A) MLE for Fisher LDA : $Y \sim \text{Bernoulli}(\pi)$

$X|Y=j \sim \mathcal{N}(\mu_j, \Sigma)$

$$\ell(\pi, \mu, \Sigma) = \log \prod_{j=1}^{N} p(y_j|\pi)\, p(x_j|y_j, \theta)$$

$$= \sum_{j=1}^{N} \left(\log p(y_j|\pi) + \log p(x_j|y_j, \theta) \right)$$

$$P(y_j = \pi) = \pi^{y_j}(1-\pi)^{1-y_j}, \quad P(x_j|y_j) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left(-\tfrac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$\Rightarrow \ell(\pi, \mu, \Sigma) = \sum_{j=1}^{N} \left(\log \pi^{y_j} + \log(1-\pi)^{1-y_j} - \log 2\pi|\Sigma|^{1/2} - \tfrac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right)$$

$$= \sum_{j=1}^{N} \left(y_j \log \pi + (1-y_j)\log(1-\pi) - \tfrac{1}{2}\log|\Sigma| - \log 2\pi - \tfrac{1}{2}(x_j-\mu)^T \Sigma^{-1}(x_j-\mu) \right)$$

MLE for π:

$$\hat{\pi}_{MLE} = \underset{\pi}{\operatorname{argmax}}\, \ell(\pi, \mu, \Sigma)$$

$$\nabla_\pi \ell(\pi, \mu, \Sigma) = \sum_j \left(\frac{y_j}{\pi} - \frac{(1-y_j)}{1-\pi} \right) = 0.$$

$$\Rightarrow \sum_j \frac{y_j}{\pi} = \sum_j \frac{(1-y_j)}{1-\pi} = \frac{\sum_j 1 - \sum_j y_j}{1-\pi} = \frac{N - \sum_j y_j}{1-\pi}$$

$$\frac{(1-\pi)\sum_j y_j}{\pi} = N - \sum_j y_j$$

$$\left(\frac{1}{\pi} - 1\right)\cancel{\sum_j y_j} = N - \cancel{\sum_j y_j}$$

$$\frac{1}{\pi}\sum_j y_j = N \Rightarrow \boxed{\hat{\pi}_{MLE} = \frac{\sum_j y_j}{N}}$$

MLE for μ_i $\quad i = \{0,1\}$.

$$\nabla_{\mu_i} \ell(\pi,\mu,\Sigma) = \nabla_{\mu_i}\left(\sum_{j=1}^{N}\sum_{k=0}^{1} \left|-\frac{1}{2}(x_j-\mu_k)^T\Sigma^{-1}(x_j-\mu_k)\right)\right) = 0$$

$$\Rightarrow \sum_{j=1}^{N} -\Sigma^{-1}(x_j-\mu_i) = 0$$

$$\sum_{j=1}^{N} x_j - N\mu_i = 0$$

$$\Rightarrow \boxed{\hat{\mu}_{i\,MLE} = \frac{\sum_{j=1} x_j}{N}}$$

MLE for Σ

$$\nabla_{\Sigma^{-1}} \ell(\pi,\mu,\Sigma) = \nabla_{\Sigma}\left(\sum_{j=1}^{N}\left|-\frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_j-\mu)\Sigma^{-1}(x_j-\mu)\right)\right) = 0$$

$$= -\frac{n}{2}\Sigma + -\frac{n}{2}\hat{\Sigma} = 0.$$

$$\Sigma^{-1}$$

$$p(y_n|\pi) = \pi^{y_n}(1-\pi)^{y_n}, \quad p(x_n|y_n,\theta)$$

$$= p(x_n|y_n=1,\mu_1,\Sigma^{-1})^{y_n} p(x_n|y_n=0,\mu_0,\Sigma^{-1})^{1-y_n}$$

$$p(x,y) = \prod_{n=1}^{N} p(y_n|\pi)\,p(x_n|y_n,\theta)$$

$$\ell(\pi,\theta) = \log p(x,y) = \sum_{n=1}^{n}\left(\log p(y_n|\pi) + \log p(x_n|y_n,\theta)\right)$$

$$\theta = (\mu_0, \mu_1, \Sigma^{-1})$$

MLE for π $\nabla_\pi \ell(\pi, \theta) = \sum_{n=1}^{n} \nabla_\pi \log(y_n | \pi) = 0.$

$$\Rightarrow \sum_{n=1}^{n} \nabla_\pi \left(y_n \log \pi + (1-y_n) \log(1-\pi) \right) = 0$$

$$\Leftrightarrow \hat{\pi}_{MLE} = \sum_{n=1}^{N} \frac{y_n}{N}$$

MLE for μ_1 $\nabla_{\mu_1} \ell(\pi, \theta) = \sum_{n=1}^{n} \nabla_{\mu_1} \log p(x_n | y_n, \theta) = 0$

$$\Rightarrow \sum_{n=1}^{N} \nabla_{\mu_1} \left(\log \left\{ p(x_n | y_n=1, \mu_1, \Sigma^{-1})^{y_n} p(x_n | y_n=0, \mu_0, \Sigma^{-1})^{1-y_n} \right\} \right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \nabla_{\mu_1} \left(y_n \log p(x_n | y_n=1, \mu_1, \Sigma^{-1}) + \underbrace{(1-y_n) \log(x_n | y_n=0, \mu_0, \Sigma^{-1})}_{=0,\ \text{does not depend on } \mu_1} \right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \nabla_{\mu_1} y_n \left(-\cancel{\log} 2\pi - \frac{1}{2} \cancel{\log} |\Sigma| - \frac{1}{2} (x_n - \mu_1)^T \Sigma^{-1} (x_n - \mu_1) \right) = 0.$$

$$\Rightarrow \sum_{n=1}^{N} y_n \Sigma^{-1} (x_n - \mu_1) = 0$$

$$\Rightarrow \sum_{n=1}^{N} y_n x_n = \sum_{n=1}^{N} \mu_1$$

$$\Rightarrow \hat{\mu}_{1\,MLE} = \frac{\sum_{n=1}^{N} y_n x_n}{\sum_{n=1}^{N} y_n}$$

For μ_0 : $y_n \mapsto 1-y_n$

MLE for Σ^{-1}

$$\ell(\pi,\mu_1,\mu_0,\Sigma^{-1}) = \sum_{n=1}^{N} \left(\log p(y_n|\pi) + \log p(x_n|y_n,\mu_1,\mu_0,\Sigma^{-1}) \right)$$

$$\Rightarrow \nabla_{\Sigma^{-1}} \ell = \sum_{n=1}^{N} \nabla_{\Sigma^{-1}} \log p(x_n|y_n,\mu_1,\mu_0,\Sigma^{-1}) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \nabla_{\Sigma^{-1}} \left(y_n \log p(x_n|y_n=1,\mu_1,\Sigma^{-1}) + (1-y_n) \log p(x_n|y_n=0,\mu_0,\Sigma^{-1}) \right) = 0.$$

$$\Rightarrow \sum_{n=1}^{N} \nabla_{\Sigma^{-1}} \left\{ y_n \left(\cancel{-\log 2\pi} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_n-\mu_1)^T \Sigma^{-1} (x_n-\mu_1) \right) + (1-y_n) \left(\right. \right.$$

$$\left. \left. \cancel{-\log 2\pi} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_n-\mu_0)^T \Sigma^{-1} (x_n-\mu_0) \right) \right\}.$$

Matrix calculus tricks: $\nabla_A \log \det A = A^{-1}$

$$\Rightarrow \nabla_{\Sigma^{-1}} \log \det \Sigma = \Sigma$$

also $(x-\mu)^T \Sigma^{-1} (x-\mu)$ is a scalar so

$$= \mathrm{Tr}\left((x-\mu)^T \Sigma^{-1} (x-\mu) \right)$$

Cyclic property of trace:

$$= \mathrm{Tr}\left(\Sigma^{-1} (x-\mu)(x-\mu)^T \right) = \langle \Sigma^{-1}, (x-\mu)(x-\mu)^T \rangle$$

$$\Rightarrow \nabla_{\Sigma^{-1}} \ell = \sum_{n=1}^{N} \left(y_n \left(-\frac{1}{2}\Sigma - \nabla_{\Sigma^{-1}} \frac{1}{2} \langle \Sigma^{-1}, (x_n-\mu_1)(x_n-\mu_1)^T \rangle \right) \right.$$

$$\left. + (1-y_n) \left(-\frac{1}{2}\Sigma - \nabla_{\Sigma^{-1}} \langle \Sigma^{-1}, (x_n-\mu_0)(x_n-\mu_0)^T \rangle \right) \right) = 0$$

$$\Rightarrow \sum_{n=1}^{N} \left(y_n \left(-\frac{1}{2}\Sigma - \frac{1}{2} \underbrace{(x_n-\mu_1)(x_n-\mu_1)^T}_{=\tilde{\mu}_1} \right) + (1-y_n) \left(-\frac{1}{2}\Sigma \right. \right.$$

$$\left. \left. -\frac{1}{2}\tilde{\mu}_0 \right) \right) = 0$$

$$\Rightarrow -\sum_{n=1}^{N} \cancel{y_n} \frac{\Sigma}{2} - \sum_{n=1}^{N} y_n \frac{\tilde{\mu}_1}{2} - \sum_{n=1}^{N} \frac{\Sigma}{2} + \cancel{\sum_{n=1}^{N} y_n \frac{\Sigma}{2}} + (1-y_n) \sum_{n=1}^{N} \tilde{\mu}_0 \; {=0}$$

$$\Rightarrow -\frac{1}{2}\sum_{n=1}^{N} y_n \tilde{\mu}_1 - \frac{N}{2}\Sigma - \frac{1}{2}\sum_{n=1}^{N}(1-y_n)\tilde{\mu}_0 = 0.$$

$$\Rightarrow \hat{\Sigma}_{MLE} = \frac{\sum_{n=1}^{N} y_n \tilde{\mu}_1 + \sum_{n=1}^{N}(1-y_n)\tilde{\mu}_0}{N}$$

$$\boxed{\hat{\Sigma}_{MLE} = \frac{\sum_{n=1}^{N} y_n (x_n-\mu_1)(x_n-\mu_1)^T + \sum_{n=1}^{N}(1-y_n)(x_n-\mu_0)(x_n-\mu_0)^T}{N}}$$

$$p(y=1|x) = \frac{p(x|y=1)\,p(y=1)}{p(x)}$$

$$p(x) = \sum_{y} p(x,y)$$

Data

	x_1	x_2	y
1	:		
2			
⋮			
n			

$X: n\times 2 \quad Y = n\times 1$

$n\times 1 \,@\, n\times 2$

$(2\times n).(n\times 1) = 2\times 1 \quad \begin{pmatrix} \cdot \\ \cdot \end{pmatrix}$

$1\times n \,.\, n\times 2 \qquad 1\times 2$

$n\times 1 \qquad 2\times 2$

$\cancel{n\times 2} \quad n\times 2$

$X - n\mu = n\times 2.$

$n\times 2 \,.\, n\times 2$

$2\times n \,.\, n\times 2 \qquad 2\times 2.$

$\sigma(w^T\tilde{x})$

$$w = \begin{pmatrix} \mu_1^T\Sigma^{-1} - \mu_0^T\Sigma^{-1} \\ \mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1 + \log\frac{\pi}{1-\pi} \end{pmatrix}, \tilde{x} = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

$$w^T\tilde{x} = (\mu_0^T\Sigma^{-1} - \mu_0^T\Sigma^{-1})X \qquad \begin{pmatrix} \vdots & \vdots \end{pmatrix}$$

1×2

$n \times 2 \quad (2 \times n)\, n \times 2 \quad = (2 \times 2) \times (n \times 2)$

$(2 \times 2) \cdot x \quad 2 \times n$

$(2 \times n)(x\ n \times 1) = 2 \times 1$

$(\ )$

$1 - N_0$

$\frac{N_1}{N} \quad \pi.$

Model: $p(x|y=j) = \mathcal{N}(\mu_j, \Sigma_j)$

$p(y) = \pi^y(1-\pi)^{1-y}$

GDA derivation

posterior $p(y=1|x) = \frac{p(x|y=1)p(y=1)}{p(x)}$ (Baye's rule)

but

$p(x) = \sum_y p(x,y) = \sum_y p(x|y)p(y)$ (marginalization ⊕ chain rule)

$$\Rightarrow p(y=1|x) = \frac{p(x|y=1)p(y=1)}{p(x|y=1)p(y=1) + p(x|y=0)p(y=0)} = \frac{\mathcal{N}(\mu_1,\Sigma_1)\pi}{\mathcal{N}(\mu_1,\Sigma_1)\pi + \mathcal{N}(\mu_0,\Sigma_0)(1-\pi)}$$

$$= \frac{1}{1+e^{-f(x)}} \equiv \sigma(x)$$

where $f(x) \equiv \log\left(\frac{\mathcal{N}(\mu_1,\Sigma_1)\pi}{\mathcal{N}(\mu_0,\Sigma_1)(1-\pi)}\right)$

Let's compute $f(x)$.

$$f(x) = \log \mathcal{N}(\mu_1,\Sigma_1) - \log \mathcal{N}(\mu_0,\Sigma_0) + \log\frac{\pi}{1-\pi}$$

$$= \cancel{\log 2\pi} - \frac{1}{2}\log|\Sigma_1| - \frac{1}{2}(x-\mu_1)^T\Sigma_1(x-\mu_1) - \cancel{\log 2\pi} + \frac{1}{2}\log|\Sigma_0|$$
$$+ \frac{1}{2}(x-\mu_0)^T\Sigma_0(x-\mu_0) + \log\frac{\pi}{1-\pi}$$

$$= \frac{1}{2}\left(x^T(\Sigma_0^{-1}-\Sigma_1^{-1})x + 2(\mu_1^T\Sigma_1^{-1} - \mu_0^T\Sigma_0^{-1})x + \log\frac{|\Sigma_0|}{|\Sigma_1|}\right.$$
$$\left. + \mu_0^T\Sigma_0\mu_0 - \mu_1^T\Sigma_1\mu_1\right) + \log\left(\frac{\pi}{1-\pi}\right)$$

$$= \omega^T\tilde{x}$$

with

$$\omega \equiv \begin{pmatrix} \frac{1}{2}(\Sigma_0^{-1}-\Sigma_1^{-1}) \\ \mu_1^T\Sigma_1^{-1} - \mu_0^T\Sigma_0^{-1} \\ \frac{1}{2}\left(\mu_0^T\Sigma_0\mu_0 - \mu_1^T\Sigma_1\mu_1 + \log\frac{|\Sigma_0|}{|\Sigma_1|}\right) + \log\frac{\pi}{1-\pi} \end{pmatrix}, \quad \tilde{x} \equiv \begin{pmatrix} x^2 \\ x \\ 1 \end{pmatrix}$$

$$\pi = \frac{N_1}{N} \qquad \frac{\pi}{N_1} = \frac{1}{N}$$

$$1 - \frac{N_1}{N} \qquad \frac{N - N_1}{N} = \frac{N_0}{N.N_0} = \frac{1}{N}$$

X: $n \times 3$, $\omega = 3 \times 1$

$(400 \times 1) \times X \; (400 \times 3)$