Ex03

Mittwoch, 28. Oktober 2020 15:09

$\omega \in \mathbb{R}^D,\quad x_i \in \mathbb{R}^D,\ y_i \in \{\pm 1\},\ i \in \{1,\dots,N\},\quad N_1 = N_{-1} = N/2$

3.1 $\frac{\partial}{\partial b} \sum_{i=1}^N (\omega^T x_i + b - y_i)^2 = 0$

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

$\Rightarrow \sum_{i=1}^N b = -\sum_{i=1}^N \omega^T x_i - y_i$

$\sum_{i=1}^N y_i = \sum_{i:y_i=1} y_i + \sum_{i:y_i=-1} y_i = N_1 - N_{-1} = 0$

$\Rightarrow \hat{b} = \frac{1}{N} \sum_{i=1}^N y_i - \omega^T x_i = \underline{\underline{-\frac{1}{N} \sum \omega^T x_i}}$

3.2 $\partial_\omega (\omega^T x) = x,\quad a \cdot (b^T \cdot c) = (a \cdot c^T) \cdot b$

$\partial_\omega \sum_{i=1}^N (\omega^T x_i + \hat{b} - y_i)^2 = 0$

$\Rightarrow 0 = \sum_{i=1}^N 2 (\omega^T x_i + \hat{b} - y_i)\, x_i$

$\sum y_i x_i = \sum_{i:y_i=1} y_i x_i + \sum_{i:y_i=-1} y_i x_i = N_1 \mu_1 - N_{-1} \mu_{-1} = \frac{N}{2} (\mu_1 - \mu_{-1})$

$\Rightarrow 0 = \sum x_i \cdot (\omega^T x_i) + \sum_j \hat{b} x_j - \sum y_i \cdot x_i$

$= \sum (x_i \cdot x_i^T) \cdot \omega - \frac{1}{N} \sum_j \left(\sum_i \omega^T x_i\right) x_j - \frac{N}{2} (\mu_1 - \mu_{-1})$

$\Rightarrow \frac{\mu_1 - \mu_{-1}}{2} = \frac{1}{N} \sum (x_i \cdot x_i^T) \omega - \frac{1}{N^2} \sum_{j,i} (x_j \cdot x_i^T)\, \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T - \frac{1}{N^2} \sum_i \left(\sum_j x_j\right) \cdot x_i^T\right] \cdot \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T - \frac{1}{N^2} \sum_i (N_1 \mu_1 + N_{-1} \mu_{-1}) \cdot x_i^T\right] \cdot \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T - \frac{1}{2N} (\mu_1 + \mu_{-1}) \sum_i x_i^T\right] \cdot \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T - \frac{1}{4} (\mu_1 + \mu_{-1})(\mu_1 + \mu_{-1})^T\right] \cdot \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T + \frac{1}{4} (\mu_1 - \mu_{-1})(\mu_1 - \mu_{-1})^T - \frac{1}{2} (\mu_1 \mu_1^T + \mu_{-1} \mu_{-1}^T)\right] \cdot \omega$

$= \left[\frac{1}{N} \sum_i x_i x_i^T - \frac{1}{2} (\mu_1 \mu_1^T + \mu_{-1} \mu_{-1}^T) + \frac{1}{4} S_B\right] \cdot \omega$

$= \underline{\underline{\left[S_\omega + \frac{1}{4} S_B\right] \cdot \omega}}$

$S_\omega = \frac{1}{N} \sum_i (x_i - \mu_{y_i})(x_i - \mu_{y_i})^T = \frac{1}{N} \sum_i x_i x_i^T - x_i \mu_{y_i}^T - \mu_{y_i} x_i^T + \mu_{y_i} \mu_{y_i}^T$

$= \frac{1}{N} \sum_i x_i x_i^T + \frac{1}{N} \left[\sum_{i:y_i=1} \mu_1 \mu_1^T - x_i \mu_1^T - \mu_1 x_i^T\right] + \frac{1}{N} \left[\sum_{i:y_i=-1} \mu_{-1} \mu_{-1}^T - x_i \mu_{-1}^T - \mu_{-1} x_i^T\right]$

$= \frac{1}{N} \sum_i x_i x_i^T + \frac{1}{N} \left[\frac{N}{2} \mu_1 \mu_1^T - \frac{N}{2} \mu_1 \mu_1^T - \frac{N}{2} \mu_1 \mu_1^T\right] + \frac{1}{N} \left[\frac{N}{2} \mu_{-1} \mu_{-1}^T - \frac{N}{2} \mu_{-1} \mu_{-1}^T - \mu_{-1} \mu_{-1}^T\right]$

$= \frac{1}{N} \sum_i x_i x_i^T - \frac{1}{2} (\mu_1 \mu_1 + \mu_{-1} \mu_{-1}^T)$

3.3 Define $c := (\mu_1 - \mu_{-1})^T \hat{\omega}$

$\Rightarrow S_B \hat{\omega} = (\mu_1 - \mu_{-1})(\mu_1 - \mu_{-1})^T \hat{\omega} = c (\mu_1 - \mu_{-1})$

$S_\omega \hat{\omega} + \frac{1}{4} S_B \hat{\omega} = S_\omega \hat{\omega} + \frac{c}{4} (\mu_1 - \mu_{-1}) = \frac{1}{2} (\mu_1 - \mu_{-1})$

$\Rightarrow S_\omega \hat{\omega} = (c' + \frac{1}{2})(\mu_1 - \mu_{-1})$

$\Rightarrow \hat{\omega} = (c' + \frac{1}{2})\, S_\omega^{-1} (\mu_{-1} - \mu_{-1}) = \underline{\underline{\tau\, S_\omega^{-1} (\mu_1 - \mu_{-1})}}$

$\tau := c' + \frac{1}{2}$