Variational Inference

x : observed

z : Latent/hidden/unobserved

$$p(z|x) = \frac{p(x,z)}{p(x)} = \frac{p(X=x,z)}{p(X=x)}$$

$$\boxed{p(x) = \sum_{z} p(x,z)}$$

$$q(z) \simeq p(z|x)$$

$$q^*(z) = \underset{q(z) \in Q}{\arg\min}\ d\big(\underline{q(z)}, \underline{p(z|x)}\big)$$

$p(x) \qquad q(x) \qquad x \in X$

① $\boxed{\min_x |p(x) - q(x)|}$ ✗ $\qquad \dfrac{1}{\sum_x p(x)\, q(x)}$

② $\boxed{\sum_{x \in X} |p(x) - q(x)|}$ ✓

$d(p, q)$

③ $\boxed{\sum_{x \in X} (p(x) - q(x))^2}$ ✓

distance $\boxed{d(x,y)}$

① $d(x,y) \geq 0$

② ~~$d(x,x) = 0$~~ $\qquad d(x,y) = 0 \Leftrightarrow x = y$

③ $d(x,y) = d(y,x)$

④ $d(x,y) + d(y,z) \geq d(x,z)$

$$d_1 = \sum_{x \in X} p(x) \, |p(x) - q(x)|$$

$$d_1 \neq d_2$$

$$d_2 = \sum_{x \in X} q(x) \, |p(x) - q(x)|$$

$$d(p, q) = \sum_{a} p(x) \, |p(x) - q(x)|$$

$$d(q, p) = \sum q(x) \, |p(x) - q(x)|$$

$$\sum_x p(x)q(x)\,|p(x)-q(x)|$$

متوسط‌گیری نیست

$p(x) > 0$

$q(x) > 0$

$$\sum_x p(x)q(x) \neq 1$$

$p(x) \cong q(x) \geq \epsilon$

$p(x) \cong q(x) \geq 1-\epsilon$

$$\sum_{x \in X} \log\left(\frac{p(x)}{q(x)}\right) \cancel{p(x)}$$

$p(x) \geq 0$

$q(x) \geq 0$

$p(x) = 10^{-1}$

$q(x) = 10^{-10}$

$$KL(p \| q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)} \geq 0$$

$$KL(q \| p) = \sum_{x \in X} q(x) \log \frac{q(x)}{p(x)} \geq 0$$

$$\boxed{q^*(z) = \underset{\underline{q(z) \in Q}}{\arg\min} \; KL\Big(q(z) \,\|\, p(z|x)\Big) \quad \checkmark}$$

$$q^*(z) = \underset{\underline{\underline{q(z) \in Q}}}{\arg\min} \; KL\Big(p(z|x) \,\|\, q(z)\Big)$$

$$KL(p\|q) = \sum_x p(x)\log\frac{p(x)}{q(x)}$$

$$\boxed{\begin{array}{c} p(x) > 0 \\ q(x) \to 0 \end{array}}$$

$p(x) \to 0$

$q(x) > 0$

$$KL(q\|p) = \sum_x q(x)\log\frac{q(x)}{p(x)}$$

q

$p(x) \to 0$

$q(x) > 0$

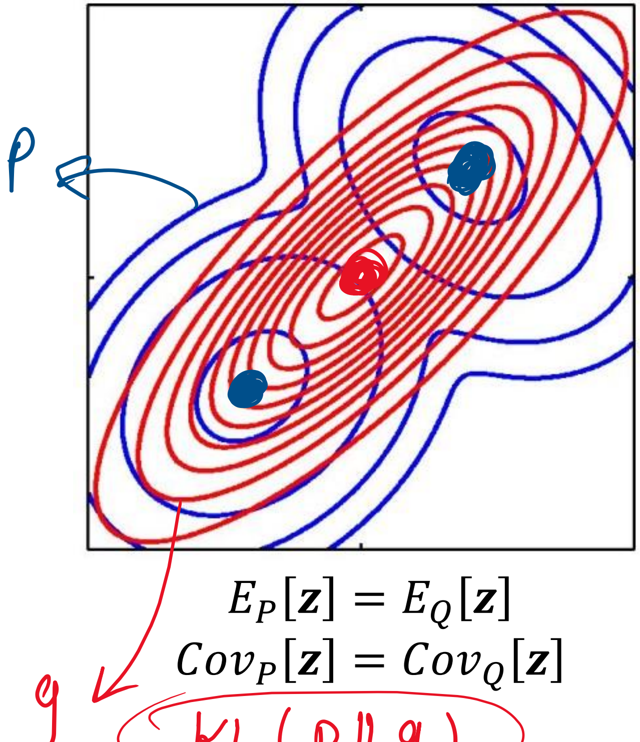
KL divergence: M-projection vs. I-projection

- Let P is mixture of two 2D Gaussians and Q be a 2D Gaussian distribution with arbitrary covariance matrix:

$$Q^* = \underset{Q}{\operatorname{argmin}} \int P(\mathbf{z}) \log \frac{P(\mathbf{z})}{Q(\mathbf{z})} d\mathbf{z}$$

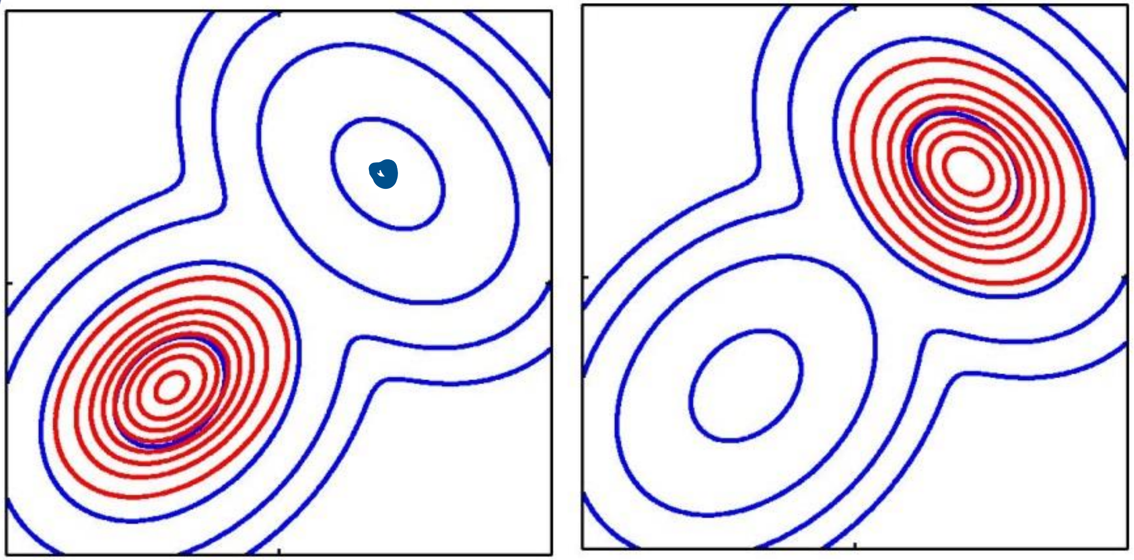
P: Blue
Q^*: Red

$$Q^* = \underset{Q}{\operatorname{argmin}} \int Q(\mathbf{z}) \log \frac{Q(\mathbf{z})}{P(\mathbf{z})} d\mathbf{z}$$



$$E_P[\mathbf{z}] = E_Q[\mathbf{z}]$$
$$Cov_P[\mathbf{z}] = Cov_Q[\mathbf{z}]$$

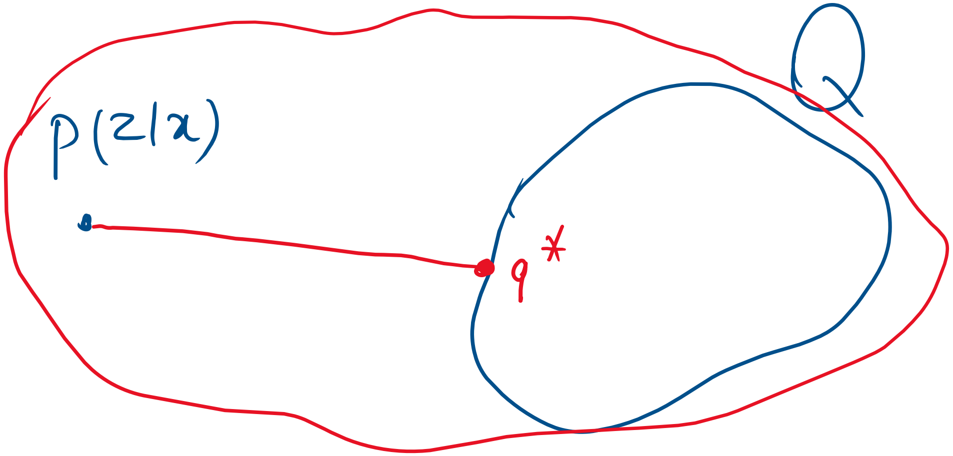
KL(p||q)



two good solutions!

KL(q||p)

[Bishop]



p(z|x)
Q
q*

$$q^*(z) = \underset{q(z)\in Q}{\arg\min}\; KL\big(q(z) \,\|\, \underbrace{p(z|x)}_{?}\big)$$

$$p(z|x) = \frac{p(x,z)}{p(x)}$$

$$KL\big(q(z) \,\|\, p(z|x)\big) = \sum_z q(z) \log \frac{q(z)}{p(z|x)}$$

$$= \sum_z q(z) \log \frac{q(z)\,p(x)}{p(x,z)} = \sum_z q(z) \Big(\underline{\log \frac{q(z)}{p(x,z)}} + \underline{\log p(x)} \Big)$$

$$= \underbrace{\sum_z q(z) \log \frac{q(z)}{p(x,z)}}_{} + \underbrace{\sum_z q(z) \log p(x)}_{}$$

$$= KL\big(q(z) \,\|\, p(x,z)\big) + \log p(x)$$

$$KL\left(q(z)\,\|\,p(z|x)\right) = KL\left(q(z)\,\|\,p(x,z)\right) + \log P(x)$$

$$\Rightarrow \underbrace{\log P(x)}_{\substack{\text{const.}\\ \text{w.r.t } q}} = \underbrace{-KL\left(q(z)\,\|\,p(x,z)\right)}_{\max \uparrow} + \underbrace{KL\left(q(z)\,\|\,p(z|x)\right)}_{\min \downarrow}$$

$$q^*(z) = \arg\min_{q(z)} KL\left(q(z)\,\|\,p(z|x)\right)$$

$$\boxed{= \arg\max_{q(z)} -KL\left(q(z)\,\|\,\underline{p(x,z)}\right)}$$

$$\boxed{\log \underbrace{P(x)}^{\text{Evidence}}} = -KL\left(q(z) \,\|\, P(x,z)\right) + \underbrace{KL\left(q(z) \,\|\, P(z|x)\right)}_{\geq 0}$$

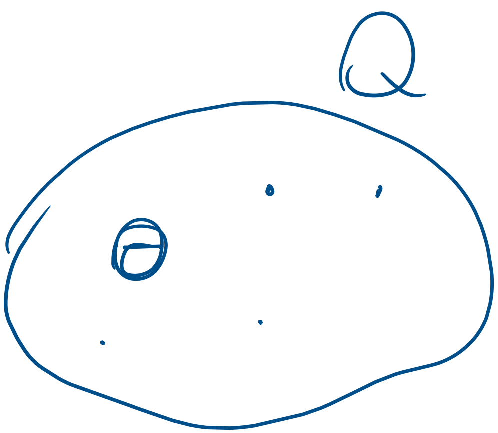
$$S = x \cup z$$

$$\underbrace{-KL\left(q(z) \,\|\, P(x,z)\right)}_{\text{Evidence Lower Bound (ELBO)}} \leq \underbrace{\boxed{\log P(x)}}_{\text{Evidence}}$$

$$\boxed{P(x) = \sum_{z} P(x,z)}$$

$$\boxed{\theta^* = \arg\max_\theta \log P(x)}$$

$$\tilde{\theta} = \arg\max_\theta \text{ELBO} = \arg\max_\theta -KL\left(q_\theta(z) \,\|\, P(x,z)\right)$$



VAE $\max_{\theta}$ ELBO

NF $\max_{\theta}$ $\underbrace{\text{Likelihood}}_{\log P(x)}$