SHORT NOTES (1)

PCA

(A.) EVD (Eigenvalue Decomposition)

$$\boxed{A = V \Lambda V^{-1}}$$

A = original square matrix

V = matrix whose columns are eigenvectors of A

Λ = diagonal matrix whose diagonal elements are eigenvalues of A.

(B.) SVD (Singular Value Decomposition)

$A = U \Sigma V^T$ $A^{m\times n} \rightarrow$ rectangular matrix

$\hookrightarrow AV = \sigma U$

V = matrix whose columns are eigenvectors of ~~AA~~ A^TA

U = matrix whose columns are eigenvectors of AA^T

V^T = transpose of V i.e. matrix whose rows are eigenvectors of A^TA.

$$[A]_{m\times n} = \underbrace{\begin{bmatrix} \uparrow & & \uparrow \\ v_1 & \cdots & v_k \\ \downarrow & & \downarrow \end{bmatrix}_{n\times k}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix}_{k\times k}}_{\Sigma} \begin{bmatrix} \leftarrow v_1 \rightarrow \\ \vdots \\ \leftarrow v_k \rightarrow \end{bmatrix}_{k\times n}$$

→ we pick only top k-large eigenvectors only for PCA.

In PCA we want:-

(1.) We want to represent the data using ~~lesser~~ fewer dimensions.

(2.) The data has high variance along these dimensions.

(3.) The dimensions has ~~no~~ minimum correlation.

Thm 6:- $X^{m \times n}$ is a matrix whose columns are zero mean then $\Sigma = \frac{1}{m} X^T X$ is the covariance matrix where, each (i, j)th element stores the covariance b/w columns i and j of X.

$$\mathrm{Cov}(X, Y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$\bar{x}$ = mean of feature x
$\bar{y}$ = " " y
n = no. of data pts

PCA FORMULA

To perform PCA on X $\quad P^T \Sigma P = D$

Step-1:- $\Sigma = X^T X$

Step-2:- P = matrix containing eigen vectors of Σ
D = diagonal matrix having eigenvalues of Σ

Step-3:- $\boxed{\hat{X} = XP}$