SMAI assignment

① a)

Defn. of eigenvalue decomposition

Let 'A' be a square $n \times n$ matrix with 'n' linearly independent eigenvectors q_i (i is from 1 to n).

Then, A can be factorized as

$$A = Q \Lambda Q^{-1}$$

where Q is a $n \times n$ square matrix whose i^{th} column is the eigenvector q_i of A and Λ is the diagonal matrix whose elements $\Lambda_{ii} = \lambda_i$.

Defn. of SVD

SVD of a $m \times n$ complex matrix M is a factorization of the form $U \Sigma V^*$ where U is a $m \times m$ complex unitary matrix, Σ is a $m \times n$ rectangular diagonal matrix and V is a $n \times n$ complex unitary matrix.

Also, if M is real, one can guarantee U and V to be real orthogonal matrices.

The SVD is more generalizable as SVD will always exist for any sort of rectangular or square matrix, whereas eigendecomposition can exist only for square matrices, and even among ~~matrices, eigen~~ some square matrices, eigenvalue decomposition may not be possible.

For eg. eigenvalue decomposition is not possible for:-

$$C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

SMAI assignment

① a)

Defn. of eigenvalue decomposition

Let 'A' be a square $n \times n$ matrix with 'n' linearly independent eigenvectors q_i (i is from 1 to n).

Then, A can be factorized as

$$A = Q \Lambda Q^{-1}$$

where Q is a $n \times n$ square matrix whose i^{th} column is the eigenvector of q_i of A and Λ is the diagonal matrix whose elements $\Lambda_{ii} = \lambda_i$.

Defn. of SVD

SVD of a $m \times n$ complex matrix M is a factorization of the form $U \Sigma V^*$ where U is a $m \times m$ complex unitary matrix, Σ is a $m \times n$ rectangular diagonal matrix and V is a $n \times n$ complex unitary matrix.
Also, if M is real, one can guarantee U and V to be real orthogonal matrices.

The SVD is more generalizable as SVD will always exist for any sort of rectangular or square matrix, whereas eigendecomposition can exist only for square matrices, and even among ~~matrix, eigen~~ some square matrices ~~square~~, eigenvalue decomposition may not be possible.

Eg. eigenvalue decomposition is not possible for:-

$$C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

b) $M = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}$ | Also, let $M = UDV^T$ — (1)

orthogonal

orthogonal

$\therefore MM^T = \underset{3\times2}{\begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}} \underset{2\times3}{\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}}$

$= \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$ — (2)

~~From (1)~~

$\cancel{MM^T = UDV^T(UDV^T)^T = UDV^T(V^T)^T D^T U^T = UD(V^TV)D^TU^T = UDD^TU^T}$

$M^TM = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} = \begin{bmatrix} 333 & 81 \\ 81 & 117 \end{bmatrix}$ — (3)

<u>To find eigenvalues of M^TM</u>

$\det(M^TM - \lambda I) = \det\left(\begin{bmatrix} 333-\lambda & 81 \\ 81 & 117-\lambda \end{bmatrix}\right) = 0$

$\Rightarrow (\lambda - 117)(\lambda - 333) - 81^2 = 0$

$\Rightarrow \lambda = 90$ and $\lambda = 360$

$MM^T = UDD^TU^T$ (since U and V are orthonormal $\Rightarrow VV^T = I$)

Calculating eigenvalues

$\det(MM^T - \lambda I) = 0$

$$\Rightarrow \begin{vmatrix} 80-\lambda & 100 & 40 \\ 100 & 170-\lambda & 140 \\ 40 & 140 & 200-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (80-\lambda)\left[(\lambda-170)(\lambda-200) - (140)^2\right]$$
$$- 100\left[(100)(200-\lambda) - 40\times 140\right]$$
$$+ 40\left[100\times 140 - (40)(170-\lambda)\right] = 0$$

on simplifying,

$$\left.\begin{aligned} &\Rightarrow -\lambda(\lambda-90)(\lambda-360) = 0 \\ &\rightarrow \lambda = 0, 90, 360 \end{aligned}\right\} \text{— (4)}$$

Now, finding corresponding eigenvectors

$$\begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\Rightarrow \lambda x = 80x + 100y + 40z \quad \text{— (5)}$$

$$\lambda y = 100x + 170y + 140z \quad \text{— (6)}$$

$$\lambda z = 40x + 140y + 200z \quad \text{— (7)}$$

On solving,
eigenvector e_1 corresponding to λ_1 i.e, 360

is $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ — (10)

eigenvector e_2 corresponding to λ_2 i.e, 90

is $\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ — (11)

eigenvector e_3 corresponding to λ_3 i.e, 0 is $\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$ — (12)

Using (10), (11), (12)
and the fact that $MM^T = UDD^TU^T$ — (13)

$\Rightarrow DD^T$ = diagonal as well.

This means
that eigendecomposition of MM^T should have it's eigenvectors as columns of U. — (15)

~~Let~~ $\cancel{D = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}}$

By wanting

Using (15) and by wanting to keep the vectors of U normalized,

$$U = \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \text{normalized } e_1 & \text{normalized } e_2 & \text{normalized } e_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix}$$

$$\Rightarrow U = \begin{bmatrix} \frac{1}{3} & 2/3 \\ 2/3 & -2/3 \\ 2/3 & 1/3 \end{bmatrix}$$

$$\Rightarrow U = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \quad - (16)$$

Therefore, now, we know by defn that D = rectangular diagonal matrix $\Rightarrow$

$$D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \\ 0 & 0 \end{bmatrix}$$

$$DD^T = \begin{bmatrix} 360 & 0 & 0 \\ 0 & 90 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \end{bmatrix} = \begin{bmatrix} d_1^2 & 0 & 0 \\ 0 & d_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow d_1 = 6\sqrt{10}, \quad d_2 = 3\sqrt{10} \quad - (17)$$

from (3), $m^T m = \begin{bmatrix} 333 & 81 \\ 81 & 117 \end{bmatrix}$

$m^T m = V D^T U^T U D V^T$
$= V D^T D V^T$

($U^T U = I$ since ~~orthogonal~~ it's vectors are orthonormal)

$\det(m^T m - \lambda I) = (333-\lambda)(117-\lambda) = 81^2$
$\Rightarrow \lambda = 360, 90$

Now, columns of V will be normalized versions of corresponding eigenvectors.

$\therefore \begin{bmatrix} 333 & 81 \\ 81 & 117 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

For $\lambda = 360$,

$$\left.\begin{aligned} 333v_1 + 81v_2 &= 360v_1 \\ 81v_1 + 117v_2 &= 360v_2 \end{aligned}\right\} \begin{aligned} &-(18) \\ &-(19) \end{aligned}$$

Using (18) and (19), $v_1 = 3v_2$ } since we want normalized

$\Rightarrow \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix}$ is an eigenvec — (20)

Now, ~~we know that U is orthonormal matrix~~
~~and V~~ is orthonormal matrix.

for $\lambda = 90$,

$$333v_1 + 81v_2 = 90v_1$$
$$\Rightarrow v_2 = -3v_1$$

$\therefore$ Since we want normalized version

$$\Rightarrow \begin{bmatrix} +1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$ is a valid eigenvector. — 21

$$\therefore V = \begin{bmatrix} 3/\sqrt{10} & +1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \quad \text{— 22}$$

Also, I have already proved earlier that

$$D = \begin{bmatrix} 6/\sqrt{10} & 0 \\ 0 & 3/\sqrt{10} \\ 0 & 0 \end{bmatrix}, \quad U = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}$$

A3

a)

Let's say H is a hypothesis.

Then prior probability of our hypothesis being true is P(H).

Now, in case some new evidence (E) is introduced, then we will get some new information, then the probability of H being true post knowing about the evidence i.e, P(H/E) is posterior probability.

The difference between prior and posterior probability is introduced due to the new information we get from a conducted experiment or measurement.

b) Let E_1 be event that someone have flu.
Let E_2 be the event that someone has symptom of headache and sore throat.

As per question,

$P(E_1) = 0.05$ — ①

$P(E_2) = 0.2$ — ②

Also, $P(E_2|E_1) = 0.9$ — ③

Using Bayes formula,

$$P(E_1|E_2) = \frac{P(E_2|E_1) \times P(E_1)}{P(E_2)}$$

$$= \frac{0.9 \times 0.05}{0.2} = \frac{9}{10} \times \frac{1}{20} \times \frac{8}{4}$$

$$= \frac{9}{40}$$

$\Rightarrow P(E_1|E_2)$ = probability of a person having flu given that ~~he~~ she already has headache etc.

$$= \frac{9}{40} = 0.225$$

A2

d = no. of dimensions
n = no. of samples

Let X = feature matrix
shape → ~~~~ $d \times n$

~~Let X on SVD = UDV^T~~
~~real~~

Here, $n = 2$.

Let features be x_1, x_2 and straight line be $x_2 = mx_1 + c$.

Let $(x_1)_i$ be value of x_1 for i^{th} sample.

Claim: After normalizing each of the features, straight line will pass through origin.

Proof: $(x_2)_i = m(x_1)_i + c \quad \forall i \rightarrow L_1$

$$E[x_2] = \frac{1}{n}\sum_{i=1}^{n}(x_2)_i, \quad E[x_1] = \frac{1}{n}\sum_{i=1}^{2}(x_1)_i$$

Clearly, $(E[x_1], E[x_2])$ lies on $y = mx + c$.

After translation of j^{th} point by subtracting mean,

Original → (x_{1j}, x_{2j})

Now → $(x_{1j} - E[x_1], x_{2j} - E[x_2])$

Now, $(E[x_1], E[x_2])$ lied on L_1.

After translation by subtracting means,

new → $(E[x_1], E[x_2]) = (0, 0)$ on line L_2

17

Page 12

Also,

$(x_{2j} - E[x_2])$

$= m x_{1j} + c - \frac{1}{n}\sum_{i=1}^{n}(m x_{1i} + c)$

$= m x_{1j} - \frac{1}{n}\sum_{i=1}^{n} m x_{1i}$

$\Rightarrow (x_{2j} - E[x_2]) = m(x_{1j} - E[x_1])$

$\Rightarrow$ all points, after translation lie on a straight line passing through origin. Hence, proved.

Let us now assume (x_1, x_2) to be normalized and be on a straight line passing through origin.

③ Now, the eigenvalue - eigenvector pairs of $\frac{1}{n} XX^T$ i.e, covariance matrix gives us potential components along which dimensionality reduction can be done.

Also, the magnitude of eigenvalues helps us know along which dimensions, the spread in data points would be maximum and hence, helps us rank the eigenvectors by usefulness.

a) Yes

③ Assuming question means equal diagonal elements, 2 eigenvectors would be same → λ.

Now, corresponding to this eigenvector λ, it is possible to have 2 linearly independent eigenvectors in its eigenspace.

Let these be v_1, v_2.

Then, if we want to reduce dimensions, projecting along either v_1 or v_2 can help.

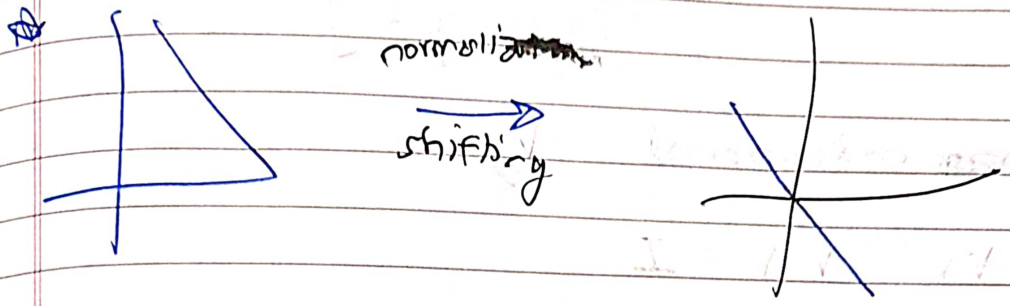
b) Yes; Let the distinct eigenvalues be
$\lambda_1 \rightarrow v_1$
$\lambda_2 \rightarrow v_2$

Without loss of generality, let $|\lambda_2| > |\lambda_1|$.

If we wish to reduce dimensions, projecting along v_2 would result in better spread.

Yes

c) I proved earlier that if lie on a straight line, then after shifting/normalizing, straight line will pass through origin.



Now,

$$X_{normalized} \rightarrow \begin{bmatrix} x_{11} & x_{12} & \cdots \\ mx_{11} & mx_{12} & \cdots \end{bmatrix}$$

Since row 2 is multiple of row 1,

$X_{normalized}$ does not have full rank.

Also, we know ~~rank of~~

$$Rank(X) = Rank(XX^T)$$ } true for all matrices

$\Rightarrow$ XX^T does not have full rank

$\Rightarrow \underbrace{\det(XX^T)}_{\text{product of eigenvalues}} = 0 \Rightarrow$ atleast one eigenvalue is zero

$\Rightarrow$ DD^T has an entirely zero row

and $rank(DD^T) = rank(D)$

$\Rightarrow$ D is not full rank.

if $V = \begin{bmatrix} \uparrow & \uparrow & \uparrow & & \uparrow \\ \hat{v}_1 & \hat{v}_2 & \hat{v}_3 & \cdots & \hat{v}_n \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{bmatrix}$

d) No.

By SVD defn, V is ~~orthonormal~~ orthogonal

$$\Rightarrow VV^T = V^TV = \begin{bmatrix} p_1^2 & 0 & 0 & 0 \\ 0 & p_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & p_n^2 \end{bmatrix}$$

For orthonormal V,

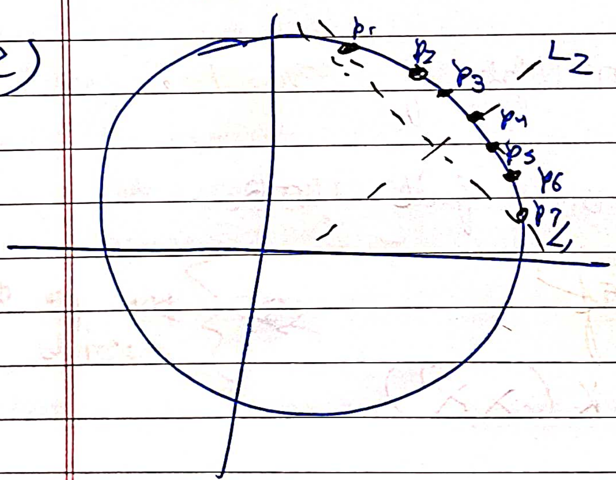
$$VV^T = V^TV = I$$

$$\Rightarrow V^{-1} = V^T$$

$\Rightarrow$ inverse exists

$\Rightarrow$ V has full-rank

e)



$\boxed{\text{No}}$

$L_1 \perp L_2$

Let 7 points be as shown.

Here, L_1 is best direction to project if needed.

But L_2 also has spread along it.

So, both will have corresponding non-zero eigenva

3) No, PCA does not account for class labels.