LDA (吳俊德實驗室提供)

S_w : within-class matrix (每一類中自己的分散程度)

S_B : between-class matrix(類與類間的分散程度)

$j = 1, 2, \ldots J$

$X^{(1)} X^{(2)} \ldots X^{(J)}$

$\downarrow \quad \downarrow \qquad \downarrow$

$\mu^{(1)} \mu^{(2)} \ldots \mu^{(J)} =$ mean vector

$\Sigma^{(1)} \Sigma^{(2)} \ldots \ \Sigma^{(J)} =$ covarience matrix

$w^{(1)} w^{(2)} \ldots \ w^{(J)} =$ weight scalar

matrix $S_w = \sum_{j=1}^{J} w^{(j)} \Sigma^{(j)}$

matrix $S_B = \sum_{j=1}^{J} w^{(j)} [\mu^{(j)} - \mu][\mu^{(j)} - \mu]^T \Rightarrow$ vector $\mu = \sum_{j=1}^{J} w^{(j)} \mu^{(j)}$

scalar $y = e^T X$

$y^{(1)} y^{(2)} \ldots y^{(J)}$

$m^{(1)} m^{(2)} \ldots m^{(J)} \Rightarrow m^{(j)} = e^T \mu^{(j)}$

$\sigma^2_{(1)} \sigma^2_{(2)} \ldots \sigma^2_{(J)} \Rightarrow \sigma^2_{(j)} = e^T \Sigma^{(j)} e$

$w^{(1)} w^{(2)} \ldots w^{(J)}$

within-class varience

$$\sigma_B^2 = \sum_{j=1}^{\mathrm{J}} w^{(j)} \sigma^2_{(j)} = \sum_{j=1}^{J} w^{(j)} [e^T \Sigma^{(j)} e] = e^T S_W e$$

between-class varience

$$\begin{aligned}
\sigma_B^2 &= \sum_{j=1}^{J} w^{(j)} [m^{(j)} - m][m^{(j)} - m]^T \\
&= \sum_{j=1}^{J} w^{(j)} [e^T \mu^{(j)} - e^T \mu][e^T \mu^{(j)} - e^T \mu]^T \\
&= e^T \left\{ \sum_{j=1}^{J} w^{(j)} [\mu^{(j)} - \mu][\mu^{(j)} - \mu]^T \right\} e \\
&= e^T S_B e
\end{aligned}$$

$$\max J(e) = \frac{e^T S_B e}{e^T S_W e} \quad \begin{matrix} \longrightarrow \text{越大越好} \\ \longrightarrow \text{越小越好} \end{matrix}$$

$e : S_W^{-1} S_B$的max eigenvector

$\rightarrow$ max eigenvalue

$$\frac{\partial J}{\partial e}=\frac{2S_Be\left(e^TS_We\right)-2S_We\left(e^TS_Be\right)}{\left(e^TS_We\right)^2}=0$$

$$S_Be\left(e^TS_We\right)=S_We\left(e^TS_Be\right)$$

$$S_B\vec{e}=\left(\frac{e^TS_Be}{e^TS_We}\right)S_W\vec{e}$$

$$\lambda=\frac{e^TS_Be}{e^TS_We}$$

$$S_B\vec{e}=\lambda S_W\vec{e}$$

$$S_W^{-1}S_B\vec{e}=\lambda\vec{e}$$

$$\boxed{\begin{aligned}Ax&=\lambda Bx\\ B^{-1}Ax&=\lambda x\end{aligned}}$$

$\vec{e}$ is $S_W^{-1}S_B$的eigenvector

symmetric

PCA: Σ covarience

$\searrow \lambda$: real ≥ 0

LDA:$S_W^{-1}S_B \Rightarrow$ non-symmetric

$$S_B=S_B^T,\quad S_W=S_W^T$$

$$\left(S_W^{-1}S_B\right)^T=S_B^T\left(S_W^{-1}\right)^T\neq S_w^{-1}S_B$$

PCA: $\max e^T\left(\overset{\uparrow}{S_W}+\overset{\uparrow}{S_B}\right)e$

LDA: $\max\dfrac{e^TS_Be}{e^TS_We}\begin{matrix}\uparrow\\ \downarrow\end{matrix}$

LDA 優於PCA為$S_B\uparrow$, $S_W\downarrow$