1. 함수, 미적분Function, Calculus - 1/3

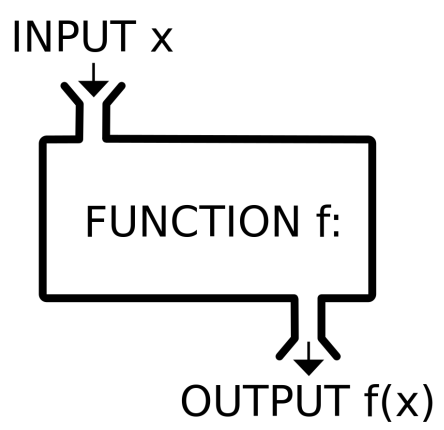
함수

스칼라scalar , 벡터vector[1]

- 스칼라
  - 양만으로 표현할 수 있는 물리량, 수학적으로는 숫자 하나
  - 예 : 온도
- 벡터
  - 양과 방향으로 표현할 수 있는 물리량
  - 예 : 속도
  - 3차원 벡터의 예 : $\mathbf{v} = (2, 3, 4)^{\mathrm{T}}$
  - 머신러닝에서는 n차원 벡터를 다룸

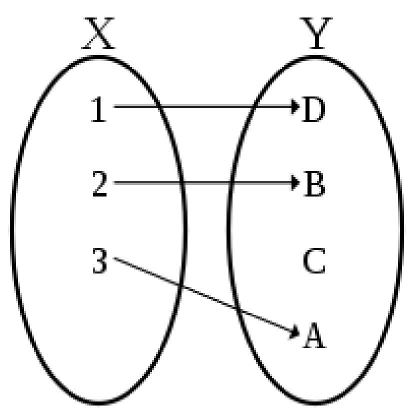
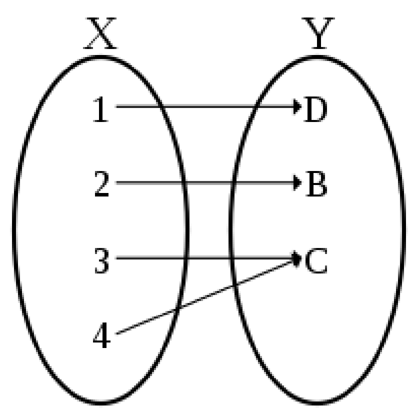
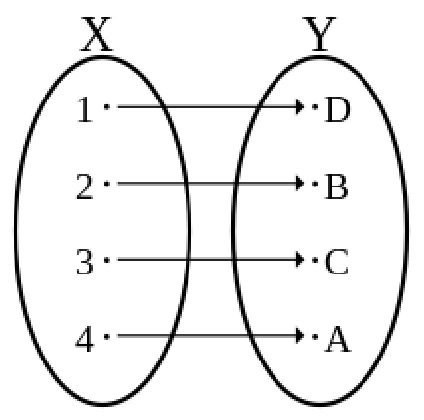
함수의 정의

- 첫 번째 집합의 임의의 한 원소를 두 번째 집합의 오직 한 원소에 대응시키는 대응 관계[2]
- 표기 1. $f : X \to Y$, X:정의역$^{\text{domain}}$, Y:공역$^{\text{codomain}}$
- 표기 2. $f(X) = Y$
- 실수 집합 $\mathbb{R}$로 쓸 때 $f : \mathbb{R} \to \mathbb{R}$는 실수를 실수로 대응시키는 함수 예: $f : x \to x^2$
- $f : \mathbb{R}^2 \to \mathbb{R}$, $f : (x_1, x_2) \to x_1 + x_2$
- 입력되는 변수 : 독립변수, 출력되는 변수 : 종속변수



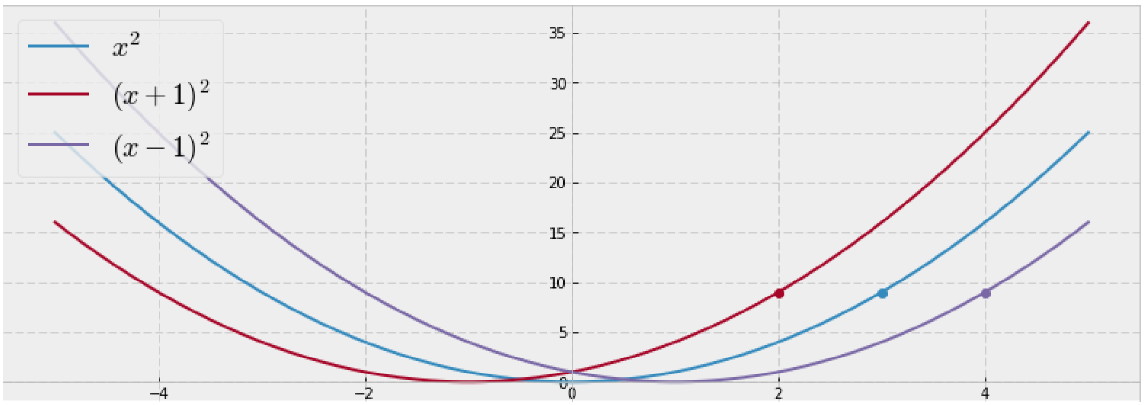
https://en.wikipedia.org/wiki/Function_(mathematics)#/media/File:Function_machine2.svg

- 사랑의 작대기(?) 와 같은 성질

		
일대일 함수(단사함수)	전사함수	일대일 대응함수(전단사함수)
https://ko.wikipedia.org/wiki/함수 public domain		

함수의 그래프

- 독립변수와 종속변수의 관계를 그림으로 표현
- 일반적으로 3변수 음함수 표현까지만 가능



- 위 그래프는 x축에서 하나의 값을 선택했을 때 그 선택된 x^*의 제곱한 값과 같은 y값에 점을 찍어 그린것이다.
- 파란점은 $y = x^2$인 관계에서 $y = 9$와 같아지는 x가 이루는 위치에 찍힌 점
- $y = (x+1)^2$인 관계에서 $y = 9$와 같아지는 x는 3보다 작은 값 $\rightarrow$ 그래프가 왼쪽으로 이동
- $y = (x-1)^2$인 관계에서 $y = 9$와 같아지는 x는 3보다 큰 값 $\rightarrow$ 그래프가 오른쪽으로 이동

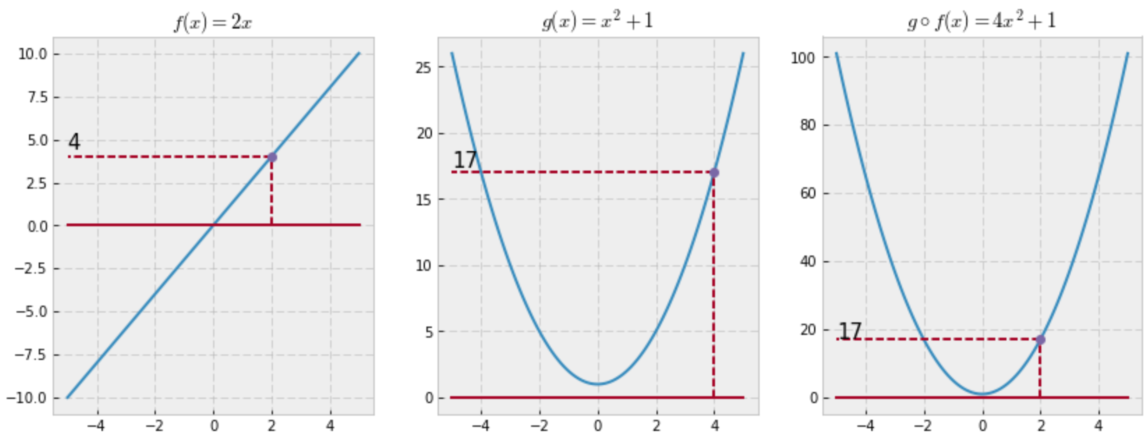
함수의 합성

- 함수 $f : X \rightarrow Y$의 공역과 함수 $g : Y \rightarrow Z$의 정의역이 같다고 할 때 다음과 같이 정의된 함수 $g \circ f$를 두 함수 f와 g의 합성이라고 한다.
- 또는 다음처럼 표기하기도 함 $g(f(X))$
- 쉽게 f의 출력이 g의 입력으로 들어감
- 뉴럴네트워크의 레이어의 출력이 다음 레이어의 입력으로 들어가는 개념과 동일

합성함수의 성질(참고)†

- 교환법칙 성립하지 않음 $f \circ g \neq g \circ f$
- 결합법칙 성립 $h \circ (g \circ f) = (h \circ g) \circ f$
- $f : X \rightarrow Y$이고 $I_X : X \rightarrow X, I_Y : Y \rightarrow Y$가 항등함수일 때 $f \circ I_X = f, I_Y \circ f = f$

- $f : x \rightarrow 2x, g : x \rightarrow x^2 + 1$ 일때
$$g \circ f = g(f(x)) = (2x)^2 + 1 = 4x + 1$$



역함수

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

역함수의 성질(참고)†

- 역함수의 역함수는 자기 자신

$$\left(f^{-1}\right)^{-1} = f$$

- 역함수와의 합성함수는 항등함수

$$\left(f^{-1} \circ f\right)(x) = f^{-1}\left(f(x)\right) = f^{-1}(y) = x$$
$$\left(f \circ f^{-1}\right)(y) = f\left(f(y)^{-1}\right) = f(x) = y$$

- 두 함수 $f : X \to Y, g : Y \to X$에 대해서 두 함수의 합성함수가 항등함수이면 역함수

$$(f \circ g)(y) = y \implies g = f^{-1}, f = g^{-1}$$
$$(g \circ f)(x) = x \implies g = f^{-1}, f = g^{-1}$$

- 두 함수 $f : X \to Y, g : Y \to Z$가 일대일 대응함수여서 그 역함수가 $f^-1(y), g^{-1}(x)$일 때

$$\begin{aligned}\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f) &= f^{-1} \circ \left(g^{-1} \circ g\right) \circ f \\ &= f^{-1} \circ I_Y \circ f \\ &= f^{-1} \circ f = I_X \\ (g \circ f) \circ \left(f^{-1} \circ g^{-1}\right) &= g \circ \left(f \circ f^{-1}\right) \circ g^{-1} \\ &= g \circ I_Y \circ g^{-1} \\ &= g \circ g^{-1} = I_Z \\ (g \circ f)^{-1} &= \left(f^{-1} \circ g^{-1}\right)\end{aligned}$$

함수의 종류[3]

- 위 소개한 스칼라, 벡터에 대해서 머신러닝에서 주로 다루는 함수들이 어떤 함수들인지 중심으로 간단히 살펴봄

일변수-실함수univariable scalar function

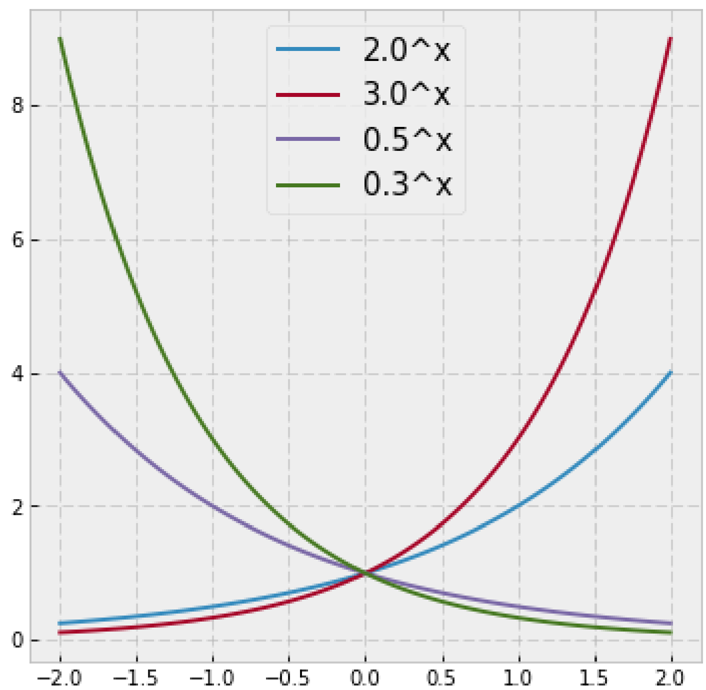
- $y = f(x) \qquad f: \mathbb{R} \to \mathbb{R}$
- 고등학교때 많이 봤던 우리가 익히 알고 있는 함수
- **다항함수**, 분수함수, **지수함수**, **로그함수**, 삼각함수
- $f(x) = x^2$

지수 함수

- 머신러닝에 자주 등장하는 대표적인 일변수-실함수인 지수, 로그 함수를 간단히 정리
- 지수함수

$$y = a^x \qquad a > 0, a \neq 1$$

- 그래프는 a에 따라 달라짐
- $a > 1$: 양의 방향으로 증가, $0 < a < 1$: 양의 방향으로 감소



몇가지 지수법칙(참고)

$$a^m a^n = a^{m+n}$$

$$a^m \div a^n = a^{m-n}$$

$$(a^m)^n = a^{mn}$$

$$(ab)^m = a^m b^m$$

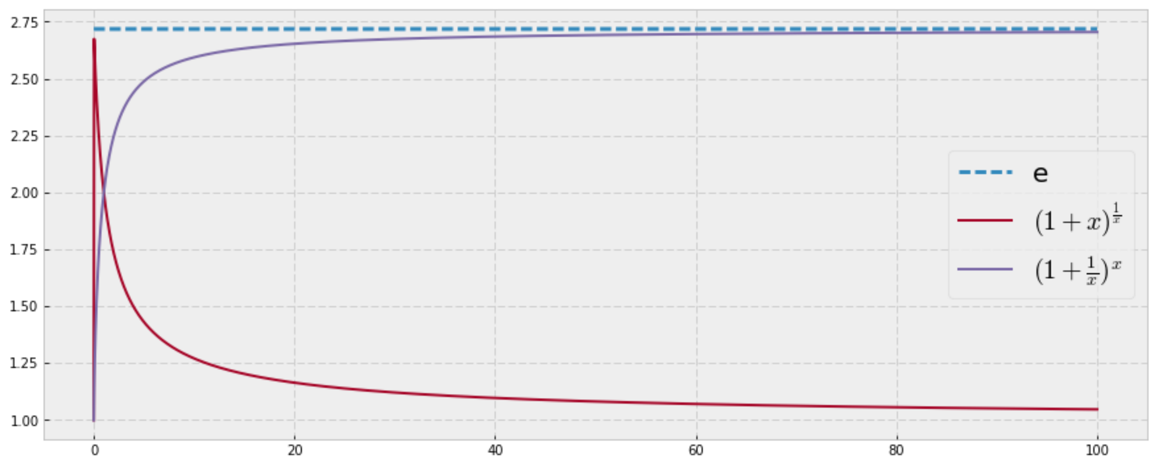
$$a^0 = 1$$

$$a^{-m} = \frac{1}{a^m} \qquad \because a^m a^{-m} = a^{m+(-m)} = a^0 = 1$$

무리수 e

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

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



로그 함수

- 로그 정의

$$\log_a x = c$$

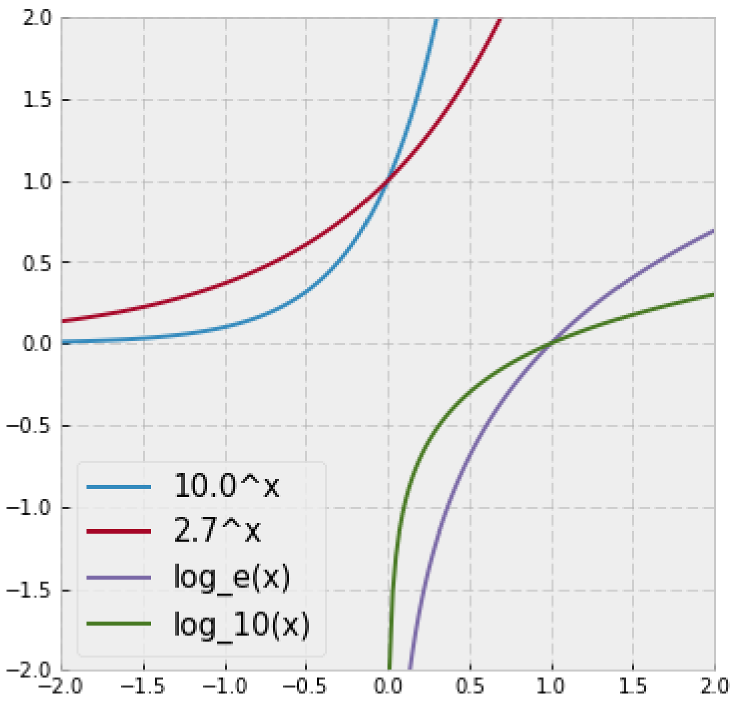
- 로그는 숫자, 위 식에서 a를 x로 만들기 위해 a의 어깨위에 거듭 제곱 되어야 하는 숫자가 로그

$$a^c = x$$

- 따라서 정의대로 써보면

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

- x를 바꿔가면서 함수처럼 생각해볼 수 있다.
- 자연로그 : 특히 a가 e인 경우 $\log_e$를 $\ln$으로 표시
- 식을 보면 로그의 출력이 지수식의 입력으로 들어가고 지수식의 출력이 로그의 입력으로 $\Longrightarrow$ 역함수 관계



로그의 성질(참고)

- $\log_a 1 = 0, \log_a a = 1 : a^{\log_a 1} = 1, a^{\log_a a} = a$
- $\log_a MN = \log_a M + \log_a N$

$\log_a MN$은 a를 MN으로 만들기 위해 어깨위에 올라가는 수

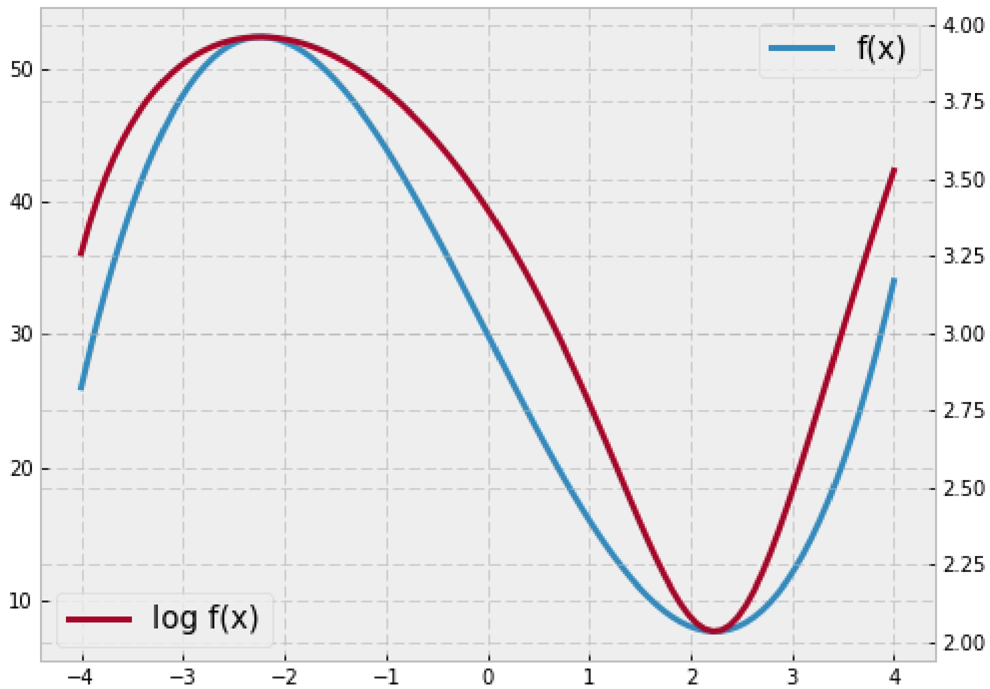
$$a^{\log_a MN} = MN$$
$$a^{\log_a M} = M, \quad a^{\log_a N} = N$$
$$a^{\log_a M} \cdot a^{\log_a N} = MN$$
$$a^{\log_a M + \log_a N} = MN$$

- $\log_a \frac{M}{N} = \log_a M - \log_a N$
- $\log_a M^k = k \log_a M$
- $\log_a b = \dfrac{\log_c b}{\log_c a}$

로그의 정의의 대로 써주면

$$a^{\log_a b} = b$$
$$\log_c \left(a^{\log_a b}\right) = \log_c b$$
$$(\log_a b) \log_c a = \log_c b \quad \because \log_a M^k = k \log_a M$$
$$\log_a b = \frac{\log_c b}{\log_c a} \quad \because \log_c a \neq 0$$

- 어떤 함수의 극점의 위치를 변화시키지 않는다.



다변수-실함수multivariable scalar function

- $z = f(x, y) \qquad f : \mathbb{R}^2 \to \mathbb{R}$
- 가장 자주 보게 될 함수
- 스칼라장scalar field를 정의한다고 말한다.
- 공간의 온도장(온도분포), 대기의 기압장(기압분포), 하지만 우리는 물리적 의미는 생각하지 않음
- $f(\mathbf{x}) = f(x_1, x_2) = 2x_1 + 3x_2 - x_1^3 - 2x_2^2$
- 공간에서의 유클리드 거리 $f(\mathbf{p}; \mathbf{p}_0) = f(x, y, z; x_0, y_0, z_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$
- **우리에게 가장 중요한 예는 목적함수objective function, 코스트함수cost function**

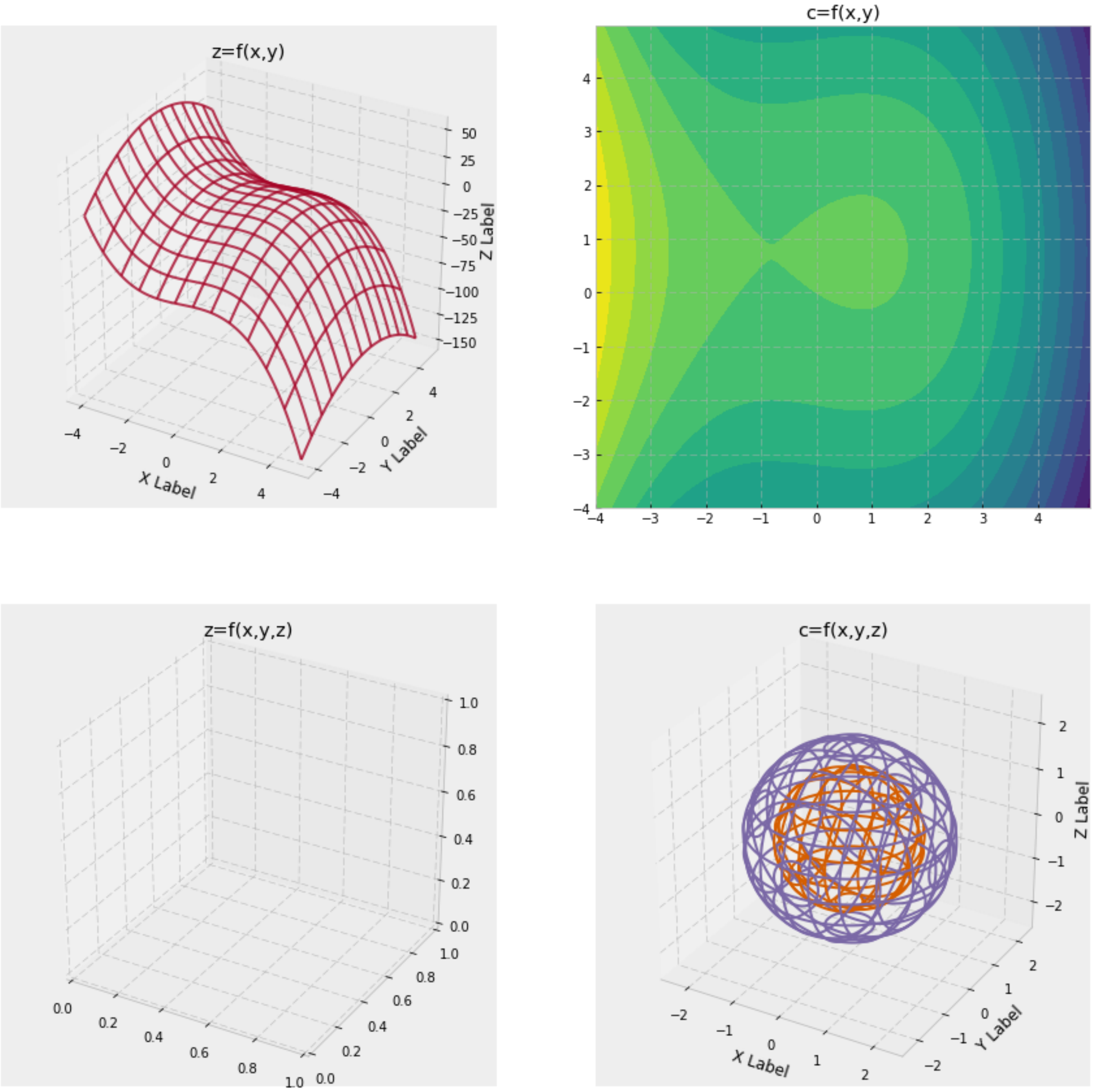
$$C(\mathbf{w}) = \frac{1}{2N} \sum_{i=1}^{N} \| y(\mathbf{x}_i; \mathbf{w}) - t \|^2$$

- 결합확률밀도함수

$$f(x, y) = \begin{cases} \frac{12}{5} x(2 - x - y), & 0 < x < 1, 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

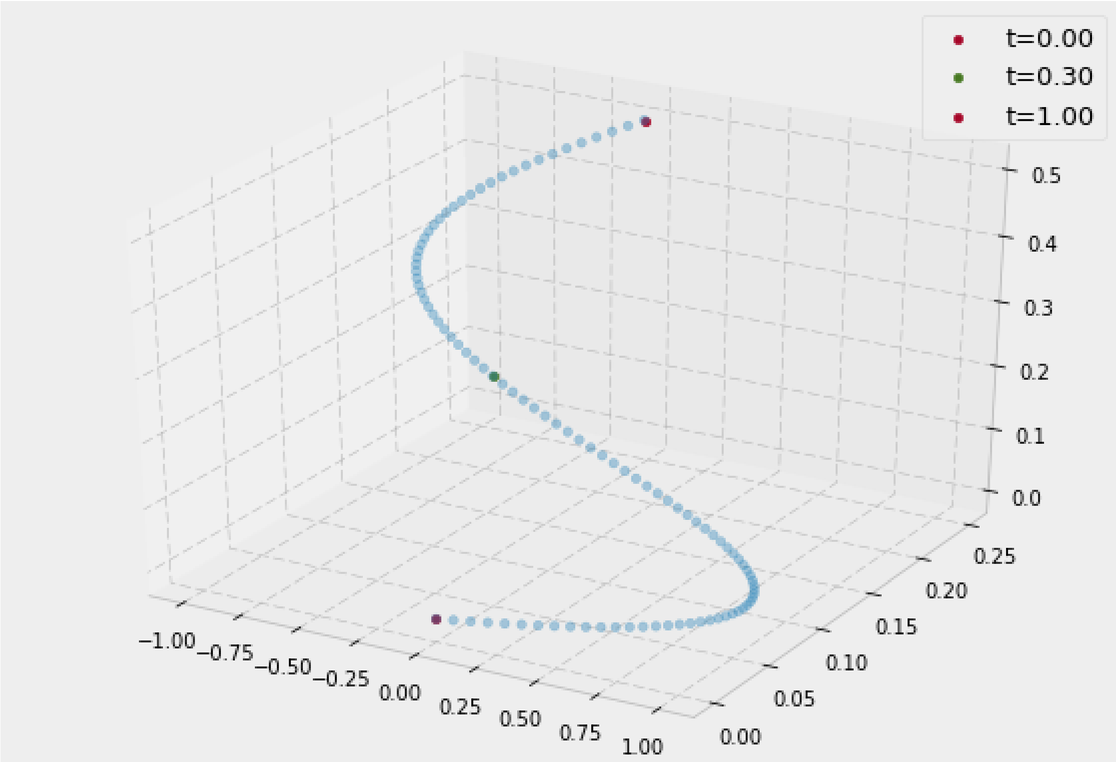
양함수와 음함수[4]

- 2변수 실함수에서의 양함수와 음함수 : $z = f(x, y)$ 와 $f(x, y) = 0$
- 3변수 실함수에서의 양함수와 음함수 : $w = f(x, y, z)$ 와 $f(x, y, z) = 0$
- 뭐가 어떻게 다른지 그림으로 확인해보기


z=f(x,y)
X Label
Y Label
Z Label
c=f(x,y)
z=f(x,y,z)
c=f(x,y,z)
X Label
Y Label
Z Label

일변수-벡터함수univariable vector function

- $\mathbf{r}(t) = (x(t), y(t), z(t))$
- 평면 또는 공간에 존재하는 곡선
- $f\left(x(t), y(t), z(t)\right) = \left(\sin(6t), \frac{1}{4}t, \frac{t^2}{2}\right)^\mathrm{T}$



다변수-벡터함수multivariable vector function

- 파라메트릭 표현parametric representation으로 3차원에 존재하는 곡면[5]

$$\mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v))^\mathrm{T} \qquad \mathbb{R}^2 \to \mathbb{R}^3$$

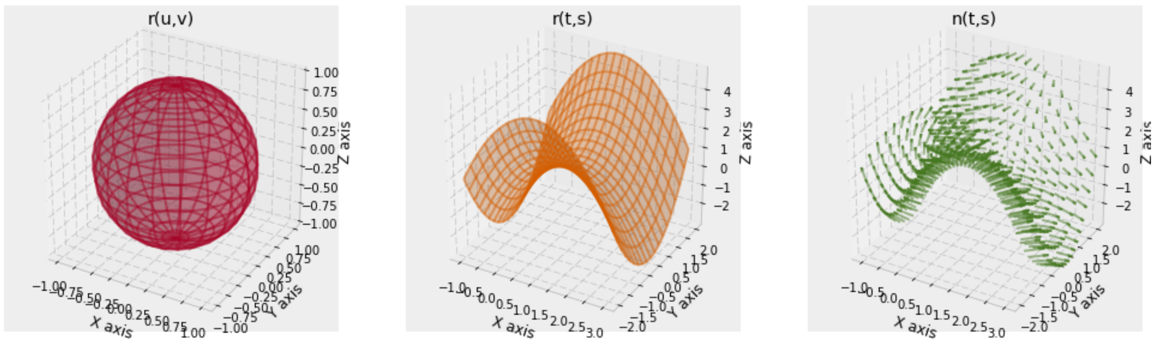
- 예[5] : $\mathbf{r}(u, v) = (\cos u \sin v, \sin u \sin v, \cos v)^\mathrm{T} \qquad 0 \le u \le 2\pi, \quad 0 \le v \le \pi$
- 곡면에서의 법선벡터장normal vector field

$$\mathbf{n}(t, s) = \frac{\dfrac{\partial \mathbf{r}}{\partial t} \times \dfrac{\partial \mathbf{r}}{\partial s}}{|\mathbf{n}|} \qquad \mathbb{R}^3 \to \mathbb{R}^3$$

- 예[6] : 곡면 $\mathbf{r}(t, s) = (t + 1, s, s^2 - t^2 + 1)^\mathrm{T}$의 법선 벡터장
- **소프트맥스 활성함수softmax activation function**

$$\sigma : \mathbb{R}^K \to [0, 1]^K$$

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



In [13]:

```
z = np.array([1.0, 2.0, 3.0, 4.0, 1.0, 2.0, 3.0])
exp_z = np.exp(z)
print("input vector z : {}".format(exp_z))

sum_exp_z = exp_z.sum()
print("sum exp(z)     : {}".format(sum_exp_z))

softmax = exp_z / sum_exp_z
print("softmax(z)     : {}".format(softmax))
print("sum softmax(z) : {:.2f}".format(softmax.sum()))
```

```
input vector z : [ 2.7183  7.3891 20.0855 54.5982  2.7183  7.3891 20.0855]
sum exp(z)     : 114.98389973429897
softmax(z)     : [ 0.0236  0.0643  0.1747  0.4748  0.0236  0.0643  0.1747]
sum softmax(z) : 1.00
```