mcem

Rob McCulloch

November 19, 2019

Monte Carlo EM

Recall the basic setup for the EM algorithm.

We have a joint model,

$$f(z, x \mid \theta)$$

where,

- x is observed.
- z is latent.

Iterates of θ: $\{\theta^t\}$.

$$Q(\theta \mid \theta^t) = E(\log(f(z, x \mid \theta))$$

where E is over $Z \mid x, \theta^t$.

$$\theta^{t+1} = \underset{\theta}{\text{argmax}}\ Q(\theta \mid \theta^t)$$

The idea of the *Monte Carlo EM* is simple.

In some cases the expectation in the E step cannot be done analytically.
This is exactly what Monte Carlo is for!

So, you simply replace the $E(\log(f(z, x \mid \theta))$ with a monte carlo approximation.
This means you will have to be able to get iid draws from $Z \mid x, \theta^t$ and then compute

$$E(\log(f(z, x \mid \theta)) \approx \frac{1}{m} \sum \log(f(z_j, x \mid \theta))$$

where $z_j \sim Z \mid x, \theta^t$, *iid*.

- read 4.3.1 of "Computational Statistics" by Givens and Hoeting
- code up examples 4.5 (EM) and 4.7 (MCEM) from Givens and Hoeting
- compare your results with the results from the exact algorithm which does the expectation analytically.

You can use the data suggested at the end of example 4.7 and see if you can get something like Figure 4.2