Exercise Sheet 11

Robin Schmidt
Probabilistic Inference & Learning

January 16, 2019

EM & Variational Inference

1. EM for MAP

2. Softer k-means

$$\log p(x|\pi,\mu,\Sigma) = \sum_i^n \log\left(\sum_j^k \pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)\right)$$

$$\nabla_{\mu_j} \log p(x|\pi,\mu,\Sigma) = -\frac{1}{2}\sum_i^n \underbrace{\frac{\pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)}{\sum_{j'}, \pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)}}_{=:r_{ji}} \Sigma\left(x_i-\mu_j\right)$$

$$\mu_j = \frac{1}{R_j}\sum_i^n r_j x_i$$

$$\nabla_{\Sigma} \log p(x|\pi,\mu,\Sigma) = -\frac{1}{2}\sum_i^n \underbrace{\frac{\pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)}{\sum_{j'}, \pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)}}_{=:r_{ji}} \left(x_i-\mu_j\right)\left(x_i-\mu_j\right)^\top$$

$$\Sigma = \frac{1}{R_j}\sum_i^n r_{ji}\left(x_i-\mu_j\right)\left(x_i-\mu_j\right)^\top$$

$$\nabla_{\pi_j}\left[\log p(x|\pi,\mu,\Sigma) + \lambda\left(\sum_j \pi_j - 1\right)\right] = \sum_i^n \frac{\mathcal{N}\left(x_i;\mu_j,\Sigma\right)}{\sum_{j'} \pi_j \mathcal{N}\left(x_i;\mu_j,\Sigma\right)} + \lambda$$

$$0 = \sum_i^n \pi_j \frac{\mathcal{N}\left(x_i;\mu_j,\Sigma\right)}{\sum_{j'} \pi_j, \mathcal{N}\left(x_i;\mu_j,\Sigma\right)} + \lambda\pi_j = \sum_i^n r_{ij} + \lambda\pi_j$$

$$\pi_j = \frac{\sum_i r_{ij}}{n} =: \frac{R_j}{n}$$

3. Gibbs' inequality

$$
\begin{aligned}
\int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx &\geq \log \int p(x) \frac{q(x)}{p(x)} dx \\
\int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx &\geq \log \int q(x) \\
\int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx &\geq \log 1 \\
\int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx &\geq 0
\end{aligned}
$$

4. Free energy for Gaussians

(a)

$$
\begin{aligned}
D_{\mathrm{KL}} \left(\mathcal{N} \left(x; 0, \sigma_q^2 \right) \| \mathcal{N} \left(x; 0, \sigma_p^2 \right) \right) &= \int q(x) \log \frac{\frac{1}{\sqrt{2\pi\sigma_q^2}} \exp\left\{ -\frac{(x-0)^2}{2\sigma_q^2} \right\}}{\frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left\{ -\frac{(x-0)^2}{2\sigma_p^2} \right\}} dx \\
&= \int q(x) \log \left(\frac{\sigma_q^2}{\sigma_p^2} \right) dx + \int q(x) \left(\frac{-(x-0)^2}{2\sigma_q^2} + \frac{(x-0)^2}{2\sigma_p^2} \right) dx \\
&= \frac{1}{2} \log \left(\frac{\sigma_q^2}{\sigma_p^2} \right) + \frac{1}{2\sigma_q^2} \left(- \int (x-0)^2 q(x) dx \right) + \frac{1}{2\sigma_p^2} \left(\int (x-0)^2 q(x) dx \right)
\end{aligned}
$$

Using the property that $\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2$:

$$
\begin{aligned}
&= \frac{1}{2} \log \left(\frac{\sigma_q^2}{\sigma_p^2} \right) - \frac{\sigma_q^2}{2\sigma_q^2} + \frac{\sigma_q^2}{2\sigma_p^2} \\
&= \frac{1}{2} \left(\log \left(\frac{\sigma_q^2}{\sigma_p^2} \right) - 1 + \frac{\sigma_q^2}{\sigma_p^2} \right)
\end{aligned}
$$