**5.1 계산 그래프**

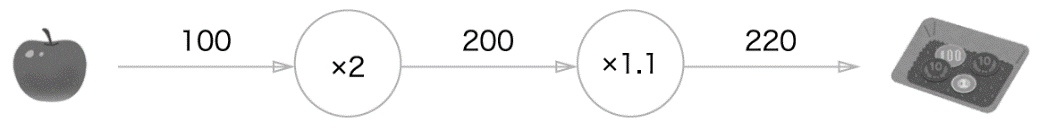
계산그래프: 계산 과정을 그래프로 나타낸 것, 복수의 노드, 엣지로 표현

**5.1.1 계산 그래프로 풀다**

문제 1: 현빈 군은 슈퍼에서 1개에 100원인 사과를 2개 샀습니다. 이때 지불 금액을 구하세요. 단 소비세가 10% 부과됩니다.

노드: 원으로 표기하고 원안에 연산 내용을 적음

화살표: 계산 결과를 화살표 위에 적어, 노드의 계산 결과가 왼쪽에서 오른쪽으로 전해짐



곱셈인 x만을 연산으로 생각할 경우, 사과의 개수와 소비세가 변수가 되어 원 밖에 표기

텍스트, 도표, 스크린샷, 라인이(가) 표시된 사진

자동 생성된 설명

문제 2: 현빈 군은 슈퍼에서 사과를 2개, 귤을 3개 샀습니다. 사과는 1개에 100원, 귤은 1개 150원입니다. 소비세가 10%일 때 지불 금액을 구하세요.

텍스트, 도표, 폰트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

1. 계산 그래프를 구성

2. 그래프에서 계산을 왼쪽에서 오른쪽으로 진행

**5.1.2 국소적 계산**

계단 그래프 특징: 국소적 계산을 전파함으로써 최종 결과를 얻음

국소적: 자신과 직접 관계된 작은 범위

국소적 계산: 전체에서 어떤 일이 벌어지든 상관없이 자신과 관계된 정보만으로 결과를 출력

텍스트, 스크린샷, 도표, 폰트이(가) 표시된 사진

자동 생성된 설명

각 노드는 자신과 관련한 계산 외에는 신경 쓸 것이 없음

**5.1.3 왜 계산 그래프로 푸는가?**

계산 그래프의 이점: 국소적 계산, 중간 계산 결과를 모두 보관 가능, 역전파를 통한 미분을 효율적으로 계산 가능

텍스트, 도표, 스크린샷, 라인이(가) 표시된 사진

자동 생성된 설명

역전파는 오른쪽에서 왼쪽으로 미분 값을 전달

**5.2 연쇄법칙**

**5.2.1 계산 그래프의 역전파**

도표, 라인, 폰트, 스케치이(가) 표시된 사진

자동 생성된 설명

신호 E에 노드의 국소적 미분을 곱한 후 다음 노드로 선달

국소적 미분: 순전파 때의 y=f(x) 계산의 미분을 구하는 것, x에 대한 ㅛ의 미분을 구하는 것

국소적인 미분을 상류에서 전달된 값에 곱해 앞쪽 노드로 전달하는 것

**5.2.2 연쇄법칙이란?**

합성함수: 여러 함수로 구성된 함수

폰트, 화이트, 상징, 타이포그래피이(가) 표시된 사진

자동 생성된 설명

연쇄법칙: 합성 함수의 미분에 대한 성질, 합성 함수를 구성하는 각 함수의 미분의 곱으로 나타낼 수 있음

폰트, 타이포그래피, 번호, 상징이(가) 표시된 사진

자동 생성된 설명 폰트, 화이트, 타이포그래피, 텍스트이(가) 표시된 사진

자동 생성된 설명

폰트, 친필, 화이트, 서예이(가) 표시된 사진

자동 생성된 설명

**5.2.3 연쇄법칙과 계산 그래프**

도표, 폰트, 라인, 그림이(가) 표시된 사진

자동 생성된 설명

도표, 라인, 그림, 스케치이(가) 표시된 사진

자동 생성된 설명

**5.3 역전파**

**5.3.1 덧셈 노드의 역전파**

폰트, 타이포그래피, 화이트, 텍스트이(가) 표시된 사진

자동 생성된 설명

도표, 라인이(가) 표시된 사진

자동 생성된 설명

역전파 때는 상류에서 전해진 상류에서 전해진 미분에 1을 곱하여 하류로 흘린다. 즉, 덧셈 노드의 역전파는 1을 곱하기만 할 뿐이므로 입력된 값을 그대로 다음 노드로 보낸다.

도표, 화이트, 폰트, 텍스트이(가) 표시된 사진

자동 생성된 설명

도표, 라인이(가) 표시된 사진

자동 생성된 설명

**5.3.2 곱셈 노드의 역전파**

폰트, 화이트, 텍스트, 타이포그래피이(가) 표시된 사진

자동 생성된 설명

도표, 라인이(가) 표시된 사진

자동 생성된 설명

곱셈 노드 역전파는 상류의 값에 순전파 때의 입력 신호들을 서로 바꾼 값을 곱해서 하류로 보낸다.

도표, 라인이(가) 표시된 사진

자동 생성된 설명

1.3\*5 = 6.5

1.3\*10 = 13

곱셈의 역전파는 순방향 입력 신호의 값이 필요함

**5.3.3 사과 쇼핑의 예**

도표, 캘리퍼스이(가) 표시된 사진

자동 생성된 설명

사과 가격은 2.2 개수는 110, 소비세는 200의 미분이다.

소비세와 사과 가격이 같은 양만큼 오르면 최종 금액에는 소비세가 200의 크기로, 사과 가격이 2.2 크기로 영향을 줄 수 있다.

**5.4 단순한 계층 구현하기**

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/layer_naive.ipynb>

**5.4.1 곱셈 계층**

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/buy_apple.ipynb>

**5.4.2 덧셈 계층**

도표, 텍스트, 지도, 평면도이(가) 표시된 사진

자동 생성된 설명

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/buy_apple_orange.ipynb>

**5.5 활성화 함수 계층 구현하기**

**5.5.1 ReLU 계층**

폰트, 화이트, 친필, 텍스트이(가) 표시된 사진

자동 생성된 설명폰트, 화이트, 도표, 타이포그래피이(가) 표시된 사진

자동 생성된 설명

순전파 때의 입력인 .x가 0보다 크면 역전파는 상류의 값을 그대로 하류로 흘림, 반면 순전파 때 x가 0 이하면 역전파 때는 하류로 신호를 보내지 않음

도표, 스크린샷, 라인이(가) 표시된 사진

자동 생성된 설명

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/5.5.1.ipynb>

**5.5.2 Sigmoid 계층**

폰트, 화이트, 라인, 상징이(가) 표시된 사진

자동 생성된 설명

텍스트, 도표, 폰트, 라인이(가) 표시된 사진

자동 생성된 설명

1단계: ‘/’ 노드, 즉 y=1/x을 미분하면 다음 식이 된다.

폰트, 도표, 친필, 화이트이(가) 표시된 사진

자동 생성된 설명

도표, 텍스트, 라인, 폰트이(가) 표시된 사진

자동 생성된 설명

2단계: ‘+’ 노드는 상류의 값을 여과 없이 하류로 내보낸다.

도표, 라인, 텍스트, 폰트이(가) 표시된 사진

자동 생성된 설명

3단계: ‘exp’ 노드는 y=exp(x) 연산을 수행

폰트, 화이트, 서예, 상징이(가) 표시된 사진

자동 생성된 설명

텍스트, 라인, 폰트, 도표이(가) 표시된 사진

자동 생성된 설명

4단계: ‘x’ 노드는 순전파 때의 값을 ‘서로 바꿔’ 곱한다.

텍스트, 라인, 도표, 폰트이(가) 표시된 사진

자동 생성된 설명

Sigmoid 노드 하나로 대체 가능하다.

텍스트, 폰트, 라인, 도표이(가) 표시된 사진

자동 생성된 설명

텍스트, 폰트, 친필, 라인이(가) 표시된 사진

자동 생성된 설명

Sigmoid 계층의 역전파는 순전파의 출력(y)만으로 계산할 수 있다.

폰트, 도표, 텍스트, 라인이(가) 표시된 사진

자동 생성된 설명

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/common/layers.py>

**5.6 Affine/Softmax 계층 구현하기**

**5.6.1 Affine 계층**

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/5.6.1.ipynb>

뉴런의 가중치의 합은 Y=np.dot(X,W)+B와 같이 계산한다. 이 Y를 활성화 함수로 변환해 다음 층으로 전파하는 것이 신경망 순전파의 흐름

도표, 텍스트, 라인, 폰트이(가) 표시된 사진

자동 생성된 설명

행렬을 사용한 역전파도 행렬의 원소마다 전개해보면 스칼라값을 사용한 지금까지의 계산 그래프와 같은 순서로 생각할 수 있다.

폰트, 타이포그래피, 화이트, 텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트, 폰트, 친필, 화이트이(가) 표시된 사진

자동 생성된 설명

텍스트, 도표, 라인, 폰트이(가) 표시된 사진

자동 생성된 설명

텍스트, 폰트, 친필, 화이트이(가) 표시된 사진

자동 생성된 설명

X와 al/aX는 같은 형상이고, W와 al/aW도 같은 형상이다.

도표, 텍스트, 폰트, 라인이(가) 표시된 사진

자동 생성된 설명

행렬 곱의 역전파는 행렬의 대응하는 차원의 원소 수가 일치하도록 곱을 조립하여 구할 수 있다.

**5.6.2 배치용 Affine 계층**

데이터 N개를 묶어 순전파 하는 경우, Affine 계층이라고 부른다.

텍스트, 도표, 폰트, 라인이(가) 표시된 사진

자동 생성된 설명

기존과 다른 부분은 입력인 X의 형상이 (N, 2)가 되었다. 역전파 때는 행렬의 형상에 주의하면 그 이전과 같이 구할 수 있다.

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/5.%EC%98%A4%EC%B0%A8%EC%97%AD%EC%A0%84%ED%8C%8C%EB%B2%95/5.6.2.ipynb>

<https://github.com/sinyebin/DeepLearning-from-Scratch/blob/master/common/layers.py>

**5.6.3 Softmax\_with\_Loss 계층**

텍스트, 도표, 라인, 평면도이(가) 표시된 사진

자동 생성된 설명

입력 이미지가 Affine 계층과 ReLU 계층을 통과하며 변환되고, 마지막 softmax 계층에 의해서 10개의 입력이 정규화 된다. 손글씨 숫자는 가짓수가 10개 이므로 Softmax 계층의 입력은 10개이다.

도표, 평면도, 지도, 라인이(가) 표시된 사진

자동 생성된 설명

Softmax\_with\_loss 계층은 다소 복잡하다.

도표, 평면도, 기술 도면, 라인이(가) 표시된 사진

자동 생성된 설명

Softmax 계층은 입력(a1, a2, a3)를 정규화하여 (y1, y2, y3)를 출력하고, Cross Entropy Error 계층은 Softmax의 출력(y1, y2, y3)와 정답 레이블 (t1, t2, t3)를 받고, 이 데이터들로부터 손실 L을 출력한다. 역전파의 결과는(y1-t1, y2-t2, y3-t3)로 softmax 계층의 출력과 정답 레이블의 차분이다. 신경망의 역전파에서는 이 차이인 오차가 앞 계층에 전해진다.

신경망의 출력과 정답 레이블의 오차를 효율적으로 앞 계층에 전달해야 한다. 앞의 (y1-t1, y2-t2, y3-t3)라는 결과는 바로 softmax 계층의 출력과 정답 레이블의 차이로, 신경망의 현재 출력과 정답 레이블의 오차를 있는 그대로 드러낸다.

**5.7 오차역전파법 구현하기**

**5.7.1 신경망 학습의 전체 그림**

전제: 신경망에는 적응 가능한 가중치와 편향이 있고, 이 가중치와 편향을 훈련 데이터에 적응하도록 조정하는 과정을 ‘학습’이라고 한다.

1단계 – 미니배치: 훈련 데이터 중 일부를 무작위로 가져온다. 이렇게 선별한 데이터를 미니배치라 하며, 그 미니배치의 손실 함수 값을 줄이는 것이 목표

2단계 – 기울기 산출: 미니배치의 손실 함수 값을 줄이기 위해 각 가중치 매개변수의 기울기를 구합니다. 기울기는 손실 함수의 값을 가장 작게 하는 방향을 제시

3단계 – 매개변수 갱신: 가중치 매개변수를 기울기 방향으로 아주 조금 갱신

4단계 – 반복: 1~3단계를 반복

**5.7.2 오차역전파법을 적용한 신경망 구현하기**

신경망 계층을 순서가 있는 딕셔너리에 보관하는 점이 중요

순전파 때는 추가한 순서대로 각 계층의 forward() 메서드를 호출하면 처리 완료

역전파 때는 계층을 반대 순서로 호출

**5.8 정리**

계산 그래프를 이용하면 계산 과정을 시각적으로 파악 가능

계산 그래프의 노드는 국소적 계산으로 구성, 국소적 계산을 조합해 전체 계산을 구성

계산 그래프의 순전파는 통상의 계산을 수행, 한편 계산 그래프의 역전파로는 각 노드의 미분을 구할 수 있음

신경망의 구성 요소를 계층으로 구현하여 기울기를 효율적으로 계산할 수 있음

수치 미분과 오차역전파법의 결과를 비교하면 오차역전파법의 구현에 잘못이 없는지 확인 가능