CSCE 421
HW 1
Jeffrey Xu
09/07/20

1.1. The gradient of a multi-variate function is a vector of the respective partials of the function. Our given function is shown below:

$$f(x, y) = x^2 + ln(x) + xy + y^3$$

The gradient vector is as shown below:

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix}$$

The gradient values are shown below:

$$f_x = 2x + x^{-1} + xy + y$$

$$f_y = x + 3y^2$$

$$\begin{bmatrix} 2x + x^{-1} + xy + y \\ x + 3y^2 \end{bmatrix}$$

The gradient value for $(1, -1)$ can be computed by plugging those values into the gradient.

$$\bigtriangledown f(1, -1) = \begin{bmatrix} 2(1) + 1^{-1} + 1(-1) + (-1) \\ 1 + 3(-1)^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

1.2. Given the definition of the gradient function, we need to calculate the gradient for the given function below:

$$f(x, y, z) = tanh(x^3y^3) + sin(z)$$

The partials of the function are shown below:

$$f_x = 3y^3x^2sech^2(x^3y^3)$$

$$f_y = 3x^3y^2sech^2(x^3y^3)$$

$$f_z = cos(z)$$

$$\bigtriangledown f = \begin{bmatrix} 3y^3x^2sech^2(x^3y^3) \\ 3x^3y^2sech^2(x^3y^3) \\ cos(z) \end{bmatrix}$$

The gradient value at $(-1, 0, \pi/2)$ is shown below:

$$\bigtriangledown f(-1, 0, \pi/2) = \begin{bmatrix} (-1)^2(0)^3sech^2((-1)^3(0)^3) \\ (-1)^2(0)^3sech^2((-1)^3(0)^3) \\ cos(\pi/2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

2.1. The following matrix multiplication computation is shown below:

$$\begin{bmatrix} 1 & -1 & 6 & 7 \\ 9 & 0 & 8 & 1 \\ -8 & 5 & 2 & 3 \\ 10 & 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 0 & -1 \\ -3 & 0 \\ 11 & 4 \end{bmatrix} =$$

$$\begin{bmatrix} 1(6) - 1(0) - 3(6) + 7(11) & 1(2) - 1(-1) + 6(0) + 7(4) \\ 6(9) + 0(0) - 3(8) + 11 & 9(2) - 1(0) + 8(0) + 4 \\ -8(6) + 5(0) - 3(2) + 3(11) & -8(2) - 5 + 2(0) + 3(4) \\ 10(6) + 4(0) - 3(0) + 11 & 2(10) - 4 + 4 \end{bmatrix} = \begin{bmatrix} 65 & 31 \\ 41 & 22 \\ -21 & -9 \\ 71 & 20 \end{bmatrix}$$

2.2.

$$\begin{bmatrix} 10 \\ 4 \\ -1 \\ 8 \end{bmatrix} \begin{bmatrix} 7 & 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10(7) & 3(10) & 0(10) & 1(10) \\ 4(7) & 3(4) & 0(4) & 1(4) \\ -7 & -3 & 0 & -1 \\ 8(7) & 3(8) & 0(8) & 1(8) \end{bmatrix} \begin{bmatrix} 70 & 30 & 0 & 10 \\ 28 & 12 & 0 & 4 \\ -7 & -3 & 0 & -1 \\ 56 & 24 & 0 & 8 \end{bmatrix}$$

2.3

$$\begin{bmatrix} 9 & -3 & 1 & 6 \end{bmatrix} \begin{bmatrix} -3 \\ 4 \\ -9 \\ 0 \end{bmatrix} = 9(-3) - 3(4) + 1(-9) + 6(0) = -48$$

3.1. We want to calculate the l_0 norm for each vector given. The two vectors are shown below:

$$\mathbf{a} = \begin{bmatrix} 5 \\ 0 \\ -1 \\ 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 9 \\ 5 \\ 2 \end{bmatrix}$$

The l_0 norm counts the number of non-zero entries in the vector. For **a**, we can see $l_0 = 3$ and for **b**, $l_0 = 4$.

3.2. The l_1 norm is the sum of the absolute values of each element in the vector:

$$l_1 = \sum_{i=1}^{n} |a_n|$$

For **a**, $l_1 = |5| + |0| + |-1| + |4| = 10$ and for **b**, $l_1 = |7| + |9| + |5| + |2| = 23$.

3.3. The l_2 norm is simply the square root of the sum of squares of each element:

$$l_2 = \sqrt{\sum_{i=1}^{n} a_i^2}$$

The values of the l_2 norm for each vector is shown below:

$$l_2(\mathbf{a}) = \sqrt{5^2 + 0^2 + (-1)^2 + 4^2} = \sqrt{42}$$
$$l_2(\mathbf{b}) = \sqrt{7^2 + 9^2 + 5^2 + 2^2} = \sqrt{159}$$

3.4. The l_∞ norm computes the absolute maximal element in the vector.

$$l_\infty = max_n |x_n|$$

For **a**, $l_\infty = 5$ and for **b**, $l_\infty = 9$.

4.1. For rolling two dices where each has six faces numbered 1 to 6, the sample space is $\mathbb{Z}[1,6] * \mathbb{Z}[1,6]$, or the cartesian product of the integers from 1 to 6 with itself. This produces pairs of integers $(1,1), (1,2), (1,3), ..., (6,6)$. Note that this notation does specify that the first coordinate denotes the value of the first dice and the second coordinate denotes the value of the second dice.

4.2. Our sample space has cardinality of 36. Therefore, the probability of any event X can be calculated with the following equation.

$$\mathbb{P}(X) = \frac{|X|}{36}$$

We need to find all the elements of the sample space that satisfy the property that the value of both dices add up to 10. We notice that $(5,5), (6,4), (4,6)$ satisfy this condition. Therefore $\mathbb{P}(\text{Sum of dices equals } 10) = \frac{3}{36} = \frac{1}{12}$.

4.3. This problem follows similar logic to the previous question. The elements of Ω that sum to 6 are $(1,5), (5,1), (2,4), (4,2), (3,3)$. There are 5 elements in this subset, therefore the probability of the sum equalling 6 is $\frac{5}{36}$.

5.1. Recall that the expectation of a continuous random variable can be computed as shown below.

$$E[X] = \int_{-\infty}^{\infty} xp(x)dx$$

We notice for the Uniform distribution, it is only defined for the interval $[a, b]$. The computation for the expectation of the Uniform distribution is shown below.

$$E[X] = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\int_a^b xdx =$$

$$\frac{1}{b-a}[\frac{1}{2}x^2]_a^b = \frac{1}{2(b-a)}[b^2 - a^2] = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$$

5.2. The standard deviation of a distribution is the square root of the variance of the distribution. The variance of a distribution can be computed using the following equation.

$$Var(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

The only extra value we need to compute is $E[X^2]$. Note, $E[f(x)] = \int f(x)p(x)dx$.

$$E[X^2] = \frac{1}{b-a}\int_a^b x^2dx = \frac{1}{b-a}[\frac{1}{3}x^3]_a^b = \frac{1}{3(b-a)}[b^3 - a^3] = \frac{(b-a)(a^2+ab+b^2)}{3(b-a)} =$$

$$\frac{a^2+ab+b^2}{3}$$

$$Var(X) = \frac{a^2+ab+b^2}{3} - (\frac{b+a}{2})^2 = \frac{a^2+ab+b^2}{3} - \frac{b^2+2ab+a^2}{4} =$$

$$\frac{4a^2 + 4ab + 4b^2}{12} - \frac{3b^2 + 6ab + 3a^2}{12} = \frac{b^2 - 2ab + a^2}{12} =$$

$$\frac{(b-a)^2}{12}$$

Therefore the standard deviation is equal to $\dfrac{b-a}{\sqrt{12}}$.

6.1. The computation for the accuracy of this detector is shown below.

$$A = \frac{37 + 55}{37 + 23 + 45 + 55} = 0.575$$

6.2. The balanced accuracy weighs each class equally to discourage skewed classification. Let $+1$ denote avocado and -1 denote no avocado.

$$A_{+1} = \frac{37}{37 + 23} = 0.617$$

$$A_{-1} = \frac{55}{45 + 55} = 0.55$$

$$A_b = \frac{A_{+1} + A_{-1}}{2} = \frac{0.617 + 0.55}{2} = 0.5833$$

6.3. The precision measures the number of true positives detected out of all positives detected.

$$P = \frac{TP}{TP + FP} = \frac{37}{37 + 45} = 0.45122$$

6.4. The recall measures the amount of true positives detected out of all positive values.

$$R = \frac{TP}{TP + FN} = \frac{37}{37 + 23} = 0.617$$

6.5. The equation for F-measure is shown below.

$$F_1 = 2\frac{PR}{P + R} = \frac{TP}{TP + 0.5(FP + FN)} = \frac{37}{37 + 0.5(45 + 23)} = 0.521$$

7.1.