P1

Personally, I value sense of humor the most among all other things. I always consider happiness to be one of the most important things in my life. By having a good sense of humor, I can not only be in a good mood myself, but I can bring laughters to my friends and family as well. With the techniques in machine learning, it's possible to create a program which can generate jokes on a certain subject and bring happiness to even more people.

P2

(a) (1) False

(2) False

(3) False

(b) False

P3

(a)

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\boldsymbol{A}_{Ref} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

So the rank of A is 1.

(b)

$$\begin{aligned} \boldsymbol{Au} + \boldsymbol{Bv} &= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 16 \end{bmatrix} \end{aligned}$$

(c)

$$\begin{aligned} \boldsymbol{u}^\intercal \boldsymbol{Av} &= \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= 20 \end{aligned}$$

(d)

$$\begin{aligned}\|u\|_2 &= \sqrt{u_1^2 + u_2^2 + \cdots + u_d^2} \\ &= \sqrt{2^2 + 1^2} \\ &= \sqrt{5}\end{aligned}$$

(e)

$$\begin{aligned}\nabla f(v) &= \frac{\partial}{\partial v_i} f(v) \\ &= \sum_{j=1}^{d} f_{ij} v_j + \sum_{k=1}^{d} f_{ki} v_k \\ &= \boldsymbol{v}^\intercal (\boldsymbol{A} + \boldsymbol{B})^\intercal + \boldsymbol{v}^\intercal (\boldsymbol{A} + \boldsymbol{B}) \\ &= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix} + \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 18 \end{bmatrix} + \begin{bmatrix} 9 & 18 \end{bmatrix} \\ &= \begin{bmatrix} 18 & 36 \end{bmatrix}\end{aligned}$$

(f)

$$\begin{aligned}F(x_1, x_2, \lambda) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \lambda(1 - {x_1}^2 - {x_2}^2) \\ &= 5{x_1}^2 + 4x_1x_2 + 8{x_2}^2 + \lambda(1 - {x_1}^2 - {x_2}^2)\end{aligned}$$

$$\begin{aligned}10x_1 + 4x_2 - 2\lambda x_1 &= 0 \\ 4x_1 + 16x_2 - 2\lambda x_2 &= 0 \\ 1 - {x_1}^2 - {x_2}^2 &= 0\end{aligned}$$

$$\begin{aligned}\lambda = 4 &\text{ or } \lambda = 9 \\ x1 = \frac{-2\sqrt{5}}{5} &\text{ or } x1 = \frac{2\sqrt{5}}{5} \\ x2 = \frac{\sqrt{5}}{5} &\text{ or } x2 = -\frac{\sqrt{5}}{5} \\ x = \begin{bmatrix} \frac{-2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{bmatrix} &\text{ or } x = \begin{bmatrix} \frac{2\sqrt{5}}{5} \\ \frac{-\sqrt{5}}{5} \end{bmatrix} \\ f_{min}(x) &= 4\end{aligned}$$

P4

(a) A and B, A and C are independent.

(b) $P(C \mid T) = \frac{0.2}{0.5} = 0.4$

(c) $P(X = 1) = \frac{11}{36}$

$P(X = 2) = \frac{9}{36} = \frac{1}{4}$

$P(X = 3) = \frac{7}{36}$

$P(X = 4) = \frac{5}{36}$

$P(X = 5) = \frac{3}{36} = \frac{1}{12}$

$P(X = 6) = \frac{1}{36}$

(d) $\mathbb{E}(X) = 1 \cdot \frac{11}{36} + 2 \cdot \frac{9}{36} + 3 \cdot \frac{7}{36} + 4 \cdot \frac{5}{36} + 5 \cdot \frac{3}{36} + 6 \cdot \frac{1}{36} = \frac{91}{36}$

(e) $\mathbb{E} = \frac{1}{P(heads)} = \frac{1}{\frac{1}{5}} = 5$

(f) $P(ARIAR \mid n = 5) = \frac{1}{4}^5 = \frac{1}{1024}$

$\mathbb{E}(ARIAR \mid n = 5) = \frac{1}{P(ARIAR \mid n = 5)} = 1024$

$\mathbb{E}(Times) = \frac{n - 4}{\mathbb{E}(ARIAR \mid n = 5)} = \frac{n - 4}{1024}$

P5

(a)

$$
\begin{aligned}
P(X \leq 1000000) = \int_0^{1000000} \lambda e^{-\lambda x} \mathrm{d}x. &= 0.5 \\
-e^{-\lambda x}\Big|_0^{1000000} &= 0.5 \\
-e^{-1000000\lambda} + e^0 &= 0.5 \\
e^{-1000000\lambda} &= 0.5 \\
\lambda &= \frac{\ln 0.5}{-1000000} \\
\lambda &= \frac{\ln 2}{1000000}
\end{aligned}
$$

(b)

$$\begin{aligned}\mathbb{E}(X) &= 0\\ \sigma &= \sqrt{\mathbb{E}(X^2) - (\mathbb{E}(X))^2} = 1\\ \mathbb{E}(Y) &= \mathbb{E}(X^2) = 1\end{aligned}$$

(c)

$$\begin{aligned}\int_0^1 \int_0^{0.5} c \mathrm{d}x_2 \mathrm{d}x_1 &= 1\\ \int_0^1 0.5c \mathrm{d}x_1 &= 1\\ c &= 2\end{aligned}$$

(d)

$$\begin{aligned}P(X_2 \geq X_1) &= \int_0^{0.5} \int_{x_1}^1 2 \mathrm{d}x_2 \mathrm{d}x_1\\ &= \int_0^{0.5} (2 - 2x_1) \mathrm{d}x_1\\ &= 2x_1 - {x_1}^2 \Big|_0^{0.5}\\ &= 0.75\end{aligned}$$

(e)

$$\begin{aligned}P(Y) &= \int_0^{0.25} \int_{2x_2}^{0.5} 2 \mathrm{d}x_1 \mathrm{d}x_2\\ &= \int_0^{0.25} (1 - 4x_2) \mathrm{d}x_1\\ &= x_2 - 2{x_2}^2 \Big|_0^{0.25}\\ &= 0.125\\ \mathbb{E}(Y) &= 0.125 \cdot 1 + 0.875 \cdot -1\\ &= -0.75\end{aligned}$$

X_1 and Y are independent.

(f) X_1 and Z are independent.
$\mathbb{E}(X_1 Z) = 0$, since $\mathbb{E}(X) = 0$