회귀 문제 기울기

● 손실 함수 : 오차제곱합

● 은닉층 활성화 함수 : 시그모이드

● 출력층 활성화 함수 : 항등 함수

분류 문제 기울기

● 손실 함수 : 교차 엔트로피 오차

● 은닉층 활성화 함수 : 시그모이드

● 출력층 활성화 함수 : 소프트맥스 함수

최적화 알고리즘

1. 확률적 경사 하강법(Stochastic Gradient Descent, SGD)

2. 모멘텀(Momentum)

3. 아다그라드 (Adaptive Gradient : AdaGrad)

4. RMSProp

5. Adam (ADAptive Moment estimation )

배치 사이즈

배치(batch):훈련 샘플 그룹

에포크(epoch): 모든 훈련 데이터를 1회 학습하는 것

1에포크는 여러개의 배치로 구성

배치 사이즈=배치에 포함된 샘플의 수

1. 배치 학습

배치 사이즈= 전체 훈련 데이터의 수

1 에포크마다 전체 훈련 데이터 오차의 평균을 구해 가중치와 bias를 수정

일반적으로 안정된 학습이 진행, 상대적으로 빠른 속도

local minima주의

훈련 데이터 샘플 수가 1000개 : 배치사이즈= 1,000, 1 에포크 가중치& bias 1회 수정

1. 온라인 학습

배치 사이즈= 1

개별 샘플마다 가중치와 bias를 수정 ∴안정성이 떨어짐

local minima 위험 낮음(이전 경사하강법 기울기 구하는 방법)

훈련 데이터 샘플 수가 1000개 : 배치사이즈= 1, 1 에포크 가중치& bias 1000회 수정

1. 미니 배치

샘플을 분할,분할된 그룹마다 가중치&bias 수정

온라인 학습보다 안정성높음(데이터 민감성 낮음)

훈련 데이터 샘플 수가 1000개 : 배치사이즈= 50, 1 에포크 가중치& bias 20회 수정

적절한 배치 사이즈 찾기어려움

행렬 연산

순전파 행렬: 배치 사이즈X뉴런(노드)

역전파행렬: