IFT-3395/6390 Fundamentals of Machine Learning
Professors: Guillaume Rabusseau / Ioannis Mitliagkas

Homework 0

Devoir 0

1. Undergraduates 1 pts Graduates 1 pts
   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.

   $$Expectation : E(X) = p_1(P = X_1) + p_2(P = X_2) + ...p_n(P = X_n)$$
   $$E(X) = 1.1/6 + 2.1/6 + 3.1/6 + 4.1/6 + 5.1/6 + 6.1/6 = 3.25$$
   $$E(X^2) = 1.1/6 + 4.1/6 + 9.1/6 + 16.1/6 + 25.1/6 + 36.1/6 = 15.16$$
   $$E(X)^2 = 3.25 * 3.25 = 10.56$$
   $$Variance : Var(x) = E(X^2) - E(X)^2 = 15.16 - 10.56 = 4.5$$

2. Undergraduates 1 pts Graduates 1 pts
   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.

   $$||v||_k = [\sum_{k=1}^{d} |v_k|^p]^{\frac{1}{p}}$$

   if p = 1, then the resulting 1-norm is the sum of all absolute values.
   if p = 2, then the resulting 2-norm is euclidean length of the vector.

   Euclidean inner product:

   $$< u, v >= u_1v_1 + u_2v_2 + u_3v_3...u_dv_d$$
   $$< A, u >= \sum_{k=1}^{d} A_{nd}u_d$$

3. *Undergraduates 1 pts Graduates 1 pts*
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. *Yes they both compute and ALGO2 is faster as the O(n)=1*

4. *Undergraduates 1 pts Graduates 1 pts*
Question. *Give the step-by-step derivation of the following derivatives:*

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

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

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

Answer.

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

$$ii)\quad f(x,\beta) = x \exp(-\beta x) \quad \frac{df}{d\beta} = -x^2 exp(-\beta x)$$

$$iii)\quad f(x) = \sin(\exp(x^2)) \quad \frac{df}{dx} = 2xexp(x^2)cos(exp(x^2))$$

5. *Undergraduates 1 pts Graduates 1 pts*
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.

$$Var(x) = E(X^2) - E(X)^2$$

$$Var(x) = (StdDeviation)^2 = 1$$

$$E(X)^2 = (Mean)^2 = \mu^2$$

$$E[X^2] = 1 + \mu^2$$