Derivations and Equations

1 Deriving the bound

The standard variational formulation is

$$\begin{aligned}
\log p(y) &= \int q(z) \log \frac{p(y|z)p(z)}{q(z)} \frac{q(z)}{p(z|y)} dz \\
&= \underbrace{\mathbb{E}_{q(z)}\big[\log p(y|z)p(z) - \log q(z)\big]}_{\mathcal{L}_1} + \mathrm{KL}\big[q(z) \| p(z|y)\big] \qquad (1)
\end{aligned}$$

The GP-LVM-CMF variational joint in the augmented latent space is:

$$q(z, f, u, X) = \prod_{i=1}^{N} \{q(z_i|f(X_i))q(f(X_i)|u)\} q(u)q(X) \qquad (2)$$

which induces an intractable variational marginal

$$q(z) = \int \prod_{i=1}^{N} \{q(z_i|f(X_i))q(f(X_i)|u)\} q(u)q(X) df du dX. \qquad (3)$$

However we can apply the variational formulation again, noting that $q(z) = \frac{q(z|f,u,X)q(f,u,X)}{q(f,u,X|z)}$, and integrating 1 wrt the auxiliary variables, to obtain a further lower bound:

$$\begin{aligned}
\log p(y) &= \int q(z|f, u, X)q(f, u, X) \log \frac{p(y|z)p(z)r(f, u, X|z)}{q(z|f, u, X)q(f, u, X)} \frac{q(f, u, X|z)}{r(f, u, X|z)} \frac{q(z)}{p(z|y)} df du dX dz \\
&= \mathbb{E}_{q(z,f,u,X)}\big[\log p(y|z)p(z)r(f, u, X|z) - \log q(z|f, X)q(f, u, X)\big] \qquad (4) \\
&\quad + \mathbb{E}_q(z)\big[\mathrm{KL}[q(f, u, X|z) \| r(f, u, X|z)]\big] + \mathrm{KL}\big[q(z) \| p(z|y)\big] \qquad (5)
\end{aligned}$$

in which the new auxiliary lower bound, $\mathcal{L}_{aux}$ is given by the expression 4, and where we have introduced the auxiliary distribution $r(f, u, X|z)$ which serves to approximate the variational posterior, $q(f, u, X|z)$, of the auxiliary variables conditioned on the latent variables.

We may re-express $\mathcal{L}_{aux}$ in a way which makes use of the analytical expression for the K-L divergence between two Gaussians, $q(f, u, x)$ and $r(f, u, X|z)$ and, in the case that the prior of the generative model, $p(z)$, is also Gaussian distributed - as is the case for the continuous latent variable MLP model we'll consider first - then the bound contains a second Guassian KL term:

$$\mathcal{L}_{aux} = \underbrace{\mathbb{E}_{q(z)}\big[\log p(y|z)\big]}_{A} - \underbrace{\mathbb{E}_{q(f,u,X)}\big[\mathrm{KL}[q(z|f,X)\|p(z)]\big]}_{B} - \underbrace{\mathbb{E}_{q(z)}\big[\mathrm{KL}[q(f,u,X)\|r(f,u,X|z)]\big]}_{C}. \tag{6}$$

2 Expressions

2.1 Generative model

2.2 Variational model

$$q(X) = \prod_{i=1}^{N} \mathcal{N}(x_{i,:} \mid 0, I_R) = \prod_{i=1}^{N}\prod_{r=1}^{R} \mathcal{N}(x_{i,r} \mid 0, 1) \tag{7}$$

$$q(u) = \prod_{m=1}^{M} \mathcal{N}(u_{:,m} \mid 0, K_{X_u X_u}) \tag{8}$$

$$q(f(X)|u) = \mathcal{N}(f(X) \mid K_{XX_u}K_{X_uX_u}^{-1}u, K_{XX} - K_{XX_u}K_{X_uX_u}^{-1}K_{X_uX}) \tag{9}$$

$$q(z|f(X)) = \prod_{i=1}^{N} \mathcal{N}(z_i \mid f(X_i), \sigma^2) \tag{10}$$

2.3 Auxiliary model

$$r(f, u, X|z) = q(f(X)|u, X)r(u, X|z) \tag{11}$$

$$r(u, X|z) = \tag{12}$$