نمره: ۱۰ + ۳۵

مسئله‌ی ۱. مقدار و بردار ویژه (۱۰ نمره)

.۱

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\det(R - \lambda I) = 0. \Rightarrow \det\begin{bmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{bmatrix} = 0.$$

$$(\cos\theta - \lambda)^2 + \sin^2\theta = 1 + \lambda^2 - 2\lambda\cos\theta = 0.$$

$$\Rightarrow \lambda^2 - 2\lambda\cos\theta + 1 = 0. \Rightarrow \lambda_1, \lambda_2 = \frac{\cos\theta \pm \sqrt{\cos^2\theta - 1}}{1}$$

$$\Rightarrow \lambda_1 = \cos\theta + j\sin\theta \quad , \quad \lambda_2 = \cos\theta - j\sin\theta$$

$$\Rightarrow |\lambda_1| = |\lambda_2| = 1$$

$$\Rightarrow \prod_{i=1}^{n} |\lambda_i| = 1$$

$$\det R(\theta) = \cos^2\theta + \sin^2\theta = 1 \Rightarrow \boxed{\det R(\theta) = \prod_{i=1}^{2} |\lambda_i|}$$

بردار ویژه : $R v = \lambda v \Rightarrow \begin{cases} v_1\cos\theta - v_2\sin\theta = v_1\cos\theta + j v_1\sin\theta \\ v_1\sin\theta + \cancel{v_2\cos\theta} = \cancel{v_2\cos\theta} + j v_2\sin\theta \end{cases}$

$$\Rightarrow \begin{cases} v_2 = -j v_1 \\ v_1 = j v_2 \end{cases} \Rightarrow \vec{v}_{(1)} = \begin{bmatrix} 1 \\ -j \end{bmatrix}$$

بردار ویژه متناظر با $\lambda = \cos\theta + j\sin\theta$

$$\begin{cases} \cancel{v_1\cos\theta} - v_2\sin\theta = \cancel{v_1\cos\theta} - j v_1\sin\theta \Rightarrow v_2 = j v_1 \\ v_1\sin\theta + \cancel{v_2\cos\theta} = \cancel{v_2\cos\theta} - j v_2\sin\theta \Rightarrow v_1 = -j v_2 \end{cases}$$

$$\Rightarrow \vec{v}_{(2)} = \begin{bmatrix} 1 \\ j \end{bmatrix}$$

بردار ویژه متناظر با

$$\lambda = \cos\theta - j\sin\theta$$

$$R(\theta) = V \Lambda V^{-1} \qquad V = \begin{bmatrix} 1 & 1 \\ -j & j \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \cos\theta + j\sin\theta & 0 \\ 0 & \cos\theta - j\sin\theta \end{bmatrix}$$

$$\hookrightarrow R^n(\theta) = (V\Lambda V^{-1})(V\Lambda V^{-1}) \cdots (V\Lambda V^{-1})$$

$$= V(\Lambda)^n V^{-1}$$

$$= V \begin{bmatrix} (\cos\theta + j\sin\theta)^n & 0 \\ 0 & (\cos\theta - j\sin\theta)^n \end{bmatrix} V^{-1}$$

$$= V \begin{bmatrix} e^{j\theta n} & 0 \\ 0 & e^{-j\theta n} \end{bmatrix} V^{-1} = V \begin{bmatrix} \cos(n\theta) + j\sin n\theta & 0 \\ 0 & \cos(n\theta) - j\sin(n\theta) \end{bmatrix} V^{-1}$$

$$= R(n\theta) \qquad \Rightarrow R^n(\theta) = R(n\theta)$$

(با جایگذاری $\alpha = n\theta$ در $R(n\theta)$ می‌بینیم که مقدار ویژه
و بردارهای ویژه همانند قسمت اول سوال هستند برای $n\theta$)

مسئله‌ی ۲. مشتق بردار و ماتریس (۲۵ نمره)

۱. توجه نمایید که استفاده‌ی شما از رویکرد Numerator و Denominator اهمیتی ندارد و تفاوت جواب شما با پاسخ‌نامه در حد Transpose ایرادی ندارد.

الف)

$$\frac{\partial}{\partial x} x^T a \ , \quad x \in \mathbb{R}^n, a \in \mathbb{R}^n$$

$$[x_1, \dots, x_n] \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = x_1 a_1 + \dots + x_n a_n = F = a_1 x_1 + \dots + a_n x_n = a^T x$$

$$\frac{\partial F}{\partial x_j} = a_j \Rightarrow \frac{\partial}{\partial x} f = a^T$$

ب)

$$tr(X) = \sum_{i=1}^{n} x_{ii} \Rightarrow \frac{\partial}{\partial y} tr(x) = \frac{\partial}{\partial y} \sum_{i=1}^{n} x_{ii}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial y} x_{ii} = tr\left(\begin{bmatrix} \frac{\partial x_{11}}{\partial y} & \cdots & \frac{\partial x_{1n}}{\partial y} \\ \vdots & & \vdots \\ \frac{x_{n1}}{\partial y} & \cdots & \frac{\partial x_{nn}}{\partial y} \end{bmatrix} \right) = tr\left(\frac{\partial}{\partial y} X\right)$$

پ) $\frac{\partial}{\partial X}\mathrm{tr}(X^T A X) = X^T(A + A^T)$

$$\mathrm{tr}(X^T A X)$$
$$= \mathrm{tr}\left(\begin{bmatrix} x_{11} & \cdots & x_{n1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \cdots & x_{nn} \end{bmatrix}\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}\right)$$
$$= \sum_{m=1}^{n}\sum_{i=1}^{n} x_{im}\sum_{j=1}^{n} x_{jm}a_{ji}$$
$$= \sum_{m=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_{im}x_{jm}a_{ji}$$

$$\frac{\partial}{\partial X}\mathrm{tr}(X^T A X)_{pq}$$
$$= \frac{\partial}{\partial x_{qp}}\sum_{m=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_{im}x_{jm}a_{ji}$$
$$= \sum_{j=1}^{n} x_{jp}a_{jq} + \sum_{i=1}^{n} x_{ip}a_{qi}$$

$$\frac{\partial}{\partial X}\mathrm{tr}(X^T A X)_{pp}$$
$$= \frac{\partial}{\partial x_{pp}}\sum_{m=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_{im}x_{jm}a_{ji}$$
$$= \sum_{j=1, j\neq p}^{n} x_{jp}a_{jp} + \sum_{i=1, i\neq p}^{n} x_{ip}a_{pi} + 2x_{pp}a_{pp}$$
$$= \sum_{j=1}^{n} x_{jp}a_{jp} + \sum_{i=1}^{n} x_{ip}a_{pi}$$

بنابراین:

$$\frac{\partial}{\partial X}\mathrm{tr}(X^T A X)_{pq} = \sum_{j=1}^{n} x_{jp}a_{jq} + \sum_{i=1}^{n} x_{ip}a_{qi} = (X^T A)_{pq} + (X^T A^T)_{pq}$$

$$\rightarrow \frac{\partial}{\partial X}\mathrm{tr}(X^T A X) = X^T(A + A^T)$$

ت) $\frac{\partial}{\partial X}\log(\det(X)) = X^{-T}$ (در این سوال اگر از رویکرد Numerator استفاده شود سمت راست دیگر نباید Transpose شود، ولی همانطورکه مطرح شد در نمره‌دهی، این مساله اهمیتی ندارد).

می‌دانیم:

$$1)\ \det(X) = \sum_{k=1}^{n} X_{ik}C_{ik}$$

$$2)\ (X^{-1})_{ij} = \frac{1}{\det(X)}C_{ji}$$

حال داریم:

$$\begin{tikzcd}
& \frac{\partial}{\partial X}\log(\det(X)) = X^{-1} \\
\frac{\partial}{\partial X_{ij}}\log(\det(X)) = \frac{1}{\det\,(X)}\frac{\partial}{\partial X_{ij}}\det(X) = \frac{1}{\det\,(X)}C_{ij} = (X^{-1})_{ji} \arrow[ur, "\text{Nominator Layout}"] \arrow[dr, "\text{Denominator Layout}"'] & \\
& \frac{\partial}{\partial X}\log(\det(X)) = X^{-T}
\end{tikzcd}$$

۲.

$$f(x_1, x_2, x_3) = (\underbrace{x_1 x_2^2 x_3^2}_{f_1},\ \underbrace{x_3^2 \sin(x_2)}_{f_2},\ \underbrace{x_1^2 e^{x_2}}_{f_3})$$

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_2^2 x_3^2 & 2x_1 x_2 x_3^2 & 2x_1 x_2^2 x_3 \\ 0 & x_3^2 \cos(x_2) & 2x_3 \sin(x_2) \\ 2x_1 e^{x_2} & x_1^2 e^{x_2} & 0 \end{bmatrix}$$

$$x = (1, \pi, -1) \Rightarrow \begin{bmatrix} \pi^2 & 2\pi & -2\pi^2 \\ 0 & -1 & 0 \\ 2e^{\pi} & e^{\pi} & 0 \end{bmatrix}$$

مسئله‌ی ۳. الگوریتم کاهش ابعاد SNE (۱۰ نمره)

۱.

$$q_{j|i} = \frac{e^{-\|y_i-y_j\|^2}}{\sum_{k\neq i} e^{-\|y_i-y_k\|^2}} = \frac{E_{ij}}{\sum_{k\neq i} E_{ik}} = \frac{E_{ij}}{Z_i} \tag{1}$$

Notice that $E_{ij} = E_{ji}$. The loss function is defined as

$$\begin{aligned} C &= \sum_{k,l\neq k} p_{l|k} \log \frac{p_{l|k}}{q_{l|k}} = \sum_{k,l\neq k} p_{l|k} \log p_{l|k} - p_{l|k} \log q_{l|k} \\ &= \sum_{k,l\neq k} p_{l|k} \log p_{l|k} - p_{l|k} \log E_{kl} + p_{l|k} \log Z_k \end{aligned} \tag{2}$$

We derive with respect to y_i. To make the derivation less cluttered, I will omit the ∂y_i term at the denominator.

$$\frac{\partial C}{\partial y_i} = \sum_{k,l\neq k} -p_{l|k} \partial \log E_{kl} + \sum_{k,l\neq k} p_{l|k} \partial \log Z_k$$

We start with the first term, noting that the derivative is non-zero when $\forall j \neq i$, $k = i$ or $l = i$

$$\sum_{k,l\neq k} -p_{l|k} \partial \log E_{kl} = \sum_{j\neq i} -p_{j|i} \partial \log E_{ij} - p_{i|j} \partial \log E_{ji} \tag{3}$$

Since $\partial E_{ij} = E_{ij}(-2(y_i - y_j))$ we have

$$\begin{aligned} &\sum_{j\neq i} -p_{j|i} \frac{\cancel{E_{ij}}}{\cancel{E_{ij}}}(-2(y_i - y_j)) - p_{i|j} \frac{\cancel{E_{ji}}}{\cancel{E_{ji}}}(2(y_j - y_i)) \\ &\qquad = 2 \sum_{j\neq i} (p_{j|i} + p_{i|j})(y_i - y_j)) \end{aligned} \tag{4}$$

We conclude with the second term. Since $\sum_{l\neq j} p_{l|j} = 1$ and Z_j does not depend on k, we can write (changing variable from l to j to make it more similar to the already computed terms)

$$\sum_{j,k\neq j} p_{k|j} \partial \log Z_j = \sum_j \partial \log Z_j$$

We conclude with the second term. Since $\sum_{l\neq j} p_{l|j} = 1$ and Z_j does not depend on k, we can write (changing variable from l to j to make it more similar to the already computed terms)

$$\sum_{j,k\neq j} p_{k|j} \partial \log Z_j = \sum_j \partial \log Z_j$$

The derivative is non-zero when $k = i$ or $j = i$ (also, in the latter case we can move Z_i inside the summation because constant)

$$\begin{aligned} &= \sum_j \frac{1}{Z_j} \sum_{k\neq j} \partial E_{jk} \\ &= \sum_{j\neq i} \frac{E_{ji}}{Z_j}(2(y_j - y_i)) + \sum_{j\neq i} \frac{E_{ij}}{Z_i}(-2(y_i - y_j)) \\ &= 2 \sum_{j\neq i} (-q_{j|i} - qi|j)(y_i - y_j) \end{aligned} \tag{5}$$

Combining eq. (4) and (5) we arrive at the final result

$$\frac{\partial C}{\partial y_i} = 2 \sum_{j\neq i} (p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j) \quad \square \tag{6}$$