MACHINE LEARNING WS17/18 (Exercise 3)

Max Simon & Debora Fieberg

# 1 LDA-Derivation from the LS Error

decision boundary in LDA given by D-1 dim. hyperplane:

$$\underbrace{\hat{\omega}}_{\text{normal vect.}}{}^T x + \underbrace{\hat{b}}_{\text{scalar}} = 0 \qquad \text{(here: } \omega \,\&\, x \text{ are column vectors!)}$$

decision rule: $\hat{y} = \text{sign}(\hat{\omega}^T x + \hat{b}) = \begin{cases} 1 & , \hat{\omega}^T x + \hat{b} > 0 \\ -1 & , \text{— " —} < 0 \end{cases}$

training phase: N datapoints $\{x_i\}_{i=1\dots N}$, $x_i \in \mathbb{R}^D$ given plus $\{y_i\}_{i=1\dots N}$, $y_i \in \{-1, 1\}$ labels.

(★) Assumption: balanced training set $N_1 = N_{-1} = \frac{N}{2}$

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

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

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

$$\Rightarrow 2Nb = -\sum_{i=1}^{N} \omega^T x_i - y_i$$

$$\Rightarrow \hat{b} = \frac{\sum_{i=1}^{N} y_i - \omega^T x_i}{2N}$$

b) Show that $\frac{\partial}{\partial \omega} \sum_{i=1}^{N} (\omega^T x_i + \hat{b} - y_i)^2 = 0 \iff (S_w + \frac{1}{4} S_B)\hat{\omega} = (\mu_1 - \mu_{-1})$

with $a := \mu_1 = \frac{1}{N_1} \sum_{i; y_i = 1} x_i$ , $b := \mu_{-1} = \frac{1}{N_{-1}} \sum_{i; y_i = -1} x_i$

and $S_B = (a - b)(a - b)^T$ between class

$S_W = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_{y_i})(x_i - \mu_{y_i})^T$ within-class.

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

$\hat{b}$ aus a)

$$\iff \sum_{i=1}^{N} (x_i - \underbrace{\tfrac{\sum x_i}{N}}_{= \frac{a+b}{2}}) \cancel{2} (\omega^T x_i + \tfrac{1}{N}(\overbrace{\cancel{\textstyle\sum_{i=1}^{N} y_i}}^{\sum y_i = 0 \text{ wegen } ★} - \omega^T x_i) - y_i) = 0$$

$= \frac{a+b}{2}$ because $a + b \stackrel{★}{=} \frac{2}{N} (\sum_{y_i=1} x_i + \sum_{y_i=-1} x_i) = \frac{2}{N} \sum x_i$

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

because $\sum y_i = 0$ because ★

$$= \sum_{i=1}^{N} x_i \omega^T x_i - x \omega^T \tfrac{a+b}{2} - x_i y_i - \tfrac{a+b}{2} \omega^T x_i + \tfrac{a+b}{2} \omega^T \tfrac{a+b}{2} + \cancel{\tfrac{a+b}{2} y_i}$$

$$\stackrel{\text{?}}{=} \sum_{i=1}^{N} \left( \left( x_i x_i^T - x_i (\tfrac{a+b}{2})^T - \tfrac{a+b}{2} x_i^T + \tfrac{a+b}{2} (\tfrac{a+b}{2})^T \right) \omega - x_i y_i \right)$$

$$\iff \sum_{i=1}^{N} \left[ x_i x_i^T - \frac{x_i a^T}{2} - \frac{x_i b^T}{2} - \frac{a x_i^T}{2} - \frac{b x_i^T}{2} + \frac{a a^T}{4} + \frac{a b^T}{4} + \frac{b a^T}{4} + \frac{b b^T}{4} \right] \stackrel{★}{=} \sum_{i=1}^{N} x_i y_i = \frac{N}{2}(a - b)$$

$$\Leftrightarrow \underbrace{\frac{1}{N}\sum_{i=1}^{N}\left[2x_ix_i^T - x_ia^T - x_ib^T - ax_i^T - bx_i^T + \frac{aa^T + ab^T + ba^T + bb^T}{2}\right]}_{\overset{?}{=}\; S_w + \frac{1}{4}S_B \;=\; \frac{1}{N}\sum_{i=1}^{N}\left[(x_i-\mu_{y_i})(x_i-\mu_{y_i})^T\right] + \frac{1}{4}(aa^T - ba^T - ab^T + aa^T)}\omega = a-b$$

$$\boxed{\text{c) Show: } \underbrace{\hat{\omega} = \tau S_w^{-1}(a-b)}_{\Updownarrow} \Leftrightarrow \left(S_w + \tfrac{1}{4}S_B\right)\hat{\omega} = a-b}$$

$$\underbrace{(S_w\tau^{-1})\hat{\omega} = a-b}$$

$$\Leftrightarrow S_w\tau^{-1} = S_w + \frac{1}{4}S_B$$

$$\Leftrightarrow \tau = \frac{S_w}{S_w + \frac{S_B}{4}} \quad \text{pos. const.}$$