# 함수의 쓰임새

함수는 숫자들어가서 숫자가 나온다

함수 대응관계의 특징

• 성질 : 사랑의 작대기

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

매트플롯리브로 그림됨

2차함수(1차원함수) 포물선인 이유

직교좌표 두개 썼기 때문

직교좌표 두개 안 쓴 경우?

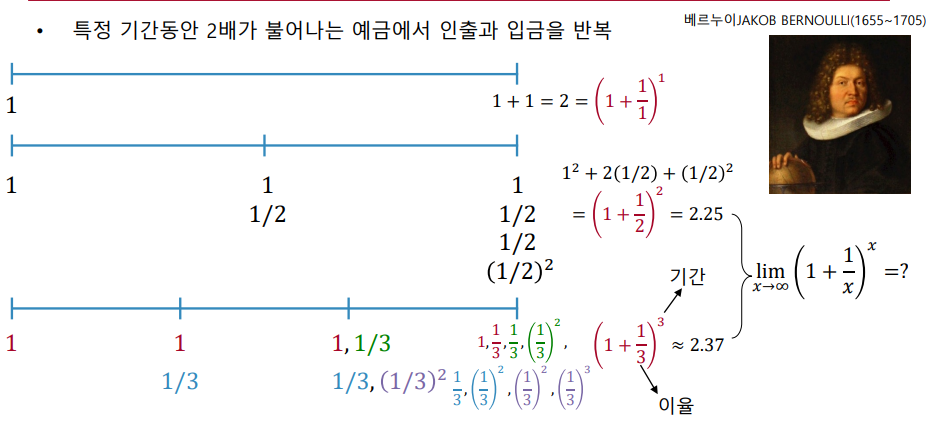
밝기로 (이후 그래디언트 그리는 방법으로 연계)

다항함수

지수법칙

자연상수 e – 복리 이자로 설명하심

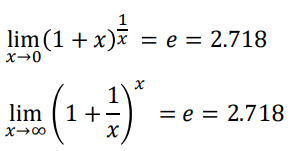
복리 이자

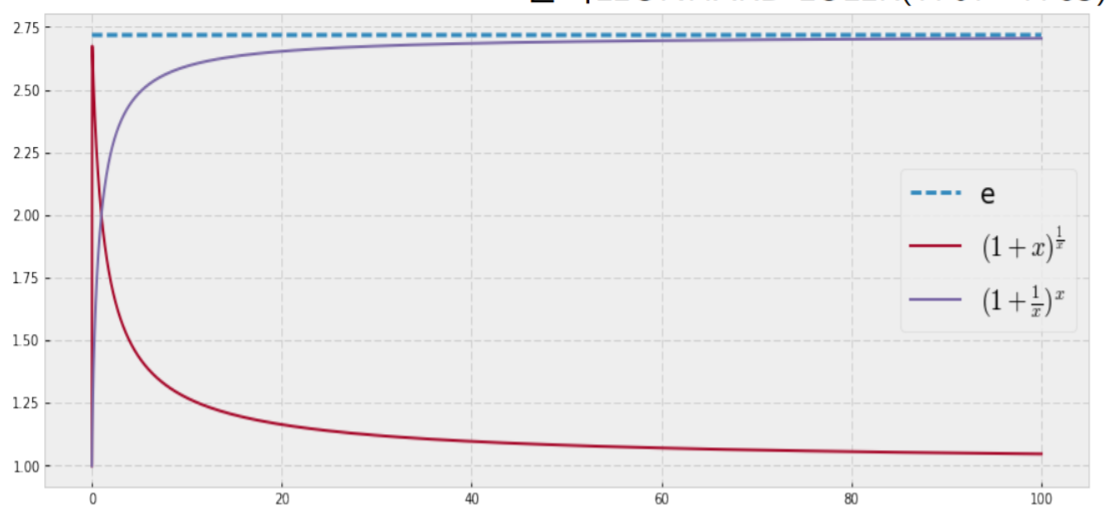


• 야코프 베르누이에 의해 아래식이 수렴함이 발견

• 라이프니츠에 의해 상수 b로 처음 사용

• 오일러에 의해 e로 처음 표기





• 𝑒의 지수함수 표기



로그함수 : 지수함수의 역함수

역함수 : 출력이 입력으로 들어가는 것

$𝑓^{−1}: 𝑌 → 𝑋, 𝑥 = 𝑓^{−1}(𝑦)$

e를 밑으로 가지면 편함 (0부터 1사이의 입력 값에 민감 – 시그모이드 함수로 감)

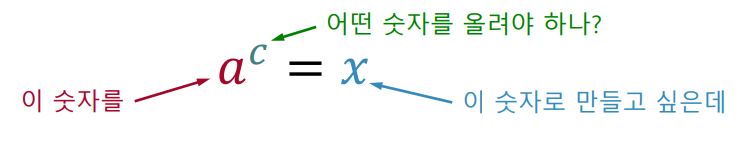
(중요)로그함수 쓰임새 : 곱하기를 더하기로 쪼갬, 볼록한 위치를 바꾸지 않음, 숫자가 커졌을 때 함수 값의 민감도를 낮춤

• 로그 정의

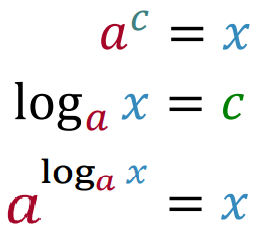


• 로그는 숫자

• 𝑎를 𝑥로 만들기 위해 𝑎의 어깨 위에 거듭 제곱 되어야 하는 숫자



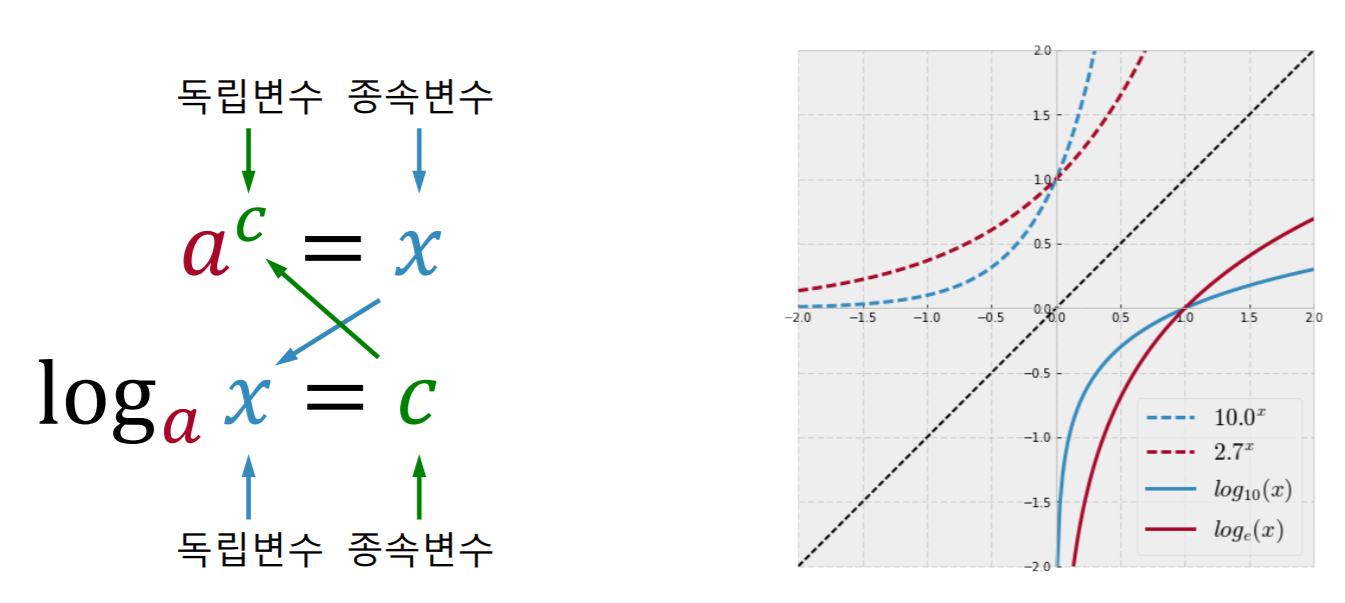
• 로그 정의대로 쓰면



• 𝑒가 𝑎자리에 오면 자연로그라 하고 ln 𝑥로 표기, 그냥 log 𝑥로 쓰기도함

• 𝑥를 바꿔가면서 함수처럼 생각로

• 로그함수, 지수함수 비교



로그함수 특징 : 입력에 대해 출력이 많이 변화하지 않음

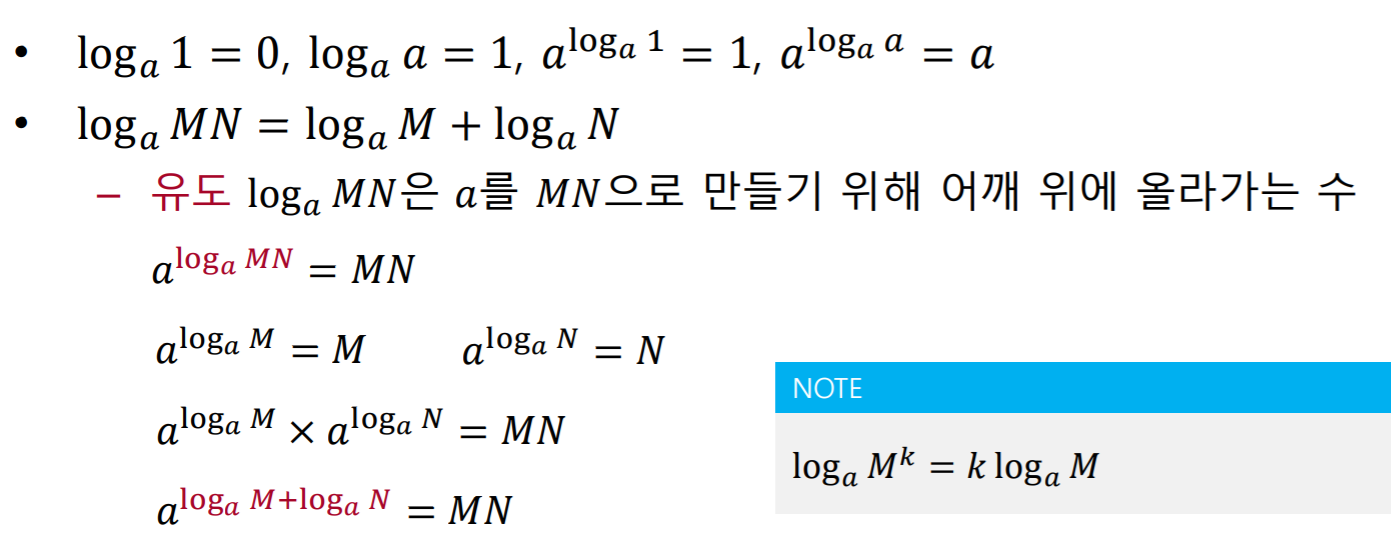
반대로 지수함수는 입력에 대해 출력이 엄청 큼

로그 성질 (정의에서 얻어낼 수 있다)

• log\_𝑎 1 = 0, log\_𝑎 𝑎 = 1, 𝑎^{log\_𝑎 1} = 1, 𝑎^{log\_𝑎 𝑎} = 𝑎

• log𝑎 𝑀𝑀𝑀𝑀 = log𝑎𝑎 𝑀𝑀 + log𝑎𝑎 𝑁𝑁

– 유도 log𝑎𝑎 𝑀𝑀𝑀𝑀은 𝑎𝑎를 𝑀𝑀𝑀𝑀으로 만들기 위해 어깨 위에 올라가는 수



최대 가능도 추정 MLE

우리가 알고 싶은 건 :

p: 확률밀도함수 (다변수함수, n변수 스칼라 함수)

* 양수가 되야 하고 (확률이니까)
* 아래 면적 다 더하면 1

N변수의 스칼라 함수

독립조건이면 다 곱하면 됨

독립이랑 가정해보자

\mu로 미분해서 0 되는 위치를 찾아야함 -> 곱미분 필요 (n이 백만개 이상이면 구하기가 힘듦)

* 전체에 로그를 씌우자!

미분당하는 함수가 로그로 다 쪼개져

다행히도 로그의 미분은 너무 쉬움

베르누이 분포이냐 카테고리 분포이냐에 따라

인경신경망은 합성함수

다변수 스칼라함수

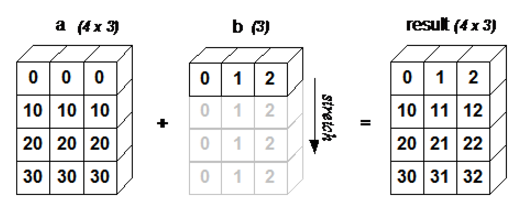
**• 신경망은 매우 많은 함수가 다음처럼 겹겹이 합성된 것이라 할 수 있다.**

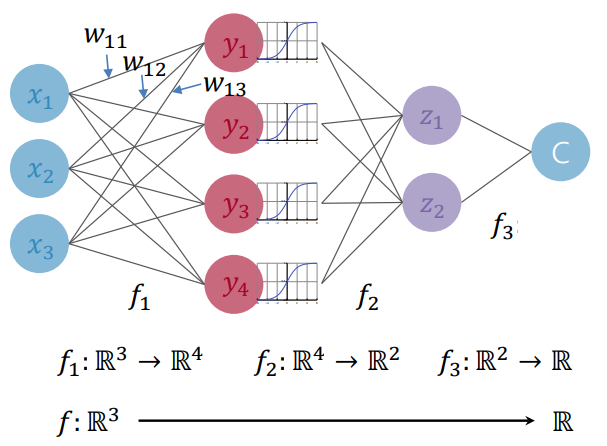
**다변수 함수**

**숫자 3개가 들어가서 숫자 4개가 나오는 다변수 벡터 함수 (3차원 벡터를 4차원 벡터로 트랜스폼 시킴)**

저 안에 신경망은 4바이3 으로 되야 함 (출력 바이 입력)

아래 사진에서 b가 열벡터로 곱하면 출력이 열벡터(4차원 벡터)가 됨



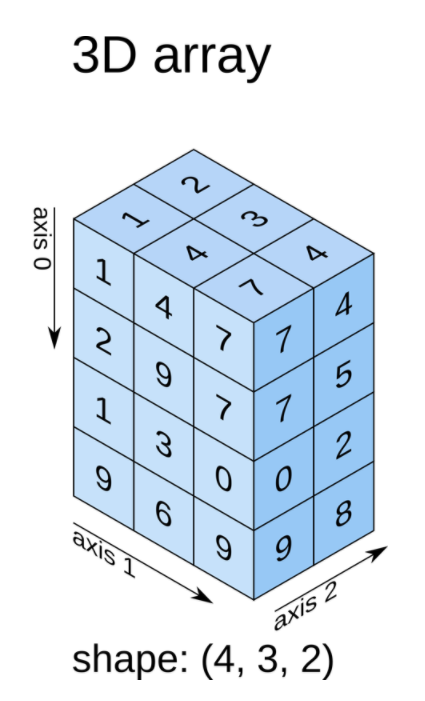


𝑓1: ℝ3 → ℝ4 𝑓2: ℝ4 → ℝ2 𝑓3: ℝ2 → ℝ

𝑓: ℝ3 → ℝ

여기서 잠깐!

벡터의 차원과 축의 차원

저렇게 생겼다하면 벡터의 차원은 3

분류기의

코스트 펑션 (크로스 엔트로피 함수)

왜 로그함수를 덮어 썼나? Gradient 확보! (두 가지 관점에서 볼 수 있음. 1.시그모이드 그대로 뽑아내면, 함수값이 증가하다가 감소하다가(또는 미분값이 0에 가까운형태)이지만 log를 씌어주면 cost함수가 지속적으로 감소하는 형태로 나타남. 2. log는 엔트로피에서 나왔음.)

크로스 엔트로피 목적함수와 손실함수가 결국 같은 의미임

배치노말리제이션 사용하면 로스함수 울퉁불퉁한걸 펴 줌 과 같은 맥락

그래디언트로 내려와야함

다변수 실함수

3차원함수로 보이는 이건2차원 함수다 했음. 우리가 보기 편하게 그린게 저런 3차원 평면 그림

• 아래 그래프에서 왼쪽 그래프는 (𝑥1, 𝑥2)로 계산된𝑓(𝑥1, 𝑥2)를 수직축(높이값)에 그린 것

• 오른쪽 그래프는 왼쪽 그래프에서 특정 높이를 만족하는 값만을 모아서 𝑥1𝑥2평면에 그린 것

• 때문에 등고선이 되며 선에 함수값이 적혀 있음

|  |  |
| --- | --- |
| plot\_surface(X1, X2, Z, …) | contour(X1, X2, Z, …) |

우리가 알던 2차원 함수($y=x^2$)는 정의역 축, 공역 축 으로 그린거

f(x,y,z) 는 공간(3차원)이 정의역이 되므로 그래프를 그릴 수 없음

그러나 함수값을 하나로 고정하면 면으로 바뀜

c = f(x,y,z)

한쪽 변이 숫자로 고정? 𝑓𝑓 𝑥𝑥, 𝑦𝑦 = 𝑐𝑐

• 2변수 실함수 𝑧𝑧 = 𝑓𝑓(𝑥𝑥, 𝑦𝑦), 𝑓𝑓 𝑥𝑥, 𝑦𝑦 = 𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐.

• 3변수 실함수 𝑤𝑤 = 𝑓𝑓(𝑥𝑥, 𝑦𝑦, 𝑧𝑧), 𝑓𝑓 𝑥𝑥, 𝑦𝑦, 𝑧𝑧 = 𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐𝑐.

• 3변수 실함수의 경우 𝑤𝑤 = 𝑓𝑓(𝑥𝑥, 𝑦𝑦, 𝑧𝑧) 형태는 그래프로 그릴 수 없음

그릴 수 없다.

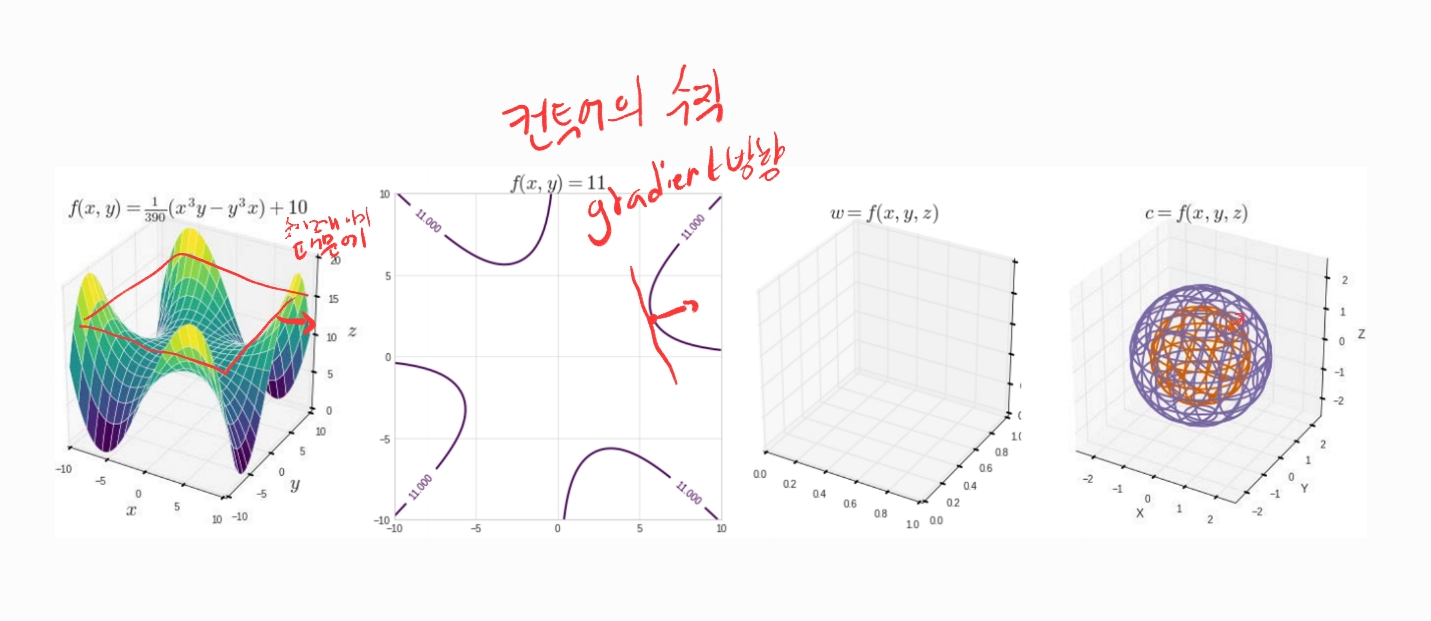
점에 함수값을 적어서 그려야하는데 축이없음

삼변수함수의 한쪽값을 고정(상수로 정의)하면 그릴 수 있음

이 상수가 커지면 구가 커짐

=> 잘랐을 때 단면이 구인거임

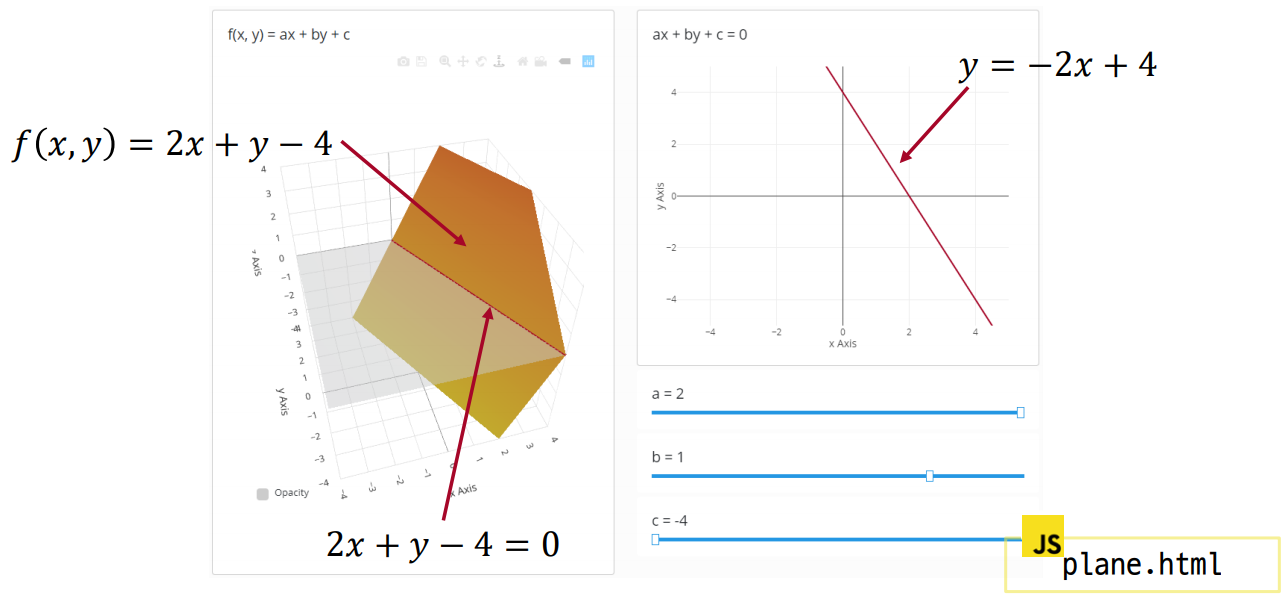
따라서 매개변수가 4개인걸 상수로 정의했을 때 (잘랐을 때)의 모양이 속이 빈 구인거



평면과 직선의 방정식

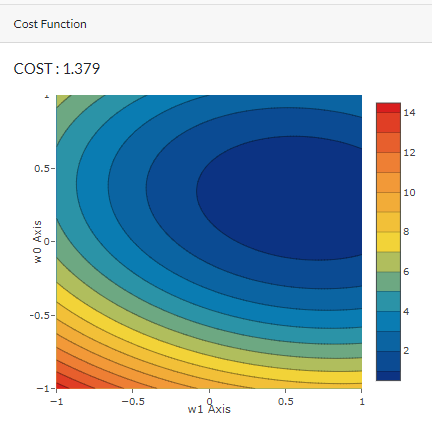
"곡면함수 = 숫자" 가 온다는 건 한 평면에서의 직교하는 숫자값이 있는 곳을 표시한건데

즉, 평면함수=숫자 하면 직선이 나옴

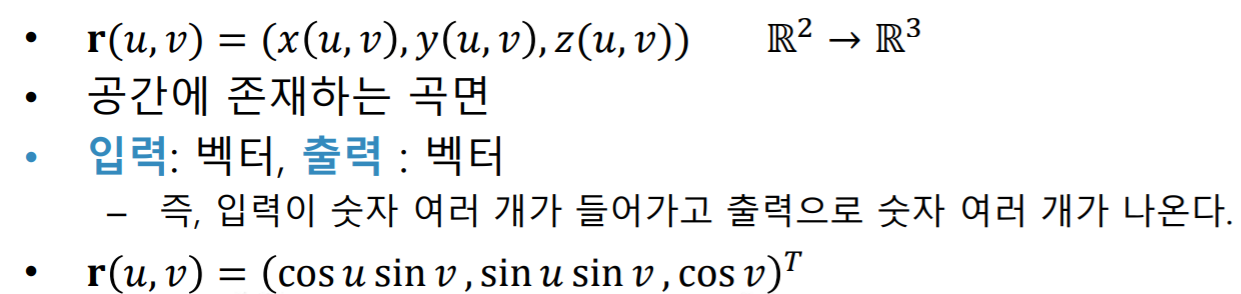


예시)

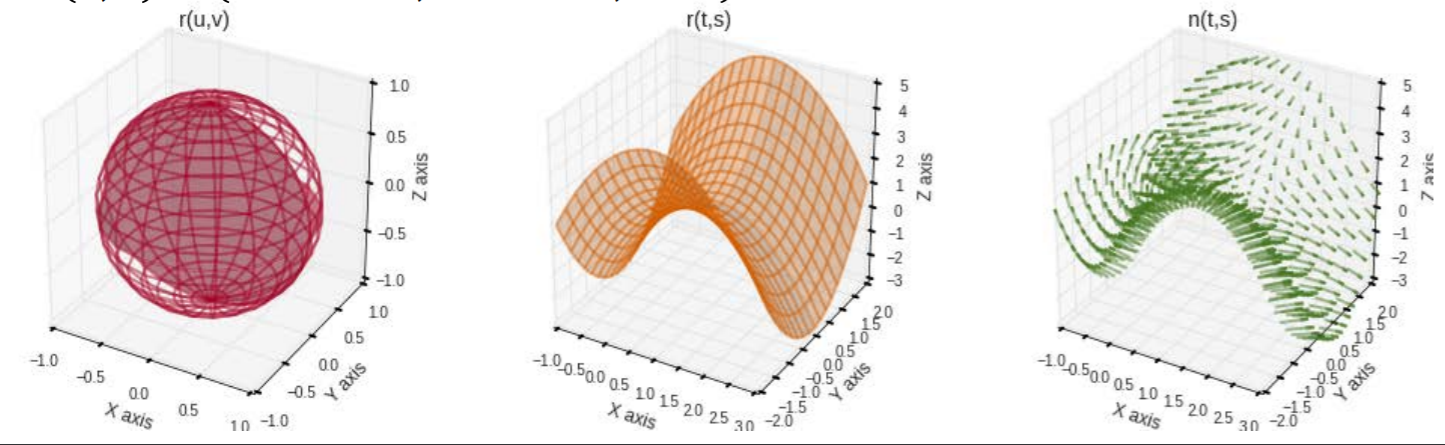
아래 cost function은 변수 2개로 그려진 평면의 단면을 자른거



다변수-벡터함수Multivariable vector func



도메인 자체가 곡면으로 말려있음  
r(u,v) 는 원점부터 벡터로 그린게 아니라, 벡터가 가르키는 지점을 점으로 나타낸 것



다변수 벡터함수의 활용

Softmax 함수를 취하면 서로 엮어서 다 더해서 1이 됨

Loss 함수

- 코스트함수<sup>cost function</sup> 또는 손실함수<sup>loss function</sup>

- 아래 손실함수는 입력 데이터 $i$개의 손실 또는 비용으로 설정한 값을 모두 더한 것으로 0이 가장 이상적인 상태

- 이 특정 $\mathbf{w}$에서 손실함수의 값을 순전파로 구하고, 역전파로 미분계수를 구한다.

$$

C(\mathbf{w}) = - \sum\_{i=1}^{N} y\_i \log[\sigma(z(\mathbf{x}\_i; \mathbf{w}))] + (1-y\_i)\log[1-\sigma(z(\mathbf{x}\_i; \mathbf{w}))]

$$

정답인 애 그대로, 아닌 애 0으로 뜸

다클래스에서 크로스 엔트로피를 쓰려면 소프트 맥스 함수를 쓰면 됨

0.8, 0.1, 0.1 이면 정답 0.8 나온 애에 대해서만 로스를 반영하면 됨

시그모이드를 쓴다면, 정답이 아닌애들을 죽이면서, 로스를 반영하기 위해서 mean square error 사용

### 소프트맥스 함수에 바이너리 크로스 엔트로피

# 미분

## 합성함수의 미분과 곱의 미분

합성함수의 미분은 각각 미분해서 곱한거고

이게 역전파 알고리즘의 기본 아이디어

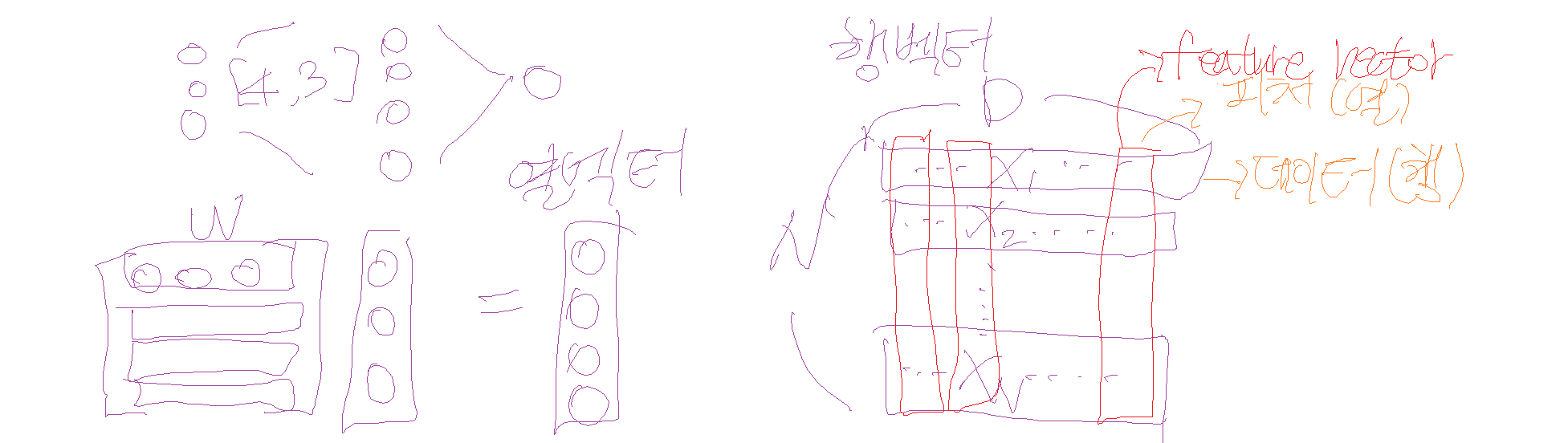
곱의 미분은 앞미분곱뒤에값 앞에값곱뒤엣값미분

## 편도함수 (편미분)

- 방향미분의 가장 특수한 미분이 편미분

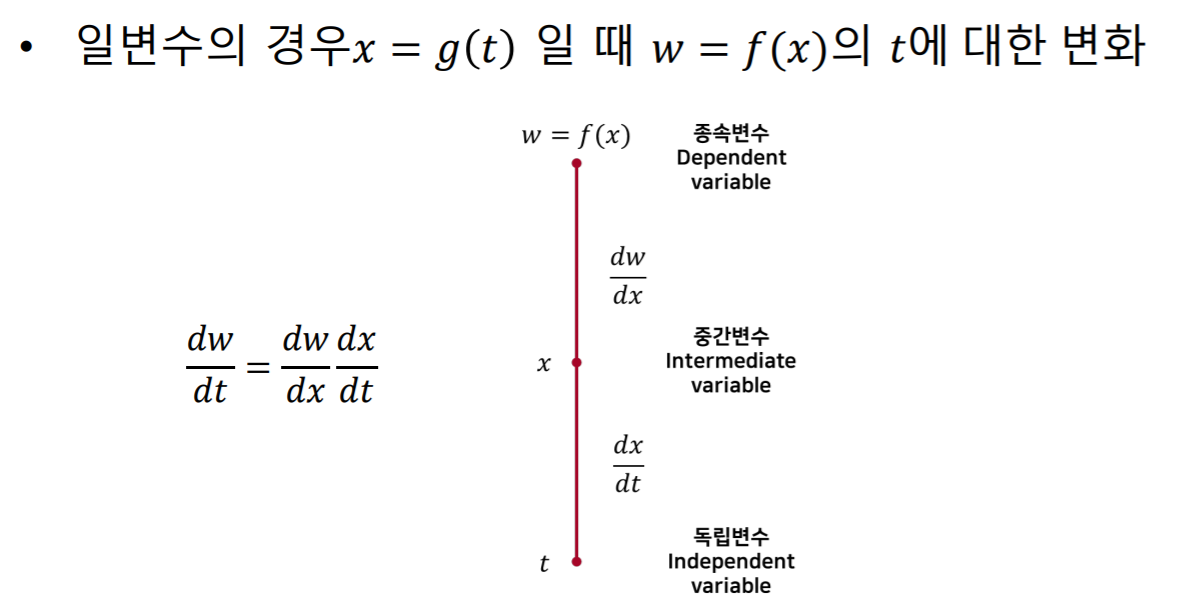
- 미분방향은 x





## 연쇄법칙

### 일변수 함수



독립변수는 종속변수에 영향을 주는 변수

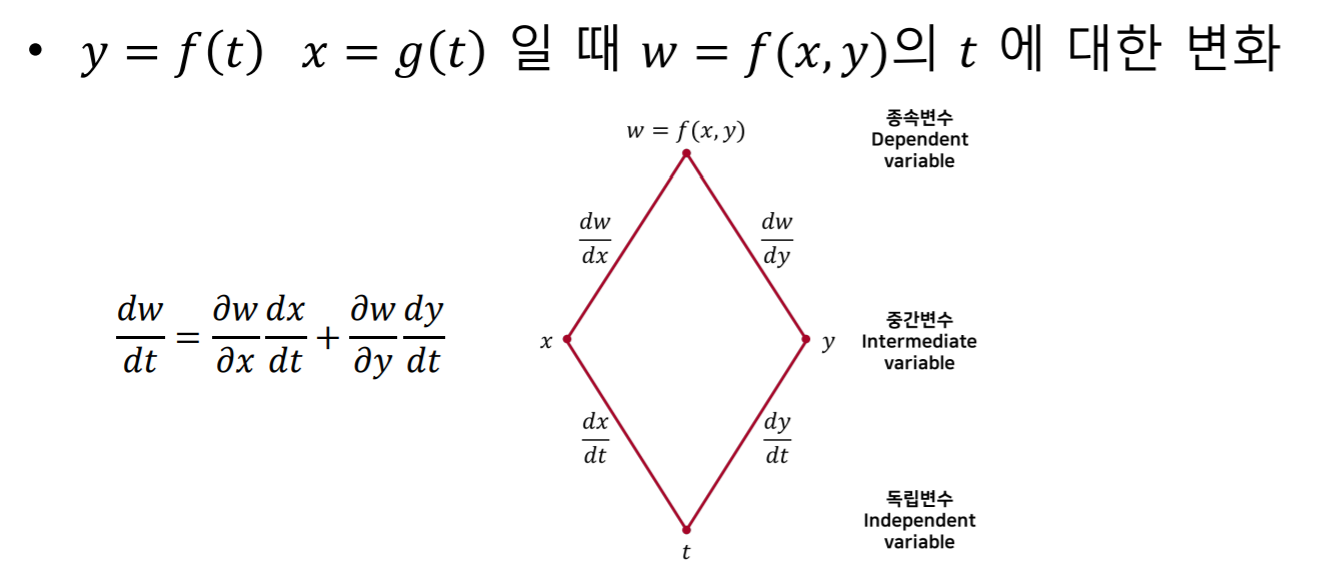
확률에서 독립과 종속 의미는 전혀상관없이

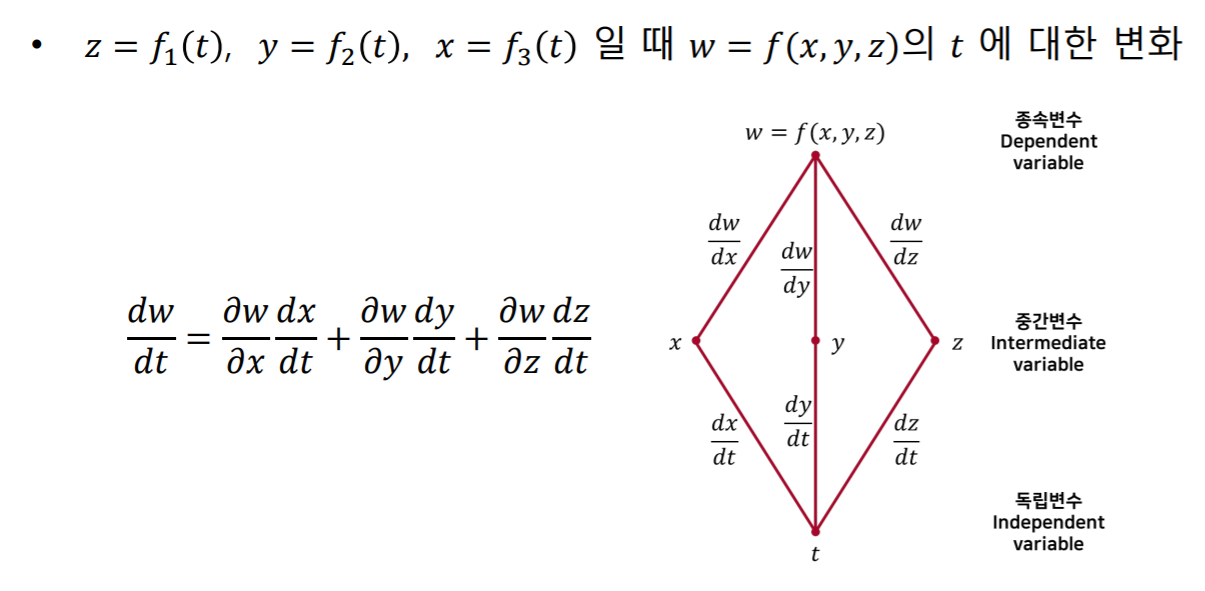
사건의 독립, 조건부확률

### 다변수 함수의 연쇄법칙

다변수함수의 미분가능성

W가 x,y에 대해서 미분가능만하면 더하기로 해주면 된다 정도만 알면 됨





## 로지스틱 함수의 미분

## 소프트맥스 함수의 미분

미분과정 : 야코비안행렬 트랜스포즈 곱하기 그레디언트

미분계수가 k곱k가되는데 얘들 끼리 묶이면 재밌는 현상

적분할때 (치환적분) 적분 도메인을 바꿔주는

단위를 바꿔주는 얘기

확률변수를 다른 확률변수로 바꿀 때

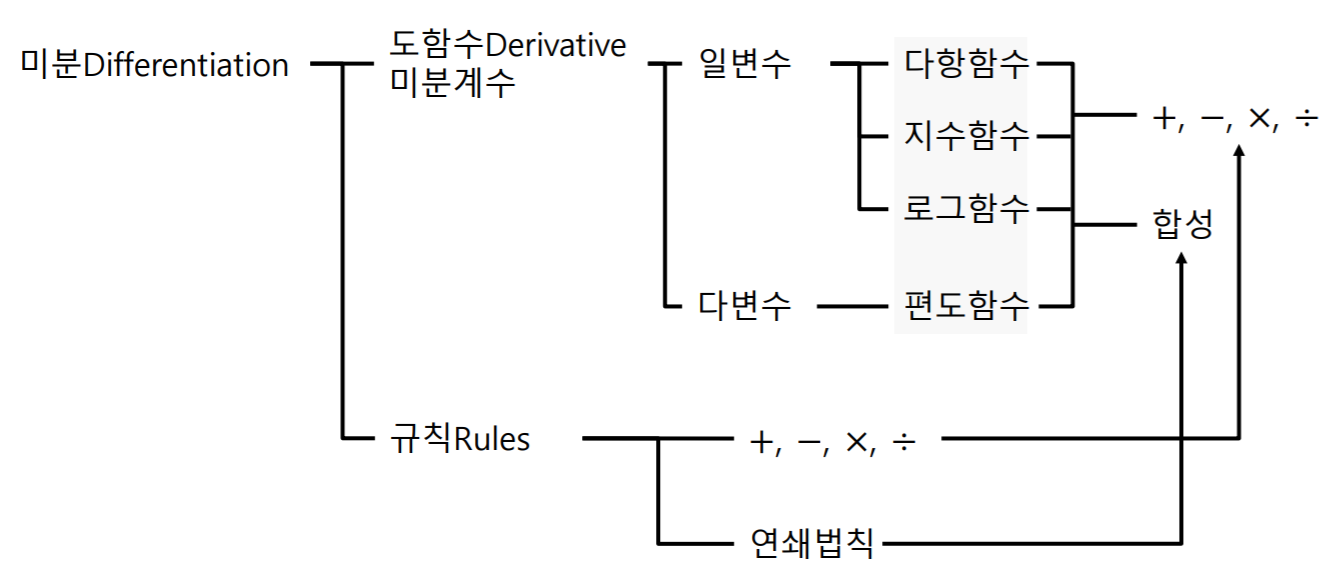
주대각선을 기준을 서로 대칭

가능한 모든 조합(변화율)을 적어놓은행렬

야코비안 행렬,

### 야코비안

총정리



# 디자인 매트릭스 정규방정식

디자인 매트릭스 곱 벡터 해서 타겟 값이 나와야 함

이항

가우시안 정규분포