2.3 Solving Systems of Linear Equations

특수해, 일반해 (particular and general solution)

간단한 예시 (특수해, 일반해)

- 실수 x, y에 대하여 다음 방정식의 해를 구하면?

특수해 (particular Solution)

- 하나만 존재하는 것이 아닌, 매우 많이 존재
- 해가 (x, y) 튜플이라고 했을 때,
  - $(3,0),(1,2),\cdots$
- 위 순서쌍 모두가 특수해가 될 수 있다.

일반해 (General Solution)

- $x = t, t$는 실수
- $y = 3 - x$

  $= 3 - t$

  $\therefore (t, 3-t), \quad t \in \mathbb{R}$

교재 예시 (특수해, 일반해)

해당 교재에서는 위 간단한 예시처럼 특수해를 거쳐 일반해를 찾는 방법을 사용합니다.

$$\begin{bmatrix} 1 & 0 & 8 & -4 \\ 0 & 1 & 2 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 42 \\ 8 \end{bmatrix}.$$

- $x_1 + 0x_2 + 8x_3 - 4x_4 = 42$

  $0x_1 + x_2 + 2x_3 + 12x_4 = 8$

- $\begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_2 + \begin{bmatrix} 8 \\ 2 \end{bmatrix} x_3 + \begin{bmatrix} -4 \\ 12 \end{bmatrix} x_4 = \begin{bmatrix} 42 \\ 8 \end{bmatrix}$

해당 시스템에서는 두 개의 방정식과 4개의 미지수가 있습니다.

따라서 우리는 무한히 많은 해가 존재한다는 것을 예상할 수 있습니다.

특수해 구하기

해당 식에서 x_3, x_4를 0이라고 했을 때 우리는 매우 쉽게 특수해를 구할 수 있습니다.

하나의 특수해는 $x_1 = 42, x_2 = 8, x_3 = 0, x_4 = 0$이 됩니다.

이는 $[42, 8, 0, 0]^T$로 표현할 수 있습니다.

일반해 구하기

다른 모든 해를 구하기 위해 $Ax = 0$을 만족하는 x를 찾습니다.

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_2 + \begin{bmatrix} \color{red}{8} \\ \color{red}{2} \end{bmatrix} x_3 + \begin{bmatrix} -4 \\ 12 \end{bmatrix} x_4 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x_3 = -1, \quad x_4 = 0$으로 두었을 때, $x_1 = 8, \quad x_2 = 2$ 가 될 것입니다.

$[8, 2, -1, 0]^T$로 표현할 수 있으며 해당 벡터에 λ(스칼라)를 곱해도 여전히 결과는 0입니다.

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

이번에는 x_4를 이용하여 같은 방법을 사용하면 아래 식을 얻을 수 있습니다.

$\lambda[-4, 12, 0, -1]^T$

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

위 두 방법을 통해 **일반해**를 다음과 같이 구할 수 있습니다.

$$\left\{ \boldsymbol{x} \in \mathbb{R}^4 : \boldsymbol{x} = \begin{bmatrix} 42 \\ 8 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 8 \\ 2 \\ -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} -4 \\ 12 \\ 0 \\ -1 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

정리 (특수해와 일반해)

3가지 과정을 통해 특수해를 통해 일반해를 구할 수 있습니다.

1. $Ax = b$ 에서 특수해를 구합니다.
2. $Ax = 0$ 에서 모든 해를 구합니다.
3. 1번 과정과 2번 과정을 통해 얻은 해를 결합하여 일반해를 구합니다.

해당 연립선형방정식은 쉬운 형태의 행렬이기에 특수해와 일반해를 찾을 수 있었습니다.
하지만 일반적인 연립선형방정식은 단순한 형태가 아닙니다.

그래서 **가우스 소거법(Gaussian elimination)**을 통해 모든 연립선형방정식을 간단한 형태로 변환하여 문제를 해결하는 방법을 사용합니다.
이때 핵심이 되는 것은
elementary transformation(기초 행 연산)을 적용해 나가는 과정입니다.

이러한 연산은 세 가지 유형으로 나뉘며, 이들을 통해 행렬을 점점 더 간단한 형태로 바꿔 갈 수 있습니다.
변환 과정은 다음과 같습니다.

- 두 방정식을 교호나한다. (행을 교환)
- 0이 아닌 상수 $\lambda \in \mathbb{R}$ 을 한 방정식에 곱한다.
- 두 방정식을 더한다.

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

먼저 위 연립선형방정식을 $Ax = b$의 형태로 변환합니다.

전에 변경했던 식보다 더 편하게 **augmented matrix(첨가행렬), $[A|b]$** 형태로 표현합니다.

- 변수 x를 생략하고, [계수 | 상수]로 표현합니다.

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

위 3가지의 변환 방법을 통해 다음과 같은 행렬을 얻을 수 있습니다.

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

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

$$\begin{array}{ccl} & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{array}\right] & \begin{array}{l} \\ \\ \\ -R_2 - R_3 \end{array} \\ \rightsquigarrow & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] & \begin{array}{l} \\ \cdot(-1) \\ \cdot(-\frac{1}{3}) \\ \\ \end{array} \\ \rightsquigarrow & \left[\begin{array}{rrrrr|r} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{array}\right] & \end{array}$$

이를 다시 연립선형방정식의 형태로 표현하면 다음과 같습니다.

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

위 식에서 a = -1 일 때, 해당 시스템은 해가 존재합니다. 특수해는 다음과 같습니다.

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

이를 통해, 일반해를 구하면 다음과 같습니다.

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

REF (Row - Echelon Form)

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

- 특성
  - 0이 아닌 모든 행은 모두 0행 위에 있다. (즉, 모두 0행은 맨 아래에 있다.)
  - 0이 아닌 행의 선행계수는 항상 위 행의 0이 아닌 첫 번째 항목의 오른쪽에 있다.
  - Pivot 아래의 모든 열 항목은 0이다.
- Pivot?

- Pivot → 행의 선행 계수 (왼쪽에서부터 0이 아닌 첫번 째 숫자)
- 각 행의 Pivot은 위에 있는 행의 Pivot의 오른쪽에 위치합니다.
- 따라서 모든 REF 형태의 연립방정식은 항상 계단(Step)구조입니다.

- Basic variables, free variables
    - REF에서 pivot에 대응되는 변수를 Basic variables라고 합니다.

      나머지 변수들을 free variables라고 합니다.

      위 REF에서는 x_1, x_3, x_4 → Basic | x_2, x_5 → Free

RREF (Reduced Row Echelon Form)

아래 조건을 성립하는 연립방정식을 RREF라고 합니다.

- REF
- 모든 Pivot 이 1
- 피벗은 해당 열에서 0이 아닌 유일한 요소일 때

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

RREF를 사용하면 $Ax = 0$ 꼴의 연립방정식의 해를 쉽게 구할 수 있습니다.

예시

> **가우스 소거법(Gaussian elimination)은 연립선형방정식을 RREF 으로 변형하는 알고리즘입니다.**

The Minus-1 Trick

The Minus-1 Trick은 동차 방정식(homogeneous system of linear equations) $Ax = 0$ 을 찾기 위한 방법입니다.

$$
\boldsymbol{A}=\begin{bmatrix}
0 & \cdots & 0 & \mathbf{1} & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & & \vdots & 0 & 0 & \cdots & 0 & \mathbf{1} & * & \cdots & * & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \mathbf{1} & * & \cdots & *
\end{bmatrix}
$$

- 행렬 $A \in \mathbb{R}^{\Bbbk \times n}$는 RREF 형태이고 다음과 같은 행렬이라고 가정. $(x \in \mathbb{R}^n)$
- $*$ 는 임의의 실수이며, 위 행렬의 어느 행에서든지 첫 번째로 0이 아닌 요소의 값은 1이며, 그 행에서 나머지 요소는 0이라고 가정
- 볼드체로 표시된 피벗을 가진 열을 $j_1, j_2, \cdots, j_k$ 는 표준 단위 벡터 $e_1, \cdots, e_k \in \mathbb{R}^{\Bbbk}$ 입니다.
- 위 행렬은 $n-k$ 행을 더하여 $n \times n$ 행렬 $\tilde{A}$ **행렬로 확장할 수 있습니다.**
- $n-k$ 행은 다음과 같은 형태를 가집니다. $[0 \ \cdots \ 0 \ -1 \ 0 \ \cdots \ 0]$

그러면, 피벗으로 -1을 포함하는 행렬 A~***A***~ 의 열은 homogeneous equation system =0***Ax***=**0** 의 해입니다. 조금 더 정확하게 말하자면, -1을 피벗으로 하는 열들은 =0***Ax***=**0** 의 solution space의 basis(기저)를 형성하며, 이를 kernel 또는 null space라고 부릅니다 → 해당 내용 정리하여 보충 자료 추가하기

Homogeneous Linaer System(동차선형계)

- 추후 정리하여 추가하기

예시 (The Minus-1 Trick)

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

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

- 위 행렬식을 → $n-k$ 행을 추가하여 5×5 형태로 변환합니다.
- 위 행렬로부터 대각 요소 -1을 포함하는 열을 가지고, 바로 $Ax = 0$ 의 해를 얻을 수 있습니다.

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

가우스 소거법을 통해 역행렬 구하기

Example 2.9 (Calculating an Inverse Matrix by Gaussian Elimination)
To determine the inverse of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.57}$$

we write down the augmented matrix

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

and use Gaussian elimination to bring it into reduced row-echelon form

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

such that the desired inverse is given as its right-hand side:

$$\boldsymbol{A}^{-1} = \begin{bmatrix} -1 & 2 & -2 & 2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 1 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}. \tag{2.58}$$

We can verify that (2.58) is indeed the inverse by performing the multiplication $\boldsymbol{A}\boldsymbol{A}^{-1}$ and observing that we recover $\boldsymbol{I}_4$.

Algorithms for Solving a System of Linear Equations

이번 장에서는 $Ax = b$ 형태의 연립선형방정식의 해를 구하는 방법에 대해 간략하게 살펴보았습니다.

해는 항상 존재한다고 가정하지만, 해가 존재 하지 않을 수도 있습니다. 해당 경우에는 **근사(approximate solution)** 해야 하며 선형 회귀 파트(ch.9)에서 다룹니다.

특수한 경우에는 역행렬을 구해 $x = A^{-1}b$ 로 해를 구할 수 있습니다. 하지만 이는 행렬 A가 정사각 행렬이면서 역행렬이 존재해야 합니다. 그렇지 않으면 A는 'linear independent columns 가 있다'라는 가정 하에 다음 변환을 사용할 수 있습니다.

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{b} \iff \boldsymbol{x} = (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{b}$$

그리고 Moore-Penrose pseudo-inverse(무어-펜로즈 유사 역행렬)를 사용하여 해를 구할 수 있습니다.

이런 접근 방법은 $A^T A$의 역행렬을 구하는데 많은 연산이 필요하다는 것입니다.

가우스 소거법(gaussian elimination)은

- 행렬식 계산
- 벡터 집합이 선형 독립이라는 것을 확인
- 행렬의 rank 계산
- 벡터 공간의 기저를 결정
- 역행렬

다양하게 활용됩니다. 하지만 이 방법 또한 행렬의 크기가 커지며 많은 연산량을 요구합니다.

이를 개선한 다양한 방법이 존재하지만, 핵심 아이디어는 반복법입니다.

- $x^{(k+1)} = Cx^{(k)} + d$
- 오차 (error) $||x^{(k+1)} - x_*||$ 를 매 반복마다 줄일 수 있는 C와 d를 탐색하며 x로 수렴

출처

- Mathmatics for Machine Learning (https://github.com/mml-book/mml-book.github.io)
- https://www.youtube.com/watch?v=E1WxLsUZHXA&list=PLDXd6V7UxqskqH5HYlXnP6sjLHfm_92-s&index=5
- https://junstar92.github.io/mml-study-note/2022/07/01/ch2-3.html
- https://blog.naver.com/walk_along/222158344290
- https://blog.naver.com/walk_along/222158345597
- https://deep-learning-study.tistory.com/527