Luke Doman

B455 - HW2

1. a. $D = X_1 \ldots X_n \sim N(\theta, \sigma_1^2)$ $\quad \theta \sim N(\mu, \sigma_2^2)$

$$\theta_{map} = \underset{\theta}{\operatorname{argmax}} P(D|\theta) P(\theta)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{-(X_i-\theta)^2}{2\sigma_1^2}\right) \cdot \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(\frac{-(\theta-\mu)^2}{2\sigma_2^2}\right)$$

$$\log \theta_{map} = \sum_{i=1}^{n} \cancel{\log\left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)} - \frac{(X_i-\theta)^2}{2\sigma_1^2} + \cancel{\log\left(\frac{1}{\sqrt{2\pi\sigma_2^2}}\right)} - \frac{(\theta-\mu)^2}{2\sigma_2^2}$$

$$\frac{d}{d\theta} = 0 = -\sum_{1}^{n} \frac{(X_i-\theta)^2}{2\sigma_1^2} - \frac{(\theta-\mu)^2}{2\sigma_2^2}$$

$$= +\sum_{1}^{n} \frac{\cancel{2}(X_i-\theta)\cancel{(-1)}}{\cancel{2}\sigma_1^2} - \frac{\cancel{2}(\theta-\mu)}{\cancel{2}\sigma_2^2}$$

$$= \sum_{1}^{n} \frac{(X_i-\theta)}{\sigma_1^2} - \frac{\theta-\mu}{\sigma_2^2}$$

$$= \frac{\sigma_2^2 \sum_{1}^{n}(X_i-\theta)}{\sigma_1^2\sigma_2^2} - \frac{\sigma_1^2(\theta-\mu)}{\sigma_1^2\sigma_2^2}$$

$$\boxed{= \frac{\sigma_2^2 \sum_{1}^{n}(X_i-\theta) + \sigma_1^2(\theta-\mu)}{\sigma_1^2\sigma_2^2}}$$

b. $D = X_1 \ldots X_n \sim N(\theta, \sigma^2)$ $\quad \theta \sim \frac{1}{2b} \exp\left(\frac{-|x-\mu|}{b} \right)$

$$\theta_{map} = \underset{\theta}{\operatorname{argmax}}\, P(D|\theta)\, P(\theta)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(X_i-\theta)^2}{2\sigma^2} \right) \frac{1}{2b} \exp\left(\frac{-|\theta-\mu|}{b} \right)$$

$$\log \theta_{map} = \sum_{1}^{n} \cancel{\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)} - \frac{(X_i-\theta)^2}{2\sigma^2} + \cancel{\log\left(\frac{1}{2b}\right)} - \frac{|\theta - \cancel{\mu}^{\,0}|}{b}$$

0 - given

$$\frac{d}{d\theta} = 0 = -\sum_{1}^{n} \frac{(X_i-\theta)^2}{2\sigma^2} - \frac{|\theta|}{b}$$

$$= +\sum_{1}^{n} \frac{\cancel{2}(X_i-\theta)\cancel{(-1)}}{\cancel{2}\sigma^2} - \frac{|\theta|}{\theta b}\cancel{(1)}$$

$$= \frac{\theta b \sum_{1}^{n}(X_i-\theta)}{\theta b \sigma^2} - \frac{\sigma^2|\theta|}{\theta b \sigma^2}$$

$$\boxed{= \frac{\theta b \sum_{1}^{n}(X_i-\theta) - \sigma^2|\theta|}{\theta b \sigma^2}}$$

c. $D = X_1 \dots X_n$ where $X_i \sim \mathcal{N}(\theta, \Sigma_0)$

$$\Sigma_0 = I \in \mathbb{R}^{d \times d}, \quad \theta \sim \mathcal{N}(\mu = 0, \Sigma = \sigma^2 I)$$

$$\theta_{map} = \underset{\theta}{\operatorname{argmax}} \; P(D \mid \theta) P(\theta)$$

$$= \prod_{i=1}^{d} \prod_{j=1}^{d} \left(\frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(\frac{-(X_{ij} - \theta)^2}{2\sigma_j^2} \right) \right) \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(\frac{-\theta}{2\sigma_i^2} \right)$$

$$\log \theta_{map} = \sum_{i=1}^{d} \sum_{j=1}^{d} \left(\cancel{\log\left(\frac{1}{\sqrt{2\pi\sigma_j^2}} \right)} - \frac{(X_{ij} - \theta)^2}{2\sigma_j^2} \right) + \cancel{\log\left(\frac{1}{\sqrt{2\pi\sigma_i^2}} \right)} - \frac{\theta}{2\sigma_i^2}$$

$$= -\sum_{i=1} \sum_{j=1} \frac{(X_{ij} - \theta)^2}{2\sigma_j^2} - \frac{\theta}{2\sigma_i^2}$$

$$\frac{\partial}{\partial \theta} = 0 = +\sum_{i=1} \sum_{j=1} \frac{\cancel{2}(X_{ij} - \theta)\cancel{(-1)}}{\cancel{2}\sigma_j^2} + \frac{\theta \cancel{(-1)}}{2\sigma_i^2}$$

$$\boxed{= \sum_{i=1}^{d} \sum_{j=1}^{d} \frac{(X_{ij} - \theta)}{\sigma_j^2} + \frac{\theta}{2\sigma_i^2}}$$

2. a. As sample size increases, so does accuracy, until a point is reached where it fails because it's a singular matrix. This is due to creating the matrix from too many identical/very similar features. We can overcome this by using singular value decomposition.

b - g. Please see implentation/comments in attached python files.