

클린업 1주차

3팀 선형대수학

황정현
고경현
김지민
반경림
전효림

‘ 선형대수 ’

통계 분석의 시작점

선형성을 바탕으로
선형변환과 그때의 공간에 대해 연구하는
대수학의 한 분야



선형성

“ 함수 f 는 선형이다 ”

가산성

임의의 수 x 와 y 에 대해 $f(x + y) = f(x) + f(y)$

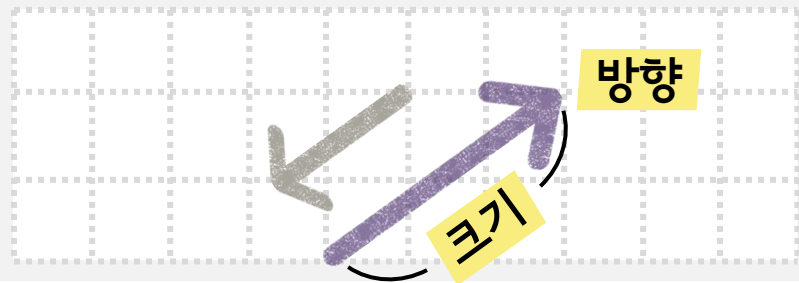
동차성

임의의 수 x 와 상수 a 에 대해 $f(ax) = a \cdot f(x)$

벡터(vector)의 개념

✓ 벡터의 물리적 의미

‘ 어느 **방향**으로 **얼마만큼**의
힘 또는 속도를 갖는지 ’
≠ 스칼라(*scalar*)



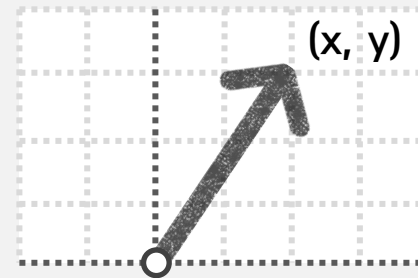
✓ 데이터분석의 기본 단위



다른 데이터를 구성하는 기본
+

데이터를 **선형적 관점**에서
이해할 수 있는 핵심 매개

✓ 선형대수학의 공간적 이해



2개의 원소



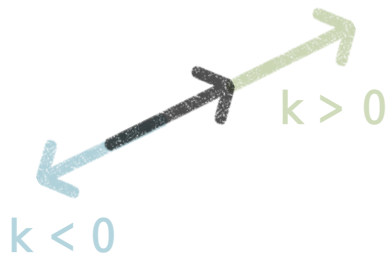
\mathbb{R}^2

- **원소**(component): 벡터를 구성하는 값
- **벡터공간**(vector space, \mathbb{R}): 벡터로 이뤄지는 공간
- **n개의 원소**로 이뤄진 벡터는 **n차원** 공간에 있음

벡터의 연산

2 기하학적으로 계산하기

상수배



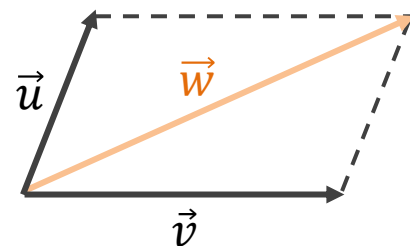
벡터에 상수 k 를 곱한다

=

벡터의 길이를 k 배 한다

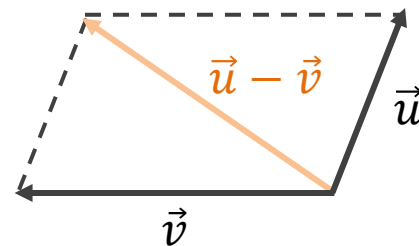
덧셈

벡터 u 에서 벡터 v 만큼
이동하면서 생기는
평행사변형의 대각선이
 $u+v$, 즉 벡터 w



뺄셈

벡터 v 에 대해 -1 배 한 뒤
벡터 u 와 더하면 벡터 $u-v$



행렬(matrix)의 개념

- 실수를 직사각형 모양의 행과 열로 배열한 형태

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

원소(element)
또는 성분(entry)

행 m 개와 열 n 개로 이루어진
 $m \times n$ 크기의 행렬

역행렬

$$\underset{\substack{n \times n \text{의} \\ \text{정방행렬}}}{\mathbf{A}} \cdot \mathbf{A}^{-1} = \underset{\substack{n \times n \text{의} \\ \text{단위행렬}}}{\mathbf{I}}$$

행렬 A 는 오직 하나의 역행렬만 가질 수 있기 때문에

Properties

2 x 2 matrix의 경우

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

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

$$(A^T)^{-1} = (A^{-1})^T$$

역행렬을 가진다

해가 유일하다 (unique)

가역하다 (invertible)

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

모두 같은 말!



선형방정식(Linear equation)



양의 정수 n 에 대하여, $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$ 형태로 표현되는 식

문자의 개수가 많거나 일반화된 해를 찾기 힘들 때

행렬과 벡터를 이용한 선형방정식의 꼴로 만들어 해결할 수 있다!

$$\begin{array}{l}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\
 \vdots \\
 a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m
 \end{array}
 \quad
 \begin{pmatrix}
 a_{11} & a_{12} & \dots & a_{1n} \\
 a_{21} & a_{22} & \dots & a_{2n} \\
 \vdots & \vdots & & \vdots \\
 a_{m1} & a_{m2} & \dots & a_{mn}
 \end{pmatrix}
 \begin{pmatrix}
 x_1 \\
 x_2 \\
 \vdots \\
 x_n
 \end{pmatrix}
 =
 \begin{pmatrix}
 b_1 \\
 b_2 \\
 \vdots \\
 b_m
 \end{pmatrix}$$

A

$m \times n$ 행렬

x

$n \times 1$ 벡터

b

$m \times 1$ 벡터

Ax = b 판별 및 해 구하기

2 해의 종류

일치한다

Consistent

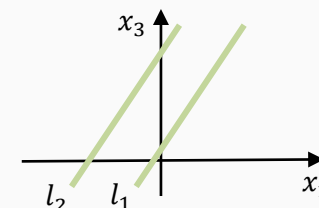
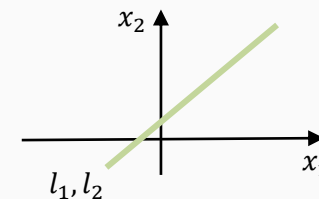
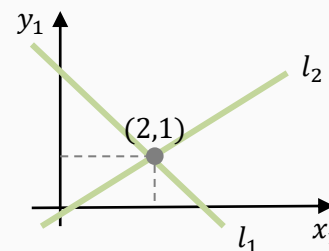
해가 유일하다

해가 무수히 많다

불일치한다

Inconsistent

해가 없다



Ax = b 판별 및 해 구하기

3 해의 판별

✓ 사전 작업

Ax = b 를 행렬 $[A \mid b]$ 로 만들고 G-J 소거법을 이용해
RREF인 $[H \mid c]$ 로 만든다

$$\begin{cases} -2x - 5y + 2z = -3 \\ x + 3y = 4 \\ y + 3z = 6 \end{cases}$$

$$\left[\begin{array}{ccc|c} -2 & -5 & 2 & -3 \\ 1 & 3 & 0 & 4 \\ 0 & 1 & 3 & 6 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

선형결합(Linear combination)과 span

- 벡터들을 상수배와 벡터 덧셈을 통해 조합하여 새로운 벡터를 얻는 연산

$$Ax = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

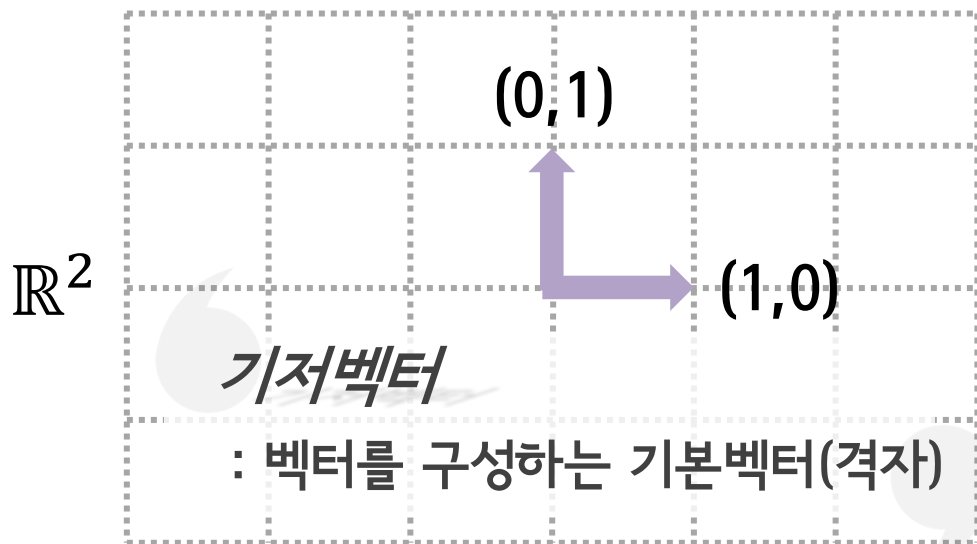
$\text{span}\{a_1 \ a_2 \ \cdots \ a_n\}$



$m \times n$ 행렬 A 와 n 개의 원소를 가진 벡터 x 를 곱할 때,

a_i 의 조합 = **span**

span의 공간적 이해



2차원 벡터공간 내
모든 벡터는 **기저벡터인**
 $(0, 1)$ 과 $(1, 0)$ 의
span(조합)을 통해
표현할 수 있음

span의 공간적 이해

다음주를 기대해주세요!

 \mathbb{R}^2 $(0, 1)$ $Ax = b$ 의 해가 존재한다 a 의 조합으로 b 를 표현할 수 있다

기저벡터

벡터 b 가 $\text{span}(a_1, a_2, \dots, a_n)$ 에 있다

: 벡터를 구성하는 기본벡터(격자)

2차원 벡터공간 내

모든 벡터는 기저벡터인

 $(0, 1)$ 과 $(1, 0)$ 의 $\text{span}(\text{조합})$ 을 통해

표현할 수 있음

선형 '변환'의 의미



선형방정식 $Ax = b$ 를 '변환'의 관점에서 이해해보자!

$$\begin{matrix} A \\ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \end{matrix} \begin{matrix} x \\ \begin{pmatrix} x \\ y \end{pmatrix} \end{matrix} = \begin{matrix} b \\ \begin{pmatrix} 0 \\ 3 \end{pmatrix} \end{matrix}$$

x라는 input에 A라는 변환을 거쳐 b라는 새로운 output 반환

선형 '변환'의 의미



INPUT $Ax = b$ 의 해를 찾고자 하는 것은 곧 OUTPUT
 A 를 곱했을 때 b 로 변환되는 벡터 x 를 찾는 것이다!

선형방정식 $Ax = b$ 를 '변환'의 관점에서 이해해보자!


$$\begin{matrix} A \\ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \end{matrix} \begin{matrix} x \\ \begin{pmatrix} x \\ y \end{pmatrix} \end{matrix} = \begin{matrix} b \\ \begin{pmatrix} 0 \\ 3 \end{pmatrix} \end{matrix}$$


x 라는 input에 A 라는 변환을 거쳐 b 라는 새로운 output 반환

‘선형’ 변환의 의미 (2) 공간적 의미

$$\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\begin{aligned} T \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= 1T \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= 1 \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \end{aligned}$$

 : 기존 기저

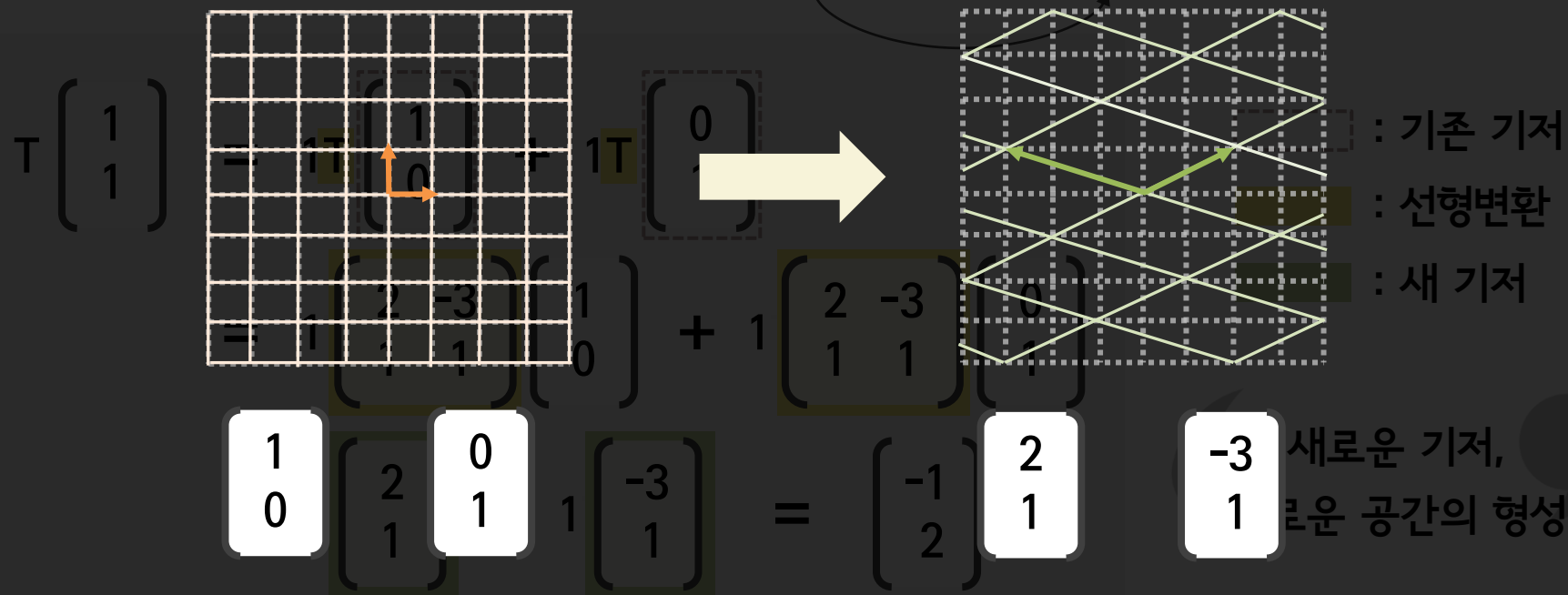
 : 선형변환

 : 새 기저

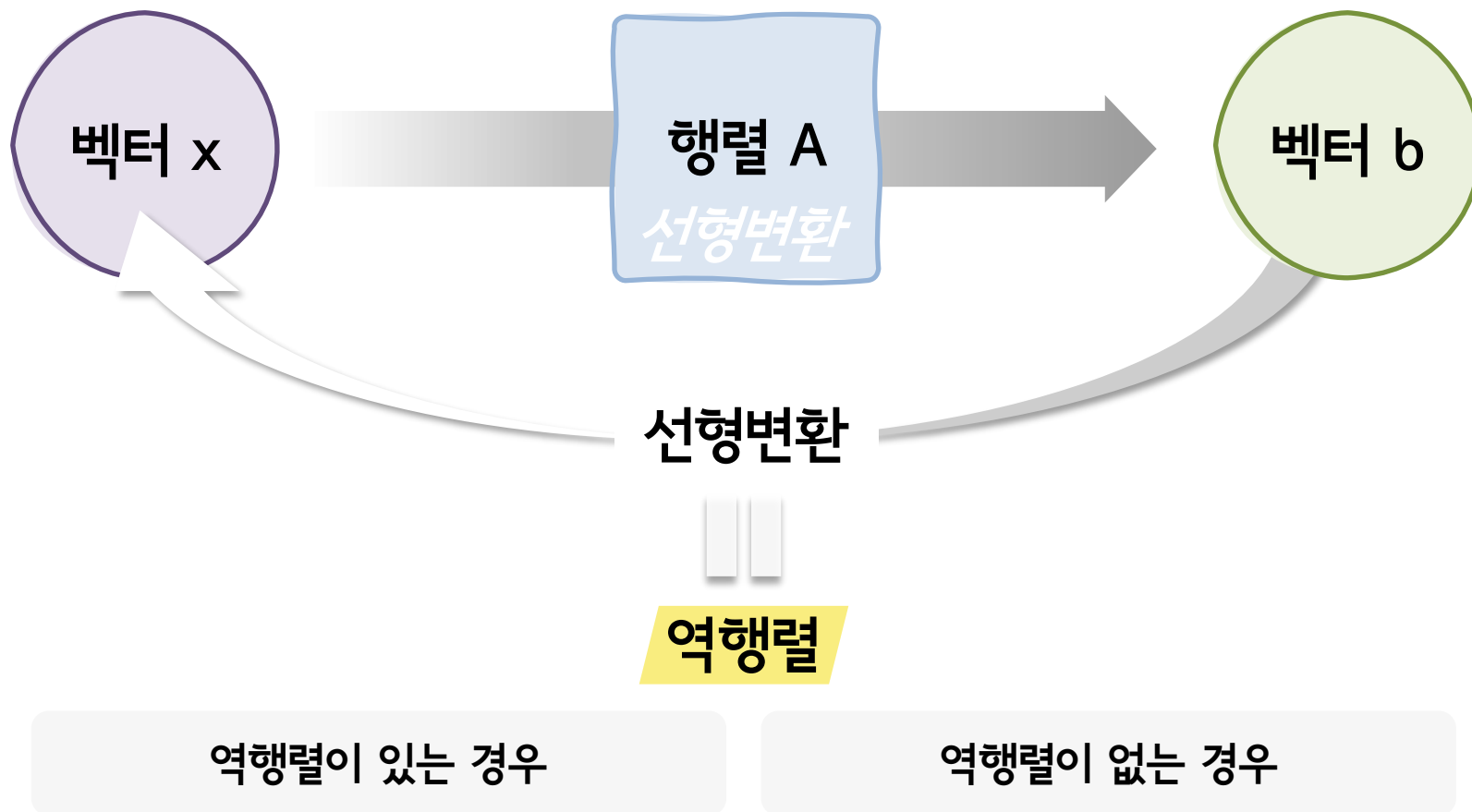
새로운 기저,
새로운 공간의 형성

‘선형’ 변환의 의미 (2) 공간적 의미

$$\begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \text{☀️} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$
 선형변환 이후 기저가 변해따!

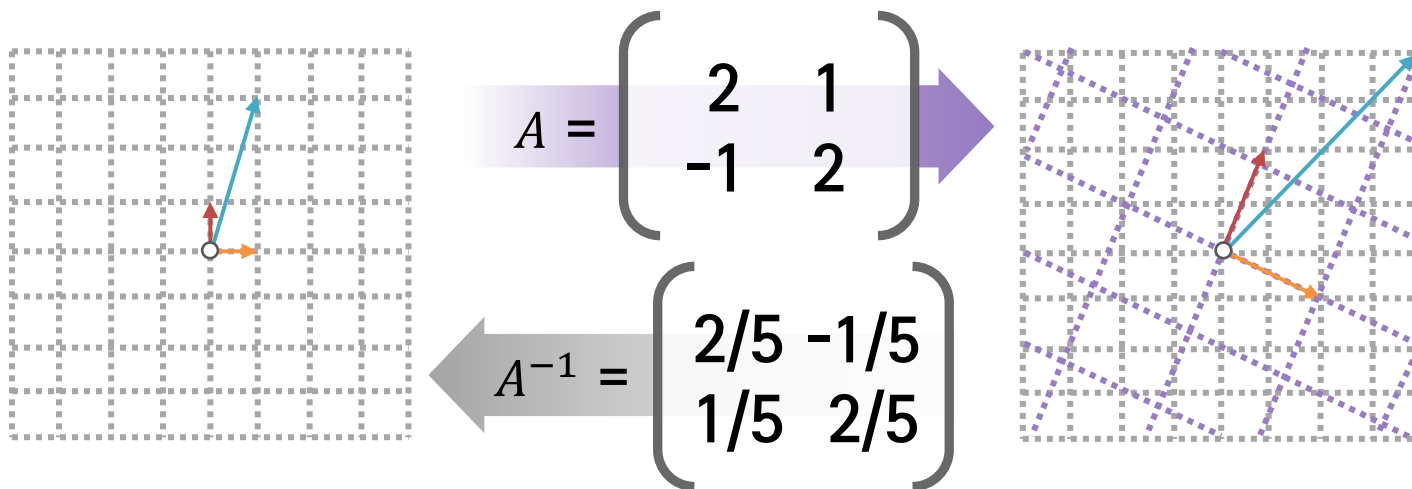


선형변환으로 역행렬 이해하기



선형변환으로 역행렬 이해하기

1 역행렬이 있는 경우



정방행렬 A 의 역행렬이 존재한다

- ⊖ $Ax = b$ 가 유일한 해를 갖는다(unique)
- ⊖ 특정 x 를 선형변환한 Ax 가 유일하다
- ⊖ x 와 Ax 가 서로 일대일 대응이다

선형변환으로 역행렬 이해하기

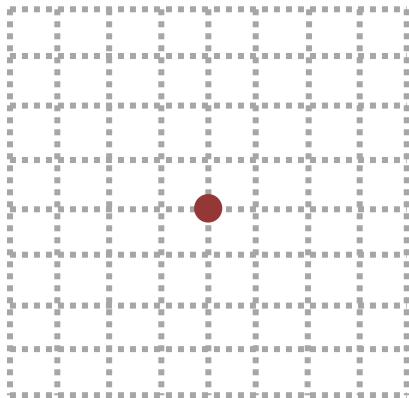
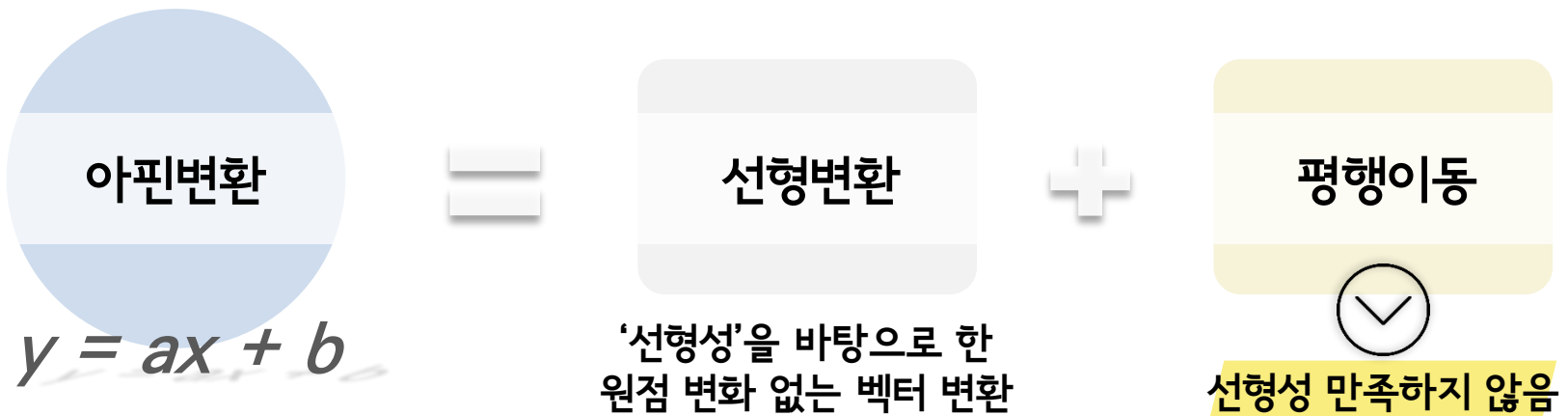
2 역행렬이 없는 경우



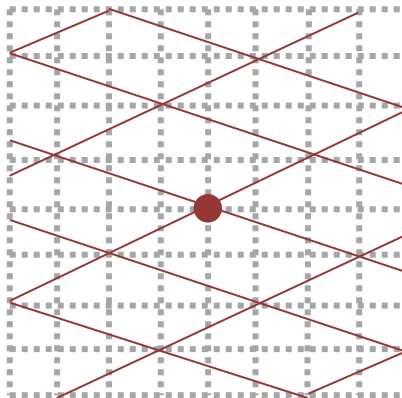
1차원 벡터를 어떠한 특정
2차원 벡터로 되돌려야 하는지
알 수 없다

- ⊖ x 와 Ax 가 서로 일대일 대응이 아니다
- ⊖ $Ax = b$ 의 해가 유일하지 않다
- ⊖ 역행렬이 존재하지 않는다

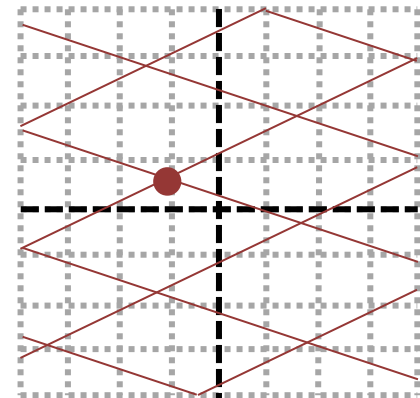
Affine transformation



변환 전

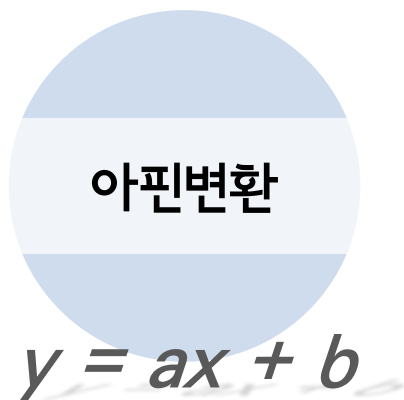


선형변환



아핀변환

Affine transformation



선형변환



평행이동



‘선형성’을 바탕으로 한
원점 변화 없는 벡터 변환

선형성 만족하지 않음

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}$$

Ax + **b** (2차원)

$$a \begin{pmatrix} 2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} c \\ d \end{pmatrix}$$

선형결합

$$\begin{pmatrix} 2 & 1 & c \\ 1 & 2 & d \end{pmatrix} \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$$

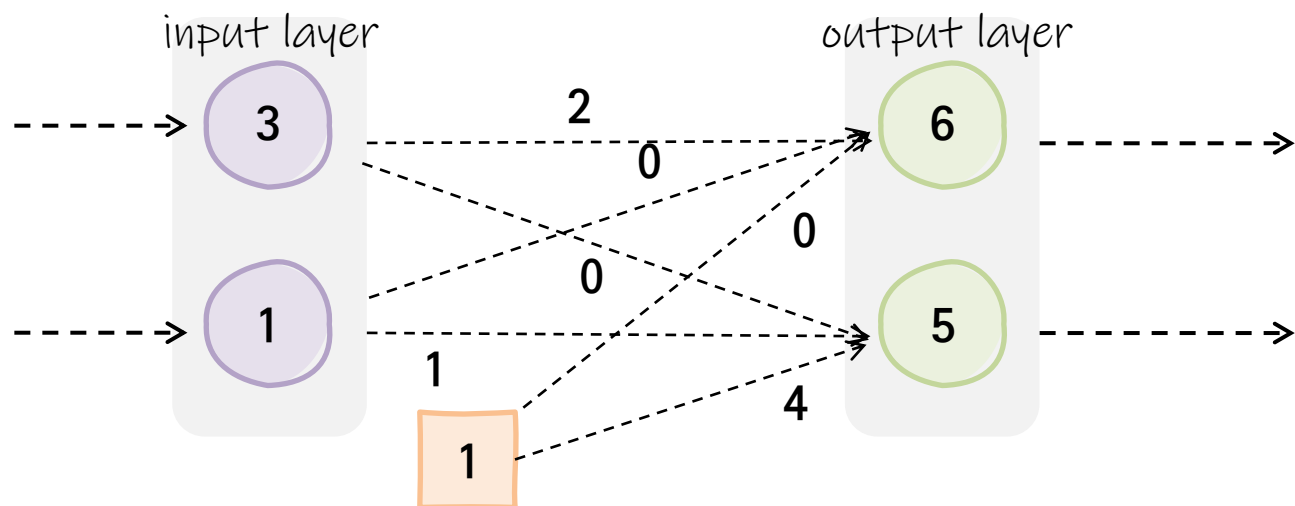
A*x* (3차원)

Affine과 딥러닝

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}$$

가중치 input layer
입력값 bias

딥러닝
모델에의
적용



선대, 딥러닝을 만나다

Affine과 딥러닝

그럼 아핀변환이 곧 딥러닝이다,
이 말인가요?



$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}$$

가중치

input layer

bias

input
vector

가중치 행렬을
통한 선형변환

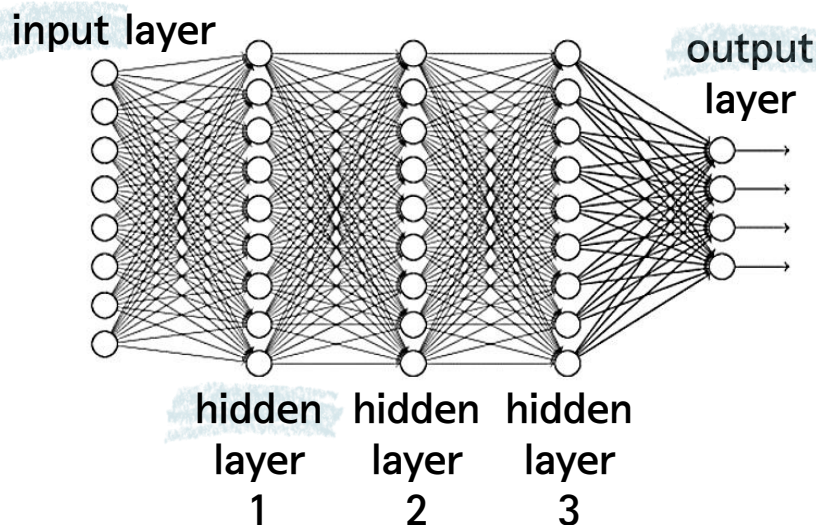
output
vector

딥러닝
모델에의
적용

딥러닝의 학습 과정 중 **bias**를 고려하는 것을
아핀변환에서의 **평행이동**처럼 해석할 수 있다!



Affine과 딥러닝



활성화 함수로 예측값 계산

손실함수로 예측과 실제의 오차 측정

손실함수를 바탕으로 가중치 업데이트

sigmoid, tanh, ...
비선형 활성화 함수를
이용하는 이유

선형함수 $f(x) = kx$ 를 이용해
n번 층을 쌓음



$k^n x$

즉, k^n 을 한 번 적용하는 것과 같아
여러 hidden layer을 쌓으며
가중치를 업데이트하는 이점이 없음