Page No.
Date: / /

Multiple classes

Discriminant functions build in two ways.

One-vs-rest

⇓

Build $k-1$ discriminant functions. Each discriminant function solves two classes classification problems.

eg (C_k vs not C_k)

One-vs-One

⇓

One discriminant functions per pair of classes.

total functions $= {}^{k}C_{2} = \frac{k(k-1)}{2}$

Dimensionality Reduction → (subtraction / minimization)

→ The no. of features/variables/columns present in a given dataset is known as dimensionality

⇓

process to reduce these features is called dimensionality reduction.

Causes

As a feature is high there is chance of overfitting, error, computational task more complicated, difficult visualization etc

⇓

Curse of Dimensionality

Is the process/way of converting the higher dimensions dataset into lesser dimensions dataset ensuring that is provides similar information.

(feature extraction,
feature selection)

Teacher's Signature ....

→ Widely used in supervised learning problem.

Application of DR
- → Noise reduction
- → cluster Analysis.
- → Data Visualization.
- → bioinformatics.
- → Speech recognition.
- → Signal processing.
- → ~~Computer Vision~~

Mainly used for the fields that deal with high dimensional data.

D R Techniques

- Feature Extraction
  ⇓
  - PCA.
  - LDA.
- Feature Selection.
- Dimensionality Reduction.
  - Components / Factors based
    ⇓
    - (i) Factor Analysis
    - (ii) PCA (Principle Component Analysis)
    - (iii) ICA. (Independent Component Analysis)
  - Projection Based.
    ⇓
    - (i) ISOMAP
    - (ii) t-SNE
    - (iii) UMAP.

    data visualization mainly / NLP etc

Benefits of applying DR
- → Less Computation training time.
- → Easily visualization
- → Reduce Overfitting case,
- → May remove multicollinearity if redundant features present

Disadvantages of DR

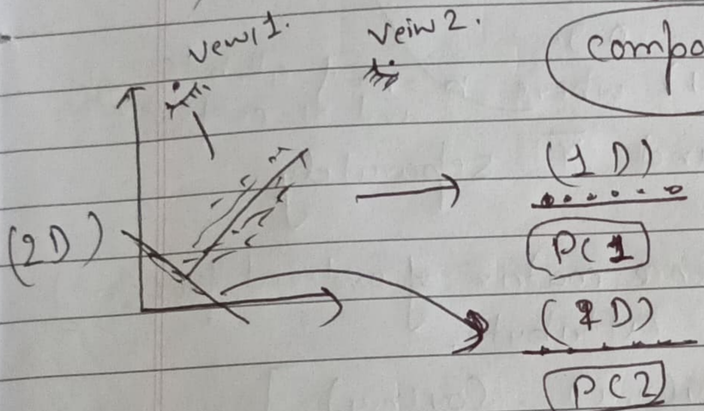
- Some of the data or feature may be lost.
- In PCA, sometimes the principal components required to consider are unknown.

① Principal Component Analysis (PCA)

- It is mainly used for overcome overfitting.
- In this we find the principle components.

⇓

apply views (different ways to understand the data's)



property 1

no. of PC ≤ no. of attributes (columns)

priority given ⇒ P(1) > P(2) > P(3) > .... PC(n)

property 2

Orthognal property

⇓

Means two thing are independents (i.e every principal component are independent to each other).

property 3

PC must be the linear combination of the original features.

Page No.
Date: / /

Ans

X	Y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9.

here $\bar{X} = 1.81$
$\bar{Y} = 1.91$

Step 0 Describe dataset. (n, N) where n = no. of attribute, and N = no. of observation

Step 1 Calculate $\bar{X}$ and $\bar{Y}$ separately.

Step 2 Create covarrience matrix of ordered pair (n^2)
here. 2 attribute,

$$\begin{bmatrix} Cov(x,x) & Cov(x,y) \\ Cov(y,x) & Cov(y,y) \end{bmatrix}_{2\times 2}.$$

where.

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

always $Cov(x,y) = Cov(y,x)$

$\therefore$

$$C = \begin{bmatrix} 0.6165 & 0.6154. \\ 0.6154 & 0.7165 \end{bmatrix}_{2\times 2}$$

Step 3 find egienvalue by form $C - \lambda I = 0$

$$\det(C - \lambda I) = 0$$

I → Identity matrix

$$\begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} 0.6165-\lambda & 0.6154 \\ 0.6154 & 0.7165-\lambda \end{bmatrix} = 0$$

here $\lambda_2 > \lambda_1$ $\Rightarrow$ $\lambda^2 - 1.333\lambda + 0.0630 = 0$

$\therefore \lambda_1 = 0.0490$ and $\lambda_2 = 1.2840$

<u>Step 4</u> Compute egienvector for each eigenvalue. by using this form. $C\nu = \lambda\nu$

$$\underbrace{C(\nu)}_{\text{egienvector}} = \lambda\nu$$

for $\lambda_1 \Rightarrow$

$$\begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 0.0490 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

solve this

$$\begin{cases} 0.6165x_1 + 0.6154y_1 = 0.0490x_1 \\ 0.6154x_1 + 0.7165y_1 = 0.0490y_1 \end{cases}$$

take any one of them.

$x_1 = -1.0845y_1$

put $y_1 = 1$ for temporary basis.

$\therefore x_1 = -1.0845$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} -1.0845 \\ 1 \end{bmatrix}$$

Normalize egien vector

$$= \sqrt{(-1.0845)^2 + (1)^2}$$

$$= \sqrt{1.17614 + 1}$$

$$= \sqrt{2.17614}$$

$$= 1.47517$$

Now,

$$x_1 = \frac{-1.0845}{1.47517} = -0.7351$$

$$y_1 = \frac{1}{1.47517} = 0.6778$$

for $\lambda_2 \Rightarrow$

$$\begin{bmatrix} 0.6165 & 0.6154 \\ 0.6154 & 0.7165 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = 1.2840 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

$0.6165x_2 + 0.6154y_2 = 1.2840x_2$

$0.6154x_2 + 0.7165y_2 = 1.2840y_2$

$$0.6678 X_2 = 0.6154 Y_2$$

$$0.6154 X_2 = 0.5675 Y_2$$

take any one of them

$$X_2 = 0.9215 Y_2$$

put $Y_2 = 1$ take as temporary basis

$$\therefore X_2 = 0.9215$$

$$\begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 0.9215 \\ 1 \end{bmatrix}$$

$$= \sqrt{(0.9215)^2 + (1)^2}$$

$$= \sqrt{0.8499 + 1}$$

$$= 1.359$$

Now,

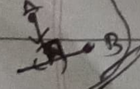
$$X_2 = \frac{0.9215}{1.359} = 0.6780$$

$$Y_2 = \frac{1}{1.359} = 0.7358$$

- It is an unsupervised algorithm. (Clustering?)
- Statistical process that converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation.

mutually exclusive

$$\vec{A} \cdot \vec{B} = 0$$



New transformation feature is called Principal Components (PC).

- Most widely used in EDA and predictive modelling

supervised learning ↓ (like dummy variable)

draw strong patterns from the given dataset by reducing the variances.

11.

Works by considering the variance of each attribute because the high attribute shows the good split between the classes and hence it reduces the dimensionality.

Orthogonal ⟶ Variables are not correleated to each other (i.e independent to each other)

#. Eigenvectors ⟶ A vector that is associated with a set of linear equations.

⟶ Also known as latent vector / characteristic vector.

⟶ Non-vector in which when a given matrix is multiplied it is equal to scalar multiple of that vector.

$A_{n \times n}$ v be non-zero vector

$\therefore \; Av = \lambda v$

⟶ egien value

Eigenvalue ⟶ A values are generally associated with eigenvectors in linear algebra.

Dimensionality, # Correlation, # Orthogonal, # Covariance matrix.

~~Steps Sorting the eign vectors (according to the priority condition)~~

Step 6 ~~P * matrix = mat~~ eigen vector^T Matrix = values

Application PCA

⟶ Computer vision.

⟶ image compression.

⟶ finding hidden pattern if data has high dimensions (Finance, data mining, Psychology etc.)

Example:

X	Y
4	11
8	4
13	5
7	14.

here n (no. of features) $= 2.$
$N =$ (no. of sample) $= 4.$

here

$\bar{X} = 8$ and $\bar{Y} = 8.5$

Now we calculate. covarriance matrix of ordered pair.

$$C \begin{bmatrix} x,x & x,y \\ y,x & y,y \end{bmatrix} = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Now we calalute egien value, eigen vector and normalized eijer vector.

$\therefore$

egien value $\Rightarrow$ $\det(C - \lambda I) = 0$

$$\det\left(\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix}\right) = 0$$

$$\lambda^2 - 37\lambda + 201 = 0 \Rightarrow \left(\cdot \frac{\sqrt{b^2-4ac}}{2a} \right) =$$

$\therefore$ $\lambda_1 = 30.38$ $\lambda_2 = 6.6151$

here $\lambda_1 > \lambda_2$.
PC1 PC2.

Now, egien vector for each λ (λ_1, λ_2).

$Cv = \lambda_1 v$ / $(C - \lambda I) v = 0$

$$\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 30.38 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$14 x_1 - 11 y_1 = 30.38 x_1$$
$$-11 x_1 + 23 y_1 = 30.38 y_1$$

$$(14 - \lambda) x_1 - 11 y_1 = 0$$
$$-11 x_1 + (23 - \lambda) y_1 = 0$$

take any one of them.

$$(14 - \lambda) x_1 = 11 y_1$$
$$x_1 = \frac{11}{14 - \lambda} y_1$$

Similarly, for λ_2 we do.

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

Now derive new dataset / PC.

PC1	PC2
P_{11}	P_{21}
P_{12}	P_{22}
P_{13}	P_{23}
P_{14}	P_{24}

here, for PC_1

$$P_{11} = e_1^T \begin{bmatrix} (x_1 - \bar{x}) \\ (y_1 - \bar{y}) \end{bmatrix}$$
$$= \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

$$P_{11} = -4.3052$$

Similarly for P_{12} P_{13} P_{14} all observations

Similarly for PC_2.

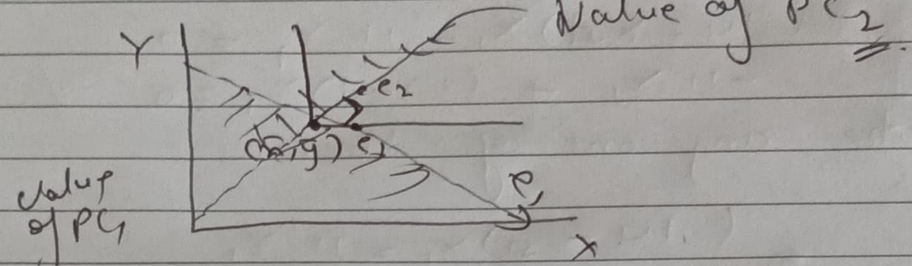
$$PC_2 \quad P_{21} = e_2^T \begin{bmatrix} x_1 - \bar{x} \\ y_1 - \bar{y} \end{bmatrix}$$

$$P_{21} = \begin{bmatrix} 0.8303 & 0.5574 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

$$= -3.3212 + 1.3935$$
$$= -1.9277.$$

" done for P_{22}, P_{23}, P_{24}

Now



② Factor Analysis

→ Also called Latent variable models

→ find the common factor between the attributes hidden behind them.

Reduce large no. of attributes in to a fewer number of factors

be a driving factor.

It also called common cause. (Latent variable)

Which columns are behaviour similarly comes in one group.

equation be

$$\text{attributes} = \alpha\left(\text{factor/driving factor}\right) + c_1$$

factor analysis equation ⇒

PCA equation ⇒ factor/driving factor (PC) $= \sum_{j=1}^{n} \alpha_n$ attributes

where $\tau_1 > \tau_2 > \dots \tau_n$

(Factor Analysis trying to find hidden factors in the attributes while PCA trying to explain most of data Variance by few attributes)

⇒

Factor

At1	At2	At3	At4	A5

factors/Driving factor

eg

Columns / observed variables

- Waiting time
- cleanliness
- Staff Behaviour.

→ Service

- Tast of Food
- Food Freshness
- Food Temperature.

→ Food Quality

Latent Variable → Variables that are not directly observed but are inferred from other variables.

→ factors / driving factor / common cause.

Lantent Variable model → Mathematical models that aim to explain observed variables in of the latent variables.

Assumption in FA

→ No outlier present in data.

→ sample size > factor. (1 : 5) factor observation

→ Correleated variables. (Barret test perform)

→ Normalized data

Teacher's Signature ....

- Purpose FA
  - → Data Reduction
  - → Latent Variable Discovery.
  - → Dimensionality Reduction.

- Types FA
  - → EFA (Exploratory Factor Analysis),
  - → CFA (Confirmatory Factor Analysis),