MS1b: SDM - Problem Sheet 3

1. (Expectation-Maximization) Assume you are interested in clustering n binary images. Each binary image $\mathbf{x}$ corresponds to a vector of p binary random variables; p being the number of pixels. We adopt a probabilistic approach and model the probability mass function of $\mathbf{x}$ using a finite mixture model

$$f(\mathbf{x}|\theta) = \sum_{k=1}^{K} \pi_k f(\mathbf{x}|\phi_k)$$

where $f(\mathbf{x}|\phi_k)$ is a product of p Bernoulli distributions of parameters $\phi_k = (\phi_k^1, ..., \phi_k^p) \in [0,1]^p$ and $\pi_1, \ldots, \pi_K$ satisfy $\pi_k \geq 0 \; \forall k$ and $\sum_{k=1}^{K} \pi_k = 1$. We want to estimate the unknown parameters $\theta = \{\pi_k, \phi_k\}_{k=1}^{K}$ given $\mathbf{x}_{1:n}$ by maximizing the associated likelihood function using the EM algorithm. Establish explicitly the EM update; i.e. the expression of the estimate $\theta^{(t)}$ at iteration t as a function of $\theta^{(t-1)}$.

Answer: Each data $\mathbf{x} = (x^1, ..., x^p) \in \{0,1\}^p$ is modeled by

$$f\left(\mathbf{x}|\,\theta\right) = \sum_{k=1}^{K} \pi_k f\left(\mathbf{x}|\,\phi_k\right)$$

with $\phi_k = (\phi_k^1, ..., \phi_k^p)$ and

$$f\left(\mathbf{x}|\,\phi_k\right) = \prod_{l=1}^{p} \left(\phi_k^l\right)^{x^l} \left(1 - \phi_k^l\right)^{1-x^l}.$$

The EM equations are simply

$$\pi_k^{(t)} = \frac{\sum_{i=1}^{n} \Pr\left(z_i = k|\,\mathbf{x}_i, \phi_k^{(t-1)}\right)}{n},$$

$$\left(\phi_k^l\right)^{(t)} = \frac{\sum_{i=1:x_i^l=1}^{n} \Pr\left(z_i = k|\,\mathbf{x}_i, \phi_k^{(t-1)}\right)}{\sum_{i=1}^{n} \Pr\left(z_i = k|\,\mathbf{x}_i, \phi_k^{(t-1)}\right)}$$

where

$$\Pr\left(z_i = k|\,\mathbf{x}_i, \phi_k^{(t-1)}\right) = \frac{\pi_k^{(t-1)} f\left(\mathbf{x}_i|\,\phi_k^{(t-1)}\right)}{f\left(\mathbf{x}_i|\,\theta^{(t-1)}\right)}.$$