PCA (吳俊德實驗室提供)

Random Vector X

$\tilde{X} = X - E[X]$

$\therefore \tilde{X}$: zero-mean vector

$y_1 = \vec{e}_1^T \tilde{X}$

$\because E[\tilde{X}] = 0 \therefore E[y_1] = 0$

$\text{Var}(y_1) = E[y_1^2]$

$= \text{E}[\vec{e}_1^T \tilde{X} (\tilde{X}^T \vec{e}_1)]$

$= \vec{e}_1^T E[\tilde{X}\tilde{X}^T] \vec{e}_1$

$= \vec{e}_1^T \Sigma_{\tilde{X}} \vec{e}_1$

$\Sigma_{\tilde{X}}$: covariance of $\tilde{X}$

$\max_{\vec{e}_1}[\vec{e}_1^T \Sigma_{\tilde{X}} \vec{e}_1]$:condition $|\vec{e}_1|^2 = \vec{e}_1^T \vec{e}_1 = 1$

Apply Lagrange Principle as follows.

$J(\vec{e}_1 \lambda_1) = \vec{e}_1^T \Sigma_{\tilde{X}} \vec{e}_1 - \lambda_1 [\vec{e}_1^T \vec{e}_1 - 1]$

$\frac{\partial J}{\partial \vec{e}_1} = 2\Sigma_{\tilde{X}} \vec{e}_1 - 2\lambda_1 \vec{e}_1 = 0$

$\Sigma_{\tilde{X}} \vec{e}_1 = \lambda_1 \vec{e}_1$

$A\vec{X} = \lambda \vec{X}$
λ is an eigenvalue of A
$\vec{X}$ is an eigenvector of A

$\therefore \vec{e}_1$ is an eigenvector of $\Sigma_{\tilde{X}} \Rightarrow$ corresponding to the largest eigenvector

λ_1 is an eigenvector of $\Sigma_{\tilde{X}}$

Max:

$$\vec{e}_1^T \Sigma_{\tilde{X}} \vec{e}_1 = \vec{e}_1^T \lambda_1 \vec{e}_1$$

$$= \vec{e}_1^T \vec{e}_1 \lambda_1$$

$$= \lambda_1$$

Let $y_2 = \vec{e}_2^T \tilde{X}$

max $\mathrm{var}(y_2)$, $\vec{e}_1 \perp \vec{e}_2 (\vec{e}_1^T \cdot \vec{e}_2 = 0)$, $\vec{e}_2^T \vec{e}_2 = 1$

$$\vec{e}_2^T \Sigma_{\tilde{X}} \vec{e}_2 - a_1 \left(\vec{e}_1^T \vec{e}_2 - 0 \right) - \lambda_2 \left(\vec{e}_2 \vec{e}_2 - 1 \right)$$

$$(1) \frac{\partial(\bullet)}{\partial \vec{e}_2} = 2\Sigma_{\tilde{X}} \vec{e}_2 - a_1 \vec{e}_1 - 2\lambda_2 \vec{e}_2 = \vec{0}$$

$$(2) \frac{\partial(\bullet)}{\partial \vec{e}_2} = \vec{e}_1^T \vec{e}_2 = 0$$

$$(3) \frac{\partial(\bullet)}{\partial \lambda_2} = \vec{e}_2^T \vec{e}_2 - 1 = 0$$

證 $a_1 = 0$

$$2\Sigma_{\tilde{X}} \vec{e}_2 - a_1 \vec{e}_1 - 2\lambda_2 \vec{e}_2 = \vec{0}$$

$$\vec{e}_1^T \left(2\Sigma_{\tilde{X}} \vec{e}_2 - a_1 \vec{e}_1 - 2\lambda_2 \vec{e}_2 \right) = 0$$

$$2\vec{e}_1^T \vec{e}_2 \Sigma_{\tilde{X}} - a_1 \vec{e}_1 \vec{e}_1^T - 2\lambda_2 \vec{e}_1^T \vec{e}_2 = \vec{0}$$

$$\therefore a_1 = 0$$

$$\therefore \Sigma_{\tilde{X}} \vec{e}_2 = \lambda_2 \vec{e}_2$$

$\vec{e}_2$ is an eigenvector

λ_2 is the 2nd largest eigenvalue

y_2 and $y_1 \Rightarrow$ uncorrelated

$\downarrow \qquad \downarrow$

$\vec{e}_2^T \tilde{X} \quad \vec{e}_1^T \tilde{X}$

$E[\left(\vec{e}_1^T \vec{X}\right)\left(\vec{e}_2^T \vec{X}\right)] = E[\vec{e}_1^T \vec{X}]E[\vec{e}_2^T \vec{X}] = 0$

$E[\vec{e}_1^T \vec{X}\vec{X}^T \vec{e}_2] = 0$

$\vec{e}_1^T \Sigma_{\tilde{X}} \vec{e}_2 = 0$

$\lambda \vec{e}_1^T \vec{e}_2 = 0$

$\because \lambda \neq 0$

$\therefore \vec{e}_1^T \vec{e}_2 = 0 \Rightarrow$ 內積為零 $\therefore$ uncorrelated

PCA: n-dim $\tilde{X} \rightarrow$ m-dim $\tilde{y}$

$\Sigma_{\tilde{X}}$: covarience of $\tilde{X}$

$\downarrow$ eigendecomposition

$\left\{ \underset{n\times1}{\vec{e}_1}, \underset{n\times1}{\vec{e}_2}, \cdots \underset{n\times1}{\vec{e}_m} \right\}$

largest $\lambda_1, \lambda_2, \cdots \lambda_m$

$$\tilde{y}_{m\times1} = \begin{bmatrix} \vec{e}_1^T \\ \vec{e}_2^T \\ \vdots \\ \vec{e}_m^T \end{bmatrix} \tilde{X}_{n\times1} \quad ,n > m$$

Lagrange Multiplier

$$\max_{\tilde{X}} f(X) : \text{subject to } g_1(x) = 0$$

$$\max_{X,\lambda_1} \tilde{f}(X, \lambda_1) = f(X) + \lambda_1 (g_1(X) - 0)$$

$$\frac{\partial \tilde{f}}{\partial X} = \frac{\partial f}{\partial X} + \lambda_1 \frac{\partial g}{\partial X}$$

$$\frac{\partial \tilde{f}}{\partial \lambda_1} = g_1(X) = 0$$

$$\text{ex} : g_1(x) = 0 \quad g_2(x) = 0 \quad g_3(x) = 0$$

$$f(X) + \lambda_1 (g_1(x) - 0) + \lambda_2 (g_2(x) - 0) + \lambda_3 (g_3(x) - 0)$$

$$\frac{\partial \tilde{f}}{\partial X} = \frac{\partial f}{\partial X} + \lambda_1 \frac{\partial g_1}{\partial X} + \lambda_2 \frac{\partial g_2}{\partial X} + \lambda_3 \frac{\partial g_3}{\partial X}$$

$$\frac{\partial \tilde{f}}{\partial \lambda_1} = g_1 = 0$$

$$\frac{\partial \tilde{f}}{\partial \lambda_1} = g_2 = 0$$

$$\frac{\partial \tilde{f}}{\partial \lambda_1} = g_3 = 0$$