- **ขั้นตอนที่ 1: หาค่าเฉลี่ยและปรับข้อมูล (Center the Data)**

P(x,y)	X	Y
A	2	1
B	4	3
C	5	5
D	7	5
Mean (ค่าเฉลี่ย)	(2 + 4 + 5 + 7) / 4 = 4.5	(1 + 3 + 5 + 5) / 4 = 3.5

P'(x,y)	X	Y
A'	2 - 4.5 = -2.5	1 - 3.5 = -2.5
B'	4 - 4.5 = -0.5	3 - 3.5 = -0.5
C'	5 - 4.5 = 0.5	5 - 3.5 = 1.5
D'	7 - 4.5 = 2.5	5 - 3.5 = 1.5

- **ขั้นตอนที่ 2: คำนวณ Covariance Matrix (C)**
- ใช้ข้อมูลที่ปรับแล้ว (Centered Data) มาคำนวณ

$$Cov(X_1, X_1) = \frac{\sum_{k=1}^{4}(x_{1k} - \overline{x_1})(x_{1k} - \overline{x_1})}{4-1}$$

$$= \frac{(x_{11} - \overline{x_1})(x_{11} - \overline{x_1}) + (x_{12} - \overline{x_1})(x_{12} - \overline{x_1}) + (x_{13} - \overline{x_1})(x_{13} - \overline{x_1}) + (x_{14} - \overline{x_1})(x_{14} - \overline{x_1})}{4-1}$$

$$= \frac{(2-4.5)(2-4.5) + (4-4.5)(4-4.5) + (5-4.5)(5-4.5) + (7-4.5)(7-4.5)}{4-1}$$

$$= \frac{(-2.5)(-2.5) + (-0.5)(-0.5) + (0.5)(0.5) + (2.5)(2.5)}{4-1} = \frac{13}{3} \approx 4.33$$

$$Cov(X_1, X_2) = \frac{\sum_{k=1}^{4}(x_{1k} - \overline{x_1})(x_{2k} - \overline{x_2})}{4-1}$$

$$= \frac{(x_{11} - \overline{x_1})(x_{21} - \overline{x_2}) + (x_{12} - \overline{x_1})(x_{22} - \overline{x_2}) + (x_{13} - \overline{x_1})(x_{23} - \overline{x_2}) + (x_{14} - \overline{x_1})(x_{24} - \overline{x_2})}{4-1}$$

$$= \frac{(2-4.5)(1-3.5) + (4-4.5)(3-3.5) + (5-4.5)(5-3.5) + (7-4.5)(5-3.5)}{4-1}$$

$$= \frac{(-2.5)(-2.5) + (-0.5)(-0.5) + (0.5)(1.5) + (2.5)(1.5)}{4-1} = \frac{11}{3} \approx 3.67$$

$$Cov(X_2, X_1) = \frac{\sum_{k=1}^{4}(x_{2k} - \overline{x_2})(x_{1k} - \overline{x_1})}{4-1}$$

$$= \frac{(x_{21} - \overline{x_2})(x_{11} - \overline{x_1}) + (x_{22} - \overline{x_2})(x_{12} - \overline{x_1}) + (x_{23} - \overline{x_2})(x_{13} - \overline{x_1}) + (x_{24} - \overline{x_2})(x_{14} - \overline{x_1})}{4-1}$$

$$= \frac{(1-3.5)(2-4.5) + (3-3.5)(4-4.5) + (5-3.5)(5-4.5) + (5-3.5)(7-4.5)}{4-1}$$

$$= \frac{(-2.5)(-2.5) + (-0.5)(-0.5) + (1.5)(0.5) + (1.5)(2.5)}{4-1} = \frac{11}{3} \approx 3.67$$

$$Cov(X_2, X_2) = \frac{\sum_{k=1}^{4}(x_{2k} - \overline{x_2})(x_{2k} - \overline{x_2})}{4-1}$$

$$= \frac{(x_{21} - \overline{x_2})(x_{21} - \overline{x_2}) + (x_{22} - \overline{x_2})(x_{22} - \overline{x_2}) + (x_{23} - \overline{x_2})(x_{23} - \overline{x_2}) + (x_{24} - \overline{x_2})(x_{24} - \overline{x_2})}{4-1}$$

$$= \frac{(1-3.5)(1-3.5) + (3-3.5)(3-3.5) + (5-3.5)(5-3.5) + (5-3.5)(5-3.5)}{4-1}$$

$$= \frac{(-2.5)(-2.5) + (-0.5)(-0.5) + (1.5)(1.5) + (1.5)(1.5)}{4-1} = \frac{11}{3} \approx 3.67$$

$$Covariance\ atrix = \begin{bmatrix} 4.33 & 3.67 \\ 3.67 & 3.67 \end{bmatrix}$$

- **ขั้นตอนที่ 3: คำนวณ Eigenvalues (λ)**
- เราต้องการหาค่า λ ที่ทำให้สมการ det(C−λI)=0 เป็นจริง

$$det\left[\begin{pmatrix}4.33 & 3.67\\3.67 & 3.67\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right] = 0$$

$$\det\left[\begin{pmatrix}4.33 & 3.67\\3.67 & 3.67\end{pmatrix} - \lambda\begin{pmatrix}\lambda & 0\\0 & \lambda\end{pmatrix}\right] = 0$$

$$\det\left[\begin{pmatrix}4.33-\lambda & 3.67\\3.67 & 3.67-\lambda\end{pmatrix}\right] = 0$$

13.47

(4.33 - λ) (3.67 - λ)

(4.33 - λ) (3.67 - λ) - 13.47 = 0

λ^2 - 8λ + 15.89 - 13.47 = 0

λ^2 - 8λ + 2.42 = 0

(λ – 7.685) (λ – 0.315) = 0

- แก้สมการกำลังสอง ได้ Eigenvalues:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\lambda = \frac{-(-8) \pm \sqrt{(-8)^2 - 4(1)(2.42)}}{2(1)}$$

$$\lambda = \frac{8 \pm \sqrt{64 - 9.68}}{2}$$

$$\lambda = \frac{8 \pm \sqrt{54.32}}{2}$$

$$\lambda = \frac{8 \pm 7.37}{2}$$

$$\lambda = \frac{8 + 7.37}{2} = \frac{15.37}{2} \approx 7.685$$

$$\lambda = \frac{8 - 7.37}{2} = \frac{0.63}{2} \approx 0.315$$

- $\lambda_1 \approx 7.685$ (Principal Eigenvalue)
- $\lambda_2 \approx 0.315$

- หา Eigenvector สำหรับ Eigenvalue ที่ใหญ่ที่สุด ($\lambda_1 \approx 7.685$)
- แก้สมการ $(C-\lambda_1 I)v_1=0$

$$\left[\begin{pmatrix} 4.33 & 3.67 \\ 3.67 & 3.67 \end{pmatrix} - \lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\left[\begin{pmatrix} 4.33 & 3.67 \\ 3.67 & 3.67 \end{pmatrix} - 7.685 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\left[\begin{pmatrix} 4.33 & 3.67 \\ 3.67 & 3.67 \end{pmatrix} - \begin{pmatrix} 7.685 & 0 \\ 0 & 7.685 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} 4.33 - 7.685 & 3.67 \\ 3.67 & 3.67 - 7.685 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} -3.355 & 3.67 \\ 3.67 & -4.015 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} -3.355x + 3.67y \\ 3.67x - 4.015y \end{pmatrix} = 0$$

$$-3.355x + 3.67y = 0$$

$$y = \frac{3.355}{3.67}x \approx 0.91x$$

$$3.67x - 4.02y = 0$$

$$y = \frac{3.67}{4.015}x \approx 0.91x$$

- การแปลงเป็นเวกเตอร์หนึ่งหน่วย (Normalization)

y ≈ 0.91x

1.เวกเตอร์ตั้งต้น: จาก y ≈ 0.91x และเราเลือก x=1 จะได้เวกเตอร์ v = [1, 0.91]

2.คำนวณขนาด (Magnitude): ใช้ทฤษฎีบทพีทาโกรัส ||v|| = sqrt(x^2 + y^2)

  - ||v|| = sqrt(1^2 + 0.91^2) = sqrt(1 + 0.8281) = sqrt(1.8281) ≈ 1.352

3.ทำให้เป็นเวกเตอร์หนึ่งหน่วย: นำแต่ละส่วนประกอบมาหารด้วยขนาดของมัน û = v / ||v||

  - x_unit = 1 / 1.352 ≈ 0.739
  - y_unit = 0.91 / 1.352 ≈ 0.673

- ผลลัพธ์: เมื่อปัดเศษจะได้ Eigenvector v_1 (PC1) ≈ [0.74, 0.67]

- **ขั้นตอนที่ 4 & 5: เลือก PC1 และแปลงข้อมูล**
- เราเลือก Eigenvector ที่มี Eigenvalue สูงสุดเป็น Principal Component 1 (PC1)
  - PC1 = [0.74, 0.67]
- แปลงข้อมูล (Project Data): นำข้อมูลที่ปรับแล้ว (Centered Data) มาคูณ (Dotproduct) กับ PC1
- สูตร: New Coordinate = A' ・ PC1
  - A" = (-2.5)(0.74) + (-2.5)(0.67) ≈ -3.525
  - B" = (-0.5)(0.74) + (-0.5)(0.67) ≈ -0.705
  - C" = (0.5)(0.74) + (1.5)(0.67) ≈ 1.375
  - D" = (2.5)(0.74) + (1.5)(0.67) ≈ 2.855
- ผลลัพธ์: เราได้ลดมิติข้อมูลจาก 2D (x, y) เหลือเพียง 1D!

P(x,y)	X	Y
E	1	5
F	2	3
G	4	2
H	5	1
Mean (ค่าเฉลี่ย)	(1 + 2 + 4 + 5) / 4 = 3	(5 + 3 + 2 + 1) / 4 = 2.75

- **ขั้นตอนที่ 1: หาค่าเฉลี่ยและปรับข้อมูล (Center the Data)**

P'(x,y)	X	Y
E '	1 - 3 = -2	5 - 2.75 = 2.25
F'	2 – 3 = -1	3 - 2.75 = 0.25
G'	4 - 3 = 1	2 - 2.75 = -0.75
H'	5 - 3 = 2	1 - 2.75 = -1.75

- **ขั้นตอนที่ 2: คำนวณ Covariance Matrix (C)**
- ใช้ข้อมูลที่ปรับแล้ว (Centered Data) มาคำนวณ

$$Cov(X_1, X_1) = \frac{\sum_{k=1}^{4}(x_{1k} - \overline{x_1})(x_{1k} - \overline{x_1})}{4-1}$$

$$= \frac{(x_{11} - \overline{x_1})(x_{11} - \overline{x_1}) + (x_{12} - \overline{x_1})(x_{12} - \overline{x_1}) + (x_{13} - \overline{x_1})(x_{13} - \overline{x_1}) + (x_{14} - \overline{x_1})(x_{14} - \overline{x_1})}{4-1}$$

$$= \frac{(1-3)(1-3) + (2-3)(2-3) + (4-3)(4-3) + (5-3)(5-3)}{4-1}$$

$$= \frac{(-2)(-2) + (-1)(-1) + (1)(1) + (2)(2)}{4-1} = \frac{8}{3} \approx 3.33$$

$$Cov(X_1, X_2) = \frac{\sum_{k=1}^{4}(x_{1k} - \overline{x_1})(x_{2k} - \overline{x_2})}{4-1}$$

$$= \frac{(x_{11} - \overline{x_1})(x_{21} - \overline{x_2}) + (x_{12} - \overline{x_1})(x_{22} - \overline{x_2}) + (x_{13} - \overline{x_1})(x_{23} - \overline{x_2}) + (x_{14} - \overline{x_1})(x_{24} - \overline{x_2})}{4-1}$$

$$= \frac{(1-3)(5-2.75) + (2-3)(3-2.75) + (4-3)(2-2.75) + (5-3)(1-2.75)}{4-1}$$

$$= \frac{(-2)(2.25) + (-1)(0.25) + (1)(-0.75) + (2)(-1.75)}{4-1} = -\frac{9}{3} \approx -3$$

$$Cov(X_2, X_1) = \frac{\sum_{k=1}^{4}(x_{2k} - \overline{x_2})(x_{1k} - \overline{x_1})}{4-1}$$

$$= \frac{(x_{21} - \overline{x_2})(x_{11} - \overline{x_1}) + (x_{22} - \overline{x_2})(x_{12} - \overline{x_1}) + (x_{23} - \overline{x_2})(x_{13} - \overline{x_1}) + (x_{24} - \overline{x_2})(x_{14} - \overline{x_1})}{4-1}$$

$$= \frac{(5 - 2.75)(1 - 3) + (3 - 2.75)(2 - 3) + (2 - 2.75)(4 - 3) + (1 - 2.75)(5 - 3)}{4-1}$$

$$= \frac{(2.25)(-2) + (0.25)(-1) + (-0.75)(1) + (-1.75)(2)}{4-1} = -\frac{9}{3} \approx -3$$

$$Cov(X_2, X_2) = \frac{\sum_{k=1}^{4}(x_{2k} - \overline{x_2})(x_{2k} - \overline{x_2})}{4-1}$$

$$= \frac{(x_{21} - \overline{x_2})(x_{21} - \overline{x_2}) + (x_{22} - \overline{x_2})(x_{22} - \overline{x_2}) + (x_{23} - \overline{x_2})(x_{23} - \overline{x_2}) + (x_{24} - \overline{x_2})(x_{24} - \overline{x_2})}{4-1}$$

$$= \frac{(5 - 2.75)(5 - 2.75) + (3 - 2.75)(3 - 2.75) + (2 - 2.75)(2 - 2.75) + (1 - 2.75)(1 - 2.75)}{4-1}$$

$$= \frac{(2.25)(2.25) + (0.25)(0.25) + (-0.75)(-0.75) + (-1.75)(-1.75)}{4-1} = \frac{8.74}{3} \approx 2.91$$

$$Covariance\ atrix = \begin{bmatrix} 3.33 & -3 \\ -3 & 2.91 \end{bmatrix}$$

- **ขั้นตอนที่ 3: คำนวณ Eigenvalues (λ)**
- เราต้องการหาค่า λ ที่ทำให้สมการ det(C−λI)=0 เป็นจริง

$$det\left[\begin{pmatrix} 3.33 & -3 \\ -3 & 2.91 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = 0$$

$$\det\left[\begin{pmatrix} 3.33 & -3 \\ -3 & 2.91 \end{pmatrix} - \lambda \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right] = 0$$

$$\det\left[\begin{pmatrix} 3.33 - \lambda & -3 \\ -3 & 2.91 - \lambda \end{pmatrix}\right] = 0$$

9

(3.33 - λ) (2.91 - λ)

$(3.33 - \lambda)(2.91 - \lambda) - 9 = 0$

$\lambda^2 - 6.24\lambda + 9.69 - 9 = 0$

$\lambda^2 - 6.24\lambda + 0.69 = 0$

$(\lambda - 6.127)(\lambda - 0.112) = 0$

- แก้สมการกำลังสอง ได้ Eigenvalues:

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\lambda = \frac{-(-6.24) \pm \sqrt{(-6.24)^2 - 4(1)(0.69)}}{2(1)}$$

$$\lambda = \frac{6.24 \pm \sqrt{38.9376 - 2.76}}{2}$$

$$\lambda = \frac{6.24 \pm \sqrt{36.1776}}{2}$$

$$\lambda = \frac{6.24 \pm 6.0148}{2}$$

$$\lambda_1 = \frac{6.24 + 6.0148}{2} = \frac{12.2548}{2} \approx 6.127$$

$$\lambda_2 = \frac{6.24 - 6.0148}{2} = \frac{0.2252}{2} \approx 0.112$$

  - $\lambda_1 \approx 6.127$ (Principal Eigenvalue)
  - $\lambda_2 \approx 0.112$

- หา Eigenvector สำหรับ Eigenvalue ที่ใหญ่ที่สุด ($\lambda_1 \approx 6.127$)
- แก้สมการ $(C-\lambda_1 I)v_1=0$

$$\left[\begin{pmatrix} 3.33 & -3 \\ -3 & 2.91 \end{pmatrix} - \lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\left[\begin{pmatrix} 3.33 & -3 \\ -3 & 2.91 \end{pmatrix} - 6.127 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\left[\begin{pmatrix} 3.33 & -3 \\ -3 & 2.91 \end{pmatrix} - \begin{pmatrix} 6.127 & 0 \\ 0 & 6.127 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} 3.33 - 6.127 & -3 \\ -3 & 2.91 - 6.127 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} -2.797 & -3 \\ -3 & -3.217 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} -2.797x - 3y \\ -3x - 3.217y \end{pmatrix} = 0$$

$$-2.797x - 3y = 0$$

$$y = \frac{2.797}{3} x \approx 0.93x$$

$$-3x - 3.217y = 0$$

$$y = \frac{3}{3.217} x \approx 0.93x$$

- การแปลงเป็นเวกเตอร์หนึ่งหน่วย (Normalization)

$$y \approx 0.93x$$

1.เวกเตอร์ตั้งต้น: จาก $y \approx 0.93x$ และเราเลือก $x=1$ จะได้เวกเตอร์ $v = [1, 0.93]$

2.คำนวณขนาด (Magnitude): ใช้ทฤษฎีบทพีทาโกรัส $||v|| = \text{sqrt}(x^2 + y^2)$

  - $||v|| = \text{sqrt}(1^2 + 0.93^2) = \text{sqrt}(1 + 0.8649) = \text{sqrt}(1.8649) \approx 1.365$

3.ทำให้เป็นเวกเตอร์หนึ่งหน่วย: นำแต่ละส่วนประกอบมาหารด้วยขนาดของมัน $\hat{u} = v / ||v||$

  - x_unit = 1 / 1.365 ≈ 0.732
  - y_unit = 0.93 / 1.365 ≈ 0.681

- ผลลัพธ์: เมื่อปัดเศษจะได้ Eigenvector v_1 (PC1) ≈ [0.73, 0.68]

- **ขั้นตอนที่ 4 & 5: เลือก PC1 และแปลงข้อมูล**
- เราเลือก Eigenvector ที่มี Eigenvalue สูงสุดเป็น Principal Component 1 (PC1)
  - PC1 = [0.73, 0.68]
- แปลงข้อมูล (Project Data): นำข้อมูลที่ปรับแล้ว (Centered Data) มาคูณ (Dotproduct) กับ PC1
- สูตร: New Coordinate = A' • PC1
  - A'' = (-2)(0.73) + (2.25)(0.68) ≈ 0.07
  - B'' = (-1)(0.73) + (0.25)(0.68) ≈ -0.56
  - C'' = (1)(0.73) + (-0.75)(0.68) ≈ 0.22
  - D'' = (2)(0.73) + (-1.75)(0.68) ≈ 0.27
- ผลลัพธ์: เราได้ลดมิติข้อมูลจาก 2D (x, y) เหลือเพียง 1D!