Problem 8.16

X is Bernoulli with $Pr(X=1) = 1 - Pr(X=0) = p$. Y given $X = j$ is normal with mean μ_j and variance σ^2.

Consider now the monotone missing-data pattern with Y completely observed but n-r values of X missing and an ignorable mechanism.

Describe the E and M steps of EM algorithm for this problem.

preparation

Let $\theta = (p, \mu_0, \mu_1, \sigma)$, we have

$$f(y_i|x_i = j; \theta) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(y_i-\mu_j)^2}{2\sigma_2}\right]$$

$$f(x_i = 1|\theta) = 1 - f(x_i = 0|\theta) = p$$

E-step

$$w_i = E[X_i|x, y, \theta^{(t)}] = \begin{cases} x_i & , i \leq r \\ \frac{f(y_i|x_i=1;\theta^{(t)})f(x_i=1;\theta^{(t)})}{f(y_i|x_i=0;\theta^{(t)})f(x_i=0;\theta^{(t)})+f(y_i|x_i=1;\theta^{(t)})f(x_i=1;\theta^{(t)})} & , i > r \end{cases}$$

M-step

$$\begin{aligned} p^{(t+1)} &= \sum_{i=1}^{n} w_i/n \\ \mu_0^{(t+1)} &= \sum_{i=1}^{n}(1-w_i)y_i/\sum_{i=1}^{n}(1-w_i) \\ \mu_1^{(t+1)} &= \sum_{i=1}^{n} w_i y_i/\sum_{i=1}^{n} w_i \\ \sigma^{(t+1)} &= \sqrt{\sum_{i=0}^{n}[(1-w_i)(y_i-\mu_0^{(t+1)})^2 + w_i(y_i-\mu_1^{(t+1)})^2]/n} \end{aligned}$$