

② 線形判別分析 (LDA)

- m : 全体の平均ベクトル
- m_i : クラス i の平均ベクトル
- w : 新たな軸のベクトル
- n_i : クラス i のデータ数

理論① 別クラスのデータをできるだけ遠ざける

→ 各クラスの平均が、全体の平均から離れていければいい。

$$\begin{aligned}
 J_B(w) &= \sum_{i=1}^C n_i \{(m_i - m)^T w\}^2 \\
 &= \sum_{i=1}^C n_i \cdot w^T (m_i - m)^T (m_i - m) w \\
 &= w^T \left\{ \sum_{i=1}^C n_i (m_i - m)^T (m_i - m) \right\} w \\
 &= w^T S_B w \Rightarrow \boxed{\text{最大化}}
 \end{aligned}$$

↪ クラス間共分散行列

理論② 同じクラスのデータをできるだけ近づける

→ 射影後のクラス内の分散を小さくする。

$$\begin{aligned}
 J_i(w) &= \sum_{x^{(i)} \in C_i} \{(x^{(i)} - m_i)^T w\}^2 \\
 \text{とすると全体は} \\
 J_w(w) &= \sum_{i=1}^C J_i(w) = \sum_{i=1}^C \sum_{x^{(i)} \in C_i} \{(x^{(i)} - m_i)^T w\}^2
 \end{aligned}$$

$$\begin{aligned}
 J_w(w) &= \sum_{i=1}^C \sum_{x^{(i)} \in C_i} \{(x^{(i)} - m_i)^T w\}^2 \\
 &= \sum_{i=1}^C \sum_{x^{(i)} \in C_i} w^T (x^{(i)} - m_i)^T (x^{(i)} - m_i) w \\
 &= w^T \left\{ \sum_{i=1}^C \sum_{x^{(i)} \in C_i} (x^{(i)} - m_i)^T (x^{(i)} - m_i) \right\} w \\
 &= w^T S_w w \Rightarrow \boxed{\text{最小化}}
 \end{aligned}$$

↪ クラス内共分散行列

これをまとめると

$$J(w) = \frac{J_B(w)}{J_w(w)} = \frac{w^T S_B w}{w^T S_w w} \Rightarrow \boxed{\text{最大化}}$$

↪ 最大化

そこで、 $J(w)$ を最大化させる w を見つけたい。

→ $J(w)$ を w で微分する。

$$\frac{\partial J(w)}{\partial w} = \frac{\partial}{\partial w} \cdot \frac{w^T S_B w}{w^T S_w w}$$

性質①

$$\frac{\partial}{\partial x} x^T A x = (A + A^T)x$$

$$= \frac{1}{(w^T S_w w)^2} \{ (w^T S_w w) (S_B + S_B^T) w - (w^T S_B w) (S_w + S_w^T) w \}$$

↪ S_B, S_w は対称行列

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

$$= \frac{2}{(w^T S_w w)^2} \{ (w^T S_w w) S_B w - (w^T S_B w) S_w w \}$$

= ①

したがって、

$$(w^T S_A w) S_B w = (w^T S_B w) S_A w$$

\Leftrightarrow

$$S_B w = \frac{w^T S_B w}{w^T S_A w} S_A w$$

\Leftrightarrow

$$S_A^{-1} S_B w = \frac{w^T S_B w}{w^T S_A w} w$$

$\frac{w^T S_B w}{w^T S_A w}$ はスカラー, $S_A^{-1} S_B$ は行列なので、

$\frac{w^T S_B w}{w^T S_A w}$ は $S_A^{-1} S_B$ の固有値であり, w は $S_A^{-1} S_B$ の固有ベクトル
