Thesis Progress 5

Eigenvalues and Eigenvector :-

If v is a vector and λ is a scaling factor and it ~~satisfied~~ satisfies the equation $Av = \lambda v$ where A is a square matrix, then v is called eigenvector and λ is called eigenvalue.

Given,

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$$

$$Av = \lambda v$$

$$\Rightarrow Av = \lambda v I \qquad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow Av - \lambda I v = 0$$

$$\Rightarrow (A - \lambda I) v = 0$$

But an eigenvector is a nonzero vector. So, eigenvector v cannot be 0.

So, $A - \lambda I = 0$

$$\text{de} \quad |A - \lambda I| = 0$$

$$\Rightarrow \left| \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$$\Rightarrow \left| \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{vmatrix} = 0$$

$$\Rightarrow (1-\lambda)(-\lambda) - 2 = 0$$

$$\Rightarrow -\lambda + \lambda^2 - 2 = 0$$

$$\Rightarrow \lambda^2 - \lambda - 2 = 0$$

$$\therefore \lambda_1 = 2 \quad \lambda_2 = -1.$$

Now,

$$Av = \lambda v$$

$$\Rightarrow \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix} \quad \left[\text{Let, } v = \begin{bmatrix} x \\ y \end{bmatrix}\right]$$

$$\Rightarrow \begin{bmatrix} x+2y \\ x \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

$\therefore x + 2y = 2x \Rightarrow x = 2y$

$\therefore x = 2y$

if $y = 1$, then $x = 2$

So, for eigenvalue $\lambda_1 = 2$,

eigenvector $v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Again,

$$Av = \lambda v$$

$$\Rightarrow \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -1 \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} x+2y \\ x \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$$

$\therefore x + 2y = -x \Rightarrow x = -y.$

$x = -y$

if $y = 1$, $x = -1$.

So, for eigenvalue $\lambda_2 = -1$,

eigenvector $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Covariance and Correlation

$$\text{var}(x) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Here,

x_i = Values of observations

$\bar{x}$ = Mean of those observations

n = Number of observations (for sample

$$SD(x) = \sqrt{\text{var}(x)}$$

$X = [2, 4, 6]$

$\bar{x} = \frac{2+4+6}{3} = 4$

$$\text{var}(x) = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3-1}$$

$$= \frac{8}{2}$$

$\therefore \text{var}(x) = 4$

$\therefore SD(x) = \sqrt{4} = 2$

~~Covarian~~

Covariance:

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

Correlation Coefficient:

$$r_{xy} = \frac{Cov(x,y)}{SD(x) \cdot SD(y)}$$

$$\therefore r_{xy} = \frac{Cov(x,y)}{\sqrt{Var(x)} \cdot \sqrt{Var(y)}}$$

~~$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{}$~~

$$r_{xy} = \frac{\frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}}{\sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \cdot \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}}}$$

$$\therefore r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

x	y
Body Weight (kg)	Body Height (cm)
61	157
62	168
73	170
74	181
82	191
86	185
$\bar{x} = 73$	$\bar{y} = 175.33$

$$\text{var}(x) = \frac{(61-73)^2 + (62-73)^2 + \cdots + (86-73)^2}{6-1}$$

$$= 103.2$$

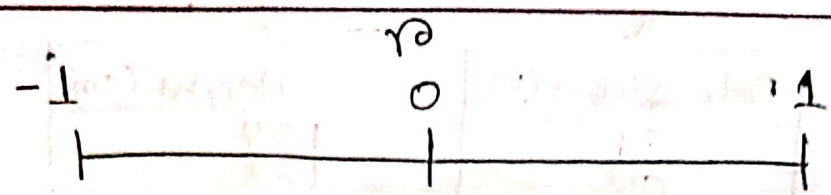
$$\text{var}(y) = \frac{(157-175.33)^2 + (168-175.33)^2 + \cdots + (185-175.33)^2}{6-1}$$

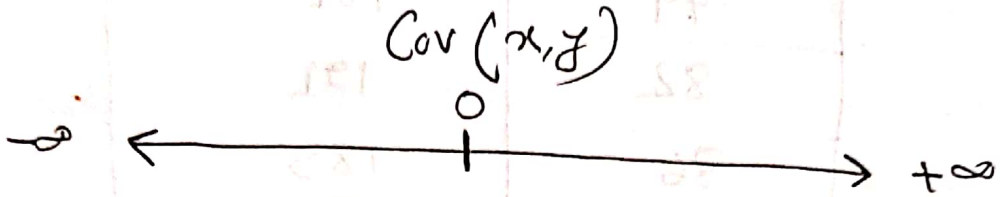
$$= 157.87$$

$$\text{Cov}(x,y) = \frac{(61-73)(157-175.33) + (62-73)(168-175.33) + \cdots}{6-1}$$

$$= 114.6$$

$$\therefore r_{xy} = \frac{\text{Cov}(x,y)}{\sqrt{\text{var}(x)}\sqrt{\text{var}(y)}} = \frac{114.6}{\sqrt{103.2}\sqrt{157.87}} = 0.898$$





Covariance Matrix: (Previous Table)

x = Body weight (kg)

y = Body Height (cm)

	x	y
x	$Cov(x,x)$	$Cov(x,y)$
y	$Cov(y,x)$	$Cov(y,y)$

$$Cov(x,x) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})}{n-1} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = var(x)$$

Similarly, $Cov(y,y) = var(y)$

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

$$= Cov(y,x)$$

∴ Covariance matrix of previous table

	x	y
x	var(x)	Cov(x,y)
y	Cov(y,x)	var(y)

⇒

	x	y
x	103.2	114.6
y	114.6	157.87

Correlation Matrix:

	x	y
x	r_{xx}	r_{xy}
y	r_{yx}	r_{yy}

$$r_{xx} = \frac{Cov(x,x)}{\sqrt{var(x)} \cdot \sqrt{var(x)}} = \frac{var(x)}{var(x)} = 1, \quad \text{So } r_{yy} = 1.$$

$r_{xy} = r_{yx}$

	x	y
x	1	r_{xy}
y	r_{yx}	1

⇒

	x	y
x	1	0.898
y	0.898	1

Standardization of data:

Standardizing data is common in many multivariate statistical methods when we have **different units** of our variable or if they **differ a lot.**

$$Z_i = x_i - \bar{x} \quad \text{(Centered data)}$$

$$Z_i = \frac{x_i - \bar{x}}{SD(x)} \quad \text{(Standardized data)}$$

Original Data

Body weight (kg)	Body Height (cm)
61	157
62	168
73	170
74	181
82	191
86	185
Mean = 73	Mean = 175.33
SD = 10.16	SD = 12.564

Centered Data

Body weight (kg)	Body Height (cm)
61-73 = -12	157 - 175.33
62 - 73 = -11	168 - 175.33
73 - 73 = 0	170 - 175.33
74 - 73 = 1	181 - 175.33 = 5.67
82 - 73 = 9	191 - 175.33 = 15.67
86 - 73 = 13	185 - 175.33 = 9.67
Mean = 0	Mean = 0
SD = 10.16	SD = 12.56

Standardized Data

Body weight	Body Height
-1.18	-1.46
-1.08	-0.58
0	-0.42
0.098	0.45
0.88	1.24
1.28	0.77
Mean = 0	Mean = 0
SD = 1	SD = 1

For standardized data: [x = Body weight, y = Body Height]

$Cov(x,y) = 0.89$, $Cov(x,x) = Var(x) = 1$.
$Cov(y,y) = Var(y) = 1$.

$$r_{xy} = \frac{Cov(x,y)}{\sqrt{Var(x)}\sqrt{Var(y)}} = \frac{Cov(x,y)}{SD(x) \cdot SD(y)} = \frac{Cov(x,y)}{1 \cdot 1}$$

$$\therefore \boxed{r_{xy} = Cov(x,y)} = 0.89$$

$r_{xx} = r_{yy} = 1$.

Covariance Matrix

	x	y
x	$Cov(x,x)$	$Cov(x,y)$
y	$Cov(y,x)$	$Cov(y,y)$

Correlation Matrix

	x	y
x	r_{xx}	r_{xy}
y	r_{yx}	r_{yy}

$\Rightarrow$

	x	y
x	1	0.89
y	0.89	1

$\Rightarrow$

	x	y
x	1	0.89
y	0.89	1

same

Principal Component Analysis (PCA):

Dealing with too many features or variables or dimensions can be intimidating and will also be time consuming for the model.

So, **dimensionality reduction** is what we want to achieve. There are mainly two categories to achieve dimensionality reduction:

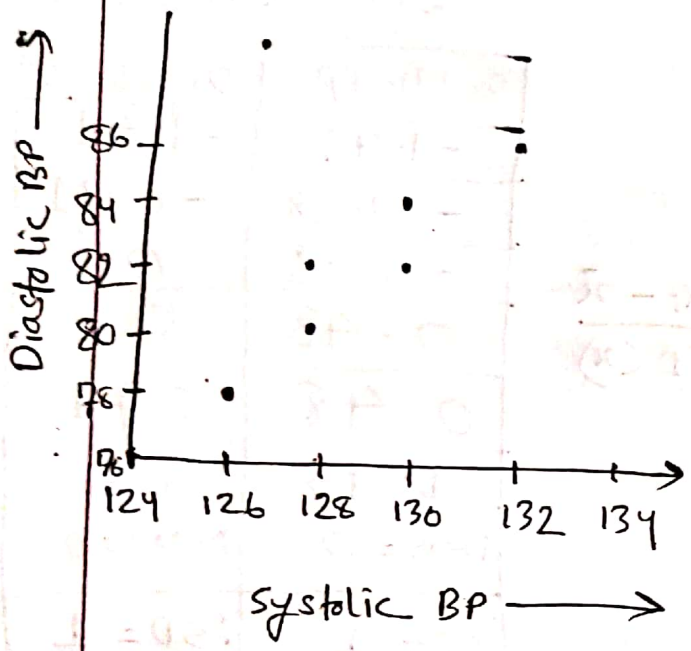
(i) Feature Elimination

(ii) Feature Extraction

Feature Elimination: Feature elimination is what it sounds like: we reduce the number of features by removing the less important features.

Feature Extraction: Suppose we have ten features. In feature extraction, we create ten "new" features where each "new" feature is the linear combination of all ten "old" feature. After that, out of ten "new" feature, we only pick few which contain significant amount of information. **PCA** is a technique for **feature extraction.**

PCA: the math step-by-step with a simple example



Systolic BP	Diastolic BP
126	78
128	80
128	82
130	82
130	84
132	86

To explain how the PCA works, we will use the following example data. We will use PCA to combine the two blood pressure variables into just one variable. based on data from six individuals.

Steps in Principal Component Analysis:

(i) Standardize the data
(ii) Calculate the covariance matrix
(iii) Calculate eigenvalues and eigenvectors
(iv) Sort the eigenvector in descending order based on eigen values and form feature vector
(v) Calculate the Principal components
(vi) Feature extraction

① Standardize the data:

Original Data

Systolic BP	Diastolic BP
126	78
128	80
128	82
130	82
130	84
132	86
Mean = 129	Mean = 82
SD = 2.1	SD = 2.83

$$z_i = \frac{x_i - \bar{x}}{SD(x)}$$

Standardized Data

Systolic BP	Diastolic BP
-1.43	-1.41
-0.48	-0.71
-0.48	0
0.48	0
0.48	0.71
1.43	1.41
Mean = 0	Mean = 0
SD = 1	SD = 1

(ii) Calculate the covariance matrix

x = Systolic BP in standardized data

y = Diastolic BP in standardized data

	x	y
x	Cov(x,x)	Cov(x,y)
y	Cov(y,x)	Cov(y,y)

$$Cov(x,x) = Var(x) = SD(x)^2 = 1^2 = 1$$

$$Cov(y,y) = Var(y) = SD(y)^2 = 1^2 = 1$$

$$\text{Cov}(x,y) = \text{Cov}(y,x) = \frac{\sum_{i=1}^{5} (x_i - \bar{x})(y_i - \bar{y})}{5} = 0.94$$

∴ Covariance matrix,

	x	y
x	1	0.94
y	0.94	1

(iii) Calculate eigenvalues and eigenvectors:

Let, $A = \begin{bmatrix} 1 & 0.94 \\ 0.94 & 1 \end{bmatrix}$

λ = eigen values

v = eigenvectors

$$Av = \lambda v \Rightarrow Av - \lambda vI = 0 \Rightarrow (A - \lambda I)v = 0$$

$$\Rightarrow -A - \lambda I = 0 \Rightarrow \left| \begin{bmatrix} 1 & 0.94 \\ 0.94 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\Rightarrow \begin{vmatrix} 1-\lambda & 0.94 \\ 0.94 & 1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (1-\lambda)^2 - (0.94)^2 = 0 \Rightarrow 1 - 2\lambda + \lambda^2 - 0.88 = 0$$

$$\Rightarrow \lambda^2 - 2\lambda + 0.11 = 0$$

$$\therefore \lambda_1 = 1.94, \quad \lambda_2 = 0.06$$

$$Av = \lambda v \Rightarrow \begin{bmatrix} 1 & 0.94 \\ 0.94 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1.94 \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x+0.94y \\ 0.94x+y \end{bmatrix} = \begin{bmatrix} 1.94x \\ 1.94y \end{bmatrix}$$

$x + 0.94y = 1.94x \Rightarrow 0.94y = 0.94x \Rightarrow y = x$

$0.94x + y = 1.94y \Rightarrow 0.94x = 0.94y \Rightarrow y = x$

if $x = 1$ then $y = 1$.

So, $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda_1 = 1.94$

$\Rightarrow v_1 = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}$, $\lambda_1 = 1.94$.

Again,

$Av = \lambda v$

$$\Rightarrow \begin{bmatrix} 1 & 0.94 \\ 0.94 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.06 \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x+0.94y \\ 0.94x+y \end{bmatrix} = \begin{bmatrix} 0.06x \\ 0.06y \end{bmatrix}$$

$x + 0.94y = 0.06x \Rightarrow 0.94y = -0.94x \Rightarrow y = -x$

$0.94x + y = 0.06y \Rightarrow 0.94x = -0.94y \Rightarrow x = -y$

if $y = 1$ then $x = -1$.

So, $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\lambda_2 = 0.06$.

$\Rightarrow v_2 = \begin{bmatrix} -0.71 \\ 0.71 \end{bmatrix}$, $\lambda_2 = 0.06$.

(iv) Sort the eigenvector in descending order based on eigenvalues and form feature vector:

$\lambda_1 = 1.94$, $v_1 = \begin{bmatrix} 0.71 \\ 0.71 \end{bmatrix}$

$\lambda_2 = 0.06$, $v_2 = \begin{bmatrix} -0.71 \\ 0.71 \end{bmatrix}$

feature vector $V = \begin{bmatrix} 0.71 & -0.71 \\ 0.71 & 0.71 \end{bmatrix}$

(v) Calculate the principal components

$V = \begin{bmatrix} 0.71 & -0.71 \\ 0.71 & 0.71 \end{bmatrix}$ $D = \begin{bmatrix} -1.43 & -1.41 \\ -0.48 & -0.71 \\ -0.48 & 0 \\ -0.48 & 0 \\ 0.48 & 0.71 \\ 1.43 & 1.41 \end{bmatrix}$

$DV = \begin{bmatrix} -1.43 & -1.41 \\ -0.48 & -0.71 \\ -0.48 & 0 \\ -0.48 & 0 \\ 0.48 & 0.71 \\ 1.43 & 1.41 \end{bmatrix} \begin{bmatrix} 0.71 & -0.71 \\ 0.71 & 0.71 \end{bmatrix} = \begin{bmatrix} -2.02 & 0.01 \\ -0.84 & -0.16 \\ -0.34 & 0.34 \\ -0.34 & -0.34 \\ 0.84 & 0.16 \\ 2.02 & -0.01 \end{bmatrix}$

PC1 PC2

$$\therefore DV = \begin{bmatrix} -2.02 & 0.01 \\ -0.84 & -0.16 \\ -0.34 & 0.34 \\ 0.34 & -0.34 \\ 0.84 & 0.16 \\ 2.02 & -0.01 \end{bmatrix}$$

Transformed data

(vi) Feature Extraction:

PC1	PC2
-2.02	0.01
-0.84	-0.16
-0.34	0.34
0.34	-0.34
0.84	0.16
2.02	-0.01
Mean = 0	Mean = 0
var = 1.96	var = 0.06

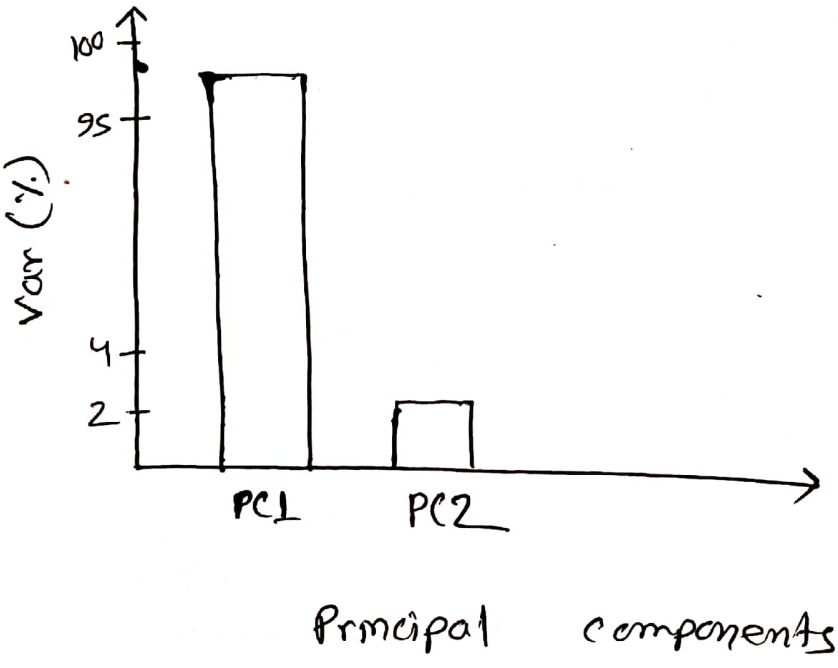
It is observed that,

variance of PC1 = updated (after sorting) eigenvalue $\lambda_1 = 1.96$ (almost)

variance of PC2 = updated (after sorting) eigenvalue $\lambda_2 = 0.06$ (almost)

Total variance = 1.96 + 0.06 = 2.02

	PC1	PC2
var	1.96	0.06
var (%)	97.03%	2.97%
cum (%)	97.03%	100%



Since variance of PC1 is 97.03%, we can eliminate PC2 that contains only 2.97% of information.