Bayesian Parameter Averaging for GLMs

W. Penny

Wellcome Trust Centre for Neuroimaging,
University College, London WC1N 3BG, UK.

April 30, 2013

Consider fitting GLMs to data from N subjects. We assume a Gaussian prior over GLM coefficients with mean μ_0 and precision Λ_0. For subject i we have data y_i, design matrix X_i and data precisions Q_i. What is the posterior over GLM coefficients given data from all subjects ? We can work this out by fitting a big GLM with design matrix

$$X = \begin{bmatrix} X_1 \\ X_2 \\ ... \\ X_N \end{bmatrix} \tag{1}$$

data

$$y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_N \end{bmatrix} \tag{2}$$

and precision matrix

$$Q = \begin{bmatrix} Q_1 & 0 & 0 & 0 \\ 0 & Q_2 & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & Q_N \end{bmatrix} \tag{3}$$

The posterior over GLM coefficients then has mean and precision given by

$$\Lambda = X^T Q X + \Lambda_0 \tag{4}$$

$$\mu = \Lambda^{-1}[X^T Q y + \Lambda_0 \mu_0] \tag{5}$$

Due to the block diagonal structure of Q this can be rewritten as

$$\Lambda = \left(\sum_{i=1}^{N} X_i^T Q_i X_i \right) + \Lambda_0 \tag{6}$$

$$\mu = \Lambda^{-1} \left[\left(\sum_{i=1}^{N} X_i^T Q_i y_i \right) + \Lambda_0 \mu_0 \right] \tag{7}$$

Individual Posteriors

Now consider separate estimation of the GLMs with the same prior for each. The posterior for the ith GLM is

$$\begin{aligned} \Lambda_i &= X_i^T Q_i X_i + \Lambda_0 && (8) \\ \mu_i &= \Lambda_i^{-1}[X_i^T Q_i y_i + \Lambda_0 \mu_0] && (9) \end{aligned}$$

Equation 9 can be re-arranged to give

$$\Lambda_i \mu_i = X_i^T Q_i y_i + \Lambda_0 \mu_0 \tag{10}$$

or

$$X_i^T Q_i y_i = \Lambda_i \mu_i - \Lambda_0 \mu_0 \tag{11}$$

So we can write

$$\sum_{i=1}^{N} X_i^T Q_i y_i = \left(\sum_{i=1}^{N} \Lambda_i \mu_i \right) - N \Lambda_0 \mu_0 \tag{12}$$

Similarly equation 8 can be re-arranged to give

$$X_i^T Q_i X_i = \Lambda_i - \Lambda_0 \tag{13}$$

so that

$$\sum_{i=1}^{N} X_i^T Q_i X_i = \left(\sum_{i=1}^{N} \Lambda_i \right) - N \Lambda_0 \tag{14}$$

Combining Individual Posteriors

Substituting 12 into equation 7 gives

$$\mu = \Lambda^{-1} \left[\left(\sum_{i=1}^{N} \Lambda_i \mu_i \right) - (N-1) \Lambda_0 \mu_0 \right] \tag{15}$$

Substituting 14 into equation 6 gives

$$\Lambda = \left(\sum_{i=1}^{N} \Lambda_i \right) - (N-1) \Lambda_0 \tag{16}$$

This generalises the approach in [1].

References

[1] C Kasess, K Stephan, A Weissenbacher, L Pezawas, E Moser, and C Windischberger. Multi-subject analyses with dynamic causal modeling. *Neuroimage*, 49(4):3065–3074, Feb 2010.