8. Least Squares Problem 소개

1) Over-determined Linear Systems

- 정의: '방정식의 개수 > 변수의 개수'인 linear system.

e.g.
$$\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \\ 50 & 5.0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 66 \\ 74 \\ 78 \\ 72 \end{bmatrix}$$

- 대부분 해가 없다.
- 해가 없더라도 least squares로 가장 근사(approximate)한 해를 찾을 수 있다.

2) Inner Product

- 두 vector $\boldsymbol{u}, \boldsymbol{v} \in R^n$가 주어졌을 때, $\boldsymbol{u}^T\boldsymbol{v}$를 vector의 inner product(내적)라고 한다.

$$u \cdot v = u^T v$$

- Inner product의 결과는 scalar(1 by 1 matrix)이다.

e.g.
$$u = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad v = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \qquad u \cdot v = u^T v = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = [14]$$

3) Inner product의 특성

(1) $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$

(2) $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$

(3) $(c\boldsymbol{u}) \cdot \boldsymbol{v} = c(\boldsymbol{u} \cdot \boldsymbol{v}) = \boldsymbol{u} \cdot (c\boldsymbol{v})$

(4) $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$ 이고, $\boldsymbol{u} = 0$일 때만 $\boldsymbol{u} \cdot \boldsymbol{u} = 0$이다.

4) Vector Norm

- 정의: Vector norm(length) $||\boldsymbol{v}||$는 $\boldsymbol{v} \cdot \boldsymbol{v}$의 제곱근이다.

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \|v\|^2 = v \cdot v$$

- 기하학적으로, vector norm은 원점으로부터의 거리를 의미한다.
- Scalar c에 대해, $c\boldsymbol{v}$의 길이는 $\boldsymbol{v}$의 길이에 $|c|$배 한 것과 동일하다.

$$\|cv\| = |c|\|v\|$$

5) Unit Vector

- 정의: 길이가 1인 vector.
- 어떤 vector를 unit vector로 만들려면(normalize, 정규화), 해당 vector의 길이로 나누면 된다.

$$u = \frac{1}{\|v\|} v$$

- 정규화된 vector는 기존 vector와 방향은 같으나 길이가 1이다.

6) 두 vector 사이의 거리

- 정의: 두 vector $\boldsymbol{u}, \boldsymbol{v} \in R^n$ 사이의 거리(dist($\boldsymbol{u}, \boldsymbol{v}$))는 vector $\boldsymbol{u} - \boldsymbol{v}$의 길이이다.

$$dist(u, v) = \|u - v\|$$

e.g. $u = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ 사이의 거리 $\Rightarrow$ $u - v = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

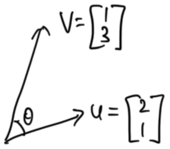
$$\|u - v\| = \sqrt{3^2 + (-1)^2} = \sqrt{10}$$

7) 두 vector 사이의 각도

- 두 vector 사이의 각도는 inner product를 사용해서 다음과 같이 구할 수 있다.

$$u \cdot v = \|u\|\|v\| \cos\theta$$

e.g.



$u \cdot v = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5$

$\|u\| = \sqrt{5}$, $\|v\| = \sqrt{10}$

$$u \cdot v = \|u\|\|v\| \cos\theta$$
$$5 = \sqrt{5} \cdot \sqrt{10} \cos\theta$$
$$\cos\theta = \frac{1}{\sqrt{2}}$$
$$\theta = 45°$$

8) Orthogonal Vectors

- 정의: Vector $\boldsymbol{u}$, $\boldsymbol{v} \in R^n$의 inner product $\boldsymbol{u} \cdot \boldsymbol{v} = 0$가 0인 경우, 두 vector는 서로에게 orthogonal(직교)하다.

pf. $u \cdot v = \|u\|\|v\| \cos\theta$ 이므로

$0 = \|u\|\|v\| \cos\theta$ 라면

$\cos\theta = 0$ 이다. $\rightarrow$ $\cos\theta = 0$ 이려면 $\theta = 90°$ 이다.

9) Least Square Solution

- 정의: Over-determined system $A\boldsymbol{x} \simeq \boldsymbol{b}$ ($A \in R^{m\times n}, \boldsymbol{b} \in R^{n},\ m \gg n$)가 주어졌을 때, least square solution $\hat{\boldsymbol{x}}$는 다음과 같이 정의된다.

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|b - Ax\|$$

e.g. $\underset{A}{\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \\ 50 & 5.0 & 1 \end{bmatrix}} \underset{x}{x} = \underset{b}{\begin{bmatrix} 66 \\ 74 \\ 78 \\ 72 \end{bmatrix}}$ 의 x에 $\underset{x_1}{\begin{bmatrix} -0.12 \\ 16 \\ -9.5 \end{bmatrix}}$, $\underset{x_2}{\begin{bmatrix} -0.4 \\ 20 \\ -20 \end{bmatrix}}$을 선정한다고 해보자.

$$Ax_1 = \begin{bmatrix} 71.3 \\ 72.2 \\ 79.9 \\ 64.5 \end{bmatrix},\quad Ax_2 = \begin{bmatrix} 66 \\ 74 \\ 78 \\ 60 \end{bmatrix}$$ 이 된다.

언뜻 보기에 기존 b와 차이가 크지 않은 x_2가 더 적절해 보이지만,
error $\|b-Ax\|$를 고려한다면 x_1은 9.55, x_2는 12로 x_1의 error가 더 작다.

$\therefore$ x_1이 x_2보다 더 적절한 근사 해라고 할 수 있다.