The Expectation-Maximization (EM) Algorithm

Jacob Goldberger

$$\log p(x;\theta) = \log \sum_z p(z,x;\theta) = \log \sum_z q(z) \frac{p(z,x;\theta)}{q(z)} \ge \sum_z q(z) \log p(x,z;\theta) - \sum_z q(z) \log q(z) = F(\theta, q)$$

$F(\theta, q)$ is called Evidence Lower BOund (ELBO) or negative free energy.

$$\log p(x;\theta) - F(\theta, q) = \sum_z q(z) \log \frac{q(z)}{p(z|x;\theta)} = KL(q(z)||p(z|x;\theta))$$

$$\Rightarrow \log p(x;\theta) = F(\theta, p(z|x;\theta)) = \max_q F(\theta, q)$$

$$\Rightarrow \theta_{ML} = \arg\max_\theta \log p(x;\theta) = \arg\max_\theta \max_q F(\theta, q)$$

Iterative EM algorithm:

- E-step: $q(z) = p(z|x;\theta_0)$
- M-step: $\theta_1 = \arg\max_\theta Q(\theta, \theta_0)$ s.t. $Q(\theta, \theta_0) = \sum_z p(z|x;\theta_0) \log p(x,z;\theta)$ is the auxiliary function:

Theorem: $p(x;\theta_1) \ge p(x;\theta_0)$

Variational EM:

$$\log p(x;\theta) = \log \sum_z p(z,x;\theta) \ge \sum_z q(z|x;\lambda) \log p(x,z;\theta) - \sum_z q(z|x;\lambda) \log q(z|x;\lambda) = F(\theta, \lambda)$$

$$\log p(x;\theta) - F(\theta, \lambda) = \sum_z q(z|x;\lambda) \log \frac{q(z|x;\lambda)}{p(z|x;\theta)} = KL(q(z|x;\lambda)||p(z|x;\theta))$$

$F(\theta, \lambda)$ is called Evidence Lower BOund (ELBO) or negative free energy. The ELBO score can be also written as:

$$F(\theta, \lambda) = \sum_z q(z|x;\lambda) \log p(x|z;\theta) - KL(q(z|x;\lambda)||p(z;\theta))$$

$$\hat{\theta} = \arg\max_\theta \log p(x;\theta) = \arg\max_\theta \max_\lambda F(\theta, \lambda) \ge \arg\max_\theta \max_{\lambda \in \Lambda} F(\theta, \lambda)$$

E-step: Find $\lambda \in \Lambda$ that maximizes $F(\theta, \lambda)$.
M-step: Find θ that maximizes $\sum_z q(z|x;\lambda) \log p(x,z;\theta)$.