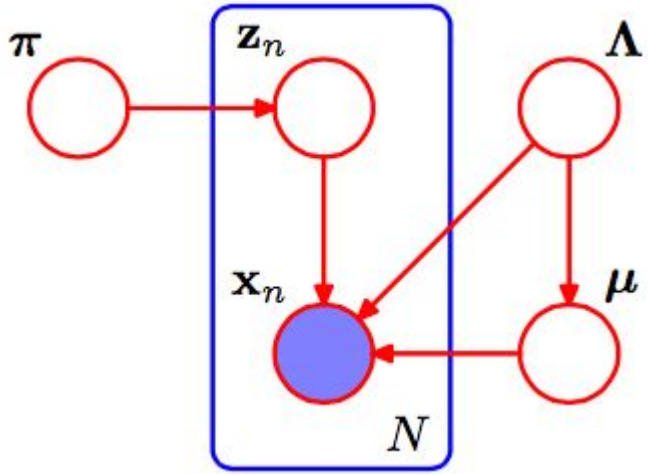
DD2434

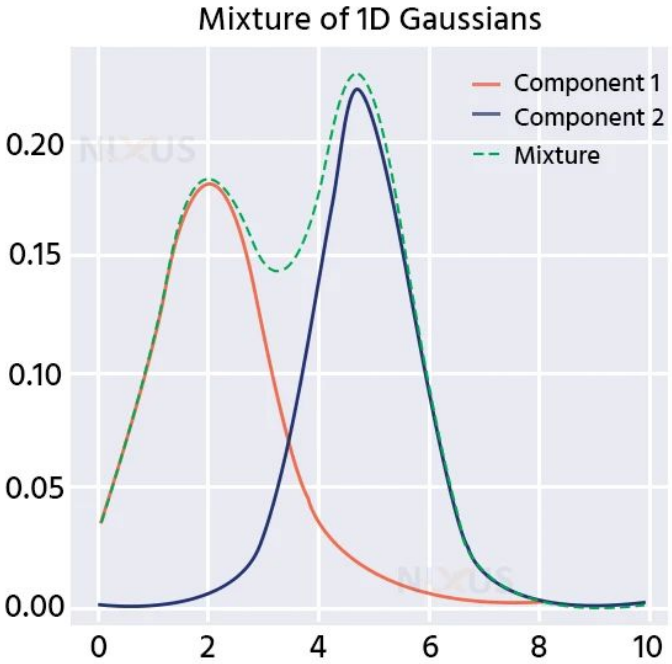
Workshop 2
Module 3

Recap

- Module 1: derived the posterior distribution for simple cases
- Module 2: framework for analyzing more complex models
- Module 3: VI - an approach infer the posterior of of complex models

Example: Gaussian Mixture model

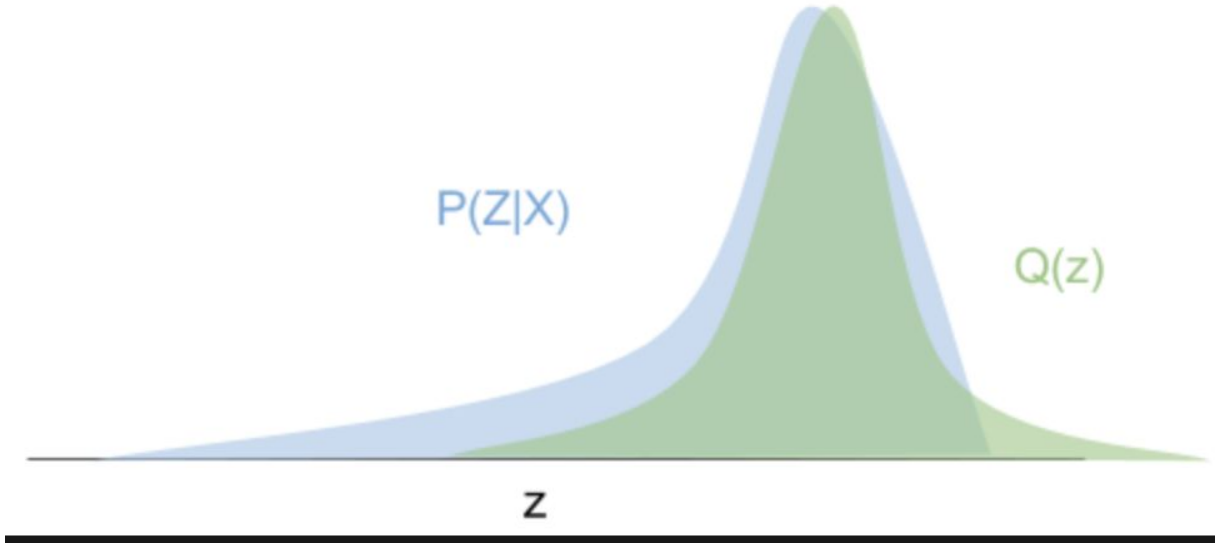




$$p(\mu_{1:K}, z_{1:n}|x_{1:n}) = \frac{\prod_k p(\mu_k) \prod_i p(z_i) p(x_i|z_i, \mu_{1:K})}{\int_{\mu_{1:K}} \sum_{z_{1:n}} \prod_k p(\mu_k) \prod_i p(z_i) p(x_i|z_i, \mu_{1:K})}$$

Variational inference

$$q = \underset{q \in \mathcal{Q}}{\operatorname{argmin}} D_{\mathbb{KL}} \left(q(\boldsymbol{z}) \parallel p_{\boldsymbol{\theta}}(\boldsymbol{z}|\boldsymbol{x}) \right)$$

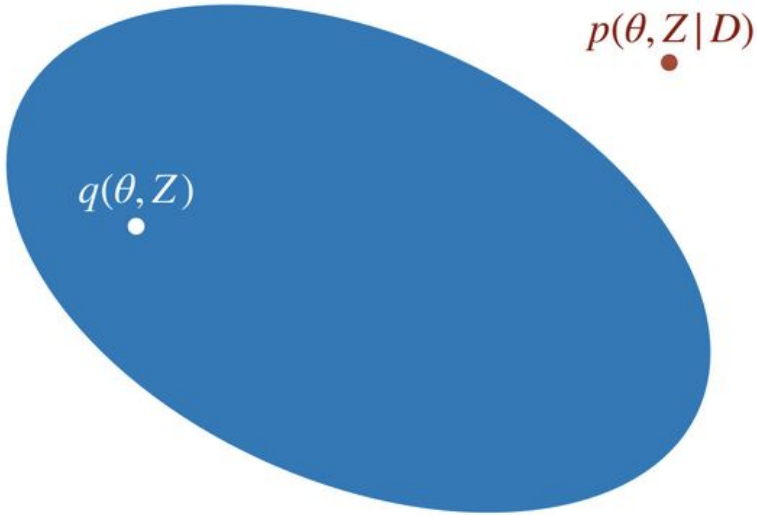


ELBO

$$\mathcal{D}_{KL}(q||p) = \log p(X) - E_{\Psi}\left[\log \frac{p(X,\Psi)}{q(\Psi)}\right]$$

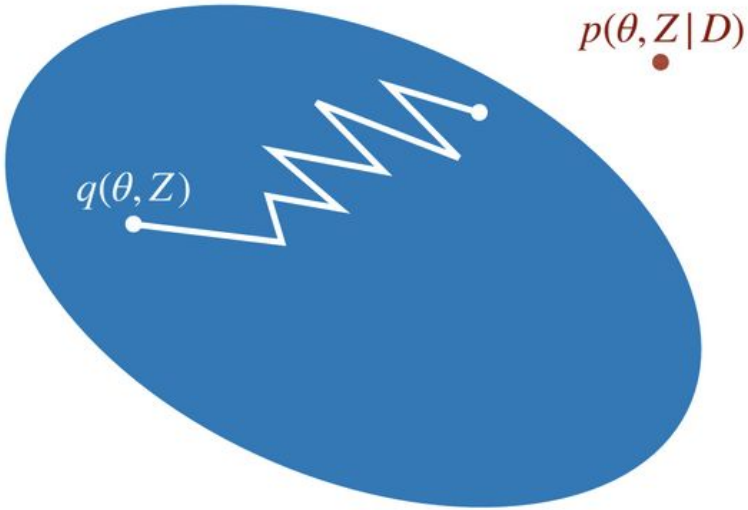
MINIMIZING KL-MAXIMIZING ELBO

$q(\theta, Z) = q_\theta(\theta) q_Z(,Z)$



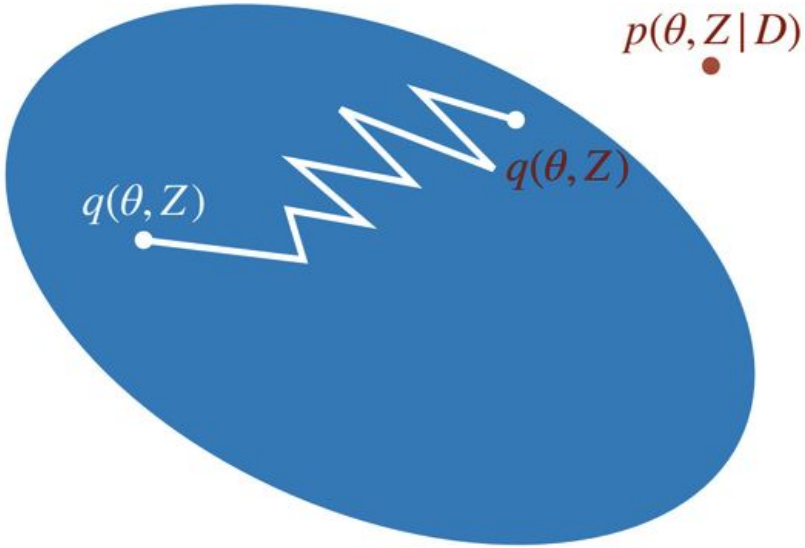
MINIMIZING KL-MAXIMIZING ELBO

$q(\theta, Z) = q_\theta(\theta) q_Z(,Z)$



MINIMIZING KL-MAXIMIZING ELBO

$q(\theta, Z) = q_\theta(\theta) q_Z(,Z)$



CAVI update equation

$$\log q^*(\Psi_k) \stackrel{+}{=} E_{-\Psi_k}[\log P(\Psi, X)]$$

CAVI Algorithm

Algorithm 1 Coordinate ascent VI

procedure APPROXIMATE POSTERIOR(X)
Initialise $q(\Psi_i)$ for $i = 1, \ldots, l$
  repeat
    for $i = 1, \ldots, l$ **do**
      Update: $\log q^*(\Psi_i) \propto E_{-\Psi_i}[\log P(\Psi, X)]$
  until convergence
  Approximate posterior: $q(\Psi) = \prod_i q^*(\Psi_i)$
return $q(\Psi)$

CAVI Algorithm

Algorithm 1 Coordinate ascent VI

procedure APPROXIMATE POSTERIOR(X)
Initialise $q(\Psi_i)$ for $i = 1, \ldots, l$
  repeat
    for $i = 1, \ldots, l$ **do**
      Update: $\log q^*(\Psi_i) \propto E_{-\Psi_i}[\log P(\Psi, X)]$
  until convergence
  Approximate posterior: $q(\Psi) = \prod_i q^*(\Psi_i)$
return $q(\Psi)$