提出課題 V：混合モデルと EM

問題文中の空欄に入る数式を選択せよ．選択した解答（(A)–(D)）は「機械学習 2024 KA240201-teams」の「一般」チャネルに出現する課題へのリンクから提出すること．

問題 1，2

負担率 $\gamma(z_{nk})$ を固定した下で，

$$\mathbb{E}_{\mathbf{Z}}[\ln p(\mathbf{X},\mathbf{Z}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\pi})] = \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\{\ln \pi_k + \ln \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k,\boldsymbol{\Sigma}_k)\} \tag{a}$$

を $\boldsymbol{\mu}_k$ と $\boldsymbol{\Sigma}_k$ について最大化しようとすると，以下で与えられる陽な解

$$\boldsymbol{\mu}_k = \frac{1}{N_k}\sum_{n=1}^{N}\gamma(z_{nk})\mathbf{x}_n \tag{b}$$

$$\boldsymbol{\Sigma}_k = \frac{1}{N_k}\sum_{n=1}^{N}\gamma(z_{nk})(\mathbf{x}_n-\boldsymbol{\mu}_k)(\mathbf{x}_n-\boldsymbol{\mu}_k)^{\mathrm{T}}, \tag{c}$$

が得られることを示す．

式 (a) を負担率 $\gamma(z_{nk})$ を固定した下で $\boldsymbol{\mu}_k$ について微分すると，

$$\frac{\partial}{\partial \boldsymbol{\mu}_k}\mathbb{E}_{\mathbf{Z}}[\ln p(\mathbf{X},\mathbf{Z}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\pi})] = \boxed{\ (1)\ }.$$

となる．これが 0 に等しいとして整理すると，式 (b) が得られる．同様に，式 (a) を負担率 $\gamma(z_{nk})$ を固定した下で $\boldsymbol{\Sigma}_k^{-1}$ について微分すると，

$$\frac{\partial}{\partial \boldsymbol{\Sigma}_k^{-1}}\mathbb{E}_{\mathbf{Z}}[\ln p(\mathbf{X},\mathbf{Z}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\pi})] = \boxed{\ (2)\ }.$$

となる．これが 0 に等しいとして整理すると，式 (c) が得られる．

問題 1. 空欄 (1) に入る数式を選択せよ．

(A) $-\sum_{n=1}^{N}\gamma(z_{nk})(\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k - \boldsymbol{\Sigma}_k^{-1}\mathbf{x}_n)$

(B) $\sum_{n=1}^{N}\gamma(z_{nk})(\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k - \boldsymbol{\Sigma}_k^{-1}\mathbf{x}_n)$

(C) $-\sum_{n=1}^{N}\gamma(z_{nk})(\boldsymbol{\mu}_k\boldsymbol{\Sigma}_k^{-1} - \mathbf{x}_n\boldsymbol{\Sigma}_k^{-1})$

(D) $\sum_{n=1}^{N}\gamma(z_{nk})(\boldsymbol{\mu}_k\boldsymbol{\Sigma}_k^{-1} - \mathbf{x}_n\boldsymbol{\Sigma}_k^{-1})$

問題 2. 空欄 (2) に入る数式を選択せよ．

(A) $-\sum_{n=1}^{N}\gamma(z_{nk})\left(\boldsymbol{\Sigma}_k - (\mathbf{x}_n-\boldsymbol{\mu}_k)(\mathbf{x}_n-\boldsymbol{\mu}_k)^{\mathrm{T}}\right)/2$

(B) $\sum_{n=1}^{N} \gamma(z_{nk}) \left(\boldsymbol{\Sigma}_k - (\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^{\mathrm{T}} \right) / 2$

(C) $-\sum_{n=1}^{N} \gamma(z_{nk}) \left(\boldsymbol{\Sigma}_k^{-1} - ((\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^{\mathrm{T}})^{-1} \right) / 2$

(D) $\sum_{n=1}^{N} \gamma(z_{nk}) \left(\boldsymbol{\Sigma}_k^{-1} - ((\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^{\mathrm{T}})^{-1} \right) / 2$

問題 3，4

混合ガウス分布について，負担率を 1 つのデータ点 $\mathbf{x}_m$ のみについてしか更新しない逐次型 EM アルゴリズムを考え，対応する新旧の負担率の値を $\gamma^{\mathrm{new}}(z_{mk})$，$\gamma^{\mathrm{old}}(z_{mk})$ とする．ここで以下を仮定する．

$$
\begin{aligned}
N_k^{\mathrm{old}} &= \sum_n \gamma^{\mathrm{old}}(z_{nk}) \\
\boldsymbol{\mu}_k^{\mathrm{old}} &= \frac{1}{N_k^{\mathrm{old}}} \sum_n \gamma^{\mathrm{old}}(z_{nk}) \mathbf{x}_n \\
\boldsymbol{\Sigma}_k^{\mathrm{old}} &= \frac{1}{N_k^{\mathrm{old}}} \sum_n \gamma^{\mathrm{old}}(z_{nk}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{\mathrm{old}})(\mathbf{x}_n - \boldsymbol{\mu}_k^{\mathrm{old}})^{\mathrm{T}} \\
\pi_k^{\mathrm{old}} &= \frac{N_k^{\mathrm{old}}}{N} = \frac{1}{N} \sum_n \gamma^{\mathrm{old}}(z_{nk}) \\
N_k^{\mathrm{new}} &= N_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) - \gamma^{\mathrm{old}}(z_{mk}) \\
\boldsymbol{\mu}_k^{\mathrm{new}} &= \boldsymbol{\mu}_k^{\mathrm{old}} + \left(\frac{\gamma^{\mathrm{new}}(z_{mk}) - \gamma^{\mathrm{old}}(z_{mk})}{N_k^{\mathrm{new}}} \right) (\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}}).
\end{aligned}
$$

負担率を逐次的に更新する場合，M ステップにおいて混合ガウスモデルの共分散行列 $\boldsymbol{\Sigma}_k$ と混合係数 π_k を更新する式は，上記の定義を用いると，それぞれ $\boxed{(3)}$，$\boxed{(4)}$ となる．

問題 3. 空欄 (3) に入る数式を選択せよ．

(A) $\boldsymbol{\Sigma}_k^{\mathrm{new}} = \boldsymbol{\Sigma}_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{new}}$
$\qquad - \gamma^{\mathrm{old}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{old}}$

(B) $\boldsymbol{\Sigma}_k^{\mathrm{new}} = \boldsymbol{\Sigma}_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{old}}$
$\qquad - \gamma^{\mathrm{old}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{old}}$

(C) $\boldsymbol{\Sigma}_k^{\mathrm{new}} = \boldsymbol{\Sigma}_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{new}}$
$\qquad - \gamma^{\mathrm{old}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N_k^{\mathrm{new}}$

(D) $\boldsymbol{\Sigma}_k^{\mathrm{new}} = \boldsymbol{\Sigma}_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{new}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N$
$\qquad - \gamma^{\mathrm{old}}(z_{mk}) \left((\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})(\mathbf{x}_m - \boldsymbol{\mu}_k^{\mathrm{old}})^{\mathrm{T}} - \boldsymbol{\Sigma}_k^{\mathrm{old}} \right) / N$

問題 4. 空欄 (4) に入る数式を選択せよ．

(A) $\pi_k^{\mathrm{new}} = \pi_k^{\mathrm{old}} + \gamma^{\mathrm{new}}(z_{mk}) / N_k^{\mathrm{new}} - \gamma^{\mathrm{old}}(z_{mk}) / N_k^{\mathrm{old}}$

(B) $\pi_k^{\text{new}} = \pi_k^{\text{old}} + \gamma^{\text{new}}(z_{mk})/N_k^{\text{old}} - \gamma^{\text{old}}(z_{mk})/N_k^{\text{old}}$

(C) $\pi_k^{\text{new}} = \pi_k^{\text{old}} + \gamma^{\text{new}}(z_{mk})/N_k^{\text{new}} - \gamma^{\text{old}}(z_{mk})/N_k^{\text{new}}$

(D) $\pi_k^{\text{new}} = \pi_k^{\text{old}} + \gamma^{\text{new}}(z_{mk})/N - \gamma^{\text{old}}(z_{mk})/N$