知識工学II 課題2　略解・ヒント

下記に示す内容は，計算結果や証明の概略・戦略を示すのみで，完全な解答にはなっていない．課題を解くための参考資料なので，記載内容をそのまま転記してレポート作成せず，計算過程や完全な証明をレポートとしてまとめること．

【課題 2-1】

(1) -4

(2) $\begin{pmatrix} 12 & 9 & -6 \\ -8 & -6 & 4 \\ 20 & 15 & -10 \end{pmatrix}$

(3) $(-5, -31, -15)^t$

(4) $(15, 19, -7)$

【課題 2-2】

(1) 行列の (i, j) 成分が，特徴データの第 i 成分と第 j 成分との共分散になっているような行列であることを言えば良い．

(2) Σ の中の項に着目する．$(\overrightarrow{x_k} - \overrightarrow{m})(\overrightarrow{x_k} - \overrightarrow{m})^t$ は正方行列であるが，その (i, j) 成分は $(x_{ki} - m_i)(x_{kj} - m_j)$ と書けることに着目して，Σ の (i, j) 成分を書き出す．すると，Σ の (i, j) 成分は，特徴データの第 i 成分と第 j 成分との共分散になっていることが示せるので，それによって証明が完結する．

【課題 2-3】

(1) 具体的な定義は代数の教科書などを調査してまとめること．尚，“実対称行列” の “実” は実数の意味であり，“対称” は対角成分を軸として行列の要素が “対称” であることを意味している．

(2) 課題 2-2(1) の内容および共分散の定義を考えると自明である．

(3) 具体的な定義は代数の教科書などを調査すること．

一般に，d 次の正方行列に d 次の縦ベクトル $\overrightarrow{x}$ を右からかけると d 次の縦ベクトルが得られる．これを $\overrightarrow{y}$ とすると，通常，$\overrightarrow{x}$ と $\overrightarrow{y}$ は異なる向きのベクトルになる．この点に着目して，固有値・固有ベクトルの定義を見ると，どのようなことが言えるか考えてみること．

(4) 具体的な定義は代数の教科書などを調査すること．

(5) まず，$\overrightarrow{e_1}, \ldots, \overrightarrow{e_d}$ を行にもつような行列 $B = (\overrightarrow{e_1} \cdots \overrightarrow{e_d})$ を考える．固有値・固有ベクトルの定義 $A\overrightarrow{e_k} = \lambda_k \overrightarrow{e_k}$ の関係式に着目しながら，ABB^t を計算し $ABB^t = B\Lambda B^t$ となることを示す．ここで Λ は，対角成分に固有値をもち非対角成分が全て 0 の行列である．一方，行列 B の性質から BB^t は単位行列になるから，結果的に $A = ABB^t = B\Lambda B^t$ となることがわかる．あとは $B\Lambda B^t$ の構造を良くみながら求める関係式を導き出せばよい．

【課題 2-4】

(1) $\overrightarrow{A} = (a_1, a_2, \ldots, a_d)^t$, $\Sigma = (\sigma_{ij})$ とするとき，$\overrightarrow{A}^t \Sigma \overrightarrow{A} - \lambda\left(\overrightarrow{A}^t \overrightarrow{A} - 1\right)$ がスカラーになることに注目する．一般に，ベクトルによる微分は勾配ベクトルと言われ，スカラー値関数 F をベクトル $\overrightarrow{A}$ で微分したものは，次のように定義される．

$$\frac{\partial F}{\partial \overrightarrow{A}} = \left(\frac{\partial F}{\partial a_1}, \frac{\partial F}{\partial a_2}, \ldots, \frac{\partial F}{\partial a_d}\right)^t$$

従って，$\overrightarrow{A}^t \Sigma \overrightarrow{A} - \lambda\left(\overrightarrow{A}^t \overrightarrow{A} - 1\right)$ を a_i したものを 0 とおけばよい．

$$\overrightarrow{A}^t \Sigma \overrightarrow{A} = \overrightarrow{A}^t \begin{pmatrix} \sum_{k=1}^{d} \sigma_{1k} a_k \\ \sum_{k=1}^{d} \sigma_{2k} a_k \\ \vdots \\ \sum_{k=1}^{d} \sigma_{dk} a_k \end{pmatrix} = a_1 \sum_{k=1}^{d} \sigma_{1k} a_k + a_2 \sum_{k=1}^{d} \sigma_{2k} a_k + \cdots a_d \sum_{k=1}^{d} \sigma_{dk} a_k$$

であるから，これを a_i で微分すると次のようになる．

$$\begin{aligned} \frac{\partial}{\partial a_i} \overrightarrow{A}^t \Sigma \overrightarrow{A} &= a_1 \sigma_{1i} + \cdots + a_{i-1} \sigma_{i-1,i} + a_{i+1} \sigma_{i+1,i} + \cdots + a_d \sigma_{di} \\ &\quad + a_i \sigma_{ii} + \sum_{k=1}^{d} \sigma_{ik} a_k \end{aligned}$$

一方，

$$\frac{\partial}{\partial a_i} \lambda\left(\overrightarrow{A}^t \overrightarrow{A} - 1\right) = 2\lambda a_i$$

従って，これらを整理することにより $\Sigma \overrightarrow{A} - \lambda \overrightarrow{A} = \overrightarrow{0}$ を得る．

(2) 前問の結果から $\Sigma \overrightarrow{A} = \lambda \overrightarrow{A}$ であることがわかる．これは固有値・固有ベクトルの定義式であり，$\overrightarrow{A}$ は Σ の固有ベクトルでなければならないことがわかる．この式の両辺に左から $\overrightarrow{A}^t$ をかけると

$$\overrightarrow{A}^t \Sigma \overrightarrow{A} = \lambda$$

となることに着目すると，$\overrightarrow{A}$ は Σ の最大固有値に対応する固有ベクトルでなければならないことがわかる．