IFT 6390 Fundamentals of Machine Learning
Ioannis Mitliagkas

Homework 0

Saurabh Bodhe - 20208545

1 Theoretical Part [5 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.

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.

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
```

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

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

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

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

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

2 Solutions

1. $$X = \{1, 2, 3, 4, 5, 6\}$$

   Assuming that the dice is fair, each of the 6 outcomes will be equally likely. Hence,
   $$p(X = x) = 1/6 \quad \forall x \in X$$

   (a) Expectation of X

   $$\begin{aligned} \mu = E[X] &= \sum_{x=1}^{6} xp(x) \\ &= \frac{1}{6}\sum_{x=1}^{6} x \\ &= \frac{21}{6} = \frac{7}{2} = 3.5 \end{aligned}$$

   (b) Variance of X

   $$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= \sum_{x=1}^{6} x^2 p(x) - \left(\frac{7}{2}\right)^2 \\ &= \frac{1}{6}\sum_{x=1}^{6} x^2 - \frac{49}{4} \\ &= \frac{1}{6} \cdot \frac{6 \times 7 \times 13}{6} - \frac{49}{4} \\ &= \frac{35}{12} = 2.917 \end{aligned}$$

2. (a) Euclidean norm of u

   $$||u||_2 = \sqrt{\sum_{i=1}^{d} u_i^2}$$

   (b) Euclidean inner product of u & v

   $$u \cdot v = \sum_{i=1}^{d} u_i v_i$$

(c) matrix-vector product Au

$$Au = C \quad where \quad c_i = \sum_{l=1}^{d} a_{il} u_l \quad i \in [1, n]$$

3. Both the algorithms compute the sum of first n natural numbers. **ALGO1** has loop which runs n times, hence has time complexity $O(n)$.
   ALGO2 has a single operation, hence has time complexity $O(1)$.
   Therefore, **ALGO2** is faster.

4. (a)

$$\begin{aligned} f(x, \beta) &= x^2 e^{-\beta x} \\ \frac{df}{dx} &= \frac{d}{dx}(x^2 e^{-\beta x}) \\ &= e^{-\beta x}(2x - \beta) \end{aligned}$$

(b)

$$\begin{aligned} f(x, \beta) &= x e^{-\beta x} \\ \frac{df}{d\beta} &= \frac{d}{d\beta}(x e^{-\beta x}) \\ &= -x^2 e^{-\beta x} \end{aligned}$$

(c)

$$\begin{aligned} f(x) &= \sin e^{x^2} \\ \frac{df}{dx} &= \frac{d}{dx}(\sin e^{x^2}) \\ &= 2x e^{x^2} \cos e^{x^2} \end{aligned}$$

5.

$$V[X] = \sigma^2 = 1^2 = 1$$

Also,

$$\begin{aligned} V[X] &= E[(X - \mu)^2] \\ &= E[X^2 + \mu^2 - 2X\mu] \\ &= E[X^2] - 2\mu E[X] + E[\mu^2] \\ &= E[X^2] - 2[E[X]]^2 + [E[X]]^2 \\ &= E[X^2] - [E[X]]^2 \\ \implies E[X^2] &= V[X] + [E[X]]^2 \\ &= 1 + \mu^2 \end{aligned}$$