ⓐ PCA（主成分分析）

☆ 2変数をはじめに考える

$$x_1 = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} \qquad x_2 = \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix}$$

サンプルサイズ n, 特徴量サイズ 2 の ベクトル

$$u_{i1} = \frac{x_{i1} - \bar{x}_1}{\sigma_1} \qquad u_{i2} = \frac{x_{i2} - \bar{x}_2}{\sigma_2}$$ ※標準化

ベクトル u_1 u_2 の線形結合を用いて 第一成分 ベクトル z_1 を定義する

$$z_1 = a_1 u_1 + a_2 u_2 \quad \cdots (*)$$

$$\left(z_{i1} = a_1 u_{i1} + a_2 u_{i2} \right)$$

☆ 目的 … データの情報を最大限残す

→ z の分散を最大化すればよい

$$V_z = \frac{\sum_{i=1}^{n} (z_{i1} - \bar{z}_1)^2}{n}$$

u_1 u_2 が平均0と定義したので $\bar{z}_1 = 0$ も同様

$$= \frac{\sum_{i=1}^{n} z_{i1}^2 - 2\sum_{i=1}^{n} z_{i1}\bar{z}_1 + \sum_{i=1}^{n} \bar{z}_1^2}{n}$$

$$= \frac{1}{n}\sum_{i=1}^{n} (a_1 u_{i1} + a_2 u_{i2})^2$$

$$= \frac{1}{n}\left\{ a_1^2 \underbrace{\sum_{i=1}^{n} u_{i1}^2}_{※1} + 2a_1a_2 \underbrace{\sum_{i=1}^{n} u_{i1}u_{i2}}_{※2} + a_2^2 \underbrace{\sum_{i=1}^{n} u_{i2}^2}_{※1} \right\}$$

※1 $$\underset{\overset{\shortparallel}{1^2}}{\sigma_{u_1}^2} = \frac{\sum_{i=1}^{n}(u_{i1} - \bar{u}_1)^2}{n} = \frac{\sum_{i=1}^{n}(u_{i1}^2 - 2u_{i1}\underset{0}{\bar{u}_1} + \underset{0}{\bar{u}_1^2})}{n} = \frac{1}{n}\sum_{i=1}^{n} u_{i1}^2$$

※2 $$\sigma_{u_1u_2} = \frac{\sum_{i=1}^{n}(u_{i1} - \bar{u}_1)(u_{i2} - \bar{u}_2)}{n} = \frac{\sum_{i=1}^{n}(u_{i1}u_{i2} - u_{i1}\underset{0}{\bar{u}_2} - u_{i2}\underset{0}{\bar{u}_1} + \underset{0}{\bar{u}_1\bar{u}_2})}{n}$$

$$\sigma_{u_1u_2} = r_{u_1u_2}\sigma_{u_1}\sigma_{u_2} = r_{u_1u_2}$$

$$= \cancel{\frac{1}{n}}\left\{ a_1^2 \cancel{n} + 2a_1a_2 r_{u_1u_2} \cancel{n} + a_2^2 \cancel{n} \right\}$$

$$= \underbrace{a_1^2 + a_2^2}_{} + 2a_1a_2 r_{u_1u_2}$$

（$a_1^2 + a_2^2$）= 重みの2乗和 $=1$ という制約を設ける

$$= 1 + 2a_1a_2 r_{u_1u_2} \quad \cdots (*)(*)$$

→ $(*)(*)$ の最大問題を ラグランジュ未定乗数法を用いて計算する

$$f(a_1, a_2, \lambda) = a_1^2 + a_2^2 + 2a_1a_2 r_{u_1u_2} - \lambda(a_1^2 + a_2^2 - 1)$$

$$\frac{\partial f}{\partial a_1} : 2a_1 + 2r_{u_1u_2}a_2 - 2\lambda a_1 = 0$$

$$\frac{\partial f}{\partial a_2} = 2a_2 + 2r_{u_1u_2}a_1 - 2\lambda a_2 = 0$$

$$\longrightarrow \underbrace{\begin{pmatrix} 1 & r_{u_1u_2} \\ r_{u_1u_2} & 1 \end{pmatrix}}_{\text{共分散行列 } V} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

→ λ に対する固有方程式 $Va = \lambda a$ を解けば a_1 a_2 の組である固有ベクトルが求まる

Thm 固有方程式の固有値 λ は 元で定義した z の分散 と等しい

[証明]

$Va = \lambda a$ に a^T を作用する

$$\longrightarrow a^T V a = a^T \lambda a = \lambda(a_1^2 + a_2^2)$$

$$\longrightarrow (a_1 \ a_2)\begin{pmatrix} 1 & r_{u_1u_2} \\ r_{u_1u_2} & 1 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda$$

$$\longrightarrow \underbrace{a_1^2 + a_2^2 + 2a_1a_2 r_{u_1u_2}}_{V_z} = \lambda$$

$$\longrightarrow V_z = \lambda$$ 証明おわり

結局, x_1 x_2 を用いて, 生成された 第一主成分は

$$z_1 = a_1 u_1 + a_2 u_2 \quad //$$

☆ 実装において

$$W = \left(\underset{\text{eigen1}}{\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}} \underset{\text{eigen2}}{\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}} \right) \qquad x = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix}$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

$$= \left(\underset{\text{第一主成分}}{\boxed{\begin{matrix} a_1x_{11} + a_2x_{12} \\ a_1x_{21} + a_2x_{22} \end{matrix}}} \ \underset{\text{第二主成分}}{\boxed{\begin{matrix} b_1x_{11} + b_2x_{12} \\ b_1x_{21} + b_2x_{22} \end{matrix}}} \right)$$

○ 第二主成分を考える。

$z_2 = b_1 \mathbf{u}_1 + b_2 \mathbf{u}_2$ と $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ で定義できるものとする

第一主成分を考えたときと同様に $b_1^2 + b_2^2 = 1$ という制約を用いる

$\longrightarrow$ z_1 と z_2 には相関がないとする

（z_1 に含まれないものを以下 z_2 z_3 … で定義するから）

$V[z_1]$, $V[z_2] \neq 0$ であることから、$\mathrm{Cov}[z_1 z_2] = 0$ である

$$\longrightarrow \frac{1}{n}\sum_{i=1}^{n}(z_{i1}-\bar{z}_1)(z_{i2}-\bar{z}_2) = 0$$

$$\longrightarrow \frac{1}{n}\left\{\sum_{i=1}^{n} z_{i1}z_{i2} - z_{i1}\underbrace{\bar{z}_2}_{=0} - z_{i2}\underbrace{\bar{z}_1}_{=0} + \underbrace{\bar{z}_1}_{=0}\underbrace{\bar{z}_2}_{=0}\right\} = 0$$

$$\longrightarrow \frac{1}{n}\sum_{i=1}^{n} z_{i1}z_{i2} = \frac{1}{n}\sum_{i=1}^{n}\{a_1u_{i1}+a_2u_{i2}\}\{b_1u_{i1}+b_2u_{i2}\} = 0$$

$$\longrightarrow \frac{1}{n}\sum_{i=1}^{n}\{a_1b_1\underbrace{u_{i1}^2}_{\text{※1}} + a_1b_2\underbrace{u_{i1}u_{i2}}_{\text{※2}} + a_2b_1\underbrace{u_{i1}u_{i2}}_{\text{※2}} + a_2b_2\underbrace{u_{i2}^2}_{\text{※1}}\} = 0$$

※1 $$\underbrace{\sigma_{u_1}^2}_{=1^2} = \frac{\sum_{i=1}^{n}(u_{i1}-\bar{u}_1)^2}{n} = \frac{\sum_{i=1}^{n}(u_{i1}^2 - 2u_{i1}\underbrace{\bar{u}_1}_{0} + \underbrace{\bar{u}_1^2}_{0})}{n} = \frac{1}{n}\sum_{i=1}^{n}u_{i1}^2$$

※2 $$\underbrace{\sigma_{u_1u_2}}_{=r_{u_1u_2}\sigma_{u_1}\sigma_{u_2}} = \frac{\sum_{i=1}^{n}(u_{i1}-\bar{u}_1)(u_{i2}-\bar{u}_2)}{n} = \frac{\sum_{i=1}^{n}(u_{i1}u_{i2} - u_{i1}\underbrace{\bar{u}_2}_{0} - u_{i2}\underbrace{\bar{u}_1}_{0} + \underbrace{\bar{u}_1\bar{u}_2}_{0})}{n}$$

$= r_{u_1u_2}$

$$\longrightarrow a_1b_1 + a_1b_2 r_{u_1u_2} + a_2b_1 r_{u_1u_2} + a_2b_2 = 0$$

$$\longrightarrow (a_1\ a_2)\begin{pmatrix} 1 & r_{u_1u_2} \\ r_{u_1u_2} & 1 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = 0$$

$$\longrightarrow \mathbf{a}^T V \mathbf{b} = 0$$

※ 第一主成分を求める際に得られた $V\mathbf{a} = \lambda_1 \mathbf{a}$ を用いた

転置

$(V\mathbf{a})^T = \lambda_1 \mathbf{a}^T$

$\longrightarrow \mathbf{a}^T V = \lambda_1 \mathbf{a}^T$

$$\longrightarrow \lambda_1 \mathbf{a}^T \mathbf{b} = 0 \quad ※$$

つまり、制約条件 $a_1b_1 + a_2b_2 = 0$

$V_{z_2} = b_1^2 + b_2^2 + 2r_{u_1u_2}b_1b_2$ の最大化を $b_1^2 + b_2^2 = 1$ と $a_1b_1 + a_2b_2 = 0$ を使って行う

$g = b_1^2 + b_2^2 + 2r_{u_1u_2}b_1b_2 - \lambda(b_1^2 + b_2^2 - 1) - \eta(a_1b_1 + a_2b_2)$ とする

$$\frac{\partial g}{\partial b_1} = 2b_1 + 2r_{u_1u_2}b_2 - 2\lambda b_1 - \eta a_1 = 0$$

$$\frac{\partial g}{\partial b_2} = 2b_2 + 2r_{u_1u_2}b_1 - 2\lambda b_2 - \eta a_2 = 0$$

$$\longrightarrow V\mathbf{b} = \lambda\mathbf{b} + \frac{\eta}{2}\mathbf{a}$$

$\mathbf{a}^T$ を左から作用

$$\longrightarrow \underbrace{\mathbf{a}^T V \mathbf{b}}_{=0} = \underbrace{\mathbf{a}\lambda\mathbf{b}}_{=0} + \frac{\eta}{2}\mathbf{a}^T\mathbf{a}$$

$$\longrightarrow \eta = 0$$

結局 $V\mathbf{b} = \lambda\mathbf{b}$ と第一主成分のときと同様