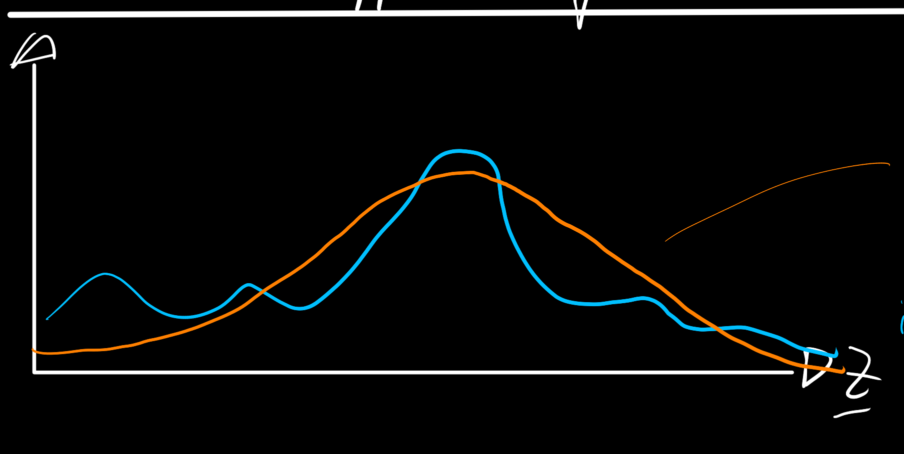
Mean Field Approach for Variational Inference



$$q^*(\underline{z}) = \underset{q(\underline{z})\in Q}{\arg\min} \left(KL(q(\underline{z}) \,||\, p(\underline{z}|D)) \right)$$

$\Updownarrow$

$$q^*(\underline{z}) = \underset{q(\underline{z})\in Q}{\arg\max} \left(\mathcal{L}(q) \right)$$

What Q?

$$\boxed{q(\underline{z}) = \prod_{d=0}^{D-1} q_d(z_d)}$$

latent vector $\in \mathbb{R}^D$

latent dimensions are independent

could also be a subdivision into smaller vectors

$\underline{z} \in \mathbb{R}^3$

$$q(\underline{z}) = q_0(z_0)\, q_1(z_1)\, q_2(z_2)$$

$$q(\underline{z}) = q_0(z_0)\, q_1(z_1)\, q_2(z_2) = \underset{q_0, q_1, q_2}{\arg\max} \left(\mathcal{L}(q(\underline{z})) \right)$$

ELBO $\quad \mathcal{L}(q) = \mathbb{E}_{\underline{z}\sim q(\underline{z})} \left[\log\left(\frac{p(\underline{z}, D)}{q(\underline{z})} \right) \right]$

$$= \int_{\underline{z}} q(\underline{z}) \cdot \log\left(\frac{p(\underline{z}, D)}{q(\underline{z})} \right) d\underline{z}$$

$$\mathcal{L}(q) = \iiint_{z_0\, z_1\, z_2} q_0(z_0) q_1(z_1) q_2(z_2) \cdot \left(\log p(\underline{z}, D) - \log q_0(z_0) - \log q_1(z_1) - \log q_2(z_2) \right) dz_0\, dz_1\, dz_2$$

$q_0 := q_0(z_0) \qquad q_1 := q_1(z_1) \qquad q_2 := q_2(z_2) \qquad p := p(\underline{z}, D)$

$$= \int_{z_0}\int_{z_1}\int_{z_2} q_0\, q_1\, q_2 \log p \; dz_0\, dz_1\, dz_2$$

$$- \int_{z_0}\int_{z_1}\int_{z_2} q_0\, q_1\, q_2 \left(\log q_0 + \log q_1 + \log q_2 \right) dz_0\, dz_1\, dz_2$$

$$= \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1 q_2 \log p \; dz_1\, dz_2 \right) dz_0$$

$$- \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1\, q_2 \left(\log q_0 + \log q_1 + \log q_2 \right) dz_1\, dz_2 \right) dz_0$$

$$= \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1 q_2 \log p \; dz_1\, dz_2 \right) dz_0$$

$$- \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1 q_2 \log q_0 \; dz_1\, dz_2 \right) dz_0$$

$$- \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1 q_2 \left(\log q_1 + \log q_2 \right) dz_1\, dz_2 \right) dz_0$$

$$= \int_{z_0} q_0 \left(\iint_{z_1 z_2} q_1 q_2 \log p \; dz_1\, dz_2 \right) dz_0$$

$$- \int_{z_0} q_0 \log q_0 \, dz_0 \cdot \underbrace{\iint_{z_1 z_2} q_1\, q_2 \; dz_1\, dz_2}_{1}$$

$$- \underbrace{\int_{z_0} q_0 \, dz_0}_{1} \iint_{z_1 z_2} q_1\, q_2 \left(\log q_1 + \log q_2 \right) dz_1\, dz_2$$

$$= \int_{z_0} q_0 \left(\underbrace{\iint_{z_1 z_2} q_1 q_2 \log p \; dz_1\, dz_2}_{\mathbb{E}_{1,2}[\log p]} \right) dz_0$$

$$- \int_{z_0} q_0 \log q_0 \, dz_0 - \underbrace{\iint_{z_1 z_2} q_1 q_2 \left(\log q_1 + \log q_2 \right) dz_1 dz_2}_{\mathbb{E}_{1,2}[\log q_1 + \log q_2]}$$

$$\mathbb{E}_{1,2}[\cdot] = \mathbb{E}_{\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim q_1(z_1) q_2(z_2)} [\cdot]$$

$$= \underbrace{\int_{z_0} q_0 \cdot \left(\mathbb{E}_{1,2}[\log p] - \log q_0 \right) dz_0}_{\text{only integration over } z_0} - \underbrace{\mathbb{E}_{1,2}[\log q_1 + \log q_2]}_{\text{does not contain } z_0}$$

$\rightarrow$ Fix q_1 & q_2

$$\rightsquigarrow \mathcal{L}(q) = \mathcal{L}(q_0) = \underbrace{\int_{z_0} \overbrace{q_0 \left(\mathbb{E}_{1,2}[\log p] - \log q_0 \right)}^{\ell} dz_0}_{I(q_0)} + \text{const.}$$

Back into optimization

$$\left[q_0^*, q_1^*, q_2^* = \underset{q_0, q_1, q_2}{\arg\max}\, \mathcal{L}(q) \right]$$

fixed q_1, q_2

$$q_0^* = \underset{q_0}{\arg\max}\, \mathcal{L}(q_0)$$

$$= \underset{q_0}{\arg\max}\, I(q_0)$$

$$\left(\text{s.t.} \int_{z_0} q_0 \, dz_0 = 1 \right)$$

$\rightarrow$ Take functional derivative and set to zero

$$\frac{\delta I(q_0)}{\delta q_0} \overset{!}{=} 0 \quad \underset{\text{Lagrange}}{\overset{\text{Euler}}{\rightsquigarrow}} \quad \frac{\partial \ell}{\partial q_0} \overset{!}{=} 0$$

$$0 \overset{!}{=} \mathbb{E}_{1,2}[\log p] - \log q_0 - \cancel{\frac{q_0}{q_0}}^{1}$$

$\Leftrightarrow$

$$\log q_0 = \mathbb{E}_{1,2}[\log p] - \cancel{1}$$

$$\boxed{\log q_0(z_0) \overset{+}{=} \underbrace{\mathbb{E}_{1,2} [\log \overbrace{p(z_0, z_1, z_2, D)}^{\text{depends on } z_0, z_1, z_2}]}_{\text{depends on } z_0}}$$

in general:

$$\boxed{\log q_j(z_j) \overset{+}{=} \mathbb{E}_{\underbrace{i \neq j}_{\text{all } z_i \text{ without } z_j}} [\log p(\underline{z}, \underline{x} = D)]}$$