Linear Algebra

주재걸
고려대학교 컴퓨터학과

Lecture Overview

- Elements in linear algebra
- Linear system
- Linear combination, vector equation,
  Four views of matrix multiplication
- Linear independence, span, and subspace
- Linear transformation
- Least squares
- Eigendecomposition
- Singular value decomposition

Linear Equation 선형 방정식

- A **linear equation** in the variables $x_1, \cdots, x_n$ is an equation that can be written in the form
변수들

$$a_1x_1+a_2x_2+\cdots+a_nx_n = b, \rightarrow \text{선형 방정식}$$

where b 상수 and the coefficients $a_1, \cdots, a_n$ 계수 are real or complex numbers that are usually known in advance.

- The above equation can be written as

내적

소문자, bold → matrix $\mathbf{a}^T\mathbf{x} = b$

where $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

$\underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}}_{a^T} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x}$

Linear System: Set of Equations

→ 선형 방정식의 집합

- A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations involving the same variables - say, $x_1, \cdots, x_n$.

Linear System Example

- Suppose we collected persons' weight, height, and life-span (e.g., how long s/he lived)

feature들 ✓ target

Person ID	Weight	Height	Is_smoking	Life-span
1	60kg	5.5ft	Yes (=1)	66
2	65kg	5.0ft	No (=0)	74
3	55kg	6.0ft	Yes (=1)	78

- We want to set up the following linear system:

$$
\left.\begin{aligned}
60x_1 + 5.5x_2 + 1\cdot x_3 &= 66\\
65x_1 + 5.0x_2 + 0\cdot x_3 &= 74\\
55x_1 + 6.0x_2 + 1\cdot x_3 &= 78
\end{aligned}\right\}
$$

를 모두 만족하는 변수 찾기

- Once we solve for x_1, x_2, and x_3, given a new person with his/her weight, height, and is_smoking, we can predict his/her life-span.

Linear System Example

- The essential information of a linear system can be written compactly using a **matrix**.
- In the following set of equations,

$$
\begin{aligned}
60x_1 + 5.5x_2 + 1\cdot x_3 &= 66\\
65x_1 + 5.0x_2 + 0\cdot x_3 &= 74\\
55x_1 + 6.0x_2 + 1\cdot x_3 &= 78
\end{aligned}
$$

- Let's collect all the coefficients on the left and form a matrix

$$A = \begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix}$$

→ 계수만 모아놓음

- Also, let's form two vectors: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}$

가중치 벡터

상수 (target)

From **Multiple** Equations to **Single Matrix** Equation

- Multiple equations can be converted into a **single** matrix equations

$$
\begin{aligned}
60x_1+5.5x_2+1\cdot x_3 &= 66\\
65x_1+5.0x_2+0\cdot x_3 &= 74\\
55x_1+6.0x_2+1\cdot x_3 &= 78
\end{aligned}
\quad\Rightarrow\quad
\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
=
\begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}
\quad\Leftarrow\quad
\begin{aligned}
\mathbf{a}_1^T\mathbf{x} &= 66\\
\mathbf{a}_2^T\mathbf{x} &= 74\\
\mathbf{a}_2^T\mathbf{x} &= 78
\end{aligned}
$$

내적

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

등식 하나로 표현

- How can we solve for $\mathbf{x}$?

Identity Matrix

항등행렬 정사각행렬

- **Definition**: An identity matrix is a square matrix whose diagonal entries are all 1's, and all the other entries are zeros. Often, we denote it as $I_n \in \mathbb{R}^{n\times n}$.
  - e.g., $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- An identity matrix I_n preserves any vector $\mathbf{x} \in \mathbb{R}^n$ after multiplying $\mathbf{x}$ by I_n :

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

Inverse Matrix

역행렬

$$\begin{cases} A^{-1}A = I_n \\ AA^{-1} \neq I_n \,? \end{cases}$$

$$\boxed{\begin{bmatrix} & \end{bmatrix}}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\boxed{\begin{bmatrix} & \end{bmatrix}} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

→ 없다

정사각행렬

- **Definition**: For a square matrix $A \in \mathbb{R}^{n\times n}$, its inverse matrix A^{-1} is defined such that

$$\underbrace{A^{-1}A}_{\text{왼}} = \underbrace{AA^{-1}}_{\text{오}} = I_n.$$

항등행렬

- For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, its inverse matrix A^{-1} is defined as

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

$$\underbrace{\boxed{\begin{bmatrix} ? \end{bmatrix}}}_{3\times 2}\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}_{2\times 3} = \underbrace{I}_{3\times 3}$$

$\neq I$

→ 일부 존재. 그러나 오른쪽에 곱하면 항등행렬 안나옴 → 역행렬 아님

Solving Linear System via Inverse Matrix

- We can now solve $A\mathbf{x} = \mathbf{b}$ as follows:

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

Solving Linear System via Inverse Matrix

- **Example:**

$$\underbrace{\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} \underset{=}{=} \underbrace{\begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}}_{\mathbf{b}} \Longrightarrow A^{-1} = \begin{bmatrix} 0.0870 & 0.0087 & -0.0870 \\ -1.1304 & 0.0870 & 1.1314 \\ 2.0000 & -1.0000 & -1.0000 \end{bmatrix}$$

- One can verify

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

최적의 coefficients

- $\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 0.0870 & 0.0087 & -0.0870 \\ -1.1304 & 0.0870 & 1.1314 \\ 2.0000 & -1.0000 & -1.0000 \end{bmatrix} \begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix} = \boxed{\begin{bmatrix} -0.4 \\ 20 \\ -20 \end{bmatrix}}$

Solving Linear System via Inverse Matrix

- Now, the life-span can be written as

$$(\text{life-span}) = -0.4 \times(\text{weight})+20 \times(\text{height})$$
$$-20 \times(\text{is\_smoking}).$$

Non-Invertible Matrix A for $A\mathbf{x} = \mathbf{b}$

- Note that **if** A is invertible, the solution is uniquely obtained as

  = A에 역행렬 존재

  unique한 값으로 (하나의 값)

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

  = A에 역행렬 존재X

- **What if** A is non-invertible, i.e., the inverse does not exist?
  - E.g., For $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, in $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, the denominator $ad - bc = 0$, so A is not invertible.

    ↳ 이면 역행렬 존재X $ad = bc \leftrightarrow a:b = c:d$

- For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $ad - bc$ is called the determinant of A, or $\boxed{\det A}$.

  판별식

$\det A = 0 \leftrightarrow$ 역행렬 존재X

$\det A \neq 0 \leftrightarrow$ 역행렬 존재

Does a Matrix Have an Inverse Matrix?

- $\det A$ determines whether A is invertible (when $\det A \neq 0$) or not (when $\det A = 0$).

- For more details on how to compute the determinant of a matrix $A \in \mathbb{R}^{n \times n}$ where $n \geq 3$, you can study the following:
  - 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/

3x3, 4x4, … 인 경우의 determinant

Inverse Matrix Larger than 2×2

- If invertible, is there any formula for computing an inverse matrix of a matrix $A \in \mathbb{R}^{n \times n}$ where $n \geq 3$?

- No, but one can compute it.
- We skip details, but you can study Gaussian elimination in Lay Ch1.2 and then study Lay Ch2.2.

Non-Invertible Matrix A for $A\mathbf{x} = \mathbf{b}$

- Back to the linear system, if A is non-invertible, $A\mathbf{x} = \mathbf{b}$ will have either no solution or infinitely many solutions.

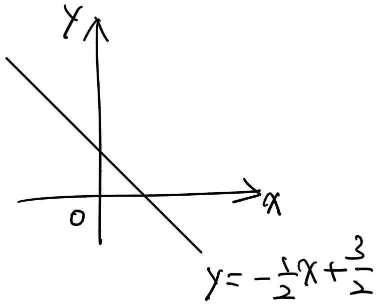
해가 없다 해가 무수히 많다

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ \cancel{6}\,7 \end{bmatrix} \qquad y = -\frac{1}{2}x + \frac{3}{2}$$

이면 no solution

$$x + 2y = 3$$
$$2x + 4y = 6$$

$\times 2$ 해가 무수히 많다



Rectangular Matrix A in $A\mathbf{x} = \mathbf{b}$

A가 직사각행렬

- **What if** A is a rectangular matrix, e.g., $A \in \mathbb{R}^{m \times n}$, where $m \neq n$?

3 n

Person ID	Weight	Height	Is_smoking	Life-span
1	60kg	5.5ft	Yes (=1)	66
2	65kg	5.0ft	No (=0)	74
3	55kg	6.0ft	Yes (=1)	78

3 m

$$\underbrace{\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 66 \\ 74 \\ 78 \end{bmatrix}}_{\mathbf{b}}$$

- Recall m = #equations and n = #variables.
- $m < n$: more variables than equations 데이터 수 < feature 수(미지수 개수) → 해가 무수히 많다 regularization
  - Usually infinitely many solutions exist (under-determined system).
- $m > n$: more equations than variables 데이터 수 > feature 수 → 해가 없다
  - Usually no solution exists (over-determined system).
- To study how to compute the solution in these general cases, check out Lay Ch1.2 and Lay Ch1.5.

⇒ 가장 근사적으로 만족하는 경우를 찾자.