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

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

   Hence, The Expected Value of a random variable is:

   (a) Expectation of X

   $E(X) = p_1(P = X_1) + p_2(P = X_2) + p_3(P = X_3) + ... + p_n(P = X_n)$
   $E(X) = 1.(1/6) + 2.(1/6) + 3.(1/6) + 4.(1/6) + 5.(1/6) + 6.(1/6) =$ **3.5**

   (b) Variance of X

   $E(X^2) = 1.(1/6)+4.(1/6)+9.(1/6)+16.(1/6)+25.(1/6)+36.(1/6) = 15.16$
   $E^2(X) = (3.5)^2 = 12.25$
   $Var(X) = E(X^2) - E^2(X) =$ **15.16 - 12.25 = 2.91**

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 of u -

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

(b) Euclidean inner product (aka dot product) -

$$u.v = \sum_{n=1}^{d} u_i v_i$$

(c) Matrix-Vector product -

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

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. Both the algorithms compute the sum of numbers from 1 to n. In the first algorithm (ALGO1) the sum is calculated using a for loop and hence the complexity is O(n). However, in the second algorithm (ALGO2) direct formula is applied and hence the time complexity is O(1). Hence, second algorithm (ALGO2) is faster than first one (ALGO1).

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

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

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

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

Answer.

(a)

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

(b)

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

(c)

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

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.

$$\begin{aligned}Var(x) &= E[X^2] - E[X]^2 \\ Var(x) &= (StdDeviation)^2 = 1 \\ E[X]^2 &= (Mean)^2 = \mu^2 \\ E[X^2] &= 1 + \mu^2\end{aligned}$$