Numerical example of one EM iteration over a Mixture of Gaussians

Lecturer: Roni Rosenfeld   *Scribe: Roni Rosenfeld*

1 Estimating Means of 2 Gaussians

This is a numerical mini-example of a single EM iteration as applies to the problem of estimating the mean of two Gaussians. The full derivation and explanation of the EM algorithm for this case can be found in many books (e.g. Mitchell's book, section 6.12.1).

Let $(x_1, x_2, x_3) = (2, 4, 7)$ be our three datapoints, presumed to have each been generated from one of two Gaussians.

The stdev of both Gaussians are given: $\sigma_1 = \sigma_2 = 1/\sqrt{2}$.
The prior over the two Gaussians in also given: $\lambda_1 = \lambda_2 = 0.5$
Let j be an index over the Gaussians, i an index over the data points, and k an index over the E-M iterations.

We are trying to derive values for the two means $\boldsymbol{\mu} = (\mu_1, \mu_2)$ that maximize the likelihood of the given data, i.e.:

$$\boldsymbol{\mu}_{ML} = \arg \max_{\mu_1, \mu_2} \prod_i \sum_j \lambda_j e^{-\frac{(x_i - \mu_j)^2}{2\sigma^2}}$$

Let us initialize the Gaussian means to some reasonable values (inside the data range, and integer valued, to make calculation easy): $\mu_1^{[0]} = 3, \mu_2^{[0]} = 6$.

Let $z_{i,j} = 1$ if x_i was generated by Gaussian j, and 0 otherwise. The $z_{i,j}$'s are our latent variables.

The likelihood function can be written as:

$$L(x_i | \mu^{[k]}) = \sum_j \lambda_j L(x_i | \mu = \mu_j^{[k]})$$

The E-Step is:

$$E[z_{i,j} | \mu^{[k]}] = \frac{\lambda_j L(x_i | \mu = \mu_j^{[k]})}{\sum_{j'} \lambda_{j'} L(x_i | \mu = \mu_{j'}^{[k]})}$$

and the M-step is:

$$\mu_j^{[k+1]} = \frac{\sum_i E[z_{i,j} | \mu^{[k]}] \cdot x_i}{\sum_i E[z_{i,j} | \mu^{[k]}]}$$

So here is the first iteration:

i	1	2	3
x_i	2.0	4.0	7.0
$L(x_i\|\mu=\mu_1^{[0]})$	$\frac{1}{\sqrt{\pi}}e^{-1}$	$\frac{1}{\sqrt{\pi}}e^{-1}$	$\frac{1}{\sqrt{\pi}}e^{-16}$
$L(x_i\|\mu=\mu_2^{[0]})$	$\frac{1}{\sqrt{\pi}}e^{-16}$	$\frac{1}{\sqrt{\pi}}e^{-4}$	$\frac{1}{\sqrt{\pi}}e^{-1}$
$L(x_i\|\mu^{[\mathbf{0}]})$	$\frac{1}{2\sqrt{\pi}}(e^{-1}+e^{-16})$	$\frac{1}{2\sqrt{\pi}}(e^{-1}+e^{-4})$	$\frac{1}{2\sqrt{\pi}}(e^{-16}+e^{-1})$
$E[z_{i,1}\|\mu^{[0]}]$	$\frac{e^{-1}}{e^{-1}+e^{-16}}\approx 1$	$\frac{e^{-1}}{e^{-1}+e^{-4}}\approx 0.953$	$\frac{e^{-16}}{e^{-1}+e^{-16}}\approx 0$
$E[z_{i,2}\|\mu^{[0]}]$	$\frac{e^{-16}}{e^{-1}+e^{-16}}\approx 0$	$\frac{e^{-4}}{e^{-1}+e^{-4}}\approx 0.047$	$\frac{e^{-1}}{e^{-1}+e^{-16}}\approx 1$

$$(1)$$

And therefore:

$$\mu_1^{[1]} \approx \frac{1*2.0+0.953*4.0+0*7.0}{1+0.953} \approx 2.976$$

and

$$\mu_2^{[1]} \approx \frac{0*2.0+0.047*4.0+1*7.0}{1+0.047} \approx 6.865$$