

WS2020-IFML-exc03

j.kammerer

December 2020

3.1

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

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

$$= \sum_{i=1}^N \frac{\partial}{\partial b} (b^2 + 2(w^T x_i)b - 2by_i) \quad (3)$$

$$= 2Nb + \sum_{i=1}^N 2(w^T x_i) - \sum_{i=1}^N 2y_i \quad (4)$$

$$= 2Nb + \sum_{i=1}^N 2(w^T x_i) \quad (5)$$

$$\iff \hat{b} = -\frac{1}{N} \sum_{i=1}^N (w^T x_i) \quad (6)$$

$$= -w^T \frac{1}{N} \sum_{i=1}^N x_i \quad (7)$$

$$= -w^T \mu \quad (8)$$

3.2

$$0 = \frac{\partial}{\partial w} \sum_{i=1}^N (w^T x_i + \hat{b} - y_i)^2 \quad (9)$$

$$= \sum_{i=1}^N \frac{\partial}{\partial w} ((w^T x_i)^2 + 2(w^T x_i)\hat{b} - 2(w^T x_i)y_i) \quad (10)$$

$$= \sum_{i=1}^N (2(w^T x_i)x_i^T + 2x_i^T \hat{b} - 2x_i^T y_i) \quad (11)$$

$$\iff 0 = \sum_{i=1}^N x_i^T w^T (x_i - \mu) - \sum_{i=1}^N x_i^T y_i \quad (12)$$

$$\iff 0 = \sum_{i=1}^N x_i w^T (x_i - \mu) - \sum_{i=1}^N x_i y_i \quad (13)$$

$$\iff 0 = \sum_{i=1}^N x_i (x_i - \mu)^T w - \sum_{i=1}^N x_i y_i \quad (14)$$

$$\iff 0 = \left(\sum_{i=1}^N x_i x_i^T - N \mu \mu^T \right) \hat{w} - \sum_{i=1}^N x_i y_i \quad (15)$$

$$(16)$$

$$S_W = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_{y_i})(x_i - \mu_{y_i})^T \quad (17)$$

$$= \frac{1}{N} \sum_{i:y_i=1} (x_i - \mu_1)(x_i - \mu_1)^T + \frac{1}{N} \sum_{i:y_i=-1} (x_i - \mu_{-1})(x_i - \mu_{-1})^T \quad (18)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i x_i^T - \left(\frac{1}{N} \sum_{i:y_i=1} x_i \right) \mu_1^T - \mu_1 \left(\frac{1}{N} \sum_{i:y_i=1} x_i^T \right) + \frac{1}{N} \sum_{i:y_i=1} \mu_1 \mu_1^T \quad (19)$$

$$- \left(\frac{1}{N} \sum_{i:y_i=-1} x_i \right) \mu_{-1}^T - \mu_{-1} \left(\frac{1}{N} \sum_{i:y_i=-1} x_i^T \right) + \frac{1}{N} \sum_{i:y_i=-1} \mu_{-1} \mu_{-1}^T \quad (20)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i x_i^T - \frac{1}{2} \mu_1 \mu_1^T - \frac{1}{2} \mu_1 \mu_1^T + \frac{1}{2} \mu_1 \mu_1^T \quad (21)$$

$$- \frac{1}{2} \mu_{-1} \mu_{-1}^T - \frac{1}{2} \mu_{-1} \mu_{-1}^T + \frac{1}{2} \mu_{-1} \mu_{-1}^T \quad (22)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i x_i^T - \frac{1}{2} \mu_1 \mu_1^T - \frac{1}{2} \mu_{-1} \mu_{-1}^T \quad (23)$$

$$S_B = (\mu_1 - \mu_{-1})(\mu_1 - \mu_{-1})^T \quad (24)$$

$$= \mu_1 \mu_1^T - \mu_1 \mu_{-1}^T - \mu_{-1} \mu_1^T + \mu_{-1} \mu_{-1}^T \quad (25)$$

$$0 = (S_W + \frac{1}{4}S_B)\hat{w} - \frac{\mu_1 - \mu_{-1}}{2} \quad (26)$$

$$= (\frac{1}{N} \sum_{i=1}^N x_i x_i^T - \frac{1}{4}(\mu_1 \mu_1^T + \mu_1 \mu_{-1}^T + \mu_{-1} \mu_1^T + \mu_{-1} \mu_{-1}^T))\hat{w} \quad (27)$$

$$- \frac{2}{N}(\sum_{i:y_i=1} x_i y_i + \sum_{i:y_i=-1} x_i y_i) \quad (28)$$

$$= (\frac{1}{N} \sum_{i=1}^N x_i x_i^T - \frac{1}{4}(\mu_1 + \mu_{-1})(\mu_1 + \mu_{-1})^T)\hat{w} - \frac{1}{N} \sum_{i=1}^N x_i y_i \quad (29)$$

$$= (\frac{1}{N} \sum_{i=1}^N x_i x_i^T - \mu \mu^T)\hat{w} - \frac{1}{N} \sum_{i=1}^N x_i y_i \quad (30)$$

$$\iff 0 = (\sum_{i=1}^N x_i x_i^T - N \mu \mu^T)\hat{w} - \sum_{i=1}^N x_i y_i \quad (31)$$

$$QED. \quad (32)$$

$$(33)$$