2023-2 DSL 정규세션 과제

ML/DL을 위한 수학 1,2

10기 정 성 오

Q1.

(1) $\det(A) = 1 \times 2 \times 3$

$= 6$

A를 활용한 linear transformation를 통해 basis vector가 바뀌었을 때 공간이 6배 만큼 늘어난다.

(2) $A\vec{x} = \lambda\vec{x}$

$(A-\lambda I)\vec{x} = \vec{0}$

$\det\left(\begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 3 \\ 0 & 0 & 3-\lambda \end{bmatrix}\right) = \vec{0}$

$(1-\lambda)(2-\lambda)(3-\lambda) = \vec{0} \rightarrow \lambda = 1, 2, 3$

i) $\lambda = 1$

$(A-\lambda I)\vec{x} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix}\vec{x} = \vec{0}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$(A-\lambda I) \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$\vec{x} = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad a \in \mathbb{R}$

ii) $\lambda = 2$

$(A-\lambda I)\vec{x} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}\vec{x} = \vec{0}$

$(A-\lambda I) \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

$\vec{x} = \begin{bmatrix} 0 \\ a \\ 0 \end{bmatrix} = a\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad a \in \mathbb{R}$

iii) $\lambda = 3$

$(A-\lambda I)\vec{x} = \begin{bmatrix} -2 & 0 & 1 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix}\vec{x} = \vec{0}$

$(A-\lambda I) \sim \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}$

$\vec{x} = \begin{bmatrix} a \\ 6a \\ 2a \end{bmatrix} = a\begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix} \quad a \in \mathbb{R}$

$\therefore$ eigen value : 1 , 2 , 3

eigen vector : $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$ → 해당 vector들의 basis가 A에 의해 바뀌어도 여전히 각 vector의 span 안에 존재한다. (eigen value를 곱해준 값이 됨.)

Q2. $B=\begin{bmatrix}0&0&0\\0&0&0\\3&0&1\end{bmatrix} \rightarrow D=P^{-1}BP$

$\det(B-\lambda I)=\begin{vmatrix}-\lambda&0&0\\0&-\lambda&0\\3&0&1-\lambda\end{vmatrix} = \lambda^2(1-\lambda)=0 \rightarrow \lambda=0 \text{ or } 1$

i) $\lambda=0$

$(B-\lambda I)\vec{x}=\begin{bmatrix}0&0&0\\0&0&0\\3&0&1\end{bmatrix}\vec{x}=\vec{0}$

$(B-\lambda I)\sim\begin{bmatrix}3&0&1\\0&0&0\\0&0&0\end{bmatrix}$

$\vec{x}=\begin{bmatrix}-3a\\b\\a\end{bmatrix}=a\begin{bmatrix}-3\\0\\1\end{bmatrix}+b\begin{bmatrix}0\\1\\0\end{bmatrix}$ $(a\in\mathbb{R},\ b\in\mathbb{R})$

ii) $\lambda=1$

$(B-\lambda I)\vec{x}=\begin{bmatrix}-1&0&0\\0&-1&0\\3&0&0\end{bmatrix}\vec{x}=\vec{0}$

$(B-\lambda I)\sim\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$

$\vec{x}=\begin{bmatrix}0\\0\\a\end{bmatrix}=a\begin{bmatrix}0\\0\\1\end{bmatrix}$

↳ $P=\begin{bmatrix}-3&0&0\\0&1&0\\1&0&1\end{bmatrix}$

$D=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$

P^{-1} 구하기

$\left[\begin{array}{ccc|ccc}-3&0&0&1&0&0\\0&1&0&0&1&0\\1&0&1&0&0&1\end{array}\right] \overset{-\frac{1}{3}\times① \rightarrow ①}{\sim} \left[\begin{array}{ccc|ccc}1&0&0&-\frac{1}{3}&0&0\\0&1&0&0&1&0\\1&0&1&0&0&1\end{array}\right] \overset{-1\times①+③\rightarrow③}{\sim} \left[\begin{array}{ccc|ccc}1&0&0&-\frac{1}{3}&0&0\\0&1&0&0&1&0\\0&0&1&\frac{1}{3}&0&1\end{array}\right]$

$P^{-1}=\begin{bmatrix}-\frac{1}{3}&0&0\\0&1&0\\\frac{1}{3}&0&1\end{bmatrix}$

$\therefore D=P^{-1}BP$

$\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}=\begin{bmatrix}-\frac{1}{3}&0&0\\0&1&0\\\frac{1}{3}&0&1\end{bmatrix}\begin{bmatrix}0&0&0\\0&0&0\\3&0&1\end{bmatrix}\begin{bmatrix}-3&0&0\\0&1&0\\1&0&1\end{bmatrix}$

Q3. $P(spam)=0.2$

$P('당첨'|spam)=0.5$

$P('당첨'|spam^c)=0.01$

$P(spam|'당첨')=\dfrac{P(spam\cap'당첨')}{P('당첨')}$

$=\dfrac{P(spam\cap'당첨')}{P(spam\cap'당첨')+P(spam^c\cap'당첨')}$

$=\dfrac{0.1}{0.1+0.008}$

$=\dfrac{100}{100+8}$

$=\dfrac{100}{108}$

※ $P('당첨|spam)$ $\dfrac{P('당첨'\cap spam)}{P(spam)}=0.5$

$P('당첨'\cap spam)=0.5\times0.2=0.1$

$P('당첨'|spam^c)=\dfrac{P('당첨'\cap spam^c)}{P(spam^c)}=0.01$

$P('당첨'\cap spam^c)=0.01\times0.8=0.008$

$\therefore \dfrac{100}{108}$

Q4.

(2) $\theta = \frac{9}{15}$

entropy $H[X] = -\theta \ln \theta - (1-\theta)\ln(1-\theta)$

$= -\frac{9}{15}\ln\frac{9}{15} - \frac{6}{15}\ln\frac{6}{15}$

$\fallingdotseq 0.673$

(3) $\theta_1 = \frac{9}{15}$, $\theta_2 = \frac{4}{15}$

$D_{KL}(\underset{\theta_1}{P} \| \underset{\theta_2}{Q}) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$

데이터 분포의 차이 $D_{KL}(P \| Q) = \frac{9}{15} \log\left(\left(\frac{9}{15}\right) / \left(\frac{4}{15}\right)\right)$

$\fallingdotseq 0.486$

Q5. $f_\beta(x) = \dfrac{\exp(\beta^T X)}{1+\exp(\beta^T X)} = \dfrac{1}{1+\exp(-\beta^T X)}$

loss function $J(\beta) = -\sum\{y_i \log(f_\beta(x_i)) + (1-y_i)\log(1-f_\beta(x_i))\} \rightarrow$ convex

convex optimization

min $J(\beta)$

s.t. $g_i(\vec{x}) \le 0 \quad (i = 1, 2, \cdots, m)$

$a_j^T \vec{x} = b_j \quad (j = 1, 2, \cdots, p)$

($f, g_1, g_2, \cdots, g_m$은 convex function)

아무리 검색해보고 찾아봐도 답을 도출하지 못했습니다...