Linear System

수업 목표

이번 수업의 핵심:

- 선형방정식과 Linear system의 정의와 예시
- 행렬을 통한 Linear system의 표현
- 항등행렬과 역행렬의 정의와 특성
- 역행렬과 행렬식을 활용한 Linear system 풀이

핵심 개념

- Linear system, 선형방정식
- 항등행렬
- 역행렬
- 행렬식

선형방정식

- 변수 $x_1, \cdots, x_n$를 가지는 선형방정식:
  - 계수(Coefficient) $a_1, \cdots, a_n$와 상수 b는 실수/허수 형태의 사전에 알려진 숫자

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

$$\mathbf{a}^\top \mathbf{x} = b \quad \text{where } \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- 예시) $3x + 5y = 4$

$$\mathbf{a} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, b = 4, \qquad \mathbf{a}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{a} = \begin{bmatrix} 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1 + 5x_2 = b = 4$$

- $x + 2y + 3xy = 5$는 선형방정식?

  →$3xy$ 항이 2차항이기 때문에 아님, 모든 항이 선형이어야 함

Linear System: 선형방정식의 집합

Linear system (선형계, System of Linear Equations)

- 변수 $x_1, \cdots, x_n$를 공유하는 여러 선형 방정식의 집합

예시) 사람들의 체중, 신장, 흡연 여부 및 그 사람이 얼마나 살았는 지를 수집

번호	체중 (kg)	신장 (cm)	흡연 여부	수명
1	65	180	흡연자 (=1)	72
2	50	160	비흡연자 (=0)	82
3	60	170	흡연자 (=1)	68

- 위의 정보를 다음과 같은 Linear system로 표현:

$$65x_1 + 180x_2 + 1 \times x_3 = 72$$
$$50x_1 + 160x_2 + 0 \times x_3 = 82$$
$$60x_1 + 170x_2 + 1 \times x_3 = 68$$

계수 (Coefficient)

- 만약 x_1, x_2, x_3을 구할 수 있다면,
  체중/키/흡연 여부로 새로운 사람의 수명을 예측해 볼 수 있음

행렬을 통한 Linear System의 표현

- Linear system의 핵심 정보들은 **행렬**을 통하여 간결하게 표현이 가능

$$\begin{aligned} 65x_1 + 180x_2 + 1{\times}x_3 &= 72 \\ 50x_1 + 160x_2 + 0{\times}x_3 &= 82 \\ 60x_1 + 170x_2 + 1{\times}x_3 &= 68 \end{aligned}$$

1. 위 수식의 모든 계수를 행렬로 구성하고, 남은 정보를 두 벡터로 표현

$$A = \begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix}$$

2. 여러 선형식이 하나의 행렬식으로 표현

$$\begin{aligned} 65x_1 + 180x_2 + 1{\times}x_3 &= 72 \\ 50x_1 + 160x_2 + 0{\times}x_3 &= 82 \\ 60x_1 + 170x_2 + 1{\times}x_3 &= 68 \end{aligned} \Longrightarrow \underbrace{\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} \underset{=}{=} \underbrace{\begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix}}_{\mathbf{b}}$$

Linear System를 통한 행렬 해석

- 하나의 행렬은 선형식들의 계수 집합으로 해석 가능

$$A = \begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix} = \begin{bmatrix} -\mathbf{a}_1^\top - \\ -\mathbf{a}_2^\top - \\ -\mathbf{a}_3^\top - \end{bmatrix}$$

- 행렬식을 분해하여 여러 선형식으로 표현

$$\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix} \Leftarrow \begin{matrix} \mathbf{a}_1^\top \mathbf{x} = 72 \\ \mathbf{a}_2^\top \mathbf{x} = 82 \\ \mathbf{a}_3^\top \mathbf{x} = 68 \end{matrix}$$

$$\begin{bmatrix} 65 & 180 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_1^\top \mathbf{x} = 72, \quad \begin{bmatrix} 50 & 160 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_2^\top \mathbf{x} = 82, \quad \begin{bmatrix} 60 & 170 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{a}_3^\top \mathbf{x} = 68$$

→ 그래서 $\mathbf{x}$를 구하는 방법은?

항등행렬

항등행렬 (Identity Matrix)

- 대각선의 요소들만 1이고 나머지 요소들은 0인 정사각행렬
- $n \times n$ 크기의 항등행렬은 $I_n \in \mathbb{R}^{n \times n}$로 표현

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- 벡터 $\mathbf{x} \in \mathbb{R}^n$에 항등행렬 I_n를 곱하면 동일한 $\mathbf{x}$가 나옴

$$\forall \mathbf{x} \in \mathbb{R}^n, \qquad I_n \mathbf{x} = \mathbf{x}$$

역행렬

역행렬 (Inverse Matrix)

- 정사각행렬 $A \in \mathbb{R}^{n\times n}$의 역행렬 A^{-1}는 다음과 같이 정의:

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

- 2×2 크기 행렬 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$의 역행렬 A^{-1}은 다음과 같음

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

예시) 행렬 $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 2}$의 역행렬 A^{-1}

$$A^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \frac{1}{1\times 4 - 2\times 3}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$$

- $AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$
- $A^{-1}A = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$

역행렬을 활용한 Linear System 풀이

- 행렬식 $A\mathbf{x} = \mathbf{b}$는 다음과 같이 풀 수 있음

$$A\mathbf{x} = \mathbf{b} \implies A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \implies I_n\mathbf{x} = A^{-1}\mathbf{b} \implies \mathbf{x} = A^{-1}\mathbf{b}$$

예시)

$$\underbrace{\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix}}_{\mathbf{b}} \qquad A^{-1} = \begin{bmatrix} 0.5333 & -0.0333 & -0.5333 \\ -0.1667 & 0.0167 & 0.1667 \\ -3.6667 & -0.8333 & 4.6667 \end{bmatrix}$$

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 0.5333 & -0.0333 & -0.5333 \\ -0.1667 & 0.0167 & 0.1667 \\ -3.6667 & -0.8333 & 4.6667 \end{bmatrix} \begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix} = \begin{bmatrix} -0.6 \\ 0.7 \\ -15 \end{bmatrix}$$

- 따라서 어떤 사람의 수명은 다음과 같이 예측:

$$(\text{수명}) = -0.6 \times (\text{체중}) + 0.7 \times (\text{신장}) - 15 \times (\text{흡연 여부})$$

역행렬이 존재하지 않는 경우

- A가 역행렬이 존재한다면(Invertible), $\mathbf{x}$는 유일한 해가 존재

$$\mathbf{x} = A^{-1}\mathbf{b}$$

- 만약 A가 역행렬이 없다면(Non-invertible)?
  - 예시) $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $A^{-1} = \frac{1}{1\times4-2\times2}\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$, 이때 분모가 $1\times4 - 2\times2 = 0$

- $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$일 때, $ab - bc$를 A의 행렬식(Determinant)이라고 부름

$$\det A = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad \text{where } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

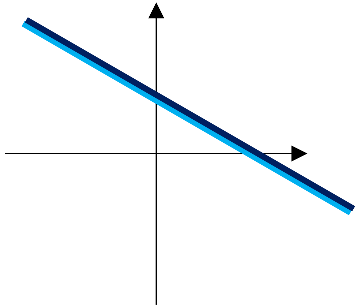
역행렬이 존재하지 않는 경우

- 만약 A가 역행렬이 없다면, $A\mathbf{x} = \mathbf{b}$는 해가 없거나, 무한히 많은 해를 가짐

1. 해가 무한히 많은 경우

$$A\mathbf{x} = \mathbf{b} \Rightarrow \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \end{bmatrix} \Rightarrow \begin{cases} x + 2y = 4 \\ 3x + 6y = 12 \end{cases}$$

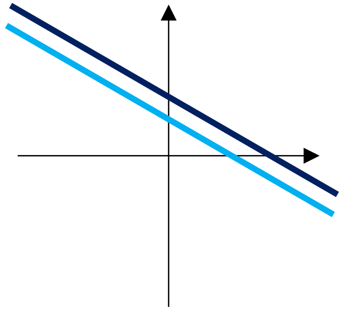
서로 겹침
(무한히 많은 해)



2. 해가 없는 경우

$$A\mathbf{x} = \mathbf{b} \Rightarrow \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 13 \end{bmatrix} \Rightarrow \begin{cases} x + 2y = 4 \\ 3x + 6y = 13 \end{cases}$$

평행
(해 없음)



행렬이 역행렬을 가지기 위한 조건

- 2×2 크기 행렬 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times2}$은 $ad - bc$가 역행렬 존재 여부를 결정
  - 때문에 $ad - bc$를 A의 행렬식(Determinant), 혹은 $\det A$ 라고 부름

- $\det A$는 A가 역행렬이 존재 여부를 결정(Determine)

- 3×3 크기 행렬 $A \in \mathbb{R}^{3\times3}$의 경우:

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix} = aei + bfg + cdh - ceg - afh - bdi$$

  - 혹은 다음과 같이 표현

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

고차원 행렬의 역행렬

- $n \geq 3$인 행렬 $A \in \mathbb{R}^{n \times n}$의 역행렬의 존재 여부 또한 행렬식으로 확인
  - 고차원 행렬의 행렬식 계산은 아래 참조

- 고차원 행렬 $A \in \mathbb{R}^{n \times n}$의 역행렬을 구하는 쉬운 방법(Closed-form)은 없음
  - 가우스 소거법(Gaussian elimination)을 활용한 계산 방법은 아래 참조

$n \geq 3$ 고차원 행렬의 행렬식 계산:
https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/lecture-18-properties-of-determinants/
https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/lecture-19-determinant-formulas-and-cofactors/

가우스 소거법을 활용한 $n \geq 3$ 고차원 행렬의 역행렬 계산:
https://home.cs.colorado.edu/~alko5368/lecturesCSCI2820/mathbook.pdf의 Chapter 1.2, 2.2를 참고

직사각행렬의 Linear System

- 만약 A가 직사각행렬, 즉 $A \in \mathbb{R}^{m \times n}, m \neq n$이라면?

번호	체중 (kg)	신장 (m)	흡연 여부	수명
1	65	180	흡연자 (=1)	72
2	50	160	비흡연자 (=0)	82
3	60	170	흡연자 (=1)	68

$$\underbrace{\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 72 \\ 82 \\ 68 \end{bmatrix}}_{\mathbf{b}}$$

- m = (식의 개수), n = (변수의 개수)라는 점을 고려
  - $m < n$: 변수가 수식보다 많음

    → Under-determined system: 일반적으로 무한히 많은 해가 존재
  - $m > n$: 수식이 변수보다 많음

    → Over-determined system: 일반적으로 해가 없음

일반적인 경우에서 해를 구하는 방법:
https://home.cs.Colorado.edu/~alko5368/lecturesCSCI2820/mathbook.pdf의 Chapter 1.2, 2.5를 참고

Under-Determined System

- Under-determined system에서는 우리 모델이 다양한 형태로 나올 수 있

번호	체중 (kg)	신장 (m)	흡연 여부	수명
1	65	180	흡연자 (=1)	72
2	50	160	비흡연자 (=0)	82

$$\begin{bmatrix}65 & 180 & 1\\ 50 & 160 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}72\\ 82\end{bmatrix}$$

$$A\mathbf{x}=\mathbf{b}$$

  - 해답 1: (수명) $= -0.6\times$(체중) $+ 0.7\times$(신장) $- 15\times$(흡연 여부)
  - 해답 2: (수명) $= -2.314\times$(체중) $+ 1.236\times$(신장) $+ 0\times$(흡연 여부)
  - 해답 3: (수명) $= -9999\times$(체중) $+ 3125.2\times$(신장) $+ 87471\times$(흡연 여부)

- 무한히 많은 해가 존재할 경우, Regularization(규제)을 활용
  - 해답 3과 같은 극단적인 해가 나오는 위험을 관리하기 위함
  - 예시) x_1, x_2, x_3의 절대값(L1 Regularization) 혹은 제곱합(L2 Regularization) 을 작게 유지

Over-Determined System

- Over-determined system에서는 일반적으로 해답이 없음

번호	체중 (kg)	신장 (m)	흡연 여부	수명
1	65	180	흡연자 (=1)	72
2	50	160	비흡연자 (=0)	82
3	60	170	흡연자 (=1)	68
4	73	174	비흡연자 (=0)	75

$$\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \\ 73 & 174 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 72 \\ 82 \\ 68 \\ 75 \end{bmatrix}$$

$$A \quad \mathbf{x} = \mathbf{b}$$

- 해가 없을 경우, 오차를 최소화하는 해답을 찾음
  - 오차의 절대값 혹은 오차의 제곱합을 최소화
  - 해답 1:

$$\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \\ 73 & 174 & 0 \end{bmatrix} \begin{bmatrix} -0.6 \\ 0.7 \\ -15 \end{bmatrix} = \begin{bmatrix} 72 \\ 82 \\ 68 \\ 78 \end{bmatrix} \Rightarrow \text{오차: 9 (제곱합)}$$

$$A \quad \mathbf{x}_1 = \mathbf{b}'$$

해답 2:

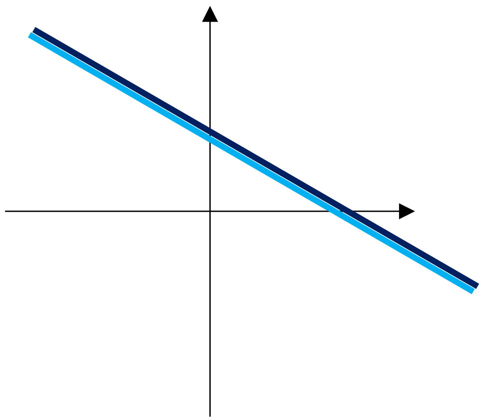
$$\begin{bmatrix} 65 & 180 & 1 \\ 50 & 160 & 0 \\ 60 & 170 & 1 \\ 73 & 174 & 0 \end{bmatrix} \begin{bmatrix} -0.76 \\ 0.75 \\ -13.7 \end{bmatrix} = \begin{bmatrix} 71.9 \\ 82 \\ 68.2 \\ 75.02 \end{bmatrix} \Rightarrow \text{오차: 0.05 (제곱합)}$$

$$A \quad \mathbf{x}_2 = \mathbf{b}'$$

요약

- Linear system의 정의와 행렬을 통한 Linear system 표현 방법 학습
- 역행렬의 특성 및 이를 활용한 Linear system 풀이
- 행렬식의 의미와 역행렬이 존재할 조건 확인
- Under/Over-determined system의 특성 및 해답 계산 소개

서로 겹침
(무한히 많은 해)



평행
(해 없음)