Sri Naga Hansi Mamidi
20012906

CS-583 Deep learning Assignment-1

Q1. $x = [5, 0, 1, -2]^T$ is a 4-dimensional vector. Calculate the following values:

1. the squared l_2-norm of x:

$$\|x\|_2^2 = \left(\sqrt{5^2+0^2+(1)^2+(-2)^2}\right)^2 = 25+1+4 = 30$$

$$\left(\because \|x\|_2 = \sqrt{\Sigma x_i^2}\right)$$

2. the l_1-norm of x:

$$\|x\|_1 = \Sigma |x_i| = 5+0+1+2 = 8$$

3. the inner product of x and a, where $a = [4, -2, 6, -1]^T$:

$$x \cdot a = \begin{bmatrix} 5 \\ 0 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \\ 6 \\ -1 \end{bmatrix} = 5(4)+0(-2)+1(6)+(-2)(-1)$$

$$= 20+6+2 = 28.$$

Q2. The matrix $A \in R^{2\times3}$ and vector $b \in R^3$ are defined in the following:

$$A = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix} \text{ and } b = \begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix}$$

Calculate the following values:

1. The matrix-vector product:

$$Ab = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix} \begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 6(-4)+1(5)+2(0) \\ -5(-4)+0(5)+(-3)(0) \end{bmatrix}$$

$$= \begin{bmatrix} -24+5 \\ 20 \end{bmatrix} = \begin{bmatrix} -19 \\ 20 \end{bmatrix}$$

2. The matrix-matrix product:

$$AA^T = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix} \begin{bmatrix} 6 & -5 \\ 1 & 0 \\ 2 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 6(6)+1(1)+2(2) & 6(-5)+1(0)+2(-3) \\ -5(6)+0(1)-3(2) & -5(-5)+0+3(3) \end{bmatrix}$$

$$= \begin{bmatrix} 36+1+4 & -30-6 \\ -30-6 & 25+9 \end{bmatrix}$$

$$= \begin{bmatrix} 41 & -36 \\ -36 & 34 \end{bmatrix}$$

Q3. Let $x = [x_1, x_2, x_3]$ and $y = \frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}$. Calculate the following:

$\frac{dy}{dx}$ at $[9, \frac{1}{2}, \frac{1}{3}]$

Sol:

$$\frac{dy}{dx} = \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}\right) \\ \frac{\partial}{\partial x_2}\left(\frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}\right) \\ \frac{\partial}{\partial x_3}\left(\frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\cancel{2}x_1}{\cancel{2}} - \frac{1}{x_3} \\ \frac{1}{x_2} \\ x_1 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 - \frac{1}{x_3} \\ \frac{1}{x_2} \\ -x_1 \end{bmatrix}$$

$$\frac{dy}{dx} \text{ at } [9, \tfrac{1}{2}, \tfrac{1}{3}] = \begin{bmatrix} 9 - \frac{1}{\frac{1}{3}} \\ \frac{1}{1/2} \\ -9 \end{bmatrix} = \begin{bmatrix} 9-3 \\ 2 \\ -9 \end{bmatrix}$$

$$= \begin{bmatrix} 6 \\ 2 \\ -9 \end{bmatrix}$$

Q4. X is an $n \times d$ matrix, y is an $n \times 1$ vector, and w is an $d \times 1$ vector. Let $f(w) = \|Xw - y\|_2^2 + \lambda \|w\|_2^2$.

Calculate the following:

$$\frac{\partial f(w)}{\partial w} = \frac{\partial}{\partial w}\left[\|Xw - y\|_2^2 + \lambda \|w\|_2^2\right]$$

$$\|Xw - y\|^2 = w^T X^T X w - 2 w^T X^T y + y^T y$$

$$\therefore \frac{\partial}{\partial w} \|Xw - y\|^2 = \frac{\partial}{\partial w}(w^T X^T X w - 2 w^T X^T y + y^T y)$$

$$= -2X^T y + 2X^T X w$$

$$= 2(X^T X w - X^T y)$$

$$\frac{\partial}{\partial w}\left(\lambda \|w\|_2^2\right) = 2\lambda w$$

$$\therefore \frac{\partial f(w)}{\partial w} = 2\left(X^T X w - X^T y + \lambda w\right)$$