Model selection for the gestalt model

1 The likelihood function

In order to evaluate the log-likelihood for cross validation steps, we have to use an approximation.

$$\begin{aligned} p(X \mid C_{1..k}) &= \prod_{n=1}^{N} \iiint_{-\infty}^{\infty} p(x_n \mid v_n, z_n) p(v_n \mid g_n) p(g_n) p(z_n) \mathrm{d}v_n \mathrm{d}g_n \mathrm{d}z_n = \\ &= \prod_{n=1}^{N} \iint_{-\infty}^{\infty} p(g_n) p(z_n) \int_{-\infty}^{\infty} p(x_n \mid v_n, z_n) p(v_n \mid g_n) \mathrm{d}v_n \mathrm{d}g_n \mathrm{d}z_n \end{aligned} \tag{1}$$

the product of the two conditionals can be rewritten as a Gaussian over v times another Gaussian

$$p(x \mid v, z) p(v \mid g) = \frac{1}{\sqrt{\det(A^T A)}} \frac{1}{z^{D_v}} \mathcal{N}(\frac{1}{z} A^+ x; 0, \frac{\sigma_x}{z^2} (A^T A)^{-1} + C_v) \mathcal{N}(v; \mu_c, C_c) \tag{2}$$

where v only appears in the second Gaussian, so integrating this formula will set that term to 1 and leave everything else as it is.

$$p(X \mid C_{1..k}) = \det(A^T A)^{-\frac{N}{2}} \prod_{n=1}^{N} \int_{0}^{\infty} p(z) \frac{1}{z^{D_v}} \int_{0}^{\infty} p(g) \mathcal{N}(\frac{1}{z} A^+ x; 0, \frac{\sigma_x}{z^2} (A^T A)^{-1} + C_v) \mathrm{d}g \mathrm{d}z \tag{3}$$

2 Approximation with sampling

The integrals may be approximated by averaging over samples from $p(g)$ and $p(z)$, leaving out terms constant in g, z and x

$$p(X \mid C_{1..k}) \sim \prod_{n=1}^{N} \frac{1}{L^2} \sum_{l_1=1}^{L} \frac{1}{z_{l_1}^{D_v}} \sum_{l_2=1}^{L} \mathcal{N}(\frac{1}{z_{l_1}} A^+ x_n; 0, \frac{\sigma_x}{z_{l_1}^2} (A^T A)^{-1} + C_v^{l_2}) \tag{4}$$

as z and g are independent, we can choose a large enough L and merge the sums, and bring the division by L outside of the product

$$p(X \mid C_{1..k}) \sim \frac{1}{L^N} \prod_{n=1}^{N} \sum_{l=1}^{L} \frac{1}{z_{l_1}^{D_v}} \mathcal{N}(\frac{1}{z_l} A^+ x_n; 0, \frac{\sigma_x}{z_l^2} (A^T A)^{-1} + C_v^l) \tag{5}$$

taking the logarithm we get

$$\log p(X \mid C_{1..k}) \sim -N \log(L) \sum_{n=1}^{N} \log \left[\sum_{l=1}^{L} \frac{1}{z_{l_1}^{D_v}} \mathcal{N}(\frac{1}{z_l} A^+ x_n; 0, \frac{\sigma_x}{z_l^2} (A^T A)^{-1} + C_v^l) \right] \tag{6}$$