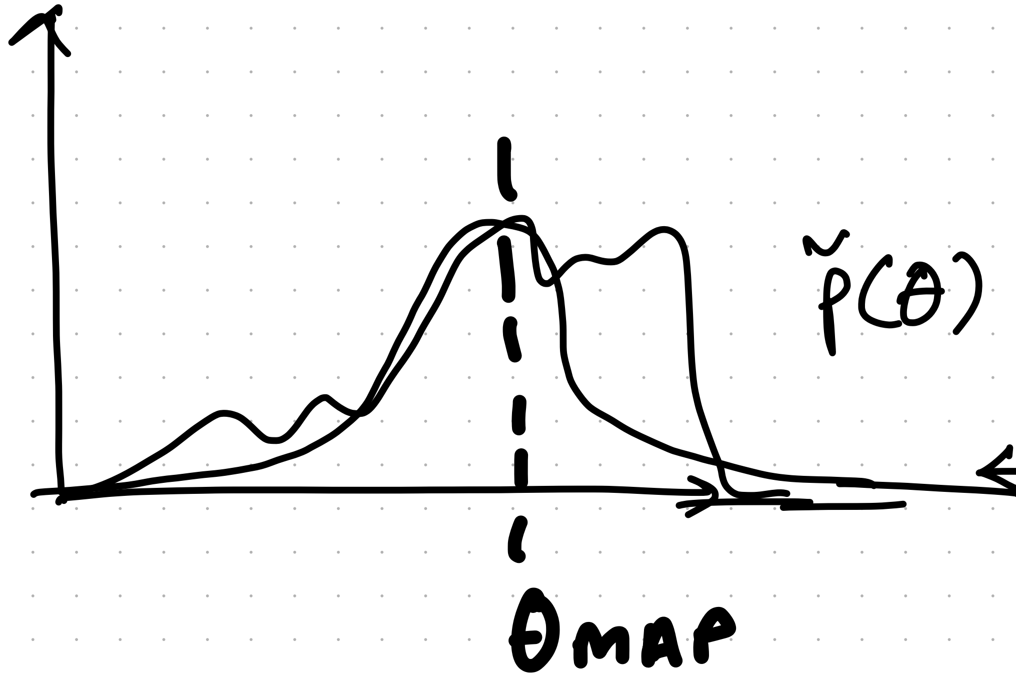
# Variational Inference

$$P(\theta|D) = \frac{P(D|\theta) \cdot P(\theta)}{P(D)}$$

$$\propto P(D|\theta)\, P(\theta) \triangleq \tilde{p}(\theta)$$

Laplace Appr.



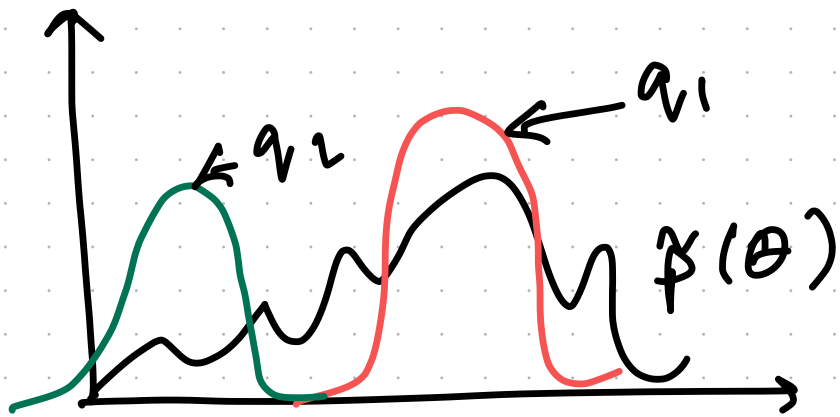

$\tilde{p}(\theta)$ $\left( p(D|\theta) \cdot p(\theta) \right)$

$q(\theta)$

$\sim \mathcal{N}(\theta \mid \theta_{MAP}, \Sigma a 6)$

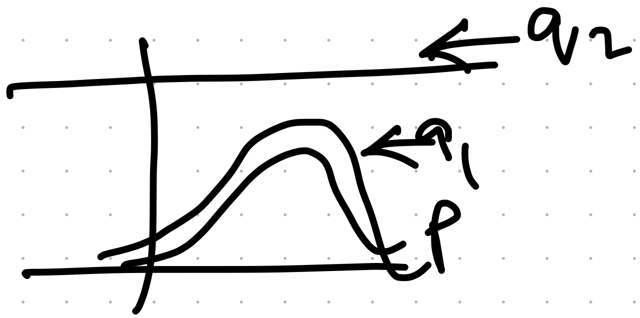
$\hookrightarrow$ $2^{nd}$ order Taylor Series Approximation

Variational Inference

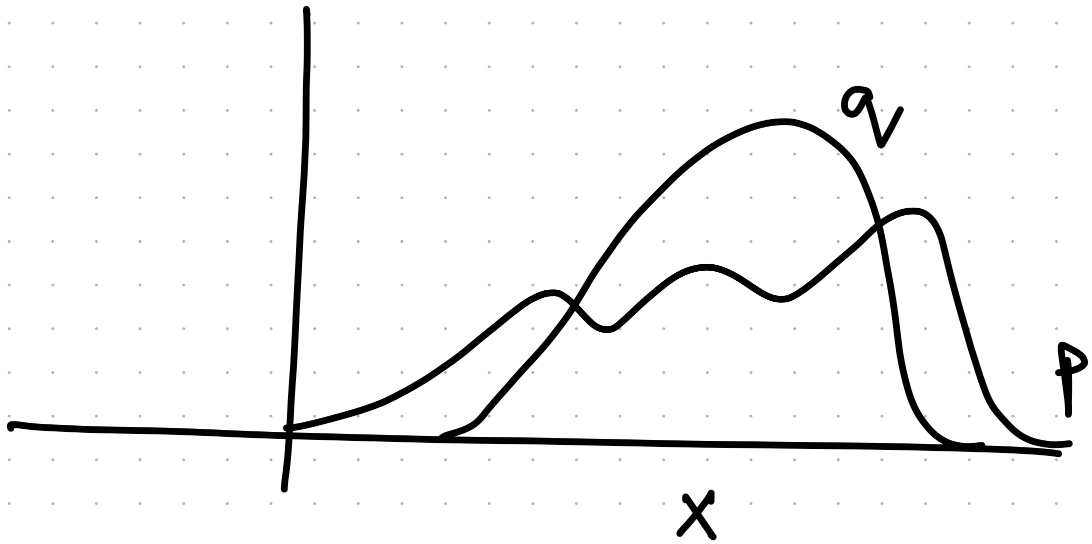


KL divergence

$$D_{KL}[q \| p] = E_{x \sim q}\left[\log\left(\frac{q(x)}{p(x)}\right)\right]$$




$$D_{KL}[q_1 \| p] << D_{KL}[q_2 \| p]$$




$q(x_i) - p(x_i)$ to be small

$\log q(x_i) - \log p(x_i)$ to be small

$\log \frac{q(x_i)}{p(x_i)}$ to be small

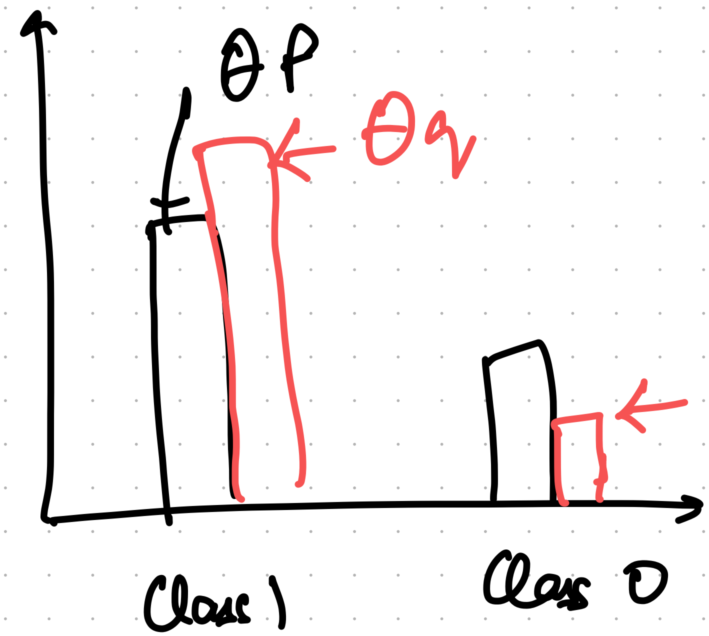
$$E_{x \sim q}\left[\log\left(\frac{q(x_i)}{p(x_i)}\right)\right]$$

$$D_{KL}[q\|p] = E_{x\sim q}\left[\log\left(\frac{q(x_i)}{p(x_i)}\right)\right]$$

$$= \sum q(x_i)\log\left(\frac{q(x_i)}{p(x_i)}\right) \text{ ; } x \text{ discrete}$$

OR

$$= \int q(x)\log\left(\frac{q(x)}{p(x)}\right)dx$$




$$D_{KL}[q||p]$$

$$q(x) = \theta_q^{x}(1-\theta_q)^{1-x}$$

$$p(x) = \theta_p^{x}(1-\theta_p)^{1-x}$$

$$D_{KL}[q||p] = \sum_{x=0}^{1} q(x) \log \frac{q(x)}{p(x)}$$

$$D_{KL}[q \| p] = \sum_{x=0}^{1} q(x) \log \frac{q(x)}{p(x)}$$

$$= q(0) \log \frac{q(0)}{p(0)} + q(1) \log \frac{q(1)}{p(1)}$$

$$= (1-\theta_q) \log \left( \frac{1-\theta_q}{1-\theta_p} \right) + \theta_q \log \frac{\theta_q}{\theta_p}$$