Q1.

(1) $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$, $\det(A) = 6 + 0 + 0 = 6$

⇒ 기존 3차원 기저 벡터가 이루는 공간이 1에서 6으로 왜곡되었다.

(2) $(A - \lambda I)\mathbf{x} = 0 \Rightarrow \det(A - \lambda I) = 0$

$$\begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 3 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(3-\lambda) + 0 + 1 \cdot 0 = 0$$

$\therefore \lambda = 1$ or $\lambda = 2$ or $\lambda = 3$

i) $\lambda = 1$

$$\begin{array}{ccc} x_1 & x_2 & x_3 \end{array}$$

$$\left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right]$$

$\downarrow$

$x_3 = 0,\ x_2 = 0$

$x_1 = t$

$\mathbf{x}_{\lambda=1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

ii) $\lambda = 2$

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

$\downarrow$

$x_3 = 0,\ x_1 = 0$

$x_2 = s$

$\mathbf{x}_{\lambda=2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

iii) $\lambda = 3$

$$\left[\begin{array}{ccc|c} -2 & 0 & 1 & 0 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\downarrow$

$x_3 = u,\ x_1 = \frac{1}{2}u,\ x_2 = 3u$

$\mathbf{x}_{\lambda=3} = \begin{bmatrix} \frac{1}{2} \\ 3 \\ 1 \end{bmatrix}$

⇒ eigenvector 는 A라는 선형변환과 관계없이 유지되는 벡터의 고유 방향을 의미하고,
eigenvalue 는 A라는 선형변환에 의해 eigenvector들의 크기가 바뀌는 정도로,
3차원이니까 각각 1배, 2배, 3배가 된다는 것을 의미한다.

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

I) P, D 구하기

$\det(B - \lambda I) = \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 3 & 0 & 1-\lambda \end{vmatrix} = (-\lambda)(-\lambda)(1-\lambda) = 0$ $\therefore \lambda_1 = 0,\ \lambda_2 = 1$

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

i) $\lambda_1 = 0$

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

$x_2 = t_1$, $x_3 = s_1$

$x_1 = -\frac{1}{3} s_1$

$X = t_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + s_1 \begin{bmatrix} -\frac{1}{3} \\ 0 \\ 1 \end{bmatrix}$

ii) $\lambda_2 = 1$

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

$x_1 = 0$, $x_2 = 0$, $x_3 = u$

$X = u \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

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

II) P^{-1} 구하기

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

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

Q3.

	spam	spam x
당첨 o	0.5	0.01
당첨 x	0.5	0.99

$P(\text{spam}) = 0.2$

$P(\text{not spam}) = 0.8$

$P(\text{당첨} \mid \text{spam}) = 0.5$

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

$$P(\text{spam} \mid \text{당첨}) = \frac{P(\text{당첨} \mid \text{spam}) \cdot P(\text{spam})}{P(\text{당첨})}$$

$$= \frac{P(\text{당첨} \mid \text{spam}) \cdot P(\text{spam})}{P(\text{당첨} \mid \text{spam}) \cdot P(\text{spam}) + P(\text{당첨} \mid \text{not spam}) \cdot P(\text{not spam})}$$

$$= \frac{0.5 \cdot 0.2}{0.5 \cdot 0.2 + 0.01 \cdot 0.8} \approx 0.9259$$

Q4. (2) <hit.csv> $P(\text{Hit}) = \frac{9}{15} = \frac{3}{5}$, $P(\text{out}) = \frac{2}{5}$

$$H_2 = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5}$$

$$= -\frac{3}{5}\log_2 3 + \frac{3}{5}\log_2 5 - \frac{2}{5}\log_2 2 + \frac{2}{5}\log_2 5$$

$$= \log_2 5 - \frac{3}{5}\log_2 3 - \frac{2}{5} \approx 0.971$$

(3) <hit_other> $\Rightarrow$ $P(\text{Hit}) = \frac{4}{15}$, $P(\text{out}) = \frac{11}{15}$

$$KL(2\|3) = H(2,3) - H(2)$$

$$= \frac{3}{5}\log_2\frac{15}{4} + \frac{2}{5}\log_2\frac{15}{11} - H_2$$

$$= \frac{3}{5}\log_2 15 - \frac{3}{5}\log_2 4 + \frac{2}{5}\log_2 15 - \frac{2}{5}\log_2 11 - H_2$$

$$= \log_2 3 + \cancel{\log_2 5} - \frac{6}{5} - \frac{2}{5}\log_2 11 - \cancel{\log_2 5} + \frac{3}{5}\log_2 3 + \frac{2}{5}$$

$$= \frac{8}{5}\log_2 3 - \frac{2}{5}\log_2 11 - \frac{4}{5} \approx 0.3522$$

Q5.

Logistic Regression의 Loss function : cross entropy

$$\min_w \left\{ \sum_{i=1}^{m} \left(-y^{(i)} \log(\hat{y}^{(i)}) - (1-y^{(i)}) \log(1-\hat{y}^{(i)}) \right) \right\} , \quad \hat{y}^{(i)} = \frac{1}{1+e^{-w^T x^{(i)}}} = f(w)$$

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

$$= -y^{(i)} \log \frac{1}{1+e^{-w^T x^{(i)}}} - (1-y^{(i)}) \left(-w^T x^{(i)} + \log \frac{1}{1+e^{-w^T x^{(i)}}} \right)$$

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

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

since $\nabla^2 f(w) \geq 0$, $J(w)$ is convex in w

$\Rightarrow \min_w J(w)$, $w \in \mathbb{R}$, $J(w)$: convex. $\mathbb{R}$: convex set.

$\therefore \nabla J(w) = 0$ implies $w = w^*$, where $w^* = \min_w J(w)$