IFT 6390 Fundamentals of Machine Learning
Ioannis Mitliagkas

Homework 0 (Matteo Esposito)

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.

   i)

$$E[X] = \sum_{i=1}^{6} P(X = i) * i = \left(\frac{1}{6} * 1 + \frac{1}{6} * 2 + \ldots + \frac{1}{6} * 6\right) = 3.5$$

   ii)

$$V[X] = E[X^2] - (E[X])^2 = \sum_{i=1}^{6} P(X = i) * i^2 - (3.5)^2$$
$$= \left(\frac{1}{6} * 1^2 + \frac{1}{6} * 2^2 + \ldots + \frac{1}{6} * 6^2\right) - (3.5)^2 = \frac{105}{36}$$

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.

   i)

$$\|u\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_d^2}$$

   ii)

$$< u, v >= \sqrt{u_1 * v_1 + u_2 * v_2 + \ldots + u_d * v_d}$$

   iii)

$$A * u = a_{11} * u_1 + a_{12} * u_2 + a_{1d} * u_d + \ldots + a_{nd} * u_d$$

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
```

- Both algorithms calculate the sum of every integer from 1 to n. Where n is the only parameter passed to the 2 algorithms. ALGO2 will be quicker with a time complexity of $\mathcal{O}(1)$ since it performs a single operation (composed of several mathematical operations), compared to ALGO1 which has a time complexity of $\mathcal{O}(n)$ since it is implemented with a loop.

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

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

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

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

i)

$$\frac{d}{dx}(x^2 \exp(-\beta x)) = 2xe^{-\beta x} - \beta x^2 e^{-\beta x}$$

ii)

$$\frac{d}{d\beta}(x \exp(-\beta x)) = -x^2 e^{-\beta x}$$

iii)

$$\frac{d}{dx}(\sin(\exp(x^2))) = 2xe^{x^2} cos(e^{x^2})$$

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 $\mathbb{E}[X^2]$.

   i)
$$E[X^2] = V(X) + E[X]^2 = 1^2 - \mu^2$$

2 Practical Part [7 points]

1. [1 points] Create a numpy array from a python list
2. [1 points] Create a numpy array of length 1 from a python number
3. [1 points] Sum two arrays elementwise
4. [1 points] Sum an array and a number
5. [1 points] Mutliple two arrays elementwise
6. [1 points] Dot product of two arrays
7. [1 points] Dot product of an array (1D) and a matrix (2D array)