<u>Principal Component Analysis</u>

1. Introduction

Principal Component Analysis (PCA) เป็นวิธีทางสถิติที่ใช้ในการแปลงตัวแปรทางสถิติที่ตั้งฉากกัน (orthogonal) หรือมีความสัมพันธ์กัน (correlated) ให้กลายเป็นตัวแปรที่ไม่สัมพันธ์กัน (uncorrelated)

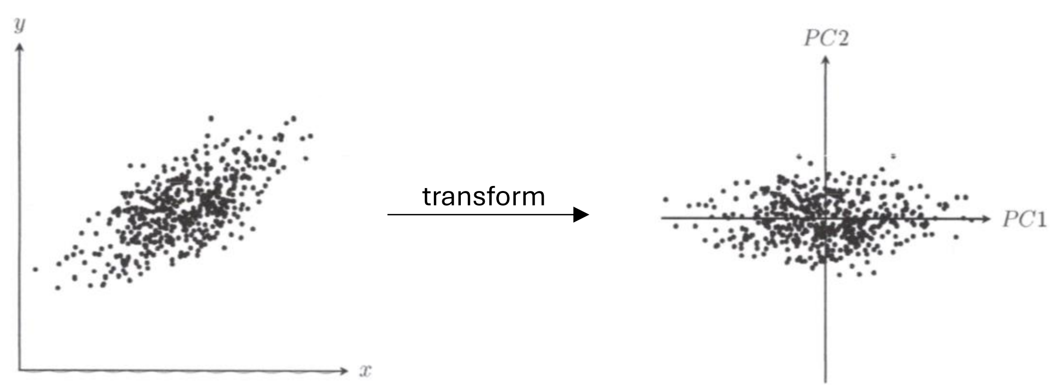

Figure 1: ลักษณะการแปลงข้อมูลใน PCA

รูปวงรีแกน สมมาตรของวงรีจะเรียกว่า “principal axis” ซึ่งคำว่า principal axis เป็นการเรียกรวมของแกนเอกและแกนโทของวงรี ในกรณีที่เป็นวงรีสองมิติก็จะมีแกน สมมาตร 2 แกน แต่ถ้าอยู่ใน $\boldsymbol{p}$ มิติก็จะมี principal axis $\boldsymbol{p}$ แกน สำหรับข้อมูลที่มีการกระจายตามรูปที่ 1 เราจะกำหนด principal component ด้วยแนวคิดเดียวกันกับ

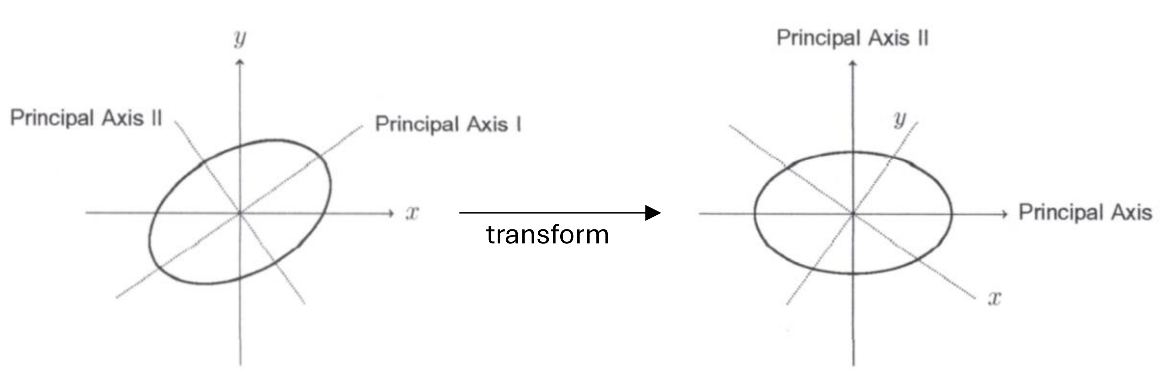

Figure 2: Principal axis ของวงรี

principal axis ของวงรี แต่เนื่องจากจุดข้อมูลไม่ได้มีแกนเอกแกนโทอย่างชัดเจน principal component จึงหาจากการ maximize variance แทน ซึ่งแกนที่จุดข้อมูลโปรเจ็กต์ลงมาแล้ว จะให้ค่า variance บนแกนมากที่สุดก็จะเป็น principal component เหมือนกับส่วนที่ป่องที่สุดของวงรีเป็น principal axis

2. Finding Principal Components

ข้อมูลก่อนที่จะนำมาทำ PCA จะอยู่ในรูปของเมทริกซ์ขนาด $n \times p$ โดยที่ n คือจำนวนจุดข้อมูลและ p เป็นจำนวน feature

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}$$

ก่อนที่จะหา principal component ให้ลบทุกฟีเจอร์ด้วยค่าเฉลี่ยเพื่อเป็นการเลื่อนข้อมูลมาอยู่บริเวณกึ่งกลางหรือ

$$\sum_{i=1}^{n} x_{i,d} = 0$$

เนื่องจากต้องการกำหนดให้เวกเตอร์ principal component vector มีทิศพุ่งออกจากจุดกำเนิด จึงต้องกำหนดเวกเตอร์หนึ่งหน่วยของ principal component ใน p มิติ ก็จะเขียนได้เป็น

$$\mathbf{v} = v_1\mathbf{x}_1 + v_2\mathbf{x}_2 + \cdots + v_p\mathbf{x}_p$$

หรือเขียนเป็นรูปเมทริกซ์ได้ว่า

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}$$

จากนั้นโปรเจ็กต์เวกเตอร์ของจุดข้อมูลลงบนแกน $\mathbf{v}$ จะได้ดังรูป

ถ้าพิจารณาจุด x_i ใดๆ เมื่อโปรเจ็กต์ลงบนแกน $\mathbf{v}$ แล้ว ระยะห่างจากจุดกำเนิดจนถึงจุดที่ข้อมูลถูกโปรเจ็กต์ก็จะหาได้จาก

$$s_i = \mathbf{v} \cdot \mathbf{x}_i$$

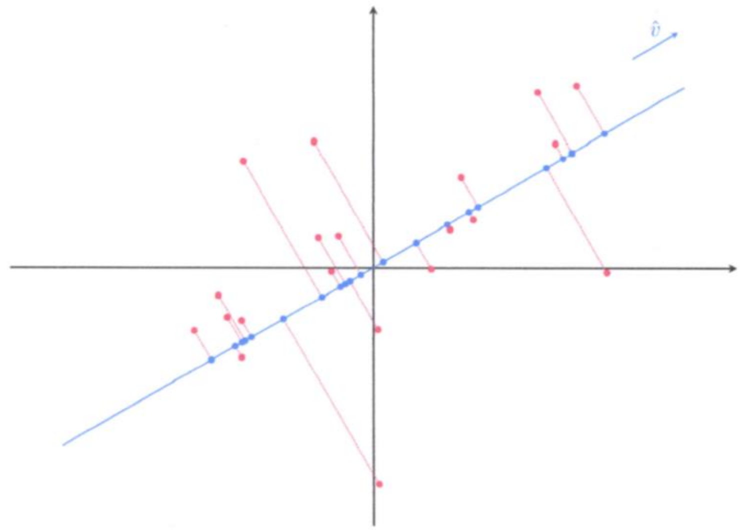

Figure 3: ลักษณะการโปรเจ็กต์จุดข้อมูลลงบนแกน

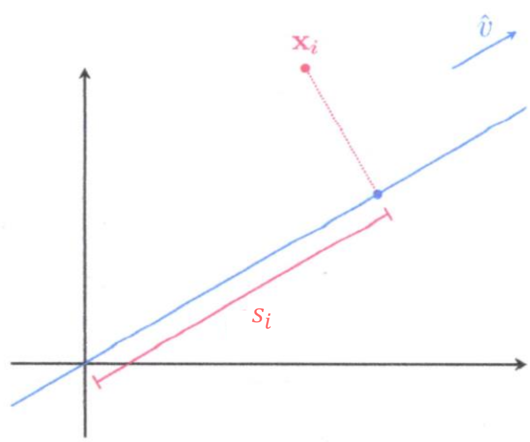

Figure 4: วิธีการวัดระยะ s_i

Variance บนแกน $\mathbf{v}$ หาได้จาก

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} s_i^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i)^2$$

Principal component จะเป็นแกนที่ทำให้เกิด variance มากที่สุด ดังนั้นก็จะได้ว่า

$$\max \sigma^2 = \max \sum_{i=1}^{n} s_i^2$$

โดยมีเงื่อนไขคือ

$$\|\mathbf{v}\|^2 = 1$$

หรือ

$$v_1^2 + v_2^2 + \cdots + v_p^2 = 1$$

ซึ่งเงื่อนไขนี้จะบังคับว่าให้เวกเตอร์ $\mathbf{v}$ เป็นเวกเตอร์หนึ่งหน่วย

สมการตั้งต้น PCA

$$\max \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i)^2$$

subject to

$$g(\mathbf{v}) = 1 - v_1^2 - v_2^2 - \cdots - v_p^2 = 0$$

ชุดสมการนี้สามารถแก้ได้โดยวิธี Lagrange Multiplier ซึ่ง Lagrangian function คือ

$$\begin{aligned} L &= \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i)^2 + \lambda(1 - \sum_{d=1}^{p} v_d^2) \\ &= \sum_{i=1}^{n} \left(v_1 x_{i,1} + v_2 x_{i,2} + \cdots + v_p x_{i,p}\right)^2 + \lambda\left(1 - v_1^2 - v_2^2 - \cdots - v_p^2\right) \end{aligned}$$

นำ L ไปหาอนุพันธ์ย่อยเทียบกับ v_d

$$\begin{aligned} 0 &= \frac{\partial L}{\partial v_d} \\ 0 &= \sum_{i=1}^{n} 2\left(v_1 x_{i,1} + v_2 x_{i,2} + \cdots + v_p x_{i,p}\right) x_{i,d} - \lambda 2 v_d \\ 0 &= 2 \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i) x_{i,d} - 2\lambda v_d \\ \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i) x_{i,d} &= \lambda v_d \end{aligned}$$

จากเทอม $\sum (\mathbf{v} \cdot \mathbf{x}_i) x_{i,d}$ กระจายผลรวมและจัดรูปให้อยู่ในเมทริกซ์

$$
\begin{aligned}
\sum_{i=1}^{n}(\mathbf{v}\cdot\mathbf{x}_i)x_{i,d} &= \left(v_1x_{1,1}+v_2x_{1,2}+\cdots+v_px_{1,p}\right)x_{1,d} \\
&\quad +\left(v_1x_{2,1}+v_2x_{2,2}+\cdots+v_px_{2,p}\right)x_{2,d} \\
&\quad +\cdots+\left(v_1x_{n,1}+v_2x_{n,2}+\cdots+v_px_{n,p}\right)x_{n,d} \\
&= \begin{bmatrix} x_{1,d} & x_{2,d} & \cdots & x_{n,p} \end{bmatrix} \begin{bmatrix} v_1x_{1,1}+v_2x_{1,2}+\cdots+v_px_{1,p} \\ v_1x_{2,1}+v_2x_{2,2}+\cdots+v_px_{2,p} \\ \vdots \\ v_1x_{n,1}+v_2x_{n,2}+\cdots+v_px_{n,p} \end{bmatrix} \\
&= \begin{bmatrix} x_{1,d} & x_{2,d} & \cdots & x_{n,p} \end{bmatrix} \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}
\end{aligned}
$$

ถ้านำมารวมกัน p สมการ

$$
\sum_{i=1}^{n}(\mathbf{v}\cdot\mathbf{x}_i)\mathbf{x}_i = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{n,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,p} & x_{2,p} & \cdots & x_{n,p} \end{bmatrix} \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}
$$

สามารถเขียนได้เป็น

$$
\sum_{i=1}^{n}(\mathbf{v}\cdot\mathbf{x}_i)\mathbf{x}_i = X^TX\mathbf{v}
$$

ถ้าจะได้คูณเมทริกซ์ X^TX

$$
X^TX = \begin{bmatrix} \sum_{i=1}^{n}\left(x_{i,1}\right)^2 & \sum_{i=1}^{n} x_{i,1}x_{i,2} & \cdots & \sum_{i=1}^{n} x_{i,1}x_{i,p} \\ \sum_{i=1}^{n} x_{i,2}x_{i,1} & \sum_{i=1}^{n}\left(x_{i,2}\right)^2 & \cdots & \sum_{i=1}^{n} x_{i,2}x_{i,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{i,p}x_{i,1} & \sum_{i=1}^{n} x_{i,p}x_{i,2} & \cdots & \sum_{i=1}^{n}\left(x_{i,p}\right)^2 \end{bmatrix}
$$

$$= n \begin{bmatrix} \mathrm{cov}_{11} & \mathrm{cov}_{12} & \cdots & \mathrm{cov}_{1p} \\ \mathrm{cov}_{12} & \mathrm{cov}_{22} & \cdots & \mathrm{cov}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}_{p1} & \mathrm{cov}_{p2} & \cdots & \mathrm{cov}_{pp} \end{bmatrix}$$

เมื่อ $\mathbf{cov}_{ij}$ คือ covariance ของ feature ที่ i กับ j ถ้า $i = j$ ก็จะได้ variance ดังนั้นในแนวทแยงของเมทริกซ์นี้ก็คือค่าความแปรปรวนของเมทริกซ์ ถ้าให้ $\sum$ เป็น covariance matrix ก็จะได้ว่า $\sum = \frac{1}{n} X^T X$ จะได้

$$\begin{aligned} X^T X\mathbf{v} - \lambda\mathbf{v} &= 0 \\ (X^T X - \lambda I)\mathbf{v} &= 0 \end{aligned}$$

ถ้าให้ $S = X^T X$ จะได้

$$(S - \lambda I)\mathbf{v} = 0$$

ปัญหานี้เป็นปัญหา eigenvalue-eigenvector ซึ่งวิธีแก้จะต้องหา eigenvalue แต่ละตัว และหา eigenvector ที่สอดคล้องกับ eigenvalue ก็จะได้เวกเตอร์ที่เป็น principal component ออกมา ตามจำนวนมิติ

3. Numerical Example

สมมติว่าข้อมูลมี 2 มิติ ดังนี้

$$X = \begin{bmatrix} -2.2 & -1.6 \\ -0.2 & 1.4 \\ 1.8 & 0.4 \\ -0.2 & -0.6 \\ 0.8 & 0.4 \end{bmatrix}$$

ลองหาค่าเฉลี่ยของแต่ละ component

$$\mu_1 = \frac{1}{5}(-2.2 - 0.2 + 1.8 - 0.2 + 0.8) = 0$$

$$\mu_2 = \frac{1}{5}(-1.6 + 1.4 + 0.4 - 0.6 + 0.4) = 0$$

เมื่อนำ X ไปทรานสโพสต์และคูณกับ X

$$
\begin{aligned}
S &= X^T X \\
&= \begin{bmatrix} -2.2 & -0.2 & 1.8 & -0.2 & 0.8 \\ -1.6 & 1.4 & 0.4 & -0.6 & 0.4 \end{bmatrix} \begin{bmatrix} -2.2 & -1.6 \\ -0.2 & 1.4 \\ 1.8 & 0.4 \\ -0.2 & -0.6 \\ 0.8 & 0.4 \end{bmatrix} \\
&= \begin{bmatrix} 8.8 & 4.4 \\ 4.4 & 5.2 \end{bmatrix}
\end{aligned}
$$

เมทริกซ์ที่ได้มานี้จะเป็นเมทริกซ์จตุรัสที่มีขนาดเท่ากับจำนวนมิติและเป็นเมทริกซ์สมมาตร จากนั้นนำเมทริกซ์นี้ไปหา eigenvalue ซึ่งจำนวน eigenvalue จะเท่ากับจำนวนมิติ ซึ่งเมทริกซ์นี้เป็นเมทริกซ์ขนาด 2 x 2 ดังนั้นจะมี eigenvalue อยู่ 2 ตัว จาก

$$|S - \lambda I| = 0$$

เมื่อ λ คือ eigenvalue ดังนั้น

$$
\begin{aligned}
|S - \lambda I| &= 0 \\
\begin{vmatrix} 8.8 & 4.4 \\ 4.4 & 5.2 \end{vmatrix} &= 0 \\
(8.8 - \lambda)(5.2 - \lambda) - (4.4)(4.4) &= 0 \\
\lambda^2 - 14\lambda + 26.4 &= 0
\end{aligned}
$$

เมื่อแก้สมการนี้ จะได้คำตอบสองคำตอบคือ

$$\lambda = \frac{35 - \sqrt{565}}{5}, \frac{35 + \sqrt{565}}{5}$$

หา eigenvector ที่สอดคล้องกับ eigenvector แต่ละตัว จาก

$$(S - \lambda I)\mathbf{v} = 0$$

เมื่อ $\hat{v}$ เป็น eigenvector

- เมื่อ $\lambda = \frac{35-\sqrt{565}}{5}$ จะได้

$$\begin{bmatrix} 8.8 - \frac{35-\sqrt{565}}{5} & 4.4 \\ 4.4 & 5.2 - \frac{35-\sqrt{565}}{5} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

จะได้ชุดสมการ

$$\begin{aligned} \left(8.8 - \frac{35-\sqrt{565}}{5}\right) v_1 + 4.4 v_2 &= 0 \\ 4.4 v_1 + \left(5.2 - \frac{35-\sqrt{565}}{5}\right) &= 0 \end{aligned}$$

หรือ

$$\begin{aligned} \left(9+\sqrt{565}\right) v_1 + 22 v_2 &= 0 \\ 22 v_1 + \left(-9+\sqrt{565}\right) v_2 &= 0 \end{aligned}$$

ถ้าคูณสมการล่างด้วยสมการทั้งสองชุดนี้เมื่อจัดรูปจะได้สมการเดียวกัน คือ

$$\left(9+\sqrt{565}\right) v_1 + 22 v_2 = 0$$

หนึ่งในคำตอบที่ทำให้ v_1 และ v_2 เป็นจริงคือ

$$\begin{aligned} v_1 &= 22 \\ v_2 &= 9+\sqrt{565} \end{aligned}$$

นั่นคือถ้า $\lambda = \frac{35-\sqrt{565}}{5} = 2.24605427$ จะได้ eigenvector เป็น

$$\hat{v} = \begin{bmatrix} -22 \\ 9+\sqrt{565} \end{bmatrix}$$

ถ้า normalize ให้มีขนาดเป็นหนึ่ง โดยการหารด้วยขนาดของเวกเตอร์จะได้

$$\mathbf{v} = \begin{bmatrix} -\frac{22}{\sqrt{1130+18\sqrt{565}}} \\ \frac{9+\sqrt{565}}{\sqrt{1130+18\sqrt{565}}} \end{bmatrix} = \begin{bmatrix} -0.5573899686 \\ 0.8302508192 \end{bmatrix}$$

- เมื่อ $\lambda = \frac{35+\sqrt{565}}{5} = 11.75394573$ จะได้

$$\begin{bmatrix} 8.8 - \frac{35+\sqrt{565}}{5} & 4.4 \\ 4.4 & 5.2 - \frac{35+\sqrt{565}}{5} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

เมื่อจัดรูปจะได้สมการ

$$\left(9-\sqrt{565}\right)v_1 + 22v_2 = 0$$

$$22v_1 - \left(9+\sqrt{565}\right)v_2 = 0$$

สมการสองสมการนี้เป็นสมการเดียวกัน eigenvector ที่เป็นไปได้คือ

$$\mathbf{v} = \begin{bmatrix} 9+\sqrt{565} \\ 22 \end{bmatrix}$$

เมื่อ normalize จะได้

$$\mathbf{v} = \begin{bmatrix} \frac{9+\sqrt{565}}{\sqrt{1130+18\sqrt{565}}} \\ \frac{22}{\sqrt{1130+18\sqrt{565}}} \end{bmatrix} = \begin{bmatrix} 0.8302508192 \\ 0.5573899686 \end{bmatrix}$$

สรุปคือ eigenvector ที่สอดคล้องกับ eigenvalue คือ

- ถ้า $\lambda = 2.24605427$ จะได้

$$\mathbf{v} = \begin{bmatrix} -0.5573899686 \\ 0.8302508192 \end{bmatrix}$$

- ถ้า $\lambda = 11.75394573$ จะได้

$$\mathbf{v} = \begin{bmatrix} 0.8302508192 \\ 0.5573899686 \end{bmatrix}$$

Eigenvector จะเป็นเวกเตอร์ที่มีทิศขนานกับ principal component และถ้าต้องการจะดูว่า principal component ไหนที่เมื่อโปรเจ็กต์ข้อมูลลงมาแล้วจะให้ค่า variance มากที่สุดให้ดูที่ค่า eigenvalue ที่สอดคล้องกับ eigenvector นั้น ถ้า eigenvalue มีค่ามากแปลว่า variance ของข้อมูลที่โปรเจ็กต์ลงบน eigenvector นั้นจะยิ่งมีค่ามาก และยิ่งมีความสำคัญอย่างมาก ส่วนแกนที่ค่า eigenvalue มีค่าน้อยหมายความว่าข้อมูลที่โปรเจ็กต์ลงบนแกนนั้นมีค่า variance น้อย ซึ่งอาจจะไม่ได้มีนัยสำคัญและสามารถที่จะไม่นำมาพิจารณาได้

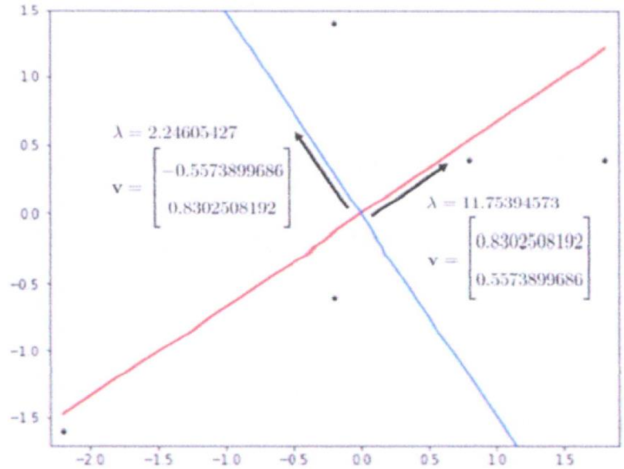

Figure 5: eigenvector และ eigenvector ของข้อมูลตัวอย่างชุดนี้

เมทริกซ์ที่นำมาใช้ในการแปลงระบบพิกัดจะหาได้โดยนำ eigenvector มาต่อกันโดยเรียงจาก eigenvector ที่มี eigenvalue ที่มีค่ามากที่สุดไปหาน้อยสุด

$$V = \begin{bmatrix} 0.8302508192 & -0.5573899686 \\ 0.5573899686 & 0.8302508192 \end{bmatrix}$$

ถ้าต้องการจะแปลงข้อมูลให้อยู่บนพิกัดของ principal component ทำได้โดยคูณกับข้อมูล

$$X' = \begin{bmatrix} -2.2 & -1.6 \\ -0.2 & 1.4 \\ 1.8 & 0.4 \\ -0.2 & -0.6 \\ 0.8 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8302508192 & -0.5573899686 \\ 0.5573899686 & 0.8302508192 \end{bmatrix}$$

$$= \begin{bmatrix} 2.71837575 & -0.10214338 \\ 0.61429579 & 1.27382914 \\ 1.71740746 & -0.67120162 \\ -0.50048415 & -0.3866725 \\ 0.88715664 & -0.11381165 \end{bmatrix}$$

ถ้าต้องการลดจำนวนมิติของข้อมูลก็ตัดหลักท้ายๆ ที่มีค่า eigenvalue น้อยๆ ทิ้งจนเหลือจำนวนหลักเท่ากับจำนวนมิติที่ต้องการ เช่นถ้าต้องการจะแปลงให้เหลือมิติเดียว

$$X' = \begin{bmatrix} -2.2 & -1.6 \\ -0.2 & 1.4 \\ 1.8 & 0.4 \\ -0.2 & -0.6 \\ 0.8 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8302508192 \\ 0.5573899686 \end{bmatrix}$$

$$= \begin{bmatrix} 2.71837575 \\ 0.61429579 \\ 1.71740746 \\ -0.50048415 \\ 0.88715664 \end{bmatrix}$$

<u>ขั้นตอน PCA</u>

1. standardize ข้อมูล: เลื่อนข้อมูลให้มาอยู่ตรงกลางและทำให้ความแปรปรวนของแต่ละหลักเป็น 1

```
Z = (X - np.mean(X, axis=0))/np.std(X, axis=0)
```

2. คำนวณหา covariance matrix

```
S = np.dot(np.transpose(Z), Z)
```

3. หา eigenvalue และ eigenvector

```
eigenvalue, eigenvector = np.linalg.eig(S)
```

4. เรียง eigenvalue และ eigenvector ตามขนาดของ eigenvalue จากมากไปน้อย

```
sorted_idx = np.argsort(eigenvalue)[::-1]
eigenvalue = eigenvalue[sorted_idx]
eigenvector = np.transpose(np.transpose(eigenvector)[sorted_idx])
```

5. สามารถเอา eigenvector ไปใช้ได้ตามต้องการ เช่นถ้าต้องการจะโปรเจ็กต์ข้อมูลลงบนระนาบ 2มิติ ก็จะได้

```
new_Z = np.dot(new_standardized_data, eigenvector[:, :2])
```