提出課題 III：線形識別モデル

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

クラスの事前確率 $p(\mathcal{C}_k) = \pi_k$ と一般的なクラスの条件付き確率密度 $p(\boldsymbol{\phi}|\mathcal{C}_k)$ によって定義される K クラス分類問題を考える．ここで，$\boldsymbol{\phi}$ は入力特徴ベクトルである．学習データ $\{\boldsymbol{\phi}_n, \mathbf{t}_n\}$ が与えられたと仮定する．ただし，$n = 1, \ldots, N$ であり，$\mathbf{t}_n$ は，1-of-K 表記法を用いた長さ K の 2 値目的変数ベクトルである．つまり，パターン n のクラスが $\mathcal{C}_k$ である場合，2 値目的変数ベクトルは構成要素 $t_{nj} = I_{jk}$ をもつ．データがこのモデルから独立に抽出されると仮定すると，その事前確率に対する最尤解が以下の式で与えられることを示す．

$$\pi_k = \frac{N_k}{N}. \tag{a}$$

ここで，N_k はクラス $\mathcal{C}_k$ に割り当てられるデータの個数である．

尤度関数は以下の式によって与えられ，

$$p(\{\boldsymbol{\phi}_n, \mathbf{t}_n\}|\{\pi_k\}) = \prod_{n=1}^{N}\prod_{k=1}^{K}\{\boxed{\quad(1)\quad}\pi_k\}^{t_{nk}}$$

対数をとると，

$$\ln p(\{\boldsymbol{\phi}_n, \mathbf{t}_n\}|\{\pi_k\}) = \sum_{n=1}^{N}\sum_{k=1}^{K} t_{nk}\{\ln \boxed{\quad(1)\quad} + \ln \pi_k\}$$

となる．π_k について対数尤度の最大化する際には，拘束条件 $\sum_{k=1}^{K}\pi_k = 1$ を満たす必要がある．これはラグランジュ乗数 λ を導入して，以下の式

$$\ln p(\{\boldsymbol{\phi}_n, \mathbf{t}_n\}|\{\pi_k\}) + \lambda\left(\sum_{k=1}^{K}\pi_k - 1\right)$$

を最大化することによって実現できる．π_k についての微分を 0 とおくと，

$$\frac{1}{\pi_k}\boxed{\quad(2)\quad} + \lambda = 0$$

となり，整理すると

$$-\pi_k\lambda = \boxed{\quad(2)\quad} = N_k$$

となる．両辺で k についての和を求めることで，$\lambda = -N$ となることがわかり，これを用いて λ を消去することで式 (a) が得られる．

問題 1. 空欄 (1) に入る数式を選択せよ．

(A) $p(\boldsymbol{\phi}_n)$

(B) $p(\mathcal{C}_k)$

(C) $p(\boldsymbol{\phi}_n, \mathcal{C}_k)$

(D) $p(\boldsymbol{\phi}_n|\mathcal{C}_k)$

問題 2. 空欄 (2) に入る数式を選択せよ．

(A) t_{nk}

(B) $\sum_{n=1}^{N} t_{nk}$

(C) $\sum_{k=1}^{K} t_{nk}$

(D) $\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk}$

問題 3，4

以下のソフトマックス関数

$$p(\mathcal{C}_k|\boldsymbol{\phi}) = y_k(\boldsymbol{\phi}) = \frac{e^{a_k}}{\sum_j e^{a_j}} \tag{b}$$

の微分が

$$\frac{\partial y_k}{\partial a_j} = y_k(I_{kj} - y_j) \tag{c}$$

によって与えれることを示す．ここで，a_k は $a_k = \mathbf{w}_k^{\mathrm{T}} \boldsymbol{\phi}$ によって定義され，I_{kj} は単位行列の要素である．

式 (b) より

$$\frac{\partial y_k}{\partial a_k} = \frac{\boxed{(3)}}{(\sum_i e^{a_i})^2} = y_k(1 - y_k),$$
$$\frac{\partial y_k}{\partial a_j} = -\frac{\boxed{(4)}}{(\sum_i e^{a_i})^2} = -y_k y_j, \quad j \neq k$$

となる．これらの結果を統合すると，式 (c) が得られる．

問題 3. 空欄 (3) に入る数式を選択せよ．

(A) $(e^{a_k} \sum_i e^{a_i}) - e^{a_k}$

(B) $(e^{a_k} \sum_i e^{a_i}) - e^{2a_k}$

(C) $(e^{2a_k} \sum_i e^{a_i}) - e^{a_k}$

(D) $(e^{2a_k}\sum_i e^{a_i}) - e^{2a_k}$

問題 4. 空欄 (4) に入る数式を選択せよ.

(A) $e^{a_k}e^{a_j}$

(B) $e^{2a_k}e^{a_j}$

(C) $e^{a_k}e^{2a_j}$

(D) $e^{2a_k}e^{2a_j}$