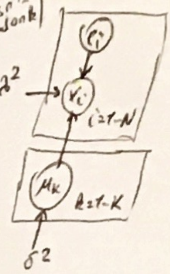
~~Gibbs Sampling~~

Assignement 1

- $q(\mu, c) = \prod_{k=1}^{K} q(\mu_k) \cdot \prod_{i=1}^{N} q(c_i)$
- $p(\mu, c, x) = \prod_{k=1}^{K} p(\mu_k) \prod_{i=1}^{N} p(c_i)\, p(x_i \mid c_i, \mu)$

mean of all means ($\frac{1}{K}\sum \mu_k$ ← doesn't depend on k)



$\mathcal{L}$

ELBO

$$\mathcal{L}(x \mid m, s^2, \phi) = \mathbb{E}_q[\log p(x, \mu, z)] - \mathbb{E}_q[\log q(\mu, c)]$$

$$= \mathbb{E}_q\left[\sum_{k=1}^{K} \log(p(\mu_k)) + \sum_{i=1}^{N} \log\Big(p(c_i)\Big) + \log\Big(p(x_i \mid c_i, \mu)\Big)\right]$$

$$- \mathbb{E}_q\left[\sum_{k=1}^{K} \log(q(\mu_k)) + \sum_{i=1}^{N} \log(q(c_i))\right]$$

$$= \underbrace{\sum_{k=1}^{K} \mathbb{E}_q[\log(p(\mu_k))]}_{1} + \underbrace{\sum_{i=1}^{N} \mathbb{E}_q[\log(p(c_i))]}_{2} + \underbrace{\sum_{i=1}^{N} \mathbb{E}_q[\log(p(x_i \mid c_i, \mu))]}_{3} - \underbrace{\sum_{k=1}^{K} \mathbb{E}_q[\log q(\mu_k)]}_{4} - \underbrace{\sum_{i=1}^{N} \mathbb{E}_q[\log(q(c_i))]}_{5}$$

We know in addition :

N ← nr of samples
K ← nr of clusters

a) the form of $q(\mu_k) = \mathcal{N}(\mu_k \mid m_k, s_k^2 I)$

b) form of $q(c_i) \sim \text{Categorical}(\phi_i)$, $\phi_i = \{\phi_{i,1} \ldots \phi_{i,K}\}$

and

c) $p(\mu_k) = \mathcal{N}(\alpha, \sigma^2 I)$

d) $p(c_i) = \text{Categorical}(\frac{1}{K}, \ldots, \frac{1}{K})$

e) $p(x_i \mid c_i, \mu) = \mathcal{N}(\mu_{c_i}, \lambda^2 I)$.

(1) $\mathbb{E}_q\left[\log p(\mu_k)\right] = \int_{\mu_1 \dots \mu_K} \sum_{c_1 - c_N} \prod_{i=1}^{N} q(c_i) \prod_{k=1}^{K} q(\mu_k) \log(p(\mu_k)) \, d\mu_1 \dots \mu_K.$

$= \int_{\mu_k} q(\mu_k) \log p(\mu_k) \, d\mu_k$

$\log p(\mu_k) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2}\frac{\mu_k^T\mu_k}{\sigma^2} + \frac{\mu_k^T\alpha}{\sigma^2} - \frac{\alpha^T\alpha}{2\sigma^2}$

$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{\alpha^T\alpha}{2\sigma^2} + \frac{\alpha^T}{\sigma^2}\overbrace{\int q(\mu_k)\mu_k \, d\mu_k}^{\mathbb{E}_q(\mu_k)} - \frac{1}{2\sigma^2}\overbrace{\int q(\mu_k)\mu_k^T\mu_k \, d\mu_k}^{\mathbb{E}_{q_k}[\mu_k^2]}$

$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{\alpha^T\alpha}{2\sigma} + \frac{\alpha^T m_k}{\sigma^2} - \frac{m_k^T m_k + s_k^2}{2\sigma^2}$

$\mathbb{E}(\mu_k^T\mu_k) = 2s_k^2 + m_k^T m_k$

$\boxed{(1) = K\left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{\alpha^T\alpha}{2\sigma}\right) + \frac{\alpha^T}{\sigma^2}\sum_{k=1}^{K} m_k - \frac{1}{2\sigma^2}\sum_{k=1}^{K} (m_k^T m_k + s_k^2)}$

(2) $\mathbb{E}_q\left[\log p(c_i)\right] = \log p(c_i) = -\log(K)$ ← since $p(c_i) = \frac{1}{K}$ $\forall c_i = 1 \dots K$

↳ $\boxed{(2) = -N\log(K)}$

(4) $\mathbb{E}_q\left[\log q(\mu_k)\right] = \int_{\mu_k} q(\mu_k) \log q(\mu_k) \, d\mu_k.$

$\log q(\mu_k) = -\frac{1}{2}\log(2\pi s_k^2) - \frac{1}{2}\frac{\mu_k^T\mu_k}{s_k^2} - \frac{m_k^T m_k}{2s_k^2} + \frac{\mu_k^T m_k}{s_k^2}$

$= -\frac{1}{2}\log(2\pi s_k^2) - \frac{m_k^T m_k}{2s_k^2} + \frac{m_k^T m_k}{s_k^2} - \frac{m_k^T m_k + s_k^2}{2s_k^2}$

$= -\frac{1}{2}\log(2\pi s_k^2) - \frac{1}{2}$

$\boxed{(4) \to -\frac{K}{2} - \frac{1}{2}\sum_{k=1}^{K}\log(2\pi s_k^2)}$

(5) $\mathbb{E}_q\left[\log(q(c_i))\right] = \sum_{c_1 - c_N}\prod_{k=1}^{K} q(c_k)\log(q(c_i))]$

$= \sum_{k=1}^{K}\phi_{i,k}\log(\phi_{i,k})$

↳ $\boxed{(5) = \sum_{i=1}^{N}\sum_{k=1}^{K}\phi_{i,k}\log(\phi_{i,k})}$

③

$$\mathbb{E}_q\left[\log(p(x_i \mid c_i, \boldsymbol{\mu}))\right] = \int \left(\sum_{c_1 \dots c_N} \left[\prod_{i=1}^{N} q(c_i)\right] \left[\prod_{k=1}^{K} q(\mu_k)\right] \log N(x_i ; \mu_{[c_i]}, \lambda^2 I) \right) d\mu_1 \dots d\mu_K)$$

$$= -\frac{1}{2}\log(2\pi\lambda^2) - \frac{1}{2}\frac{x_i^T x_i}{\lambda^2} + \int \sum_{c_i} q(c_i)\, q(\mu_{[c_i]}) \left(\frac{x_i^T \mu_{[c_i]}}{\lambda^2} - \frac{\mu_{[c_i]}^T \mu_{[c_i]}}{2\lambda^2} \right) d\mu_{[c_i]}$$

$$= -\frac{1}{2}\log(2\pi\lambda^2) - \frac{1}{2}\frac{x_i^T x_i}{\lambda^2} + \int \sum_{k=1}^{K} \phi_{i,k}\, q(\mu_k) \left(\frac{x_i^T \mu_k}{\lambda^2} - \frac{\mu_k^T \mu_k}{2\lambda^2} \right) d\mu_k.$$

$$= \text{— '' —} + \sum_{k=1}^{K} \phi_{i,k} . \int q(\mu_k) \left(\frac{x_i^T \mu_k}{\lambda^2} - \frac{\mu_k^T \mu_k}{2\lambda^2} \right) d\mu_k.$$

$$= -\frac{1}{2}\log(2\pi\lambda^2) - \frac{1}{2}\frac{x_i^T x_i}{\lambda^2} + \sum_{k=1}^{K} \phi_{i,k} \left(\frac{x_i^T m_k}{\lambda^2} - \frac{2 s_k^2 + m_k^T m_k}{2\lambda^2} \right)$$

$$\hookrightarrow \quad ③ = -\frac{N}{2}\log(2\pi\lambda^2) - \frac{1}{2\lambda^2}\sum_{i=1}^{N} x_i^T x_i + \sum_{k=1}^{K} \frac{m_k^T \sum_{i=1}^{N} x_i \phi_{i,k}}{\lambda^2} - \underbrace{\sum_{i=1}^{n}\sum_{k=1}^{K} \frac{2 s_k^2 + m_k^T m_k}{2\lambda^2}\phi_{i,k}}_{\frac{1}{2\lambda^2}\sum_{k=1}^{K}(2 s_k^2 + m_k^T m_k)\sum_{i=1}^{n}\phi_{i,k}}$$

$$\log N(x_i ; \mu_{[c_i]}, \lambda^2 I) = -\frac{1}{2}\log(2\pi\lambda^2) - \frac{1}{2}\frac{x_i^T x_i}{\lambda^2} + \frac{\mu_{[c_i]}^T x_i}{\lambda^2} - \frac{\mu_{[c_i]}^T \mu_{[c_i]}}{2\lambda^2}$$

Assignement 2

we have $q^*(c_i) \propto \exp\{\mathbb{E}_{q(c_{-i},\mu)}[\log p(\overset{\text{all data}}{x}, c_i, \overset{\text{all means}}{c_{-i}}, \mu)]\}$

Where the log distribution is:

$$\log(p(x, c_i, c_{-i}, \mu)) = \log p(\mu) + \sum_{j \neq i}(\log p(c_j) + \log p(x_j | c_j, \mu)) + \log p(c_i) + \log p(x_i | c_i, \mu)$$

$$p(\mu, c, x) = \prod_{k=1}^{K} p(\mu_k) \prod_{i=1}^{N} p(c_i)\, p(x_i | c_i, \mu)$$

After ignoring everything that doesn't depend on c_i and taking the expectation (which has no effect) we get:

$$q^*(c_i) \propto \exp\{\log(p(c_i)) + E[\log(p(x_i | c_i, \mu))]\}$$

$$q(c_i) \propto \exp\{\log(\tfrac{1}{K}) + \mathbb{E}_{q(c_i,\mu)}[\log p(x_i | c_i, \mu)]\}$$

$$q(c_i) \propto \exp\left(2\frac{x_i^T m_{(c_i)}}{2\lambda^2} - p\left(\frac{s_{(c_i)}^2}{\lambda^2} + \frac{m_{(c_i)}^T m_{(c_i)}}{2\lambda^2}\right)\right)$$

$$\Rightarrow \phi_{i,k} \propto \exp\left(\frac{x_i^T m_k}{\lambda^2} - \underbrace{p\left(\frac{2 s_k^2 + m_k^T m_k}{2\lambda^2}\right)}_{\mathbb{E}_{q(\mu_k)}(\mu_k^T \mu_k)}\right)$$

$\boxed{c_i}$

★ $\mathbb{E}_{q(c\ldots)}$

$$\mathbb{E}_{q(c_{-i},\mu)}[\log p(x_i | c_i, \mu)] \propto \int q(\mu_{(c_i)}) \left(\frac{2 x_i^T \mu_{(c_i)}}{2\lambda^2} - \frac{\mu_{(c_i)}^T \mu_{(c_i)}}{2\lambda^2}\right) d\mu_{(c_i)} \cancel{- \frac{p}{2}\log(\lambda^2)} \cancel{- \frac{x^T x}{2\lambda^2}}$$

$$\propto \frac{2 x_i^T m_{(c_i)}}{2\lambda^2} - p\left(\frac{s_{(c_i)}^2}{\lambda^2} + \frac{m_{(c_i)}^T m_{(c_i)}}{2\lambda^2}\right).$$

Assignement 3

$$q^*(\mu_k) \propto \exp\left(\log p(\mu_k) + \sum_{i=1}^{N} \mathbb{E}_{q(\mu_{-k}, c_i)}[\log p(x_i | c_i, \mu)]\right).$$

$$= \exp\left(-\frac{\mu_k^T \mu_k}{2\delta^2} + \sum_{i=1}^{N} \phi_{i,k}\left(-\frac{(x_i - \mu_k)^T (x_i - \mu_k)}{2\lambda^2}\right) + \text{const}\right)$$

$$= \exp\left(\frac{-\mu_k^T \mu_k}{2\delta^2} \overbrace{- \sum_{i=1}^{N} \phi_{i,k} \frac{x_i^T x_i}{2\lambda^2}}^{\text{const}} + 2\sum_{i=1}^{N} \phi_{i,k} \frac{x_i^T \mu_k}{2\lambda^2} - \frac{\mu_k^T \mu_k}{2\lambda^2} \sum_{i=1}^{N} \phi_{i,k} + \text{const}\right)$$

$$= \exp\left(-\frac{\mu_k^T \mu_k}{2}\left(\frac{1}{\delta^2} + \frac{\sum_{i=1}^{N} \phi_{i,k}}{\lambda^2}\right) + \frac{2\left(\sum_{i=1}^{N} \phi_{i,k} x_i\right)^T \mu_k}{2\lambda^2} + \text{const}\right).$$

$$\frac{-\mu_k^T \mu_k}{2\delta^2}\left(1 + \frac{\delta^2}{\lambda^2}\sum^{N} \phi_{i,k}\right)$$

On the other hand we have

$$N(m_k, s_k^2 I) \propto \exp\left(-\frac{(\mu_k - m_k)^T (\mu_k - m_k)}{2 s_k^2}\right)$$

$$\propto \exp\left(-\frac{\mu_k^T \mu_k}{2 s_k^2} + 2\frac{\mu_k^T m_k}{2 s_k^2} + \frac{m_k^T m_k}{2 s_k^2}\right)$$

$$\Rightarrow \frac{1}{2 s_k^2} = \frac{1}{2\delta^2}\left(1 + \frac{\delta^2}{\lambda^2}\sum_{i=1}^{N} \phi_{i,k}\right) \Rightarrow \boxed{s_k^2 = \frac{\delta^2}{1 + \sum_{i=1}^{N} \phi_{i,k} \frac{\delta^2}{\lambda^2}}}$$

$$= \frac{1}{1/\delta^2 + \sum_{i=1}^{N} \phi_{i,k} \cdot \frac{1}{\lambda^2}}$$

$$\boxed{\cancel{m_k = \sum_{i=1}^{N} \phi_{i,k} x_i}}$$

$$\frac{2\left(\sum_{i=1}^{N} \phi_{i,k} x_i\right)^T \mu_k}{2\lambda^2} = \frac{2\mu_k^T m_k}{2 s_k^2}$$

$$\Rightarrow \boxed{m_k = s_k \sum_{i=1}^{N} \phi_{i,k} x_i}$$

!!