Advanced Data Analysis
Homework Week - 12

Aswin Vijay

Question 2

We are asked to derive the following pairwise expressions for Local Fischer Discriminant Analysis, given the number of samples as n and number of samples of class l being n_l. The within-class scatter matrix S^w and the between-class scatter matrix S^b are,

$$\begin{aligned} S^w &= \frac{1}{2}\sum_{i,i'=1}^{n} Q^w_{i,i'}(x_i - x_{i'})(x_i - x_{i'})^T \\ S^b &= \frac{1}{2}\sum_{i,i'=1}^{n} Q^b_{i,i'}(x_i - x_{i'})(x_i - x_{i'})^T \end{aligned} \tag{1}$$

where,

$$\begin{aligned} Q^w_{i,i'} &= \begin{cases} \frac{1}{n_l} > 0 & (y_i = y_{i'} = l) \\ 0 & (y_i \neq y_{i'}) \end{cases} \\ Q^b_{i,i'} &= \begin{cases} \frac{1}{n} - \frac{1}{n_l} > 0 & (y_i = y_{i'} = l) \\ \frac{1}{n} & (y_i \neq y_{i'}) \end{cases} \end{aligned} \tag{2}$$

Derivation

From Fischer Discriminant Analysis we have the following scatter matrices given total number of classes c,

$$\begin{aligned} S^w &= \sum_{l=1}^{c}\sum_{i:y_i=l}(x_i - \mu_l)(x_i - \mu_l)^T \\ S^b &= \sum_{i=1}^{c} n_l(\mu_l - \mu)(\mu_l - \mu)^T \end{aligned} \tag{3}$$

where μ_l is the smaple mean of class l,

$$\begin{aligned} \mu_l &= \frac{1}{n_l}\sum_{i:y_i=l} x_i \\ \mu &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{l=1}^{c} n_l\mu_l \end{aligned} \tag{4}$$

Starting with S^w in Eq.3 and substituting the class sample mean μ_l we get,

$$
\begin{aligned}
S^w &= \sum_{l=1}^{c} \sum_{i:y_i=l} (x_i - \mu_l)(x_i - \mu_l)^T \\
&= \sum_{l=1}^{c} \sum_{i:y_i=l} \left(x_i - \frac{1}{n_l} \sum_{i':y_{i'}=l} x_{i'} \right) \left(x_i - \frac{1}{n_l} \sum_{i':y_{i'}=l} x_{i'} \right)^T \\
&= \sum_{i=1}^{n} x_i x_i^T - \sum_{l=1}^{c} \frac{1}{n_l} \sum_{i,i':y_i=y_{i'}=l} x_i x_{i'}^T \\
&= \sum_{i=1}^{n} \left(\sum_{i'=1}^{n} Q^w_{i,i'} \right) x_i x_i^T - \sum_{i,i'=1}^{n} Q^w_{i,i'} x_i x_{i'}^T \quad \text{Using Eq. 2} \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} Q^w_{i,i'} (x_i x_i^T - x_i x_{i'}^T - x_{i'}^T x_i + x_{i'}^T x_{i'}^T) \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} Q^w_{i,i'} (x_i - x_{i'})(x_i - x_{i'})^T
\end{aligned}
\tag{5}
$$

Using mixture scatter matrix formula we have,

$$
\begin{aligned}
S^m &= S^w + S^b \\
&= \sum_{i=1}^{n} (x_i - \mu)(x_{i'} - \mu)
\end{aligned}
\tag{6}
$$

Using 6 we derive S^b as,

$$
\begin{aligned}
S_b &= \sum_{i=l}^{n} x_i x_i^T - \frac{1}{n} \sum_{i,i'=1}^{n} x_i x_{i'}^T - S^w \\
&= \sum_{i=1}^{n} \left(\sum_{i'=1}^{n} \frac{1}{n} \right) x_i x_i^T - \sum_{i,i'=1}^{n} \frac{1}{n} x_i x_{i'}^T - S^w \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} \left(\frac{1}{n} - Q^w_{i,i'} \right) (x_i - x_{i'})(x_i - x_{i'})^T \\
&= \frac{1}{2} \sum_{i,i'=1}^{n} Q^b_{i,i'} (x_i - x_{i'})(x_i - x_{i'})^T
\end{aligned}
\tag{7}
$$

Thus we derive both the required scatter matrices for LFDA.

References

- *Masashi Sugiyama. 2007. Dimensionality Reduction of Multimodal Labeled Data by Local Fisher Discriminant Analysis. J. Mach. Learn. Res. 8 (5/1/2007), 1027–1061.*