Written Examination Answers

Exercise 1.

(1) For each $k \in \{1, \dots, K\}$, $i \in \{1, \dots, N\}$,

$$\begin{aligned} \frac{\partial \Phi}{\partial \mu_{k,i}} &= \sum_{n=1}^{N} z_{n,k} \frac{\partial}{\partial \mu_{k,i}} \left((x_{n,i} - \mu_{k,i})^2 \right) \\ &= \sum_{n=1}^{N} z_{n,k} \times 2(x_{n,i} - \mu_{k,i}) \,. \end{aligned}$$

So

$$\nabla \Phi(\mu_1, \dots, \mu_K) = 2 \sum_{n=1}^{N} \begin{pmatrix} z_{n,1}(x_n - \mu_1) \\ \vdots \\ z_{n,K}(x_n - \mu_K) \end{pmatrix} \in (\mathbb{R}^N)^K \,.$$

(2) The only critical point of Φ is in:

$$\forall k \,,\, \mu_k = \frac{\sum_{n=1}^{N} z_{n,k} x_n}{\sum_{n=1}^{N} z_{n,k}} \,.$$

As Φ is convex, this is the absolute minimum.

Exercise 2. We compute

$$\begin{aligned} \boldsymbol{w}^T \boldsymbol{X}^T \boldsymbol{X} \boldsymbol{w} &= [w_0, w_1] \begin{bmatrix} x_{1,1} & x_{2,1} & \dots & x_{N,1} \\ x_{1,2} & x_{2,2} & \dots & x_{N,2} \end{bmatrix} \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ \vdots & \vdots \\ x_{N,1} & x_{N,2} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \\ &= \begin{bmatrix} w_0 x_{1,1} & w_0 x_{2,1} & \dots & w_0 x_{N,1} \\ w_1 x_{1,2} & w_1 x_{2,2} & \dots & w_1 x_{N,2} \end{bmatrix} \begin{bmatrix} w_0 x_{1,1} & w_1 x_{1,2} \\ w_0 x_{2,1} & w_1 x_{2,2} \\ \vdots & \vdots \\ w_0 x_{N,1} & w_1 x_{N,2} \end{bmatrix} \\ &= w_0^2 \sum_{n=1}^{N} x_{n,1}^2 + 2 w_0 w_1 \sum_{n=1}^{N} x_{n,1} x_{n,2} + w_1^2 \sum_{n=1}^{N} x_{n,2}^2 \end{aligned}$$

Exercise 3.

(1) We set $X_i(\omega) = x_i$ $(i \in \{1, 2, \dots, N\})$. We compute (using the i.i.d. assumption) for any $x_1, \dots x_N \in \{0, 1\}$,

$$L(\theta) = \mathbb{P}(X_1 = x_1, \dots, X_N = x_N | \theta) = \prod_{n=1}^{N} \theta^{x_n} (1-\theta)^{1-x_n} = \theta^{N_1} (1-\theta)^{N_0} \,.$$

(2) We work on the log of L. We have

$$\begin{aligned} \frac{d}{d\theta}(\log(L(\theta)) &= \frac{d}{d\theta} \left(N_1 \log(\theta) + N_0 \log(1-\theta) \right) \\ &= \sum_{n=1}^{N} \left(\frac{N_1}{\theta} - \frac{N_0}{1-\theta} \right) \,. \end{aligned}$$

We set $\mathcal{L}(\theta) = \log(L(\theta))$. We observe that $\mathcal{L}$ is concave and has exactly one critical point in

$$\widehat{\theta} = \frac{N_1}{N_1 + N_0} \,.$$

So this is the absolute maximum (and hence is the MLE of θ).