CS446: Machine Learning, Fall 2017, Homework 4

Name: Lanxiao Bai (lbai5)

Worked individually

Problem 1

Solution:

$$
\begin{aligned}
\log L_I(\Psi) &= \log p(\mathbf{y} \mid \Psi) \\
&= \log \sum_{k=1}^{2} \pi_k \mathcal{N}(\mathbf{y} \mid \Psi) \\
&= \log \sum_{k=1}^{2} \pi_k \mathcal{N}(\mathbf{y} \mid \Psi) \\
&= \sum_{i=1}^{n} \log \left\{ \sum_{k=1}^{2} \log \pi_k \mathcal{N}(\mathbf{y}_i \mid \Psi) \right\}
\end{aligned}
$$

Problem 2

Solution:

$$
\begin{aligned}
\log L_C(\Psi) &= \log p(\mathbf{W} \mid \Psi) \\
&= \log \mathcal{N}(\mathbf{W} \mid \Psi) \\
&= \log \prod_{i=1}^{n} \mathcal{N}(\mathbf{W}_i \mid \Psi) \\
&= \sum_{i=1}^{n} \log \mathcal{N}(\mathbf{W}_i \mid \Psi) \\
&= \sum_{i=1}^{n} \log \frac{1}{2\pi} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\mathbf{W}_i - \mu)^T \Sigma^{-1} (\mathbf{W}_i - \mu) \right) \\
&= \sum_{i=1}^{n} \left\{ -\log 2\pi - \frac{1}{2} \log \xi - \frac{1}{2} Tr\left(\Sigma^{-1} (\mathbf{W}_i - \mu)(\mathbf{W}_i - \mu)^T \right) \right\} \\
&= -n \log 2\pi - \frac{n}{2} \log \xi - \frac{1}{2} \sum_{i=1}^{n} Tr\left(\Sigma^{-1} (\mathbf{W}_i - \mu)(\mathbf{W}_i - \mu)^T \right) \\
&= -n \log 2\pi - \frac{n}{2} \log \xi - \frac{1}{2} Tr\left(\Sigma^{-1} \sum_{i=1}^{n} [(\mathbf{W}_i - \mu)(\mathbf{W}_i - \mu)^T] \right)
\end{aligned}
$$

$$= -n \log 2\pi - \frac{n}{2} \log \xi - \frac{1}{2} Tr \left(\frac{1}{\xi} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \sum_{i=1}^{n} [(\mathbf{W}_i - \mu)(\mathbf{W}_i - \mu)^T] \right)$$
$$= -n \log 2\pi - \frac{n}{2} \log \xi - \frac{1}{2} Tr \left(\frac{1}{\xi} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \begin{bmatrix} T_{11} - \mu_1^2 & T_{12} - \mu_1\mu_2 \\ T_{12} - \mu_1\mu_2 & T_{22} - \mu_2^2 \end{bmatrix} \right)$$

Problem 3

Solution: Since

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{W}_i$$

and

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{W}_i - \hat{\mu})(\mathbf{W}_i - \hat{\mu})^T$$

we have

$$\begin{aligned} \hat{\mu}_1 &= \frac{1}{n} \sum_{i=1}^{n} w_{1i} \\ &= \frac{T_1}{n} \end{aligned}$$

$$\begin{aligned} \hat{\mu}_2 &= \frac{1}{n} \sum_{i=1}^{n} w_{2i} \\ &= \frac{T_2}{n} \end{aligned}$$

$$\begin{aligned} \hat{\Sigma} &= \frac{1}{n} \sum_{i=1}^{n} [w_{i1} - \hat{\mu}_1, w_{i2} - \hat{\mu}_2]^T [w_{i1} - \hat{\mu}_1, w_{i2} - \hat{\mu}_2] \\ &= \frac{1}{n} \sum_{i=1}^{n} \begin{bmatrix} (w_{i1} - \hat{\mu}_1)^2 & (w_{i1} - \hat{\mu}_1)(w_{i2} - \hat{\mu}_2) \\ (w_{i1} - \hat{\mu}_1)(w_{i2} - \hat{\mu}_2) & (w_{i1} - \hat{\mu}_1)^2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n} \sum_{i=1}^{n} (w_{i1} - \hat{\mu}_1)^2 & \frac{1}{n} \sum_{i=1}^{n} (w_{i1} - \hat{\mu}_1)(w_{i2} - \hat{\mu}_2) \\ \frac{1}{n} \sum_{i=1}^{n} (w_{i1} - \hat{\mu}_1)(w_{i2} - \hat{\mu}_2) & \frac{1}{n} \sum_{i=1}^{n} (w_{i1} - \hat{\mu}_1)^2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{E}[(w_1 - \hat{\mu}_1)^2] & \mathbb{E}[(w_1 - \hat{\mu}_1)(w_2 - \hat{\mu}_2)] \\ \mathbb{E}[(w_1 - \hat{\mu}_1)(w_2 - \hat{\mu}_2)] & \mathbb{E}[(w_2 - \hat{\mu}_2)^2] \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{E}[w_1^2] - \hat{\mu}_1^2 & \mathbb{E}[w_1 w_2] - \hat{\mu}_1\hat{\mu}_2 \\ \mathbb{E}[w_1 w_2] - \hat{\mu}_1\hat{\mu}_2 & \mathbb{E}[w_2^2] - \hat{\mu}_2^2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{T_{11}}{n} - \frac{T_1^2}{n^2} & \frac{T_{12}}{n} - \frac{T_1 T_2}{n^2} \\ \frac{T_{12}}{n} - \frac{T_1 T_2}{n^2} & \frac{T_{22}}{n} - \frac{T_2^2}{n^2} \end{bmatrix} \end{aligned}$$

We get that

$$\hat{\sigma}_{11} = \frac{T_{11}}{n} - \frac{T_1^2}{n^2}$$
$$\hat{\sigma}_{12} = \frac{T_{12}}{n} - \frac{T_1 T_2}{n^2}$$
$$\hat{\sigma}_{22} = \frac{T_{22}}{n} - \frac{T_2^2}{n^2}$$

Problem 4

Solution:

$$\begin{aligned}
T_{11}^{(k)} &= \mathbb{E}_{\Psi^{(k)}}[T_{11} \mid \mathbf{y}] \\
&= \sum_{i=1}^{m} w_{i1}^2 + \sum_{i=m+1}^{m+m_1} \mathbb{E}_{\Psi^{(k)}}[w_{i1}^2 \mid w_{i2}] + \sum_{i=m+m_1+1}^{n} w_{i1}^2 \\
&= \sum_{i=1}^{m} w_{i1}^2 + \sum_{i=m+1}^{m+m_1} (\sigma_{11}^{(k)}(1-\rho^{2(k)}) + (\mu_1^{(k)} + \sigma_{12}^{(k)}\sigma_{22}^{-1(k)}(w_{i2}-\mu_2^{(k)}))^2) + \sum_{i=m+m_1+1}^{n} w_{i1}^2
\end{aligned}$$

$$\begin{aligned}
T_{22}^{(k)} &= \mathbb{E}_{\Psi^{(k)}}[T_{22} \mid \mathbf{y}] \\
&= \sum_{i=1}^{m_1} w_{i2}^2 + \sum_{i=m+m_1+1}^{n} \mathbb{E}_{\Psi^{(k)}}[w_{i2}^2 \mid w_{i1}] \\
&= \sum_{i=1}^{m_1} w_{i1}^2 + \sum_{i=m+m_1+1}^{n} (\sigma_{22}^{(k)}(1-\rho^{2(k)}) + (\mu_2^{(k)} + \sigma_{12}^{(k)}\sigma_{11}^{-1(k)}(w_{i1}-\mu_1^{(k)}))^2)
\end{aligned}$$

$$\begin{aligned}
T_{12}^{(k)} &= T_{21}^{(k)} = \mathbb{E}_{\Psi^{(k)}}[T_{12} \mid \mathbf{y}] \\
&= \sum_{i=1}^{m} w_{i1}w_{i2} + \sum_{i=m+1}^{m+m_1} \mathbb{E}_{\Psi^{(k)}}[w_{i1} \mid w_{i2}]w_{i2} + \sum_{i=m+m_1+1}^{n} w_{i1}\mathbb{E}_{\Psi^{(k)}}[w_{i2}^2 \mid w_{i1}] \\
&= \sum_{i=1}^{m} w_{i1}w_{i2} + \sum_{i=m+1}^{m+m_1} [(\sigma_{11}^{(k)}(1-\rho^{2(k)}) + (\mu_1^{(k)} + \sigma_{12}^{(k)}\sigma_{22}^{-1(k)}(w_{i2}-\mu_2^{(k)}))]w_{i2} + \\
&\sum_{i=m+m_1+1}^{n} [w_{i1}(\sigma_{22}^{(k)}(1-\rho^{2(k)}) + (\mu_2^{(k)} + \sigma_{12}^{(k)}\sigma_{11}^{-1(k)}(w_{i1}-\mu_1^{(k)}))]
\end{aligned}$$

Problem 5

Solution:

$Q(\Psi; \Psi^{(k)}) = \mathbb{E}_{\Psi^{(k)}}[\log(L_C(\Psi)) \mid \mathbf{y}]$

$$= -n\log 2\pi - \frac{n}{2}\log\xi - \frac{1}{2}Tr\left(\frac{1}{\xi}\begin{bmatrix}\sigma_{22}^{(k)} & -\sigma_{12}^{(k)}\\ -\sigma_{12}^{(k)} & \sigma_{11}^{(k)}\end{bmatrix}\begin{bmatrix}T_{11}^{(k)} - \mu_1^{2(k)} & T_{12}^{(k)} - \mu_1^{(k)}\mu_2^{(k)}\\ T_{12}^{(k)} - \mu_1^{(k)}\mu_2^{(k)} & T_{22}^{(k)} - \mu_2^{2(k)}\end{bmatrix}\right)$$

Problem 6

Solution:

$$\hat{\mu}_1^{(k+1)} = \frac{T_1^{(k)}}{n}$$

$$\hat{\mu}_2^{(k+1)} = \frac{T_2^{(k)}}{n}$$

$$\hat{\sigma}_{11}^{(k+1)} = \frac{T_{11}^{(k)}}{n} - \frac{T_1^{2(k)}}{n^2}$$
$$\hat{\sigma}_{12}^{(k+1)} = \frac{T_{12}^{(k)}}{n} - \frac{T_1^{(k)}T_2^{(k)}}{n^2}$$
$$\hat{\sigma}_{22}^{(k+1)} = \frac{T_{22}^{(k)}}{n} - \frac{T_2^{2(k)}}{n^2}$$