Notation & Formulas for EM Examples

Example 1. (`Multivariate t`)

In this example, the observed data vectors are the rows of an $n \times p$ data matrix, $Y_{(i)} \sim \mathcal{N}(\mu, \Sigma/\tau_i)$ conditionally given τ_i, where $\tau_i \sim \chi^2_\nu/\nu$ and (Y_i, τ_i) are *iid*. Here ν is regarded as konwn, and $\vartheta = (\mu, \Sigma)$ is the unknown parameter.

For the EM algorithm, you are to use the conditional distribution

$$f_{\tau_i|Y_i}(t|y) \sim \Gamma(\frac{\nu+p}{2}, \frac{\nu+\Delta}{2})$$

where $\Delta = \Delta(y, \mu, \Sigma) = (y-\mu)'\Sigma^{-1}(y-\mu)$, to justify in full details the EM iterative step:

$$\mu_{new} = \sum_{i=1}^{n} Y_i/(\nu + \Delta(Y_i, \mu_{old}, \Sigma_{old}))$$

$$\Sigma_{new} = \frac{1}{n}\sum_{i=1}^{n} \frac{\nu+p}{\nu + \Delta(Y_i, \mu_{old}, \Sigma_{old})} (Y_i - \mu_{new})^{\otimes 2}$$

Example 2. (`Multivariate Normal with Variance` $D + \sigma^2 1^{\otimes 2}$)

Here the data are $Y_{(i)} \sim \mathcal{N}(\mu, D + \sigma^2 1^{\otimes 2})$ with $D = \text{diag}(\mathbf{d})$ and unknown parameter $\vartheta = (\mu, \mathbf{d}, \sigma^2)$. We regard the missing data as U_i in the representation

$$Y_{(i)} = Z_{(i)} + U_i \mathbf{1} \quad , \qquad Z_{(i)}, U_{(i)} \quad \text{independent}$$

$$Z_{(i)} \sim \mathcal{N}(\mu, D) \quad , \qquad U_i \sim \mathcal{N}(0, \sigma^2)$$

Note first that if $\mu_{old} = \hat{\mu}$ initially, then the EM iteration will preserve the same value of μ_{new} in all subsequent iterations. Beyond showing this, your task is to justify in full details the EM iterative step: in terms of the auxiliary parameter $\alpha = \alpha(\mathbf{d}) = tr(D) = \sum_{j=1}^p d_j^{-1}$,

$$\sigma^2_{new} = \frac{\sigma^2_{old}}{1 + \alpha_{old}\sigma^2_{old}} \left(1 + \frac{\sigma^2_{old}}{1 + \alpha_{old}\sigma^2_{old}} \mathbf{1}' D^{-1}_{old} S_Y D^{-1}_{old} \mathbf{1}\right)$$

$$d_{new,k} = S_{Y,kk} + \sigma^2_{new} - \frac{2\sigma^2_{old}}{1 + \alpha_{old}\sigma^2_{old}} \sum_{l=1}^{p} S_{Y,kl}/d_{old,l}$$