Machine Learning Homework 02

Qiu Yihang

Apr.30, 2022

We change the classification of $\hat{\mathbf{x}}$ from $\hat{\mathbf{x}} \in H_i$ to $\hat{\mathbf{x}} \in H_j$.

Notations: $\mathbf{m}_k \triangleq \sum_{\mathbf{x} \in H_k} \mathbf{x}$. n_k is the number of data samples in H_k before the change.

$J_k = \sum_{\mathbf{x} \in H_k} \|\mathbf{x} - \mathbf{m}_i\|^2$. J_k^* is the J_k after the change.

Show that $J_i^* = J_i - \frac{n_i}{n_i - 1} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2$.

Proof. Use α^* to denote the α after the change.

We have

$$\mathbf{m}_i^* = \frac{n_i \mathbf{m}_i - \hat{\mathbf{x}}}{n_i - 1} = \mathbf{m}_i + \frac{n_i \mathbf{m}_i - \hat{\mathbf{x}}}{n_i - 1} - \mathbf{m}_i = \mathbf{m}_i - \frac{\hat{\mathbf{x}} - \mathbf{m}_i}{n_i - 1}.$$

Thus,

$$\begin{aligned}
J_i^* &= \sum_{\mathbf{x} \in H_i^*} \|\mathbf{x} - \mathbf{m}_i^*\|^2 = \sum_{\mathbf{x} \in H_i} \|\mathbf{x} - \mathbf{m}_i^*\|^2 - \|\hat{\mathbf{x}} - \mathbf{m}_i^*\|^2 \\
&= \sum_{\mathbf{x} \in H_i} \left\| \mathbf{x} - \mathbf{m}_i + \frac{\hat{\mathbf{x}} - \mathbf{m}_i}{n_i - 1} \right\|^2 - \left\| \frac{n_i}{n_i - 1} (\hat{\mathbf{x}} - \mathbf{m}_i) \right\|^2 \\
&= \sum_{\mathbf{x} \in H_i} \|\mathbf{x} - \mathbf{m}_i\|^2 + \frac{2}{n_i - 1} \sum_{\mathbf{x} \in H_i} (\hat{\mathbf{x}} - \mathbf{m}_i)^T (\mathbf{x} - \mathbf{m}_i) + \frac{1}{(n_i - 1)^2} \sum_{\mathbf{x} \in H_i} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2 \\
&\quad - \frac{n_i^2}{(n_i - 1)^2} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2 \\
&= J_i + \frac{2}{n_i - 1} (\hat{\mathbf{x}} - \mathbf{m}_i)^T \sum_{\mathbf{x} \in H_i} (\mathbf{x} - \mathbf{m}_i) + \frac{n_i(1 - n_i)}{(n_i - 1)^2} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2 \\
&= J_i + 0 - \frac{n_i}{n_i - 1} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2 \\
&= J_i - \frac{n_i}{n_i - 1} \|\hat{\mathbf{x}} - \mathbf{m}_i\|^2
\end{aligned}$$

Qed. ■