RANDOM EFFECTS AS A WEIGHTED AVERAGE

The model:

$$y_{ig} = X_{ig}\beta + \nu_g + \varepsilon_{ig}\,, i = 1, \ldots, n,\ g = 1, \ldots, M \tag{1}$$

$$\nu_g \mid X \sim N(0, \tau^2\,, \qquad \varepsilon_{ig} \sim N(0, \sigma^2)\,. \tag{2}$$

In addition we assume that conditional on X (the full $nm \times k$ matrix of X_{ig} values) the ε and ν vectors are jointly normal with diagonal covariance matrix.

As is explained in earlier notes, in this model the GLS estimate of β with known σ^2 and τ^2 takes the form of a weighted average of the between and within regression estimates, where the within estimate is the fixed effects estimator (or equivalently the estimate based on the deviations from group means of y and X) and the between estimate is the estimate based on group-mean data (i.e. the M data points generated by taking means across i for each g). The formula derived in the earlier notes was

$$\hat{\beta}_{GLS} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (\sigma^{-2}\tilde{X}'\tilde{X} + \delta^2\bar{X}'\bar{X})^{-1}(\sigma^{-2}\tilde{X}'\tilde{X}\hat{\beta}_w + \delta^2\bar{X}'\bar{X}\hat{\beta}_b)\,. \tag{3}$$

The earlier notes did not give an explicit formula for δ as a function of σ^2 and τ^2 or for the likelihood function as a function of the between and within residual sums of squares. These notes fill in these gaps.

Since the residual covariance matrix has the form

$$\Omega = I_M \otimes \tilde{\Omega} = I_M \otimes (\sigma^2 I_n + \tau^2\mathbf{1})\,, \tag{4}$$

the GLS estimator can be described as OLS on transformed data, where y and X are pre-multiplied by W

$$W = I_m \otimes \tilde{W} = I_m \otimes \left(\sigma^{-1}(I - \tfrac{1}{n}\mathbf{1}) + \tfrac{\delta}{n}\mathbf{1}\right) \tag{5}$$

$$\tilde{W}^2 = \tilde{\Omega}^{-1} = \sigma^{-2}(I - \tfrac{1}{n}\mathbf{1}) + \tfrac{\delta^2}{n}\mathbf{1} \tag{6}$$

$$\therefore \quad \tilde{W}^2\tilde{\Omega} = \left(\sigma^{-2}(I - \tfrac{1}{n}\mathbf{1}) + \tfrac{\delta^2}{n}\mathbf{1}\right)(\sigma^2 I + \tau^2\mathbf{1}) = I \tag{7}$$

$$\therefore -\frac{1}{n} + \frac{\delta^2\sigma^2}{n} + \tau^2\delta^2 = 0\,. \tag{8}$$

From this we can conclude that $\delta^2 = 1/(\sigma^2 + n\tau^2)$.

The $\tilde{\Omega}$ matrix has $n-1$ eigenvalues of σ^2 (corresponding to eigenvectors that sum to one) and 1 eigenvalue of $\tau^2 n + \sigma^2$). The full Ω matrix therefore has $M(n-1)$ eigenvalues of 1 and M of $\tau^2 n + \sigma^2$. The log likelihood function can be written as

$$-\frac{Mn}{2}\log(2\pi) - \frac{M}{2}\log(\tau^2 n + \sigma^2) - \frac{M(n-1)}{2}\log(\sigma^2) - \frac{\tilde{u}'\tilde{u}}{2\sigma^2} - \frac{\bar{u}'\bar{u}}{2(\tau^2 + \sigma^2/n)}\,, \tag{9}$$

where $\tilde{u}$ are the residuals from the within regression and $\bar{u}$ are the residuals from the between regression.

The conditional posterior distribution on β given σ^2 and τ^2 is $N(\hat{\beta}_{GLS}, (X'\Omega^{-1}X)^{-1})$, where

$$\hat{\beta}_{GLS} = (\sigma^{-2}\tilde{X}'\tilde{X} + (\sigma^2 + \tau^2 n)^{-1}\bar{X}'\bar{X})^{-1}(\sigma^{-2}\tilde{X}'\tilde{X}\hat{\beta}_w + (\sigma^2 + \tau^2 n)^{-1}\bar{X}'\bar{X}\hat{\beta}_b) \quad (10)$$

$$X'\Omega^{-1}X = \sigma^{-2}\tilde{X}'\tilde{X} + (\sigma^2 + \tau^2 n)^{-1}\bar{X}'\bar{X} \quad (11)$$

Conditional on β,

$$\sigma^{-2} \sim \text{Gamma}(\tfrac{1}{2}M(n-1) - 1, \tfrac{1}{2}\tilde{u}'\tilde{u}) \quad (12)$$

$$(\tau^2 + \sigma^2)^{-1} \sim \text{Gamma}(\tfrac{1}{2}M - 1, \tfrac{1}{2}\bar{u}'\bar{u}) \,, \quad (13)$$

with these two random variables independent. Of course this means that σ^2 and τ^2 themselves are dependent.

These results suggest a particularly simple way to sample from the posterior on σ^2, τ^2 and β. Assuming we have some initial estimates — for example by applying OLS to estimate β and estimating $\tau^2 + \sigma^2/n$ as $\bar{u}'\bar{u}/M$, σ^2 as $\tilde{u}'\tilde{u}/(Mn)$:

(1) Draw β from its normal conditional posterior above, and use the draw to construct $\tilde{u}$ and $\bar{u}$.
(2) Draw σ^{-2} and $\tau^2 + \sigma^2/n$ from their conditional posteriors above.
(3) Return to 1.

With this scheme, $\bar{X}'\bar{X}$, $\tilde{X}'\tilde{X}$, β_w, and β_b can be computed once, before the iterations start. The MCMC sampled values are constructed by reweighting these objects.

Important note about singularity: There may be variables (like a time trend) that show no variation in group means, as well as variables that are constant within states. Thus either $\bar{X}'\bar{X}$ or $\tilde{X}'\tilde{X}$ or both could be singular, making $\hat{\beta}_w$ and/or $\hat{\beta}_b$ undefined. But the expression (10) can also be written in a form that is insensitive to these singularities:

$$\hat{\beta}_{GLS} = (\sigma^{-2}\tilde{X}'\tilde{X} + (\sigma^2 + \tau^2 n)^{-1}\bar{X}'\bar{X})^{-1}(\sigma^{-2}\tilde{X}'\tilde{y} + (\sigma^2 + \tau^2 n)^{-1}\bar{X}'\bar{y}) \,. \quad (14)$$

It is this form that should be used for actual computation.

Note also that in forming $\tilde{u}$ and $\bar{u}$, one is using the same β value for both, and this β is always well defined. The fact that, e.g., a trend variable shows no variation in state means does not create a problem when forming the between residuals $\bar{u}$. The coefficient on trend just becomes a contribution to the constant term.