Principal Components Analysis (PCA)

Variables (因) (果)

$X_1 \quad X_2 \quad \cdots \quad X_p \qquad Y$

N

$\underbrace{\qquad\qquad\qquad}_{p}$

Samples

$X_1, X_2, \ldots, X_p \to$ Machine Learning $\to Y$

變數太多

變數存在相關性

⇒ 負面的影響

→ Pre-Processing

$X_1 \quad X_2 \quad \cdots \quad X_p$

建立 → $\underbrace{Z_1 \quad Z_2}_{q} \quad \cdots \quad Z_p$

→ Z_i 不相關

→ $\sigma^2_{Z_1} > \sigma^2_{Z_2} > \cdots > \sigma^2_{Z_p}$

Z的變異數 Var(Z)

↳ 訊息能量

信

→ 總訊息能量不變

$\boxed{q << p}$

q個主成份

p 個主成份 —→ 總訊息能量不變

風險：損失部份訊息

$$\sigma_{X_1}^2+\sigma_{X_2}^2+\cdots+\sigma_{X_p}^2 = \sigma_{Z_1}^2+\sigma_{Z_2}^2+\cdots+\sigma_{Z_p}^2$$

$$\frac{\sigma_{Z_1}^2+\sigma_{Z_2}^2+\cdots+\sigma_{Z_q}^2}{\sum_{i=1}^{p}\sigma_{X_i}^2} = 70\%,\ 80\%\cdots$$

$\sum_{i=1}^{p}\sigma_{X_i}^2$ ← 總訊息能量

70% ↓ 60%

新變數

Z_i 從何而來？

Linear Combination

$$\begin{aligned} Z_1 &= \alpha_{11}X_1+\alpha_{12}X_2+\cdots+\alpha_{1p}X_p \\ Z_2 &= \alpha_{21}X_1+\alpha_{22}X_2+\cdots+\alpha_{2p}X_p \\ &\ \vdots \\ Z_p &= \alpha_{p1}X_1+\alpha_{p2}X_2+\cdots+\alpha_{pp}X_p \end{aligned}$$

→ 線性組合

使得 ① $E(Z_iZ_j)=0,\ i\neq j$

（假設 $E(Z_i)=0$）

② $\sigma_{Z_1}^2>\sigma_{Z_2}^2\cdots>\sigma_{Z_p}^2$

$$\underset{p\times1}{\underline{Z}}=\begin{bmatrix} Z_1\\ Z_2\\ \vdots\\ Z_p\end{bmatrix}=\begin{bmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p}\\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pp}\end{bmatrix}\begin{bmatrix}X_1\\ X_2\\ \vdots\\ X_p\end{bmatrix}=\underset{p\times p}{A}\ \underset{p\times 1}{\underline{X}}$$

LaTeX \cdots

LaTeX \cdots

New $\underset{\sim}{z} = \underset{\sim}{A}\, \underset{\sim}{x}$ OLD

轉換矩陣 $p \times p$

Cov($\underset{\sim}{X}$)

$A = ?$ α_{ij} ?

$\underset{\sim}{z}$ 的共變異矩陣 Covariance Matrix

$$\Sigma_{\underset{\sim}{z}} = \mathrm{Cov}(\underset{\sim}{z}) = E(\underset{\sim}{z}\,\underset{\sim}{z}^T)$$

numpy.cov

$$= E\left(\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix}\begin{bmatrix} z_1 & z_2 & \cdots & z_p \end{bmatrix}\right) = E\left(\begin{bmatrix} z_1^2 & z_1 z_2 & \cdots & z_1 z_p \\ z_2 z_1 & z_2^2 & \cdots & z_2 z_p \\ \vdots & \vdots & & \\ z_p z_1 & z_p z_2 & \cdots & z_p z_p \end{bmatrix}\right)$$

$$= \begin{bmatrix} \overset{\sigma_{z_1}^2}{E(z_1^2)} & 0 & \cdots & 0 \\ 0 & E(z_2^2)\ \sigma_{z_2}^2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \overset{\sigma_{z_p}^2}{E(z_p^2)} \end{bmatrix}$$ 對角矩陣

$\sigma_{z_i}^2 = Var(z_i)$

$$\underset{\sim}{z} \longrightarrow \mathrm{Cov}(z) = \begin{bmatrix} \sigma_{z_1}^2 & & 0 \\ & \sigma_{z_2}^2 & \\ 0 & & \ddots \ \sigma_{z_p}^2 \end{bmatrix}$$

$\underset{\sim}{A}\,\underset{\sim}{x}$

? ↑ cov($\underset{\sim}{x}$) 並非對角矩陣

$\boxed{E(\underset{\sim}{z}) = \underset{\sim}{0}}$

$$\Sigma_z = \mathrm{cov}(\underset{\sim}{z}) = E(\underset{\sim}{z}\,\underset{\sim}{z}^T)$$

$$\Sigma_{\underline{z}} = \text{cov}(\underline{z}) = E(\underline{z}\,\underline{z}^T)$$

$$E\left((\underline{z}-E(\underline{z}))(\underline{z}-E(\underline{z}))^T\right)$$

$$= E(A\underline{x}(A\underline{x})^T) = E(A\underline{x}\,\underline{x}^TA^T)$$

$$= A\,\underbrace{E(\underline{x}\,\underline{x}^T)}\,A^T = A\Sigma_{\underline{x}}A^T$$

原變數的共變異矩陣

$$\Sigma_{\underline{z}} = A\Sigma_{\underline{x}}A^T$$

$$\begin{bmatrix} \lambda & 0 \\ 0 & \ddots \end{bmatrix} = ?\ [\ ]\ ?$$

Recall:

$$\underbrace{P^{-1}}\underset{n\times n}{Q}\underbrace{P} = D \leftarrow \begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix}$$

Diagonal matrix

$$P = [\underbrace{\underline{v}_1\ \ \underline{v}_2\ \cdots\ \underline{v}_n}], \quad \text{where } \boxed{Q\underline{v}_k = \lambda_k\underline{v}_k}$$

eigenvectors of Q

特徵向量

Q^{100}

$$P^{-1}QP = D$$

$$Q\underline{v}_k = \lambda_k\underline{v}_k \leftarrow \text{eigen-decomposition}$$

$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$

$k = 1, 2, \cdots, n$

$k = 1, 2, \cdots, n$

$$Q[\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_n] = [\lambda_1\underline{v}_1\ \lambda_2\underline{v}_2 \cdots \lambda_n\underline{v}_n]$$

eigenvalues

$$= [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_n]\underbrace{\begin{bmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{bmatrix}}_{D}$$

大

$$QP = PD$$

$$P^{-1}QP = P^{-1}PD = D$$

對比

$$\begin{cases}\Sigma_{\underline{z}} = A\Sigma_{\underline{x}}A^T \\ D = P^{-1}QP\end{cases}$$

原變數 $\underline{x}$ 的共變異矩陣

$$\Rightarrow \quad A^T = [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_p]$$

特徵向量

where $\Sigma_{\underline{x}}\underline{v}_k = \lambda_k\underline{v}_k$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$$

$$\Sigma_{\underline{z}} = \begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_p\end{bmatrix} = \begin{bmatrix}\sigma_{z_1}^2 & & & 0\\ & \sigma_{z_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{z_p}^2\end{bmatrix}$$

結論：

$$\underline{z} = A\underline{x} \rightarrow \Sigma_{\underline{z}} = \begin{bmatrix}\sigma_{z_1}^2 & & & 0\\ & \sigma_{z_2}^2 & & \\ & & \ddots & \\ 0 & & & \sigma_{z_p}^2\end{bmatrix}$$

大

$$A^T = [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_p], \quad \Sigma_{\underline{x}}\underline{v}_k = \lambda_k\underline{v}_k$$

$$A = \begin{bmatrix}\underline{v}_1^T \\ \underline{v}_2^T \\ \vdots\end{bmatrix}$$

$$\hookrightarrow A = \begin{bmatrix} \underline{v}_1^T \\ \underline{v}_2^T \\ \vdots \\ \underline{v}_p^T \end{bmatrix}$$

$$\underline{z} = A\underline{x} = \begin{bmatrix} \underline{v}_1^T \\ \underline{v}_2^T \\ \vdots \\ \underline{v}_p^T \end{bmatrix} \underline{x} = \begin{bmatrix} \underline{v}_1^T\underline{x} \\ \underline{v}_2^T\underline{x} \\ \vdots \\ \underline{v}_p^T\underline{x} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix}$$

$$z_i = \underline{v}_i^T \underline{x}$$

$$= \begin{bmatrix} \underline{v}_i(1) & \underline{v}_i(2) & \cdots & \underline{v}_i(p) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

第i個 新變數

$$z_i = \underline{v}_i(1)\underline{x_1} + \underline{v}_i(2)\underline{x_2} + \cdots + \underline{v}_i(p)\underline{x_p}$$

$\underline{v}_i$: 第 i 個 eigenvector of $\Sigma_{\underline{x}}$

file name

SML_HW_pca .ipynb

.pdf

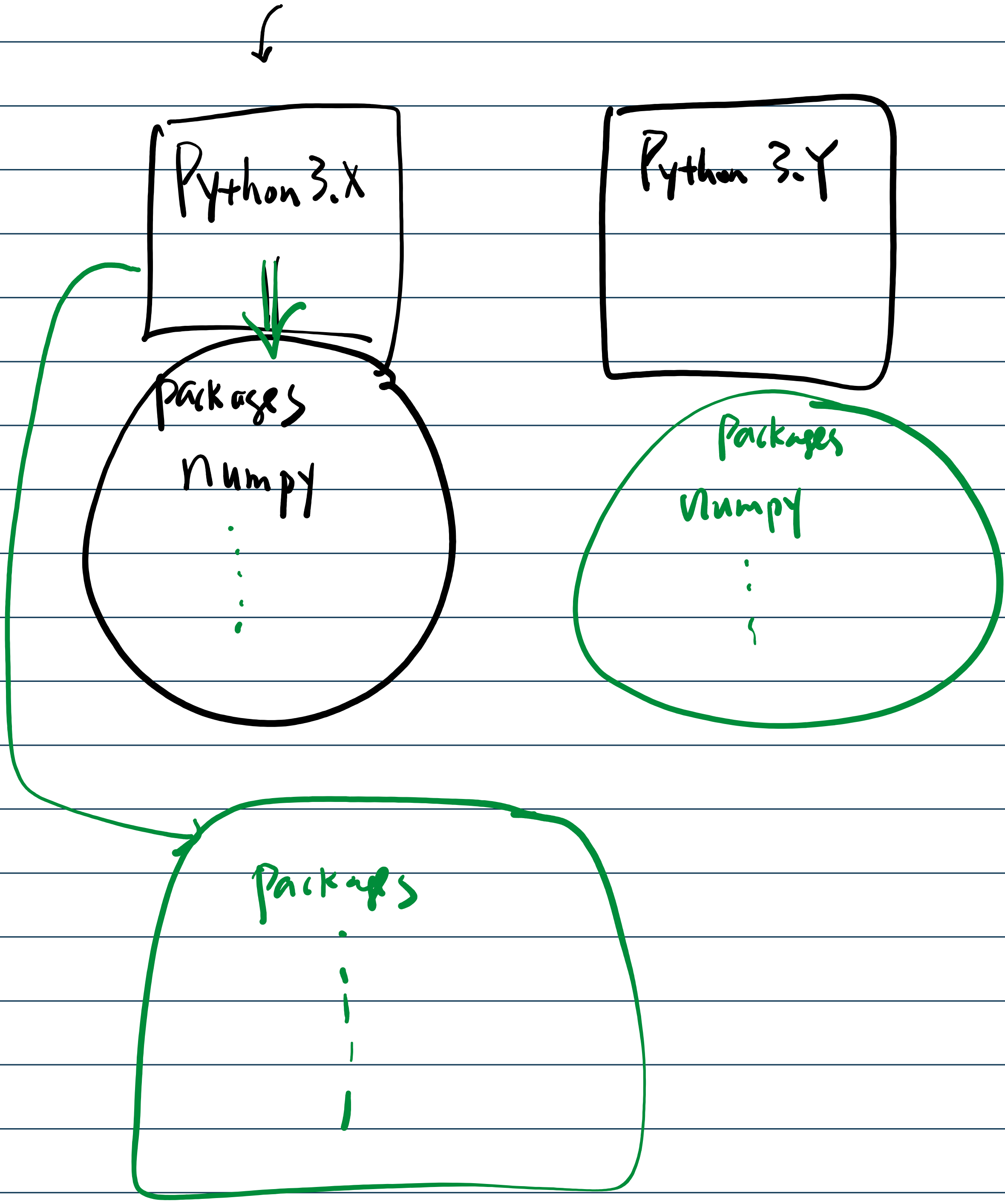
SML_Inclass_0226

Inclass_0226

□□□.ipynb

1

AbcDefGhi


Python 3.X
Packages
numpy
Python 3.Y
Packages
numpy
Packages