Variational Inference

总的目标是**求隐变量的后验分布**$p(z|x)$，但用贝叶斯公式难以求解

$$
\begin{aligned}
p(z|x) &= \frac{p(x,z)}{p(x)} \\
&= \frac{p(x,z)}{\int p(x,z)\mathrm{d}z}
\end{aligned}
$$

因此用参数化的分布族$q(z;\lambda)$来拟合

$$
\min_{\lambda} KL(q(z;\lambda)||p(z|x)) \tag{1}
$$

上式不能显示求解，因此需要转化。

注意到

$$
\begin{aligned}
\log p(x) &= \log p(x,z) - \log p(z|x) \\
&= \log \frac{p(x,z)}{q(z;\lambda)} - \log \frac{p(z|x)}{q(z;\lambda)}
\end{aligned}
$$

两边对$q(z;\lambda)$取期望得

$$
\begin{aligned}
\mathbb{E}_{q(z;\lambda)} \log p(x) &= \mathbb{E}_{q(z;\lambda)} \log \frac{p(x,z)}{q(z;\lambda)} - \mathbb{E}_{q(z;\lambda)} \log \frac{p(z|x)}{q(z;\lambda)} \\
&= KL(q(z;\lambda)||p(z|x)) + \mathbb{E}_{q(z;\lambda)} \log \frac{p(x,z)}{q(z;\lambda)}
\end{aligned}
$$

也就是

$$
\log p(x) = KL(q(z;\lambda)||p(z|x)) + \mathbb{E}_{q(z;\lambda)} \log \frac{p(x,z)}{q(z;\lambda)}
$$

于是(1)转化为

$$\min_\lambda KL(q(z;\lambda)||p(z|x)) \iff \max_\lambda \mathbb{E}_{q(z;\lambda)} \log \frac{p(x,z)}{q(z;\lambda)}$$

变分推断的目标于是归结为

$$\max_\lambda \mathbb{E}_{q(z;\lambda)}[\log p(x,z) - \log q(z;\lambda)]$$

称$\mathbb{E}_{q(z;\lambda)}[\log p(x,z) - \log q(z;\lambda)]$为**ELBO**

ELBO有如下性质

$$\log p(x) = KL(q(z;\lambda)||p(z|x)) + \text{ELBO}$$

由KL散度非负性，有

$$\log p(x) \geq \text{ELBO}$$

因此也可用ELBO来最大化对数似然