first-order conditions for block coordinate ascent lead to the updates

$$\mu_\zeta \leftarrow \left(\Omega_0^{-1} + H\omega\Upsilon\right)^{-1}\left(\Omega_0^{-1}\beta_0 + \omega\Upsilon\sum_{h=1}^{H}\mu_h\right), \tag{39}$$

$$\Sigma_\zeta \leftarrow \left(\Omega_0^{-1} + H\omega\Upsilon\right)^{-1}. \tag{40}$$

By inspection, $\Sigma_\zeta \succeq 0$, so this constraint need not be explicitly enforced. Note the similarity to conjugate posterior updating: on the precision scale, Σ_ζ is the sum of the prior precision matrix Ω_0^{-1} and H copies of the variational posterior mean $\omega\Upsilon$ for Ω^{-1}. Similarly, μ_ζ is a precision-weighted convex combination of the prior vector β_0 and the empirical average of the variational posterior means $\mu_{1:H}$ for $\beta_{1:H}$.

The updates for Υ and ω are similarly straightforward to derive; we obtain

$$\omega \leftarrow \nu + H, \tag{41}$$

$$\Upsilon \leftarrow \left(S^{-1} + \sum_{h=1}^{H}\left(\Sigma_h + (\mu_\zeta - \mu_h)(\mu_\zeta - \mu_h)^\top\right) + H\Sigma_\zeta\right)^{-1}. \tag{42}$$

Notice that the solution (41) for ω involves only the constants ν and H. We compute ω once in advance, leaving it unchanged during the variational optimization.

B An application of the delta method

Let $f(v)$ be a function from $\mathbb{R}^K$ to $\mathbb{R}$. According to the multivariate delta method for moments (Bickel and Doksum 2007),

$$\mathbb{E}f(V) \approx f(\mathbb{E}V) + \frac{1}{2}\operatorname{tr}\left[\left(\frac{\partial f(\mathbb{E}V)}{\partial v \partial v^\top}\right)\operatorname{Cov}(V)\right]. \tag{43}$$

Consider the case

$$f(v) = \log\left(1^\top \exp(xv)\right), \tag{44}$$