행렬과 선형대수학(중고등학생과 일반인을 위한 글)

세 숫자(x_1, x_2, x_3)를 입력하면 두 숫자(y_1, y_2)가 나오는 블랙박스가 있다고 하자. (벡터가 뭔지 안다면, 이 두 숫자를 하나의 이차원 벡터로 생각해도 좋다.) 예를 들어 (x_1, x_2, x_3)으로 (1, 2, 0)를 입력했더니 (y_1, y_2)로 (1, 4)가 나왔다고 하자.

이제, 이 블랙박스를 여러번 조작함으로써, 이 기계가 아주 특별한 다음의 성질들을 만족한다는 것을 발견해냈다고 하자

첫째로 각 숫자의 n 배가 되는 숫자를 입력하면 n 배의 값이 나왔다고 하자.

예를 들어 (x_1, x_2, x_3)에 (1,2,0)의 두 배인 (2,4,0)을 입력했더니 (1,4)의 두 배인 (2,8)이 나왔다는 것이다. 다른 예를 들자면 (x_1, x_2, x_3)에 (1,2,0)의 세 배인 (3,6,0)을 입력했더니 (1,4)의 세 배인 (3,12)가 나왔다는 것이다.

그럼 (x_1, x_2, x_3)에 (0, 0, 0)을 입력했을 때에는 어떤 값이 나왔을까?

답은 (0, 0)이다. 왜냐하면 (0, 0, 0)은 (1, 2, 0)의 0배이고 (1, 4)의 0배는 (0, 0)이기 때문이다.

둘째로 (x_1, x_2, x_3)에 $(a_1 + b_1, a_2 + b_2, a_3 + b_3)$를 입력시켰을 때 나오는 값은 (x_1, x_2, x_3)에 (a_1, a_2, a_3)를 입력시켰을 때의 값에 (x_1, x_2, x_3)에 (b_1, b_2, b_3)를 입력시켰을 때의 값을 더한 값이 나온다고 하자. 예를 들어 앞의 예처럼 (1, 2, 0)을 넣었을 때 (1, 4)가 나왔고 한 예를 더 들어 (3, 4, -1)을 입력했을 때 (4, 6)이 나왔다면 (4=1+3. 6=2+4, -1=0-1)을 입력했ㅇ르 때 (5, 10)이 나왔다는 것이다.

이러한 성질들을 만족하는 블랙박스, 혹은 함수를 선형 함수(linear function) 또는 선형 연산자(linear operator)라고 한다. (선형 연산자가 벡터에서 벡터로 가는 연산자라는 것을 상기해라 우리 예의 경우는 3차원 벡터에서 2차원 벡터로 가는 연산자다.)

수식을 표현하면 다음과 같이 쓸 수 있다

$$f(x_1, x_2, x_3) \ = (y_1, y_2).$$

즉:

$$f\,(1,\ 2,\ 0) = (1,\ 4)$$

$$f\,(3,\ 4,\ -1) = (4,\ 6)$$

위 두 가지 성질을 수학적으로 다시 표현해 쓰면, 첫번째 성질은 다음과 같이 써 질 수 있고

$$f\,(nx_1, nx_2, nx_3) = (ny_1, ny_2)$$

두번째 성질은 다음과 같이 써 질 수 있다.

$$f\,(a_1 + b_1, a_2 + b_2, a_3 + b_3) = f\,(a_1, a_2, a_3) + f(b_1, b_2, b_3)$$

이제, 다음과 같은 질문을 해 보자. 만약, 우리가 이 블랙 박스가 선형 함수 혹은 선형 연산자라는 사실을 먼저 알고 있었더라면, 이 블랙박스에 어떤 값을 넣었을 때 어떤 값이 나오는지 이 블랙박스를 몇 번 조작해 보면 완전히 알 수 있을까?

답은 세 번이다. 이 답을 이해하기 위해서는 선형 함수의 이 두 가지 성질이 아주 중요하다. 이 성질을 이용해 우리는 다음과 같이 쓸 수 있다.

$$f(x_1,x_2,x_3)=f(x_1+0,x_2+0,0+x_3)=f(x_1,x_2,0)+f(0,0,x_3)$$
$$=f(x_1+0,0+x_2,0+0)+f(0,0,x_3)=f(x_1,0,0)+f(0,x_2,0)+f(0,0,x_3)$$

위 식에서 선형함수의 두 번째 성질을 이용했다. 이제 첫번째 성질을 이용하면

$$\begin{aligned} f(x_1,x_2,x_3) &= f(x_1,0,0)+f(0,x_2,0)+f(0,0,x_3) \\ &= x_1 f(1,0,0)+x_2 f(0,1,0)+x_3 f(0,0,1) \end{aligned}$$

즉 $f(x_1,x_2,x_3)$는 $f(1,0,0)$, $f(0,1,0)$, $f(0,0,1)$를 알면 완전히 결정되는 것이다.

이 때 $f(1,0,0)$의 값을 (f_{11},f_{21})이라고 하고 $f(0,1,0)$의 값을 (f_{12},f_{22})라고 하고 $f(0,0,1)$의 값을 (f_{13},f_{23})이라고 하자.

그럼 다음을 얻는다

$$f\,(x_1,x_2,x_3)=\ x_1\ \ (f_{11},\ f_{21})+\ x_2\ \ (f_{12},\ f_{22})+x_3\ \ (f_{13},\ f_{23})$$
$$=(x_1f_{11}+x_2f_{12}+x_3f_{13},\ x_1f_{21}+x_2f_{22}+x_3f_{23})$$

이 관계식은 "행렬"이라는 것을 이용하면 다음과 같이 간단히 표기할 수 있다.

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

여기 y_1,y_2는 물론 각각 "$x_1f_{11}+x_2f_{12}+x_3f_{13}$," 와 "$x_1f_{21}+x_2f_{22}+x_3f_{23}$."가 된다.

즉, 위의 관계식이 행렬 곱셈의 정의이다. 내가 처음 행렬을 배웠을 때에는 행렬이 선형 함수나 선형 연산자를 간단히 표현하는 방식이라는 것을 몰라서 왜 행렬이 이런 이상한 방식으로 곱셈이 정의되나 알지 못했었다.

즉 "$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix}$"를 "$F$"라 하고 "$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$"를 "$X$"라 하고 "$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$"를 "$Y$"라 하고 위 식을 다시 쓰면:

$$FX=Y \tag{1}$$

가 된다. 혹은, 다음과 같이 쓸 수 있다.

$$\sum_{b=1}^{3} f_{ab}x_b=y_a$$

위에서 F와 같이 2개의 행과 3개의 열을 가진 행렬을 "2×3" 행렬이라고 하고, X와 같이 3개의 행과 1개의 열을 가진 행렬을 "3×1" 행렬이라고 한다. 마찬가지로, Y는 물론 "2×1" 행렬이다.

행렬 곱셈은 더 일반적인 경우에도 가능하다. 즉, $FX = Y$라고 할 때 꼭 X나 Y의 열의 개수가 하나일 필요가 없는 것이다. 더 일반적인 행렬의 곱셈 공식을 얻기 위해 X에 선형 함수를 두 번 가해 준 것을 고려해 보자.

예를 들어 g가 두 숫자를 집어 넣었을 때 네 숫자가 나오는 선형 연산자라면

(즉

$$g(y_1, y_2) = (z_1, z_2, z_3, z_4)$$

) 다음과 같은 두 연산자 g와 f의 조합을 생각해볼 수 있다.

$$g\left(f\left(x_1, x_2, x_3\right)\right) = g(y_1, y_2) = (z_1, z_2, z_3, z_4)$$

이것을 행렬로 표기하면:

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$$

다시 말해,

$$GFX = GY = Z \tag{2}$$

만약 H를 GF라고 한다면,즉, 세 개의 숫자를 넣었을 때 네 개의 숫자가 나오는 선형 연산자라고 생각하면, 우리는 h를 다음과 같은 방식을 이용해 얻을 수 있다.

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11}x_1 + f_{12}x_2 + f_{13}x_3 \\ f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3 \end{bmatrix}$$

$$= \begin{bmatrix} g_{11}\left(f_{11}x_1 + f_{12}x_2 + f_{13}x_3\right) + g_{12}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{21}\left(f_{11}x_1 + f_{12}x_2 + f_{13}x_3\right) + g_{22}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{31}\left(f_{11}x_1 + f_{12}x_2 + f_{13}x_3\right) + g_{32}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{41}\left(f_{11}x_1 + f_{12}x_2 + f_{13}x_3\right) + g_{42}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \end{bmatrix}$$

$$= \begin{bmatrix} g_{11}f_{11} + g_{12}f_{21} & g_{11}f_{12} + g_{12}f_{22} & g_{11}f_{13} + g_{12}f_{23} \\ g_{21}f_{11} + g_{22}f_{21} & g_{21}f_{12} + g_{22}f_{22} & g_{21}f_{13} + g_{22}f_{23} \\ g_{31}f_{11} + g_{32}f_{21} & g_{31}f_{12} + g_{32}f_{22} & g_{31}f_{13} + g_{32}f_{23} \\ g_{41}f_{11} + g_{42}f_{21} & g_{41}f_{12} + g_{42}f_{22} & g_{41}f_{13} + g_{42}f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \\ h_{41} & h_{42} & h_{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

다시 말해,

$$GFX = G(FX) = (GFX) = (GF)X = HX \tag{3}$$

즉, 다른 표기법을 쓰면,

$$\sum_{a=1}^{2}\sum_{b=1}^{3} g_{ca} f_{ab} x_b = z_c = \sum_{b=1}^{3} h_{cb} x_b$$

따라서

$$\sum_{a=1}^{2}\sum_{b=1}^{3} g_{ca} f_{ab} = z_c = \sum_{b=1}^{3} h_{cb}$$

즉 G기 4×2 행렬이고 F가 2×3 행렬이라면, $H = GF$는 4×3 행렬이 되고 그 값은 위의 식으로 주어진다.

일반적으로 "A"가 "$m\times p$" 행렬이고, "B"가 "$p\times n$" 행렬이면 "AB"는 "$m\times n$" 행렬이다.

행렬 A와 행렬 B를 곱하기 위해서는 이 조건이 만족되어야 한다. 예를 들어 3×4 행렬 "A"와 5×3 행렬 "B"가 있다면 AB는 정의되지 않는다. 왜냐하면 4는 5와 다른 숫자이기 때문이다.

또한, 재미있는 사실은 행렬에서는 보통의 숫자와는 달리 행렬 A와 B에 대해 AB가 BA와 일반적으로는 같지 않다는 것이다. 예를 들어:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \tag{4}$$

라면

$$AB = \begin{bmatrix} 1\times 5 + 2\times 7 & 1\times 6 + 2\times 8 \\ 3\times 5 + 4\times 7 & 3\times 6 + 4\times 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \tag{5}$$

반면에,

$$BA = \begin{bmatrix} 5\times 1 + 6\times 3 & 5\times 2 + 6\times 4 \\ 7\times 1 + 8\times 3 & 7\times 2 + 8\times 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \tag{6}$$

이 사실은 양자역학에서 아주 중요하다. 물리학자들에 따르면 위치, 운동량, 에너지 등과 같이 모든 관측 가능한 양은 그에 해당하는 선형 연산자, 혹은 행렬이 있다고 한다. 이 때, 위치에 해당하는 행렬을 X, 운동량(물체의 질량에 속도를 곱한 것을 운동량이라고 한다)에 해당하는 행렬을 P라고 한다면, XP와 PX가 같지 않다고 한다. 이 사실로부터 하이젠베르크의 불확정성원리를 유도할 수 있다. 즉, 한 물체의 위치를 정확히 측정하려면 할 수록, 그 물체의 운동량을 정확히 측정하는 것이 불가능하고, 물체의 운동량을 정확히 측정하려면 할 수록, 그 물체의 위치를 정확하게 측정하는 것이 불가능하다.

끝으로, 이러한 행렬과 선형함수나 선형연산자를 다루는 학문을 선형대수학이라고 한다. 선형대수학은 양자역학의 기초가 된다. 사실 선형대수는 실질적으로 이공계의 모든 학문에서 중요한 위치를 차지한다. 이러한 이유로, 이공계 대학생은 보통 1학년 2학기 때 선형대수학을 깊이 있게 공부한다.

연습문제 1.
선형연산자 T가 벡터에 다음과 같이 작용한다고 하자:

$$T \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2v_2 \\ v_1 \\ 3v_3 \\ v_1 + v_2 \end{bmatrix}$$

T를 행렬로 나타내라

연습문제 2.
만약 A와 B가 같은 크기의 선형연산자라고 할 때 (즉, 같은 개수의 행과 열) 그 들의 합 $C = A + B$를 다음 식에 의해 정의될 수 있다.

$$Cx = Ax + Bx \tag{7}$$

여기서 x는 선형연산자에 입력하는 숫자들이다. (즉, 벡터) 다시 말해, 이 식은 x가 입력되었을 때 C가 무엇을 출력해는지 결정해준다. 이런 경우, 각각의 성분별로 쓰면 다음이 성립된다는 것을 스스로 확인해라.

$$C_{ij} = A_{ij} + B_{ij} \tag{8}$$

만일

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \tag{9}$$

라면 C는 무엇인가?

연습문제 3.
비슷한 방식으로, 우리는 행렬에 스칼라(즉, 숫자) 곱을 할 수 있다. 예를 들어, 만약 D가 행렬이고 c가 숫자이며, 그들의 곱이 $G = cD$라면,

$$Gx = c(Dx) \tag{10}$$

을 얻는다. 각각의 성분별로 쓰면, 다음이 얻어진다는 것을 스스로 확인해라.

$$G_{ij} = cD_{ij} \tag{11}$$

(8)와 (11)를 이용하여 다음을 계산하라:

$$3 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} =? \tag{12}$$