IFT 6390 fundamentals of Machine learning

Homework 0

Faezeh Pouya Mehr

1.

i)

$E(X) = \sum_{i=1}^{6} i * P(x = i) = \sum 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = \frac{7}{2} = 3.5$

ii)

$Var(X) = E(X^2) - E(X)^2 = \sum_{i=1}^{6} i^2 * p\ (x = i) - (3.5)^2 = \sum 1 * \frac{1}{6} + 4 * \frac{1}{6} + 9 * \frac{1}{6} + 16 * \frac{1}{6} + 25 * \frac{1}{6} + 36 * \frac{1}{6} - (3.5)^2 = \frac{91}{6} - (3.5)^2 = 2.91$

2.

Euclidian norm:

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

Dot product:

$U^T V = \sum_{i=1}^{d} u_i v_i$

Matrix-vector product:

$Au = B$ where $B_i = \sum_{j=1}^{d} A_{i,j} * u_i$ where $B \epsilon R^n$ and $1 <= I <= n$

3. they both compute sum of numbers from 1 to n. the second algorithm is faster since it just needs to compute basic +,*and / operations in O(1) however algorithm 1 performs + operation n times, makes it o(n) complexity.

4.

$F(x,\beta) = x^2 e^{-\beta x}$

$\frac{df}{dx} = 2xe^{-\beta x}$ - b $x^2 e^{-\beta x}$

$F(x,\beta) = xe^{-\beta x}$

$\frac{df}{dx} = e^{-\beta x}$ - b $xe^{-\beta x}$

$F(x,\beta) = \sin(e^{x^2})$

$\frac{df}{dx} = 2xe^{x^2} \cos(e^{x^2})$

5.

$Var(x)=1 = E(X^2) - E(X)^2 \Rightarrow E(X^2) = \mu + 1$