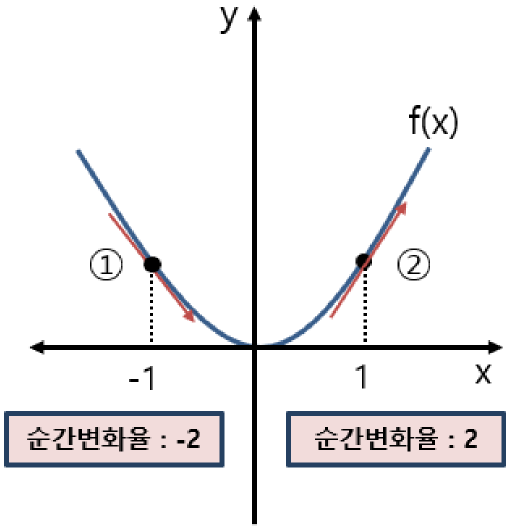
05

# 함수 근사법

2. 경사하강법

# 경사하강법 스칼라와 벡터

## 스칼라(Scalar)와 벡터(Vector)



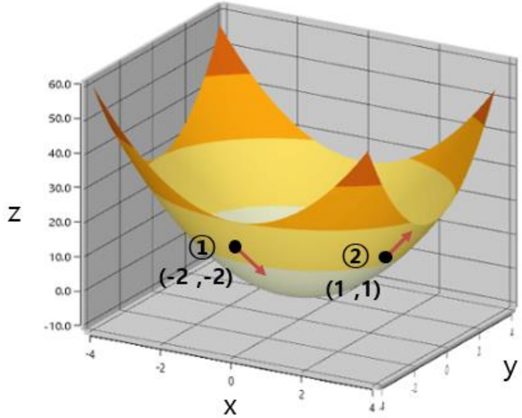

$f(x) = x^2$

$f(x)' = 2x$

- 스칼라 : 크기만 있고 방향이 없는 데이터
  ex) 몸무게, 수학점수, 키
- 벡터 : 크기와 방향을 모두 가지고 있는 데이터
  ex) 자기력, 속도, 가속도
- 스칼라를 미분 또는 편미분하면 벡터로 변하며 방향성을 가진다

# 경사하강법 그래디언트

## 그래디언트(Gradient)



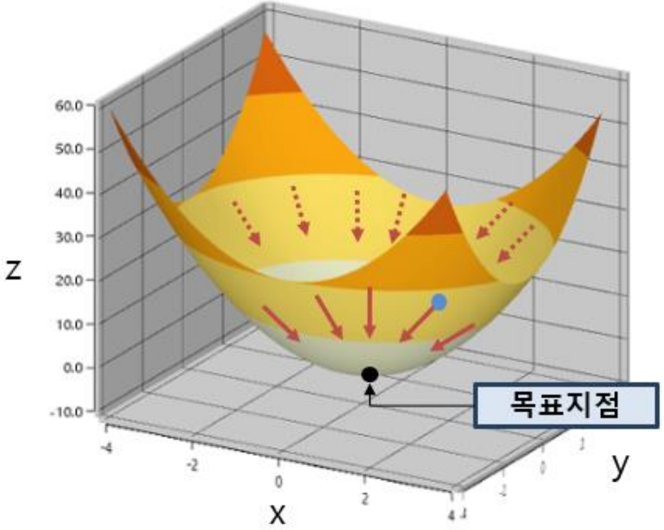

| 함 수 | $z = f(x,y) = 2x^2 + 2y^2$ |
|---|---|
| 그래디언트 | $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (4x, 4y)$ |

- 그래디언트는 공간의 기울기를 의미
- 그래디언트는 모든 변수에 대한 편미분을 각각 구하고 행렬의 형태로 구성
- 점(2)에서 (x, y) = (1 , 1)이라면 그래디언트 값은 (4, 4)가 된다. x와 y모두 값이 늘어나는 방향으로 4의 크기를 가지고 값이 변화한다는 것

# 경사하강법 기본개념



## 경사하강법(Gradient Decent)




- 경사하강법 : 반복적인 계산을 통해 미분가능한 함수의 가장 작은 값(극소점: local minimum)을 찾아내는 최적화 알고리즘
- 경사상승법 : 이와 반대로 함수의 가장 큰 값(극대점: local maximum)을 찾아내는 최적화 알고리즘

| | |
|---|---|
| 함 수 | $z = f(x,y) = 2x^2 + 2y^2$ |
| 그래디언트 | $\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ |
| 변수 증분 | $\Delta (x, y) = -\frac{1}{2} \propto \nabla f$ |

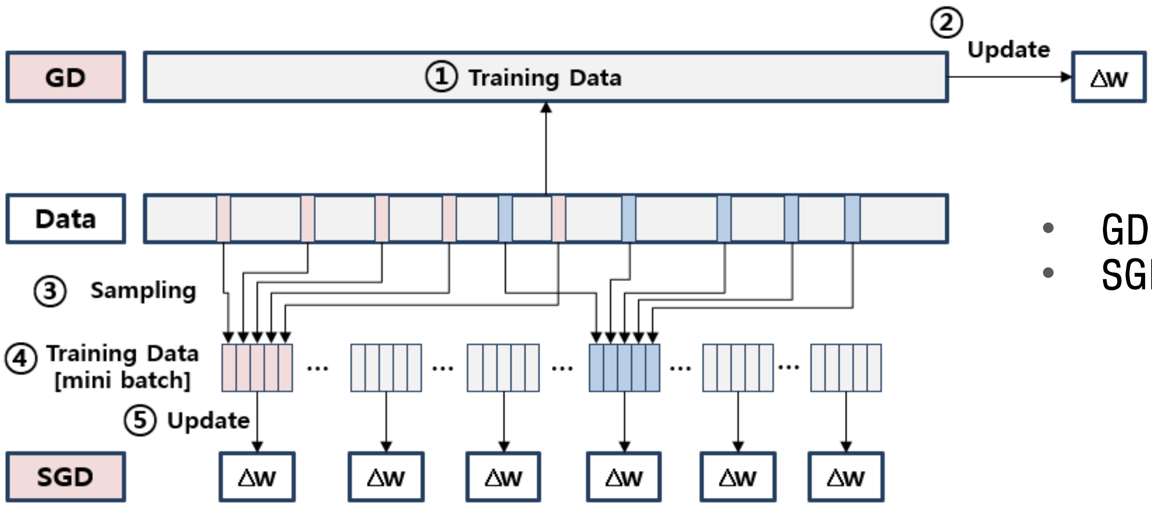
경사하강법

# 확률적 경사하강법



## 확률적 경사하강법(SGD: Stochastic Gradient Decent)

- 전체 데이터가 아닌 일부 데이터만 학습하고 좀 더 빨리 값을 수정해 가는 알고리즘




- GD : 전체 데이터
- SGD : 일부 데이터

# 경사하강법 표기법



## 강화학습에서 GD와 SGD 표기법

- 데이비드 실버(David Silver) 교수님이 작성한 강의교재에 기초한 표기법 사용

| | |
|---|---|
| 일반수식 : 편미분 | $\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ |
| 강화학습 : 편미분 | $\nabla_w J(w) = \left(\frac{\partial J(w)}{\partial w_1}, \cdots, \frac{\partial J(w)}{\partial w_n}\right)$ |
| 일반수식 : 경사하강법 | $\Delta (x, y) = -\frac{1}{2} \propto \nabla f$ |
| 강화학습 : 경사하강법 | $\Delta w = -\frac{1}{2} \propto \nabla_w J(w)$ |

https://github.com/multicore-it/rl
https://github.com/multicore-it/rl