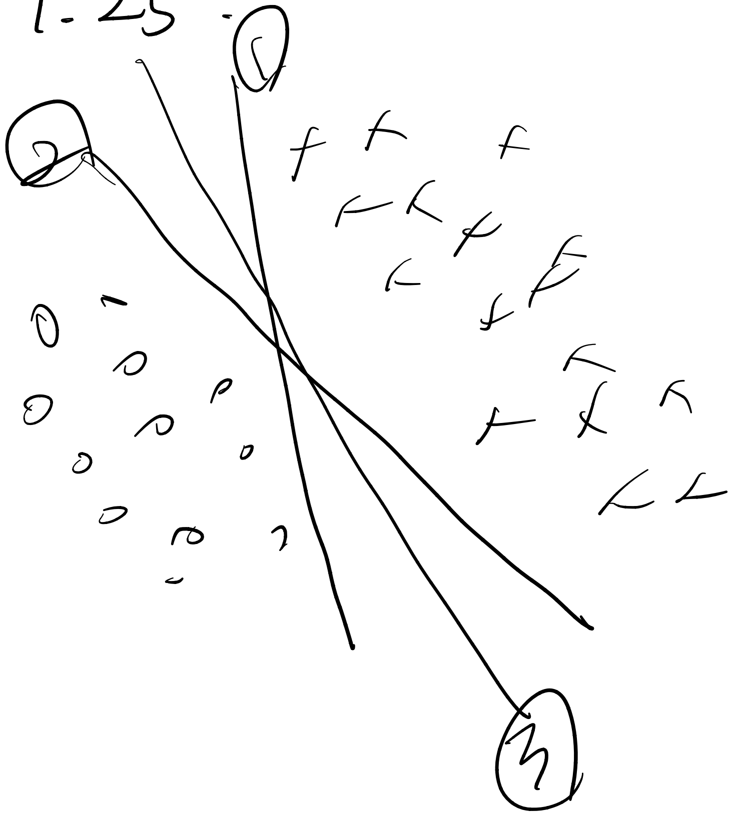
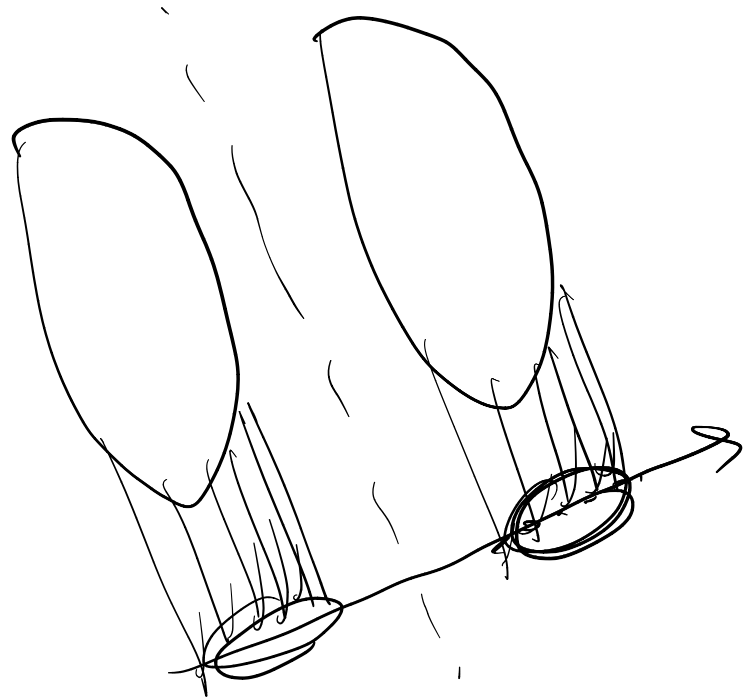
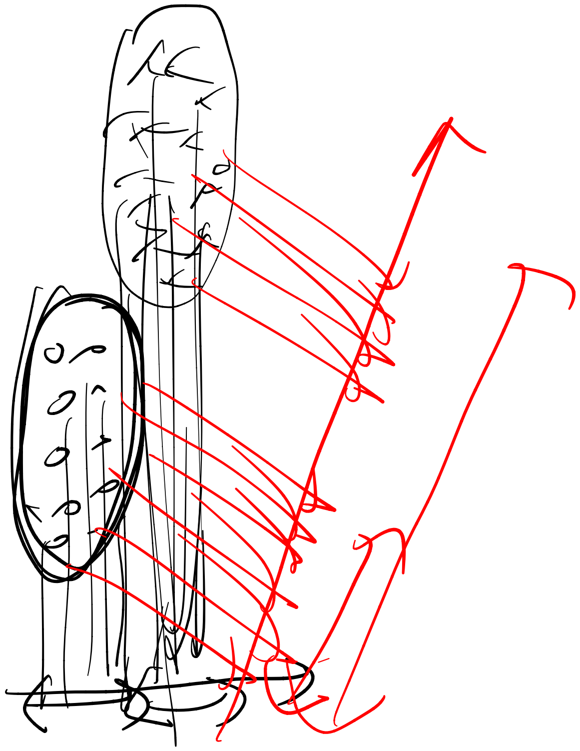
25. 9.25.







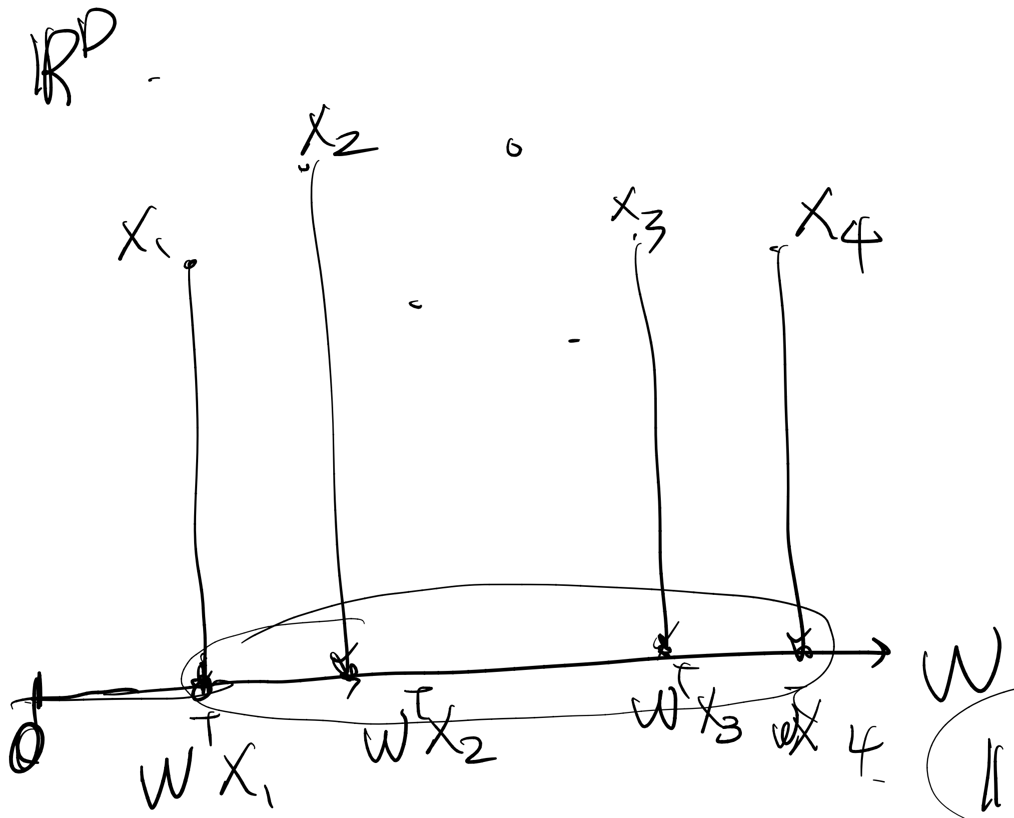
Fisher Discriminant Analysis (FDA)

(Linear " " , LDA)

---

① Variance in the projected 1-d space

$$D = \{x_i\}_{i=1}^{N} \quad x_i \in \mathbb{R}^p$$



$\|w\| = 1$

$$\hat{m} = \frac{1}{N} \sum_{i=1}^{N} w^T x_i$$

$$= w^T \underbrace{\left[\frac{1}{N} \sum_{i=1}^{N} x_i \right]}_{\hat{\mu}}$$

$$= w^T \hat{\mu}$$

$$\mathrm{Var}(w) = \frac{1}{N}\sum_{i=1}^{N}(w^T x_i - \hat{m})^2$$

$$= \frac{1}{N}\sum_{i=1}^{N}(w^T x_i - w^T\hat{\mu})^2$$

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

$$= w^T \underbrace{\left[\frac{1}{N}\sum (x_i - \hat{\mu})(x_i - \hat{\mu})^T\right]}_{\Sigma \in \mathbb{R}^{D\times D}} w$$

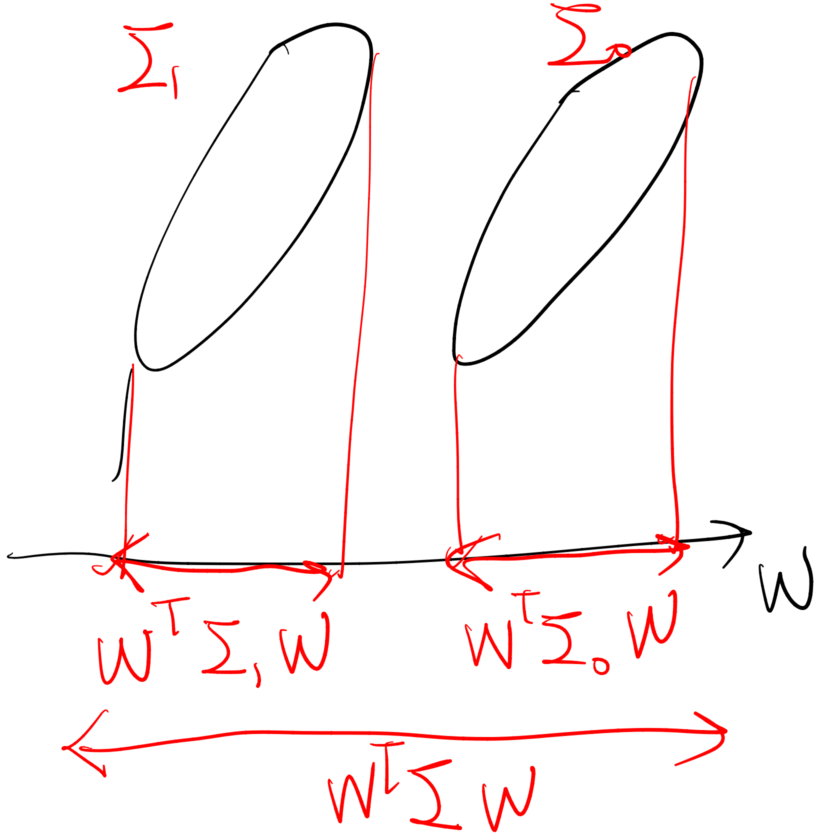
$$= w^T \Sigma w$$

$\boxed{\text{FDA}}$

$D = \{(x_i, y_i)\}_{i=1}^{N}$ $\quad x_i \in \mathbb{R}^D, \quad y_i \in \{0,1\}.$

$|\{i \mid y_i = 1\}| = N_1 \quad |\{i \mid y_i = 0\}| = N_0, \quad N_1 + N_0 = N.$

Σ



$$\max_{w} L(w) = \frac{w^T \Sigma w}{\left(\frac{N_1}{N}\right) w^T \Sigma_1 w + \left(\frac{N_0}{N}\right) w^T \Sigma_0 w}$$

$$\Sigma = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T \qquad \mu = \frac{1}{N} \sum x_i$$

$$\Sigma_1 = \frac{1}{N_1} \sum_{\{i \mid y_i = 1\}} (x_i - \mu_1)(x_i - \mu_1)^T \qquad \mu_1 = \frac{1}{N_1} \sum_{\{i \mid y_i = 1\}} x_i$$

$$\Sigma_0 = \frac{1}{N_0} \sum_{\{i \mid y_i = 0\}} (x_i - \mu_0)(x_i - \mu_0)^T \qquad \mu_0 = \frac{1}{N_0} \sum_{\{i \mid y_i = 0\}} x_i$$

$$\max L(w) = \frac{w^T \Sigma w}{\left(\frac{N_1}{N}\right) w^T \Sigma_1 w + \left(\frac{N_0}{N}\right) w^T \Sigma_0 w} = \frac{w^T \Sigma w}{w^T \Sigma_{10} w}$$

$$\Sigma_{10} = \frac{N_1}{N}\Sigma_1 + \frac{N_0}{N}\Sigma_0$$

$$\frac{dL(w)}{dw} = \frac{1}{(w^T \Sigma_{10} w)^2} \left[\underbrace{\frac{d}{dw}(w^T \Sigma w)}_{=2\Sigma w} \cdot (w^T \Sigma_{10} w) - \underbrace{\frac{d}{dw}(w^T \Sigma_{10} w)}_{=2\Sigma_{10} w} \cdot w^T \Sigma w \right]$$

$$= \frac{1}{(w^T \Sigma_{10} \Sigma)^2} \left(2 \Sigma w \cdot w^T \Sigma_{10} w - 2 \Sigma_{10} w \cdot w^T \Sigma w \right)$$

$$\boxed{\frac{dL}{dW} = 0} \qquad \frac{dL}{dW} = \frac{1}{(W^T \Sigma_{10} W)^2} \left[\cancel{2}\, \Sigma W \cdot W^T \Sigma_{10} W - \cancel{2}\, \Sigma_{10} W \cdot W^T \Sigma W \right)$$

$$= 0$$

$$\boxed{\Sigma \cdot W - \Sigma_{10} W \cdot \frac{W^T \Sigma W}{W^T \Sigma_{10} W} = 0}$$

$$\Sigma \cdot W - \left(\frac{W^T \Sigma W}{W^T \Sigma_{10} W} \right) \Sigma_{10} \cdot W = 0$$