# 1. (d)

$\Sigma = [\sigma^2]$ $\qquad\qquad\qquad x \in \mathbb{R},\ \Sigma \in \mathbb{R}$

$|\Sigma| = \sigma^2$

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

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

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

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})(x^{(i)} - \mu_{y^{(i)}})^T \xrightarrow[\text{1 dim}]{} \frac{1}{m}\sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})^2$$

$$\left\{\begin{aligned} l(\phi, \mu_0, \mu_1, \Sigma) &= \log \prod_{i=1}^{m} p(x^{(i)}, y^{(i)}; \phi, \mu_0, \mu_1, \Sigma) \\ &= \log \prod_{i=1}^{m} p(x^{(i)} | y^{(i)}; \phi, \mu_0, \mu_1, \Sigma)\, p(y^{(i)}; \phi_1) \end{aligned}\right.$$

By maximizing $l$ with respect to four parameters, prove that maximum likelihood estimates of $\phi, \mu_0, \mu_1, \Sigma$ are indeed as given in the formula above

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

$$= \sum_{i=1}^{m} \log \left\{ p(x^{(i)} | y^{(i)}; \phi, \mu_0, \mu_1, \Sigma)\, p(y^{(i)}; \phi) \right\}$$

$$= \sum_{i=1}^{m} \log \left\{ p(x^{(i)} | y^{(i)}; \phi, \mu_0, \mu_1, \Sigma) \right\} + \sum_{i=1}^{m} \log p(y^{(i)}; \phi)$$

$$= \sum_{i=1}^{m} \log \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} (x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right\} + \sum_{i=1}^{m} \log \phi^{1\{y^{(i)}=1\}} (1-\phi)^{1-1\{y^{(i)}=1\}}$$

$$= \sum_{i=1}^{m} \left[ \left( \log \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \right) + \left( -\frac{1}{2} (x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right) \right] + \sum_{i=1}^{m} \log \phi^{1\{y^{(i)}=1\}} (1-\phi)^{1-1\{y^{(i)}=1\}}$$

$$= \sum_{i=1}^{m} \left[ \log \left( \frac{1}{2\pi} \right)^{\frac{n}{2}} + \log \left( \frac{1}{|\Sigma|} \right)^{1/2} \right] + \sum_{i=1}^{m} \left( -\frac{1}{2} (x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right) + \sum_{i=1}^{m} 1\{y^{(i)}=1\} \log \phi +$$

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

$$= \frac{mn}{2} \log \frac{1}{2\pi} + \frac{m}{2} \log \left( \frac{1}{|\Sigma|} \right) + \sum_{i=1}^{m} \left( -\frac{1}{2} (x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right) + \sum_{i=1}^{m} 1\{y^{(i)}=1\} \log \phi + (m - 1\{y^{(i)}=1\}) \log (1-\phi)$$

$$\log \mathcal{L} = -\frac{mn}{2} \log 2\pi - \frac{m}{2} \log |\Sigma| + \sum_{i=1}^{m} \left( -\frac{1}{2} (x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right) + \sum_{i=1}^{m} 1\{y^{(i)}=1\} \log \phi + (m - \sum_{i=1}^{m} 1\{y^{(i)}=1\}) \log (1-\phi)$$

$$\frac{\partial \log \mathcal{L}}{\partial \phi} = \frac{1}{\phi} \times \sum_{i=1}^{m} 1(y^{(i)}=1) + \frac{1}{1-\phi} \times (-1) \times \left( m - \sum_{i=1}^{m} 1\{y^{(i)}=1\} \right) = 0$$

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

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

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

$$\log \mathcal{L} = -\frac{mn}{2}\log 2\pi - \frac{m}{2}\log|\Sigma| + \sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^T\Sigma^{-1}(x^{(i)}-\mu_{y^{(i)}})\right) + \sum_{i=1}^{m}1\{y^{(i)}=1\}\log\phi + \left(m-\sum_{i=1}^{m}1\{y^{(i)}=1\}\right)\log(1-\phi)$$

$$\frac{\partial \log \mathcal{L}}{\partial \mu_0} = \frac{\partial}{\partial \mu_0}\left\{\sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^T\Sigma^{-1}(x^{(i)}-\mu_{y^{(i)}})\right)\right\} = 0$$

$$= \frac{\partial}{\partial \mu_0}\left\{\sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^2\Sigma^{-1}\right)\right\} = 0$$

$$\frac{\partial}{\partial \mu_0}\left[-\frac{1}{2}\left\{\sum_{i=1}^{m}1\{y^{(i)}=1\}(x^{(i)}-\mu_1)^2 + \left(m-\sum_{i=1}^{m}1\{y^{(i)}=1\}\right)(x^{(i)}-\mu_0)^2\right\}\right] = 0$$

$$\frac{\partial}{\partial \mu_0}\sum_{i=1}^{m}1\{y^{(i)}=0\}(x^{(i)}-\mu_0)^2 = 0$$

$$\frac{\partial}{\partial \mu_0}\sum_{i=1}^{m}1\{y^{(i)}=0\}\left((x^{(i)})^2-2\mu_0 x^{(i)}+\mu_0^2\right) = 0$$

$$\sum_{i=1}^{m}1\{y^{(i)}=0\}(-2x^{(i)}+2\mu_0) = 0$$

$$\sum_{i=1}^{m}1\{y^{(i)}=0\}(2x^{(i)}) = \sum_{i=1}^{m}1\{y^{(i)}=0\}\,2\mu_0$$

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

✓

$$\log \mathcal{L} = -\frac{mn}{2}\log 2\pi - \frac{m}{2}\log|\Sigma| + \sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^T\Sigma^{-1}(x^{(i)}-\mu_{y^{(i)}})\right) + \sum_{i=1}^{m}1\{y^{(i)}=1\}\log\phi + \left(m-\sum_{i=1}^{m}1\{y^{(i)}=1\}\right)\log(1-\phi)$$

$$\frac{\partial \log\mathcal{L}}{\partial \mu_1} = \frac{\partial}{\partial \mu_1}\left\{\sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})^T\Sigma^{-1}(x^{(i)}-\mu_{y^{(i)}})\right\} = 0$$

$$= \frac{\partial}{\partial \mu_1}\left\{\sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})^2\Sigma^{-1}\right\} = 0$$

$$= \frac{\partial}{\partial \mu_1}\left\{\sum_{i=1}^{m}1\{y^{(i)}=1\}(x^{(i)}-\mu_{y^{(i)}})^2 + \sum_{i=1}^{m}1\{y^{(i)}=0\}(x^{(i)}-\mu_{y^{(i)}})^2\right\} = 0$$

$$= \frac{\partial}{\partial \mu_1}\left\{\sum_{i=1}^{m}1\{y^{(i)}=1\}\left((x^{(i)})^2 - 2x^{(i)}\mu_{y^{(i)}} + (\mu_{y^{(i)}})^2\right) = 0\right.$$

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

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

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

$\frac{1}{x}$

$$\frac{\partial \log\mathcal{L}}{\partial \Sigma} = \frac{\partial}{\partial \Sigma}\left\{-\frac{m}{2}\log|\Sigma| + \sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^T\Sigma^{-1}(x^{(i)}-\mu_{y^{(i)}})\right)\right\}$$

$$\Sigma = \sigma^2 \rightarrow \frac{\partial}{\partial \sigma^2}\left\{-\frac{m}{2}\log\sigma^2 + \sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^2 \times \frac{1}{\sigma^2}\right)\right\} = 0$$

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

$$-\frac{m}{2}\times\frac{1}{\sigma^2} = \sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu_{y^{(i)}})^2\right)\frac{1}{\sigma^4}$$

$$\frac{m}{2}\sigma^2 = \frac{1}{2}\sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})^2$$

$$m\sigma^2 = \sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})^2$$

$$\sigma^2 = \frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})^2 = \boxed{\frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu_{y^{(i)}})(x^{(i)}-\mu_{y^{(i)}})^T}$$