1. (c)

Gaussian Discriminant Analysis

$$p(y) = \begin{cases} \phi & \text{if } y=1 \\ 1-\phi & \text{if } y=0 \end{cases}$$

$\phi \in \mathbb{R}$
$\mu_0, \mu_1 \in \mathbb{R}^n$
$\Sigma \in \mathbb{R}^{n \times n}$

$$p(x|y=0) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0)\right)$$

$$p(x|y=1) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma^{-1} (x-\mu_1)\right)$$

To show that GDA results in a linear decision boundary,
show that posterior distribution can be written as:

$$p(y=1 \mid x ; \phi, \mu_0, \mu_1, \Sigma) = \frac{1}{1+\exp(-(\theta^T x + \theta_0))} \qquad \theta \in \mathbb{R}^n, \ \theta_0 \in \mathbb{R}$$

For convenience of notations, params $= \phi, \mu_0, \mu_1, \Sigma$

$$p(y=1 \mid x ; \text{params}) = \frac{p(x \mid y=1 ; \text{params}) \boxed{p(y=1 ; \text{params})}}{\boxed{p(x ; \text{params})}}$$

$\leftarrow \phi$

$$p(x \mid y=0 ; \text{params}) \times \underbrace{p(y=0 ; \text{params})}_{1-\phi} + p(x \mid y=1 ; \text{params}) \times \underbrace{p(y=1 ; \text{params})}_{\phi}$$

$$= \frac{p(x \mid y=1 ; \text{params})\, \phi}{(1-\phi)\, p(x \mid y=0 ; \text{params}) + \phi \times p(x \mid y=1 ; \text{params})}$$

$$= \frac{\frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma^{-1} (x-\mu_1)\right) \phi}{(1-\phi) \times \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0)\right) + \phi \times \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_1)^T \Sigma^{-1} (x-\mu_1)\right)}$$

$$p(y=1\mid x; \text{params}) = \frac{\phi \times \exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)}{\phi \times \exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right) + (1-\phi)\times \exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)}$$

$$= \frac{1}{1 + \frac{(1-\phi)}{\phi} \times \left\{ \frac{\exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)}{\exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)} \right\}}$$

$$= \frac{1}{1 + \frac{(1-\phi)}{\phi} \times \exp\left\{ -\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) + \frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) \right\}}$$

$$= \frac{1}{1 + \exp\left\{ \boxed{\log\frac{1-\phi}{\phi} - \frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) + \frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)} \right\}}$$

$$= -(\theta^T x + \theta_0)$$

$\phi \in \mathbb{R}$

$\mu_0, \mu_1 \in \mathbb{R}^n$

$\Sigma \in \mathbb{R}^{n\times n}$

$$\boxed{\log\frac{1-\phi}{\phi} \boxed{- \frac{1}{2}\underbrace{\underbrace{(x-\mu_0)^T}_{1\times n}\underbrace{\Sigma^{-1}}_{n\times n}\underbrace{(x-\mu_0)}_{n\times 1}}_{1\times 1} + \frac{1}{2}\underbrace{\underbrace{(x-\mu_1)^T}_{1\times n}\underbrace{\Sigma^{-1}}_{n\times n}\underbrace{(x-\mu_1)}_{n\times 1}}_{1\times 1}}}$$

$$-\frac{1}{2}\left\{ (x^T-\mu_0^T)\Sigma^{-1}(x-\mu_0) - (x^T-\mu_1^T)\Sigma^{-1}(x-\mu_1) \right\}$$

$$= -\frac{1}{2}\left\{ x^T\Sigma^{-1}(x-\mu_0) - \mu_0^T\Sigma^{-1}(x-\mu_0) - \left(x^T\Sigma^{-1}(x-\mu_1) - \mu_1^T\Sigma^{-1}(x-\mu_1)\right) \right\}$$

$$= -\frac{1}{2}\left\{ x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_0 - \mu_0^T\Sigma^{-1}x + \mu_0^T\Sigma^{-1}\mu_0 - \left(x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_1 - \mu_1^T\Sigma^{-1}x + \mu_1^T\Sigma^{-1}\mu_1\right) \right\}$$

$$= -\frac{1}{2}\left\{ x^T\Sigma^{-1}x - x^T\Sigma^{-1}\mu_0 - \mu_0^T\Sigma^{-1}x + \mu_0^T\Sigma^{-1}\mu_0 - x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu_1 + \mu_1^T\Sigma^{-1}x - \mu_1^T\Sigma^{-1}\mu_1 \right\}$$

$$= -\frac{1}{2}\left\{ \cancel{x^T\Sigma^{-1}x} - \underbrace{x^T\Sigma^{-1}\mu_0 - \mu_0^T\Sigma^{-1}x}_{=} + \mu_0^T\Sigma^{-1}\mu_0 - \cancel{x^T\Sigma^{-1}x} + \underbrace{x^T\Sigma^{-1}\mu_1 + \mu_1^T\Sigma^{-1}x}_{=} - \mu_1^T\Sigma^{-1}\mu_1 \right\}$$

$$= -\frac{1}{2}\left\{ -2\times \mu_0^T\Sigma^{-1}x + \mu_0^T\Sigma^{-1}\mu_0 + 2\times \mu_1^T\Sigma^{-1}x - \mu_1^T\Sigma^{-1}\mu_1 \right\}$$

$$= -\frac{1}{2}\left\{ -2\left(\mu_0^T\Sigma^{-1}x - \mu_1^T\Sigma^{-1}x\right) + \mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1 \right\}$$

$$= \left(\mu_0^T\Sigma^{-1}x - \mu_1^T\Sigma^{-1}x\right) - \frac{1}{2}\left(\mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right) \leftarrow \boxed{\;}$$

$$\frac{1}{1+\exp\left(\log\frac{1-\phi}{\phi} + \boxed{\;}\right)} \rightarrow \frac{1}{1+\exp\left\{ \log\frac{1-\phi}{\phi} + (\mu_0-\mu_1)^T\Sigma^{-1}x - \frac{1}{2}\left(\mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right)\right\}}$$

$$\rightarrow \frac{1}{1+\exp\left[-\left\{-\log\frac{1-\phi}{\phi} - (\mu_0-\mu_1)^T\Sigma^{-1}x + \frac{1}{2}\left(\mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right)\right\}\right]}$$

$$\theta^T = -(\mu_0-\mu_1)^T\Sigma^{-1} \qquad \theta \in \mathbb{R}^n$$

$$\Sigma : \text{symmetric}$$

$$\therefore\ \theta = \left\{-(\mu_0-\mu_1)^T\Sigma^{-1}\right\}^T$$
$$= -\Sigma^{-1}(\mu_0-\mu_1)$$
$$= \Sigma^{-1}(\mu_1-\mu_0) \checkmark$$

$$\theta_0 = \frac{1}{2}\left(\mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right) - \log\frac{1-\phi}{\phi} \checkmark$$

c.f.)

$$\frac{1}{2}(\mu_0+\mu_1)^T\Sigma^{-1}(\mu_0-\mu_1)$$

$$= \frac{1}{2}(\mu_0^T+\mu_1^T)\Sigma^{-1}(\mu_0-\mu_1)$$

$$= \frac{1}{2}\left\{\mu_0^T\Sigma^{-1}(\mu_0-\mu_1) + \mu_1^T\Sigma^{-1}(\mu_0-\mu_1)\right\}$$

$$= \frac{1}{2}\left\{\mu_0^T\Sigma^{-1}\mu_0 - \mu_0^T\Sigma^{-1}\mu_1 + \mu_1^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right\}$$

$$= \frac{1}{2}\left\{\mu_0^T\Sigma^{-1}\mu_0 - \mu_1^T\Sigma^{-1}\mu_1\right\}$$