

Ahmed B. Iqbal

* checked from soln

a)
$$\begin{aligned} P(W) &= \sum_i \sum_k r_{ik} \log P(x_i | \theta_k) \\ &= \sum_i \sum_k r_{ik} \sum_j x_{ij} \log \mu_{kj} + (1 - x_{ij}) \log (1 - \mu_{kj}) \end{aligned}$$

i: datapoint index, k: component, j: dimension index
of D dimension vectors

$$\begin{aligned} \frac{\partial L}{\partial \mu_{kj}} &= \sum_i r_{ik} \left(\frac{x_{ij}}{\mu_{kj}} - \frac{1 - x_{ij}}{1 - \mu_{kj}} \right) \\ &= \sum_i r_{ik} \left(\frac{x_{ij} - \mu_{kj}}{\mu_{kj}(1 - \mu_{kj})} \right) = \frac{1}{\mu_{kj}(1 - \mu_{kj})} \sum_i r_{ik} (x_{ij} - \mu_{kj}) \\ &= 0 \end{aligned}$$

Optimality condition $\sum_i r_{ik} x_{ij} = \mu_{kj} \sum_i r_{ik}$

b) checked solution

$$\begin{aligned} l(\mu) &= \sum_i \sum_k r_{ik} \log P(x_i | \mu_k) + \log P(N_k) \\ &= \sum_i \sum_k r_{ik} \left(\sum_j x_{ij} \log \mu_{kj} + (1 - x_{ij}) \log (1 - \mu_{kj}) \right) \\ &\quad + \\ &\quad (a-1) \log N_k + (b-1) \log (1 - N_k) \end{aligned}$$

* checked soln:

$$\begin{aligned} \frac{\Delta l}{\Delta \mu} &= \sum_j \left(\frac{r_{ik} x_{ij} + a - 1}{N_{kj}} - \frac{r_{ik} (1 - x_{ij}) + b - 1}{1 - N_{kj}} \right) \\ &= \frac{\sum_i r_{ik} x_{ij} - r_{ik} N_{kj} + a - 1 - N_{kj} a + N_{kj} - N_{kj} b + N_{kj}}{N_{kj} (1 - N_{kj})} \\ &= \frac{\sum_i r_{ik} x_{ij} - \left(\sum_i r_{ik} + a + b - 2 \right) N_{kj} + a - 1}{N_{kj} (1 - N_{kj})} = 0 \end{aligned}$$

Optimality condition:

$$\begin{aligned} &\sum_i r_{ik} x_{ij} + a - 1 \\ &= \sum_i r_{ik} x_{ij} + a - 1 = \left(\sum_i r_{ik} + a + b - 2 \right) N_{kj} \quad \checkmark \end{aligned}$$

2) checked soln.

$$\text{prox}_{\gamma}(x)_i = \begin{cases} x_i - \gamma & x_i > \gamma \\ 0 & |x_i| \leq \gamma \\ x_i + \gamma & x_i < -\gamma \end{cases}$$

Hence, each iterate

$$x_{i+1} = \text{prox}_{\gamma}(x_i - \gamma \nabla f(x_i)) \quad \gamma: \text{learning rate}$$

$$\nabla \|x\|_1 = \text{sign}(x_i)$$

$$\begin{aligned} \nabla \|Ax - b\|_2^2 + \lambda \|x\|_1 &= \nabla x^T A^T A x - 2b^T A x + b^T b + \lambda \|x\|_1 \\ &= \underbrace{2A^T A x - 2b^T A + \lambda \text{sign}(x)}_{\text{checked soln}} + b^T b \end{aligned}$$