Assignment 3

Anna Pavlova

April 2016

1

$$\mathrm{KL}(p||q) = \mathbb{E}_p \log \frac{p}{q} = \mathbb{E}_p \log p - \mathbb{E}_p \log q = -\mathbb{H}p - \sum_{i=1}^{N} \mathbb{E}_{p(\mathbf{Z}_i)} \log q(\mathbf{Z}_i) =$$

$$= -\mathbb{H}p(\mathbf{Z}) - \sum_{i=1}^{N} \mathbb{H}p(\mathbf{Z}_i) + \sum_{i=1}^{N} \mathrm{KL}(p(\mathbf{Z}_i)||q(\mathbf{Z}_i))$$

From this identity we obtain that:

$$q^*(\mathbf{Z}_i) = p(\mathbf{Z}_i) = \int p(\mathbf{Z}) \prod_{j \neq i} d\mathbf{Z}_j$$

However, according to the problem, the technique of Lagrange multipliers should be used, so let us find minimum of $\mathrm{KL}(p||q)$ using the approach (it will be equal to the initial problem statement):

$$\delta \Big[\mathrm{KL}(p||q) - \lambda \Big(\int q(\mathbf{Z}) d\mathbf{Z} - 1 \Big) \Big] = - \int \frac{p(\mathbf{Z})}{q(\mathbf{Z})} \delta q(\mathbf{Z}) d\mathbf{Z} - \lambda \int \delta q(\mathbf{Z}) d\mathbf{Z} =$$

$$= - \int \Big(\frac{p(\mathbf{Z})}{q(\mathbf{Z})} + \lambda \Big) \delta q(\mathbf{Z}) d\mathbf{Z} = 0$$

$$q(\mathbf{Z}) = -\frac{1}{\lambda} q(\mathbf{Z})$$

We see that $q(\mathbf{Z}) \propto p(\mathbf{Z})$ and using the fact that both p and q are distributions we conclude that $p = q$.

2

Let us derive the ELBO:

$$\mathcal{L}(q) = \mathbb{E}_{q(\mathbf{Z})q(\boldsymbol{\theta})} \log \frac{p(\mathbf{Z}, \boldsymbol{\theta})}{q(\mathbf{Z})q(\theta)} = \mathbb{E}_{q(\mathbf{Z})} \log \frac{p(\mathbf{Z}, \boldsymbol{\theta}_\mathbf{0})}{q(\mathbf{Z})} + \mathbb{H}q(\boldsymbol{\theta})$$

$\mathbb{H}q(\boldsymbol{\theta})$ is infinite but constant, because entropy does not depend on shift $\boldsymbol{\theta}_0$. So we can define a new function:

$$\tilde{\mathcal{L}}(q, \boldsymbol{\theta}_0) = \mathbb{E}_{q(\mathbf{Z})} \log \frac{p(\mathbf{Z}, \boldsymbol{\theta}_{\mathbf{0}})}{q(\mathbf{Z})} = \mathcal{L}(q) + \text{const}$$

According to deriving of EM algorithm, optimizing it by $q(\boldsymbol{z})$ is equivalent to E-step, by $\boldsymbol{\theta}_0$ to M-step.