

명예학회 cv 4, 5강

역전파와 신경망

이전까진 분류기를 정의하는 방법

가장 낮은 손실에 해당하는 매개변수 찾기

손실함수를 최소화하는 것이 목적 - W 에 대한 L 의 기울기

미분을 통한 기울기 파악 - 실수하기 쉬운 환경

■ 분석적 기울기를 도출하고 이를 사용하는 동시에 수치적 기울기를 사용하여 구현을 확인
프레임워크(계산그래프)

-복잡한 함수에 대한 기울기 계산 4:00

계산 그래프를 통하여 함수를 표현할 수 있으면 역전파를 사용하는 데 용이함

- 복잡한 신경망을 다룰 때 유용하게 쓰임

역전파의 작동 원리

기본적으로 함수의 형태 / 함수를 취하고, 계산그래프의 형식으로 계산 - 6:30

계산 그래프를 반복하여 기울기를 계산(단순한 형태) - 체인 룰

각 노드는 주변 환경만 인식하게 됨 - 로컬 그래디언트 기록 가능

- 복잡한 미적분학 대체 가능

끝에서 처음으로 돌아가는 방식의 그래디언트 - z 에 최종 손실에 대한 기울기 계산

각 노드에서 계산한 로컬 그래디언트로 추적하고, 역전파 중에서 그래디언트의 값을 알 수 있음

--- 임의로 함수의 모든 변수에 대한 기울기를 계산하는 방법들

- 최적화를 위해 지난 강의 내용의 전체 프레임워크를 활용

벡터 값에서의 적용

벡터의 그래디언트는 항상 원래 벡터와 동일한 크기 - 함수의 최종 출력에 얼마나 영향을 미치는지 알려주는 지표

코드에서의 적용

신경망을 사용하여 작업하게 되면 계산이 매우 복잡해지기에 모든 매개변수에 대해 그래디언트

공식을 직접 작성하는 것엔 무리가 있음

- 이러한 그래디언트를 얻기 위해 역전파를 사용 - 신경망의 핵심 기술

연쇄 법칙을 재귀적으로 적용하여 기울기를 계산

신경망

선형 함수에 대한 예시를 위주로 보았음 - 함수(레이어)를 중첩하여 계산

비선형 함수를 여러 개 중첩하는 방식 58:00

특정 클래스의 최종 점수를 얻기 위해 여러 템플릿을 통합

합성곱 신경망

완전 연결 레이어

완전히 연결된 레이어에 대해 벡터의 영역에서 작업 (32 32 3 -> 3072 1)

10 3072의 가중치

3차원 입력을 유지한 채로 가중치를 적용 - 필터 사용(동일한 깊이 사용)

가장자리 - 복잡한 종류의 기능(모서리, 얼룩 등) - 얼룩의 개념과 더 유사한 개념 - 네트워크 구성

풀링과 패딩

이전과 동일한 입력 크기를 유지하기 위해 제로 패딩 사용 (정보의 손실 방지, 출력의 크기 유지)

완전 연결 신경망과의 차이

풀링 레이어의 역할

표현을 간략하게 하여 관리에 용이 - 다운샘플링(깊이의 변동 없음)

최대 풀링

렐루 레이어는 기본 신경망과 함께 이전에 가졌던 것과 동일, 다운샘플링 적용 시 풀링 레이어 사용