CAVI

Wednesday, November 2, 2022 11:05 AM

At the end of the Algorithm (CAVI):

Use $q_{\hat{\theta}}(\vec{z}) = \prod_{t=1}^{T} q_t(z_t ; \hat{\theta})$ as approximation to $p_\alpha(\vec{z}|\vec{x})$

Three cases to consider:

① variable marginals: $q_t(z_t;\hat{\theta}) \approx p_\alpha(z_t|\vec{x}) = \sum_{\vec{z}_{\neg t}} p_\alpha(z_t, \vec{z}_{\neg t}|\vec{x})$

② two-variable marginal:

$$q(z_t, z_s; \hat{\theta}) = \underline{q(z_t;\hat{\theta})\, q(z_s;\hat{\theta})} \approx p_\alpha(z_t, z_s|\vec{x})$$

Q1 in L17 poll

③ three-variable marginal:

$$q(z_t, z_s, z_r) = q(z_t)\,q(z_s)\,q(z_r)$$

Why? * thanks to our M.F. assumption

$$q(z_1, z_2) = \sum_{z_3}\sum_{z_4}\cdots\sum_{z_T} q(z_1, z_2, \dots, z_T)$$

$$= \sum_{z_3}\cdots\sum_{z_T} \overbrace{\prod_{t=1}^{T} q_t(z_t)}^{q_1(z_1)q_2(z_2)\cdots q_T(z_T)}$$

$$= q_1(z_1)\,q_2(z_2) \overbrace{\sum_{z_3} q_3(z_3) \overbrace{\sum_{z_4} q_4(z_4) \cdots \overbrace{\sum_{z_T} q_T(z_T)}^{1}}^{1}}^{1}$$

$$= q_1(z_1)\, q_2(z_2)$$

CAVI Update Rule

Suppose we are given a factor graph over discrete random variables, $p_\alpha(z|x)$

$\Longrightarrow$ full conditionals: $p_\alpha(z_t|z_{\neg t}, \vec{x}) \propto \underbrace{\prod_{c\in N(z_t)} \Psi_c(\vec{z}_c)}$

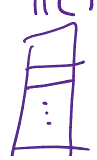
$\Longrightarrow$ full conditionals are categorical distributions

Case #1: Exponential Family

↳ not covered today...

Case #2: Multinomial

$\pi(\cdot|\vec{z}_{\neg t}, \vec{x})$

- z_t is discrete and finite valued
- suppose $p(z_t|\vec{z}_{\neg t}, \vec{x}) \triangleq \pi(z_t|\vec{z}_{\neg t}, \vec{x}) =$ 

- z_t is discrete and finite valued
- suppose $p(z_t|\vec{z}_{\neg t},\vec{x}) \triangleq \pi(z_t|\vec{z}_{\neg t},\vec{x}) = \begin{bmatrix} \\ \vdots \\ \\ \end{bmatrix}$
- then the CAVI updates are

$$q_t(z_t) = \frac{1}{Z}\exp\left(E_{q_{\neg t}}\left[\log p_\alpha(z_t|\vec{z}_{\neg t},\vec{x})\right]\right)$$

normalizing constant for q_t

$$= \frac{1}{Z}\exp\left(E_{q_{\neg t}}\left[\log \pi(z_t|\vec{z}_{\neg t},\vec{x})\right]\right)$$

$$= \frac{1}{Z}\exp\left(\sum_{\vec{z}_{\neg t}} q(\vec{z}_{\neg t}) \log \pi(z_t|\vec{z}_{\neg t},\vec{x})\right) \qquad \prod_{s:t\neq s} q_s(z_s)$$

for a factor graph

$$= \frac{1}{Z}\exp\left(C + \sum_{\vec{z}_{MB(z_t)}} \prod_{s\in MB(z_t)} q_s(z_s) \log\left[\prod_{c\in N(z_t)} \psi_c(\vec{z}_c)\right]\right)$$

$$= \frac{1}{Z}\exp\left(\sum_{\vec{z}_{MB(z_t)}} \underbrace{\prod_{s\in MB(z_t)} q_s(z_s)}\sum_{c\in N(z_t)} \log\psi_c(\vec{z}_c)\right)$$

easy to compute assuming #neighbors is not too large

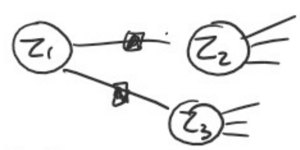
Where is C?

$$\frac{1}{Z}\exp(a+b) = \frac{\exp(a)\exp(b)}{Z} = \frac{\exp(b)}{Z'}$$

$$= \frac{1}{Z}\exp\left(\sum_{\vec{z}_{MB(z_t)}} \sum_{c\in N(z_t)} \log\psi_c(\vec{z}_c) \prod_{s\in MB(z_t)} q_s(z_s)\right)$$

Pairwise MRF Update Rule

example from slide...



$$q_1(z_1) \propto \exp\left(\sum_{z_2}\sum_{z_3} q_2(z_2)\, q_3(z_3) \log \psi_{12}(z_1,z_2) \log \psi_{13}(z_1,z_3)\right)$$

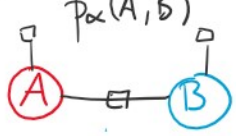
$$= \Big(\underbrace{\sum_{z_2} q_2(z_2)\log\psi_{12}(z_1,z_2)}\Big)\Big(\underbrace{\sum_{z_3} q_3(z_3)\log\psi_{13}(z_1,z_3)}\Big)\Big)$$

$$= \mu_{2\to 1}(z_1)\ \mu_{3\to 1}(z_1)$$

Ex: Two Variable Factor Graph

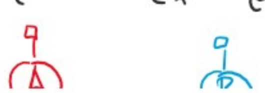
Joint:

$p_\alpha(A,B)$

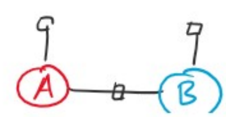


M.F. Approx

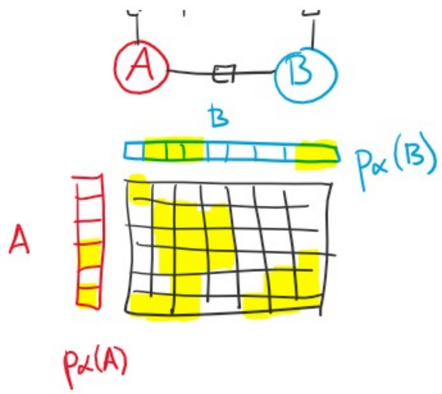
$q_\theta(A,B) = q_{\theta_A}(A)\, q_{\theta_B}(B)$

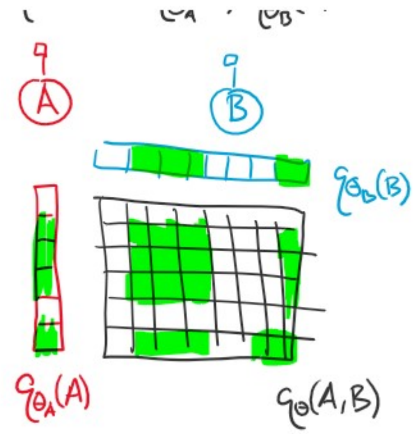


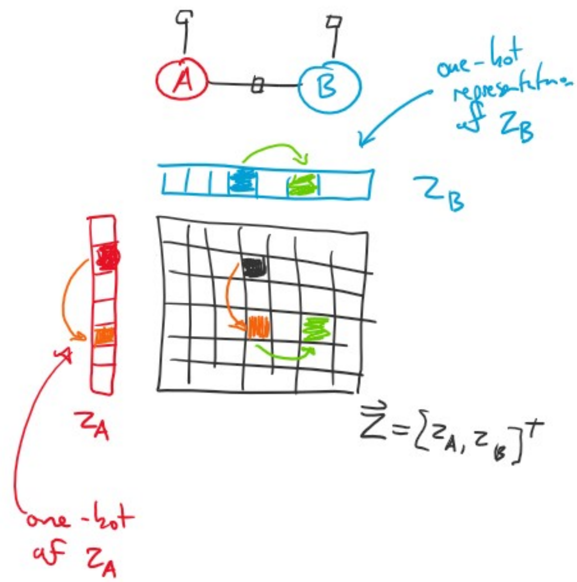
Gibbs Sampler



one-hot representation


A
B
B
pα(B)
A
pα(A)


A
B
qθB(B)
qθA(A)
qθ(A,B)


A
B
one-hot representation of ZB
ZB
ZA
one-hot of ZA
Z = [zA, zB]T