Derivative of the posterior of the gestalt model

1 Rules of differentiation

Assuming that y and a are vectors and M is a symmetric matrix of appropriate dimension, and f is a scalar function, and s is a scalar variable.

$$\frac{\partial}{\partial y} y^T M y = 2My \tag{1}$$

$$\frac{\partial}{\partial y} a^T y = a \tag{2}$$

$$\frac{\partial}{\partial M} y^T M^{-1} y = -M^{-1} y y^T M^{-1} \tag{3}$$

$$\frac{\partial}{\partial M} \log\det M = M^{-1} \tag{4}$$

$$\frac{\partial}{\partial s} f(M(s)) = \mathrm{Tr}\left[\frac{\partial f}{\partial M}\frac{\partial M}{\partial s}\right] \tag{5}$$

$$\frac{\partial}{\partial s} sM = M \tag{6}$$

2 Form of the full posterior

$$p(v,g,z \mid x) = p(x \mid v,g,z)p(v \mid g)p(g)p(z)\frac{1}{p(x)} \sim p(x \mid v,z)p(v \mid g)p(g)p(z) \tag{7}$$

so the log-posterior will be the following, up to an additive constant, using Gamma priors over g and z defined by shape and scale parameters:

$$\log p(v,g,z \mid x) \sim \log p(x \mid v,z) + \log p(v \mid g) + \log p(g) + \log p(z) = \\ = \log \mathcal{N}(x; zAv, \sigma_x I) + \log \mathcal{N}(v; 0, C_v) + \log \mathrm{Gam}(g; sh_g, sc_g) + \log \mathrm{Gam}(z; sh_z, sc_z) \tag{8}$$

logarithms of the used pdfs look as follows:

$$\log \mathcal{N}(y; \mu, C) = -\frac{1}{2}\left[\log(2\pi) + \log\det(C) + (y-\mu)^T C^{-1}(y-\mu)\right] \quad (9)$$

$$\log \mathrm{Gam}(y; sh, sc) = \log(1) - \log(\Gamma(sh)) - sh\log(sc) + (sh-1)\log(y) - \frac{y}{sc} \quad (10)$$

so discarding all terms that are constant w.r.t. all three variables, the log-posterior is composed as follows:

$$\begin{aligned}\log p(v, g, z \mid x) \sim -\frac{1}{2\sigma_x}(x - zAv)^T(x - zAv) - \\ -\frac{1}{2}\left[\log\left(\det\left(C_v\right)\right) + v^T C_v^{-1} v\right] + \\ + \sum_{j=1}^{K}\left[(sh_g - 1)\log(g_j) - \frac{g_j}{sc_g}\right] + (sh_z - 1)\log(z) - \frac{z}{sc_z}\end{aligned} \quad (11)$$

expanding the quadratic form in the first term

$$\begin{aligned}(x - zAv)^T(x - zAv) &= x^T x - zv^T A^T x + z^2 v^T A^T Av - zx^T Av = \\ &= x^T x - 2zx^T Av + z^2 v^T A^T Av\end{aligned} \quad (12)$$

as $zx^T Av$ is a scalar, thus equal to its transpose. Discarding the term not dependent on any variables of the posterior we get

$$\begin{aligned}\log p(v, g, z \mid x) \sim -\frac{z}{2\sigma_x}\left(zv^T A^T Av - 2x^T Av\right) - \\ -\frac{1}{2}\left[\log\left(\det\left(C_v\right)\right) + v^T C_v^{-1} v\right] + \\ + \sum_{j=1}^{K}\left[(sh_g - 1)\log(g_j) - \frac{g_j}{sc_g}\right] + \\ + (sh_z - 1)\log(z) - \frac{z}{sc_z}\end{aligned} \quad (13)$$

3 Derivative in v

$$\log p(v, g, z \mid x) \sim -\frac{1}{2}\left[\frac{z^2}{\sigma_x} v^T A^T Av - \frac{2z}{\sigma_x} x^T Av + v^T C_v^{-1} v\right] + f_1(g, z) \quad (14)$$

lumping the two quadratic forms together

$$\log p(v, g, z \mid x) \sim \frac{z}{\sigma_x} x^T Av - \frac{1}{2} v^T \left[\frac{z^2}{\sigma_x} A^T A + C_v^{-1}\right] v + f_1(g, z) \quad (15)$$

Taking the derivative using Eq. 1 and 2 we get

$$\frac{\partial}{\partial v} \log p(v, g, z \mid x) = \frac{z}{\sigma_x} A^T x - \left[\frac{z^2}{\sigma_x} A^T A + C_v^{-1} \right] v \tag{16}$$

4 Derivative in g

$$\begin{aligned} \log p(v, g, z \mid x) \sim & -\frac{1}{2} \left[\log\det\left(C_v\right) + v^T C_v^{-1} v \right] + \\ & + \sum_{j=1}^{K} \left[(sh_g - 1) \log(g_j) - \frac{g_j}{sc_g} \right] + f_2(v, z) \end{aligned} \tag{17}$$

Taking the derivative w.r.t. a single g_i using Eq. 5 we get

$$\begin{aligned} \frac{\partial}{\partial g_i} \log p(v, g, z \mid x) = & -\frac{1}{2} \mathrm{Tr} \left[\frac{\partial}{\partial C_v} \left[\log\det\left(C_v\right) + v^T C_v^{-1} v \right] \frac{\partial C_v}{\partial g_i} \right] + \\ & + \frac{\partial}{\partial g_i} \left[(sh_g - 1) \log(g_i) - \frac{g_i}{sc_g} \right] \end{aligned} \tag{18}$$

using Eq. 4, 3 and 6 we arrive to

$$\frac{\partial}{\partial g_i} \log p(v, g, z \mid x) = -\frac{1}{2} \mathrm{Tr} \left[\left[C_v^{-1} - C_v^{-1} v v^T C_v^{-1} \right] C_i \right] + \frac{sh_g - 1}{g_i} - \frac{1}{sc_g} \tag{19}$$

5 Derivative in z

$$\log p(v, g, z \mid x) \sim -\frac{z}{2\sigma_x} \left(z v^T A^T A v - 2 x^T A v \right) + (sh_z - 1) \log(z) - \frac{z}{sc_z} + f_3(g, v) \tag{20}$$

$$\frac{\partial}{\partial z} \log p(v, g, z \mid x) = \frac{1}{\sigma_x} \left[x^T A v - z v^T A^T A v \right] + \frac{sh_z - 1}{z} - \frac{1}{sc_z} \tag{21}$$