Fundamentals of machine Learning exercise 3

② a) minimize $C = \sum_{i=1}^{N} (w^T x_i + b - y_i)^2$ w.r.t. b

$$\Rightarrow \frac{\partial}{\partial b} \sum_{i=1}^{N} (w^T x_i + b - y_i)^2 \overset{!}{=} 0$$

$$= 2 \sum_{i=1}^{N} (w^T x_i + b - y_i) = 0 \quad \mid b \text{ indep. of } i$$

$$\Leftrightarrow N b + \sum_{i=1}^{N} (w^T x_i \underbrace{- y_i}_{=0}) = 0 \quad \mid \#(y_i = 1) = \#(y_i = -1) = \frac{N}{2}$$

$$\Rightarrow N b + \sum_{i=1}^{N} w^T x_i = 0$$

$$\Leftrightarrow \underline{\underline{b = - w^T \cdot \mu}} \quad , \text{with } \mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

b) minimize C w.r.t w and the constraint $b = -w^T \mu$

$$\Rightarrow \frac{\partial}{\partial w} \sum_{i=1}^{N} (w^T x_i + b - y_i)^2 = 0$$

$$\Leftrightarrow \sum_{i=1}^{N} x_i (w^T x_i + b - y_i) = 0 \quad \mid \text{use constraint}$$

$$\Rightarrow \sum_{i=1}^{N} x_i (w^T (x_i - \mu) - y_i) = 0$$

$$\Leftrightarrow \sum_{i=1}^{N} x_i w^T (x_i - \mu) = \sum_{i=1}^{N} x_i y_i \quad : ①$$

<u>RHS</u>

$$\sum_{i=1}^{N} x_i y_i = \sum_{i: y = 1} x_i - \sum_{i: y = -1} x_i$$

$$= \underline{\underline{\frac{N}{2} (\mu_1 - \mu_2)}} \quad \text{with } \mu_k = \frac{2}{N} \sum_{i: y = k} x_i , \quad \mu_2 := \mu_{-1}$$

<u>LHS</u>

$$\sum_{i=1}^{N} [x_i w^T x_i - x_i w^T \mu]$$

$$= \sum_{i=1}^{N} x_i w^T x_i - N \mu w^T \mu \quad \mid \mu = \frac{\mu_1 + \mu_2}{2}$$

$$= \sum_{i=1} x_i w^T x_i - \frac{N}{4} ((\mu_1 + \mu_2) w^T (\mu_1 + \mu_2))$$

$\rightsquigarrow$ 1: $\sum_{i=1}^{N} x_i (w^T x_i) - \frac{N}{4}(\mu_1 + \mu_2) w^T (\mu_1 + \mu_2) = \frac{N}{2}(\mu_1 + \mu_2)$

c) show 1 $\Leftrightarrow (S_w + \frac{1}{4} S_B) w = \frac{\mu_1 + \mu_2}{2}$

RHS

$C \frac{\mu_1 + \mu_2}{2} \overset{!}{=} \frac{N}{2}(\mu_1 + \mu_2) \Rightarrow C = N$

$\Rightarrow LHS_{1} \overset{!}{=} N(S_w + \frac{1}{4} S_B w)$

$\rightsquigarrow$ calculate $N(S_w + \frac{1}{4} S_B w)$ and compare to LHS of 1

$S_B w$

$$S_B w := (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T w$$
$$= \underline{\underline{\mu_1 \mu_1^T w + \mu_2 \mu_2^T w - \mu_1 \mu_2^T w - \mu_2 \mu_1^T w}} =: 2$$

$S_w w$

$$S_w w := \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_{y_i})(x_i - \mu_{y_i})^T w \quad \Big| \; \mu_{y_i} = \begin{cases} \mu_1 & \text{for } y_i = 1 \\ \mu_2 & \text{for } y_i = -1 \end{cases}$$
$$= \frac{1}{N}\Big(\underbrace{\sum_{i:y_i=1} (x_i - \mu_1)(x_i - \mu_1)^T w}_{=: 4} + \sum_{i:y_i=-1} (x_i - \mu_2)(x_i - \mu_2)^T w \Big) =: 3$$

$$\rightsquigarrow 4: \quad = \sum_{i:y_i=1} \Big[\underbrace{(x_i - \mu_1) x_i^T w}_{=: 5} - \underbrace{(x_i - \mu_1) \mu_1^T w}_{=: 6} \Big]$$

$$\rightsquigarrow 5: \quad \sum_{i:y_i=1} [x_i x_i^T w - \mu_1 x_i^T w] \quad \Big| \sum_{i:y_i=1} x_i^T = \frac{N}{2} \mu_1^T$$
$$= \sum_{i:y_i=1} x_i x_i^T w - \frac{N}{2} \mu_1 \mu_1^T w$$

$$6: \quad \sum_{i:y_i=1} [x_i \mu_1^T w - \mu_1 \mu_1^T w] \quad | \text{ analogue}$$
$$= \frac{N}{2} \mu_1 \mu_1^T w - \frac{N}{2} \mu_1 \mu_1^T w = 0$$

$\Rightarrow 5 = 4 = \sum_{i:y_i=1} x_i x_i^T w - \frac{N}{2} \mu_1 \mu_1^T w$, analogue for $\sum_{i:y_i=-1}$ - part of 3

$$\Rightarrow S_w w = \frac{1}{N}\left(\sum_{i:y_i=1} x_i x_i^T w + \sum_{i:y_i=-1} x_i x_i w - \frac{N}{2}\left(\mu_1\mu_1^T + \mu_2\mu_2^T\right) w\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T w - \frac{1}{2}\left(\mu_1\mu_1^T + \mu_2\mu_2^T\right) w$$

Since we now calculated easyly computable forms of S_w and S_B we can simplify $N(S_w + \frac{1}{4} S_B) w$

$$N(S_w + \frac{1}{4} S_B) w = \sum_{i=1}^{N} x_i x_i^T w - \frac{N}{2}\left(\mu_1\mu_1^T + \mu_2\mu_2^T\right) w$$

$$+ \frac{1}{4}N(\mu_1\mu_1^T w) + \frac{1}{4}(\mu_2\mu_2^T w)\cdot N$$

$$- \frac{1}{4}N(\mu_1\mu_2^T w) - \frac{1}{4}N(\mu_2\mu_1^T w)$$

$$= \sum_{i=1}^{N} x_i x_i^T w - \frac{N}{4}\left(\mu_1\mu_1^T + \mu_2\mu_2^T + \mu_1\mu_2^T + \mu_2\mu_1^T\right) w$$

$$= \sum_{i=1}^{N} x_i x_i^T w - \frac{N}{4}(\mu_1+\mu_2)\, w^T(\mu_1+\mu_2)$$

which is the exact same expression as the LHS of of 1 ∎