Additional Material for Technical Milestone Report

Joseph Feest

January 2024

1 Derivations of Update Rules and Associated Moments of Variables

For our specific algorithm, we note the independence and conditionality within the variational distribution and can refine our update rule to:

$$\begin{aligned}\log(q^*_{\mathbf{Z_j}}(\mathbf{Z}_j)) = \mathbb{E}_{\not\subseteq\lambda}[\log(p(Y|\mathbf{\Omega})) + \log(p(\mathbf{\Omega}|\boldsymbol{\delta},\tau))+ \\ \log(p(\delta|\rho)) + \log(p(\tau)) + \log(p(\rho))] + \text{const},\end{aligned} \tag{1}$$

where again $\mathbf{Z} = \{\mathbf{\Omega}, \boldsymbol{\delta}, \rho, \tau\}$. For the calculation of each variable's update step, we can now omit any distributions that do not include said variable.

1.1 Deriving the update rule for $q^*_\tau(\tau)$

Using equation 1 as a reference, and absorbing all terms that don't involve τ into the constant, we start with:

$$\log(q^*_\tau(\tau)) = \mathbb{E}_\tau[\log(p(\mathbf{\Omega}|\delta,\tau) + \log(p(\tau)] + \text{const}. \tag{2}$$

Remembering that τ is distributed as Gamma(a,b), we can assert $\log(p(\tau))$ as:

$$\log(p(\tau)) = \log(\frac{b^a}{\Gamma(a)}\tau^{a-1}e^{-b\tau}) = (a-1)\log(\tau) - b\tau + \text{const}. \tag{3}$$

When considering $p(\mathbf{\Omega}|\delta,\tau)$, the diagonal entries of $\mathbf{\Omega}$ have to be considered differently to the non-diagonal entries. Furthermore, $\mathbf{\Omega}$ is a symmetric matrix and as such we only need to consider half of the non-diagonal entries. Therefore, summations are made with all entries $\mathbf{\Omega}_{ij}$ where $i < j$ (note that summing all entries $\mathbf{\Omega}_{ij}$ where $i < j$ would give the same result.) The diagonal entries $\mathbf{\Omega}_{ii} \sim \exp(\lambda/2)$ are independent of τ and are, therefore, eventually absorbed by the constant. The equation is:

$$\log(p(\mathbf{\Omega}|\delta,\tau) = \sum_i \frac{\lambda}{2} + \sum_{i<j}\Bigg[\delta_{ij}\Bigg\{-\frac{1}{2}\log(\frac{v_1^2}{\tau}) - \frac{\mathbf{\Omega}_{ij}^2\tau}{2v_1^2}\Bigg\} + (1-\delta_{ij})\Bigg\{-\frac{1}{2}\log(\frac{v_0^2}{\tau}) - \frac{\mathbf{\Omega}_{ij}^2\tau}{2v_0^2}\Bigg\}\Bigg], \tag{4}$$

which, when simplified, becomes:

$$\log(p(\mathbf{\Omega}|\delta,\tau) = \frac{\lambda P}{2} - \delta_{ij}^{(1)}\log(\nu_1) - (1-\delta_{ij}^{(1)})\log(\nu_0) - \frac{\mathbf{\Omega}_{ij}^2\tau}{2}E_{ij}^{(1)} + \frac{P(P-1)}{4}\log(\tau). \tag{5}$$

Applying equations 4 and 5 to equation 2:

$$\log(q^*_\tau(\tau)) = \mathbb{E}_{\tau'}[(\frac{P(P-1)}{4} + a - 1)\log(\tau) - (\frac{\mathbf{\Omega}_{ij}^2\tau}{2}E_{ij}^{(1)} + b)\tau]. \tag{6}$$

And therefore the update rules become:

$$q^*_\tau(\tau) \sim \text{Gamma}(\alpha_\tau, \beta_\tau), \quad \alpha_\tau = \frac{P(P-1)}{4} + a, \quad \beta_\tau = \frac{\mathbf{\Omega}_{ij}}{2}E_{ij}^{(1)} + b. \tag{7}$$

1.2 Deriving the update rule for $q^*_\delta(\delta)$

Using equation 1 as a reference, and absorbing all terms that don't involve δ_{ij} into the constant, we start with:

$$\log(q^*_\delta(\delta)) = \mathbb{E}_{\delta'}[\log(p(\boldsymbol{\Omega}|\delta_{ij},\tau)) + \log(p(\delta|\rho))] + \text{const.} \tag{8}$$

$\log(p(\boldsymbol{\Omega}|\delta_{ij},\tau))$ has already been found in equation 5 and will be re-used here. For $\log(p(\delta|\rho))$, which is $\sim \text{Bern}(\rho)$, we have the following result:

$$\log(p(\boldsymbol{\Omega}|\delta_{ij},\tau)) = \sum_{i<j} \delta_{ij}\log(\rho) + \sum_{i<j}(1-\delta_{ij})\log(1-\rho). \tag{9}$$

Therefore:

$$\begin{aligned} \log(q^*_{\delta_{ij}}(\delta_{ij}) = \mathbb{E}_{\delta'_{ij}}[\sum_{i<j} -\delta_{ij}\{\log(\nu_1)\} - \sum_{i<j}(1-\delta_{ij})\log(\nu_0) \\ -\frac{\boldsymbol{\Omega}^2_{ij}\tau}{2}E^{(1)}_{ij} - \sum_{i<j}\delta\log(\rho) - \sum_{i<j}(1-\delta_{ij})\log(1-\rho)] \end{aligned} \tag{10}$$

1.3 Deriving the update rule for $q^*_\rho(\rho)$

For the update of ρ, again we start by taking all terms involving ρ from equation 1. All terms not involving ρ are absorbed into the constant.

$$\log(q^*_\rho(\rho)) = \mathbb{E}_{\rho'}[\log(p(\delta|\rho)) + \log(p(\rho))] + \text{const.} \tag{11}$$

Taking the natural log of each term gives us:

$$\log(p(\delta|\rho)) = \sum_{i<j}\delta_{ij}\log(\rho) + \sum_{i<j}(1-\delta_{ij})\log(1-\rho) \tag{12}$$

$$\log(p(\rho)) = (a_\rho - 1)\log(\rho) + (b_\rho - 1)\log(1-\rho) \tag{13}$$

Combining equations 12 and 13:

$$\log(q^*_\rho(\rho)) = \mathbb{E}_{\rho'}[(\sum_{i<j}\delta^{(1)}_{ij} + a_\rho - 1)\log(\rho) + (\sum_{i<j}(1-\delta_{ij}) + b_\rho - 1)\log(1-\rho)] + \text{const.} \tag{14}$$

The expectation term on the right hand side of equation 14 is equivalent to the expectation of a natural logarithm of a Beta distribution. Therefore we can define $q^*_\rho(\rho)$ and the update rules for α_ρ and β_ρ as:

$$q^*_\rho(\rho) \sim \text{Beta}(\alpha_\rho, \beta_\rho), \quad \alpha_\rho = \sum_{i<j}\delta^{(1)}_{ij} + a_\rho, \quad \beta_\rho = \sum_{i<j}(1-\delta^{(1)}_{ij}) + b_\rho. \tag{15}$$

2 Further Decomposition of the ELBO

The overall ELBO can be calculated by summing the ELBOs of each individual variable. The ELBO is defined as:

$$\mathcal{L}(q) = \mathcal{L}_y(\mathbf{Y} \mid \boldsymbol{\Omega}, \boldsymbol{\delta}, \tau) + \mathcal{L}_{\boldsymbol{\Omega}}(\boldsymbol{\Omega} \mid \boldsymbol{\delta}, \tau) + \sum_{i<j}\mathcal{L}_\delta(\delta_{ij} \mid \rho) + \mathcal{L}_\tau(\tau) + \mathcal{L}_\rho(\rho) \tag{16}$$

where

$$\mathcal{L}_y(\mathbf{Y} \mid \boldsymbol{\Omega}) \quad = \quad \mathbb{E}_q\Big[\log p(\mathbf{Y} \mid \boldsymbol{\Omega})\Big] = \frac{N}{2}\log(|\boldsymbol{\Omega}|) - \frac{1}{2}\text{tr}(\boldsymbol{Y}^{\boldsymbol{T}}\boldsymbol{Y}\boldsymbol{\Omega}),$$

$$
\begin{aligned}
\mathcal{L}_{\boldsymbol{\Omega}}(\boldsymbol{\Omega} \mid \boldsymbol{\delta}, \tau) &= \mathbb{E}_q\left[\sum_{i=1}^{P} \log p(\boldsymbol{\Omega}_{ii}) + \sum_{i<j} \log p(\boldsymbol{\Omega}_{ij} \mid \delta_{ij}, \tau)\right] \\
&= -\frac{\lambda}{2}\sum_{i=1}^{P} \boldsymbol{\Omega}_{ii} - \sum_{i<j}\left[\delta_{ij}^{(1)} \log \nu_1 + (1-\delta_{ij}^{(1)}) \log \nu_0\right] - \frac{\tau^{(1)}}{2}\sum_{i<j} \boldsymbol{\Omega}_{ij}^2 E_{ij}^{(1)} + \frac{P(P-1)}{4}\Big(\log \tau\Big)^{(1)}, \\
\mathcal{L}_{\delta}(\delta_{ij} \mid \rho) &= \mathbb{E}_q\Big[\log p(\delta_{ij} \mid \rho)\Big] - \mathbb{E}_q\Big[\log q(\delta_{ij})\Big], \\
&= \delta_{ij}^{(1)}\Big(\log \rho\Big)^{(1)} + (1-\delta_{ij}^{(1)})\Big(\log(1-\rho)\Big)^{(1)} - \delta_{ij}^{(1)}\Big(\log \delta_{ij}\Big)^{(1)} - (1-\delta_{ij}^{(1)})\Big(\log(1-\delta_{ij})\Big)^{(1)}, \\
\mathcal{L}_{\tau}(\tau) &= \mathbb{E}_q\Big[\log p(\tau)\Big] - \mathbb{E}_q\Big[\log q(\tau)\Big] \\
&= (a-\alpha_\tau)(\log \tau)^{(1)} - (b-\beta_\tau)\tau^{(1)} - \alpha_\tau \log \beta_\tau + \log \Gamma(\alpha_\tau), \\
\mathcal{L}_{\rho}(\rho) &= \mathbb{E}_q\Big[\log p(\rho)\Big] - \mathbb{E}_q\Big[\log q(\rho)\Big] \\
&= (a_\rho - 1)\Big(\log \rho\Big)^{(1)} + (b_\rho - 1)\Big(\log(1-\rho)\Big)^{(1)} \\
&\quad + \log \mathcal{B}(\alpha_\rho, \beta_\rho) - (\alpha_\rho - 1)\Big(\log \rho\Big)^{(1)} - (\beta_\rho - 1)\Big(\log(1-\rho)\Big)^{(1)},
\end{aligned}
$$

where $\cdot^{(m)}$ denotes the mth moment of the variable, and

$$
E_{ij}^{(1)} := \frac{\delta_{ij}^{(1)}}{\nu_1^2} + \frac{1-\delta_{ij}^{(1)}}{\nu_0^2}.
$$