Chapter 2. Linear Algebra.

도입

개념 및 정의.

대수학 (algebra)

집합 & 원소

: 수학에서 어떤 대상을 다룰때, `그 대상이 무엇인지'를 정의하고, 그 대상들에 대한 연산, 체계, 특성 등을 연구하는 것.

선형 대수학 (Linear Algebra)

: `벡터'를 대상으로 함.

즉, 선형 대수학은 `벡터의 정의', `벡터를 다루는 (연산하는) 과정 등에서 발견하게되는 특성 등을 살펴보는 분야이다.

벡터.

- 합해도 원래 대상의 범주 안에 있고 (덧셈에 대해 닫혀있고)
- 실수 값인 스칼라를 곱해도 다시 원래 대상의 범주 안에 있는 (스칼라 곱이 정의되는) 대상들을 의미함.

선형 대수학의 연립 일차방정식을 공부하는 이유.

데이터 과학이나 머신러닝을 공부하다보면 테이블로 정리된 데이터를 보게된다. 이 테이블은 수많은 성분들 (colums)로 이루어져 있으며, 이러한 성분들을 가장 적절하고 간단하게 수치화하여 가공하는 전처리 과정을 거친 후, 얻어진 정돈된 각 성분에 가중치를 곱해서 더한 형태로 모델을 만들곤 합니다.

↳ 일차식의 형태를 띈다.

∴ 연립 일차 방정식을 다루는 것은 머신러닝의 첫걸음이라 볼수 있다.

연립 일차 방정식의 유형과 해를 구하는 법에서 시작해보자.

2.1 Systems of Linear Equations

(연립 일차 방정식)

연립 일차 방정식 이란?

일차 연립 방정식 예.

Example 2.1

A company produces products $N_1, \dots, N_n$ for which resources $R_1, \dots, R_m$ are required. To produce a unit of product N_j, a_{ij} units of resource R_i are needed, where $i = 1, \dots, m$ and $j = 1, \dots, n$.

The objective is to find an optimal production plan, i.e., a plan of how many units x_j of product N_j should be produced if a total of b_i units of resource R_i are available and (ideally) no resources are left over.

If we produce $x_1, \dots, x_n$ units of the corresponding products, we need

a total of

$$a_{i1}x_1 + \cdots + a_{in}x_n \tag{2.2}$$

many units of resource R_i. An optimal production plan $(x_1, \dots, x_n) \in \mathbb{R}^n$, therefore, has to satisfy the following system of equations:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}, \tag{2.3}$$

where $a_{ij} \in \mathbb{R}$ and $b_i \in \mathbb{R}$.

우리가 각 제품. $N_1, N_2, N_3, \cdots N_n$을

$x_1, x_2, \cdots, x_n$개 생산한다고 할때,

R_1 이라는 자원이 얼마나 쓰이는가?

N_1을 x_1개 생산 $= a_{11} \times x_1$

N_2을 x_2개 생산 $= a_{12} \times x_2$

$\vdots$

$+$ N_n을 x_n개 생산 $= a_{1n} \times x_n$.

R_1의 총 사용량. $= a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n$

즉, R_i의 총 사용량 $= a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$

이때, 자원 R_i는 b_i 만큼만 있으므로, 각 자원의 사용량은 다음과 같다.

위와 같은 식을 연립방정식이라고 하고, $x_1, x_2, \cdots x_n$을 미지수라고 한다.

이러한 식을 만족하는 x_i의 값들인 튜플 $(x_1, x_2, \cdots, x_n)$을 일차 연립 방정식의 해 (solution)라고 한다.

연립 방정식의 해의 종류

① 해가 없는 경우

머신러닝을 배워보면 변수의 개수보다 식의 개수가 훨씬 많은 경우가 많아요.

모든 식을 만족하는 해가 존재하지 않을 때, 어떤 값을 가장 적절한 근사해로 찾아내는가의 문제가 훨씬 중요해요.

② 해가 유일하게 하나 존재하는 경우

③ 해가 무수히 많은 경우. →

$$\begin{array}{rcrcrcrl} x_1 & + & x_2 & + & x_3 & = & 3 & (1) \\ x_1 & - & x_2 & + & 2x_3 & = & 2 & (2) \\ 2x_1 & & & + & 3x_3 & = & 5 & (3) \end{array}$$

→ $\left(\frac{5}{2} - \frac{3}{2}a, \frac{1}{2} + \frac{1}{2}a, a\right), \quad a \in \mathbb{R}$

연립 일차 방정식의 행렬 도입.

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned},$$

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Ch 2.2 Matrices : 행렬의 기초

행렬은 선형대수학에서 다루는 벡터의 가장 대표적인 예로,
선형대수학 전체에 걸쳐 가장 중요한 역할을 합니다.
행렬은 **연립일차방정식**을 표현하는데 사용하기도하고,
또한 **linear Mapping**을 나타내기도 합니다 (2.7)

행렬이란?

- $(m \times n)$ 행렬이란 m개의 행과 n개의 열로 이루어진 직사각형 형태의 수의 나열.
- 연립방정식에서 행은 수많은 속성들의 비율을 곱한 조합을, 열은 하나의 성분이나 특성이 어떻게 분배되는지를 알려줘요.

행렬의 연산

덧셈)

곱셈) $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times k}$ $\rightarrow$ $C = AB \in \mathbb{R}^{m \times k}$

✓ 행렬의 곱셈은 교환법칙이 성립하지 않는다.

단위행렬) $n \times n$ 행렬에 대해. 대각선에 해당하는 값들은 모두 1, 나머지는 0.

$IA = AI = A$.

역행렬) $A^{-1} = B$ 일때, $AB = BA = I$

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\quad A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

판별식.

전치행렬) $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ 일때 $A = B^T$, $B = A^T$

대칭행렬) 대각선 성분을 기준으로 대칭

$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \rightarrow A = A^T$

대칭행렬의 합은 항상 대칭이지만, 곱은 대칭이 되지 않는다.

$AA^{-1} = I = A^{-1}A$

$(AB)^{-1} = B^{-1}A^{-1}$

$(A+B)^{-1} \neq A^{-1} + B^{-1}$

$(A^T)^T = A$

$(A+B)^T = A^T + B^T$

$(AB)^T = B^T A^T$

ch 2.3 Solving System of Linear Equations

연립방정식 $Ax = B$ 의 해를 찾는 방법. 〈특수해와 일반해의 개념〉

– 특히 미지수 많음 식이 충분하지 않은 상황 즉, 해가 무수히 많을때 어떻게 해를 표현하는가?

해들 중 하나 / 해들 간의 차이

– by 특수해와 일반해.

예) $\begin{bmatrix} 1 & \boxed{0} & 8 & -4 \\ 0 & \boxed{1} & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ \boxed{x_2} \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \end{bmatrix}$ …㉠ (c_i, x_i)

문자는 4개 인데 식은 2개 밖에 없음.

∴ 무수히 많은 해가 존재한다.

$\therefore \sum_{i=0}^{4} c_i x_i = b$

i) 특수해 찾기.

$$\begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1 + & \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_2 + & \begin{bmatrix} 8 \\ 2 \end{bmatrix} x_3 + & \begin{bmatrix} -4 \\ 12 \end{bmatrix} x_4 \end{array} \quad \cdots ㉡$$

가장 간단한 열 c_1, c_2 를 이용해서 ㉠의 우변 성분을 표현하면

$$b = \begin{bmatrix} 42 \\ 8 \end{bmatrix} = 42 \underset{c_1}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} + 8 \underset{c_2}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}$$

$$\therefore \text{특수해} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \\ 0 \\ 0 \end{bmatrix}.$$

→ 이는 해당 연립 일차 방정식의 유일한 해가 아님.

☆?

ii) 일반해 찾기

c_1과 c_2를 이용하여 c_3와 c_4를 표현해보자.

$$c_3 \Rightarrow \begin{bmatrix} 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} 8 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} 2$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}(-8) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}(-2) + \begin{bmatrix} 8 \\ 2 \end{bmatrix} \cdot 1 + \begin{bmatrix} -4 \\ 12 \end{bmatrix} \cdot 0 = 0$$

$\therefore 0 = -8c_1 + -2c_2 + c_3 + 0 \cdot c_4$ 이고.

$(x_1, x_2, x_3, x_4) = (-8, -2, 1, 0)$

$$\underbrace{\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix}}_{A} \left(\lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right) = \lambda_1 (8\boldsymbol{c}_1 + 2\boldsymbol{c}_2 - \boldsymbol{c}_3) = \mathbf{0}. \quad (2.41)$$

일반해.

$Ax = 0$ — 일반해.

c_4도 마찬가지

$$\underbrace{\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix}}_{A} \left(\lambda_2 \begin{bmatrix} -4 \\ 12 \\ 0 \\ -1 \end{bmatrix} \right) = \lambda_2 (-4\boldsymbol{c}_1 + 12\boldsymbol{c}_2 - \boldsymbol{c}_4) = \mathbf{0}$$

$Ax = 0$

따라서 $Ax = B$의 해는.

$$\left\{ \boldsymbol{x} \in \mathbb{R}^4 : \boldsymbol{x} = \underset{\text{특수해}}{\begin{bmatrix} 42 \\ 8 \\ 0 \\ 0 \end{bmatrix}} + \lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} + \lambda_2 \underset{\text{일반해}}{\begin{bmatrix} -4 \\ 12 \\ 0 \\ -1 \end{bmatrix}}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}. \quad (2.43)$$

$A(x_1 + x_2 + x_3) = B$.

$Ax_1 + \underbrace{Ax_2 + Ax_3}_{0} = B$

$Ax = b$ 의 방정식에서 2개의 해 x_1, x_2를 찾았을 때.

$Ax_1 = b$, $Ax_2 = b$.

$Ax_1 - Ax_2 = b - b$

$A\underbrace{(x_1 - x_2)}_{\alpha} = 0$ → $Ax = b$의 해들 간의 차이 ⇒ $Ax = 0$ 의 해. α.

$\therefore x_1 - x_2$는 $Ax = 0$ 의 하나의 해가 됨.

$x_1 - x_2 = \alpha$. $x_1 = x_2 + \alpha$

$\therefore Ax = b$ 의 특수해를 x_0, $Ax = 0$ 의 특수해를 α 라 할때.

$Ax = b$의 일반해는 $\underbrace{x_0}_{\text{특수해}} + \underbrace{\lambda\alpha}_{\text{해들 간의 차이}}$의 형태로 나타남.

즉 $Ax = b$의 해들 사이에는 $Ax = 0$ 만큼의 차이가 존재

이 예제를 쉽게 풀수 있었던 이유는 '계수가 간단했기 때문'임.

그러나 대부분의 방정식은 계수가 간단하지 않은 경우가 많다.

따라서 복잡한 계수들을 간단하게 만들기 위해 '가우스 소거법'을 이용.

Ch 2.3 Solving System of Linear Equations

목표: 어떻게 특수해와 일반해를 잘 구할 수 있는지.

by row-echolon form, reduced row echelon form.

Ch 2.3.2) Elementary Transformations.

- 복잡한 행렬을 간단하게 만드는 변환

key point

✓ 식 끼리의 가감법, 양변을 같은 수로 곱하기 등의 방법으로 계수는 간단하게 바꾼다.

✓ 해는 변화시키지 않는다.

예제)

$$\begin{array}{rcrcrcrcrcr} -2x_1 & + & 4x_2 & - & 2x_3 & - & x_4 & + & 4x_5 & = & -3 \\ 4x_1 & - & 8x_2 & + & 3x_3 & - & 3x_4 & + & x_5 & = & 2 \\ x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ x_1 & - & 2x_2 & & & - & 3x_4 & + & 4x_5 & = & a \end{array} .$$

좌변의 계수들은 행렬로 나타내되, 행렬 맨 뒤에 한줄을 긋고 우변의 값들도 함께 표현.

argumented Matrix (첨가행렬)

$$\left[\begin{array}{rrrrr|r} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] \begin{array}{l} \text{Swap with } R_3 \\ \\ \text{Swap with } R_1 \\ \\ \end{array}$$

가능하면 맨 앞에 1이 있는 열이 맨 위에 있는게 좋다.

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ -2 & 4 & -2 & -1 & 4 & -3 \\ 1 & -2 & 0 & -3 & 4 & a \end{array}\right] \begin{array}{l} \\ -4R_1 \\ +2R_1 \\ -R_1 \end{array}$$

간단한 형태로 만들기위해 여기를 0으로 만들려는 것. ∴ 1행의 값들을 적절히 실수배 한다음 빼준다.

0이 아닌 처음 나타나는 숫자를 1로 나타내주기 위한 단계

마지막 행의 맨 앞 0이 아닌 숫자가 0이 되도록 하기 위해 $-R_2$

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{array}\right] \begin{array}{l} \\ \\ \\ -R_2 - R_3 \end{array}$$

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] \begin{array}{l} \\ \cdot(-1) \\ \cdot(-\frac{1}{3}) \\ \\ \end{array}$$

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right]$$

그럼 관련의 식은 다음과 같이 바뀐다.

$$\begin{array}{rcrcrcrcrcr} x_1 & - & 2x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ & & & & x_3 & - & x_4 & + & 3x_5 & = & -2 \\ & & & & & & x_4 & - & 2x_5 & = & 1 \\ & & & & & & & & 0 & = & a+1 \end{array}$$

식1) $a=-1$

식1)을 통해 $a=-1$ 임을 알 수 있다.

그리고 이 식을 만족시키는 특수해 하나를 찾아보면

$$\begin{bmatrix} 1 & -2 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} x_4 + \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} x_5 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$$

2 0 −1 1 0

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} .$$

을 구할 수 있고.

앞장에서와 마찬가지로 일반해로 찾으면.

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

를 찾아낼 수 있다.

Row Echelon Form 과 연립방정식의 해.

위와 같은 방법 (행 위치 바꾸기, 실수배하기)를 통해 단순하게 찾아낸 행렬의 형태를 Row-Echelon Form 이라고 한다.

다음의 조건을 만족해야한다.

✓ 모든 성분이 0인 행은 맨 아래에 위치합니다.

그 외의 행들은 0이 아닌 값들을 적어도 하나 포함합니다.

✓ 0이 아닌 값들을 포함하는 열은 첫 행을 제외하고는,

모든 행들이 앞부분에 0부터 시작 합니다.

각 행에서 0이 아닌 값들을 pivot 이라고 한다.

예1)

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right]$$

이 row-echelon form 에서 pivot들.

이 pivot 들과 곱해지는 변수들은 x_1, x_3, x_4 가 되는데 이러한 변수들은 basic variables, 그 외의 나머지 변수들은 free variables 라고 한다.

이러한 row-echelon form을 활용하면 특수해를 좀 더 쉽게 찾을 수 있다. 위 행렬에서 pivot를 포함하고 있는 column을 잘 조합해서 우측 값들을 만족하면 된다.

$$\underset{2}{\lambda_1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \underset{-1}{\lambda_2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \underset{1}{\lambda_3} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

basic variables : $\overset{x_1}{\underset{2}{\lambda_1}}, \overset{x_3}{\underset{-1}{\lambda_2}}, \overset{x_4}{\underset{1}{\lambda_3}}$

\+

pivot이 포함되지 않는 free variables 로 두면. [2, 0, −1, 1, 0] 이라는 특수해를 가질수 있고.

+) reduced row echelon form.

- 식을 좀더 정리할 수 있다.

- 조건 ✓ row - echelon form 이면서.
  ✓ 모든 pivot이 1 이고
  ✓ pivot은 그 열에서 유일하게 0이 아닌 값일 때

이 3가지 조건을 모두 만족하면 reduced row echelon form 이라 부른다.

예를 들면)

$$\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right]$$

↳ reduced row echelon form은 아니다. 이를 좀더 정리하면.

? 어떻게?

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$$

이렇게 정리가되면 reduced row echelon form이 된다.

↳ 이 reduced row echelon을 이용해서 $Ax=0$을 만족하는 일반해를 찾는다.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$$

pivot이 아닌 2번째 열은 첫번째 열을 -3배해서 더해주면 0으로 만들 수 있다.

따라서 $Ax=0$의 특수해 중 하나는 $[-3,1,0,0,0] = -[3,-1,0,0,0]$이 된다.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} x_4 + \begin{bmatrix} 3 \\ 9 \\ -4 \end{bmatrix} x_5 = 0$$

$-3 \quad 1 \quad 0 \quad 0 \quad 0$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}$$

맨 마지막 열은 pivot 열인 1,3,4 번째 열을 -3배, -9배, 4배 해서 더해주면 0을 만들 수 있다.

$[-3,0,9,4,1]$ 로 하나의 $Ax=0$ 의 특수해가 된다.

∴ $Ax=0$의 모든 해는

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

가 된다.

* The minus 1 tric.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}.$$

pivot이 아닌 2,5열에 pivot 역할을 하는 열 2개를 끼워준다. 단, pivot 자리에 -1을 준다.

↓

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

→

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

이게 원래 $Ax=0$의 해와 같다.

Ch 2.3, 2.4 Gaussian Elimination and Vector Space.

#3일 – Vector space.

nxn 행렬에서 (n≥3) 어떻게 역행렬을 찾을 것인가?

⇒ **Gaussian Elimination.**

방법은 Reduced Row Echelon Form을 구하는 과정과 동일.

$$A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ -R_1 \\ -R_1 \\ -R_1 \end{array}$$

첫번째 행을 제외한 맨 앞 열은 0으로 만들기

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & -2 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ -2R_2 \\ -R_2 \end{array}$$

3번째, 4번째 행의 2열을 모두 0으로 만들어준다.

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ -R_4 \\ \\ \end{array}$$

3번째 행의 3번째 열을 1로 만들기 위해

4번째 행을 빼준다.

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} -2R_3 \\ +2R_3 \\ \\ -R_3 \end{array}$$

가장 간단. 이걸 이용한다.

⇓

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 2 \end{array}\right]$$

A^{-1} (항상 역행렬이 존재하지는 않는다)

좌측 행렬이 단위행렬로 정리됨.

2.4 Vector Space: 벡터 공간.

- 벡터 : – 연산 (덧셈)에 대해 닫혀있고.
  - 스칼라 곱에 대해서 닫혀 있음.

↓

좀더 formal 한 정의를 위해 **group (군)**의 개념을 도입.

- Group

Definition 2.7 (Group). Consider a set $\mathcal{G}$ and an operation $\otimes : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ defined on $\mathcal{G}$. Then $G := (\mathcal{G}, \otimes)$ is called a *group* if the following hold:

1. *Closure* of $\mathcal{G}$ under $\otimes$: $\forall x, y \in \mathcal{G} : x \otimes y \in \mathcal{G}$ 닫혀있음.
2. *Associativity:* $\forall x, y, z \in \mathcal{G} : (x \otimes y) \otimes z = x \otimes (y \otimes z)$ 분배법칙
3. *Neutral element:* $\exists e \in \mathcal{G}\, \forall x \in \mathcal{G} : x \otimes e = x$ and $e \otimes x = x$ 항등원
4. *Inverse element:* $\forall x \in \mathcal{G}\, \exists y \in \mathcal{G} : x \otimes y = e$ and $y \otimes x = e$, where e is 역원. the neutral element. We often write x^{-1} to denote the inverse element of x.

조건을 만족할 때 그 집합은 ⊗에 대하여 group 이라함.
이때 group G의 임의의 x, y에 대하여 $x \otimes y = y \otimes x$가 보장될 경우
(즉 연산에 대하여 교환법칙이 성립할 경우) 이를 abelian group 이라고 합니다.
다시 행렬로 돌아와서 역행렬이 존재하는 nxn 행렬들의 집합은
행렬의 곱에 대해서 group이 된다.
⇒ 이러한 group을 general linear group 이라하며 GL(n, R) 이라 칭한다

단, 행렬의 곱은 AB ≠ BA 이므로 GL(n, R)은 abelian 이 아니다.

현재까지 다룬 group.

inner operation : G의 원소에만 적용하는 연산

outer operation : G 외의 집합의 원소에 대해서도 적용 가능.

↳ ex) multiplication of a vector $x \in G$ by a scalar $\lambda \in R$

벡터 공간 정의)

Definition 2.9 (Vector Space). A real-valued *vector space* $V = (\mathcal{V}, +, \cdot)$ is a set $\mathcal{V}$ with two operations

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V} \quad \text{inner op} \tag{2.62}$$
$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V} \quad \text{outer op} \tag{2.63}$$

where

1. $(\mathcal{V}, +)$ is an Abelian group
2. Distributivity:
   1. $\forall \lambda \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y}$
   2. $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : (\lambda + \psi) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \psi \cdot \boldsymbol{x}$
3. Associativity (outer operation): $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : \lambda \cdot (\psi \cdot \boldsymbol{x}) = (\lambda\psi) \cdot \boldsymbol{x}$
4. Neutral element with respect to the outer operation: $\forall \boldsymbol{x} \in \mathcal{V} : 1 \cdot \boldsymbol{x} = \boldsymbol{x}$

위 조건을 만족하는 x를 vector 라 한다.

The elements $\boldsymbol{x} \in V$ are called *vectors*. The neutral element of $(\mathcal{V}, +)$ is the zero vector $\boldsymbol{0} = [0, \dots, 0]^\top$, and the inner operation $+$ is called *vector addition*. The elements $\lambda \in \mathbb{R}$ are called *scalars* and the outer operation $\cdot$ is a *multiplication by scalars*. Note that a scalar product is something different, and we will get to this in Section 3.2.

벡터 공간 예)

Example 2.11 (Vector Spaces)
Let us have a look at some important examples:

- $\mathcal{V} = \mathbb{R}^n, n \in \mathbb{N}$ is a vector space with operations defined as follows:
  - Addition: $\boldsymbol{x} + \boldsymbol{y} = (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$
  - Multiplication by scalars: $\lambda \boldsymbol{x} = \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$ for all $\lambda \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^n$
- $\mathcal{V} = \mathbb{R}^{m \times n}, m, n \in \mathbb{N}$ is a vector space with
  - Addition: $\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$ is defined elementwise for all $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{V}$
  - Multiplication by scalars: $\lambda \boldsymbol{A} = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}$ as defined in Section 2.2. Remember that $\mathbb{R}^{m \times n}$ is equivalent to $\mathbb{R}^{mn}$.
- $\mathcal{V} = \mathbb{C}$, with the standard definition of addition of complex numbers.

Vector Subspaces (벡터 부분 공간)

- vector space에 정의된 연산을 vector subspace에 적용했을 때
  그 결과가 vector subspace를 절대 벗어나지 않는 경우.

ex) vector space = 정수, scalar = 정수

vector subspace = 짝수의 집합.

↳ 아무리 더하고 빼고 정수배를 하더라도 계속 짝수에 머무르게 됨.
이러한 의미에서, vector subspace는 closed (닫혀있다) 라고 말한다.
⇒ dimension reduction (차원 축소)

Definition 2.10 (Vector Subspace). Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{U} \subseteq \mathcal{V}, \mathcal{U} \neq \emptyset$. Then $U = (\mathcal{U}, +, \cdot)$ is called *vector subspace* of V (or *linear subspace*) if U is a vector space with the vector space operations $+$ and $\cdot$ restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$. We write $U \subseteq V$ to denote a subspace U of V.

U 또한 V의 두 연산 +, •에 대해 vector space가 된다.
+: U×U (이항 연산), •: R×U (실수와 U에 대한 스칼라의 곱)

U가 V의 부분집합일 때 vector subspace가 되는지 아닌지를 확인하기 위해서는 다음 2개만 확인하면 됨.

① U는 공집합이 아니며, 0은 U의 원소이다.

② U의 closure.
: 스칼라의 곱에 대해서도, 벡터의 덧셈에 대해서도 닫혀있다.

vector subspace의 예시)

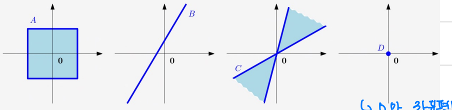
Example 2.12 (Vector Subspaces)
Let us have a look at some examples:

sub space.

- For every vector space V, the trivial subspaces are V itself and $\{\mathbf{0}\}$.
- Only example D in Figure 2.6 is a subspace of $\mathbb{R}^2$ (with the usual inner/outer operations). In A and C, the closure property is violated; B does not contain $\mathbf{0}$.
- The solution set of a homogeneous system of linear equations $\boldsymbol{Ax} = \mathbf{0}$ with n unknowns $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ is a subspace of $\mathbb{R}^n$.
- The solution of an inhomogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$, $\boldsymbol{b} \neq \mathbf{0}$ is not a subspace of $\mathbb{R}^n$. (∵ 0이 포함 X)
- The intersection of arbitrarily many subspaces is a subspace itself.

의 해의 집합은 Rn의 subspace가 됨.

subspace들의 임의의 교집합에 대한 결과도 subspace가 된다.



closure X　　0을 포함X　　closure X

↳ D만 좌표평면의 sub space가 된다.

2.5 Linear Independence

Linear Combination 정의)

Definition 2.11 (Linear Combination)**.** Consider a vector space V and a finite number of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$. Then, every $\boldsymbol{v} \in V$ of the form

$$\boldsymbol{v} = \lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i \in V \tag{2.65}$$

with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ is a *linear combination* of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.

linear independence vs linear dependence .

Definition 2.12 (Linear (In)dependence)**.** Let us consider a vector space V with $k \in \mathbb{N}$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$. If there is a non-trivial linear combination, such that $\boldsymbol{0} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ with at least one $\lambda_i \neq 0$, the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ are *linearly dependent*. If only the trivial solution exists, i.e., $\lambda_1 = \ldots = \lambda_k = 0$ the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ are *linearly independent*.

만약 $x_1, x_2, x_3, \cdots, x_k$ 가 linearly dependent 하다면.
어느 하나의 벡터 x_i 는 x_i를 제외한 다른 벡터들의 조합으로 표현 가능.
즉, 다른 벡터들의 값에 따라 영향을 받게된다.

linearly independent 하다면. 불필요한 중복이 없는 상태들로 구성이 되어 있다는 뜻.

<linearly Independence의 성질>

- k개의 벡터들은 linearly dependent하거나 linearly independent 하거나 둘 중 하나입니다. 또 다른 옵션은 없습니다.
- x1, x2, ... , xn의 벡터 중 하나가 0-벡터이면 linearly dependent합니다. 만약 이 중 두 벡터가 같아도 마찬가지입니다. (즉, 불필요한 벡터가 중복되어 포함되어있는 상태란 의미입니다.)
- 만약 벡터 집합 {x1, x2, ... , xn}이 linearly dependent하다고 한다면 그 중 하나는 다른 벡터들의 선형 조합으로 표현이 가능합니다. 특히 벡터 중 중 하나가 다른 벡터의 실수배일 때도 마찬가지입니다.
- 이 벡터들이 linearly dependent인지 아닌지는 Gaussian Elimination을 통해 쉽게 알 수 있습니다. 이 벡터들로 행렬을 만들어놓고 row-echelon form(Reduced Row Echelon Form까지 갈 필요도 없습니다)을 만들어보면 0만으로 나타나는 행이 나오는지 아닌지를 보면 됩니다.

- Row Echelon Form의 형태에서 pivot 열들은 linearly independent합니다.
- Row Echelon Form의 형태에서 pivot이 아닌 벡터들은 pivot 열들의 일차 결합으로 표현이 가능합니다. (일반해를 구할 때 Ax=0을 구하는 과정에서 이러한 방법을 이미 사용했습니다.)

만약 행렬의 모든 열 벡터들이 linearly independent 하다는 것과 모든 열들이 pivot 열이라고 하는 것은 필요충분조건입니다. 즉 하나라도 pivot이 아닌 열이 있다고 한다면 열 벡터들은 linearly dependent하게 됩니다. 다음의 예시를 통해서 살펴봅시다.

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix}$$

이 벡터들은 linearly independent 할까요?
이를 알아보기 위해 먼저 정의를 통해 살펴봅시다.

$$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \lambda_3 \boldsymbol{x}_3 = \lambda_1 \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \end{bmatrix} = \boldsymbol{0}$$

각 벡터 앞에 $\lambda_1, \lambda_2, \lambda_3$ 를 각각 곱해서
그 벡터들의 합이 0벡터가 되었을 때
각 $\lambda_1=\lambda_2=\lambda_3=0$ 이 됨을 보이면 되겠죠?

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

그런데 이 벡터들을 열로 하는 행렬을 만들어서
Row Echelon Form을 만들었더니 각각의 열이
모두 pivot이 되어있습니다. 따라서 처음에 살펴본
벡터들은 linearly independent 하다는 걸 알 수 있어요.