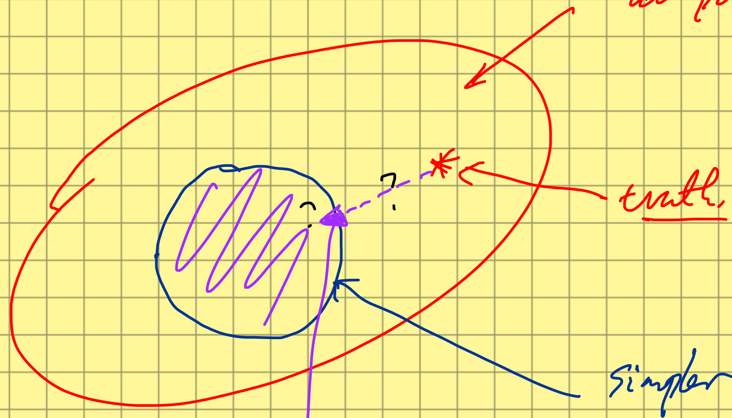


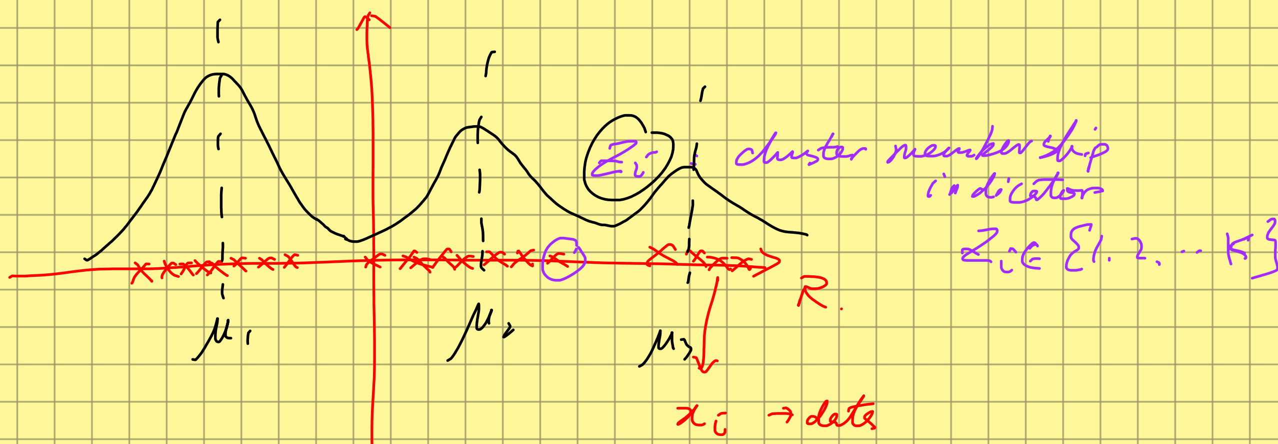
variational distributions

$\{ q : q(\underset{\sim}{z} ; \underset{\sim}{\nu}) \}_{\underset{\sim}{\nu}}$

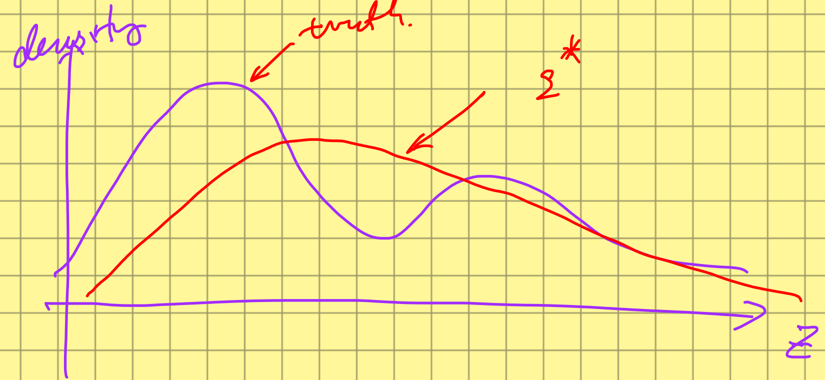
$q^*(\underset{\sim}{z}) = q^*(\underset{\sim}{z} ; \underset{\sim}{\nu}^*)$.

Key Question
1) How to measure the difference between distn.
2) How to find the closest one?

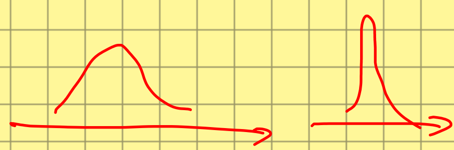
Gaussian Mixture



Caveats about V.I.

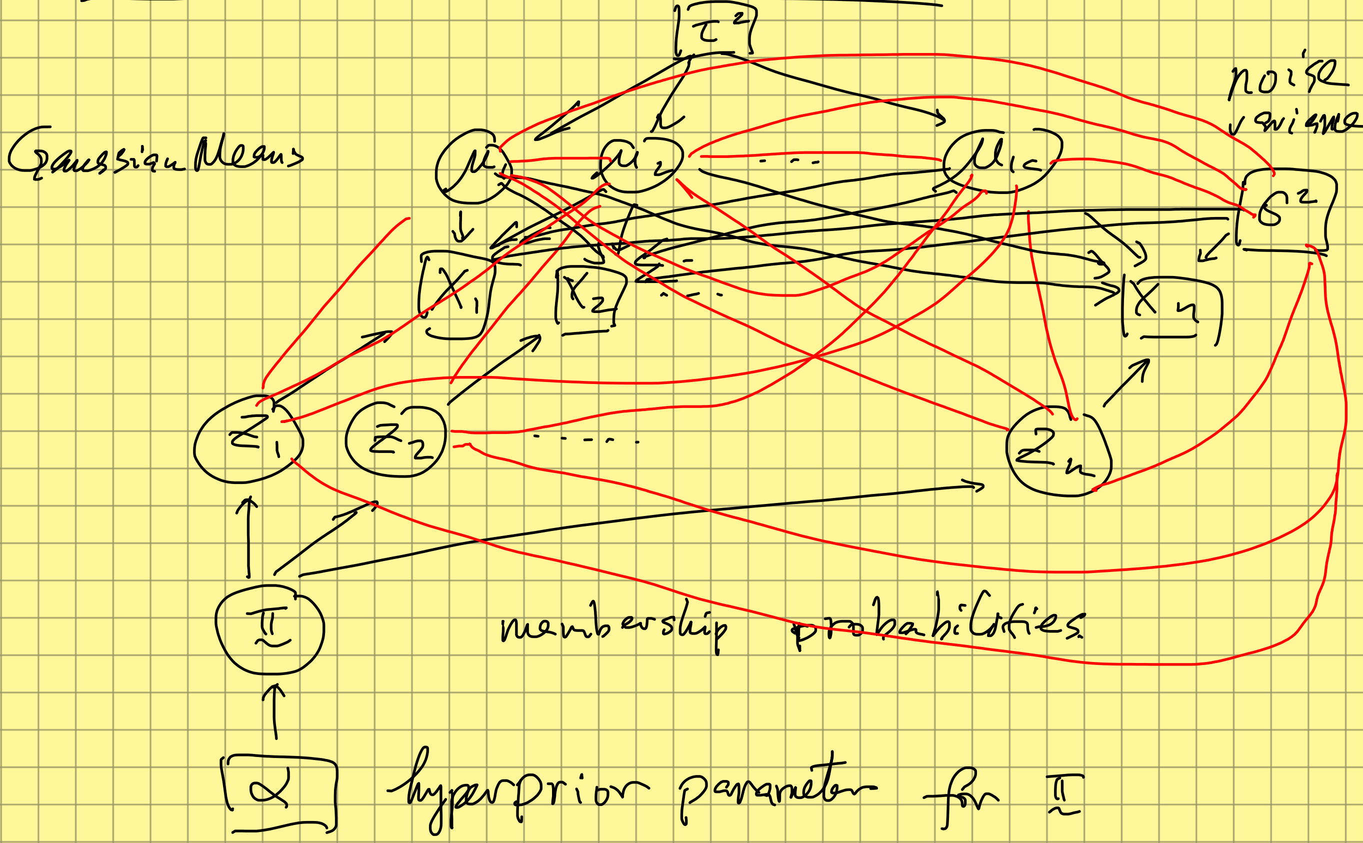


suppose we choose a family of unimodal distn,



...

DAG Gaussian Mixture Model



$$[z_1 \mid z_{-1}, \underset{\sim}{\mu}, \underset{\sim}{\pi}, \underset{\sim}{X}, \tau^2, \sigma^2, \alpha]$$

$$pr(\underline{\underline{\underset{\sim}{z}, \underset{\sim}{\mu}}} \mid \underset{\sim}{X}, \tau^2, \sigma^2, \alpha)$$

NOT INDEPENDENT

On Coordinate Ascent update

Q: why use $p(z_k \mid z_{-k}, x)$ when updating the marginal dist'n $q(z_k)$?

A: For computational simplicity, putting z_k at the last position in the generic chain rule factorization so that z_k only appears in ELB summation once!

Suppose to approximate $p(z_1, z_2, z_3 \mid x)$, we use a family $q(z_1, z_2, z_3) = \prod_{j=1}^{3} q(z_j)$.

We update $q(z_1)$, then $q(z_2)$, then $q(z_3)$.

1) update $q(z_1)$

$$ELB(q) = E_q(\log p(z_1, z_2, z_3 \mid x)) - E_q(\log q(z_1, z_2, z_3))$$

$$= E_q[\log \Pr(z_2) + \log \Pr(z_3 \mid z_2) + \log \Pr(z_1 \mid z_2, z_3)]$$

$$- E_{q(z_1)} \log q(z_1) - E_{q(z_2)} \log q(z_2) - E_{q(z_3)} \log q(z_3)$$

These are the terms only involving $q(z_1)$

Do update $q^*(z_1) = \cdots$

2) Update $q(z_2)$: same strategy but now, we REWRITE as

$$E_q[\log P(z_1) + \log \Pr(z_3 \mid z_1) + \log \Pr(z_2 \mid z_1, z_3)]$$

Again only one term involves $q(z_2)$!!

Update $q^*(z_2) \cdots$

3) Update $q(z_3) \cdots$ by similar tricks.

Therefore by putting the margin z_k at the last position in the chain rule, ONE term from the chain rule is enough.