경사하강법 순한 맛

미분

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

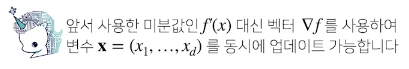
자동 생성된 설명

입력이 1변수가 아니라 다변수(벡터)인 경우에는 편미분을 사용

텍스트이(가) 표시된 사진

자동 생성된 설명

역삼각형 기호는 nabla 라고 한다



텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

L2 norm을 사용 하나 L2 norm의 제곱을 사용하나 동일

그래서 제곱을 사용하면 앞의 2가 없어짐

텍스트이(가) 표시된 사진

자동 생성된 설명

GD alg 에서는 epoch 과 lr이 중요한 parameter

Convex func에서만 유효 linear regression

경사하강법 매운맛

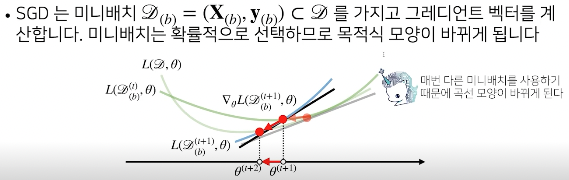
Stochastic GD (SGD) 확률적 경사하강법

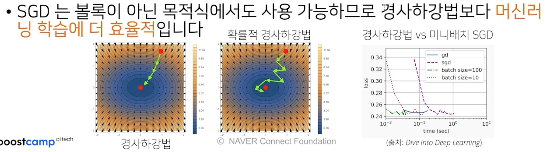
모든 데이터를 사용하여 업데이트하는 대신 데이터 한개 또는 일부 활용하여 업데이트

Non-convex func에 유효

텍스트이(가) 표시된 사진

자동 생성된 설명

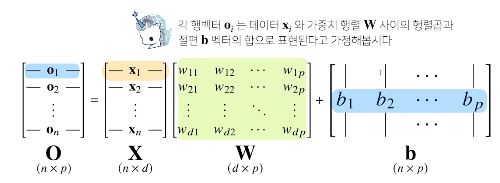


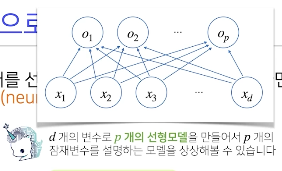


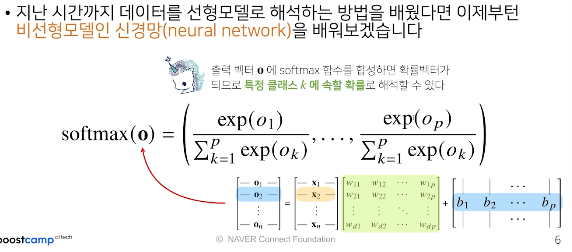
SGD의 원리는 필수불가결적. 미니배치

딥러닝 학습방법 이해하기

NN은 nonlinear model







텍스트이(가) 표시된 사진

자동 생성된 설명

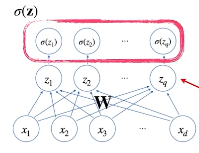
Numerator 분모

Denumerator 너무 큰 벡터가 들어오면 오버플로우 발생 가능 따라서 np.max를 빼줌

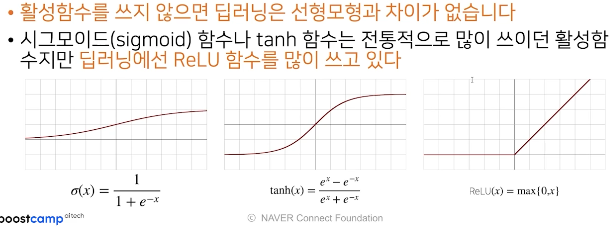
이것을 하면 오버플로우를 방지함과 동시에 원래 softmax값과 동일한 결과

텍스트이(가) 표시된 사진

자동 생성된 설명

 텍스트이(가) 표시된 사진

자동 생성된 설명



Multi-layer perceptron MLP

텍스트, 화이트보드이(가) 표시된 사진

자동 생성된 설명

층을 여러개 쌓을 수록 필요한 뉴런의 숫자는 훨씬 빨리 줄어든다

BackPropagation Alg

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

AI Math 6 확률론 맛보기

확률변수, 조건부확률, 기댓값 등은 확률론의 매우 기초적인 내용이며 이를 정확히 이해하셔야 바로 다음 강의에서 배우실 통계학으로 이어질 수 있습니다. 기댓값을 계산하는 방법, 특히 확률분포를 모를 때 몬테카를로 방법을 통해 기댓값을 계산하는 방법 등은 머신러닝에서 매우 빈번하게 사용되므로 충분히 공부하시고 넘어가시기 바랍니다.

**Further Question**

몬테카를로 방법을 활용하여 원주율에 대한 근삿값을 어떻게 구할 수 있을까요?

데이터공간 xy라 했을때 D는 데이터공간에서 데이터를 추출하는 분포



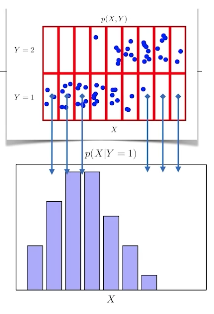
확률변수는 이산형discrete와 연속형continuous로 구분됨



텍스트이(가) 표시된 사진

자동 생성된 설명

P(X)는 주변확률분포로 y에 대한 정보를 주진 않습니다.

 텍스트이(가) 표시된 사진

자동 생성된 설명



텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명

몬테카를로 샘플링 방법은 변수 유형 (이산형, 연속형) 에 상관없이 사용할 수 있다.