Chapter 10 Variational Inference

Yue Yu

10.1 Variational Inference

(10.6) – (10.9)

If we focus on the factor $q_j(\mathbf{Z}_j)$,

$$\begin{aligned}
\mathcal{L}(q) &= \int q(\mathbf{Z}) \ln \left\{ \frac{p(\mathbf{X}, \mathbf{Z})}{q(\mathbf{Z})} \right\} \mathrm{d}\mathbf{Z} \\
&= \int \prod_i q_i \left\{ \ln p(\mathbf{X}, \mathbf{Z}) - \sum_i \ln q_i \right\} \mathbf{dZ} \\
&= \int q_j \Bigg\{ \underbrace{\int \cdots \int \ln p(\mathbf{X}, \mathbf{Z}) \prod_{i \neq j} q_i \, \mathrm{d}\mathbf{Z}_i}_{\mathbb{E}_{i \neq j}[\ln p(\mathbf{X}, \mathbf{Z})]} \Bigg\} \mathrm{d}\mathbf{Z}_j - \int q_j \ln q_j \, \mathrm{d}\mathbf{Z}_j + \text{const} \\
&= -\mathrm{KL}(q_j \| q_j^*) + \text{const}
\end{aligned}$$

where we defined a new distribution q_j^* that satisfies

$$\ln q_j^* = \mathbb{E}_{i \neq j}[\ln p(\mathbf{X}, \mathbf{Z})] + \text{const}.$$

Since KL divergence is non-negative, the lower bound $\mathcal{L}(q)$ reaches the maximum when $\mathrm{KL}(q_j \| q_j^*) = 0$. In other words, when the optimal value is just q_j^*.