IFT 6390 Fundamentals of Machine Learning
Ioannis Mitliagkas

Homework 0

Solutions

1. **Question.** 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.

   Answer

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

   (a) $\mathbb{E}[x]$

   $$\mathbb{E}[x] = \sum_{i=1}^{k} x_i p_i = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = 3.5 \quad (1)$$

   (b) $\sigma^2(x)$

   $$\begin{aligned} &\sigma^2(x) = \mathbb{E}[x^2] - \mathbb{E}[x]^2 \\ &\mathbb{E}[x^2] = 1 * \frac{1}{6} + 4 * \frac{1}{6} + 9 * \frac{1}{6} + 16 * \frac{1}{6} + 25 * \frac{1}{6} + 36 * \frac{1}{6} = 15.16 \\ &\mathbb{E}[x]^2 = (3.5)^2 = 12.25 \\ &\sigma^2(x) = 15.16 - 12.25 = 2.91 \end{aligned} \quad (2)$$

2. **Question.** 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.

   Answer.

   (a) Euclidean norm:

   $$\|u\| = \sqrt{\sum_{n=1}^{d} u_i^2} \quad (3)$$

(b) Dot product:

$$u \bullet v = \sum_{n=1}^{d} u_i v_i \tag{4}$$

(c) Matrix-Vector product:

$$Au = \sum_{n=1}^{d} A_{i,j} u_j \tag{5}$$

3. **Question.** Consider the two algorithms below. What do they compute and which algorithm is faster?

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

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

Answer. Time complexity of **ALGO1** $= O(n)$ while time complexity of **ALGO2** $= O(1)$, hence the latter is faster than the former.

4. **Question.** Give the step-by-step derivation of the following derivatives:

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

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

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

Answer.

(a)

$$\begin{aligned}
\frac{d(x^2 exp(-\beta x)}{dx} &= x^2 \frac{d(exp(-\beta x)}{dx} + exp(-\beta x)\frac{d(x^2)}{dx} \\
&= x^2 exp(-\beta x)\frac{d(-\beta x)}{dx} + 2x exp(-\beta x) \\
&= -\beta x^2 exp(-\beta x) + 2x exp(-\beta x) \\
&= x exp(-\beta x)[2 - \beta x]
\end{aligned} \tag{6}$$

(b)

$$\begin{aligned} \frac{d(xexp(-\beta x)}{d\beta} &= x\frac{d(exp(-\beta x)}{d\beta} \\ &= xexp(-\beta x)\frac{d(-\beta x)}{d\beta} \\ &= -x^2 exp(-\beta x) \end{aligned} \tag{7}$$

(c)

$$\begin{aligned} \frac{d(sin(exp(x^2)}{dx} &= cos(exp(x^2))\frac{d(exp(x^2)}{dx} \\ &= cos(exp(x^2))exp(x^2)\frac{d(x^2)}{dx} \\ &= cos(exp(x^2))exp(x^2)(2x) \\ &= (2x)exp(x^2)cos(exp(x^2)) \end{aligned} \tag{8}$$

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

   Answer.

$$\begin{aligned} &\sigma^2(x) = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = 1 \\ &\mathbb{E}[x]^2 = \mu^2 \\ &\mathbb{E}[x^2] = 1 + \mu^2 \end{aligned} \tag{9}$$