EP Set Up and Algorithm

We wish to split the sky map into $J = 360$ bins, regressing FUV (y_{ij}) on i100 (x_{ij}) within each bin j. To start we'll fit a constant slope β but a varying intercept a_j:

$$y_{ij} \sim N(a_j + bx_{ij}, \sigma^2) \qquad \alpha_j \sim N(\theta, \tau^2)$$

Letting $K = 6$ and using the notation $\phi = (b, \log\sigma, \theta, \log\tau)$, we have that the posterior can be written as:

$$p(\mathbf{a}, b, \theta, \log\tau, \log\sigma | y) = f(\phi)\prod_{k=1}^{6} f_k(\phi, \mathbf{a}_{(k)}) = f(\phi)\prod_{k=1}^{6} f_k(\mathbf{a}_{(k)}|\phi) f_k(\phi)$$

$$= p(\phi)\prod_{k=1}^{6} p(\mathbf{a}_{(k)}|\theta, \log\tau) p(y|\mathbf{a}_{(k)}, b, \log\sigma)$$

Then, utilizing a flat prior over ϕ, we get that for our approximation, $g_0(\phi)$ is a constant, so that our approximation then becomes:

$$g(\phi) = \prod_{k=1}^{6} g_k(\phi) = \prod_{k=1}^{6} N(\phi|\mu_k, \Sigma_k)$$

Our initial approximation then becomes, for all k, $\mu_k = \mathbf{0}$ and $\Sigma_k = 10^2\mathbf{I}/6$ (note that the dimension of μ_k is equal to the dimension of ϕ, which is 4). The algorithm then proceeds as follows:

(a) Compute the cavity distribution $g_{-k}(\phi) = \frac{g(\phi)}{g_k(\phi)} = \prod_{k' \neq k} N(\phi|\mu_{k'}, \Sigma_{k'}) = N(\mu_{-k}, \Sigma_{-k})$ where:

$$\Sigma_{-k}^{-1} = \sum_{k' \neq k} \Sigma_{k'}^{-1}$$

$$\Sigma_{-k}^{-1}\mu_{-k} = \sum_{k' \neq k} \Sigma_{k'}^{-1}\mu_{k'}$$

(b.i) Approximate[1] the tilted distribution $g_{\backslash k}(\mathbf{a}_{(k)}, \phi) \approx N(\mu^*_{\backslash k}, \Sigma^*_{\backslash k})$ where:

$$g_{\backslash k}(\mathbf{a}_{(k)}) = g_{-k}(\phi) p(\mathbf{a}_{(k)}, y_{(k)}|\phi) = N(\phi|\mu_{-k}, \Sigma_{-k}) \prod_{j \in (k)} N(a_j|\theta, \tau) p(y_j|a_j, b, \sigma^2)$$

(b.ii) Approximate[2] the marginal tilted distribution $g_{\backslash k}(\phi) = \int g_{\backslash k}(\mathbf{a}_{(k)}, \phi) d\mathbf{a}_{(\mathbf{k})} \approx N(\mu_{\backslash k}, \Sigma_{\backslash k})$.

(c) Update[3] the site distribution $g_k^{new} = \frac{g_{\backslash k}(\phi)}{g_{-k}(\phi)} = N(\mu_k^{new}, \Sigma_k^{new})$ where:

$$(\Sigma_k^{new})^{-1} = (\Sigma_{\backslash k})^{-1} - \Sigma_{-k}^{-1}$$

$$(\Sigma_k^{new})^{-1}\mu_k^{new} = (\Sigma_{\backslash k})^{-1}\mu_{\backslash k} - \Sigma_{-k}^{-1}\mu_{-k}$$

(d) Update $g(\phi)$ in serial or parallel.

[1] This could possibly be done in Stan, though unsure how to approximate 64x64-dimensional covariance matrix.
[2] This could possibly be done in Stan, though unsure how to approximate 4x4-dimensional covariance matrix.
[3] This may result in non-positive-definite inverse matrices.

Complications

(b.i) This could be run in Stan as mentioned in the paper. Approximating $\mu^*_{\setminus k}$ would be easy as we would just take the means of each parameters' simulations. However, I'm unsure if we could use sample covariances of the simulations to approximate $\Sigma^*_{\setminus k}$.

(b.ii) Upon running the MCMC simulation in (b.i), we could once again simply approximate $\mu_{\setminus k}$ by taking the means of only the four ϕ parameters' simulations. (This actually obviates the need to calculate $\mu^*_{\setminus k}$, the mean of the tilted distribution itself.) However, once again, I'm unsure if we could use sample covariances of the simulations to approximate $\Sigma_{\setminus k}$.

(c) When I programmed EP for the bioassay example, after a few iterations I was getting non-positive-definite inverse matrices for both steps. I'm guessing this will end up being an issue here as well.