머신 러닝·딥러닝에 필요한 수학 기초 with 파이썬

Functions

함수

조준우
metamath@gmail.com

수 체계

- **자연수 $\mathbb{N}$**
  - 사물을 셀 때나 순서를 매길 때 사용하는 수 1, 2, 3, …
- **정수 $\mathbb{Z}$**
  - 자연수에 0과 음수를 더한 것으로 …, -2, -1, 0, 1, 2, … 같은 수
- **유리수 $\mathbb{Q}$**
  - 분자, 분모로 정수를 갖는 분수로 나타낼 수 있는 수
  - Rational number: 이치에 맞는 수?
- **무리수 $\mathbb{I}$**
  - Irrational number : 이치에 맞지 않는 수?
  - ratio: 비율
  - ir- 접두어로 반대의 의미 -> 비율이 없는 수 또는 비율로 나타낼 수 없는 수
  - e: 2.718…, π: 3.14…
- **실수 $\mathbb{R}$**
  - 유리수 + 무리수

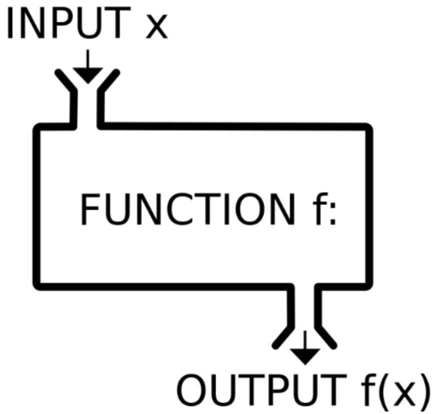
Quantities(measurement tools)

- **스칼라**Scalar :
  - Scale/skeil/ : ~을 자로 재다., 얼마만큼 자(저울)로 재다.
  - 양만으로 표시할 수 있는 물리량
  - 수직선상의 점 하나로 정의
  - 온도, 속력, 부피, 넓이
  - 숫자 하나, 1cm, 10℃, 20m^2

- **벡터**Vector
  - Vehere : 라틴어로 "운반하다", 어떤 물체를 얼마만큼 어디로 운반하다.
  - 양과 방향으로 표현할 수 있는 물리량
  - 공간의 두 점 또는 원점과 한 점으로 정의
  - 속도, 힘
  - 숫자 여러 개, $\mathbf{v} = (1,2,3)^T$
  - 볼드체 소문자로 표시

함수란?

- **함수** : 첫 번째 집합의 임의의 한 원소를 두 번째 집합의 오직 한 원소에 대응시키는 대응 관계
- **독립변수** : 입력되는 변수
- **종속변수** : 출력되는 변수
- 입력되는 독립변수가 **어떤 형태로든 연산**이 되어 출력으로 나온다.

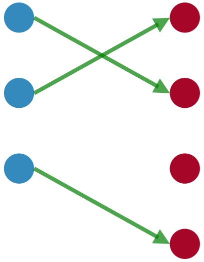


함수의 표기

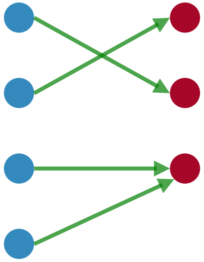
- **표기**
  - $f : X \rightarrow Y$, X : 정의역domain, Y : 공역codomain
  - $f(X) = Y$
  - 실수 집합 : $\mathbb{R}$로 쓸 때 $f : \mathbb{R} \rightarrow \mathbb{R}$는 실수를 실수로 대응시키는 함수, 예 : $f : x \rightarrow x^2$
  - $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) = x_1 + x_2$

함수 대응관계의 특징

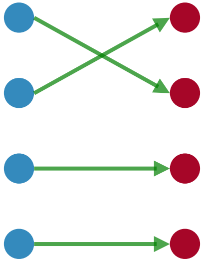
- **성질** : 사랑의 작대기



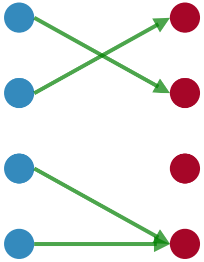
일대일 함수
(단사함수 Injection)



전사 함수
(Surjection)



일대일 대응 함수
(전단사함수 Bijection)

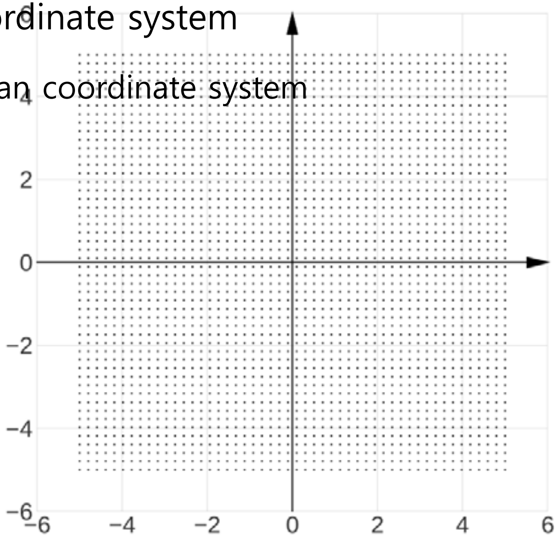


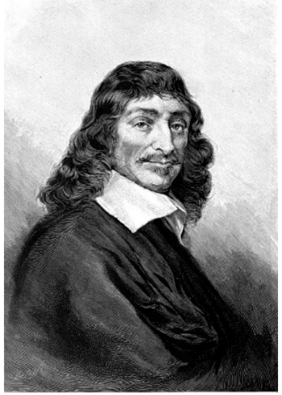
일반 함수
(Function)

함수의 표현: 좌표계

- 좌표계 coordinate system
  - 공간에 존재하는 대상을 고유한 숫자로 나타내기 위해 사용하는 시스템

- 직교 좌표계 rectangular coordinate system
  - 데카르트 좌표계 cartesian coordinate system

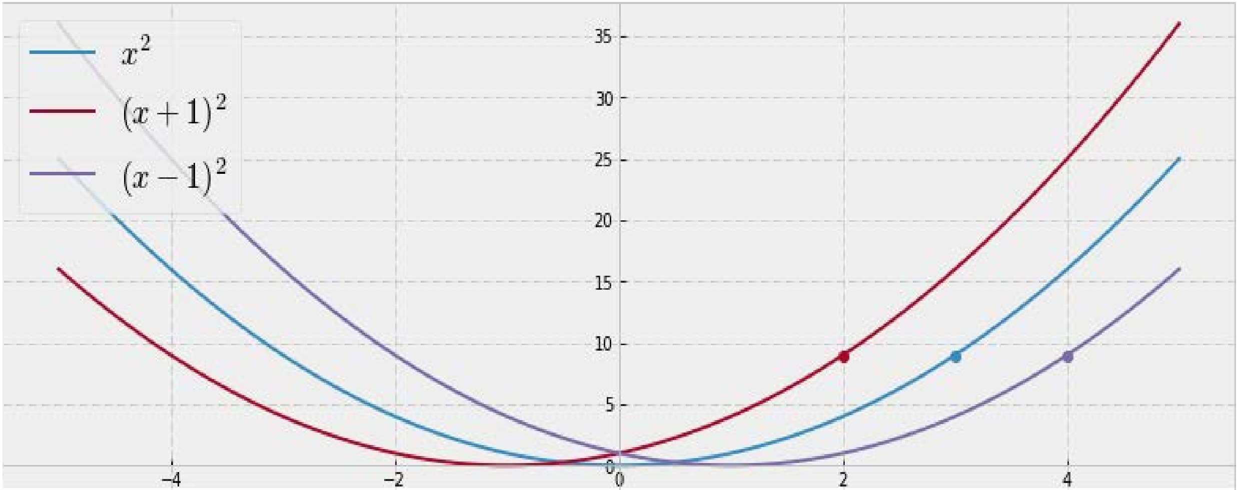




데카르트RENÉ DESCARTES(1596~1650)

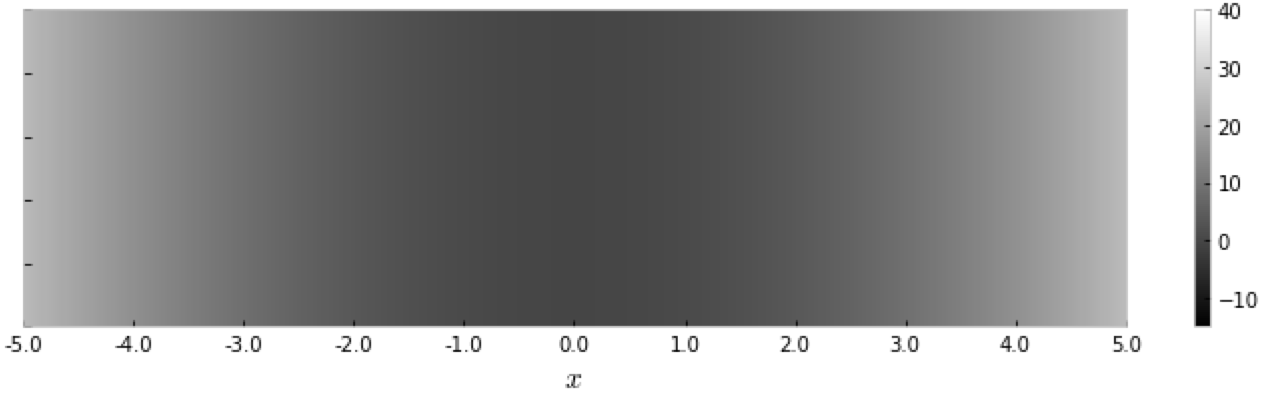
함수의 그래프

- 독립변수와 종속변수의 관계를 그림으로 표현
- x축에서 하나의 값을 선택 → 함수 → 출력
- 출력= y 인 수평선과 입력의 수직선과 교차점을 찍어 그린 것
- 파란 점은 $y = x^2$인 관계에서 $y = 9$와 같은 x 가 이루는 위치에 찍힌 점
- 그래프의 이동은 이 관계를 생각하면 이해 가능



2차 함수는 포물선?

- 꼭 포물선이어야 할까?
- 데카르트의 유산을 따르지 않는다면?
- 밝기로 나타낸 $y = x^2$



합의 기호 $\sum$

- 수열 : 정의역이 자연수 전체 집합 $\mathbb{N}$이고 공역이 실수 전체 집합 $\mathbb{R}$인 함수

$$f: \mathbb{N} \to \mathbb{R}$$

- 수열 2, 4, 6, 8, …는 $f(1) = 2$, $f(2) = 4$, $f(3) = 6$처럼 숫자의 위치, 출력이 주어진 숫자인 함수
- 수열의 첫째 항부터 제 n항 까지의 합

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + a_3 + \cdots + a_n$$

합의 기호 $\sum$의 성질

$$\sum_{k=1}^{n}(a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$$

$$\sum_{k=1}^{n}(a_k - b_k) = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n} b_k$$

$$\sum_{k=1}^{n} ca_k = c\sum_{k=1}^{n} a_k$$

$$\sum_{k=1}^{n} c = cn$$

거듭제곱

- 주어진 수를 n번 곱하는 연산

$$x^n = \underbrace{x \times x \times \cdots \times x}_{n}$$

다항함수

- **단항식** : 2, x, x2, x/3등 상수 또는 변수의 항 하나로 이루어진 식
- **다항식 :** 단항식 여러 개의 더하기로 연결된 식
  - $x^2 + 3x$
  - 다항식의 차수는 가장 높은 차수

차수

$$x^3 + 2x^2 + x + 4$$

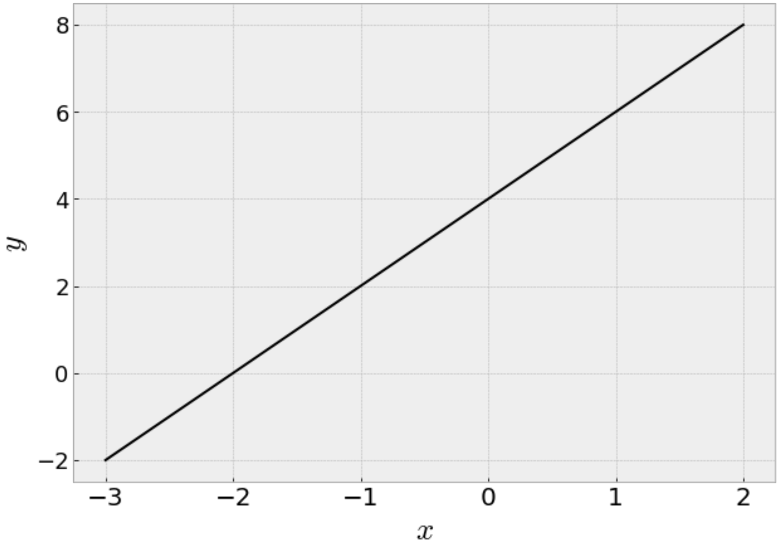
계수 상수항

- **다항함수** : 다항식으로 구성된 함수
  - $f(x) = x^2 + 2x + 3$

그래프 그리기

- **matplotlib** : Numpy 배열을 기반으로 만들어진 데이터 시각화 라이브러리
- 기본 플로팅
  - x, y 값을 가지고 2차원 평면에 플롯
  - plot(x, y, options...) 형식

```
import numpy
import matplotlib.pyplot as plt

fig = plt.figure(figsize=(10,7))
ax = fig.add_subplot(1, 1, 1)

x = np.linspace(-3, 2, 10)
y = 2*x+4
ax.plot(x, y, 'k')

plt.show()
```

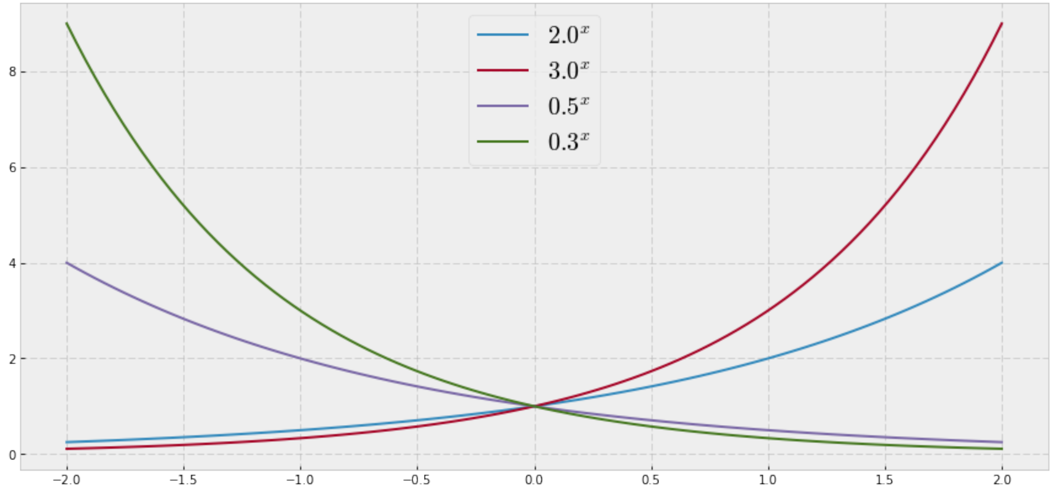


(거듭)제곱근

- **제곱근** : 어떤 수의 제곱근square root은 제곱하여 그 수가 되는 수https://ko.wikipedia.org/wiki/제곱근
- 표기 : 제곱하여 양수 a가 되는 실수 $\pm\sqrt{a}$
- 9의 제곱근 : $\pm\sqrt{9} = \pm 3$
- 제곱근 9 : $\sqrt{9} = 3$
- **거듭제곱근** : 제곱근과 마찬가지로 어떤 수의 n거듭제곱근은 n번 거듭제곱하여 그 수가 되는 수
- 방정식 형태로 나타내면 $x^n = a$를 만족시키는 x를 a의 n제곱근이라 한다.
- 표기 : n번 거듭제곱하여 양수 a가 되는 실수 $\pm\sqrt[n]{a}$ (n이 짝수), $\sqrt[n]{a}$ (n이 홀수)

지수함수

- **정의** : $y = a^x,\ a > 0, a \neq 1$
- **그래프**
  - $a > 1$: 양의 방향으로 갈 수록 증가
  - $0 < a < 1$: 양의 방향으로 갈 수록 감소



지수법칙

몇 가지 지수법칙

a, b : 실수 m, n : 양의 정수

1) $a^m a^n = a^{m+n}$ 2) $a^m \div a^n = a^{m-n}$

3) $(a^m)^n = a^{mn}$ 4) $(ab)^m = a^m b^m$

$a \neq 0$ m : 양의 정수

5) $a^0 = 1$ $\because \dfrac{a^m}{a^m} = a^{m-m} = a^0 = 1$

6) $a^{-m} = \dfrac{1}{a^m}$ $\because a^m a^{-m} = a^{m+(-m)} = a^0 = 1$

$a > 0$ $m, n(n > 0)$: 정수

7) $a^{\frac{m}{n}} = \sqrt[n]{a^m}$

$\because \left(a^{\frac{m}{n}}\right)^n = a^{\frac{m}{n}n} = a^m$

$\left(\sqrt[n]{a^m}\right)^n = a^m$

복리 이자

- 특정 기간동안 2배가 불어나는 예금에서 인출과 입금을 반복

베르누이JAKOB BERNOULLI(1655~1705)

1

$$1 + 1 = 2 = \left(1 + \frac{1}{1}\right)^1$$

1 | 1 | 1

| 1/2 | 1/2

| | 1/2

| | $(1/2)^2$

$$1^2 + 2(1/2) + (1/2)^2$$

$$= \left(1 + \frac{1}{2}\right)^2 = 2.25$$

1 | 1 | $1, 1/3$ | $1, \frac{1}{3}, \frac{1}{3}, \left(\frac{1}{3}\right)^2,$

| 1/3 | $1/3, (1/3)^2$ | $\frac{1}{3}, \left(\frac{1}{3}\right)^2, \left(\frac{1}{3}\right)^2, \left(\frac{1}{3}\right)^3$

$$\left(1 + \frac{1}{3}\right)^3 \approx 2.37$$

기간

이율

$$\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = ?$$

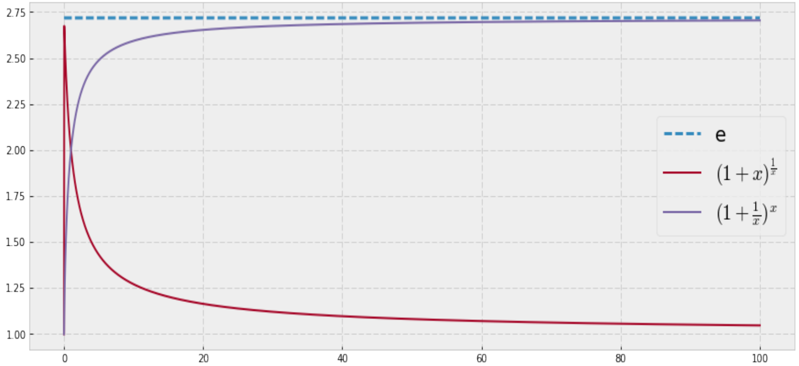
자연상수 e

- 야코프 베르누이에 의해 아래식이 수렴함이 발견
- 라이프니츠에 의해 상수 b로 처음 사용
- 오일러에 의해 e로 처음 표기

오일러LEONHARD EULER(1707~1783)

$$\lim_{x \to 0}(1+x)^{\frac{1}{x}} = e = 2.718$$

$$\lim_{x \to \infty}\left(1+\frac{1}{x}\right)^{x} = e = 2.718$$



- e의 지수함수 표기

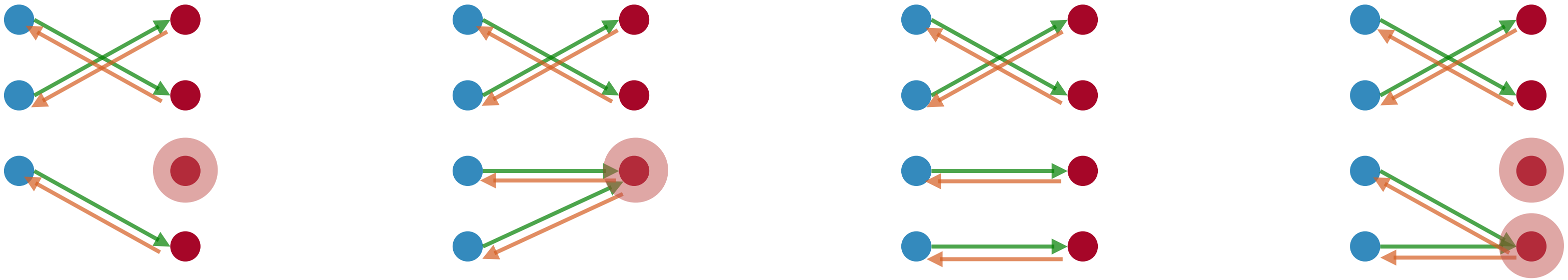
$$f(x) = e^{x} = \exp(x)$$

역함수Inverse func.

- **정의** : $f: X \to Y$일 때 $g: Y \to X$인 함수가 있어서 $f(x) = y$일 때 $g(y) = x$를 만족하는 함수
- **표기법** : $f^{-1}: Y \to X,\ x = f^{-1}(y)$
- 함수와 그의 역함수의 그래프는 $y = x$에 대칭
- 지수함수와 로그함수 그래프로 확인

역함수Inverse func.

- **존재** : 일대일 대응에서만 존재



일대일 함수 (단사함수 Injection) | 전사 함수 (Surjection) | 일대일 대응 함수 (전단사함수 Bijection) | 일반 함수 (Function)

화씨온도→ 섭씨온도

$$C = \frac{5}{9}(F - 32)$$

섭씨온도→ 화씨온도

$$F = \frac{9}{5}C + 32$$

로그함수

- 로그 정의

$$\log_a x = c$$

- 로그는 **숫자**
- a를 x로 만들기 위해 a의 어깨 위에 거듭 제곱 되어야 하는 **숫자**

어떤 숫자를 올려야 하나?

$$a^c = x$$

이 숫자를

이 숫자로 만들고 싶은데

로그함수

- 로그 정의대로 쓰면

$$a^c = x$$

$$\log_a x = c$$

$$a^{\log_a x} = x$$

- e가 a자리에 오면 자연로그라 하고 $\ln x$로 표기, 그냥 $\log x$로 쓰기도 함
- x를 바꿔가면서 함수처럼 생각

로그함수

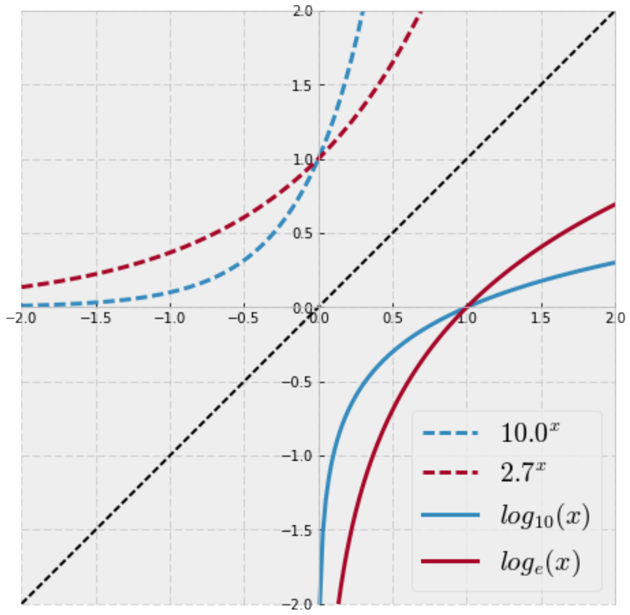
- **그래프** : 로그함수 - 지수함수 역함수 관계

독립변수 종속변수

$$a^c = x$$

$$\log_a x = c$$

독립변수 종속변수



로그성질

- **성질** : 정의를 잘 이용하면 다음을 보일 수 있다.

- $\log_a 1 = 0,\ \log_a a = 1,\ a^{\log_a 1} = 1,\ a^{\log_a a} = a$
- $\log_a MN = \log_a M + \log_a N$
  - 유도 $\log_a MN$은 a를 MN으로 만들기 위해 어깨 위에 올라가는 수

$$a^{\log_a MN} = MN$$

$$a^{\log_a M} = M \qquad a^{\log_a N} = N$$

$$a^{\log_a M} \times a^{\log_a N} = MN$$

$$a^{\log_a M + \log_a N} = MN$$

NOTE

$$\log_a M^k = k \log_a M$$

로그성질

- $\log_a \frac{M}{N} = \log_a M - \log_a N$
  - 유도 로그 정의대로 써주면

$$a^{\log_a \frac{M}{N}} = \frac{M}{N}$$

$$a^{\log_a M} = M \qquad a^{\log_a N} = N$$

$$\frac{a^{\log_a M}}{a^{\log_a N}} = \frac{M}{N}$$

$$a^{\log_a M - \log_a N} = \frac{M}{N}$$

로그성질

- $\log_a b = \dfrac{\log_c b}{\log_c a}(a > 0, a \neq 1\,)$
  - 유도 로그 정의대로 써주면

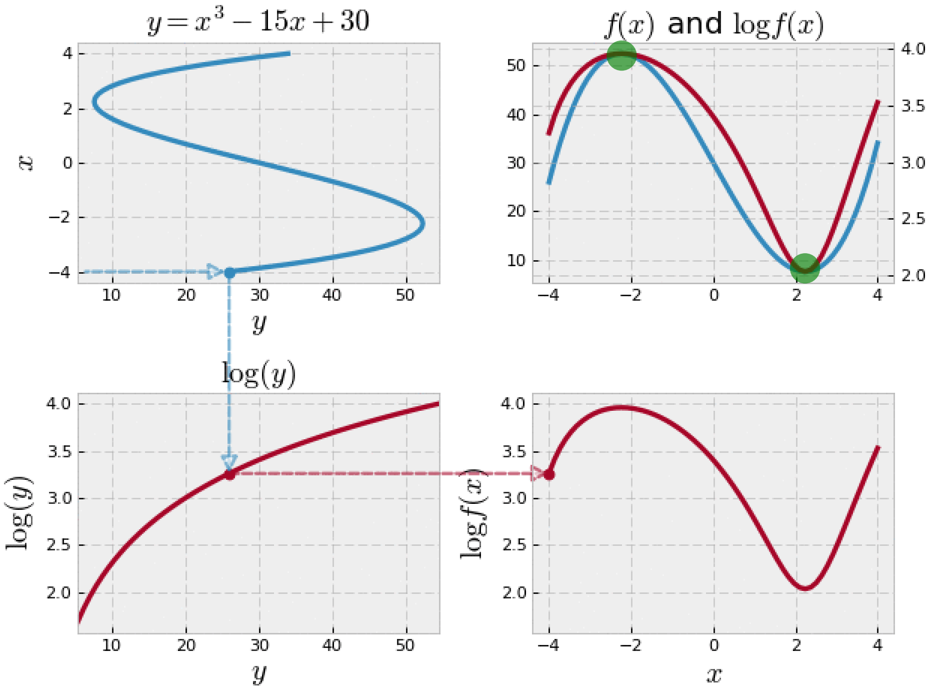
$$a^{\log_a b} = b \qquad \log_c(a^{\log_a b}) = \log_c b$$

$$\log_a b \log_c a = \log_c b \qquad \because \log_a M^k = k\log_a M$$

$$\log_a b = \frac{\log_c b}{\log_c a} \qquad \because \log_c a \neq 0$$

로그성질

- 볼록한 지점(극점)의 위치를 변화시키지 않는다.

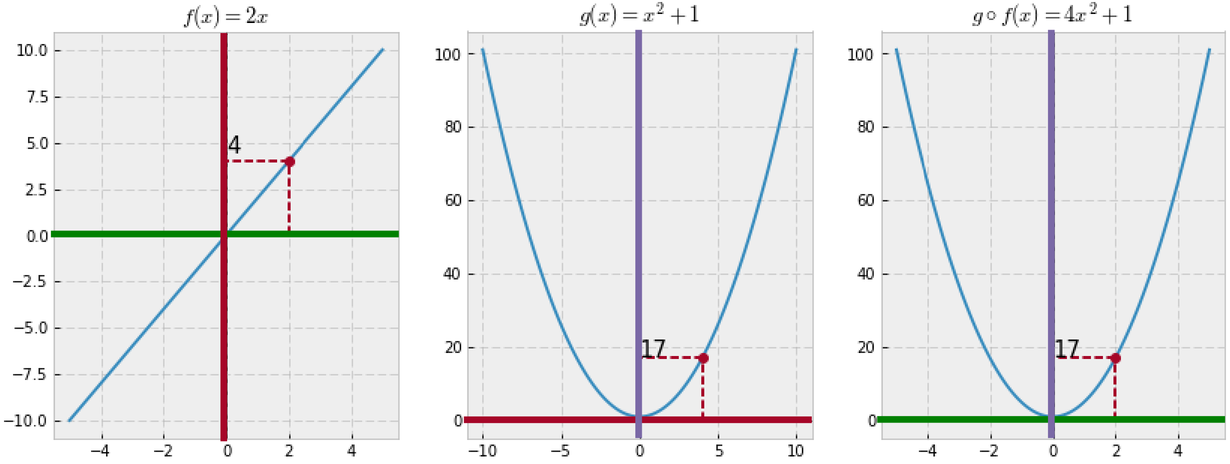


함수의 합성Composite Function

- **정의** : 함수 $f: X \to Y$의 공역과 함수 $g: Y \to Z$의 정의역이 같다고 할 때 다음과 같이 정의된 함수 $g \circ f$를 두 함수 f와 g의 합성이라고 한다.

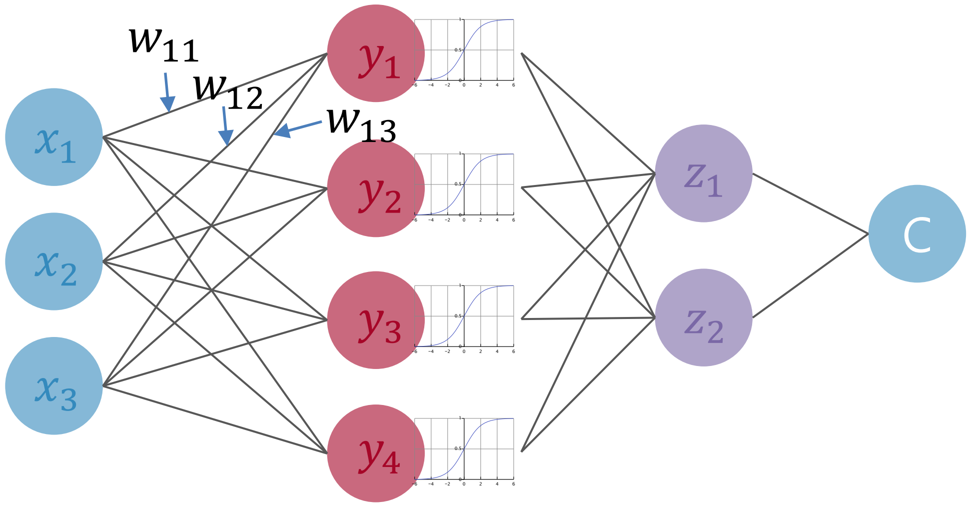
- **표기법** : $g(f(x))$: 직관적, $g \circ f$: 덜 직관적

- **입출력 관점** : f의 출력이 g의 입력으로 들어감



인공신경망은 합성함수

- 신경망은 매우 많은 함수가 다음처럼 겹겹이 합성된 것이라 할 수 있다.



$$f_1: \mathbb{R}^3 \to \mathbb{R}^4 \qquad f_2: \mathbb{R}^4 \to \mathbb{R}^2 \qquad f_3: \mathbb{R}^2 \to \mathbb{R}$$

$$f: \mathbb{R}^3 \longrightarrow \mathbb{R}$$

함수의 종류

- 일반적인 함수의 종류
  - 다항함수, 분수함수, 지수함수, 로그함수, 삼각함수, ……
  - $f(x) = 3x + 2$, $f(x) = \frac{3x^2+2}{x+1}$, $f(x) = 2^x$, $f(x) = \log 3x$
- 이 수업에서 함수의 종류
  - 입력과 출력의 형태에 따른 분류

일변수-실함수
univariable scalar func.

다변수-실함수
multivariable scalar func.

일변수-벡터함수
univariable vector func.

다변수-벡터함수
multivariable vector func.

일변수-실함수Univariable scalar func.

- $y = f(x)$, $f: \mathbb{R} \rightarrow \mathbb{R}$

- 고등학교에서 많이 봤던 우리가 익히 알고 있는 함수
  - **다항함수**, 분수함수, **지수함수**, **로그함수**, 삼각함수

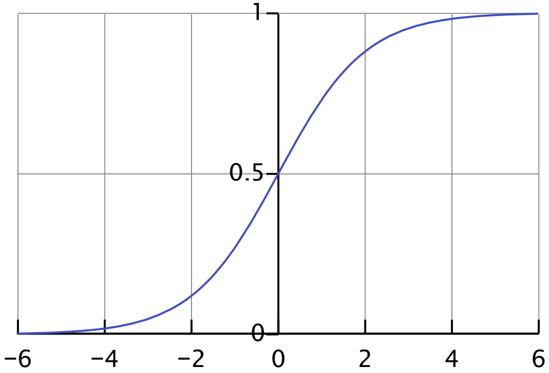
- **입력**: 스칼라, **출력** : 스칼라
  - 즉, 입력이 숫자 하나가 들어가고 출력으로 숫자 하나가 나온다.

- $f(x) = x^2$, $f(x) = 2^x$, $f(x) = \log 3x$

로지스틱 함수Logistic func.

- 신경망에서 뉴런의 활성을 결정하는 활성함수로 사용

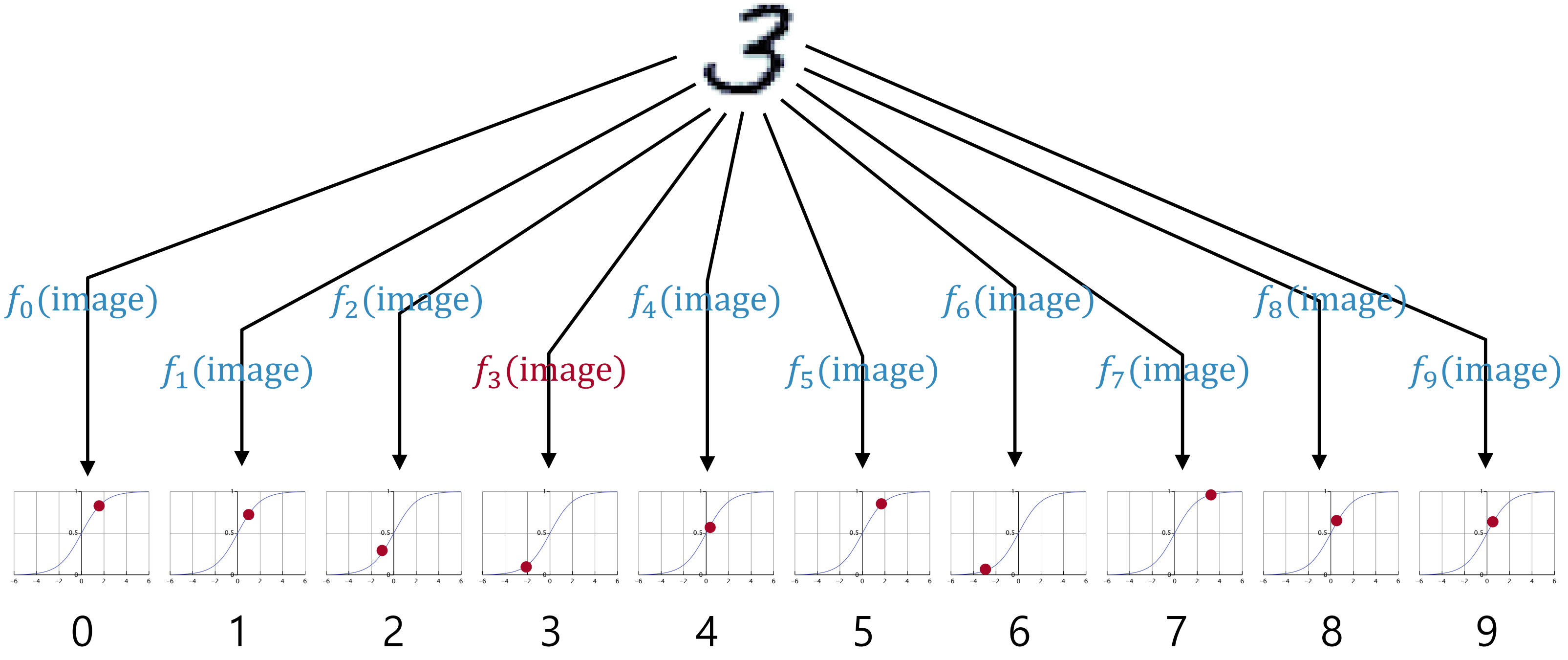
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



By Qef (Created from scratch with gnuplot) [Public domain], via Wikimedia Commons

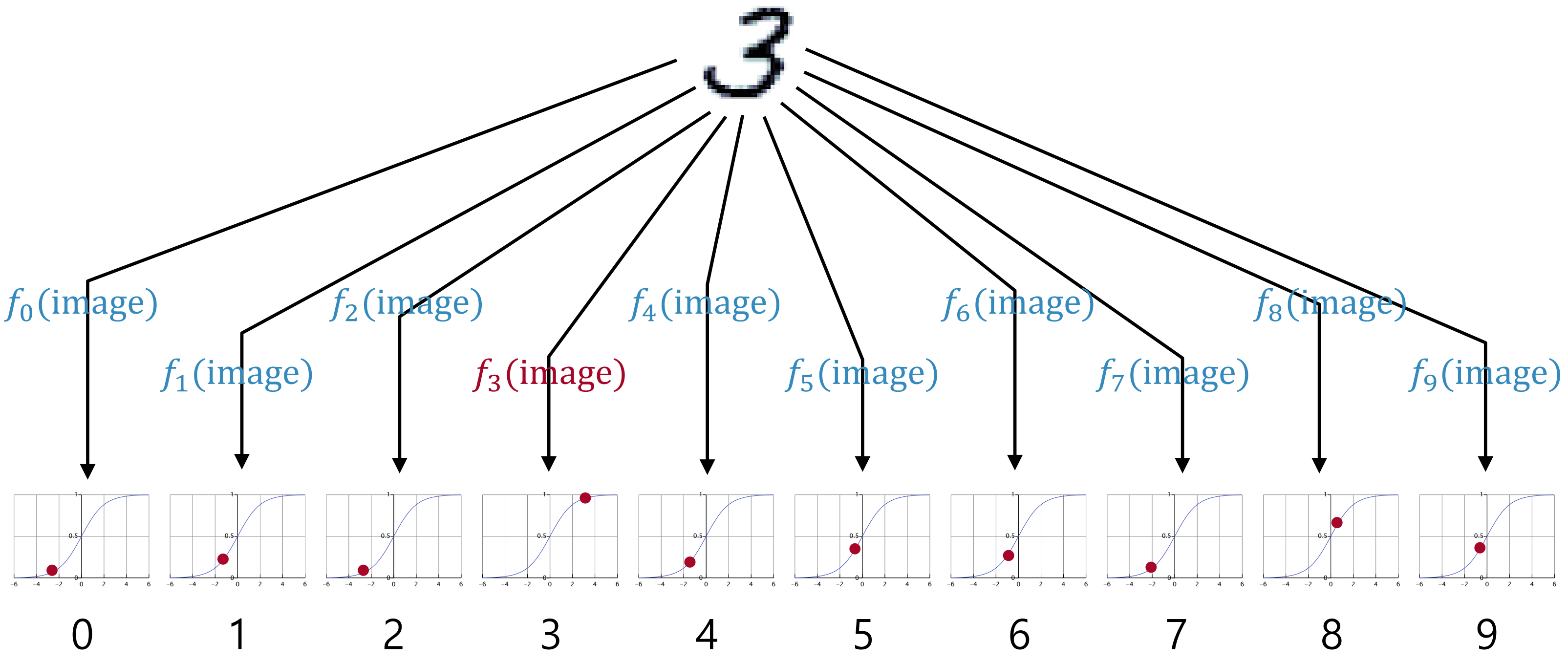
로지스틱 함수의 활용

- BAD : $f_3(\text{image})$함수가 1에 가까운 값을 출력해야 함



로지스틱 함수의 활용

- GOOD : $f_3(\text{image}) \approx 1,\ f_{i\neq 3}(\text{image}) \approx 0$



지수, 로그함수 활용

- 머신러닝 분류 문제에 있어서 평가 함수로 사용

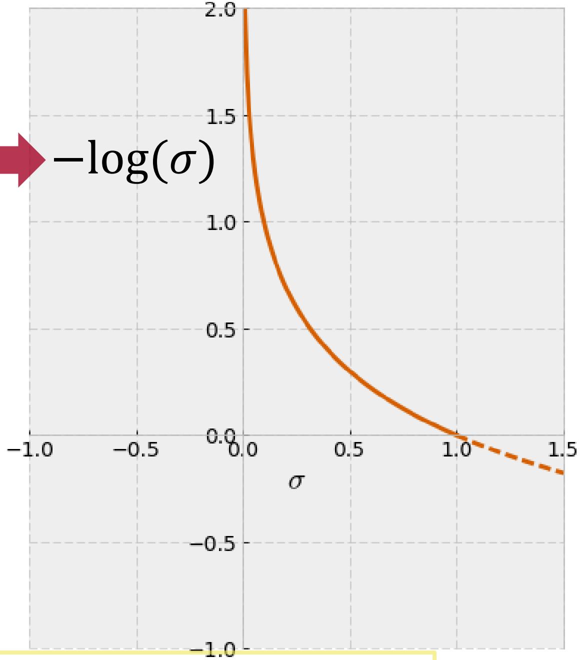
0

1

$\mathbf{x} \in \mathbb{R}^n$

$z: \mathbb{R}^n \to \mathbb{R}$ $\sigma(z) \in (0,1)$ $-\log(\sigma)$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



JS metamath1.github.io/noviceml/toyclassifier.html

다변수-실함수Multivariable scalar func.

- $z = f(x, y)$, $f: \mathbb{R}^2 \to \mathbb{R}$

- 스칼라 장을 정의
  - 공간에서의 온도장, 대기의 기압장

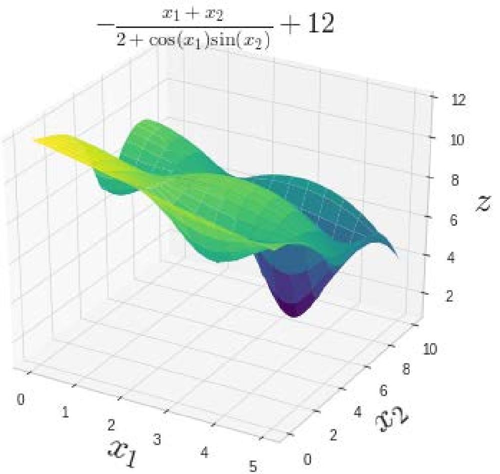
- **입력**: 벡터, **출력** : 스칼라
  - 즉, 입력이 숫자 여러 개가 들어가고 출력으로 숫자 하나가 나온다.

- $f(\boldsymbol{p}; \mathbf{p_0}) = f(x, y, z;\ x_0, y_0, z_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$

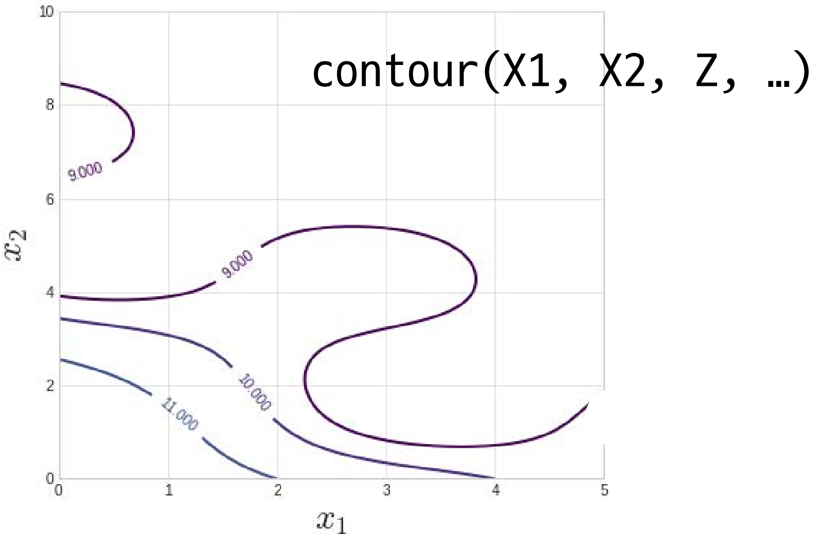
- $C(\mathbf{w}) = \frac{1}{2N} \sum_{i=1}^{N} \|y(\mathbf{w}; \mathbf{x}_i) - t\|^2$
- **수업에서 중요하게 다루는 점** : 다변수-실함수의 편미분

다변수-실함수Multivariable scalar func.

- 아래 그래프에서 왼쪽 그래프는 (x_1, x_2)로 계산된$f(x_1, x_2)$를 수직축(높이값)에 그린 것
- 오른쪽 그래프는 왼쪽 그래프에서 특정 높이를 만족하는 값만을 모아서 x_1x_2평면에 그린 것
- 때문에 등고선이 되며 선에 함수값이 적혀 있음



plot_surface(X1, X2, Z, …)

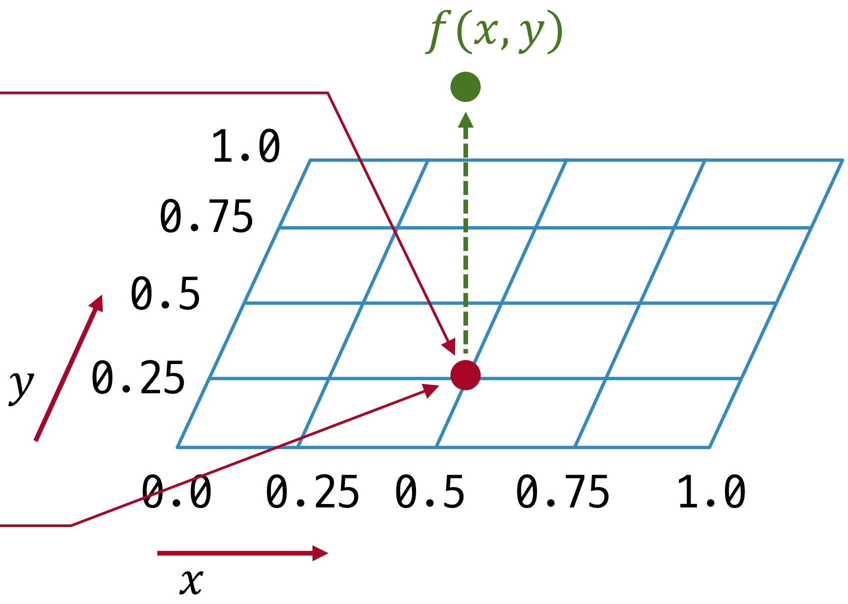


등고선 그리기|Plot Contours

- 2변수 실함수 그래프 그리기. $z = f(x, y)$, $f: \mathbb{R}^2 \to \mathbb{R}$
- numpy.meshgrid을 이용해서 xy평면에 적절히 그리드를 생성

function.ipynb

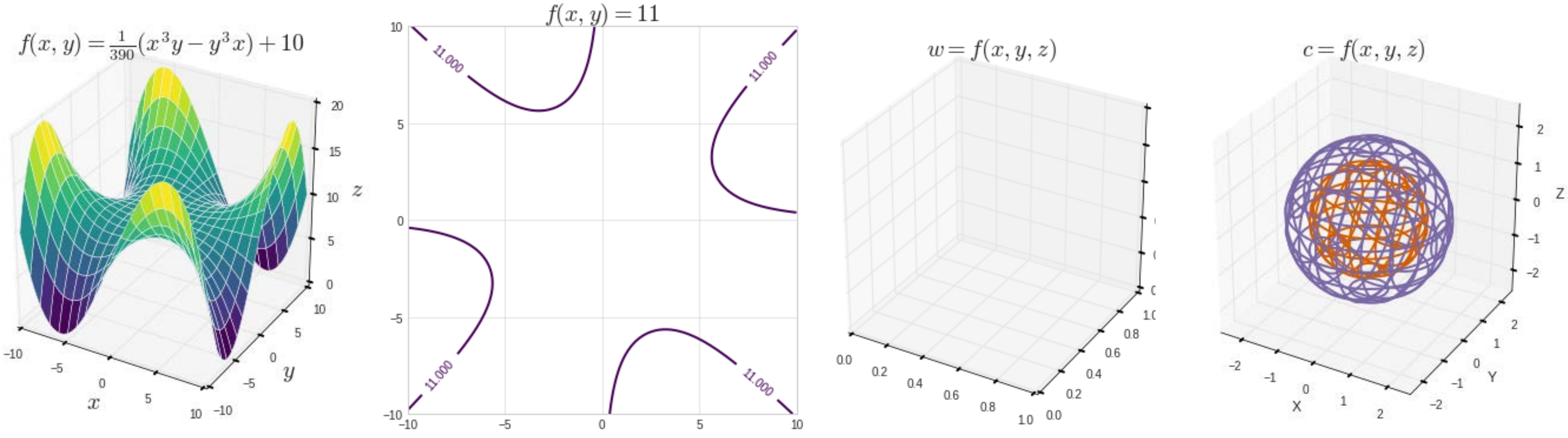
```python
x = np.linspace(0, 1, 5)
y = np.linspace(0, 1, 5)
X, Y = np.meshgrid(x, y)

print(X)
[[ 0.    0.25  0.5   0.75  1.  ]
 [ 0.    0.25  0.5   0.75  1.  ]
 [ 0.    0.25  0.5   0.75  1.  ]
 [ 0.    0.25  0.5   0.75  1.  ]
 [ 0.    0.25  0.5   0.75  1.  ]]

print(Y)
[[ 0.    0.    0.    0.    0.  ]
 [ 0.25  0.25  0.25  0.25  0.25]
 [ 0.5   0.5   0.5   0.5   0.5 ]
 [ 0.75  0.75  0.75  0.75  0.75]
 [ 1.    1.    1.    1.    1.  ]]
```



한쪽 변이 숫자로 고정? $f(x, y) = c$

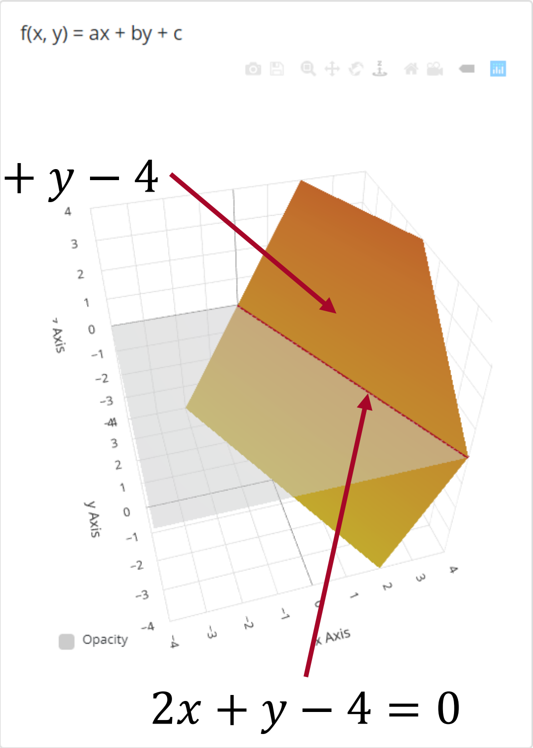
- 2변수 실함수 $z = f(x, y)$, $f(x, y) = const.$
- 3변수 실함수 $w = f(x, y, z)$, $f(x, y, z) = const.$
- 3변수 실함수의 경우 $w = f(x, y, z)$ 형태는 그래프로 그릴 수 없음

JS metamath1.github.io/noviceml/contour.html



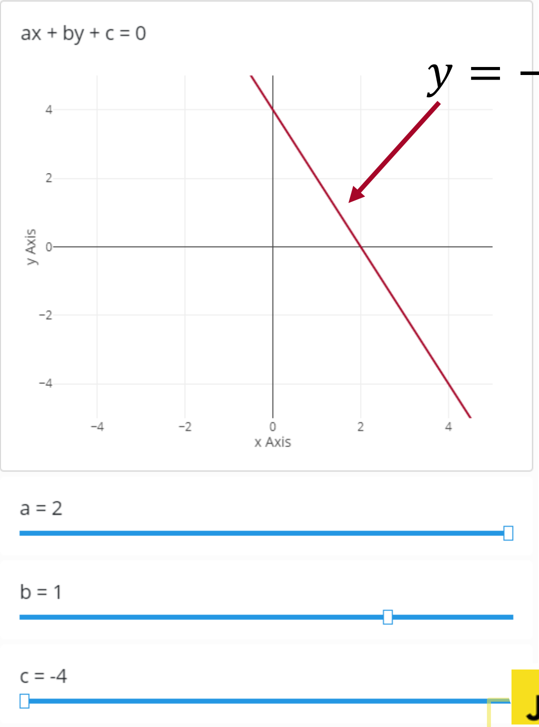
평면과 직선의 방정식

$$f(x, y) = 2x + y - 4$$



$$2x + y - 4 = 0$$

$$y = -2x + 4$$



JS plane.html

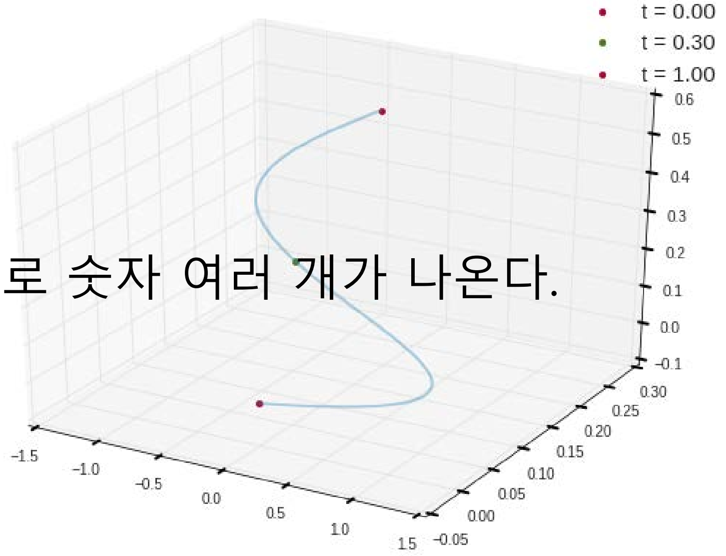
일변수-벡터함수Univariable vector func.

- $\mathbf{r}(t) = (x(t), y(t), z(t))$

- 평면 또는 공간에 존재하는 곡선
  - 시간에 따른 물체의 이동 경로

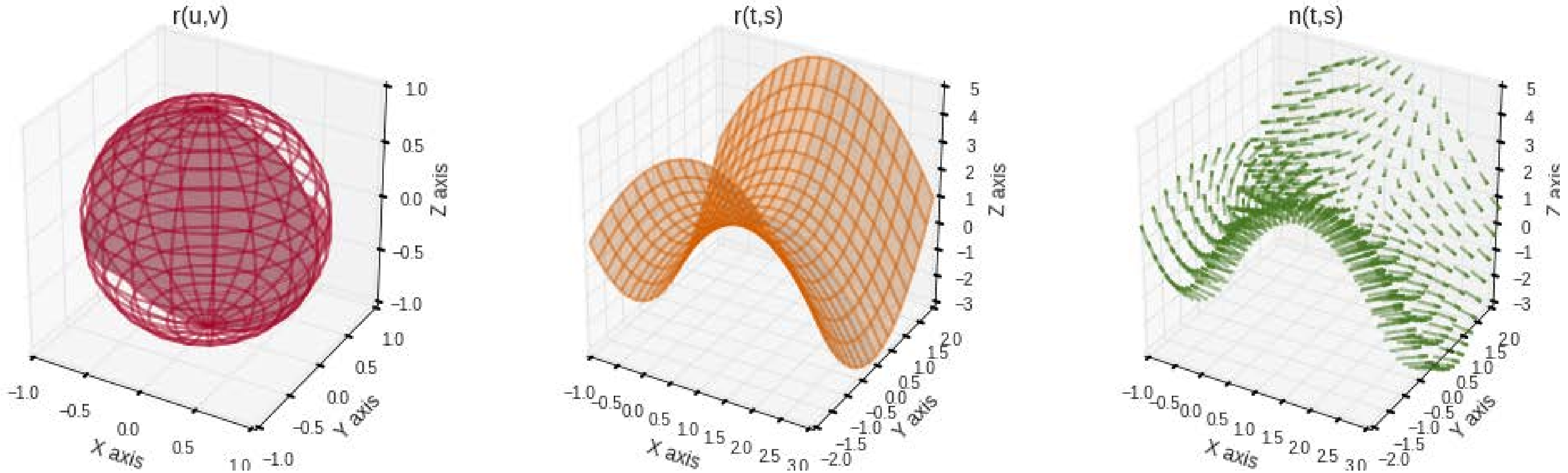
- **입력**: 스칼라, **출력** : 벡터
  - 즉, 입력이 숫자 하나가 들어가고 출력으로 숫자 여러 개가 나온다.

- $f(x(t), y(t), z(t)) = \left(\sin(6t), \frac{1}{4}t, \frac{t^2}{2}\right)$



다변수-벡터함수Multivariable vector func.

- $\mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v))$ $\mathbb{R}^2 \to \mathbb{R}^3$
- 공간에 존재하는 곡면
- **입력**: 벡터, **출력** : 벡터
  - 즉, 입력이 숫자 여러 개가 들어가고 출력으로 숫자 여러 개가 나온다.
- $\mathbf{r}(u, v) = (\cos u \sin v\,, \sin u \sin v\,, \cos v)^T$



다변수-벡터함수의 활용

- $softmax: \mathbb{R}^K \rightarrow [0,1]^K$: 실수 k개가 0에서 1사이의 숫자 k개로 매핑

Scores

$\boldsymbol{z}$

Softmax

Probabilities

$\boldsymbol{s}$

$$softmax_j(\mathbf{z}) = \frac{\exp(z_j)}{\sum_{k=1}^{K} \exp(z_i)}$$

Scores $\boldsymbol{z}$		Probabilities $\boldsymbol{s}$
1.0	→	0.0641
2.5	→	0.2875
3.0	→	0.474
2.0	→	0.1744

\+

1.0

다변수-벡터함수의 활용

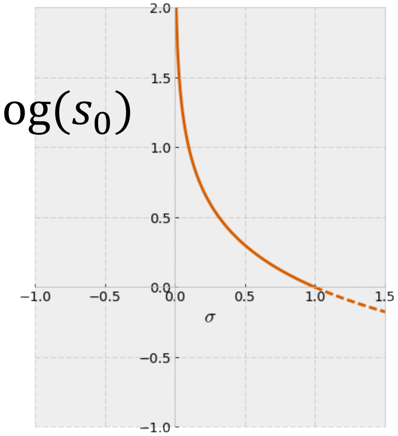
- $softmax: \mathbb{R}^K \rightarrow [0,1]^K$: 실수 k개가 0에서 1사이의 숫자 k개로 매핑

0

1

2

$z_0: \mathbb{R}^n \rightarrow \mathbb{R}$ → s_0 → $-\log(s_0)$



$z_1: \mathbb{R}^n \rightarrow \mathbb{R}$ → s_1

$z_2: \mathbb{R}^n \rightarrow \mathbb{R}$ → s_2

$\mathbf{x} \in \mathbb{R}^n$

$$s_j = softmax(\mathbf{z}) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$