MTH 602 Scientific Machine Learning

Homework 1

9/22/2025

S. M. Mahfuzul Hasan

02181922

1. Given,

$$\phi(x_1, x_2, x_3) = \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}$$

(a)

$$\begin{aligned}
\vec{F} = \nabla\phi &= \left(\frac{\partial}{\partial x_1}\hat{i} + \frac{\partial}{\partial x_2}\hat{j} + \frac{\partial}{\partial x_3}\hat{k}\right)\frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \\
&= \frac{-1}{2\sqrt{\left(x_1^2 + x_2^2 + x_3^2\right)^3}}\left(2x_1\hat{i} + 2x_2\hat{j} + 2x_3\hat{k}\right) \\
&= \frac{-1}{\sqrt{\left(x_1^2 + x_2^2 + x_3^2\right)^3}}\left(x_1\hat{i} + x_2\hat{j} + x_3\hat{k}\right) \qquad \text{(Ans.)}
\end{aligned}$$

(b) At (1, 0, 0),

$$\begin{aligned}
\vec{F}(1,0,0) &= \frac{-1}{\sqrt{\left(1^2 + 0^2 + 0^2\right)^3}}\left(1\hat{i} + 0\hat{j} + 0\hat{k}\right) \\
&= -\hat{i} \qquad \text{(Ans.)}
\end{aligned}$$

Direction: along (-)ve x_1 axis.
Magnitude: $\|\vec{F}(1,0,0)\|_2^2 = \sqrt{(-1)^2} = 1$

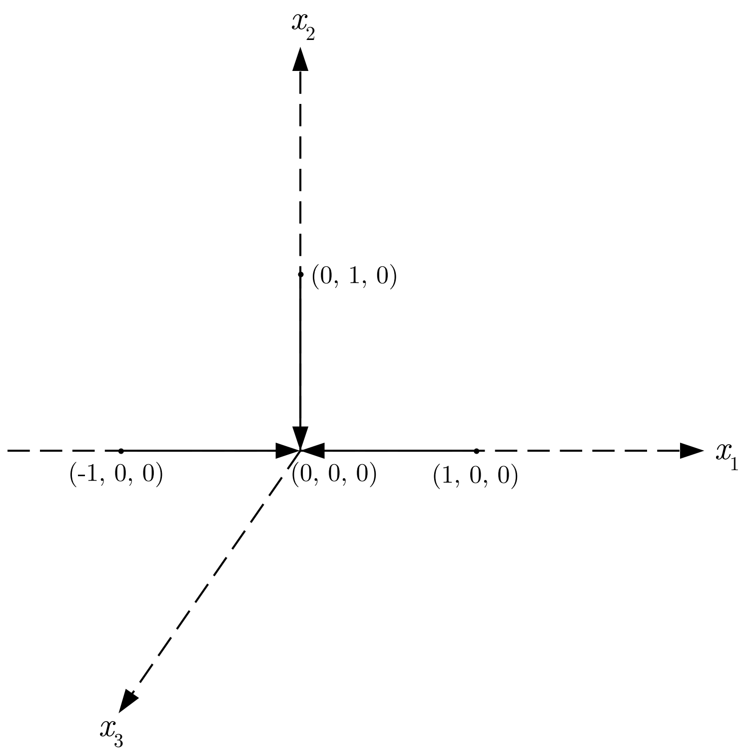

Figure 1: Sketch of $\vec{F}$ at (1, 0, 0), (-1, 0, 0) and (0, 1, 0).

At (-1, 0, 0),

$$\vec{F}(-1,0,0) = \frac{-1}{\sqrt{\left((-1)^2+0^2+0^2\right)^3}}\left(-1\hat{i}+0\hat{j}+0\hat{k}\right)$$
$$= \hat{i} \quad \text{(Ans.)}$$

Direction: along (+)ve x_1 axis.
Magnitude: $\|\vec{F}(-1,0,0)\|_2^2 = \sqrt{1^2} = 1$

At (0, 1, 0),

$$\vec{F}(0,1,0) = \frac{-1}{\sqrt{(0^2+1^2+0^2)^3}}\left(0\hat{i}+1\hat{j}+0\hat{k}\right)$$
$$= -\hat{j} \quad \text{(Ans.)}$$

Direction: along (-)ve x_2 axis.
Magnitude: $\|\vec{F}(0,1,0)\|_2^2 = \sqrt{(1)^2} = 1$

2. Given,

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T\vec{y}, \ \text{ where } \vec{x}, \vec{y} \in \mathbb{R}^N$$
$$Q^TQ = I, \ \text{ where } Q \in \mathbb{R}^{N\times N}$$

(a)

$$\text{L.H.S} = \langle Q\vec{x}, Q\vec{y} \rangle = (Q\vec{x})^T(Q\vec{y}) = \left(\vec{x}^TQ^T\right)(Q\vec{y}) = \vec{x}^T\left(Q^TQ\right)\vec{y}$$
$$= \vec{x}^TI\vec{y} = \vec{x}^T\vec{y} = \langle \vec{x}, \vec{y} \rangle = \text{R.H.S} \quad \text{(Showed)}$$

(b)

$$\|Q\vec{x}\|_2 = \sqrt{(Q\vec{x})^T(Q\vec{x})} = \sqrt{(\vec{x}^TQ^T)(Q\vec{x})} = \sqrt{\vec{x}^T(Q^TQ)\vec{x}}$$
$$= \sqrt{\vec{x}^TI\vec{x}} = \sqrt{\vec{x}^T\vec{x}} = \|\vec{x}\|_2$$

3. To verify,

$$\|\vec{x}\|_\infty \le \|\vec{x}\|_2$$
$$\|\vec{x}\|_2 \le \sqrt{N}\|\vec{x}\|_\infty$$

(a)

$$\|\vec{x}\|_\infty = \max_{1\le i\le N} |x_i| \tag{1}$$

Let's arrange the elements of $\vec{x}$ in such a way that the maximum absolute valued element sits at N^{th} position. Then equation (1) becomes-

$$\|\vec{x}\|_\infty = |x_N| \tag{2}$$

Now,

$$
\begin{aligned}
\|\vec{x}\|_2 &= \sqrt{\sum_{i=1}^{N} |x_i|^2} = \sqrt{\sum_{i=1}^{N-1} |x_i|^2 + |x_N|^2} \\
\Rightarrow \|\vec{x}\|_2 &\geq \sqrt{|x_N|^2}, \quad \text{since, } \sum_{i=1}^{N-1} |x_i|^2 \geq 0 \\
\Rightarrow \|\vec{x}\|_2 &\geq \sqrt{(\|\vec{x}\|_\infty)^2}, \quad \text{from equation (2)} \\
\Rightarrow \|\vec{x}\|_\infty &\leq \|\vec{x}\|_2 \quad \text{(Verified)}
\end{aligned}
$$

(b) We know,

$$
|x_i| \leq \max_{1 \leq i \leq N} |x_i|, \quad \text{for any } i \in [1, N] \tag{3}
$$

Let's assume again that the maximum absolute valued element is x_N. Then equation (3) can be written as,

$$
\begin{aligned}
&|x_i|^2 \leq |x_N|^2 \\
&\Rightarrow \sum_{i=1}^{N} |x_i|^2 \leq \sum_{i=1}^{N} |x_N|^2 = N|x_N|^2 \\
&\Rightarrow \sqrt{\sum_{i=1}^{N} |x_i|^2} \leq \sqrt{N|x_N|^2} \\
&\Rightarrow \|\vec{x}\|_2 \leq \sqrt{N}\|\vec{x}\|_\infty \quad \text{(Verified)}
\end{aligned}
$$

4. Let,

$$
A \in \mathbb{R}^{N \times (M+1)}, \vec{y} \in \mathbb{R}^N, \vec{\omega} \in \mathbb{R}^{M+1}
$$

(a)

$$
\begin{aligned}
\|\vec{y} - A\vec{\omega}\|_2^2 &= (\vec{y} - A\vec{\omega})^T (\vec{y} - A\vec{\omega}) \\
&= \left(\vec{y}^T - (A\vec{\omega})^T\right)(\vec{y} - A\vec{\omega}), \quad \text{since, } (X - Y)^T = X^T - Y^T \\
&= \vec{y}^T\vec{y} - \vec{y}^T(A\vec{\omega}) - (A\vec{\omega})^T\vec{y} + (A\vec{\omega})^T(A\vec{\omega}) \\
&= \vec{y}^T\vec{y} - \left(\vec{y}^T(A\vec{\omega})\right)^T - (A\vec{\omega})^T\vec{y} + (A\vec{\omega})^T(A\vec{\omega}), \\
&\qquad\qquad \text{since, } \vec{y}^T(A\vec{\omega}) \text{ is a scalar quantity} \\
&= \vec{y}^T\vec{y} - (A\vec{\omega})^T\left(\vec{y}^T\right)^T - (A\vec{\omega})^T\vec{y} + (A\vec{\omega})^T(A\vec{\omega}), \\
&\qquad\qquad \text{since, } (XY)^T = Y^T X^T \\
&= \vec{y}^T\vec{y} - (A\vec{\omega})^T\vec{y} - (A\vec{\omega})^T\vec{y} + (A\vec{\omega})^T(A\vec{\omega}), \\
&\qquad\qquad \text{since, } \left(X^T\right)^T = X \\
&= \vec{y}^T\vec{y} - 2(A\vec{\omega})^T\vec{y} + (A\vec{\omega})^T(A\vec{\omega}) \\
&= \vec{\omega}^T A^T A\vec{\omega} - 2\vec{\omega}^T A^T\vec{y} + \vec{y}^T\vec{y} \quad \text{(Showed)}
\end{aligned}
$$

Note: X and Y are representative matrices.

(b)

$$\nabla_{\vec{\omega}}\|\vec{y}-A\vec{\omega}\|_2^2 = \nabla_{\vec{\omega}}\left(\vec{\omega}^T A^T A\vec{\omega} - 2\vec{\omega}^T A^T\vec{y} + \vec{y}^T\vec{y}\right) \tag{4}$$

$$\begin{aligned}
\nabla_{\vec{\omega}}\left(\vec{\omega}^T A^T A\vec{\omega}\right) &= (A^T A + \left(A^T A\right)^T)\vec{\omega}, \\
&\qquad \text{using } \nabla_{\vec{\omega}}\left(\vec{\omega}^T B\vec{\omega}\right) = (B+B^T)\vec{\omega}, \text{ where } B = A^T A \text{ here} \\
&= \left(A^T A + A^T A\right)\vec{\omega} = 2A^T A\vec{\omega} \\
\nabla_{\vec{\omega}}\left(\vec{\omega}^T A^T\vec{y}\right) &= A^T\vec{y}, \\
&\qquad \text{using } \nabla_{\vec{\omega}}\left(\vec{\omega}^T\vec{y'}\right) = \vec{y'}, \text{ where } \vec{y'} = A^T\vec{y} \text{ here} \\
\nabla_{\vec{\omega}}\left(\vec{y}^T\vec{y}\right) &= 0, \qquad \text{since, } \vec{y} \text{ does not depend on } \vec{\omega}
\end{aligned}$$

So, equation (4) becomes-

$$\nabla_{\vec{\omega}}\|\vec{y}-A\vec{\omega}\|_2^2 = 2A^T A\vec{\omega} - 2A^T\vec{y} + 0 = 2A^T A\vec{\omega} - 2A^T\vec{y} \quad \text{(Showed)} \tag{5}$$

5. Projection matrix is defined as,

$$P = A\left(A^T A\right)^{-1} A^T$$

(a)

$$\begin{aligned}
L.H.S = P(P\vec{y}) &= A\left(A^T A\right)^{-1} A^T\left(A\left(A^T A\right)^{-1} A^T\vec{y}\right) \\
&= A\left(\left(A^T A\right)^{-1} A^T A\right)\left(A^T A\right)^{-1} A^T\vec{y} \\
&= AI\left(A^T A\right)^{-1} A^T\vec{y}, \\
&\qquad \text{since, } A^T A \text{ is a square matrix } \left(\in \mathbb{R}^{(M+1)\times(M+1)}\right) \text{ and invertible} \\
&= A\left(A^T A\right)^{-1} A^T\vec{y} = P\vec{y} = R.H.S \quad \text{(Showed)}
\end{aligned}$$

(b)

$$\begin{aligned}
L.H.S = P\vec{y}_M &= A\left(A^T A\right)^{-1} A^T\vec{y}_M \\
&= A\left(A^T A\right)^{-1} A^T(A\vec{\omega}) \\
&= A\left(\left(A^T A\right)^{-1} A^T A\right)\vec{\omega} \\
&= AI\vec{\omega} = A\vec{\omega} = \vec{y}_M = R.H.S \quad \text{(Showed)}
\end{aligned}$$

(c) The condition for least squares solution as derived in equation (5) is

$$\begin{aligned}
\nabla_{\vec{\omega}}\|\vec{r}\|_2^2 = \nabla_{\vec{\omega}}\|\vec{y}-\vec{y}_M\|_2^2 &= 2A^T A\vec{\omega}_* - 2A^T\vec{y} = 0 \\
&\Rightarrow A^T A\vec{\omega}_* - A^T\vec{y} = 0
\end{aligned} \tag{6}$$

Now,

$$\begin{aligned}
L.H.S = P\vec{r} &= A\left(A^T A\right)^{-1} A^T\left(\vec{y}-\vec{y}_M\right) \\
&= A\left(A^T A\right)^{-1}\left(A^T\vec{y} - A^T\vec{y}_M\right) \\
&= A\left(A^T A\right)^{-1}\left(A^T\vec{y} - A^T A\vec{\omega}_*\right), \qquad \text{since, } \vec{y}_M = A\vec{\omega}_* \\
&= A\left(A^T A\right)^{-1}(0), \qquad \text{from equation (6)} \\
&= 0 = R.H.S \quad \text{(Showed)}
\end{aligned}$$