Converting a population partameter into a random variable with small variance

Belhal Karimi

1 Convert the parameter into random variables with variance that tends to infinity

The initial model $p(y_i, \psi_i; \alpha, \theta)$ becomes $p(y_i, \psi_i, \alpha; \theta)$.

The new model can now be written as:

$$p(y_i|\psi_i, \alpha; \theta)p(\psi_i; \theta)\pi(\alpha) \tag{1}$$

where $\pi(\alpha)$ is a prior on our new random variable α.
The challenge is to see if we keep the same convergence properties in the SAEM algorithm when we operate that change.

Let's take a simple example: a logistic regression:

$$f(t) = \frac{1}{\beta + e^{-\psi_i t}} + a_i\epsilon_i \tag{2}$$

Hypothesis:

$$y_i \sim N(\frac{1}{\beta + e^{-\psi_i t}}, a_i^2) \tag{3}$$

$$log(\psi_i) \sim N(log(\psi_p op), \sigma^2) \tag{4}$$

The parameter θ that we are looking for is: $\theta = (\begin{bmatrix} \psi_{pop} \\ \beta \end{bmatrix}, \omega, a_i)$

1.1 Fixed population parameter

The model is written as

$$p(y, \psi; \theta) = p(y|\psi; \theta)p(\psi; \theta) \tag{5}$$

$$(\psi_i)^{k+1} \sim p(\psi_i|y_i, \theta^k; \beta^k) \tag{6}$$

$$\theta^{k+1}, \beta^{k+1} = max(p(y, \psi^{k+1}, \theta, \beta)) \tag{7}$$

1.2 Random population parameter

In the second algorithm, β is a random variable such as $\beta \sim N(\beta^*, \sigma_\beta)$
The model becomes:

$$p(y, \psi, \beta; \theta) = p(y|\psi, \beta; \theta)p(\psi; \theta)\pi(\beta) \tag{8}$$
$$\tag{9}$$

with $\theta = (\psi_{pop}, \omega, a_i)$

If we then look at what is happening in the SAEM algorithm, nothing changes in the simulation phase where the latent data are drawn from same the conditional $p(\psi_i|y_i)$.
Yet, the new algorithm requires to draw a new β at each iteration from the conditional $p(\beta|y_i, \beta_k, (\beta^*)_k)$.

$$(\psi_i)^{k+1} \sim p(\psi_i|y_i, \theta^k; \beta^k) + \beta^{k+1} \tag{10}$$
$$\beta^{k+1} \sim p(\beta|y_i, \beta^k; (\beta^*)^k) \tag{11}$$

As a matter of fact, if we consider that the complete data likelihood belongs to the exponential family. Thus can be written as:

$$f(y, \psi, \beta; \theta) = exp^{-\phi(\theta) + <\Phi(\theta), E(S(y,\psi))|y,\theta^k,\beta^k> + B(\psi,\beta)} \tag{12}$$

Algorithm 1. *EM Algorithm*

Initial value θ_0
$\theta \leftarrow \theta_0$
for $k \leftarrow 1$ *to* K ***do***
  $Q_k(\theta|\theta_{k-1}) \leftarrow \mathrm{E}(\log p(y, \psi; \theta)|y; \theta_{k-1})$
  $\theta_k \leftarrow {}_{\theta} Q_k(\theta|\theta_{k-1})$
end for
return θ_K

end