CS 583

Homework 1

John Graham

I pledge my honor that I have abided by the Stevens Honor System

---

Q1 With $x = \begin{bmatrix} 4 & 0 & -1 & 2 \end{bmatrix}^T$, calculate the following:

1. Squared ℓ_2-norm of x:

$$\begin{aligned} \|x\|_2^2 &= ((\sum_{i=1}^{4} x_i^2)^{1/2})^2 \\ &= \sum_{i=1}^{4} x_i^2 \\ &= 4^2 + 0^2 + -1^2 + 2^2 \\ &= 16 + 0 + 1 + 4 \end{aligned}$$

$$\|x\|_2^2 = 21$$

2. ℓ_1-norm of x:

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{4} |x_i| \\ &= 4 + 0 + |-1| + 2 \end{aligned}$$

$$\|x\|^1 = 7$$

3. inner product of x and a, $a = \begin{bmatrix} 0 & -2 & 6 & 3 \end{bmatrix}^T$

$$\begin{aligned} x \cdot a &= (4 * 0) + (0 * -2) + (-1 * 6) + (2 * 3) \\ &= 0 + 0 - 6 + 6 \end{aligned}$$

$$x \cdot a = 0$$

Q2 Calculate the following with $A \in \mathbb{R}^{2\times 3}$ and $b \in \mathbb{R}^3$:

$$\mathbf{A} = \begin{bmatrix} 6 & 1 & -2 \\ -5 & 0 & 8 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

1. $\mathbf{Ab} = \begin{bmatrix} 13 \\ -21 \end{bmatrix}$

2. $\mathbf{AA^T} = \begin{bmatrix} 6 & 1 & -2 \\ -5 & 0 & 8 \end{bmatrix} \begin{bmatrix} 6 & -5 \\ 1 & 0 \\ -2 & 8 \end{bmatrix} = \begin{bmatrix} 41 & -46 \\ -46 & 89 \end{bmatrix}$

Q3 With $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ and $y = \frac{x_1^2}{2} + \log_e x_2 - \frac{x_1}{x_3} = \frac{1}{2}x_1 + \log_e x_2 - x_1 x_3^{-1}$, calculate $\frac{dy}{d\mathbf{x}}$ at $\mathbf{x} = \begin{bmatrix} 7 & \frac{1}{4} & \frac{1}{3} \end{bmatrix}$:

$$\frac{dy}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial y}{x_1} & \frac{\partial y}{x_2} & \frac{\partial y}{x_3} \end{bmatrix}$$

$$\frac{\partial y}{x_1} = x_1 - x_3^{-1}$$
$$\frac{\partial y}{x_2} = x_2^{-1}$$
$$\frac{\partial y}{x_3} = x_1 x_3^{-2}$$

$$\begin{aligned} \frac{dy}{d\mathbf{x}} &= \begin{bmatrix} x_1 - x_3^{-1} & x_2^{-1} & x_1 x_3^{-2} \end{bmatrix} \\ &= \begin{bmatrix} 7 - (\frac{1}{3})^{-1} & (\frac{1}{4})^{-1} & 7(\frac{1}{3})^{-2} \end{bmatrix} \\ &= \begin{bmatrix} 7 - 3 & 4 & 7(9) \end{bmatrix} \end{aligned}$$

$$\frac{dy}{d\mathbf{x}} = \begin{bmatrix} 4 & 4 & 63 \end{bmatrix}$$

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

Derivative of some squared ℓ_2 norm of $\mathbf{x}$:

$$\frac{\partial}{\partial \mathbf{x}}(\|\mathbf{x}\|_2^2) = 2\mathbf{x}$$

Separate the derivatives of each term of $f(\mathbf{w})$:

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

Starting with $\|\mathbf{X}\mathbf{w}-\mathbf{y}\|_2^2$, using the chain rule

$$\frac{\partial}{\partial \mathbf{w}}(\|\mathbf{X}\mathbf{w}-\mathbf{y}\|_2^2) = \mathbf{X} * (2(\mathbf{X}\mathbf{w}-\mathbf{y})) = 2(\mathbf{X}^T\mathbf{X}\mathbf{w}-\mathbf{X}^T\mathbf{y})$$

Now $\lambda\|\mathbf{w}\|_2^2$:

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

Finally, we have:

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