선형과 비선형

- 선형(linear)이란, 어떤 집합의 원소가 다른 원소의 선형결합으로 나타낼 수 있는 것

→ 선형 결합(linear combination) : 집합 A의 원소 $x_1, x_2, \cdots, x_n$에 상수 $a_1, a_2, \cdots, a_n$을 각각 곱해서 더하는 것 (1차 결합)

ex) $x_1 = x_2 \times a_2 + \cdots + x_n a_n$

- 이를 통해 얻어진 새로운 원소 $a_1x_1 + a_2x_2 + \cdots + a_nx_n$도 집합 A에 속함 ⇒ 집합이 선형이다.

ex) 1차식, 1차함수 ...

출력값 $y = ax + b$ 입력값 ← 고정된 계수 $= ax^1 + bx^0$

- 선형이 아닌 것을 비선형이라고 함. (Nonlinear)

선형방정식과 선형시스템

ex) 선형방정식 : $\sqrt{2}x + 3y = 1$ ← (x, y)에 대해 선형 (곱해지는 상수)

비선형방정식 : $2\sqrt{x} + 3y = 1$ ← (x, y)에 대해 비선형 * $(\sqrt{x}, y)$에 대해 선형

→ 앞에서의 선형의 정의에 맞지 않음 · x에 root가 있음

선형방정식 (Linear Equation)

- $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ (곱해지는 상수)

- 변수/미지수(Unknown variables) : $x_1, \cdots, x_n$
- 계수 (coefficients) : $a_1, \cdots, a_n$
- 해 (solution) : 방정식을 참이 되게 하는 순서쌍 $(x_1, \cdots, x_n)$

↳ 여러개, 없을수도 있음 ⇒ 위의 방정식이 참이 되게 하는 수

선형시스템(Linear System) ⇒ 선형방정식의 여러개가 모여 구성

- 변수 $x_1, \cdots, x_n$이 주어져 있을 때, 이 변수들에 대한 유한개의 선형방정식의 집합
- 선형시스템의 모든 해(순서쌍)의 집합을 해집합이라고 함 (→ 선형시스템을 참이 되게 하는 수의 집합)

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\}$$

m개의 선형방정식이 모여서 선형시스템을 이룸. (만족하는 해)

- 두 선형시스템이 똑같은 해집합을 갖는다면, 두 선형시스템은 서로 동치(Equivalent).

⇒ 위의 예시에서 a_{m1}과 b_m이 서로 다른 값을 가지는 선형시스템이어도 해집합이 같을 수 있음 (표현하는 계수들은 다르지만 동치일 수 있음)

→ Equivalent same ⇒ 본질적으로 두 선형시스템이 같다. (계수가 같은) (같 ×)

벡터와 행렬

- 벡터 (Vector) : 숫자를 순서대로 나열한 것 (순서바꿈 ×) ⇒ [1, 2, 3] ≠ [3, 2, 1]
  - 가로로 나열하면 행벡터, 세로로 나열하면 열벡터

열 4개 행 1개 행렬, 행벡터 ← ex) $\underline{X} = [1, 2, 3, 4] \in \mathbb{R}^4$ (실수 4개로 이루어진 벡터)

  - 벡터는 행이나 열이 1개인 행렬로 이해할 수 있음.
  - 원소의 개수에 따라 n차원 벡터라고 부름.

- 행렬(Matrix) : 숫자를 행과 열로 구분하여 나열한 것
  - 행벡터를 세로로 쌓은 것, 열벡터를 가로로 쌓은 것으로 이해할 수 있음.

ex) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \in \underbrace{\mathbb{R}}_{\text{실수 행렬}}{}^{3 \times 2}$ (3행 2열 행렬)

열로 쌓음
행로 쌓음

행렬의 연산

- 행렬의 덧셈 : 크기가 같은 행렬은 동일한 위치(행, 열)에 위치한 원소끼리 더할 수 있음
- 행렬의 곱셈 (AB = C) ⇒ A행렬과 B행렬의 차원이 같아야 함
  - 앞의 행렬의 i-행벡터와 뒤의 행렬의 j-열벡터를 내적하여 얻은 스칼라가 결과 행렬의 (i, j)원소가 됨
  - 앞의 행렬의 행벡터의 길이와 뒤의 행렬의 열벡터의 길이가 같아야만 곱셈 연산이 가능
  - 교환법칙이 성립하지 않음. $AB \neq BA$
- 행렬의 전치 (Transpose) (A^T)
  - 행렬의 행과 열을 교환하는 것

ex) A^T : $A \in \mathbb{R}^{m \times n} \rightarrow A^T \in \mathbb{R}^{n \times m}$

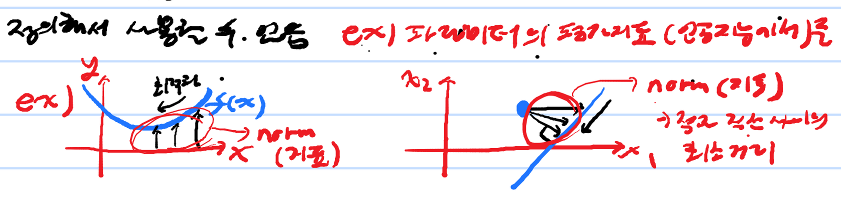
$(A^T)^T = A$, $(c \cdot A)^T = c \cdot A^T$, $(AB)^T = B^T \cdot A^T$

- 행렬의 거듭제곱 (Power)
  - 정방행렬 (square Matrix) A에 대해, A^n은 A를 n번 곱하는 것 (→ 행과 열의 길이가 같음)

ex) $A^0 = I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$, $A^k = A^{k-1} \cdot A$

노름 (Norm) ⇒ 입력된 벡터나 행렬의 넓이를 (norm으로 보는 거리로 나타내지는 모습)

- 일종의 연산으로, 벡터나 행렬의 크기를 표현할 때 사용됨 (↳ "일종의 거리")
- 벡터나 행렬을 입력으로 받아서 스칼라 값을 출력하는 함수
- 수학적으로, 특정 조건을 만족하는 함수는 norm이라고 부를 수 있으므로, 다양하게 정의해서 사용할 수 있음 ex) 파라미터의 크기제한 (인공지능에서)를 설계시 많이 사용

ex)


유클리드 노름(Euclidean norm) (일반적인 2, 3차원 공간에서의 거리 측정)

- 2차 노름, L2 norm이라고도 부름
- 2차원 벡터의 Euclidean norm은 2차원 좌표평면상에서의 벡터의 길이와 같음
- $\underline{X} = [x_1, x_2, \cdots, x_n]^T$의 Euclidean norm은 $\|\underline{X}\|$와 같이 표현 (vector, 열벡터)

ex) $\|\underline{X}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

P-norm

- Euclidean norm의 지수를 일반화한 것

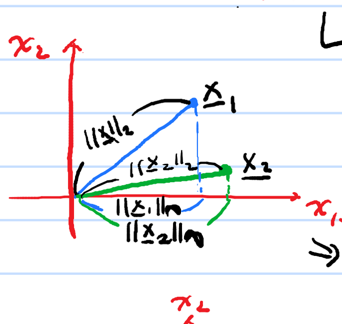
ex) $\|\underline{X}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ (P = 정수) (P = 2면 Euclidean norm)

거듭제곱 후 / p의 제곱

∞-norm (Infinite norm)

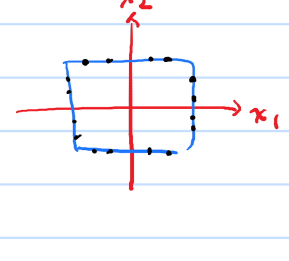
ex) $\|\underline{X}\|_\infty = \max(|x_1|, |x_2|, \cdots, |x_n|)$

↳ 각 원소의 크기 중에 가장 큰 원소를 갖는 원소의 크기 = ∞-norm

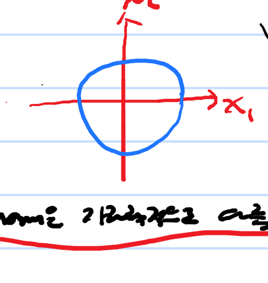


⇒ 어떤 norm을 사용하는지에 따라 크기가 달라질 수 있음.

⇒ $\|\underline{x}_1\|_2 > \|\underline{x}_2\|_2$ / $\|\underline{x}_1\|_\infty < \|\underline{x}_2\|_\infty$



⇒ 정사각형 내부의 모든 점들의 ∞-norm은 같음 (크기가 모두 같음)



⇒ 원 내부의 모든 점들의 Euclidean norm은 같음 (크기가 모두 같음)

⇒ ∞-norm과 Euclidean norm은 기하학적으로 다름

Frobenius norm

- 행렬의 노름 중 하나.

ex) $\|A\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$ where $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ (가로세로 같은 A, i행 j열의 원소)

↳ 행렬의 모든 원소들을 제곱해서 더해서 root를 씌움 ⇒ Euclidean norm의 확장