Introduction to Machine Learning (CSCI-UA.473): Homework 0

Instructor: Sumit Chopra

September 13, 2021

Questions

Probability and Calculus

Question 1 (10 points) Two people take turns trying to sink a basketball into a net. Person 1 succeeds with probability 1/3 and Person 2 succeeds with the probability 1/4. Whoever succeeds first wins the game and the game is over. Assuming that Person 1 takes the first shot, what is the probability that Person 1 wins the game? Please derive your answer.

$$
\begin{aligned}
P(\text{Person 1 wins on 1st shot}) &= \frac{1}{3} \\
P(\text{Person 1 wins on 2nd shot}) &= \left(1-\frac{1}{3}\right)\left(1-\frac{1}{4}\right)\frac{1}{3} \\
P(\text{Person 1 wins on 3rd shot}) &= \left(1-\frac{1}{3}\right)^2\left(1-\frac{1}{4}\right)^2\frac{1}{3} \\
\text{Following this pattern:} & \\
P(\text{Person 1 wins on } n\text{-th shot}) &= \left(1-\frac{1}{3}\right)^{n-1}\left(1-\frac{1}{4}\right)^{n-1}\frac{1}{3}
\end{aligned}
$$

$$
\begin{aligned}
P(\text{Person 1 wins the game}) &= \sum_{n=1}^{\infty}\left(1-\frac{1}{3}\right)^{n-1}\left(1-\frac{1}{4}\right)^{n-1}\frac{1}{3} \\
&= \lim_{k\to\infty}\frac{1}{3}\times 2^{1-k}\times\left(2^k-1\right) \\
&= \frac{2}{3} \\
&= \boxed{0.66667}
\end{aligned}
$$

Question 2 (10 points) You know that 1% of the population have COVID. You also know that 90% of the people who have COVID get a positive test result and 10% of people who do not have COVID also test positive. What is the probability that you have COVID given that you tested positive?

Let TP be the event that a person is tested positive, C be the event that a person has COVID.

Answer: Given $P(C) = 0.01, (TP|C) = 0.9, (TP|C^C) = 0.1$.
We compute

$$P(NC) = P(C^C) = 1 - 0.01 = 0.99$$

$$
\begin{aligned}
P(TP) &= P(TP|C)P(C) + P(TP|C^C)P(C^C) \\
&= 0.9 \times 0.01 + 0.1 \times 0.99 \\
&= 0.108
\end{aligned}
$$

$$P(C|TP) = \frac{P(TP|C)P(C)}{P(TP)} = \frac{0.9 \times 0.01}{0.108} = \boxed{0.083333}$$

Question 3 (10 points) Let the function $f(x)$ be defined as:

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{(1+x)} & \text{otherwise} \end{cases}$$

Is $f(x)$ a PDF? If yes, then prove that it is a PDF. If no, then prove that it is not a PDF.

Answer:

$$\int_0^\infty \frac{1}{x+1} dx = \log(x+1)|_0^\infty = \infty \neq 1$$

Since the integral does not converge, $f(x)$ is not a PDF.

Question 4 (10 points) Assume that X and Y are two independent random variables and both have the same density function:

$$f(x) = \begin{cases} 2x & \text{if} \quad 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of $\mathbb{P}(X+Y \leq 1)$?

Answer:

$$\begin{aligned} \mathbb{P}(X+Y \leq 1) &= F_{X+Y} \\ &= \int_0^1 \int_0^{1-y} (2x)(2y) dx dy \\ &= \boxed{\frac{1}{6}} \end{aligned}$$

Question 5 (10 points) Let X be a random variable which belongs to a Uniform distribution between 0 and $1: X \sim \text{Unif}(0,1)$. Let $Y = g(X) = e^X$. What is the value of $\mathbb{E}(Y)$?

Answer: Given

$$f(x) = \frac{1}{1-0} = 1 \text{ for } 0 \leq x \leq 1, \; g(x) = e^x$$

By the law of the unconscious statistician

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(g(X)) \\ &= \int_0^1 g(x) f(x) dx \\ &= \int_0^1 e^x 1 dx \\ &= e - 1 \\ &= \boxed{1.7183} \end{aligned}$$

Question 6 (10 points) Suppose that the number of errors per computer program has a Poisson distribution with mean 5 . We have 125 program submissions. Let $X_1, X_2, \ldots, X_{125}$ denote the number of errors in the programs. What is the value of $\mathbb{P}\left(\bar{X}_n < 5.5\right)$?

Answer: Since $n = 125$ is large, we can apply normal approximation with $\mu = \sigma^2 = \lambda = 5$ to the sampling Poisson distribution.

$$z^* = \frac{x - \mu}{\sqrt{\sigma^2/n}} = \frac{5.5 - 5}{\sqrt{5/125}} = 2.5 \text{ at } x = 5.5$$

$$\mathbb{P}(z < z^*) = \boxed{0.99379}$$

Question 7 (10 points) Let $X_n = f\left(W_n, X_{n-1}\right)$ for $n = 1, \ldots, P$, for some function $f()$. Let us define the value of variable E as

$$E = \left\| C - X_P \right\|^2$$

for some constant C. What is the value of the gradient $\frac{\partial E}{\partial X_0}$?

Answer: Applying Chain Rule

$$\begin{aligned}
\frac{\partial E}{\partial X_0} &= \frac{\partial \left\| C - X_P \right\|^2}{\partial X_0} \\
&= \frac{\partial (C - f(W_n, X_P - 1))}{\partial X_0} \\
&= \boxed{\frac{\partial f}{\partial X_0}(W_n, X_P - 1)\frac{\partial f}{\partial X_0}(W_n, X_P - 2)\cdots\frac{\partial f}{\partial X_0}(W_n, X_1)}
\end{aligned}$$

Question 8 (10 points) Suppose

$$f(x,y) = \frac{\sqrt{x+y}+\sqrt{x-y}}{\sqrt{x+y}-\sqrt{x-y}} + \sqrt{\sqrt{x+y}-\sqrt{x-y}}.$$

What is the value of the expression

$$2y\frac{\partial^2 f}{\partial x^2} + 4x\frac{\partial^2 f}{\partial x \partial y} + 2y\frac{\partial^2 f}{\partial y^2} + 2\frac{\partial f}{\partial y}$$

at the point where $x = 5$ and $y = 4$?

Answer:

$$2y\frac{\partial^2 f}{\partial x^2} + 4x\frac{\partial^2 f}{\partial x \partial y} + 2y\frac{\partial^2 f}{\partial y^2} + 2\frac{\partial f}{\partial y}$$

$$= 4x\left(-\frac{\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)^2}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2} + \frac{2\left(\sqrt{x-y}+\sqrt{x+y}\right)\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^3} \right.$$

$$- \frac{\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)}{4\left(\sqrt{x+y}-\sqrt{x-y}\right)^{3/2}} + \frac{\frac{1}{4(x-y)^{3/2}}-\frac{1}{4(x+y)^{3/2}}}{\sqrt{x+y}-\sqrt{x-y}} - \frac{\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)^2}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2}$$

$$\left. - \frac{\left(-\frac{1}{4(x+y)^{3/2}}-\frac{1}{4(x-y)^{3/2}}\right)\left(\sqrt{x-y}+\sqrt{x+y}\right)}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2} + \frac{-\frac{1}{4(x+y)^{3/2}}-\frac{1}{4(x-y)^{3/2}}}{2\sqrt{\sqrt{x+y}-\sqrt{x-y}}} \right)$$

$$+ 2\left(-\frac{\left(\sqrt{x-y}+\sqrt{x+y}\right)\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2} + \frac{\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}}{2\sqrt{\sqrt{x+y}-\sqrt{x-y}}} + \frac{\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}}{\sqrt{x+y}-\sqrt{x-y}} \right)$$

$$+2y\left(-\frac{\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)^2}{4\left(\sqrt{x+y}-\sqrt{x-y}\right)^{3/2}} - \frac{2\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2} + \left(\sqrt{x-y}+\sqrt{x+y}\right)\left(\frac{2\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)^2}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^3} - \frac{1}{4(x-y} \right.\right.$$

$$+2y\left(-\frac{\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)^2}{4\left(\sqrt{x+y}-\sqrt{x-y}\right)^{3/2}} - \frac{2\left(\frac{1}{2\sqrt{x+y}}+\frac{1}{2\sqrt{x-y}}\right)\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^2} + \left(\sqrt{x-y}+\sqrt{x+y}\right)\left(\frac{2\left(\frac{1}{2\sqrt{x+y}}-\frac{1}{2\sqrt{x-y}}\right)^2}{\left(\sqrt{x+y}-\sqrt{x-y}\right)^3} - \frac{1}{4(x-y} \right.\right.$$

$$= 2\left(\frac{1}{3\sqrt{2}} - \frac{5}{6}\right) + 8\left(\frac{5}{27} + \frac{7}{108\sqrt{2}}\right) + 20\left(\frac{1}{54} - \frac{11}{108\sqrt{2}}\right) + 8\left(\frac{23}{216\sqrt{2}} - \frac{4}{27}\right)$$

$$= \boxed{-1} \text{ at } x = 5, y = 4$$

Linear Algebra

Question 9 (10 points) What is an eigenvalue of a matrix? What is an eigenvector of a matrix? Describe one method (any method) you would use to compute both of them. Use the above described method to compute the eigenvalues of the matrix:

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 0 \\ -2 & 2 & 1 \end{bmatrix}$$

Definitions: An eigenvector of an $n \times n$ matrix A is a nonzero vector $\vec{x}$, such that $A\vec{x} = \lambda\vec{x}$ for a scalar eigenvalue λ.

Geometrically, when a matrix is applied onto a vector as a linear transformation, an eigenvector of the matrix points to the direction stretched or compressed by a factor of its corresponding eigenvalue .

Method: To find the eigenvectors and eigenvalues of a matrix A, we first solve $\det(A - \lambda I)$ for λ (eigenvalue), then for each λ, solve $(A - \lambda_1 I)\vec{x} = \vec{0}$ for $\vec{x}$ (eigenvector).

Computation: Given matrix $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 0 \\ -2 & 2 & 1 \end{bmatrix}$

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

Solve $\det(A - \lambda I) = 0$ for λ (Eigenvalues).

$$\Rightarrow \boxed{\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 2}$$

For $\lambda_1 = 1$:

$$A - \lambda_1 I = \begin{bmatrix} 0 & 0 & -1 \\ 1 & -1 & 0 \\ -2 & 2 & 0 \end{bmatrix}$$

Solve $(A - \lambda_1 I)\vec{x} = \vec{0}$ for $\vec{x}$ (Eigenvectors):

$$A - \lambda_1 I = 0 \Rightarrow \left[\begin{array}{ccc|c} 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \begin{cases} x_1 - x_2 & = 0 \\ x_3 & = 0 \end{cases} \Rightarrow \begin{cases} x_1 & = x_2 \\ x_3 & = 0 \end{cases}$$

$$\Rightarrow \boxed{\vec{x}_1 = (1, 1, 0)}$$

For $\lambda_2 = -1$:

$$A - \lambda_2 I = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -2 & 2 & 2 \end{bmatrix} = 0 \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1/2 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \Rightarrow \begin{cases} x_1 - \frac{1}{2}x_3 & = 0 \\ x_2 + \frac{1}{2}x_3 & = 0 \\ x_3 & = 1 \end{cases}$$

$$\Rightarrow \boxed{\vec{x}_2 = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)}$$

For $\lambda_3 = 2$:

$$A - \lambda_3 I = \begin{bmatrix} -1 & 0 & -1 \\ 1 & -2 & 0 \\ -2 & 2 & -1 \end{bmatrix} = 0 \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{cases} x_1 - +x_3 & = 0 \\ x_2 - \frac{1}{2}x_3 & = 0 \\ x_3 & = x_3 \end{cases}$$

$$\Rightarrow \boxed{\vec{x}_3 = \left(-1, -\frac{1}{2}, 1\right)} \text{ assuming } x_3 = 1$$

Question 10 (10 points) Let $X = (x_1, \ldots, x_k)$ for some fixed k, be a random variable whose probability density function is defined as:

$$f(x) = \left(\begin{array}{c} n \\ x_1, \ldots, x_k \end{array} \right) p_1^{x_1}, \ldots, p_k^{x_k}$$

where

$$\left(\begin{array}{c} n \\ x_1, \ldots, x_k \end{array} \right) = \frac{n!}{x_1!, \ldots x_k!}$$

Also $p_j \geq 0$ for all $j = \{1, \ldots, k\}$ and $\sum_{j=1}^{k} p_j = 1$. What is the value of $\mathbb{E}(X)$ and $\mathbb{V}(X)$?

Answer: Observe that $f(x)$ is in the shape of a Multinomial probability mass function

$$\mathbb{E}(X_i) = np_i$$
$$\mathbb{E}(X) = \{\mathbb{E}(X_1), ..., \mathbb{E}(X_k)\} = \{np_1, ..., np_k\}$$

$$\mathbb{V}(X_i) = np_i(1 - p_i)$$
$$\mathbb{V}(X) = \{\mathbb{V}(X_1), ..., \mathbb{V}(X_k)\} = \{np_1(1 - p_1), ..., np_k(1 - p_k)\}$$