.