<assignment_math for ML/DL> - 10기 유현동

Q1.

(1) $\det\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}\right) = 1 \times \begin{vmatrix} 2 & 3 \\ 0 & 3 \end{vmatrix} = 6$

determinant의 기하학적 의미는 linear transformation을 통해서 basis vector가 바뀌었을 때 공간이 얼마나 늘어나고 줄어드는지를 의미한다는 것이다. determinant의 값은 basis vector의 직선으로 이루어진 공간의 넓이(부피)이다.

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

$$\det(\lambda I - A) = \begin{vmatrix} \lambda-1 & 0 & -1 \\ 0 & \lambda-2 & -3 \\ 0 & 0 & \lambda-3 \end{vmatrix} = (\lambda-1) \times \begin{vmatrix} \lambda-2 & -3 \\ 0 & \lambda-3 \end{vmatrix}$$

$$= (\lambda-1)(\lambda-2)(\lambda-3)$$

A의 고유값은 $\lambda=1$, $\lambda=2$, $\lambda=3$ 이다.
각 고유값의 고유벡터를 구해보면 다음과 같다

i) $\lambda=1$ 일 때

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

$x_3 = 0$. $x_2 = 3x_3 = 0$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} t$, $t \in \mathbb{R}$.

$\therefore$ $\lambda=1$의 고유벡터는 $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ 이다.

ii) $\lambda=2$일 때

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

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}s\ ,\quad s\in\mathbb{R}$$

$\therefore$ $\lambda=2$의 고유벡터는 $\begin{bmatrix}0\\1\\0\end{bmatrix}$ 이다.

iii) $\lambda=3$일 때

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

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\3\\1\end{bmatrix}u\ ,\quad u\in\mathbb{R}$$

$\therefore$ $\lambda=3$일 때 고유벡터는 $\begin{bmatrix}\frac{1}{2}\\3\\1\end{bmatrix}$ 이다

Q2. $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ 를 대각화하여도 다음과 같다

i) 고유값 구하기.

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

$\therefore \lambda = 0, \ \lambda = 1$

ii) 고유벡터 구하기

① $\lambda = 0$ 일 때

$(\lambda I - B)x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad . \quad x_3 = 3x_1$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} l + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} m \quad . \quad l \in \mathbb{R}, m \in \mathbb{R}$

$\therefore \lambda = 0$ 일 때 고유벡터는 $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ 과 $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ 이다.

② $\lambda = 1$ 일 때

$(\lambda I - B)x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad . \quad x_1 = x_2 = 0$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} n \quad . \quad n \in \mathbb{R}$

$\therefore \lambda = 1$ 일 때 고유벡터는 $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ 이다

iii) 행렬 P 만들기

P는 B의 고유벡터들로 이루어진 행렬이다

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

P의 역행렬을 구해보면

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{array}\right]$$

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

iv) 대각화

$$P^{-1}BP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{\text{let}}{=} D$$

Q3. $P(\text{spam}) = 0.2$, $P(\text{not spam}) = 0.8$

$P(\text{당첨} \mid \text{spam}) = 0.5$, $P(\text{당첨} \mid \text{not spam}) = 0.01$

$$P(\text{당첨}) = P(\text{당첨} \mid \text{spam}) \times P(\text{spam}) + P(\text{당첨} \mid \text{not spam}) \times P(\text{not spam})$$
$$= 0.5 \times 0.2 + 0.01 \times 0.8 = 0.108$$

$$\therefore P(\text{spam} \mid \text{당첨}) = \frac{P(\text{spam}) \cdot P(\text{당첨} \mid \text{spam})}{P(\text{당첨})} = \frac{0.2 \times 0.5}{0.108} \fallingdotseq 0.9259$$

Q4.

(2). "hit.csv" 에서 스윙 결과가 hit인 경우의 확률은 $\hat{p} = \frac{9}{15}$ 이다.

X : 스윙결과

$$H(X) = -\frac{9}{15}\log_2\frac{9}{15} - \frac{6}{15}\log_2\frac{6}{15} = 0.971$$

(3). "hit_other.csv" 에서 스윙결과가 hit인 확률은 $\hat{q} = \frac{4}{15}$ 이다.

첫 번째 팀과 두 번째 팀의 스윙 결과의 데이터 분포의 차이는

KL divergence를 이용해서 구할 수 있다.

$$D(p\|q) = \frac{9}{15}\log_2\frac{\frac{9}{15}}{\frac{4}{15}} + \frac{6}{15}\log_2\frac{\frac{6}{15}}{\frac{11}{15}} = \frac{9}{15}\log_2\frac{9}{4} + \frac{6}{15}\log_2\frac{6}{11}$$

$$= 0.352$$

Q5. logistic regression의 cost function : $-y\log\hat{y} - (1-y)\log(1-\hat{y})$

(cross entropy of $y, \hat{y}$)

$$\min\sum_{i=1}^{m} -y^{(i)}\log(\hat{y}^{(i)}) - (1-y^{(i)})\log(1-\hat{y}^{(i)})$$

$$= \min_{w}\underbrace{\sum_{i=1}^{m} -y^{(i)}\cdot\log\frac{1}{1+e^{-w^T x^{(i)}}} - (1-y^{(i)})\log\frac{e^{-w^T x^{(i)}}}{1+e^{-w^T x^{(i)}}}}_{\stackrel{let}{=} J(w)}$$

J(w)가 convex임을 만족하려면 J(w)의 두 항이 모두 convex해야 한다.

$-\log\frac{e^{-w^T x^{(i)}}}{1+e^{-w^T x^{(i)}}} = w^T x - \log\frac{1}{1+e^{-w^T x^{(i)}}}$ 이고 Hessian이

positive semi definite이므로 convex하다.

따라서 J(w)가 convex하다면 $\nabla J(w) = 0$ 이고 이때의 w는 최적값인

w^*를 의미한다.