$$L(q,\theta) = E_q[\log P(x,z) - \log q]$$
$$= E_q[\log P(x,z)] - \underbrace{E_q[\log q]}_{H[q]}$$

~~MCMC: Markov Chain and Monte Carlo~~

(十二) 变分推断

Variational Inference (变分推断)

矩阵角度优化问题
1. 回归: $f(w) = w^T x$ (模型)
   loss function: $L(w) = \sum_{i=1}^{N} \| w^T x_i - y_i \|$
   $\hat{w} = \arg\min_w L(w)$ (策略)
   解法: ① 解析向下, $\frac{\partial L(w)}{\partial w} = 0 \Rightarrow w^* = (X^T X)^{-1} X^T Y$
   ② 梯度下降: GD
2. 分类
   SVM
3. EM

Inference
- 精确推断
- 近似推断
  - 确定性近似 → VI
  - 随机近似 → MCMC, MH, Gibbs.

贝叶斯角度 → 积分问题

$$P(\theta|x) = \frac{P(\theta)P(x|\theta)}{P(x)} \rightarrow \int_\theta P(x|\theta)P(\theta)\,d\theta$$

贝叶斯 inference (推断)

贝叶斯决策: 对新样本预测 $P(\tilde{x}|X)$.

$$P(\tilde{x}|X) = \int P(\tilde{x}|\theta)P(\theta|X)\,d\theta = \int P(\tilde{x}|\theta)\frac{P(\theta)P(X|\theta)}{P(X)}\,d\theta$$