宿題 1

標本 $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^n$ を考える。$\boldsymbol{x}$ は中心化されているとする。つまり，

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i = \boldsymbol{0} \tag{1}$$

とする。このとき，散布行列は，

$$\boldsymbol{C} = \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}} \tag{2}$$

クラス内散布行列は，

$$\boldsymbol{S}^{(\mathrm{w})} = \sum_{y=1}^c \sum_{i:y_i=y} \left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)^{\mathrm{T}} \tag{3}$$

クラス間散布行列は，

$$\boldsymbol{S}^{(\mathrm{b})} = \sum_{y=1}^c n_y \boldsymbol{\mu}_y \boldsymbol{\mu}_y^{\mathrm{T}} \tag{4}$$

である。ここで，n_y はクラス y の標本数であり，

$$\boldsymbol{\mu}_y = \frac{1}{n_y}\sum_{i:y_i=y} \boldsymbol{x}_i \tag{5}$$

である。

いま，散布行列について，

$$\boldsymbol{C} = \sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}} = \frac{1}{2}\left(\sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}} + \sum_{i'=1}^n \boldsymbol{x}_{i'} \boldsymbol{x}_{i'}^{\mathrm{T}}\right)$$

と変形できる。

$$\sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}} = \sum_{i'=1}^n \frac{1}{n}\left(\sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}}\right) = \frac{1}{n}\sum_{i,i'=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}}$$

であることから，

$$\boldsymbol{C} = \frac{1}{2}\cdot\frac{1}{n}\sum_{i,i'}^n \left(\boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}} + \boldsymbol{x}_{i'} \boldsymbol{x}_{i'}^{\mathrm{T}}\right)$$

となる。ここで，式 (1) から，

$$\begin{aligned}
\sum_{i,i'=1}^n \boldsymbol{x}_i \boldsymbol{x}_{i'}^{\mathrm{T}} &= \sum_{i=1}^n \boldsymbol{x}_i \left(\sum_{i'=1}^n \boldsymbol{x}_{i'}\right)^{\mathrm{T}} + \sum_{i'=1}^n \left(\sum_{i=1}^n \boldsymbol{x}_i\right) \boldsymbol{x}_{i'}^{\mathrm{T}} \\
&= 2\sum_{i=1}^n \boldsymbol{x}_i \sum_{i'=1}^n \boldsymbol{x}_{i'}^{\mathrm{T}} \\
&= 2n\boldsymbol{\mu}\cdot n\boldsymbol{\mu}^{\mathrm{T}} \\
&= \boldsymbol{0}
\end{aligned}$$

が成立することを用いると，結局，

$$
\begin{aligned}
\boldsymbol{C} &= \frac{1}{2}\cdot\frac{1}{n}\sum_{i,i'=1}^{n}\left(\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - 2\boldsymbol{x}_i\boldsymbol{x}_{i'}^{\mathrm{T}} + \boldsymbol{x}_{i'}\boldsymbol{x}_{i'}^{\mathrm{T}}\right) \\
&= \frac{1}{2}\sum_{i,i'=1}^{n}\frac{1}{n}(\boldsymbol{x}_i - \boldsymbol{x}_{i'})(\boldsymbol{x}_i - \boldsymbol{x}_{i'})^{\mathrm{T}}
\end{aligned} \tag{6}
$$

得る。

同様に，クラス内散布行列は，

$$
\begin{aligned}
\boldsymbol{S}^{(\mathrm{w})} &= \sum_{y=1}^{c}\sum_{i:y_i=y}\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)^{\mathrm{T}} \\
&= \frac{1}{2}\sum_{y=1}^{c}\frac{1}{n_y}\sum_{i,i':y_{i,i'}=y}\left\{\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)^{\mathrm{T}} + \left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)^{\mathrm{T}}\right\}
\end{aligned}
$$

と変形でき，ここで，式 (5) から，

$$
\begin{aligned}
\sum_{i,i':y_{i,i'}=y}\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)^{\mathrm{T}} &= 2\sum_{i:y_i=y}\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\cdot\sum_{i':y'_i=y}\left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)^{\mathrm{T}} \\
&= 2\left(n_y\boldsymbol{\mu}_y - n_y\boldsymbol{\mu}_y\right)\left(n_y\boldsymbol{\mu}_y - n_y\boldsymbol{\mu}_y\right)^{\mathrm{T}} \\
&= \boldsymbol{0}
\end{aligned}
$$

となることから，結局，

$$
\begin{aligned}
\boldsymbol{S}^{(\mathrm{w})} &= \frac{1}{2}\sum_{y=1}^{c}\frac{1}{n_y}\sum_{i,i':y_{i,i'}=y}\left\{\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)^{\mathrm{T}} - 2\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)^{\mathrm{T}} + \left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)^{\mathrm{T}}\right\} \\
&= \frac{1}{2}\sum_{y=1}^{c}\frac{1}{n_y}\sum_{i,i':y_{i,i'}=y}\left\{\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right) - \left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)\right\}\left\{\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right) - \left(\boldsymbol{x}_{i'} - \boldsymbol{\mu}_y\right)\right\}^{\mathrm{T}} \\
&= \frac{1}{2}\sum_{y=1}^{c}\frac{1}{n_y}\sum_{i,i':y_{i,i'}=y}(\boldsymbol{x}_i - \boldsymbol{x}_{i'})(\boldsymbol{x}_i - \boldsymbol{x}_{i'})^{\mathrm{T}}
\end{aligned}
$$

となる。これは，

$$
Q_{i,i'}^{(\mathrm{w})} = \begin{cases} 1/n_y & (y_i = y_{i'} = y) \\ 0 & (y_i \neq y_{i'}) \end{cases} \tag{7}
$$

なる $Q_{i,i'}^{(\mathrm{w})}$ を用いて，

$$
\boldsymbol{S}^{(\mathrm{w})} = \frac{1}{2}\sum_{i,i'=1}^{n} Q_{i,i'}^{(\mathrm{w})}(\boldsymbol{x}_i - \boldsymbol{x}_{i'})(\boldsymbol{x}_i - \boldsymbol{x}_{i'})^{\mathrm{T}} \tag{8}
$$

と表せる。

いま，$\boldsymbol{S}^{(\mathrm{w})}$ について変形すると，

$$
\begin{aligned}
\boldsymbol{S}^{(\mathrm{w})} &= \sum_{y=1}^{c} \sum_{i:y_i=y} \left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)\left(\boldsymbol{x}_i - \boldsymbol{\mu}_y\right)^{\mathrm{T}} \\
&= \sum_{y=1}^{c} \sum_{i:y_i=y} \left(\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - 2\boldsymbol{\mu}_y\boldsymbol{x}_i^{\mathrm{T}} + \boldsymbol{\mu}_y\boldsymbol{\mu}_y^{\mathrm{T}}\right) \\
&= \sum_{i=1}^{n} \boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - \sum_{y=1}^{c} 2\boldsymbol{\mu}_y \sum_{i:y_i=y} \boldsymbol{x}_i^{\mathrm{T}} + \sum_{y=1}^{c} n_y\boldsymbol{\mu}_y\boldsymbol{\mu}_y^{\mathrm{T}} \\
&= \sum_{i=1}^{n} \boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - \sum_{y=1}^{c} 2\boldsymbol{\mu}_y n_y\boldsymbol{\mu}_y^{\mathrm{T}} + \sum_{y=1}^{c} n_y\boldsymbol{\mu}_y\boldsymbol{\mu}_y^{\mathrm{T}} \\
&= \sum_{i=1}^{n} \boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - \sum_{y=1}^{c} n_y\boldsymbol{\mu}_y\boldsymbol{\mu}_y^{\mathrm{T}}
\end{aligned}
$$

となり，式 (2)，(4) とから，

$$
\boldsymbol{C} = \boldsymbol{S}^{(\mathrm{w})} + \boldsymbol{S}^{(\mathrm{b})} \tag{9}
$$

が成立する。

式 (6)，(8)，(9) から，

$$
\begin{aligned}
\boldsymbol{S}^{(\mathrm{b})} &= \boldsymbol{C} - \boldsymbol{S}^{(\mathrm{w})} \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} \left(\frac{1}{n} - Q_{i,i'}^{(\mathrm{w})}\right)(\boldsymbol{x}_i - \boldsymbol{x}_{i'})(\boldsymbol{x}_i - \boldsymbol{x}_{i'})^{\mathrm{T}} \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} Q_{i,i'}^{(\mathrm{b})}(\boldsymbol{x}_i - \boldsymbol{x}_{i'})(\boldsymbol{x}_i - \boldsymbol{x}_{i'})^{\mathrm{T}}
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
Q_{i,i'}^{(\mathrm{b})} &= \frac{1}{n} - Q_{i,i'}^{(\mathrm{w})} \\
&= \begin{cases} 1/n - 1/n_y & (y_i = y_{i'} = y) \\ 1/n & (y_i \neq y_{i'}) \end{cases}
\end{aligned} \tag{11}
$$

となる。