CS3390: Foundations of Machine Learning **2023**

Lecture 6: 8 September 2023

Instructor: P. K. Srijith *Scribe: Gautam Singh*

6.1 Linear Discriminant Analysis

In the following, r_t is an indicator variable to select elements belonging to a certain class.

1. Between class scatter is given by

$$(m_1 - m_2)^2 = \left(\mathbf{w}^\top \mathbf{m_1} - \mathbf{w}^\top \mathbf{m_2}\right)^2 \tag{6.1}$$
$$= \mathbf{w}^\top (\mathbf{m_1} - \mathbf{m_2}) (\mathbf{m_1} - \mathbf{m_2})^\top \mathbf{w} \tag{6.2}$$
$$= \mathbf{w}^\top \mathbf{S_B} \mathbf{w} \tag{6.3}$$

   where we define

$$\mathbf{S_B} \triangleq (\mathbf{m_1} - \mathbf{m_2}) (\mathbf{m_1} - \mathbf{m_2})^\top . \tag{6.4}$$

2. For any class, within class scatter is given by

$$s_i^2 = \sum_t \left(\mathbf{w}^\top \mathbf{x_t} - m_1\right)^2 r_t \tag{6.5}$$
$$= \sum_t \mathbf{w}^\top (\mathbf{x_t} - \mathbf{m_1}) (\mathbf{x_t} - \mathbf{m_1})^\top \mathbf{w} r_t \tag{6.6}$$
$$= \mathbf{w}^\top \mathbf{S_i} \mathbf{w} \tag{6.7}$$

   where we define

$$S_i \triangleq \sum_t (\mathbf{x_t} - \mathbf{m_1}) (\mathbf{x_t} - \mathbf{m_1})^\top r_t. \tag{6.8}$$

3. For multiple classes, we define

$$\mathbf{S_w} \triangleq \sum_i \mathbf{S_i}. \tag{6.9}$$