ELBO Lower Bound

Thursday 7 November 2024 20:06

We wish to calculate a lower bound on $\log p(x)$, since to calculate $\log p(x)$ we need $\log p(x) = \log \int p(x,z)\,dz$ for all latent variables z, which is intractable. So instead we have:

$$\log p(x) = \log \int_z p(x,z)\,dz = \log \int_z q(z) \frac{p(x,z)}{q(z)}\,dz$$ where

$q(z)$ is our variational prob. distribution. Then by defn.

$$\log p(x) = \log \mathbb{E}_{z\sim q(z)}\left(\frac{p(x,z)}{q(z)}\right)$$

Now note Jensen's inequality: for a real concave function f,

$f(\mathbb{E}[X]) \geq \mathbb{E}[f(X)]$ (proof: intuitive graphically, needs measure theory)

$f(x) = \log x$ is a concave function, so we can get

$$\log p(x) \geq \mathbb{E}_{z\sim q(z)} \log\left(\frac{p(x,z)}{q(z)}\right)$$

$p(x,z) = p(x|z)p(z)$ $\therefore \log p(x) \geq \mathbb{E}_{z\sim q(z)} \log\left(\frac{p(x|z)p(z)}{q(z)}\right)$

$$= \mathbb{E}_{z\sim q(z)} \log p(x|z) + \mathbb{E}_{z\sim q(z)} \log\frac{p(z)}{q(z)}$$
$$= \mathbb{E}_{z\sim q(z)} \log p(x|z) - \mathbb{E}_{z\sim q(z)} \log\frac{q(z)}{p(z)}$$

$$= \mathbb{E}_{z\sim q(z)} \log p(x|z) - D_{KL}(q(z)\,||\,p(z)) = ELBO$$

where $D_{KL}(q(z)\,||\,p(z)) = \mathbb{E}_{z\sim q(z)} \log \frac{q(z)}{p(z)}$ is the

<u>Kullback-Liebler Divergence</u>

- A type of statistical distance
- Expected excess surprise from using Q as model instead of P
- 'relative entropy'; 'squared distance'; 'divergence' - see info. theory.

In the case of ML problems, usually we care about:

$$\log p(x) \geq ELBO = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x)\,||\,p(z))$$

ELBO more tractable quantity