(d) For this part of the problem, you may assume n (the dimension of x) is 1, so that

$$\Sigma = [\sigma^2]$$
$$|\Sigma| = \sigma^2$$

$\Rightarrow$ In gaussian discriminant analysis:

The log-likelihood of data is given by

$$l(\phi, \mu_0, \mu_1, \Sigma) = \log \prod_{i=1}^{m} P(x^{(i)}, y^{(i)}; \phi, \mu_0, \mu_1, \Sigma)$$

$$= \log \prod_{i=1}^{m} P(x^{(i)} \mid y^{(i)}; \mu_0, \mu_1, \Sigma)\, P(y^{(i)}; \phi)$$

$$= \sum_{i=1}^{m} \log \underbrace{P(x^{(i)} \mid y^{(i)}; \mu_0 \mu_1 \Sigma)} + \log \underbrace{P(y^{(i)}; \phi)}$$

$$\frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x^{(i)} - \mu^{(i)})^T \Sigma^{-1} (x^{(i)} - \mu^{(i)})\right)$$

where $\mu^{(i)} = \mu_1^{1\{y^{(i)}=1\}} \mu_0^{1\{y^{(i)}=0\}}$

$$\phi^{1\{y^{(i)}=1\}} (1-\phi)^{1\{y^{(i)}=0\}}$$

Let $l(\phi, \mu_0, \mu_1, \Sigma) = \sum_{i=1}^{m} f(x^{(i)}, y^{(i)}, \mu_0, \mu_1, \Sigma, \phi)$

where, $f = \log P(x^{(i)} \mid y^{(i)}; \mu_0 \mu_1 \Sigma) + \log P(y^{(i)}; \phi)$

$$f = \log\left(\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(\frac{-1}{2}(x^{(i)} - \mu_1)^T \Sigma^{-1} (x^{(i)} - \mu_1)\right)\right)$$

$$+ \log \phi \quad \text{if } y^{(i)} = 1$$

$$f = \log\left(\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(\frac{1}{2}(x^{(i)} - \mu_0)^T \Sigma^{-1} (x^{(i)} - \mu_0)\right)\right)$$

$$+ \log(1-\phi) \quad \text{if } y^{(i)} = 0$$

$$f = \begin{cases} \log\left(\frac{1}{(2\pi)^{n/2}\Sigma^{1/2}} \exp\left(-\left(\frac{x^{(i)2} + \mu_1^2 - 2x^{(i)}\mu_1}{2\Sigma}\right)\right)\right) + \log\phi & \text{if } y^{(i)} = 1 \\ \log\left(\frac{1}{(2\pi)^{n/2}\Sigma^{1/2}} \exp\left(-\left(\frac{x^{(i)2} + \mu_0^2 - 2x^{(i)}\mu_0}{2\Sigma}\right)\right)\right) + \log(1-\phi) & \text{if } y^{(i)} = 0 \end{cases}$$

$$l = \sum_{i=1}^{m} f$$

$$\frac{\delta f}{\delta \phi} = \begin{cases} \frac{1}{\phi} & \text{if } y^{(i)} = 1 \\ \frac{-1}{1-\phi} & \text{if } y^{(i)} = 0 \end{cases}$$

$$\frac{\delta f}{\delta \mu_0} = \begin{cases} 0 & \text{if } y^{(i)} = 1 \\ -\frac{1}{2\Sigma}\left(2\mu_0 - 2x^{(i)}\right) & \text{if } y^{(i)} = 0 \end{cases}$$

$$\rightarrow \left(\frac{x^{(i)} - \mu_0}{\Sigma}\right)$$

$$\frac{\delta f}{\delta \mu_1} = \begin{cases} \frac{x^{(i)} - \mu_1}{\Sigma} & \text{if } y^{(i)} = 1 \\ 0 & \text{if } y^{(i)} = 0 \end{cases}$$

$$\frac{\delta f}{\delta \Sigma} = \begin{cases} \frac{1}{2\Sigma}\left(\frac{x^{(i)^2} + \mu_1^2 - 2x^{(i)}\mu_1}{\Sigma} - 1\right) & \text{if } y^{(i)} = 1 \\ \frac{1}{2\Sigma}\left(\frac{x^{(i)^2} + \mu_0^2 - 2x^{(i)}\mu_0}{\Sigma} - 1\right) & \text{if } y^{(i)} = 0 \end{cases}$$

⇒ For maxima:

$$\frac{\delta l}{\delta \phi} = 0 \ \& \ \frac{\delta l}{\delta \mu_0} = 0 \ \& \ \frac{\delta l}{\delta \mu_1} = 0 \ \& \ \frac{\delta l}{\delta \Sigma} = 0$$

① $\frac{\delta l}{\delta \phi} = 0$

$$\Rightarrow \sum_{i=1}^{m} \left(\frac{1}{\phi}\right)^{1\{y^{(i)}=1\}} \left(\frac{-1}{1-\phi}\right)^{1\{y^{(i)}=1\}} = 0$$

$$\sum_{i=1}^{m} 1\{y^{(i)}=1\} = a$$

$$\text{L.t} \quad \sum_{i=1}^{m} 1\{y^{(i)}=0\} = b$$

$$\Rightarrow \frac{a}{\phi} + \frac{-b}{1-\phi} = 0 \Rightarrow \frac{a}{\phi} = \frac{b}{1-\phi}$$

$$\Rightarrow a - a\phi = b\phi \Rightarrow \phi = \frac{a}{\underbrace{a+b}_{m}}$$

$$\boxed{\phi = \frac{1}{m} \sum_{i=1}^{m} 1\{y^{(i)}=1\}}$$

$\cancel{② \ \frac{\delta l}{\delta \mu_0} = 0}$

$$\cancel{\Rightarrow \sum_{i=1}^{m} \left(x^{(i)} - \mu_0\right)^{1\{y^{(i)}=0\}}} \quad \sum b(x^{(i)} - \mu_0)$$

② $$\frac{\delta l}{\delta \mu_0} = 0$$

$$= \sum_{i=1}^{m} 1\{y^{(i)}=0\} x^{(i)} - b\mu_0 = 0$$

$$\Rightarrow \boxed{\mu_0 = \frac{\sum_{i=1}^{m} 1\{y^{(i)}=0\} x^{(i)}}{\sum_{i=1}^{m} 1\{y^{(i)}=0\}}}$$

③ $$\frac{\delta l}{\delta \mu_1} = 0$$

$$\boxed{\mu_1 = \frac{\sum_{i=1}^{m} 1\{y^{(i)}=1\} x^{(i)}}{\sum_{i=1}^{m} 1\{y^{(i)}=1\}}}$$

④ $$\frac{\delta l}{\delta \Sigma} = 0$$

$$\sum_{i=1}^{m} \left(x^{(i)^2} + \mu_{y^{(i)}}^2 + 2x^{(i)} \mu_{y^{(i)}} - \Sigma \right) = 0$$

$$m\Sigma = \sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})(x^{(i)} - \mu_{y^{(i)}})^T$$

Date

Page

$$\Rightarrow \boxed{\Sigma = \frac{1}{m} \sum_{i=1}^{m} \left(x^{(i)} - \mu_{y^{(i)}}\right)\left(x^{(i)} - \mu_{y^{(i)}}\right)^T}$$