Machine Learning

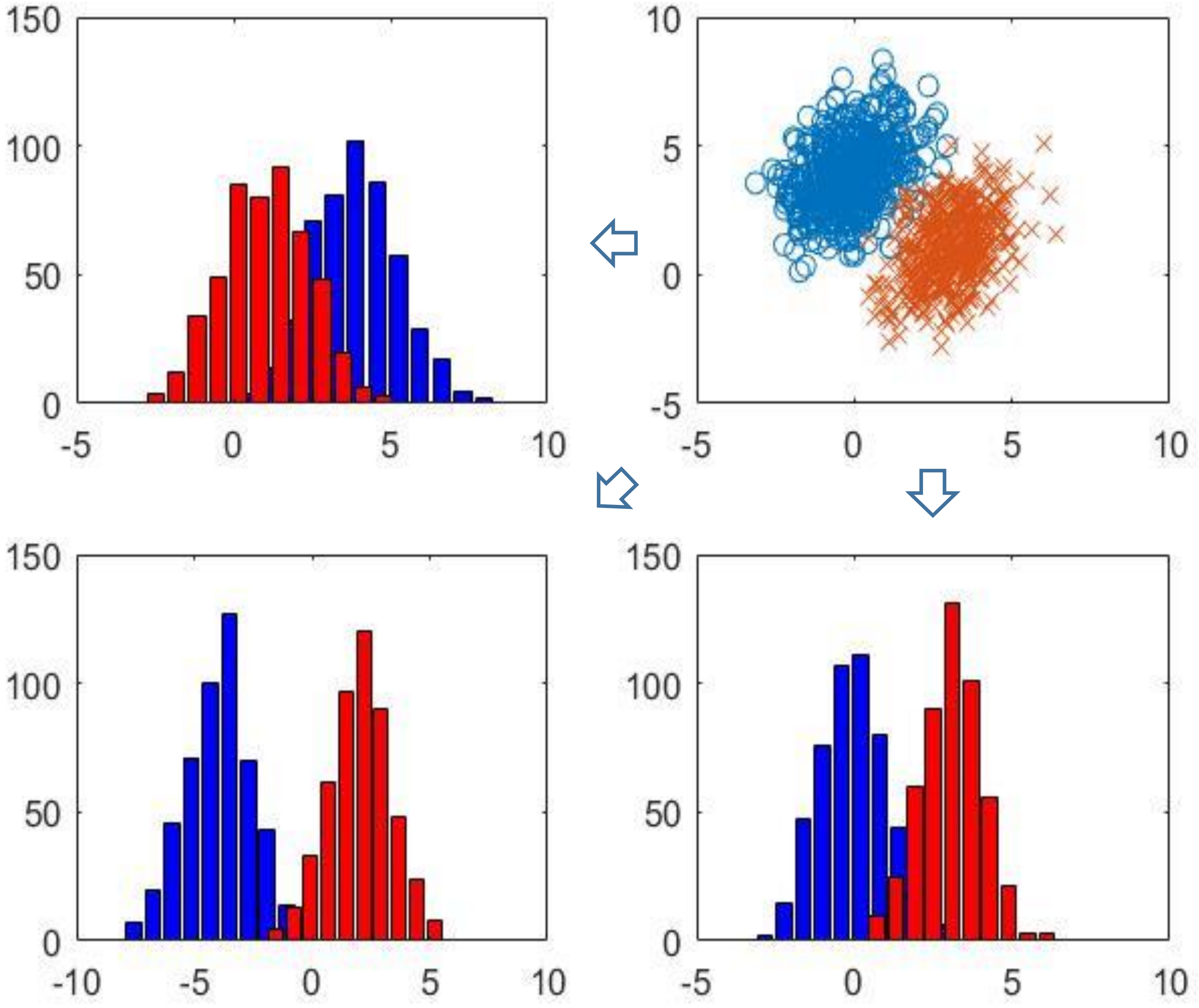
Contents

3.1. Linear Discriminant Analysis (LDA)

- Problem: Given a set of data points with class labels, find the best set of basis vectors for projecting the data points such that classification is improved.

- Idea: Form the projection (maximizing the variability across the different classes, while minimizing the variability within each class)

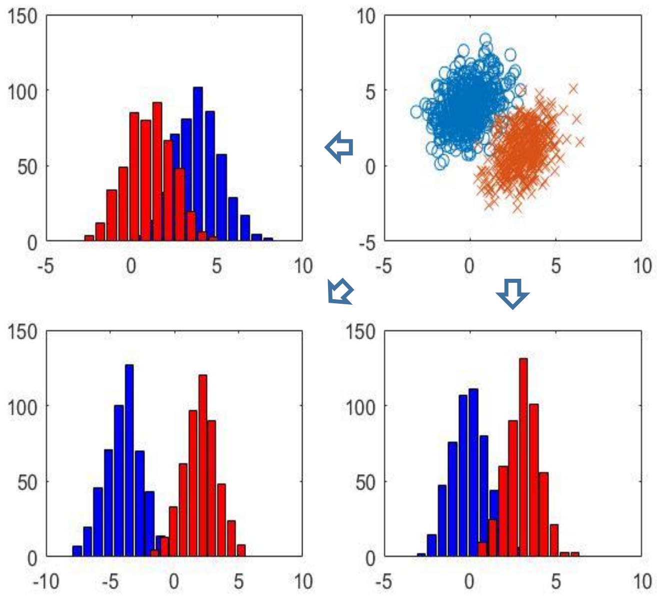
Which one is the best projection?



LDA: Maximize Difference of Means

- Considering the simple problem of projecting onto one dimension
- The two classes should be well separated in this single dimension
- A simple idea is to maximize the difference of the means in the projected space
- What is a problem with this solution?

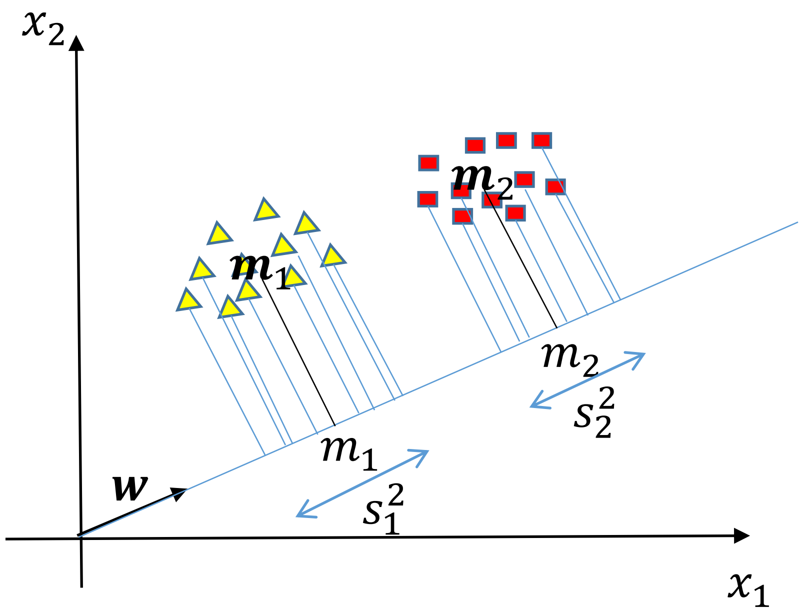
...Overlap of classes after the projection



LDA: A Better Idea

- Fisher's idea: Fisher's linear discriminant

1) 투영 후 각 클래스별 평균 간의 거리는 최대가 되어야 한다. 2) 투영 후 각 클래스 내부의 분산은 최소가 되어, 클래스 간 겹침이 최소가 되도록 한다.



3.2. Projection with Basis Vectors

- Point A = $(a_1, a_2)^T$
- Basis Vector **w** = $(w_1, w_2)^T$
- Projection $\boldsymbol{w}^T A = w_1\ a_1 + w_2\ a_2$

참조:

T(transpose)는 행렬 혹은 벡터에서 행과 열의 위치를 교환하는 이항 연산자이다. 즉,

$$(a_1, a_2)^T = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{3.2.4}$$

이고

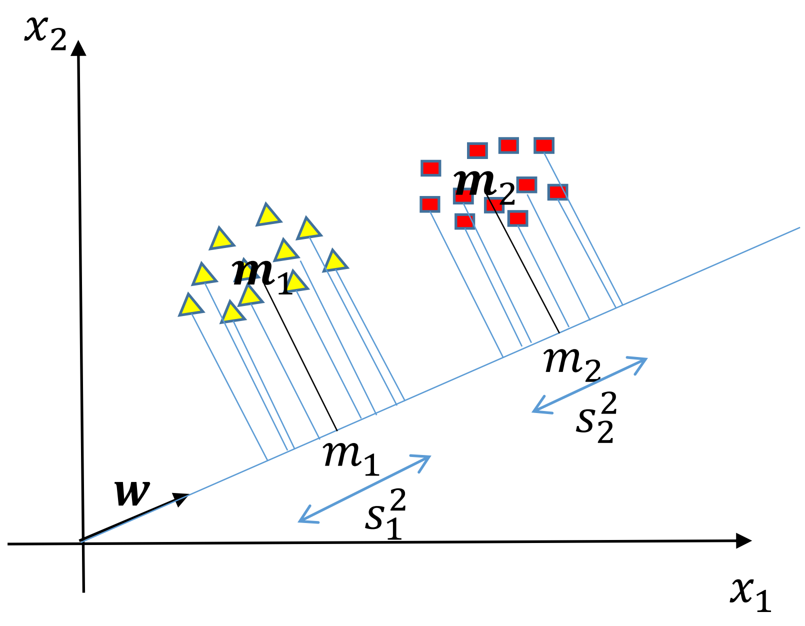
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \tag{3.2.5}$$

이다. 또한 행렬 A와 B의 곱에 대하여 이항 연산자 T를 적용하면

$$(AB)^T = B^T A^T \tag{3.2.6}$$

이 된다.

3.3. LDA Formulation



Maximizing

$$J(\boldsymbol{w}) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2} \quad (3.3.1)$$

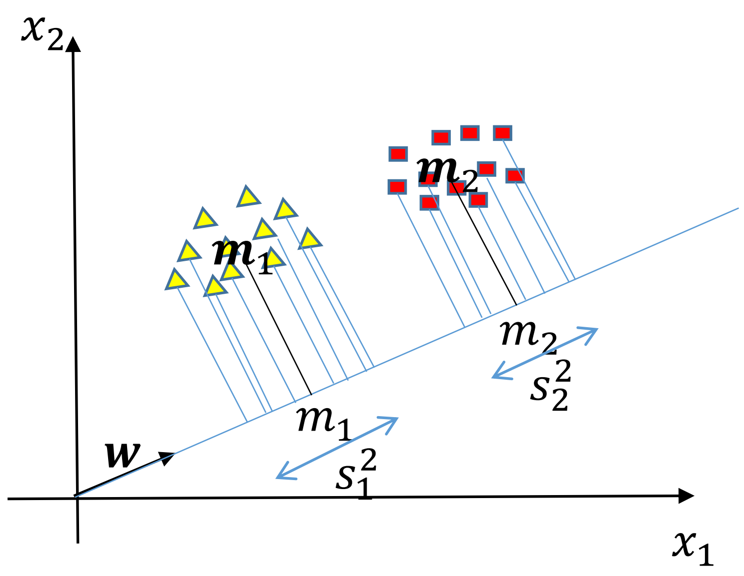
$$m_1 = \frac{\sum_{i \in C_1} \boldsymbol{w}^T \boldsymbol{x}^i}{N_1} = \frac{\sum_t \boldsymbol{w}^T \boldsymbol{x}^t r^t}{\sum_t r^t} = \boldsymbol{w}^T \frac{\sum_t \boldsymbol{x}^t r^t}{\sum_t r^t} = \boldsymbol{w}^T \boldsymbol{m}_1 \quad (3.3.2)$$

$$s_1^2 = \sum_t (\boldsymbol{w}^T \boldsymbol{x}^t - m_1)^2 r^t \quad (3.3.3)$$

Between-Class Scatter (numerator)

$$\begin{aligned}(m_1 - m_2)^2 &= (\boldsymbol{w}^T \boldsymbol{m}_1 - \boldsymbol{w}^T \boldsymbol{m}_2)^2 = (\boldsymbol{w}^T (\boldsymbol{m}_1 - \boldsymbol{m}_2))^2 \quad (3.3.4)\\ &= \boldsymbol{w}^T (\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \boldsymbol{w} = \boldsymbol{w}^T S_B \boldsymbol{w}\end{aligned}$$

$$S_B = (\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \quad (3.3.5)$$



$$J(\boldsymbol{w}) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2} \tag{3.3.1}$$

Within-Class Scatter (denominator)

$$s_1^2 = \sum_t (\boldsymbol{w}^T \boldsymbol{x}^t - m_1)^2 r^t = \sum_t (\boldsymbol{w}^T \boldsymbol{x}^t - \boldsymbol{w}^T \boldsymbol{m}_1)^2 r^t = \sum_t (\boldsymbol{w}^T (\boldsymbol{x}^t - \boldsymbol{m}_1))^2 r^t \tag{3.3.6}$$
$$= \sum_t \boldsymbol{w}^T (\boldsymbol{x}^t - \boldsymbol{m}_1)(\boldsymbol{x}^t - \boldsymbol{m}_1)^T \boldsymbol{w} r^t = \boldsymbol{w}^T S_1 \boldsymbol{w}$$

$$S_1 = \sum_t (\boldsymbol{x}^t - \boldsymbol{m}_1)(\boldsymbol{x}^t - \boldsymbol{m}_1)^T r^t \tag{3.3.7}$$

$$s_1^2 + s_2^2 = \boldsymbol{w}^T S_W \boldsymbol{w} \tag{3.3.8}$$

$$S_W = S_1 + S_2 \tag{3.3.9}$$

$$J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T S_B \boldsymbol{w}}{\boldsymbol{w}^T S_W \boldsymbol{w}} \tag{3.3.10}$$

참조: 벡터의 미분

d차원 벡터 $\boldsymbol{w}=(w_1, w_2, \ldots, w_d)^T$에 대한 미분 가능한 스칼라 함수 g가

$$g(w_1, w_2, \ldots, w_d) = g(\boldsymbol{w}) \tag{3.3.16}$$

로 주어졌다. 그러면, 함수 g의 벡터 $\boldsymbol{w}$에 대한 미분은

$$\frac{\partial g}{\partial \boldsymbol{w}} = \begin{pmatrix} \frac{\partial g}{\partial w_1} \\ \vdots \\ \frac{\partial g}{\partial w_d} \end{pmatrix} \tag{3.3.17}$$

이다. 또한 2차 미분은

$$\frac{\partial^2 g}{\partial \boldsymbol{w}^2} = \begin{pmatrix} \frac{\partial^2 g}{\partial w_1^2} & \cdots & \frac{\partial g}{\partial w_1 \partial w_d} \\ \vdots & & \vdots \\ \frac{\partial g}{\partial w_d \partial w_1} & \cdots & \frac{\partial^2 g}{\partial w_d^2} \end{pmatrix} \tag{3.3.18}$$

이 된다. 벡터 값을 지닌 함수 g가

이 된다. 벡터 값을 지닌 함수 g가

$$\boldsymbol{g}(\boldsymbol{w}) = \begin{pmatrix} \boldsymbol{g}_1(\boldsymbol{w}) \\ \vdots \\ \boldsymbol{g}_n(\boldsymbol{w}) \end{pmatrix} \tag{3.3.19}$$

로 주어지면 g의 $\boldsymbol{w}$에 대한 Jacobian 행렬은

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{w}} = \begin{pmatrix} \frac{\partial g_1}{\partial w_1} & \cdots & \frac{\partial g_n}{\partial w_1} \\ \vdots & & \vdots \\ \frac{\partial g_1}{\partial w_d} & \cdots & \frac{\partial g_n}{\partial w_d} \end{pmatrix} \tag{3.3.20}$$

이다. 또한

$$\frac{\partial f(\boldsymbol{w})g(\boldsymbol{w})}{\partial \boldsymbol{w}} = \frac{\partial f(\boldsymbol{w})}{\partial \boldsymbol{w}} g(\boldsymbol{w}) + f(\boldsymbol{w}) \frac{\partial g(\boldsymbol{w})}{\partial \boldsymbol{w}} \tag{3.3.21}$$

$$\frac{\partial f(\boldsymbol{w})/g(\boldsymbol{w})}{\partial \boldsymbol{w}} = \left[\frac{\partial f(\boldsymbol{w})}{\partial \boldsymbol{w}} g(\boldsymbol{w}) - f(\boldsymbol{w}) \frac{\partial g(\boldsymbol{w})}{\partial \boldsymbol{w}} \right] / g^2(\boldsymbol{w}) \tag{3.3.22}$$

$$\frac{\partial f(g(\boldsymbol{w}))}{\partial \boldsymbol{w}} = f'(g(\boldsymbol{w})) \frac{\partial g(\boldsymbol{w})}{\partial \boldsymbol{w}} \tag{3.3.23}$$

이다.

예제 3.3-1

행렬 $\boldsymbol{A}=(a_{ij})$가 $d\times d$ 정방행렬이고, 다음 함수

$$g(\boldsymbol{w})=\boldsymbol{w}^T\boldsymbol{A}\boldsymbol{w}=\sum_{i=1}^{d}\sum_{j=1}^{d}w_iw_ja_{ij}$$

에 대하여 $\dfrac{\partial g}{\partial \boldsymbol{w}}$를 구하여라.

Maximizing

$$J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T S_B \boldsymbol{w}}{\boldsymbol{w}^T S_W \boldsymbol{w}} \tag{3.3.10}$$

From $\frac{\partial J(\boldsymbol{w})}{\partial \boldsymbol{w}} = 0$,

$$2 S_B \boldsymbol{w} (\boldsymbol{w}^T S_W \boldsymbol{w}) - (\boldsymbol{w}^T S_B \boldsymbol{w}) 2 S_W \boldsymbol{w} = 0 \tag{3.3.11}$$

$$S_B \boldsymbol{w} (\boldsymbol{w}^T S_W \boldsymbol{w}) = (\boldsymbol{w}^T S_B \boldsymbol{w}) S_W \boldsymbol{w} \tag{3.3.12}$$

$$S_B \boldsymbol{w} = \frac{(\boldsymbol{w}^T S_W \boldsymbol{w})}{(\boldsymbol{w}^T S_B \boldsymbol{w})} S_W \boldsymbol{w} = \lambda S_W \boldsymbol{w} \tag{3.3.13}$$

$$S_B = (\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \quad (3.3.5) \quad \Longrightarrow \quad S_B \boldsymbol{w} = (\boldsymbol{m}_1 - \boldsymbol{m}_2) \boxed{(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \boldsymbol{w}} \tag{3.3.14}$$

scalar

$$\boldsymbol{w} = S_W^{-1} (\boldsymbol{m}_1 - \boldsymbol{m}_2) \tag{3.3.15}$$

More than 2 Classes (K > 2)

- Projection from d-dim space to (c-1)-dim space:
  - $y_i = \mathbf{w}_i^T \mathbf{x}, \quad i = 1, \ldots, c-1$
  - $\mathbf{y} = W^T \mathbf{x}$ (W = d×(c-1) matrix, w_i is the i-th column vector)
- Within-class scatter matrix for C_i

  $S_i = \sum_t r_i^t (x^t - m_i)(x^t - m_i)^T$ where $r_i^t = 1$ if $x^t \in C_i$ 0 otherwise
- The total within-class scatter

  $S_w = \sum_{i=1}^K S_i$
- The between-scatter matrix

  $S_B = \sum_{i=1}^K N_i \, (m_i - m)(m_i - m)^T$ where $N_i = \sum_t r_i^t$
- Find the matrix W that maximize

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|}$$

- Solution: the largest eigenvectors of $S_W^{-1} S_B$