Assignment 2

Roll number: 2021701015

Question 1   Eigen values & Eigen Vectors

A) Among Eigen value decomposition and Singular Value Decomposition, which is more generalizable to matrices & why?

Ans: The answer to the presented question could be that Singular Value Decomposition is more generalizable to matrices than Eigen value decomposition due to the following reasons:

(i) SVD can be applied to any matrix even if it is rectangular. But eigenvalue decomposition can work only for square matrix.

As the question asks for a method more generalizable, SVD can be considered more generalizable.

(ii) The vectors obtained in the eigendecomposition matrix are not necessarily orthogonal, i.e the change of basis is not just simple rotation. whereas the vectors in the matrices U & V of SVD are orthonormal and represent rotations.

(iii) It is seen that entries of D in eigenvalue decomposition can be complex number i.e +ve, -ve or imaginary. & whereas entries of diagonal matrix Σ are all real & non negative numbers.

(iv) Sometimes even for certain square matrices eigendecomposition does not exist

B) Show the method and find the Singular Value Decomposition of the following matrix : ②

$$M = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \qquad M^T = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

First find $MM^T = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$

$$\begin{bmatrix} 80-\lambda & 100 & 40 \\ 100 & 170-\lambda & 140 \\ 40 & 140 & 200-\lambda \end{bmatrix} \qquad \lambda_1 = 360 \quad \lambda_2 = 90 \quad \lambda_3 = 0$$

Find respective eigen vectors $v_1 = \begin{bmatrix} 1/2 \\ 1 \\ 1 \end{bmatrix}$ $v_2 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$ $v_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$

$\sigma_1 = 6\sqrt{10}$ $\sigma_2 = 3\sqrt{10}$ are the square root of eigen values.

$\therefore$ From σ_1 & σ_2 we obtain $\Sigma = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \\ 0 & 0 \end{bmatrix}$

U can be found by $U = \{u_1, u_2, u_3\}$ & $u_1 = \frac{1}{\sigma_1} M v_1$ similarly u_2 & u_3.

$\therefore u_1 = \frac{1}{6\sqrt{10}} \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}\begin{bmatrix} 1/2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$ similarly $u_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$

$u_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$

$v_1 = \frac{1}{\sigma_1} \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}^T \cdot u_1 = \begin{bmatrix} \frac{3\sqrt{10}}{10} \\ \frac{\sqrt{10}}{10} \end{bmatrix}$ & $v_2 = \begin{bmatrix} \frac{\sqrt{10}}{10} \\ \frac{-3\sqrt{10}}{10} \end{bmatrix}$

Question 2 LDA & PCA

[A] Suppose you want to apply PCA to your data X which is in 2D and you decompose X as VDV^T. Then, which of the following are correct:

(a) PCA can be useful if all elements of D are equal (True)
(b) PCA can be useful if all elements of D are not equal (False)
(c) D is not full-rank if all points in X lie on a straight line (True)
(d) V is not full-rank if all points in X lie on a straight line (True)
(e) D is not full-rank if all points in X lie on a circle (False)

[B] True or false.

PCA will project the datapoints (multiclass) on a line which preserve information useful for data classification.

Ans: False. PCA never consider class information. It simply projects all the data point irrespective of their classes on the line along which, ~~lasos~~ there is maximum variance.

Question 3 Bayes theorem.

[A]. What is the difference between prior & posterior probabilities?

Prior probability represents the ~~[illegible]~~ original belief whereas the posterior probability takes new information into consideration.

A prior probability is the probability that an observation will belong into a group before we collect the data. And a posterior probability is the probability of assigning observations to group the given data.

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

$P(A)$ → Prior Probability (Prob of A in general).

$P(A|B)$ → Posterior probability (Prob of A after event B has been observed)

Date :

[B] Lets say that you are at your work one day and have just finished lunch. You suddenly feel horrible and find yourself lying down. Maybe it is because one of your friend was recently sick with flu. You have a headache and sore throat, & you know that people with flu have the same symptoms roughly 90% of time. In other words, 90% of people with the flu ~~dott~~ have the same symptoms you currently have.

Wanting to gain a little more information you roll over, grab your phone & search google. You find a reputable article that says that only 5% of the population will get flu in a given year. Or prob of having flu in general is only 5%. You then spot one more statistic that says 20% of the population in a given year will have a headache & sore throat at any given time.

What is the probability of you having a flu given you have a sore throat and a headache?

$$P(F/HS) = \frac{P(HS/F)\,P(F)}{P(HS)}$$

$P(HS/F) = \frac{90}{100}$ $\quad P(F) = \frac{5}{100}$ $\quad P(HS) = \frac{20}{100}$

$$P(F/HS) = \frac{\frac{90}{\cancel{100}} \times \frac{5}{\cancel{100}}}{\frac{20}{\cancel{100}}} = \frac{9\cancel{0} \times 5}{2\cancel{0}} =$$

$$= \frac{0.9 \times 0.05}{0.2} = 0.225\ \%$$

Please consider answer in Pencil