The EM Algorithm (see pp 440-441 of Bishop (2006))

Given a joint distribution $p(X, Z \mid \theta)$ over observed variables X and latent variables Z, governed by parameters θ, the goal is to maximize the likelihood function $p(X \mid \theta)$ with respect to θ.

1. Choose an initial setting for the parameters θ^{old}
2. E Step: Evaluate $p(Z \mid X, \theta^{old})$
3. M Step: Evaluate θ^{new} given by

$$\theta^{new} = \underset{\theta}{\arg\max}\, Q(\theta, \theta^{old}) \qquad (9.32)$$

where

$$Q(\theta, \theta^{old}) = \sum_{Z} p(Z \mid X, \theta^{old}) \ln p(X, Z \mid \theta)$$

4. Check for convergence of either the log-likelihood or the parameter values. If the convergence criterion is not met, then let

$$\theta^{old} \leftarrow \theta^{new} \qquad (9.34)$$

and return to Step 2.

EM Algorithm for the Gaussian Distribution Case:

$$\left.\begin{array}{l} X_i \sim N(\mu_1, \sigma_1^2) \\ D_i \sim N(\mu_2, \sigma_2^2) \end{array}\right\} \quad S_i = \min(X_i, D_i) \quad \text{for } i = 1, \ldots, N.$$

Densities of X_i and D_i are given as follows:

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right\} = \frac{1}{\sigma_1}\,\phi\left(\frac{x-\mu_1}{\sigma_1}\right)$$

$$f_D(x) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right\} = \frac{1}{\sigma_2}\,\phi\left(\frac{x-\mu_2}{\sigma_2}\right),$$

where ϕ is the normal pdf, and Φ is the normal cdf. Because $S_i = \min(X_i, D_i)$, we write

$$\mathbb{P}(S_i \le s) = \mathbb{P}(X_i \le s)\,\mathbb{P}(D_i \le s) = \Phi\left(\frac{s-\mu_1}{\sigma_1}\right)\Phi\left(\frac{s-\mu_2}{\sigma_2}\right).$$

Its density, denoted by $f_S(s)$, is given as follows:

$$f_S(s) = \frac{d}{ds}\mathbb{P}(S \le s)$$

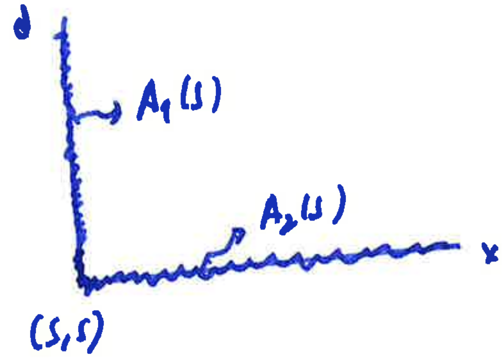
$$= \frac{1}{\sigma_1}\,\phi\left(\frac{s-\mu_1}{\sigma_1}\right)\Phi\left(\frac{s-\mu_2}{\sigma_2}\right) + \frac{1}{\sigma_2}\,\phi\left(\frac{s-\mu_2}{\sigma_2}\right)\Phi\left(\frac{s-\mu_1}{\sigma_1}\right)$$

E-Step. For the E-Step, we need to calculate $f(x,D\mid s,\mu^{old},\sigma^{old})$.

To this end, let $A(s) = A_1(s) \cup A_2(s)$ where

$$A_1(s) = \{(s,d) : d \geq s\}$$

$$A_2(s) = \{(x,s) : x \geq s\}$$



Then we write

$$f(x,D\mid s,\mu^{old},\sigma^{old}) \propto \begin{cases} \dfrac{f(x,D,s\mid \mu^{old},\sigma^{old})}{f(s\mid \mu^{old},\sigma^{old})} & \text{if } (x,D)\in A(s) \\[2ex] 0 & \text{otherwise} \end{cases}$$

For $(x,D)\in A(s)$, we have that

$$f(x,D\mid s,\mu^{old},\sigma^{old}) \propto$$

$$\frac{\dfrac{1}{\sigma_1^{old}}\dfrac{1}{\sigma_2^{old}}\,\phi\left(\dfrac{x_1-\mu_1^{old}}{\sigma_1^{old}}\right)\phi\left(\dfrac{D-\mu_2^{old}}{\sigma_2^{old}}\right)}{\dfrac{1}{\sigma_1^{old}}\,\phi\left(\dfrac{s-\mu_1^{old}}{\sigma_1^{old}}\right)\Phi\left(\dfrac{s-\mu_2^{old}}{\sigma_2^{old}}\right)+\dfrac{1}{\sigma_2^{old}}\,\phi\left(\dfrac{s-\mu_2^{old}}{\sigma_2^{old}}\right)\Phi\left(\dfrac{s-\mu_1^{old}}{\sigma_1^{old}}\right)}$$

4/

M-Step. Define $Q_i(\mu,\sigma,\mu^{old},\sigma^{old})$ as follows:

$$Q_i(\mu,\sigma,\mu^{old},\sigma^{old}) \propto \int_{(x,D)\in A(s)} f(x,D|s_i,\mu^{old},\sigma^{old}) \ln f(x,D,s_i|\mu,\sigma)$$

$$= \int_{A_1(s)} f(s_i,D|s_i,\mu^{old},\sigma^{old}) \ln f(s_i,D,s_i|\mu,\sigma)\, dD$$
$$+ \int_{A_2(s)} f(x,s_i|s_i,\mu^{old},\sigma^{old}) \ln f(x,s_i,s_i|\mu,\sigma)\, dx$$

$$= \int_s^\infty f(s_i,D|s_i,\mu^{old},\sigma^{old}) \ln f(s_i,D,s_i|\mu,\sigma)\, dD$$
$$+ \int_s^\infty f(x,s_i|s_i,\mu^{old},\sigma^{old}) \ln f(x,s_i,s_i|\mu,\sigma)\, dx \qquad (*)$$

Note that for $(x,D)\in A(s)$,

$$\ln f(x,D,s|\mu,\sigma) = \ln\left(f_x(x) f_D(D)\right)$$
$$= -\ln\sigma_1 - \ln\sqrt{2\pi} - \frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \ln\sigma_2 - \ln\sqrt{2\pi} - \frac{1}{2}\left(\frac{D-\mu_2}{\sigma_2}\right)^2$$
$$= -2\ln\sqrt{2\pi} - \ln\sigma_1 - \ln\sigma_2 - \frac{x^2}{2\sigma_1} + \frac{x\mu_1}{\sigma_1^2} - \frac{\mu_1^2}{2\sigma_1^2} - \frac{D^2}{2\sigma^2} + \frac{D\mu_2}{\sigma_2^2} - \frac{\mu_2^2}{2\sigma_2^2}$$

Substituting this into $(*)$ while ignoring the term $-2\ln\sqrt{2\pi}$ that do not depend on (μ,σ) gives the following:

$$Q_i(\mu, \sigma, \mu^{old}, \sigma^{old}) =$$

$$\int_{s_i}^{\infty} f(s_i, D \mid s_i, \mu^{old}, \sigma^{old}) \left[-\ln \sigma_1 - \ln \sigma_2 - \frac{s_i^2}{2\sigma_1^2} + \frac{s_i \mu_1}{\sigma_1^2} - \frac{\mu_1^2}{2\sigma_1^2} - \frac{D^2}{2\sigma_2^2} + \frac{D\mu_2}{\sigma_2^2} - \frac{\mu_2^2}{2\sigma_2^2} \right] dD$$

$$+ \int_{s_i}^{\infty} f(x, s_i \mid s_i, \mu^{old}, \sigma^{old}) \left[-\ln \sigma_1 - \ln \sigma_2 - \frac{x^2}{2\sigma_1^2} + \frac{x\mu_1}{\sigma_1^2} - \frac{\mu_1^2}{2\sigma_1^2} - \frac{s_i^2}{2\sigma_2^2} + \frac{s_i \mu_2}{\sigma_2^2} - \frac{\mu_2^2}{2\sigma_2^2} \right] dx$$

Let

$$D_0^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} f(s_i, D \mid s_i, \mu^{old}, \sigma^{old})\, dD,$$

$$D_1^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} D f(s_i, D \mid s_i, \mu^{old}, \sigma^{old})\, dD,$$

$$D_2^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} D^2 f(s_i, D \mid s_i, \mu^{old}, \sigma^{old})\, dD,$$

$$X_0^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} f(x, s_i \mid s_i, \mu^{old}, \sigma^{old})\, dx,$$

$$X_1^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} x f(x, s_i \mid s_i, \mu^{old}, \sigma^{old})\, dx,$$

$$X_2^i(\mu^{old}, \sigma^{old}) = \int_{s_i}^{\infty} x^2 f(x, s_i \mid s_i, \mu^{old}, \sigma^{old})\, dx.$$

Using these definitions, we write

$$Q_i(\mu, \sigma, \mu^{old}, \sigma^{old}) = \left(-\ln\sigma_1 - \ln\sigma_2 - \frac{s_i^2}{2\sigma_1^2} + \frac{s_i\mu_1}{\sigma_1^2} - \frac{\mu_1^2}{2\sigma_1^2} - \frac{\mu_2^2}{2\sigma_2^2}\right) D_0^i$$

$$+ \frac{\mu_2}{\sigma_2^2} D_1^i - \frac{D_2^i}{2\sigma_2^2} + \left(-\ln\sigma_1 - \ln\sigma_2 - \frac{\mu_1^2}{2\sigma_1^2} - \frac{s_i^2}{2\sigma_2^2} + \frac{s_i\mu_2}{\sigma_2^2} - \frac{\mu_2^2}{2\sigma_2^2}\right) X_0^i$$

$$+ \frac{\mu_1}{\sigma_1^2} X_1^i - \frac{X_2^i}{2\sigma_1^2}.$$

After rearranging the terms, we have that

$$Q_i(\mu, \sigma, \mu^{old}, \sigma^{old}) = -\ln\sigma_1 (D_0^i + X_0^i) - \frac{\mu_1^2}{2\sigma_1^2}(D_0^i + X_0^i) + \frac{\mu_1}{\sigma_1^2}(X_1^i + s_i)$$

$$-\frac{1}{2\sigma_1^2}(s_i^2 + X_2^i) - \ln\sigma_2(D_0^i + X_0^i) - \frac{\mu_2^2}{2\sigma_2^2}(D_0^i + X_0^i) + \frac{\mu_2}{\sigma_2^2}(D_2^i + s_i)$$

$$-\frac{1}{2\sigma_2^2}(D_2^i + s_i^2).$$

Recall that we wish to maximize

$$\sum_{i=1}^N Q_i(\mu, \sigma, \mu^{old}, \sigma^{old}) = -\ln\sigma_1 \sum_{i=1}^N (D_0^i + X_0^i) - \frac{\mu_1^2}{2\sigma_1^2}\sum_{i=1}^N (D_0^i + X_0^i)$$

$$+ \frac{\mu_1}{\sigma_1^2}\sum_{i=1}^N (X_1^i + s_i) - \frac{1}{2\sigma_1^2}\sum_{i=1}^N (s_i^2 + X_2^i)$$

$$-\ln\sigma_2 \sum_{i=1}^N (D_0^i + X_0^i) - \frac{\mu_2^2}{2\sigma_2^2}\sum_{i=1}^N (D_0^i + X_0^i) + \frac{\mu_2}{\sigma_2^2}\sum_{i=1}^N (D_1^i + s_i)$$

$$-\frac{1}{2\sigma_2^2}\sum_{i=1}^N (D_2^i + s_i^2)$$

7/

Clearly, the problem decomposes across (μ_1, σ_1) vs. (μ_2, σ_2). So we consider them separately. FOCs wrt μ_1 gives:

$$-\frac{\mu_1}{\sigma_1^2} \sum_{i=1}^{N} (D_0^i + X_0^i) + \frac{1}{\sigma_1^2} \sum_{i=1}^{N} (X_1^i + S_i) = 0 \Rightarrow$$

$$\boxed{\mu_1^{new} = \frac{\sum_{i=1}^{N} (X_1^i + S_i)}{\sum_{i=1}^{N} (D_0^i + X_0^i)}} \quad (**)$$

FOCs wrt σ_1 gives

$$-\frac{1}{\sigma_1} \sum_{i=1}^{N} (D_0^i + X_0^i) + \frac{\mu_1^2}{\sigma_1^3} \sum_{i=1}^{N} (D_0^i + X_0^i) - \frac{2\mu_1}{\sigma_1^3} \sum_{i=1}^{N} (X_1^i + S_i)$$

$$+ \frac{1}{\sigma_1^3} \sum_{i=1}^{N} (S_i^2 + X_2^i) = 0$$

$$\frac{1}{\sigma_1^2} \left[\mu_1^2 \sum_{i=1}^{N} (D_0^i + X_0^i) - 2\mu_1 \sum_{i=1}^{N} (X_1^i + S_i) + \sum_{i=1}^{N} (S_i^2 + X_2^i) \right] = \sum_{i=1}^{N} (D_0^i + X_0^i)$$

Substituting μ_1^{new} from $(**)$ gives

$$\frac{1}{\sigma_1^2} \left[\frac{\left(\sum_{i=1}^{N} (X_1^i + S_i)\right)^2}{\left(\sum_{i=1}^{N} (D_0^i + X_0^i)\right)^2} \cdot \sum_{i=1}^{N} (D_0^i + X_0^i) - 2\frac{\left(\sum_{i=1}^{N} (X_1^i + S_i)\right)^2}{\sum_{i=1}^{N} (D_0^i + X_0^i)} + \sum_{i=1}^{N} (S_i^2 + X_2^i) \right] = \sum_{i=1}^{N} (D_0^i + X_0^i)$$

8/

$$\frac{1}{\sigma_1^2}\left[\sum_{i=1}^{N}(S_i^2 + X_2^i) - \frac{\left(\sum_{i=1}^{N}(X_1^i + S_i)\right)^2}{\sum_{i=1}^{N}(D_0^i + X_0^i)}\right] = \sum_{i=1}^{N}(D_0^i + X_0^i)$$

$$(\sigma_1^2)^{new} = \frac{\sum_{i=1}^{N}(S_i^2 + X_2^i)}{\sum_{i=1}^{N}(D_0^i + X_0^i)} - \left(\frac{\sum_{i=1}^{N}(X_1^i + S_i)}{\sum_{i=1}^{N}(D_0^i + X_0^i)}\right)^2 \qquad (***)$$

Similarly, we have that

$$\mu_2^{new} = \frac{\sum_{i=1}^{N}(D_1^i + S_i)}{\sum_{i=1}^{N}(D_0^i + X_0^i)}$$

$$(\sigma_2^2)^{new} = \frac{\sum_{i=1}^{N}(S_i^2 + D_2^i)}{\sum_{i=1}^{N}(D_0^i + X_0^i)} - \left(\frac{\sum_{i=1}^{N}(D_1^i + S_i)}{\sum_{i=1}^{N}(D_0^i + X_0^i)}\right)^2$$

We will show below that

$$D_0^i = X_0^i = 1 \quad \forall i.$$

Thus, our estimates for μ^{new}, σ^{new} are as follows:

$$\mu_1^{new} = \frac{1}{2N} \sum_{i=1}^{N} \left(X_1^i + S_i \right) \qquad \mu_2^{new} = \frac{1}{2N} \sum_{i=1}^{N} \left(D_1^i + S_i \right)$$

$$(\sigma_1^2)^{new} = \frac{1}{2N} \sum_{i=1}^{N} \left(X_2^i + S_i^2 \right) - \left(\frac{1}{2N} \sum_{i=1}^{N} \left(X_1^i + S_i \right) \right)^2$$

$$(\sigma_2^2)^{new} = \frac{1}{2N} \sum_{i=1}^{N} \left(D_2^i + S_i^2 \right) - \left(\frac{1}{2N} \sum_{i=1}^{N} \left(D_1^i + S_i \right) \right)^2$$

Next, we derive easy-to-compute expressions for $D_0^i, D_1^i, D_2^i, X_0^i, X_1^i, X_2^i$ using the following auxiliary lemma that considers indefinite integrals:

Lemma 1

i) $\int \phi(x)\,dx = \Phi(x) + C_0$

ii) $\int x\,\phi(x)\,dx = -\phi(x) + C_1$

iii) $\int x^2 \phi(x)\,dx = \Phi(x) - x\,\phi(x) + C_2,$

where C_0, C_1 and C_2 are constants.

Lemma 2. We have that $D_0^i = 1$,

$$D_1^i = \mu_2^{old} + \frac{\sigma_2^{old}}{1 - \Phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right)} \phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right)$$

$$D_2^i = (\mu_2^{old})^2 + (\sigma_2^{old})^2 + \frac{\mu_2^{old}\,\sigma_2^{old}}{1 - \Phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right)} \phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right) + \frac{(\sigma_2^{old})^2}{1 - \Phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right)} \phi\left(\frac{s_i - \mu_2^{old}}{\sigma_2^{old}}\right) s_i .$$

Lemma 3 We have that $X_0^i = 1$,

$$X_1^i = \mu_1^{old} + \frac{\sigma_1^{old}}{1 - \Phi\left(\frac{s_i - \mu_1^{old}}{\sigma_1^{old}}\right)} \phi\left(\frac{s_i - \mu_1^{old}}{\sigma_1^{old}}\right)$$

$$X_2^i = (\mu_1^{old})^2 + (\sigma_1^{old})^2 + \frac{\mu_1^{old}\,\sigma_1^{old}}{1 - \Phi\left(\frac{s_i - \mu_1^{old}}{\sigma_1^{old}}\right)} \phi\left(\frac{s_i - \mu_1^{old}}{\sigma_1^{old}}\right) + \frac{(\sigma_1^{old})^2}{1 - \Phi\left(\frac{s_1 - \mu_1^{old}}{\sigma_1^{old}}\right)} \phi\left(\frac{s_i - \mu_1^{old}}{\sigma_1^{old}}\right) s_i$$