IFT 6390 Fundamentals of Machine Learning
Guillaume Lam

Homework 0

1 Theoretical Part [6 points]

1. [1 points] Let X be a random variable representing the outcome of a single roll of a 6-sided dice. Show the steps for the calculation of i) the expectation of X and ii) the variance of X.

   Solution:
   i) $\mathbb{E}(X) = \sum xp(x) = \mu$

$$\mu = 1 * \frac{1}{6} + 2 * \frac{1}{6} + ... + 6 * \frac{1}{6}$$
$$\mu = \frac{21}{6} = 3.5$$

   ii) $\mathbb{V}(X) = \sum (x-\mu)^2 p(x) = \sigma^2$

$$\begin{aligned}\sigma^2 &= (1-3.5)^2 * \frac{1}{6} + (2-3.5)^2 * \frac{1}{6} + ... + (6-3.5)^2 * \frac{1}{6} \\ &= \frac{6.25}{6} + \frac{2.25}{6} + \frac{0.25}{6} + \frac{0.25}{6} + \frac{2.25}{6} + \frac{6.25}{6} \\ \sigma^2 &= \frac{17.5}{6} = 2.91\overline{6}\end{aligned}$$

2. [1 points] Let $u, v \in \mathbb{R}^d$ be two vectors and let $A \in \mathbb{R}^{n \times d}$ be a matrix. Give the formulas for the euclidean norm of u, for the euclidean inner product (aka dot product) between u and v, and for the matrix-vector product Au.

   Solution:
   Euclidiean norm of u: $||u|| = \sqrt{\sum_{i=0}^{d-1} u_i^2}$

Euclidean inner product btwn u and v: $u \cdot v = \sum_{i=0}^{d-1} u_i v_i$

Matrix-vector product Au: $Au = \begin{bmatrix} \sum_{i=0}^{d-1} A_{0i} u_i \\ \sum_{i=0}^{d-1} A_{1i} u_i \\ \vdots \\ \sum_{i=0}^{d-1} A_{n-1i} u_i \end{bmatrix}$

3. [1 points] Consider the two algorithms below. What do they compute and which algorithm is faster?
Observez les deux algorithms ci-dessous. Que calculent-ils et lequel est le plus rapide ?

ALGO1(n)
```
result = 0
for i = 1...n
  result = result + i
return result
```

ALGO2(n)
```
return (n + 1) * n/2
```

Solution:

ALGO1(n)
`Time complexity:` $O(n)$
`Goal: sum of 1 to` n

ALGO2(n)
`Time complexity:` $O(1)$
`Goal: average of` n^2 `and` n

ALGO2 is faster than ALGO1

4. [1 points] Give the step-by-step derivation of the following derivatives:

$$\text{i)} \quad \frac{df}{dx} = ?, \quad \text{where } f(x, \beta) = x^2 \exp(-\beta x)$$

$$\text{ii)} \quad \frac{df}{d\beta} = ?, \quad \text{where } f(x, \beta) = x \exp(-\beta x)$$

$$\text{iii)} \quad \frac{df}{dx} = ?, \quad \text{where } f(x) = \sin(\exp(x^2))$$

Solution:
i)

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(x^2 \exp(-\beta x)) \\ &= \frac{\partial}{\partial x}(x^2) * \exp(-\beta x) + \frac{\partial}{\partial x}(\exp(-\beta x)) * x^2 \\ &= 2x \exp(-\beta x) - \beta x^2 \exp(-\beta x) \\ &= (2x - \beta x^2) \exp(-\beta x) \end{aligned}$$

ii)

$$
\begin{aligned}
\frac{\partial f}{\partial \beta} &= \frac{\partial}{\partial \beta}(x \exp(-\beta x)) \\
&= \frac{\partial}{\partial \beta}(x) * \exp(-\beta x) + \frac{\partial}{\partial \beta}(\exp(-\beta x)) * x \\
&= 0 * \exp(-\beta x) + x * -x * \exp(-\beta x) \\
&= -x^2 \exp(-\beta x)
\end{aligned}
$$

iii)

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(\sin(\exp(x^2))) \\
&= \cos(\exp(x^2)) * \frac{\partial}{\partial x}(\exp(x^2)) \\
&= \cos(\exp(x^2)) * \exp(x^2) * \frac{\partial}{\partial x}(x^2) \\
&= \cos(\exp(x^2)) * \exp(x^2) * 2x \\
&= 2x \exp(x^2) \cos(\exp(x^2))
\end{aligned}
$$

5. [1 points] Let $X \sim N(\mu, 1)$, that is the random variable X is distributed according to a Gaussian with mean μ and standard deviation 1. Show how you can calculate the second moment of X, given by $\mathbb{E}[X^2]$.

   Solution:
   By using the definition of variance: $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$, and the characteristics of a Gaussian distribution, we can replace the values as such:

$$
\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\
1 &= \mathbb{E}(X^2) - (\mu)^2 \\
\mathbb{E}(X^2) &= 1 - \mu^2
\end{aligned}
$$