CS 583A Spring23

Sanjeet Vinod Jain
CWID: 20012768

Q1] $x = [5, 0, 1, -2]^T$ $\quad a = [4, -2, 6, -1]^T$

1.) $\|x\|_2^2 = \sum_{i=1}^{n} x_i^2 = 5^2 + 0^2 + 1^2 + (-2)^2$

$= 25 + 1 + 4$

$= 30$

2) l_1 norm $= \|x\|_1 = \sum_{i=1}^{n} |x_i| = 5 + 0 + 1 + 2$

$= 8$

3) inner product of x and $a \rightarrow a^T x$.

$$a^T x = \left([4, -2, 6, -1]^T\right)^T \cdot [5, 0, 1, -2]^T$$

$$= [4, -2, 6, -1]\begin{bmatrix} 5 \\ 0 \\ 1 \\ -2 \end{bmatrix} = 20 + 0 + (6)(1) + (-1)(-2)$$

$$= 20 + 6 + 2$$

$$= 28$$

Q2] $A = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix}_{2\times3}$ $\quad b = \begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix}_{\cancel{2\times1}}$

1) $Ab = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix}\begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} (6)(-4) + (1)(5) + (2)(0) \\ (-5)(-4) + 0 + 0 \end{bmatrix}_{2\times1}$

$$= \begin{bmatrix} -24 + 5 + 0 \\ 20 \end{bmatrix} = \begin{bmatrix} -19 \\ 20 \end{bmatrix}$$

2) $AA^T = \begin{bmatrix} 6 & 1 & 2 \\ -5 & 0 & -3 \end{bmatrix}_{2\times3} \begin{bmatrix} 6 & -5 \\ 1 & 0 \\ 2 & -3 \end{bmatrix}_{3\times2}$

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

$$= \begin{bmatrix} 36+1+4 & -30+0-6 \\ -30+0-6 & 25+0+9 \end{bmatrix} = \begin{bmatrix} 41 & -36 \\ -36 & 34 \end{bmatrix}$$

Q 3) $x = [x_1, x_2, x_3]$ $\quad y = \frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}$

$\frac{\partial y}{\partial x}$ at $x = [9, 1/2, 1/3]$

$$\frac{\partial y}{\partial x} = \frac{\partial}{\partial x}\left[\frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3}\right]$$

$$= \frac{\partial}{\partial x}\left(\frac{x_1^2}{2}\right) + \frac{\partial}{\partial x}(\log_e x_2) - \frac{\partial}{\partial x}\left(\frac{x_1}{x_3}\right)$$

$$\frac{\partial y}{\partial x} = \left[\frac{\partial y}{\partial x_1}, \frac{\partial y}{\partial x_2}, \frac{\partial y}{\partial x_3}\right]$$

$$\Rightarrow \frac{\partial y}{\partial x_1} = \frac{\partial}{\partial x_1}(x_1^2/2) + \frac{\partial}{\partial x_1}(\log_e x_2) - \frac{\partial}{\partial x}(x_1/x_3)$$

$$= \frac{2x_1}{2} + 0 - \frac{1}{x_3} = x_1 - \frac{1}{x_3}$$

at $\Rightarrow \partial y = 9 - \frac{1}{\ }$ $= 9-3 = 6$ — (1)

Now $\frac{\partial y}{\partial x_2} = \frac{\partial (x_1^2/2)}{\partial x_2} + \frac{\partial \log_e x_2}{\partial x_2} - \frac{\partial (x_1/x_3)}{\partial x_2}$

$= \frac{\partial \log_e x_2}{\partial x_2} = \frac{1}{x_2} = \frac{1}{1/2} = 2$ — ②

Now $\frac{\partial y}{\partial x_3} = \frac{-\partial (x_1/x_3)}{\partial x_3} = -x_1 \left(\frac{\partial (1/x_3)}{\partial x_3} \right)$

$= -x_1 \left(\frac{-1}{x_3^2} \right) = \frac{x_1}{x_3^2} = \frac{9}{(1/3)^2} = 9 \times 9 = 81$ — ③

Combining ①, ② & ③ we get

$\frac{\partial y}{\partial x} = [6, 2, 81]$

Q4] $f(w) = ||Xw - y||_2^2 + \lambda ||w||_2^2$

Generally squared l_2 norm for vector y given as

$$||y||_2^2 = \sum_{k=1}^{n} y_k^2$$

Taking derivative $\frac{\partial ||y||_2^2}{\partial y_j} = \frac{\partial}{\partial y_j}\left[\sum_{k=1}^{n} y_k^2\right]$

$$= \cancel{\sum_{k=1}^{n} \frac{\partial}{\partial y_j} y_j^2 = 2y_j}$$

$$= \sum_{k=1}^{n} \underbrace{\frac{\partial}{\partial y_j}\left(y_k^2\right)}_{\substack{=0 \text{ if } j \neq k \\ =2y_j \text{ else}}} = 2y_j$$

$\therefore \frac{\partial ||w||_2^2}{\partial w} = 2w$ & $\frac{\partial ||w||_2^2}{\partial w_j} = 2w_j$

$$\Rightarrow \frac{\partial (||Xw - y||_2^2)}{\partial w_i} = \frac{\partial}{\partial w_i}\left(\sum_{j=1}^{n} (x_j w_i - y_j)^2\right)$$

$$= \frac{\partial}{\partial w_i} \sum_{j=1}^{n} (x_j(w_i - y)^2 =$$

$$= \sum_{j=1}^{n} \left(\frac{\partial}{\partial w_i}(x_j w_i - y_j)^2\right) = \sum_{j=1}^{n} 2(x_j w_i - y_j)\frac{\partial (x_j w_i)}{\partial w_i}$$

$$= \sum_{j=1}^{n} (2x_j(x_j w_i - y_j)$$

$$\Rightarrow \frac{\partial \left(\|Xw - y\|_2^2\right)}{\partial w} = \begin{bmatrix} \sum_{j=1}^{n} \left(2x_j^i \left(x_j^i w_1 - y_j^i\right)\right) \\ \vdots \\ \vdots \\ \sum_{j=1}^{n} \left(2x_j^i \left(x_j^i w_l - y_j^i\right)\right) \end{bmatrix}$$

Also $$\frac{\partial}{\partial w} \lambda \|w\|_2^2 = \lambda \begin{bmatrix} 2w_0 \\ \vdots \\ 2w_l \end{bmatrix} = 2\lambda \begin{bmatrix} w_0 \\ \vdots \\ w_l \end{bmatrix}$$

$$\Rightarrow \frac{\partial f(w)}{\partial w} = \begin{bmatrix} \sum_{j=1}^{n} \left[2(x_j^i)\left(x_j^i w_0 - y_j^i\right)\right] \\ \vdots \\ \vdots \\ \sum_{j=1}^{n} \left(2x_j^i \left(x_j^i w_l - y_j^i\right)\right) \end{bmatrix} + 2\lambda \begin{bmatrix} w_0 \\ \vdots \\ \vdots \\ w_l \end{bmatrix}$$