# 线性判别分析推导

$$X=(x_1\ x_2 \cdots x_N)^T=\begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{pmatrix}_{N\times p}$$

$$Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}_{N\times 1}$$

$$\{(x_i, y_i)\}_{i=1}^{N}, \quad x_i \in \mathbb{R}^p, \quad y_i \in \{\underset{C_1}{+1}, \underset{C_2}{-1}\}$$

$$x_{c_1}=\{x_i \mid y_i=+1\}, \quad x_{c_2}=\{x_i \mid y_i=-1\}$$

$$|x_{c_1}|=N_1, \quad |x_{c_2}|=N_2, \quad N_1+N_2=N$$

类内小，类间大

$$z_i = w^T x_i \qquad \|w\|=1$$

$$\bar{z} = \frac{1}{N}\sum_{i=1}^{N} w^T x_i$$

$$S_z = \frac{1}{N}\sum_{i=1}^{N}(z_i-\bar{z})(z_i-\bar{z})^T$$

$$= \frac{1}{N}\sum_{i=1}^{N}(w^T x_i-\bar{z})(w^T x_i-\bar{z})^T$$

$C_1$: $$\bar{z}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} w^T x_i$$

$$s_1 = \frac{1}{N_1}\sum_{i=1}^{N}(w^T x_i-\bar{z}_1)(w^T x_i-\bar{z}_1)^T =$$

$C_2$: $$\bar{z}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} w^T x_i$$

$$S_2 = \frac{1}{N_2}\sum_{i=1}^{N}(w^T x_i-\bar{z}_2)(w^T x_i-\bar{z}_2)^T$$

类间: $(\bar{z}_1-\bar{z}_2)^2$

类内: $s_1+S_2$

目标函数: $$J(w)=\frac{(\bar{z}_1-\bar{z}_2)^2}{s_1+S_2}$$

$$\hat{w} = \arg\max_{w} J(w)$$

$$(\bar{z}_1-\bar{z}_2)^2 = \left[\frac{1}{N_1}\sum w^T x_i - \frac{1}{N_2}\sum w^T x_i\right]^2$$

$$= \left\{w^T \cdot \left(\frac{1}{N_1}\sum x_i - \frac{1}{N_2}\sum x_i\right)\right\}^2$$

$$= \left[w^T \cdot (\bar{x}_1-\bar{x}_2)\right]^2 = w^T(\bar{x}_1-\bar{x}_2)(\bar{x}_1-\bar{x}_2)^T \cdot w$$

$$s_1+S_2 = \frac{1}{N_1}\sum\left(w^T x_i - \frac{1}{N_1}\sum w^T x_i\right)\left(w^T x_i - \frac{1}{N_1}\sum w^T x_i\right) + \cdots$$

$$= \frac{1}{N_1}\sum w^T (x_i-\bar{x}_1)(x_i-\bar{x}_1)^T w \quad + \cdots$$

$$= w^T \cdot \left[\frac{1}{N_1}\sum_{i=1}^{N_1}(x_i-\bar{x}_1)(x_i-\bar{x}_1)^T\right] \cdot w \quad + \cdots$$

$$= w^T \cdot S_1 \cdot w + w^T \cdot S_2 \cdot w$$

$$= w^T (S_1 + S_2) \cdot w$$

$$J(w) = \frac{w^T (\bar{x}_1 - \bar{x}_2)(\bar{x}_1 - \bar{x}_2) w}{w^T (S_1 + S_2) w}$$

$$= \frac{w^T \cdot S_b \cdot w}{w^T \cdot S_w \cdot w}$$

$S_b$: between-class 类间方差

$S_w$: within-class 类内方差

$$\frac{\partial J(w)}{\partial w} = 2 S_b \cdot w (w^T S_w \cdot w)^{-1} - w^T S_b \cdot w \cdot (w^T S_w \cdot w)^{-2} \cdot 2 S_w \cdot w$$

$$= 0$$

$$\Rightarrow S_b \cdot w \cdot (w^T \cdot S_w \cdot w)^{-1} - w^T S_b \cdot w \cdot (w^T S_w \cdot w)^{-2} \cdot S_w \cdot w$$

$$\Rightarrow S_b \cdot w \cdot (w^T \cdot S_w \cdot w) = w^T \cdot S_b \cdot w \cdot S_w \cdot w$$

$$\Rightarrow \underbrace{w^T \cdot S_b \cdot w}_{} \cdot S_w \cdot w = S_b \cdot w \cdot \underbrace{w^T S_w \cdot w}_{}$$

实数

$$\Rightarrow S_w \cdot w = \boxed{\frac{w^T \cdot S_w \cdot w}{w^T \cdot S_b \cdot w}} \cdot S_b \cdot w$$

$$\Rightarrow w \propto S_w^{-1} \cdot \underbrace{S_b \cdot w}_{(\bar{x}_1 - \bar{x}_2) \underbrace{(\bar{x}_1 - \bar{x}_2)^T \cdot w}_{1 \times 1}}$$

$$\propto \boxed{S_w^{-1} \cdot (\bar{x}_1 - \bar{x}_2)}$$

$S_w$ 对角，$S_w^{-1} \propto I$

$w \propto (\bar{x}_1 - \bar{x}_2)$