Homework 2

Andrey Savinov

September 29, 2019

1 Problem

Hidden variables are $Z = \{z_{nk}\}_{n=1,k=1}^{N,K}$, $U = \{u_{nk}\}_{n=1,k=1}^{N,K}$ and $X = \{x_{nk}\}_{n=1,k=1}^{N,K}$. Global variables are $\Theta = \{\pi_k, \nu_k, W_k, \mu_k\}_{k=1}^{K}$.

- E-step

$$q(x_n, u_n, z_n) = \arg\max_q \mathcal{L}(q, \Theta) = p(x_n, u_n, x_n|y_n, \Theta)$$

$$p(x_n, u_n, z_n|y_n, \Theta) = P(y_n|x_n, u_n, z_n, \Theta)P(x_n, u_n, z_n|\Theta) = p(y_n|x_n, u_n, z_n, \Theta)p(x_n, u_n, z_n|\Theta) =$$
$$= p(y_n|x_n, u_n, z_n)p(x_n|u_n, z_n)p(u_n|z_n)p(z_n)$$

$$q_n(k = i) = q_{ni} = \frac{1}{Z}\pi_i \ Gamma(u_{ni}, \frac{\nu_i}{2}, \frac{\nu_i}{2})\mathcal{N}(x_{ni}|0, u_{ni}^{-1}I_d)\mathcal{N}(y_n|W_i x_{ni} + \mu_i, u_{ni}^{-1}I_D)$$

$$Z = \sum_{i=1}^{K} q_{ni}$$

- M-step

$$\Theta = \arg\max_\Theta \mathcal{L}(q, \Theta) = \arg\max_\Theta E_q \log p(y_n, x_n, u_n, z_n|\Theta)$$

$$E_q \log p(y_n, x_n, u_n, z_n|\Theta) = \sum_{n=1}^{N} E_q \log p(y_n, x_n, u_n, z_n|\Theta) = \sum_{n=1}^{N}\sum_{k=1}^{K} q_{nk} \log p(y_n|x_{nk}, u_{nk}, z_{nk}, \Theta)p(x_{nk}, u_{nk}, z_{nk}|\Theta) =$$
$$= \sum_{n=1}^{N}\sum_{k=1}^{K} q_{nk} \log p(y_n|x_{nk}, u_{nk}, z_{nk}, \Theta)p(x_{nk}|u_{nk}, z_{nk}, \Theta)p(u_{nk}|z_{nk}, \Theta)p(z_{nk}|\Theta) =$$
$$= \sum_{n=1}^{N}\sum_{k=1}^{K} \pi_k \ Gamma(u_{nk}, \frac{\nu_k}{2}, \frac{\nu_k}{2})\mathcal{N}(x_{nk}|0, u_{nk}^{-1}I_d)\mathcal{N}(y_n|W_k x_{nk} + \mu_k, u_{nk}^{-1}I_D)\cdot$$
$$\cdot \log \pi_k \ Gamma(u_{nk}, \frac{\nu_k}{2}, \frac{\nu_k}{2})\mathcal{N}(x_{nk}|0, u_{nk}^{-1}I_d)\mathcal{N}(y_n|W_k x_{nk} + \mu_k, u_{nk}^{-1}I_D)$$

This expression should be optimized using first order condition with respect to global variables $\Theta = \{\pi_k, \nu_k, W_k, \mu_k\}_{k=1}^{K}$.

2 Problem

Z is normalization constant. $\phi(x)$ is log-concave

$$f(\mu, \Sigma) = Z \int_X \log(\phi(\omega^T x)) \exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu))dx$$

$$f(\lambda\mu_1 + (1 - \lambda)\mu_2, \Sigma) = Z \int_X \log(\phi(\omega^T x)) \exp(-\frac{1}{2}(x - \lambda\mu_1 - (1 - \lambda)\mu_2)^T \Sigma^{-1}(x - \lambda\mu_1 - (1 - \lambda)\mu_2))dx$$

$$y = x - \lambda\mu_1 - (1 - \lambda)\mu_2, \ x = y + \lambda\mu_1 + (1 - \lambda)\mu_2, \ dx = dy$$

$$f(\lambda\mu_1 + (1 - \lambda)\mu_2, \Sigma) = Z \int_X \log(\phi(\omega^T [y + \lambda\mu_1 + (1 - \lambda)\mu_2])) \exp(-\frac{1}{2}y^T \Sigma^{-1} y)dy \geq$$

$$Z \geq \int_X [Z_1 \lambda \log(\phi(\omega^T[y+\mu_1])) + Z_2(1-\lambda)\log(\phi(\omega^T[y+\mu_2]))] \exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy =$$

$$= Z_1\lambda \int_X \log(\phi(\omega^T[y+\mu_1])) \exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy + Z_2(1-\lambda)\int_X \log(\phi(\omega^T[y+\mu_2])) \exp(-\frac{1}{2}y^T\Sigma^{-1}y)dy$$

$$= \lambda f(\mu_1, \Sigma) + (1-\lambda) f(\mu_2, \Sigma)$$

Hence, $f(\mu, \Sigma)$ is concave over μ.

$$f(\mu, \lambda\Sigma_1 + (1-\lambda)\Sigma_2) = Z\int_X \log(\phi(\omega^T x)) \exp(-\frac{1}{2}(x-\mu)^T(\lambda\Sigma_1 + (1-\lambda)\Sigma_2)^{-1}(x-\mu))dx$$

$$y = \sqrt{(\lambda\Sigma_1 + (1-\lambda)\Sigma_2)^{-1}}(x-\mu),\ x = \sqrt{\lambda\Sigma_1 + (1-\lambda)\Sigma_2}y + \mu,\ dx = \sqrt{\lambda\Sigma_1 + (1-\lambda)\Sigma_2}dy$$

$$f(\mu, \lambda\Sigma_1 + (1-\lambda)\Sigma_2) = Z\int_X \log(\phi(\omega^T[\sqrt{\lambda\Sigma_1 + (1-\lambda)\Sigma_2}y + \mu])) \exp(-\frac{1}{2}y^Ty)\sqrt{(\lambda\Sigma_1 + (1-\lambda)\Sigma_2)}dy \geq$$

$\sqrt{\cdot}$ is concave function, consequently

$$\geq Z_1\lambda \int_X \log(\phi(\omega^T[\sqrt{\Sigma_1}y + \mu])) \exp(-\frac{1}{2}y^Ty)\sqrt{\Sigma_1}dy + Z_2(1-\lambda)\int_X \log(\phi(\omega^T[\sqrt{\Sigma_2}y + \mu])) \exp(-\frac{1}{2}y^Ty)\sqrt{\Sigma_2}dy =$$

$$= \lambda f(\mu, \Sigma_1) + (1-\lambda) f(\mu, \Sigma_2)$$

Hence, $f(\mu, \Sigma)$ is concave over Σ.