Derivation of Backpropagation

DS School

https://dsschool.co.kr

Derivation of Backpropagation

Multi-layer neural network에서는
한 층의 weight가 아닌 여러 층의 weight에 대한 변화량을 구해야 한다

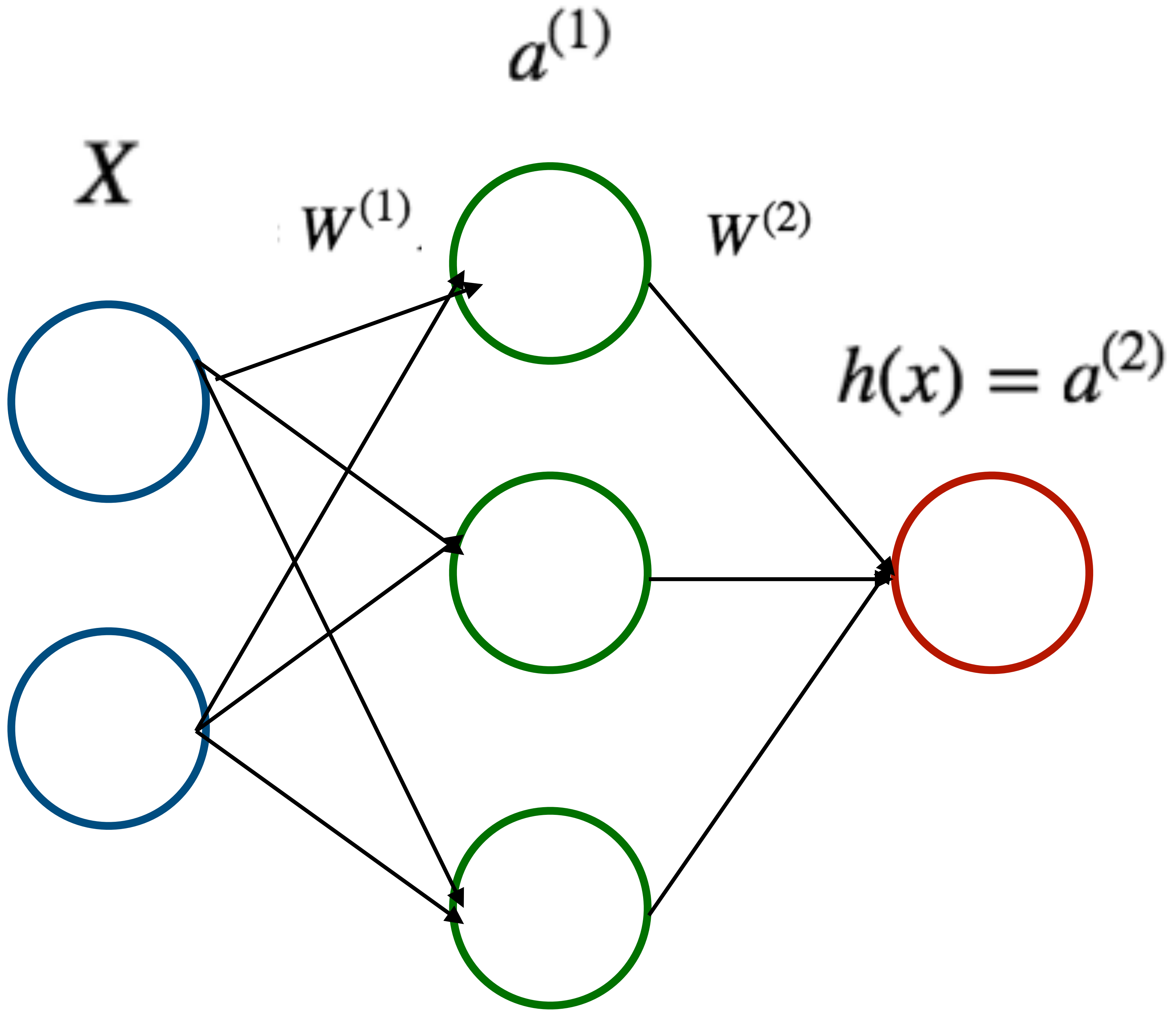
$$z^{(1)} = W^{(1)}X + b^{(1)}$$

$$a^{(1)} = sigmoid(z^{(1)})$$

$$z^{(2)} = W^{(2)}a^{(1)} + b^{(2)}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$h(x) = a^{(2)}$$



$$L(x) = \frac{1}{m}\sum_{i=0}^{m}\left(-y^{(i)}\log\left(h(x^{(i)})\right) - (1-y^{(i)})\log\left(1-h(x^{(i)})\right)\right)$$

Derivation of Backpropagation

Multi-layer neural network에서는
한 층의 weight가 아닌 여러 층의 weight에 대한 변화량을 구해야 한다

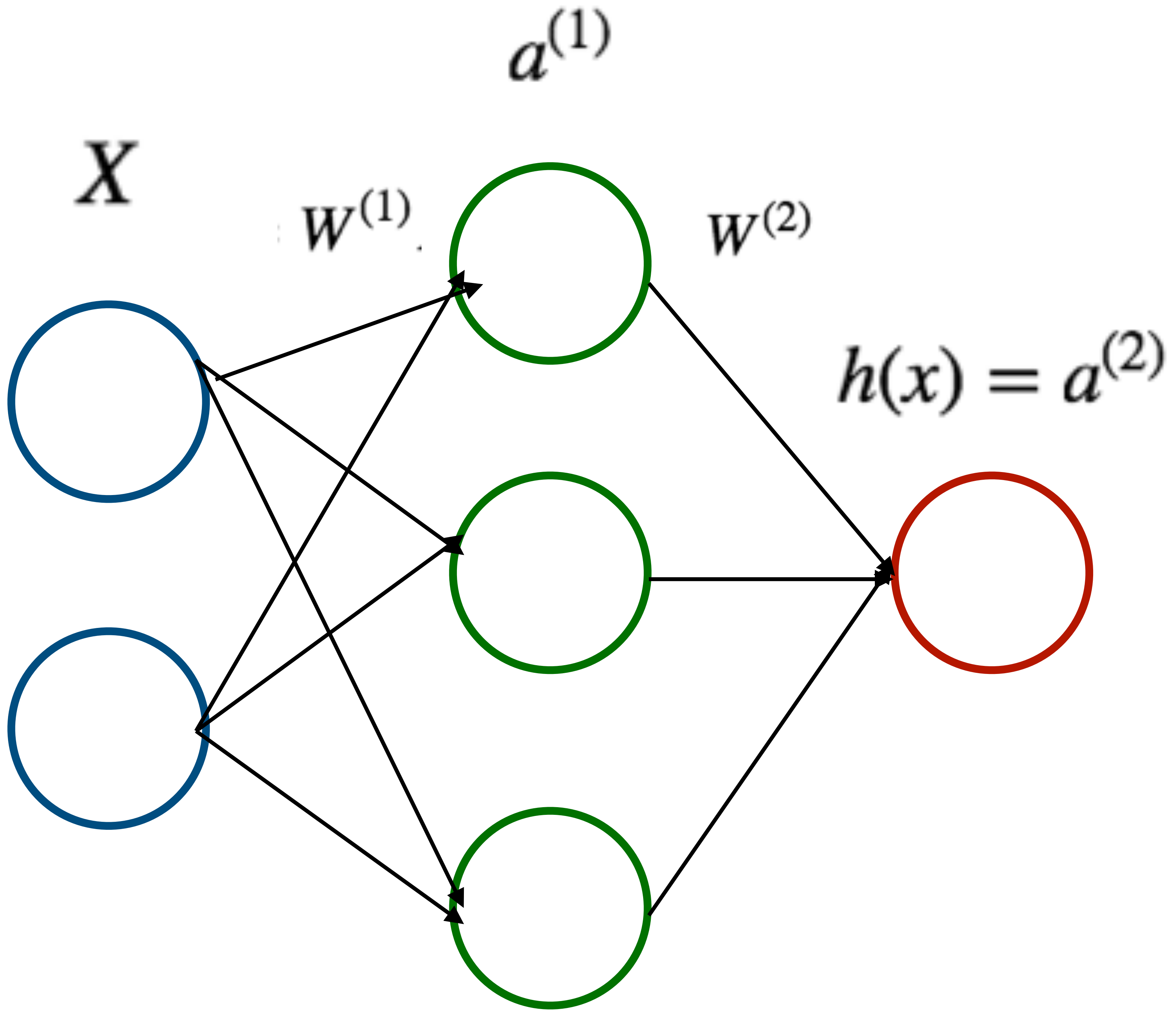
$$z^{(1)} = W^{(1)}X + b^{(1)}$$

$$a^{(1)} = sigmoid(z^{(1)})$$

$$z^{(2)} = W^{(2)}a^{(1)} + b^{(2)}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$h(x) = a^{(2)}$$

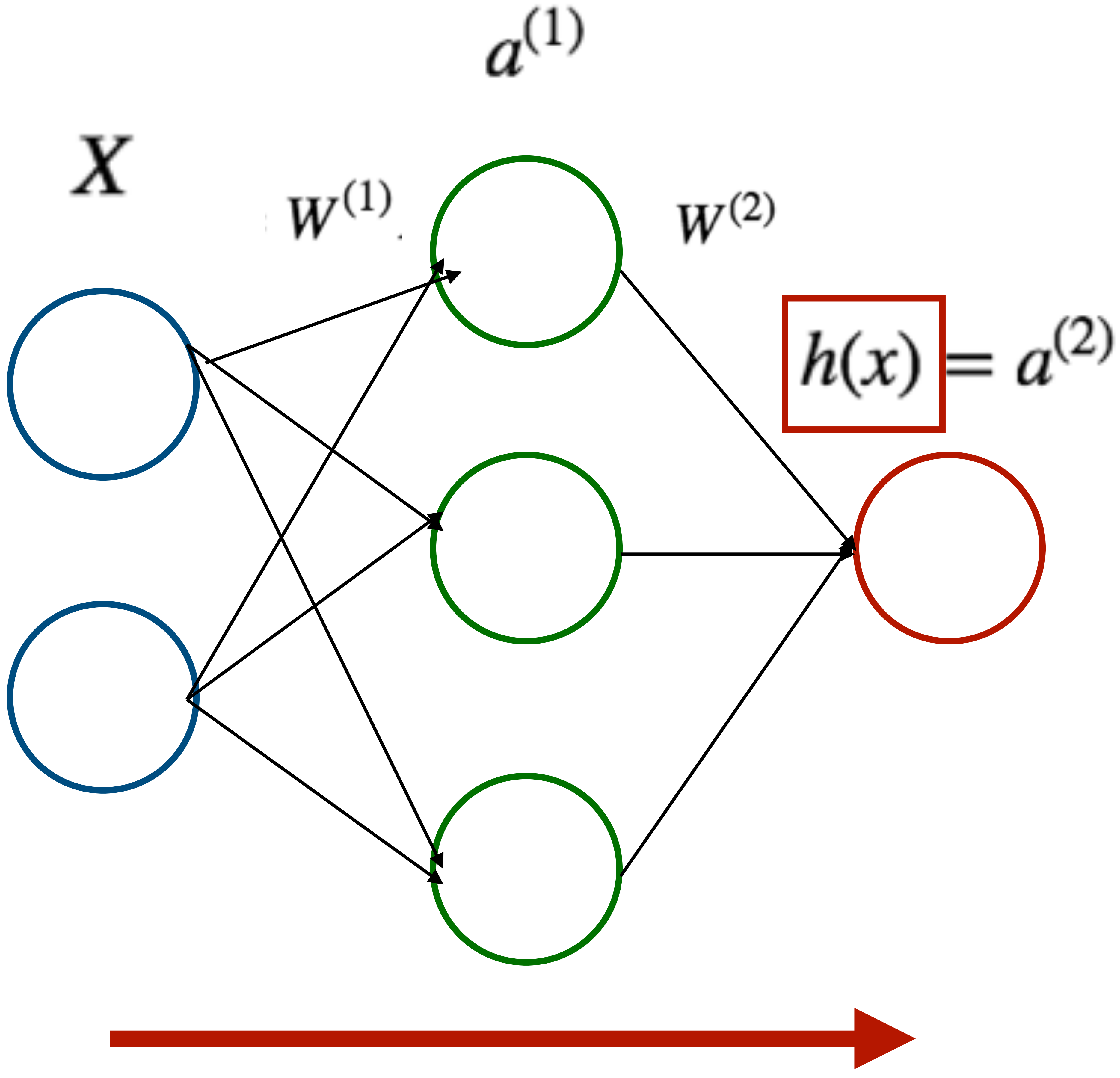


$$L(x) = \frac{1}{m}\sum_{i=0}^{m}\left(-y^{(i)}\log\left(h(x^{(i)})\right) - (1-y^{(i)})\log\left(1-h(x^{(i)})\right)\right)$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = ? \qquad \frac{\partial L(x)}{\partial W^{(1)}} = ?$$

Derivation of Backpropagation

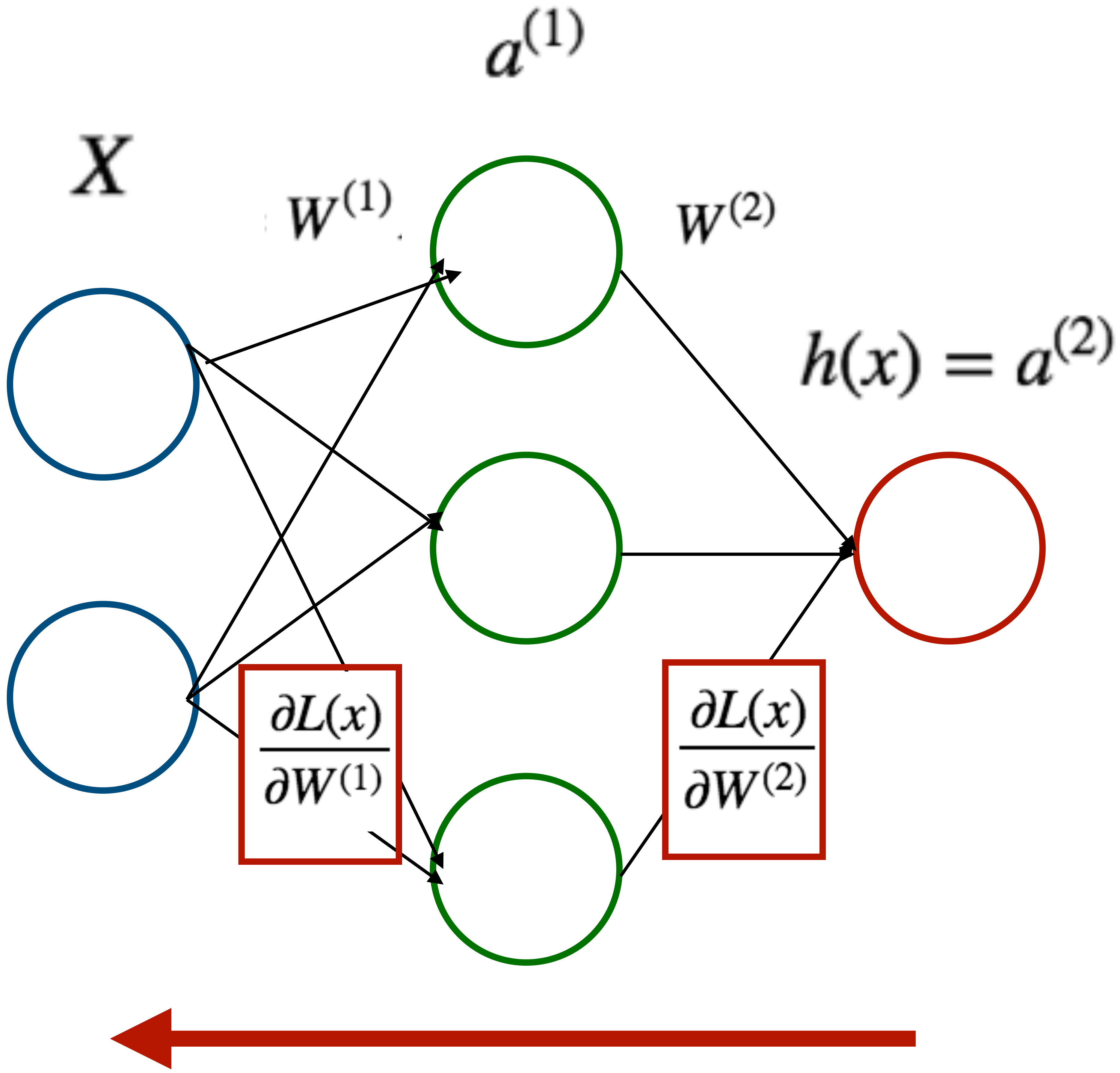
Neural Network에서 예측을 할 때는 언제나 input에서 output으로 연산을 하는데
이와 정 반대로 output에서 input으로 하나하나 연산을 하면서 weight의 변화량을 구하는 것을 back propagation 이라고 한다



input(X)을 가지고 output까지 연산하는 과정을 forward propagation 이라고 하며,
forward propagation 결과를 통해 우리는 예측값 h(x) 를 구할 수 있다

Derivation of Backpropagation

Neural Network에서 예측을 할 때는 언제나 input에서 output으로 연산을 하는데
이와 정 반대로 output에서 input으로 하나하나 연산을 하면서 weight의 변화량을 구하는 것을 back propagation 이라고 한다

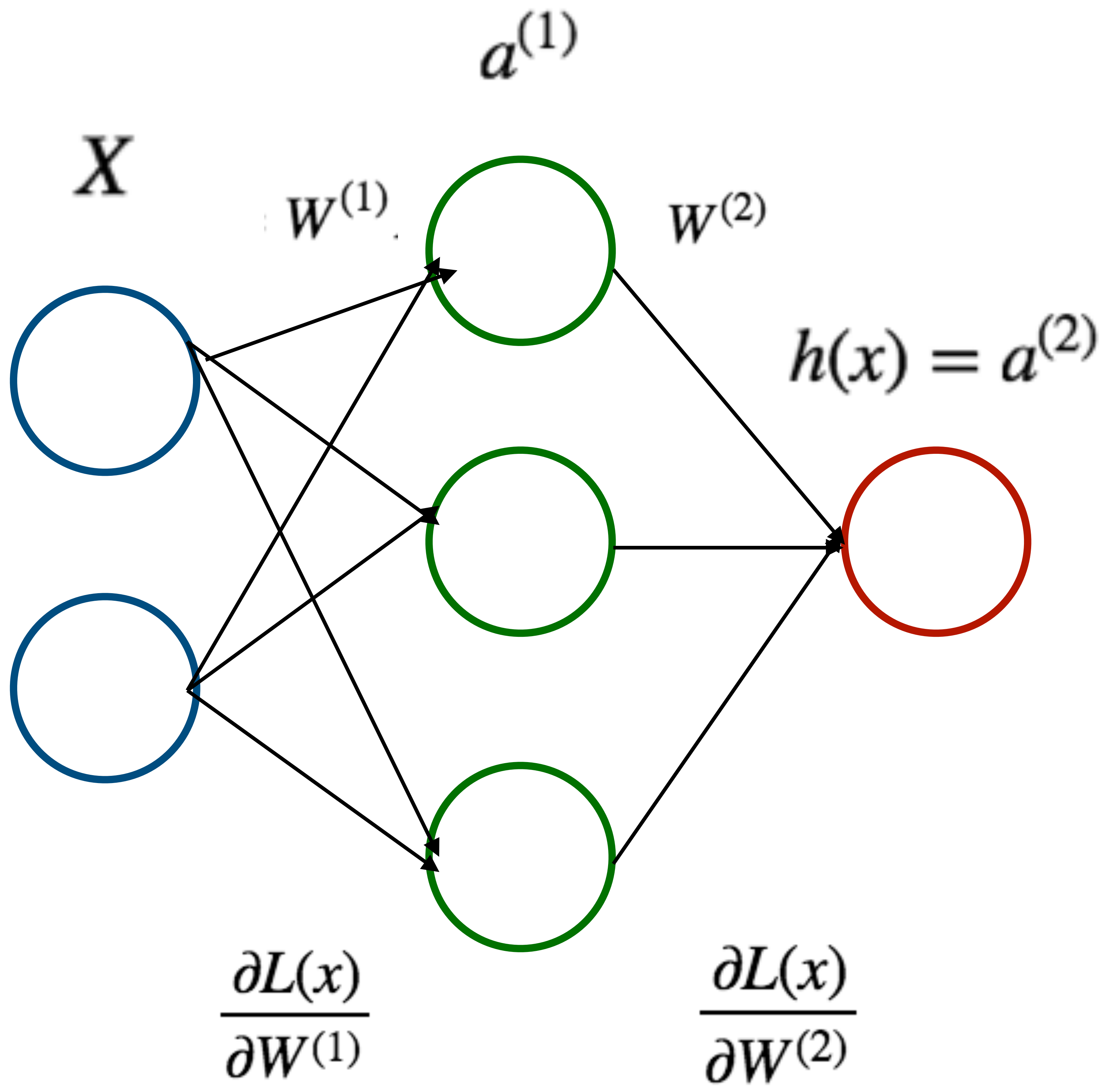


그리고 정 반대 방식으로, 예측값 h(x)와 정답 y를 가지고
output에서 input으로 연산을 해 나가면 각 레이어의 weight의 변화량을 구할 수 있다.
이를 전문용어로 back propagation 이라고 한다.

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} =$$

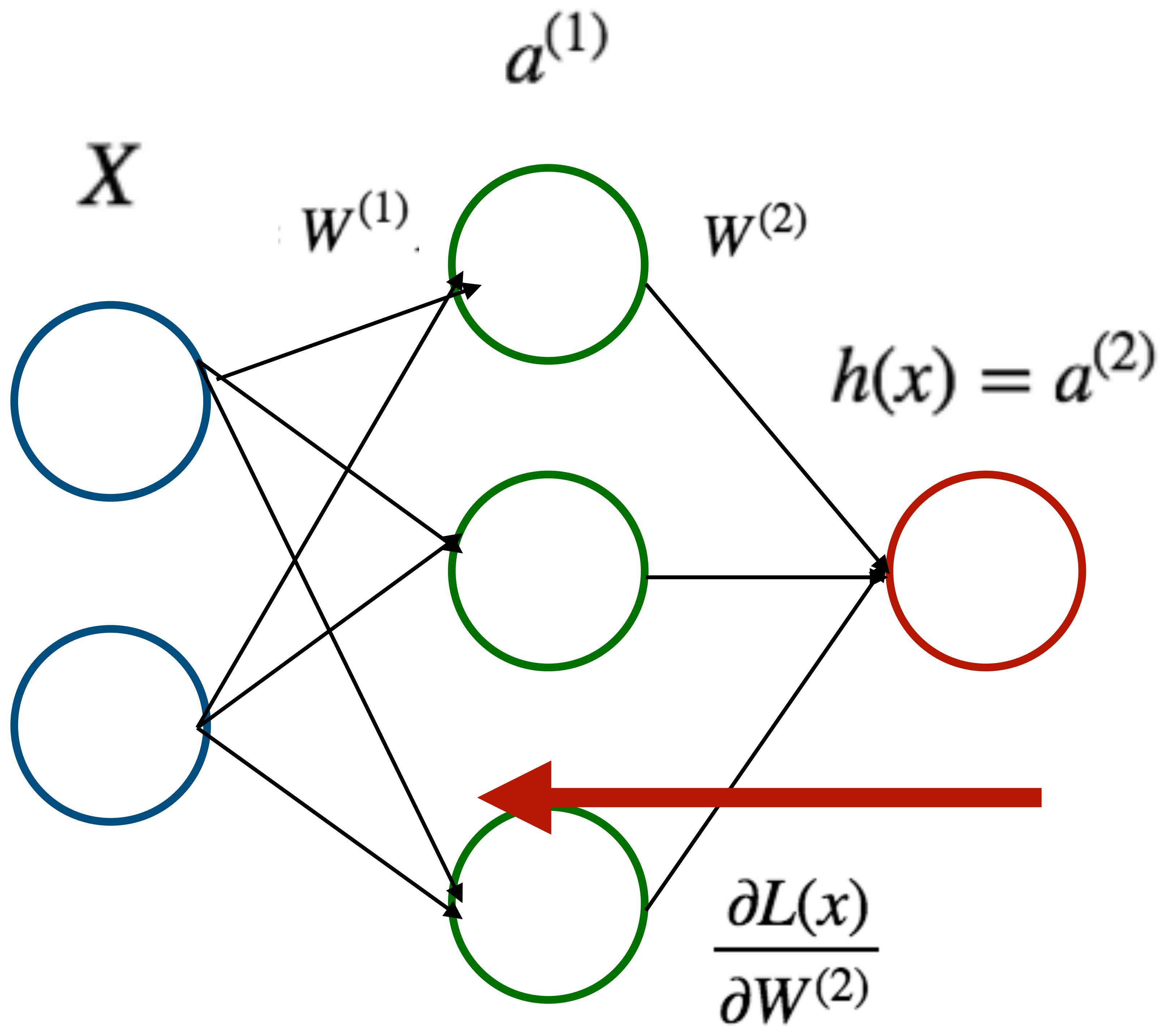


Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, **chain rule**을 활용해 여러 개로 쪼갤 수 있다
그래서 **Loss Function**의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 **backpropation**의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial W^{(2)}}}$$

합성함수의 미분(chain rule)을 활용해
두 개로 나눌 수 있다.

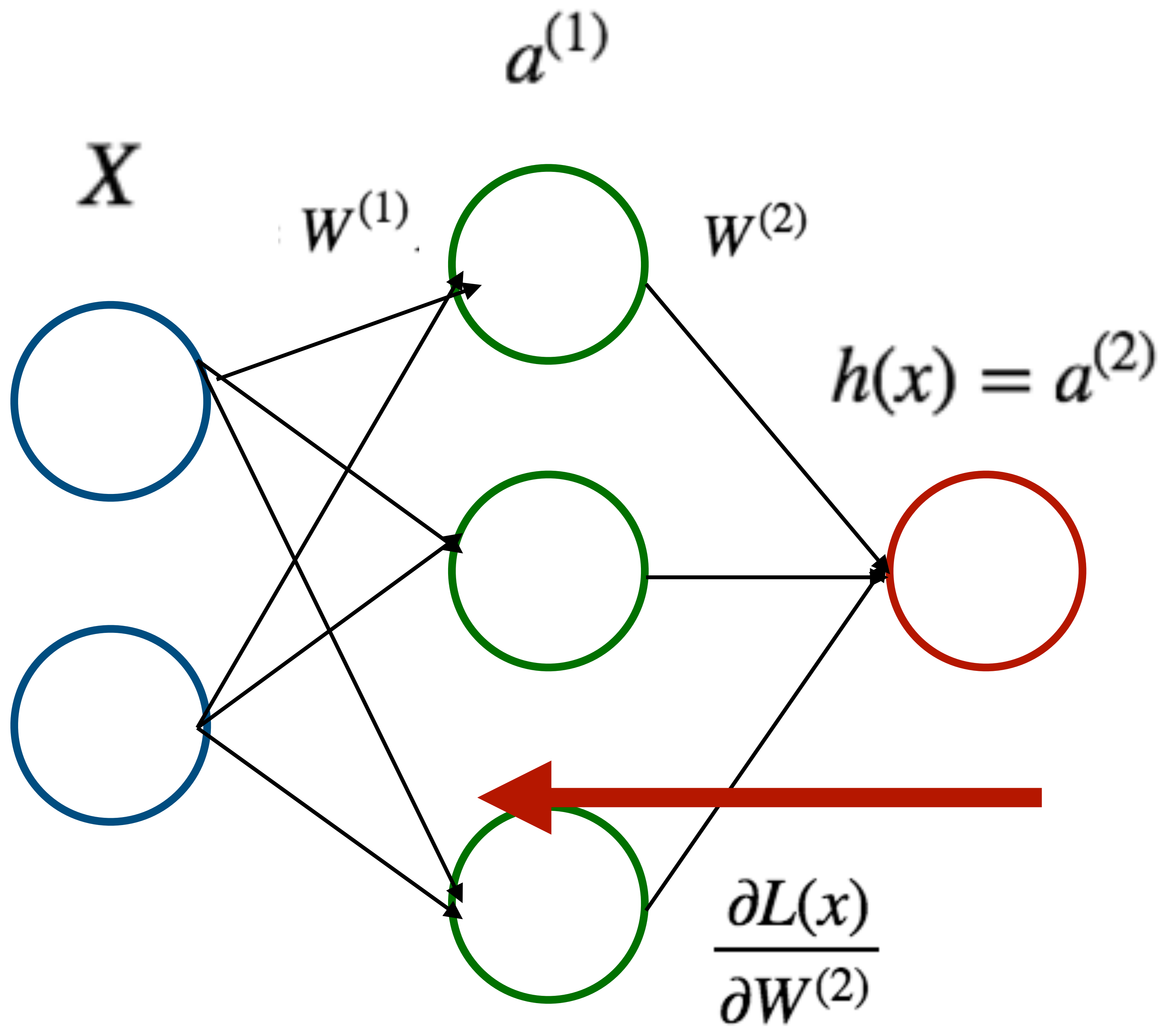


Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial W^{(2)}}$$
$$= \boxed{\frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}}$$

합성함수의 미분을 한 번 더 활용하면
세 개로 나눌 수 있다.

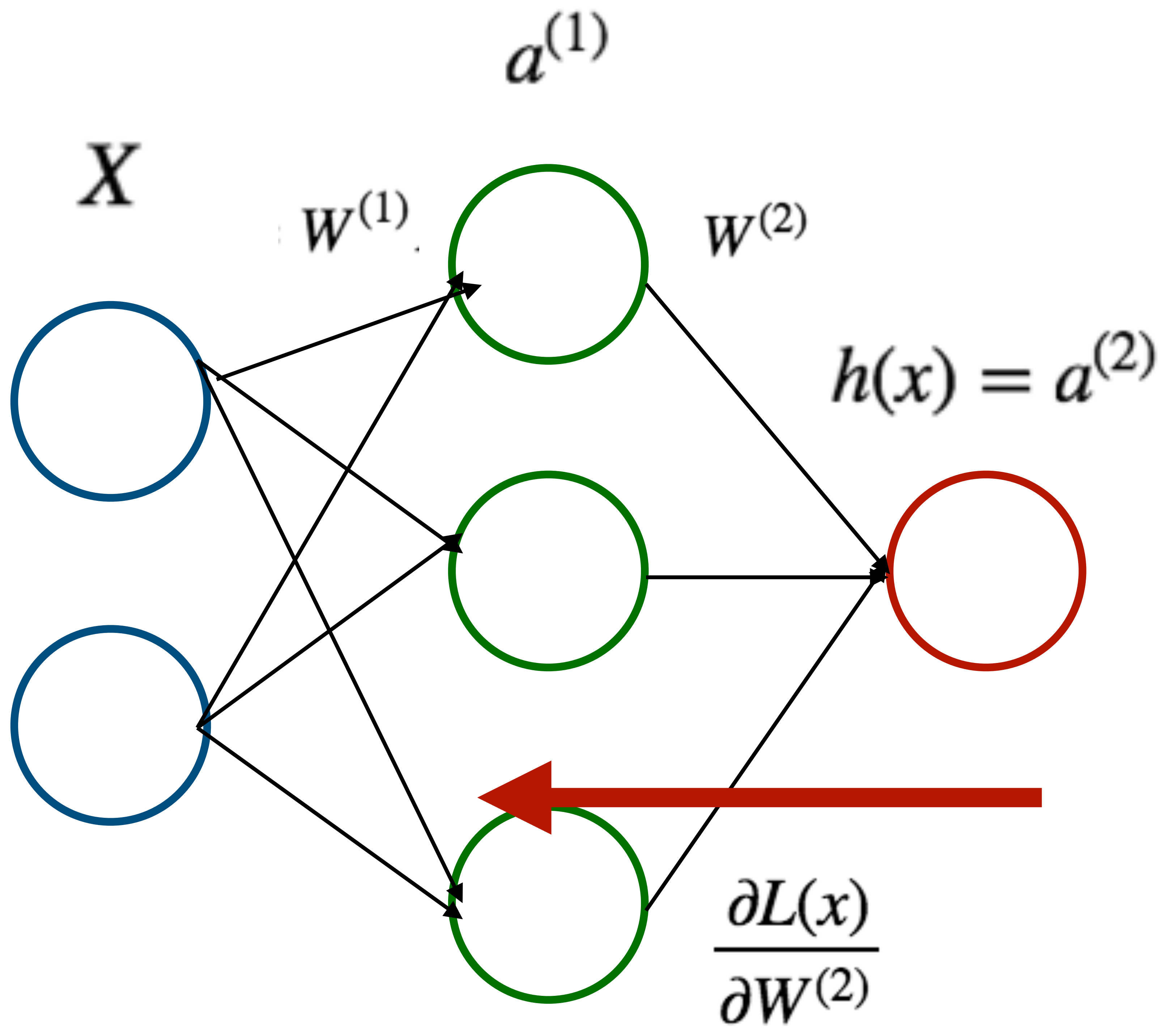


Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial W^{(2)}}$$
$$= \boxed{\frac{\partial L(x)}{\partial a^{(2)}}}\boxed{\frac{\partial a^{(2)}}{\partial z^{(2)}}}\boxed{\frac{\partial z^{(2)}}{\partial W^{(2)}}}$$

세 개를 각각 구한 뒤 곱하면
하나를 구하는 것 보다 훨씬 더 편하게 변화량을 구할 수 있다

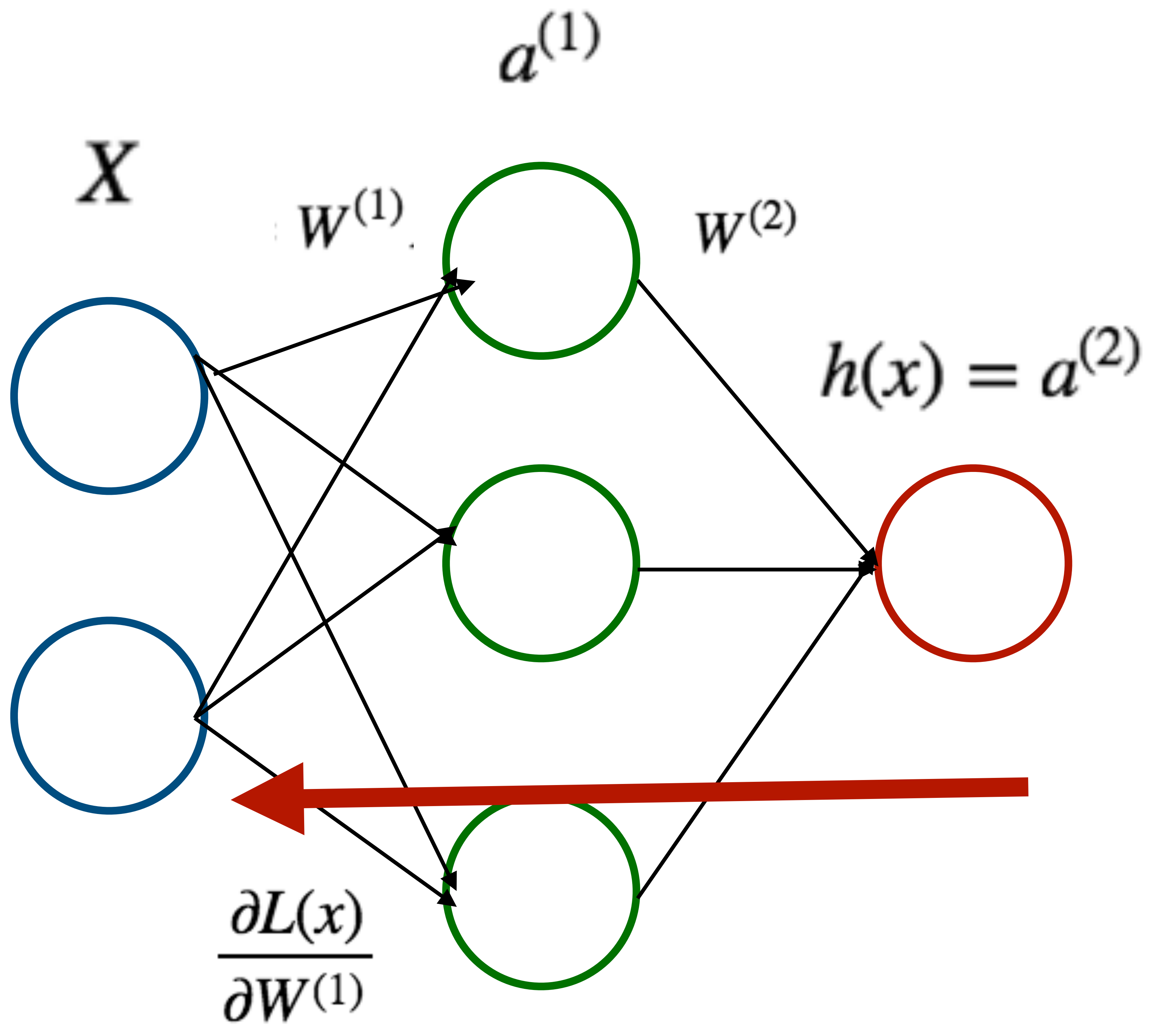


Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}\frac{\partial L(x)}{\partial W^{(1)}} &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial W^{(1)}} \\ &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(1)}} \\ &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial a^{(1)}}\frac{\partial a^{(1)}}{\partial W^{(1)}} \\ &= \boxed{\frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial a^{(1)}}\frac{\partial a^{(1)}}{\partial z^{(1)}}\frac{\partial z^{(1)}}{\partial W^{(1)}}}\end{aligned}$$



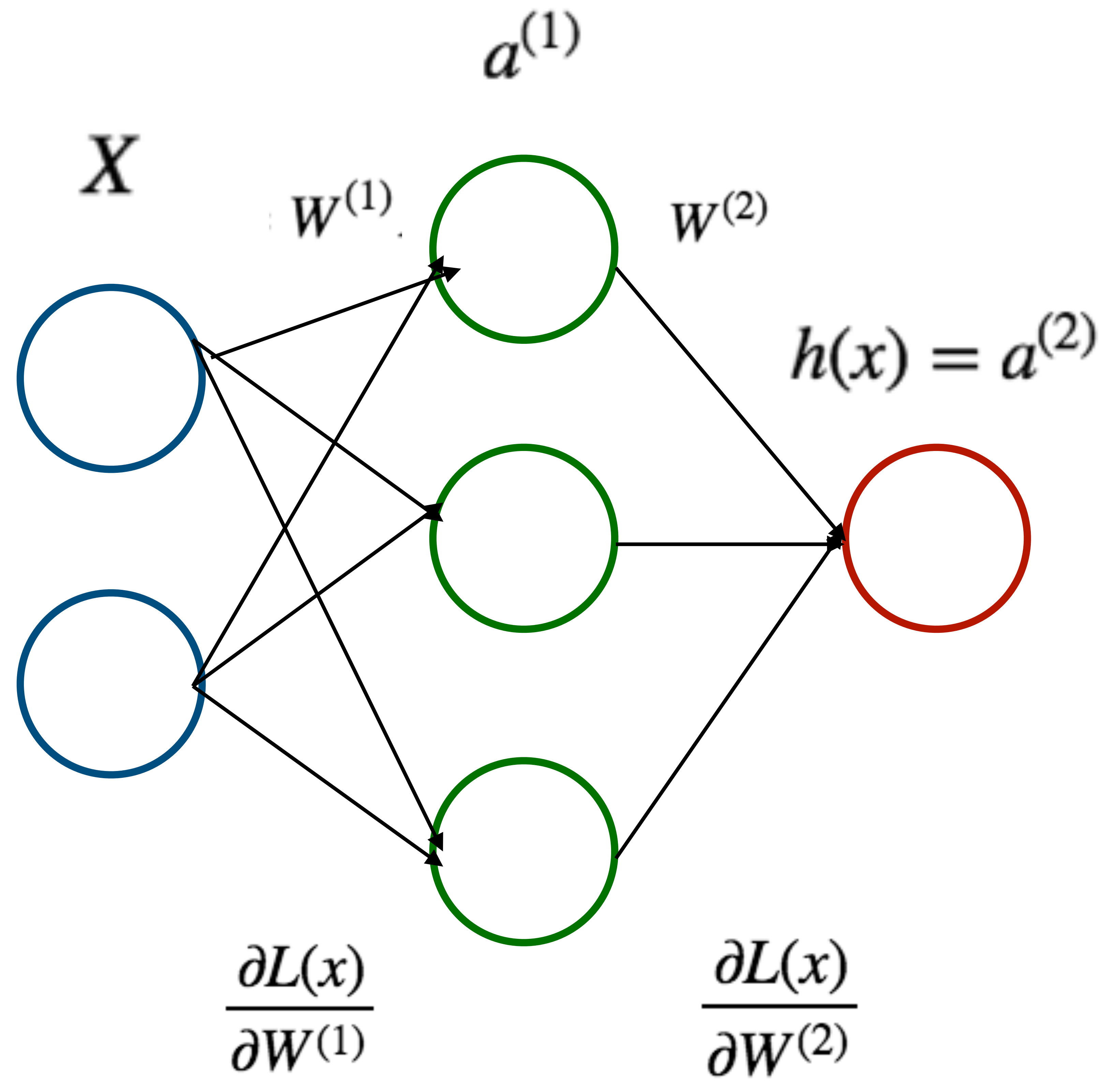
비슷한 방식을 하나 앞의 레이어에도 적용 가능하다

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}}\frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}\frac{\partial L(x)}{\partial W^{(1)}} &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial W^{(1)}} \\ &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(1)}} \\ &= \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial a^{(1)}}\frac{\partial a^{(1)}}{\partial W^{(1)}} \\ &= \boxed{\frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}}\frac{\partial z^{(2)}}{\partial a^{(1)}}\frac{\partial a^{(1)}}{\partial z^{(1)}}\frac{\partial z^{(1)}}{\partial W^{(1)}}\end{aligned}$$



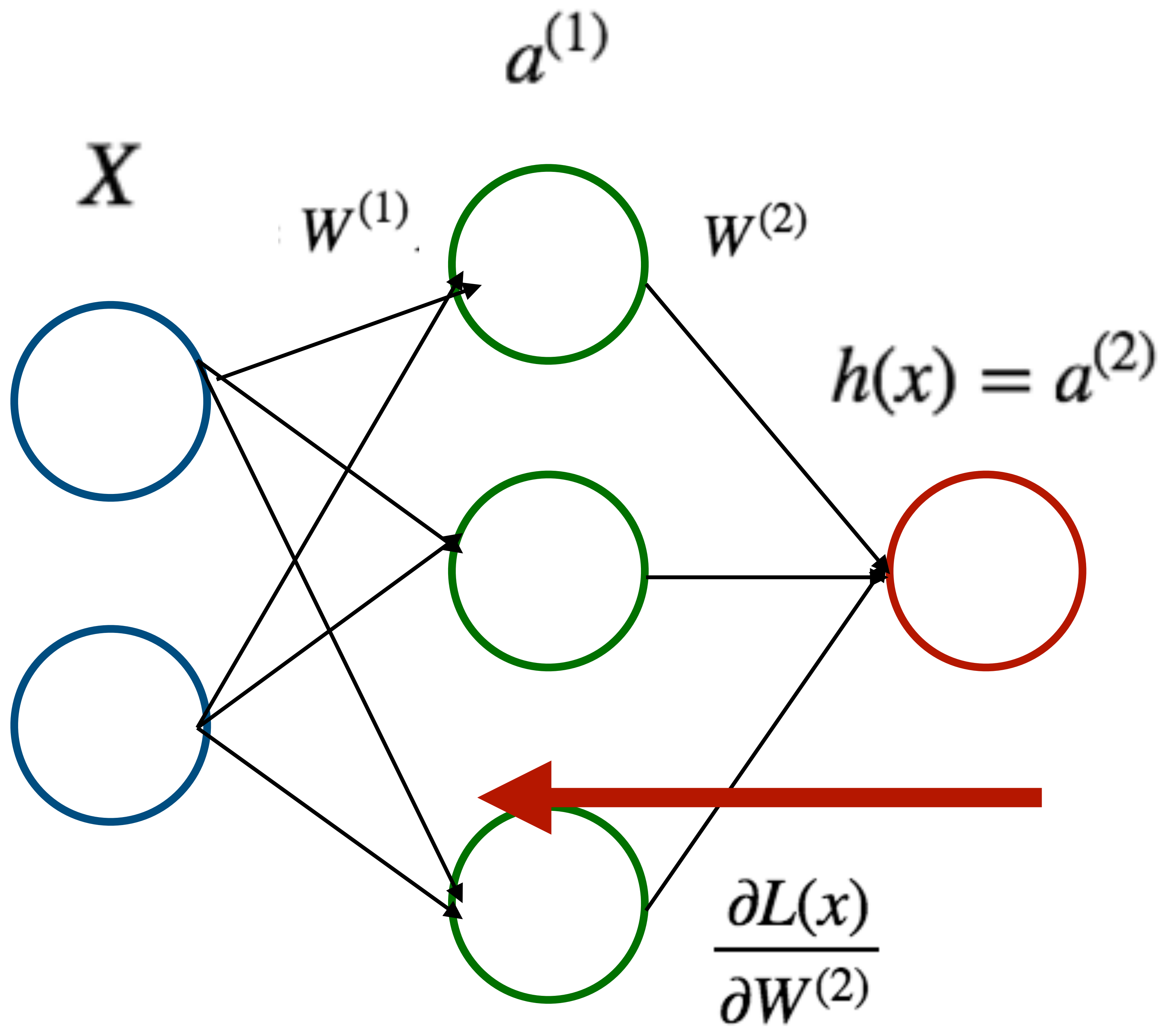
또한 W2를 구할 때 이미 계산한 부분을 W1을 구할 때도 활용할 수 있다.
그러므로 계산을 효율적으로 할 수 있는, 일종의 캐싱(caching) 효과가 일어난다.

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial W^{(2)}}$$
$$= \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

다시 되돌아가서,
위 세 부분을 하나하나 구한 뒤 곱해서 하나로 합쳐보도록 하자

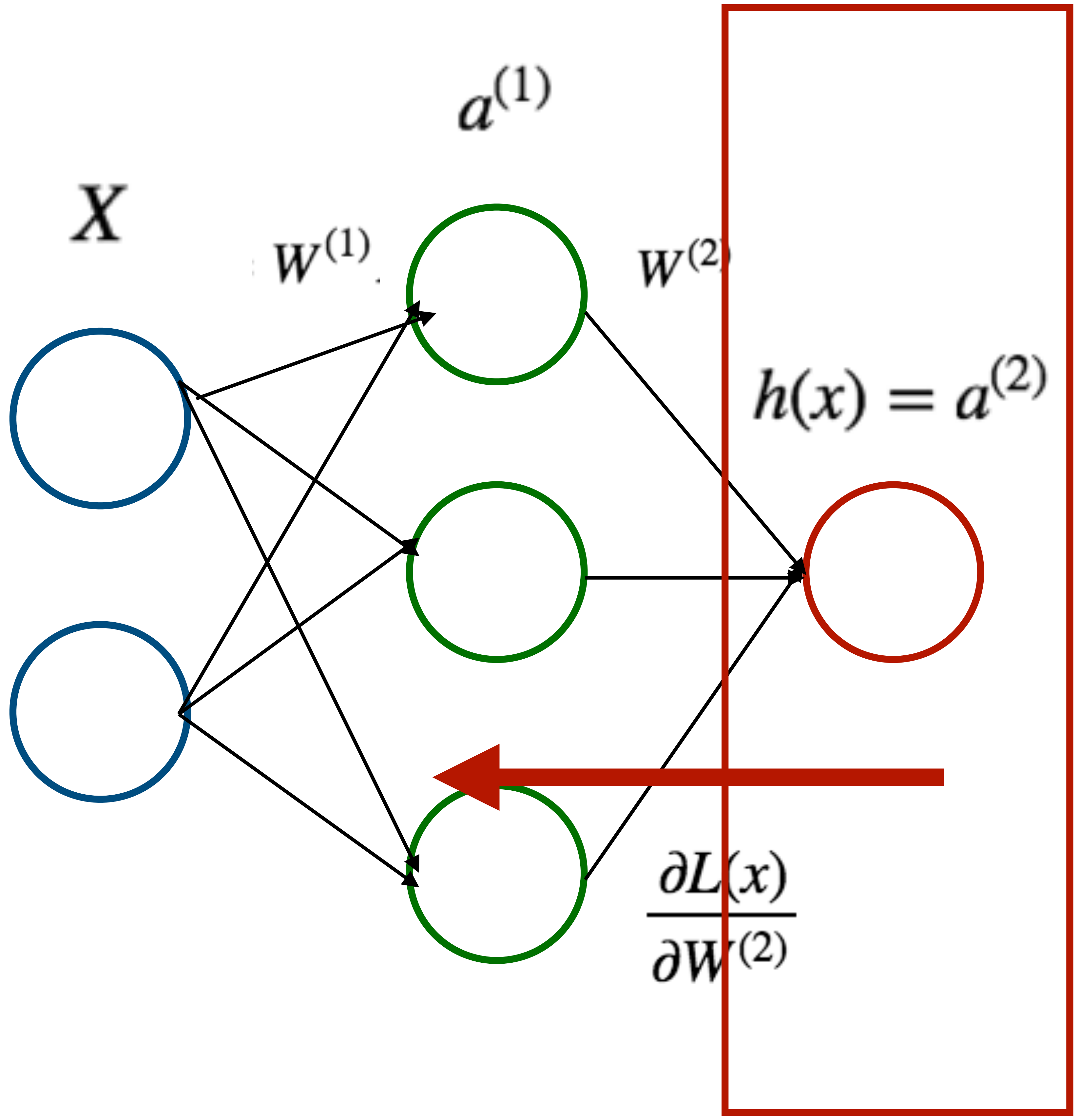


Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = \frac{1}{m} \sum_{i=0}^{m} \left(-y^{(i)} \log\left(h(x^{(i)})\right) - (1 - y^{(i)}) \log\left(1 - h(x^{(i)})\right) \right)$$



시각화하자면 이 부분

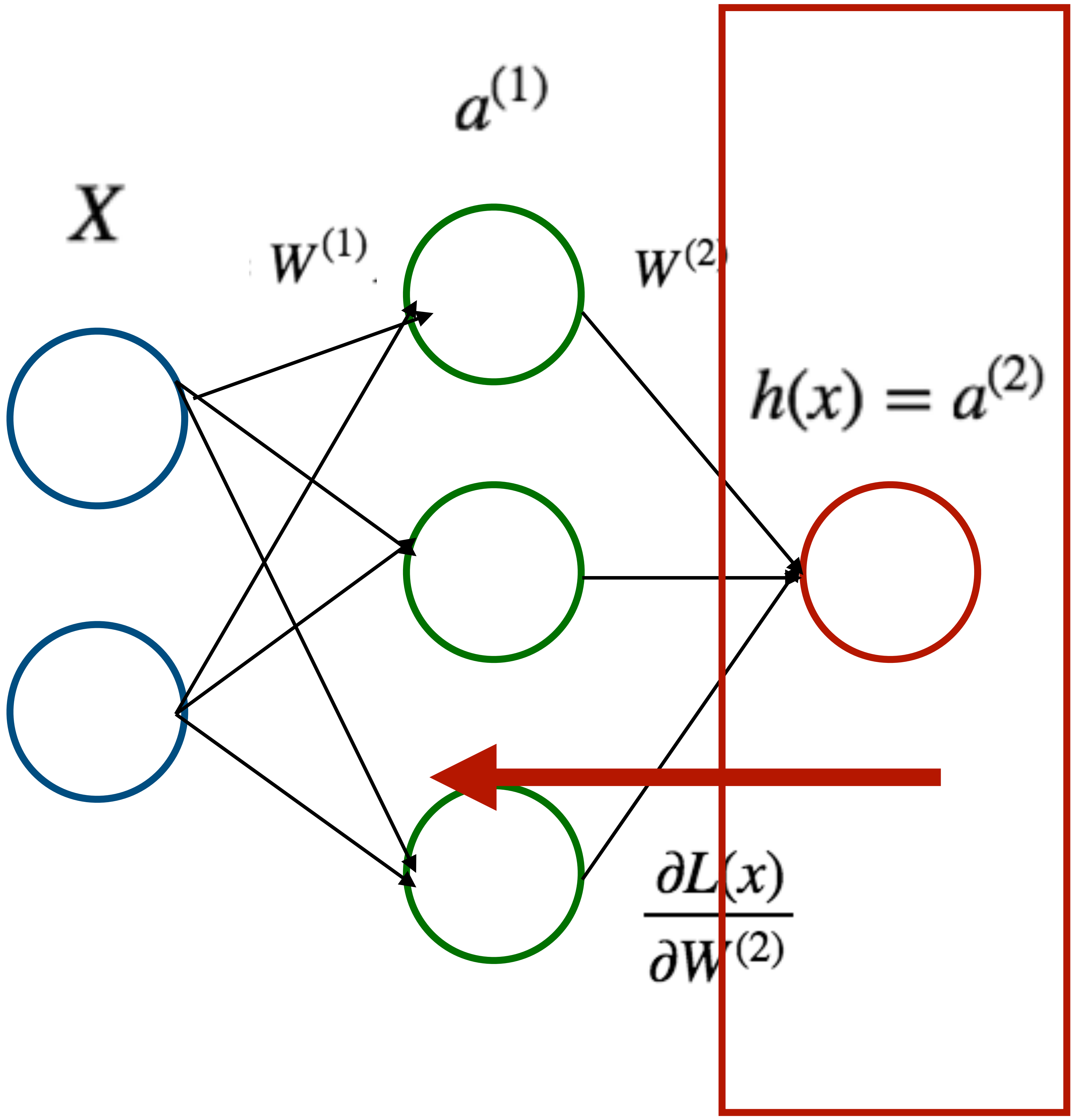
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = \boxed{\frac{1}{m}\sum_{i=0}^{m}} \left(-y^{(i)} \log\left(h(x^{(i)})\right) - (1-y^{(i)}) \log\left(1-h(x^{(i)})\right) \right)$$

변함없이 이 부분은 평균(mean)을 나타내므로
우리가 편미분 유도를 이해하는데 중요하지 않다



시각화하자면 이 부분

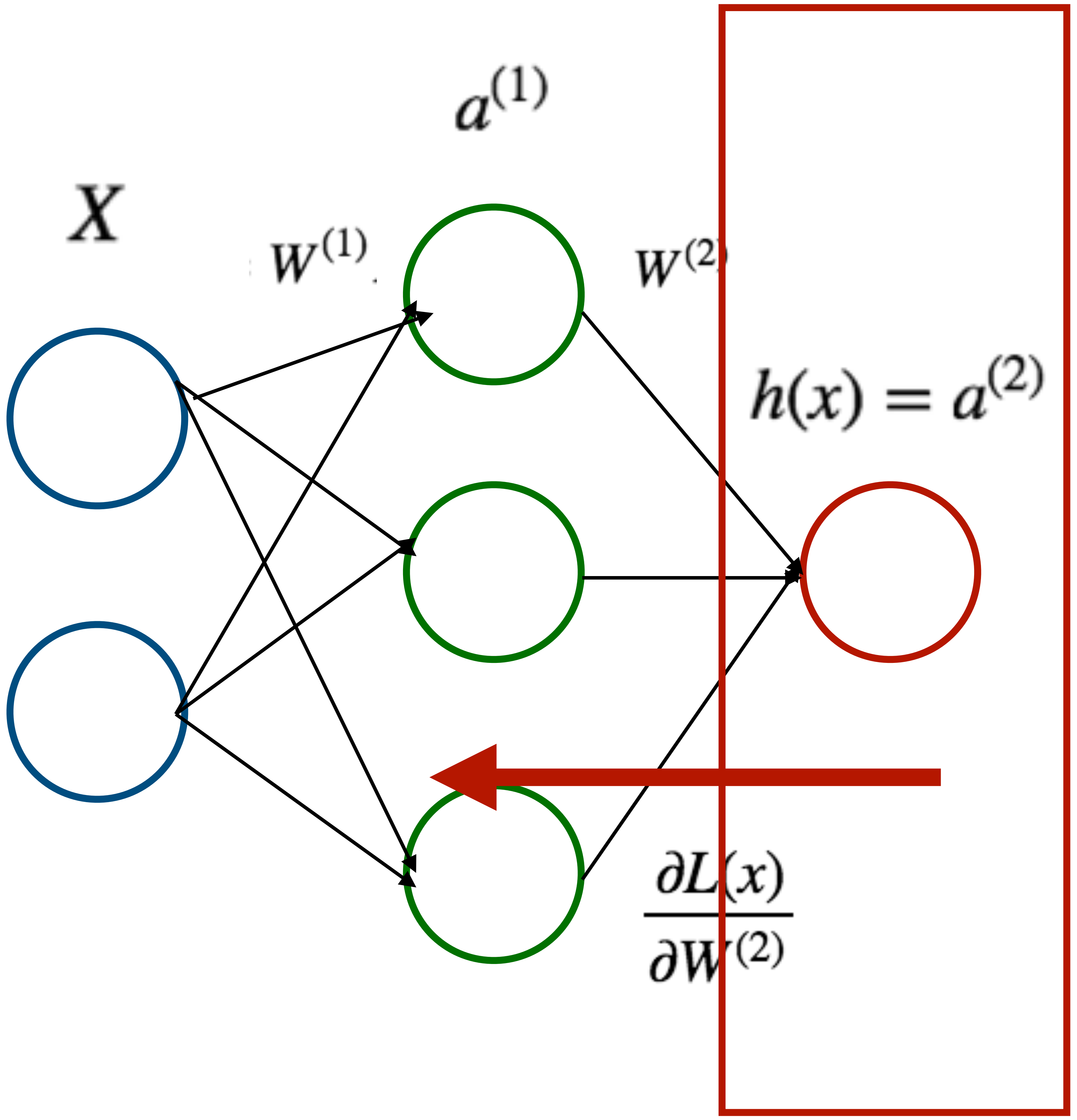
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = \frac{1}{m} \sum_{i=0}^{m} \left(-y^{(i)} \log \left(h(x^{(i)}) \right) - (1 - y^{(i)}) \log \left(1 - h(x^{(i)}) \right) \right)$$

$$L(x) = -y \log h(x) - (1 - y) \log \left(1 - h(x) \right)$$



시각화하자면 이 부분

Derivation of Backpropagation

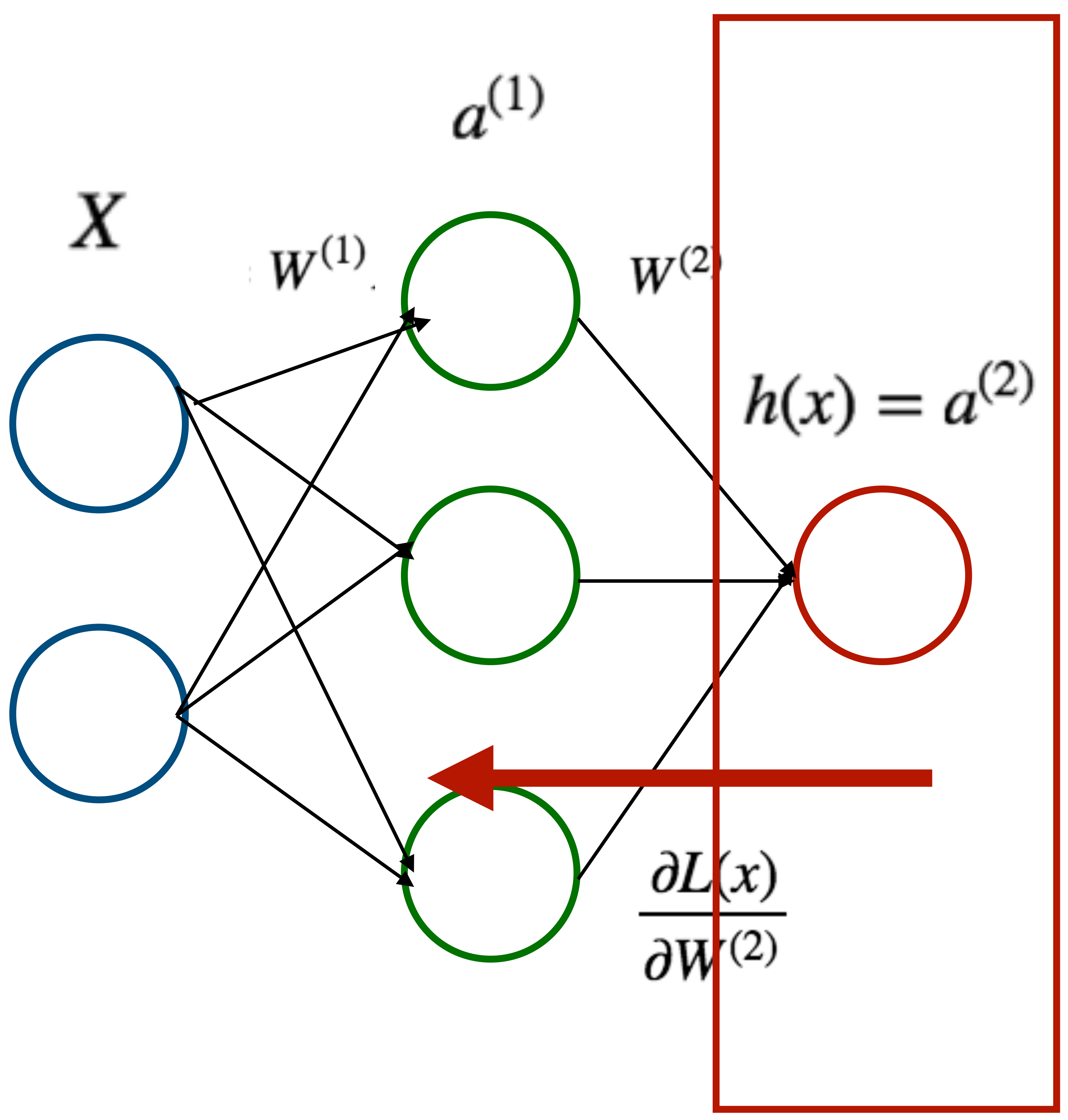
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = \frac{1}{m}\sum_{i=0}^{m}\left(-y^{(i)}\log\left(h(x^{(i)})\right) - (1-y^{(i)})\log\left(1-h(x^{(i)})\right)\right)$$

$$L(x) = -y\log \boxed{h(x)} - (1-y)\log\left(1-h(x)\right)$$

공식상으로 h(x)와 a2는 동일하기 때문에
a2로 바꿔줄 수 있다



시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

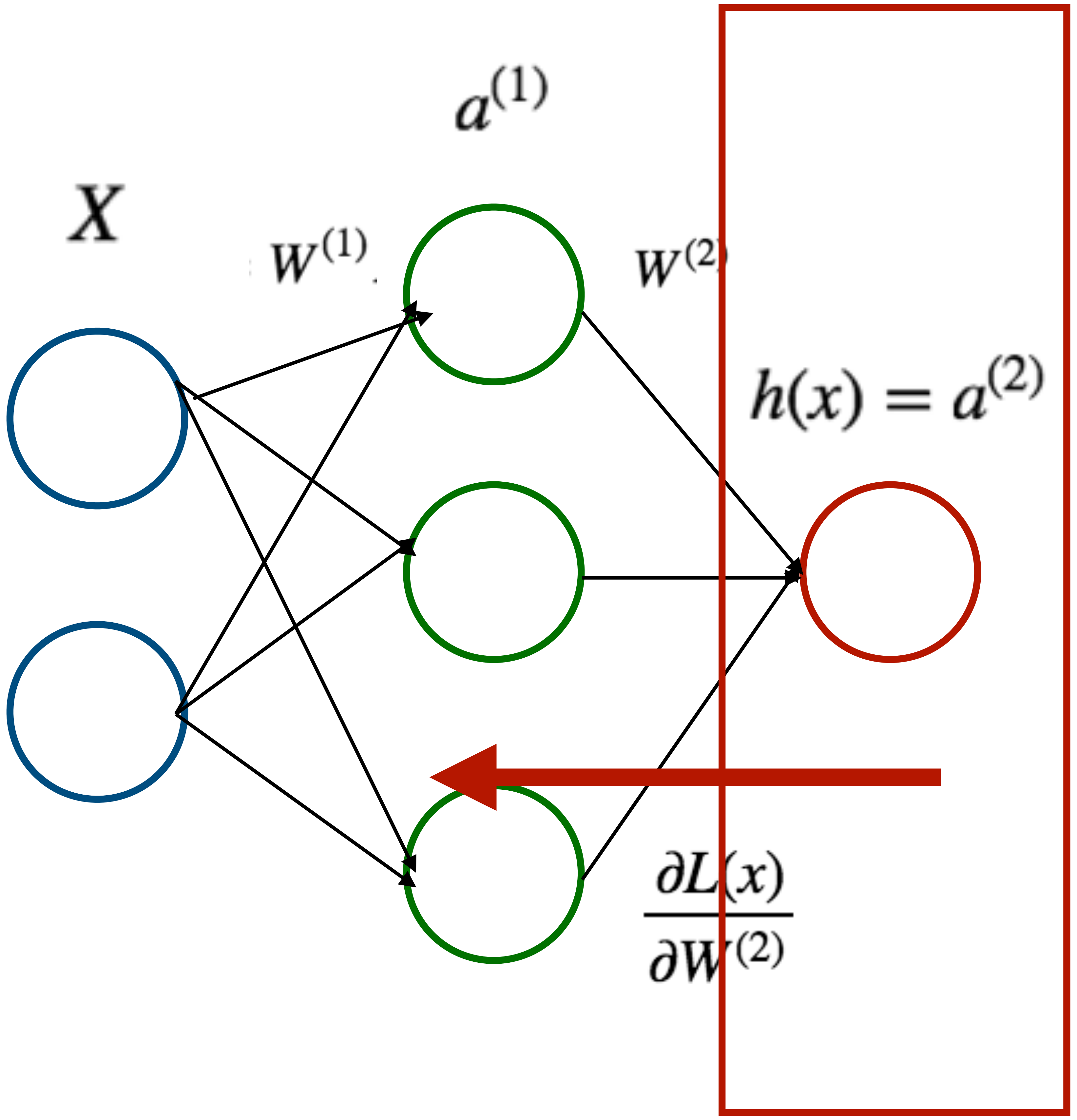
$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = \frac{1}{m}\sum_{i=0}^{m}\left(-y^{(i)}\log\left(h(x^{(i)})\right) - (1-y^{(i)})\log\left(1-h(x^{(i)})\right)\right)$$

$$L(x) = -y\log h(x) - (1-y)\log\left(1-h(x)\right)$$

$$L(x) = -y\log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)$$

훨씬 공식을 유도하기 쉽도록 변경하였다



시각화하자면 이 부분

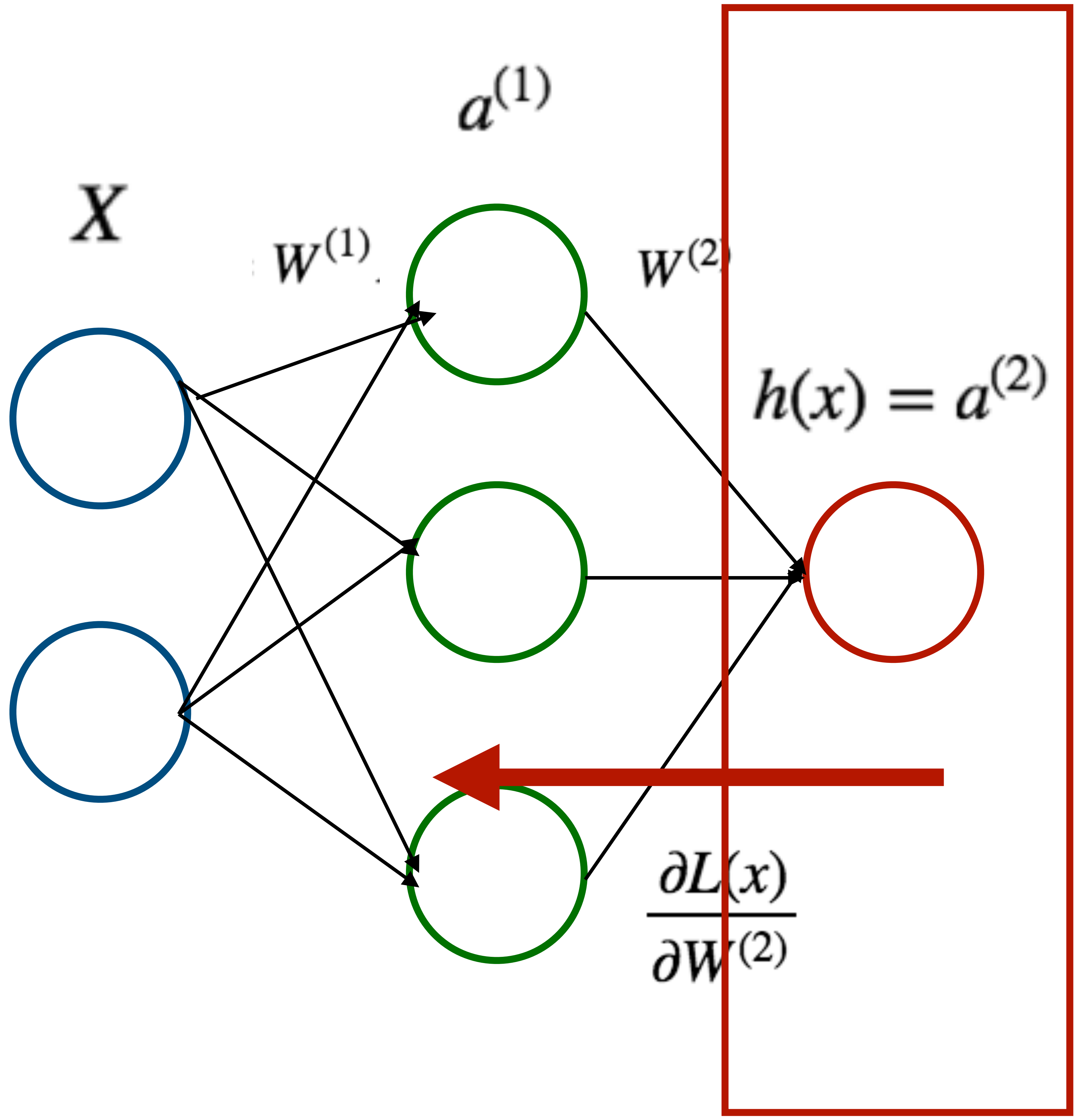
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y)\log\left(1 - a^{(2)}\right)$$

$$\frac{\partial L(x)}{\partial a^{(2)}} =$$



시각화하자면 이 부분

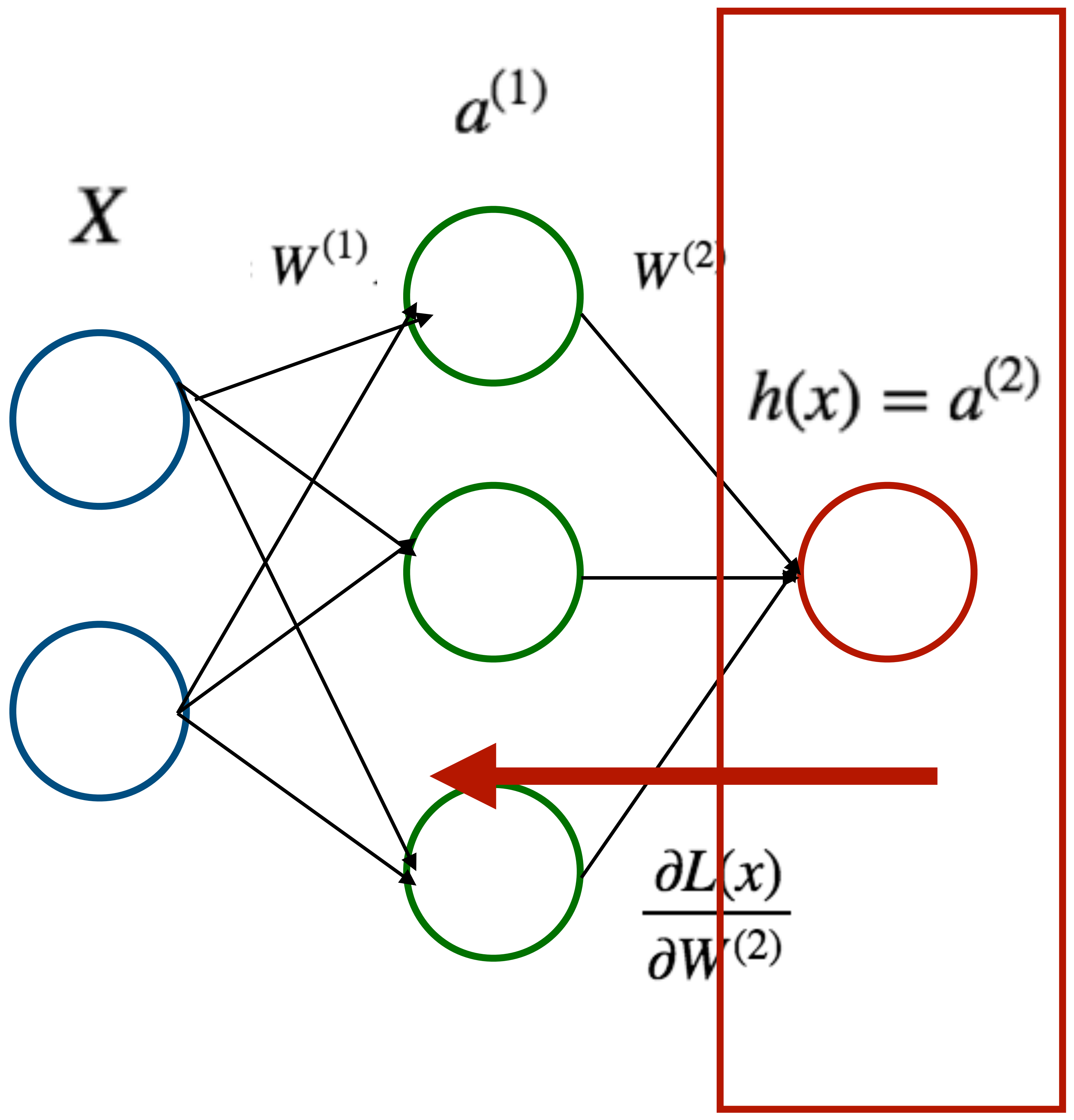
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)\right)$$



시각화하자면 이 부분

Derivation of Backpropagation

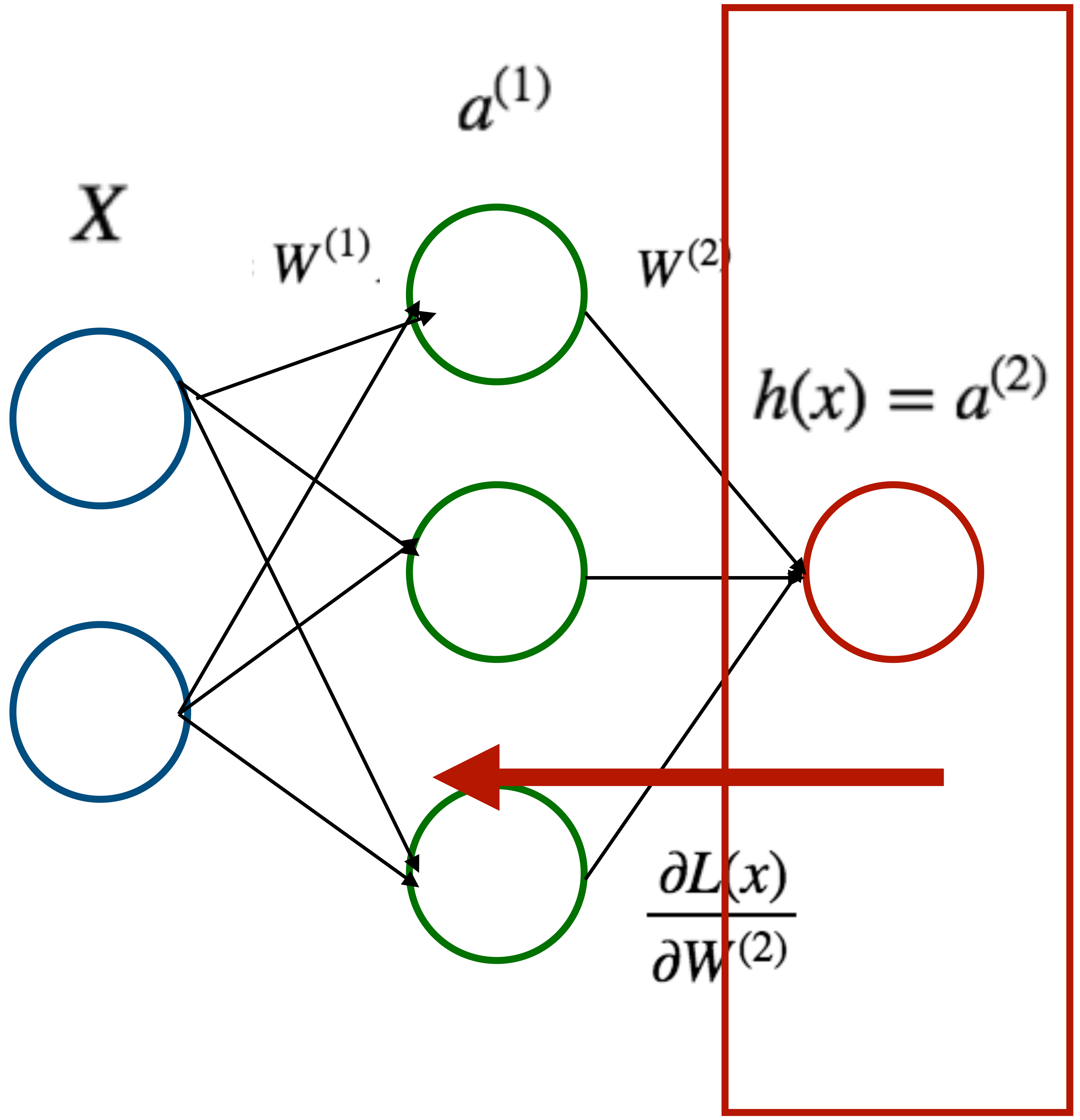
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right) \right)$$

편미분 기호는 괄호 안으로 넣을 수 있다



시각화하자면 이 부분

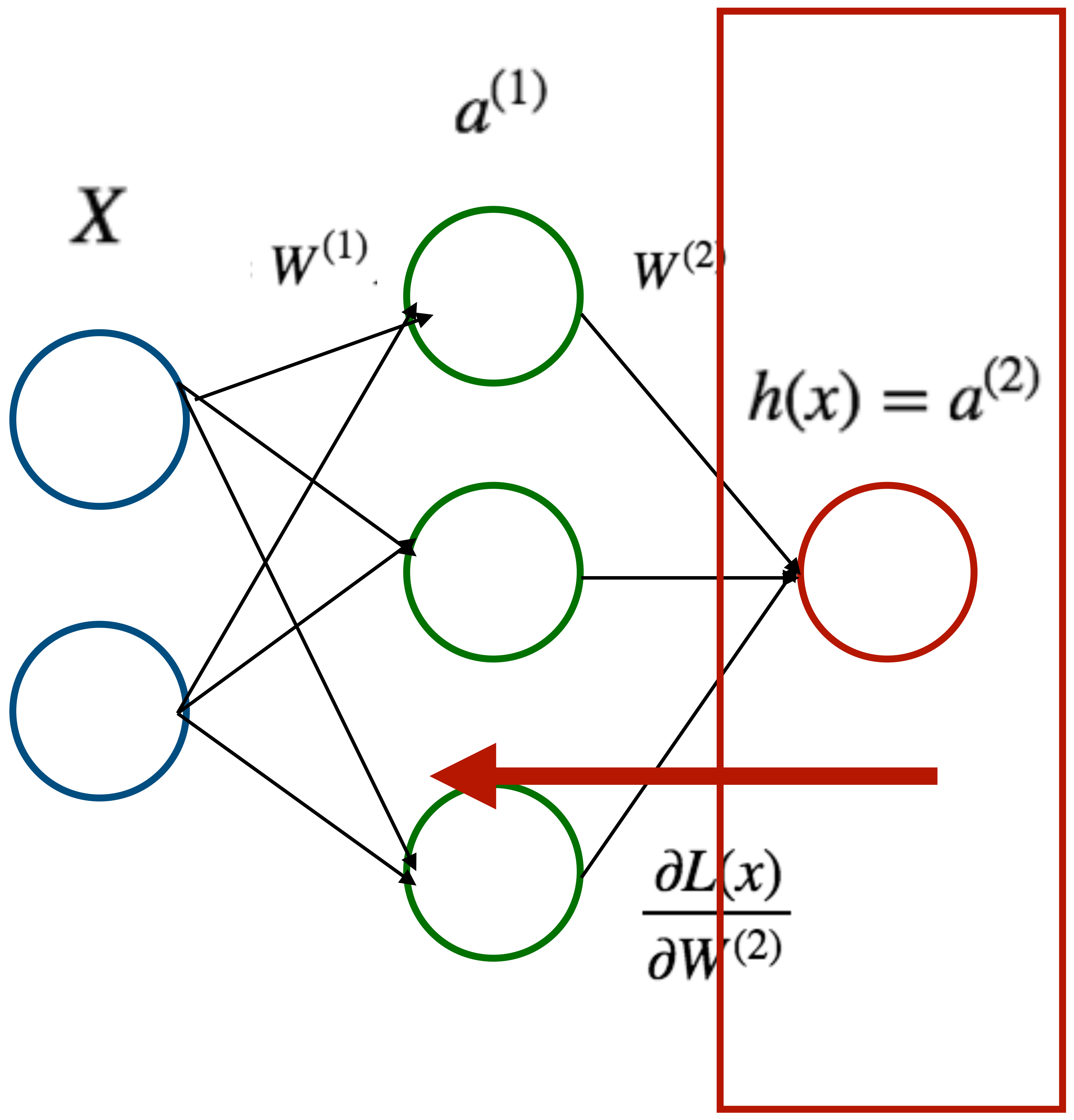
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log\left(1 - a^{(2)}\right)$$

$$\begin{aligned}\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}}\left(-y \log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right) \\ &= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}}(1-y)\log\left(1-a^{(2)}\right)\end{aligned}$$



시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

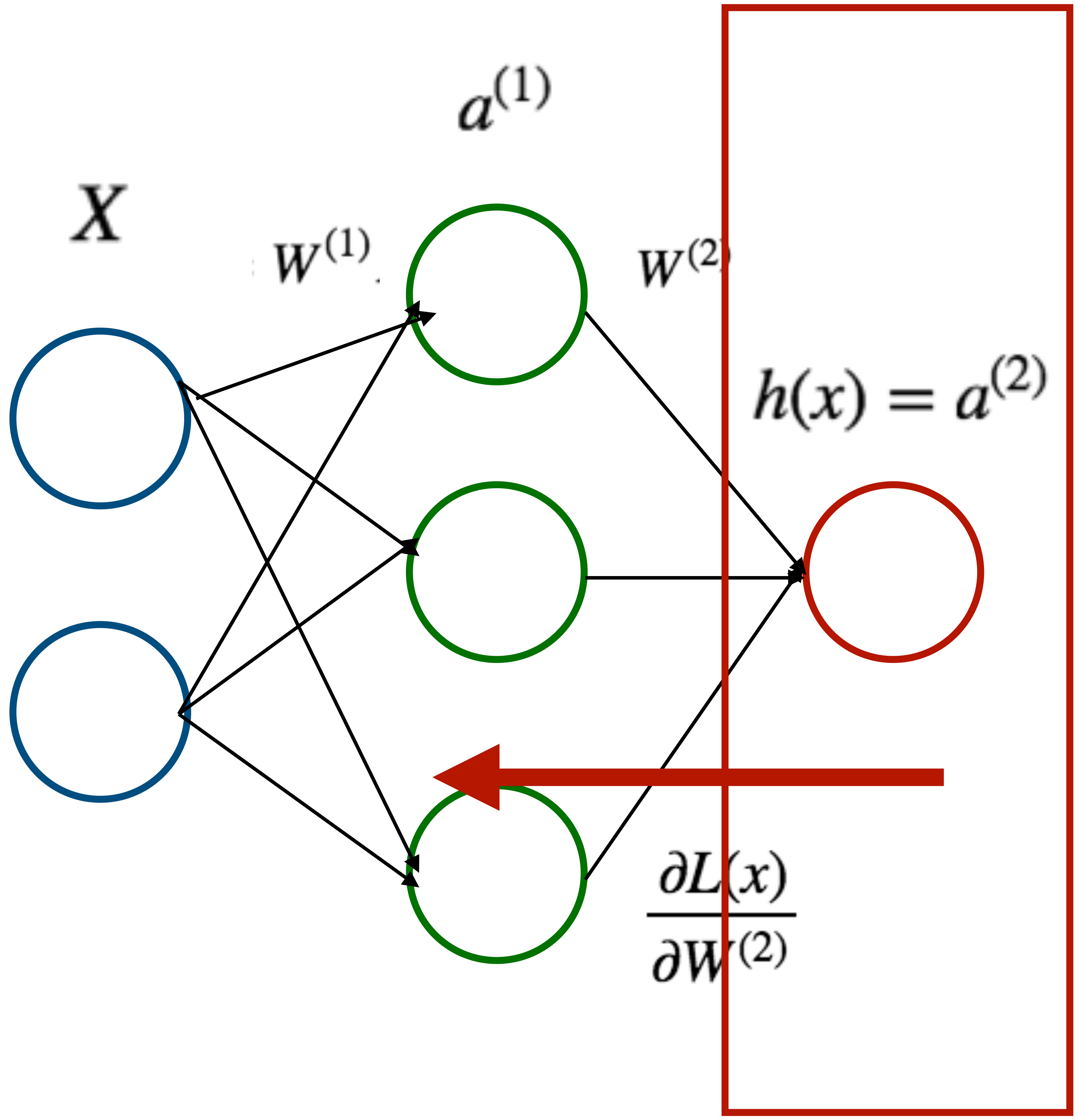
$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y)\log\left(1 - a^{(2)}\right)$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{\partial}{\partial a^{(2)}}\left(-y\log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right)$$

$$= \frac{\partial}{\partial a^{(2)}} \boxed{-y} \log a^{(2)} - \frac{\partial}{\partial a^{(2)}} \boxed{(1-y)} \log\left(1-a^{(2)}\right)$$

상수(y) 는 편미분 기호 바깥으로 뺄 수 있다



시각화하자면 이 부분

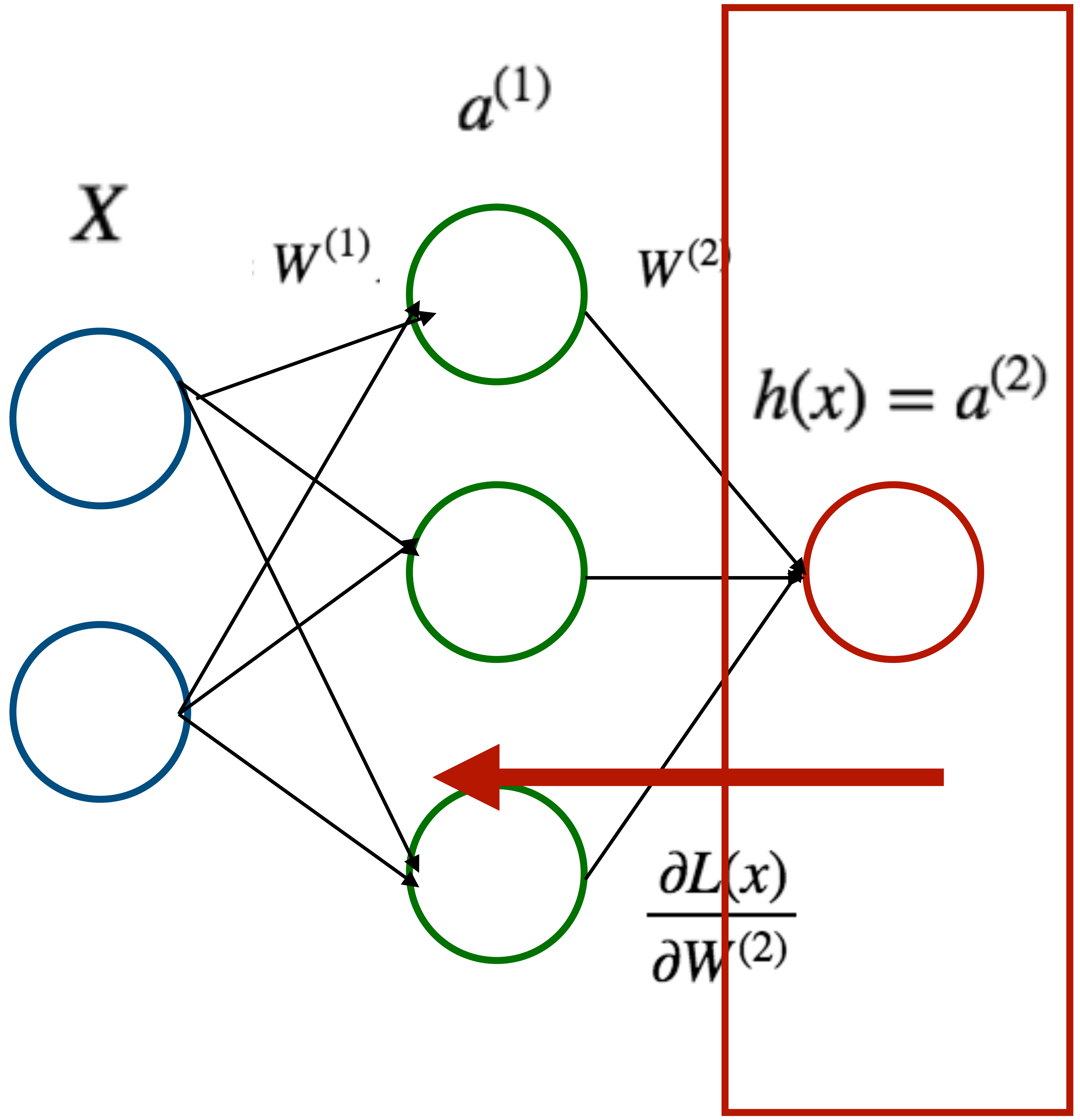
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right) \right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}} (1-y) \log \left(1 - a^{(2)}\right) \\
&= -y \frac{\partial}{\partial a^{(2)}} \log a^{(2)} - (1-y) \frac{\partial}{\partial a^{(2)}} \log \left(1 - a^{(2)}\right)
\end{aligned}$$



시각화하자면 이 부분

Derivation of Backpropagation

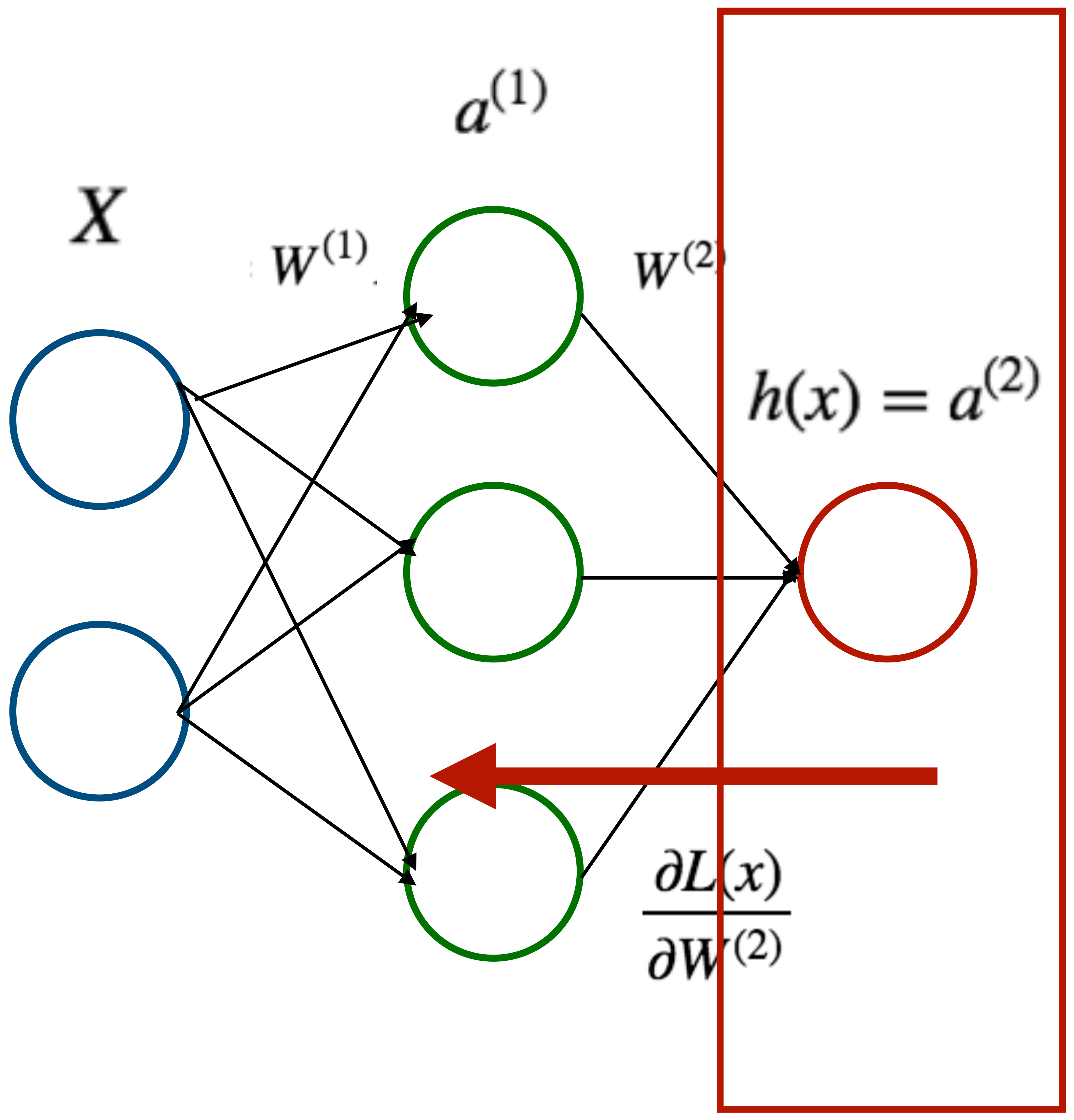
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}}\left(-y\log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right) \\
&= \frac{\partial}{\partial a^{(2)}} - y\log a^{(2)} - \frac{\partial}{\partial a^{(2)}}(1-y)\log\left(1-a^{(2)}\right) \\
&= -y\boxed{\frac{\partial}{\partial a^{(2)}}\log a^{(2)}} - (1-y)\boxed{\frac{\partial}{\partial a^{(2)}}\log\left(1-a^{(2)}\right)}
\end{aligned}$$

로그를 편미분한다 로그 + chain rule을 활용한다



시각화하자면 이 부분

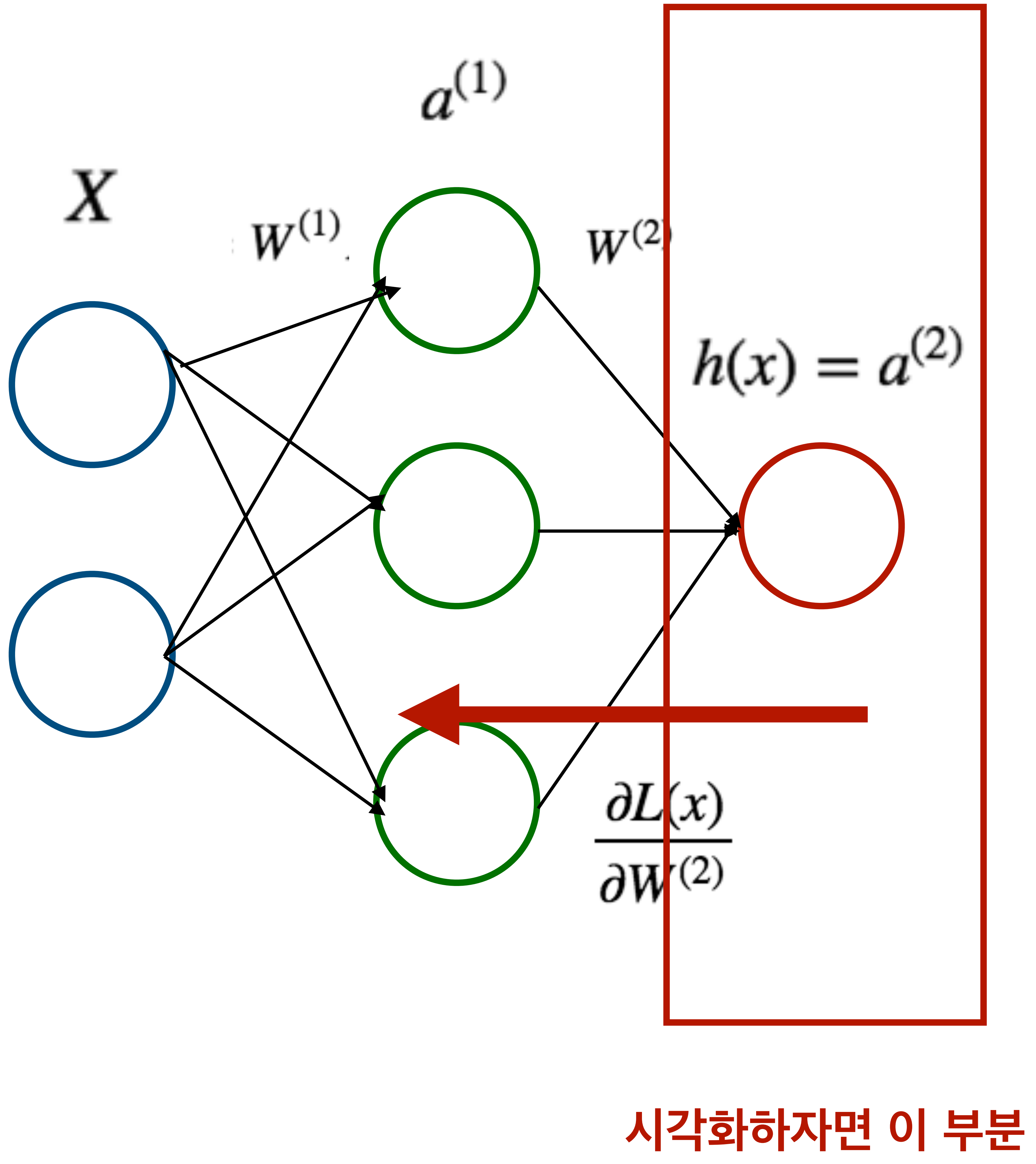
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right) \right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}} (1-y) \log \left(1 - a^{(2)}\right) \\
&= -y \frac{\partial}{\partial a^{(2)}} \log a^{(2)} - (1-y) \frac{\partial}{\partial a^{(2)}} \log \left(1 - a^{(2)}\right) \\
&= -y \frac{1}{a^{(2)}} - (1-y) \frac{1}{1 - a^{(2)}} \frac{\partial}{\partial a^{(2)}} \left(1 - a^{(2)}\right)
\end{aligned}$$



Derivation of Backpropagation

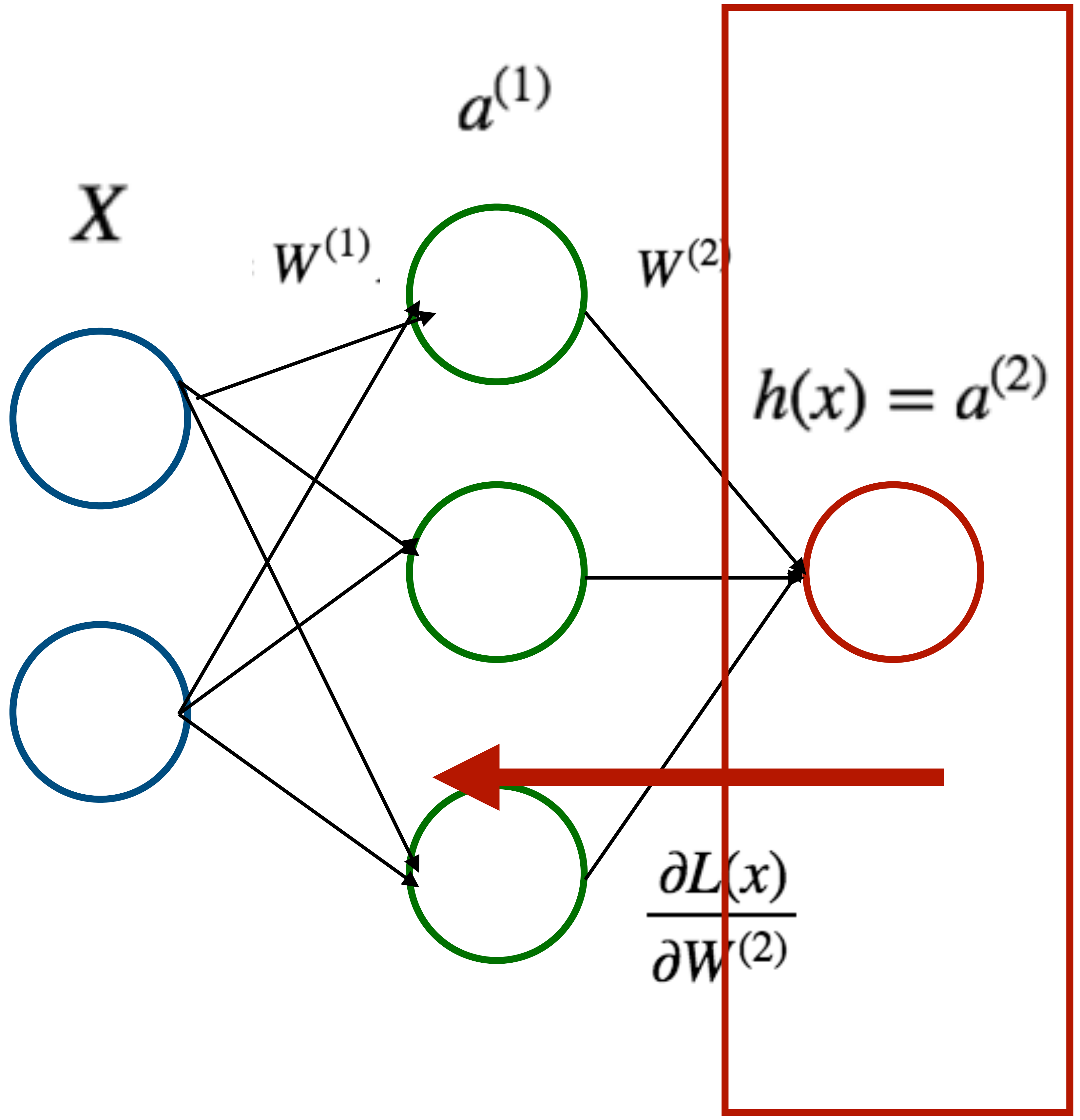
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log\left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}}\left(-y \log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}}(1-y)\log\left(1-a^{(2)}\right) \\
&= -y\frac{\partial}{\partial a^{(2)}}\log a^{(2)} - (1-y)\frac{\partial}{\partial a^{(2)}}\log\left(1-a^{(2)}\right) \\
&= -y\frac{1}{a^{(2)}} - (1-y)\frac{1}{1-a^{(2)}} \boxed{\frac{\partial}{\partial a^{(2)}}\left(1-a^{(2)}\right)}
\end{aligned}$$

상수는 사라지고,
자기 자신은 1이 된다



시각화하자면 이 부분

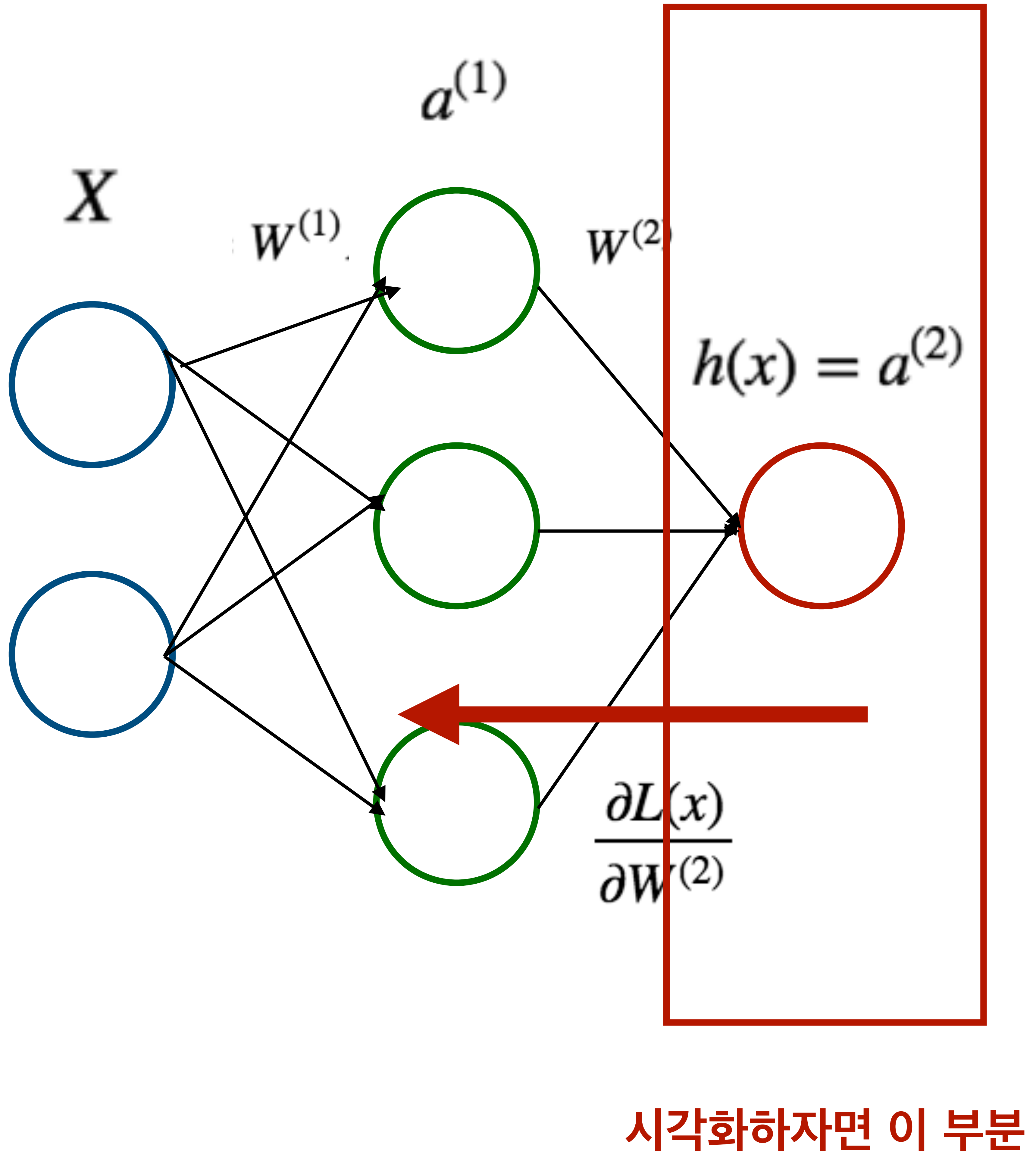
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log\left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}}\left(-y \log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}}(1-y)\log\left(1-a^{(2)}\right) \\
&= -y\frac{\partial}{\partial a^{(2)}}\log a^{(2)} - (1-y)\frac{\partial}{\partial a^{(2)}}\log\left(1-a^{(2)}\right) \\
&= -y\frac{1}{a^{(2)}} - (1-y)\frac{1}{1-a^{(2)}}\frac{\partial}{\partial a^{(2)}}\left(1-a^{(2)}\right) \\
&= -y\frac{1}{a^{(2)}} - (1-y)\frac{1}{1-a^{(2)}}(-1)
\end{aligned}$$



Derivation of Backpropagation

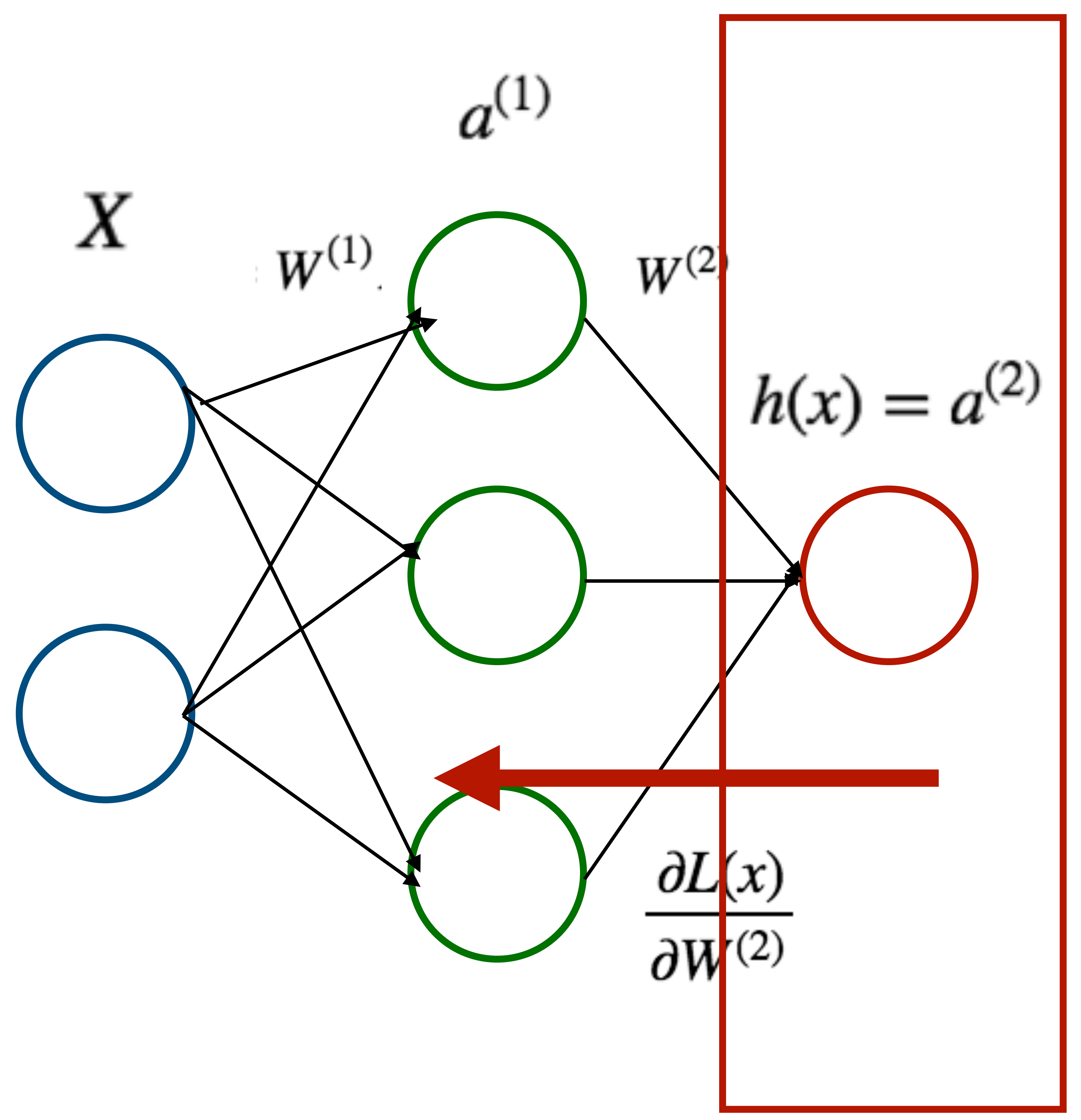
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right) \right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}} (1-y) \log \left(1 - a^{(2)}\right) \\
&= -y \frac{\partial}{\partial a^{(2)}} \log a^{(2)} - (1-y) \frac{\partial}{\partial a^{(2)}} \log \left(1 - a^{(2)}\right) \\
&= -y \frac{1}{a^{(2)}} - (1-y) \frac{1}{1-a^{(2)}} \frac{\partial}{\partial a^{(2)}} \left(1 - a^{(2)}\right) \\
&= -y \frac{1}{a^{(2)}} \boxed{- (1} - y) \frac{1}{1-a^{(2)}} \boxed{(-1)}
\end{aligned}$$

-1를 없애고 앞에
마이너스 부호를 뒤집는다.



시각화하자면 이 부분

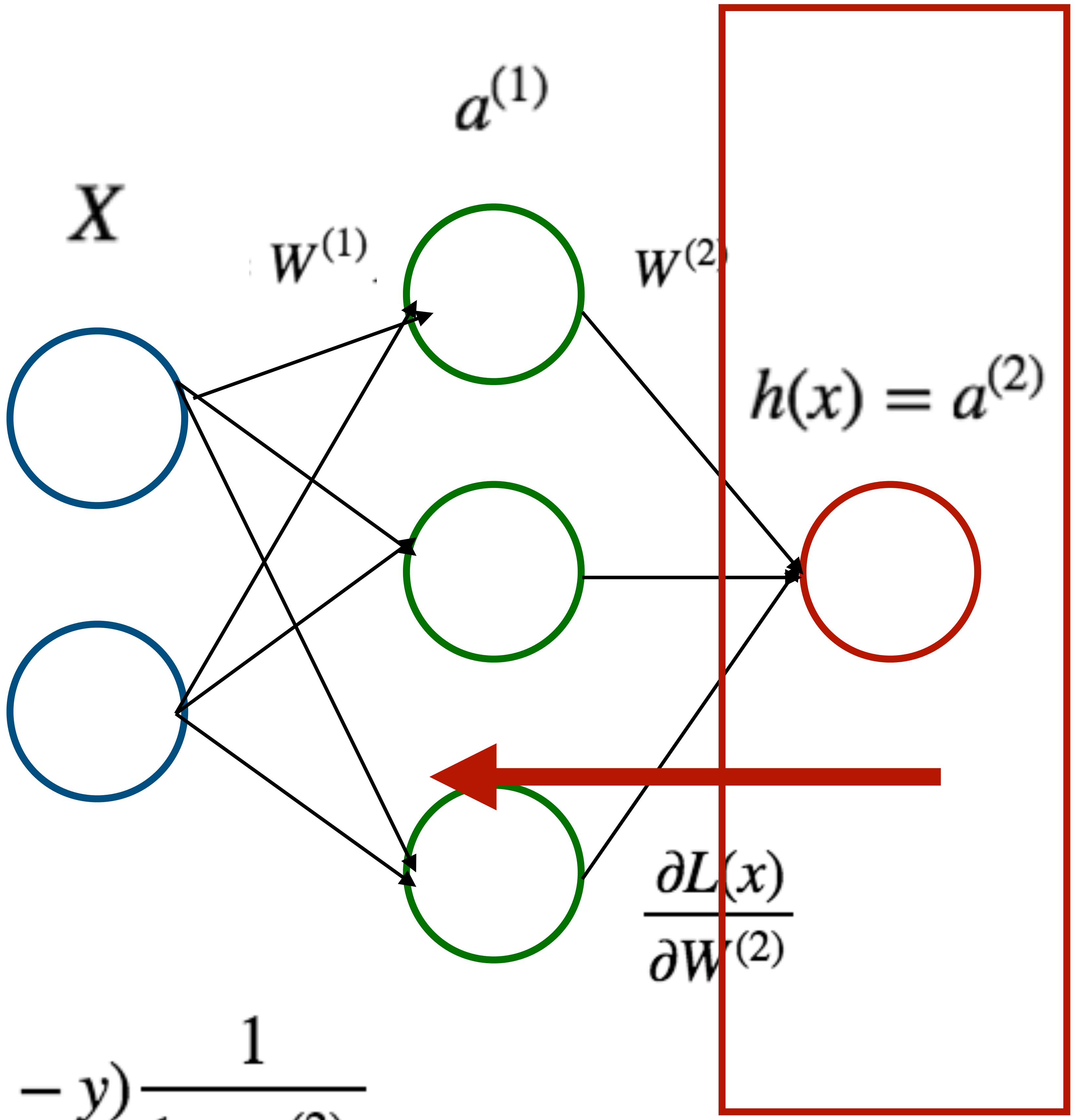
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y)\log\left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}}\left(-y\log a^{(2)} - (1-y)\log\left(1-a^{(2)}\right)\right) \\
&= \frac{\partial}{\partial a^{(2)}} - y\log a^{(2)} - \frac{\partial}{\partial a^{(2)}}(1-y)\log\left(1-a^{(2)}\right) \\
&= -y\frac{\partial}{\partial a^{(2)}}\log a^{(2)} - (1-y)\frac{\partial}{\partial a^{(2)}}\log\left(1-a^{(2)}\right) \\
&= -y\frac{1}{a^{(2)}} - (1-y)\frac{1}{1-a^{(2)}}\frac{\partial}{\partial a^{(2)}}\left(1-a^{(2)}\right) \qquad = -y\frac{1}{a^{(2)}} + (1-y)\frac{1}{1-a^{(2)}} \\
&= -y\frac{1}{a^{(2)}} - (1-y)\frac{1}{1-a^{(2)}}(-1)
\end{aligned}$$



시각화하자면 이 부분

Derivation of Backpropagation

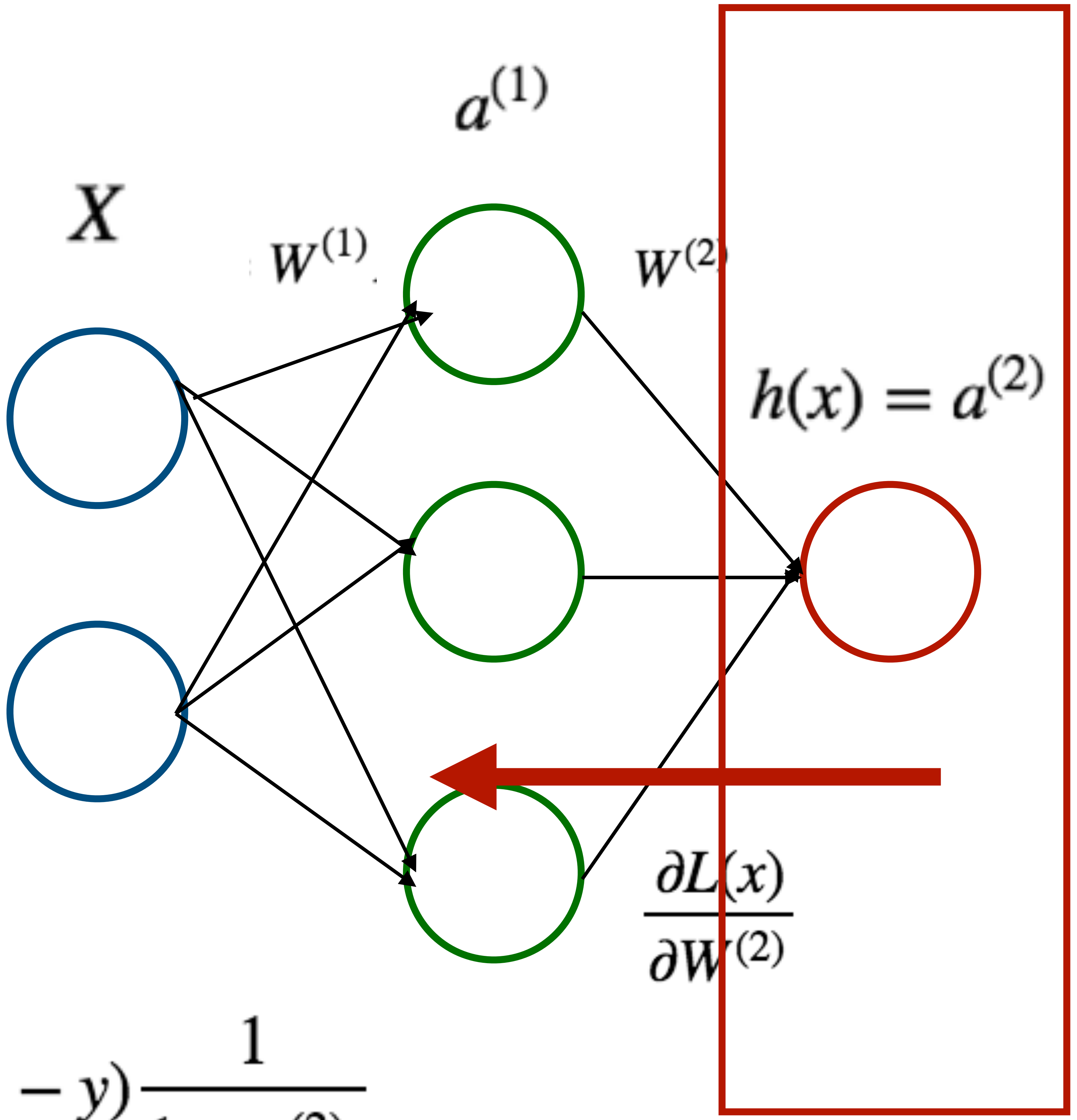
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$L(x) = -y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial a^{(2)}} &= \frac{\partial}{\partial a^{(2)}} \left(-y \log a^{(2)} - (1-y) \log \left(1 - a^{(2)}\right)\right) \\
&= \frac{\partial}{\partial a^{(2)}} - y \log a^{(2)} - \frac{\partial}{\partial a^{(2)}} (1-y) \log \left(1 - a^{(2)}\right) \\
&= -y \frac{\partial}{\partial a^{(2)}} \log a^{(2)} - (1-y) \frac{\partial}{\partial a^{(2)}} \log \left(1 - a^{(2)}\right) \\
&= -y \frac{1}{a^{(2)}} - (1-y) \frac{1}{1 - a^{(2)}} \frac{\partial}{\partial a^{(2)}} \left(1 - a^{(2)}\right) \\
&= -y \frac{1}{a^{(2)}} - (1-y) \frac{1}{1 - a^{(2)}} (-1)
\end{aligned}$$

$$\begin{aligned}
&= -y \frac{1}{a^{(2)}} + (1-y) \frac{1}{1 - a^{(2)}} \\
&= -\frac{y}{a^{(2)}} + \frac{(1-y)}{1 - a^{(2)}}
\end{aligned}$$



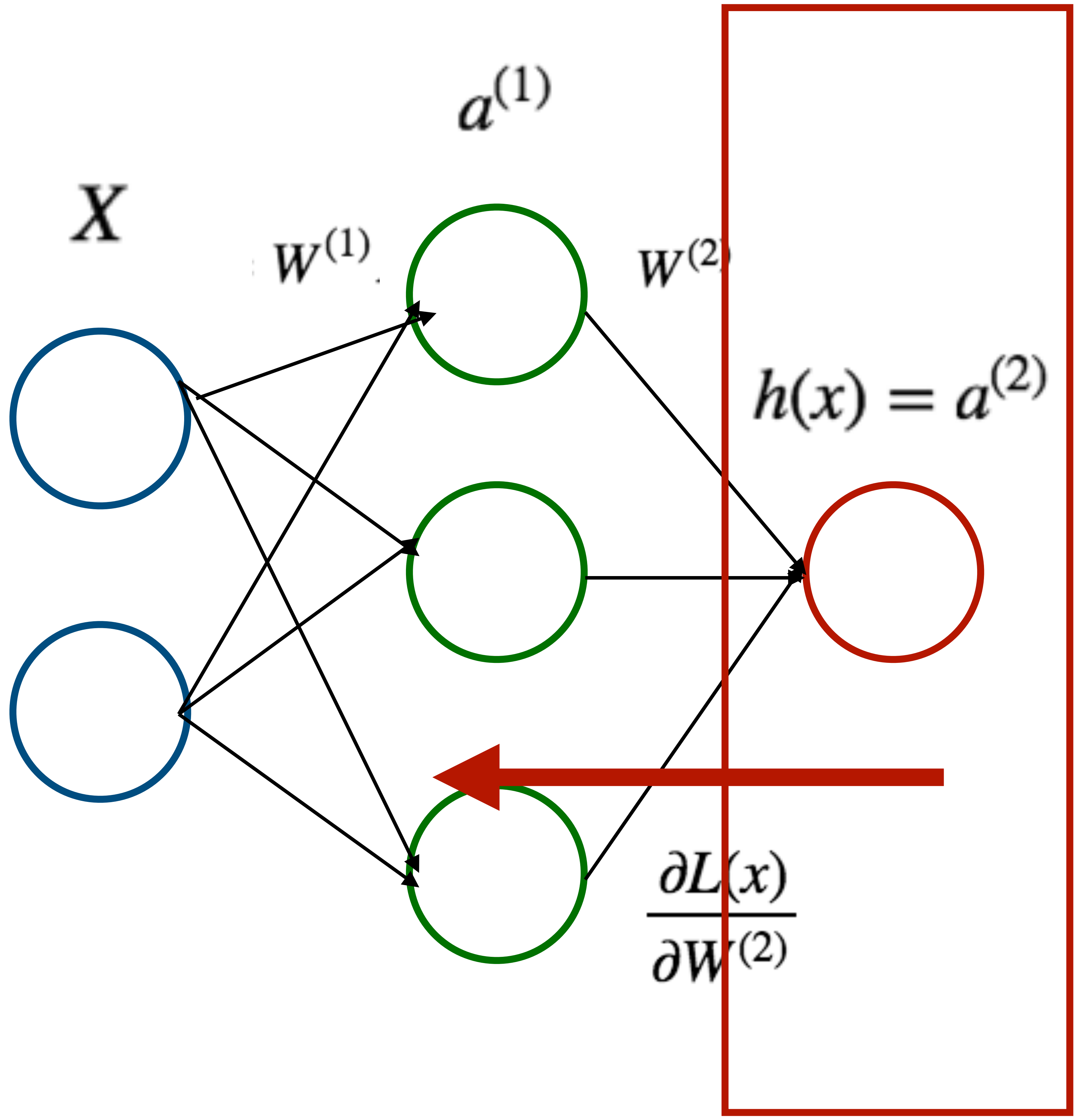
시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}}$$



시각화하자면 이 부분

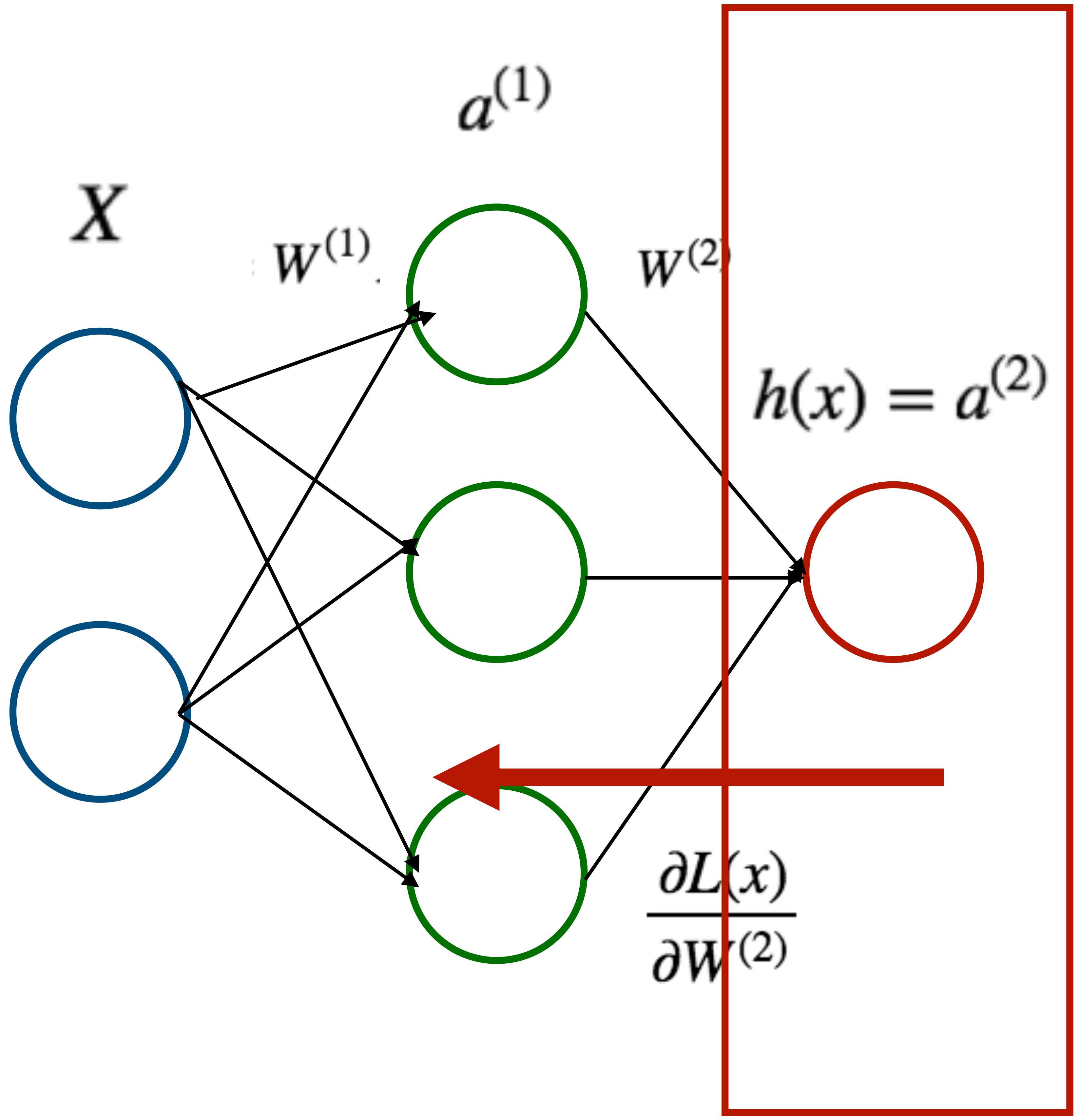
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}}$$

$$= \frac{-y(1-a^{(2)})}{a^{(2)}(1-a^{(2)})} + \frac{a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})}$$



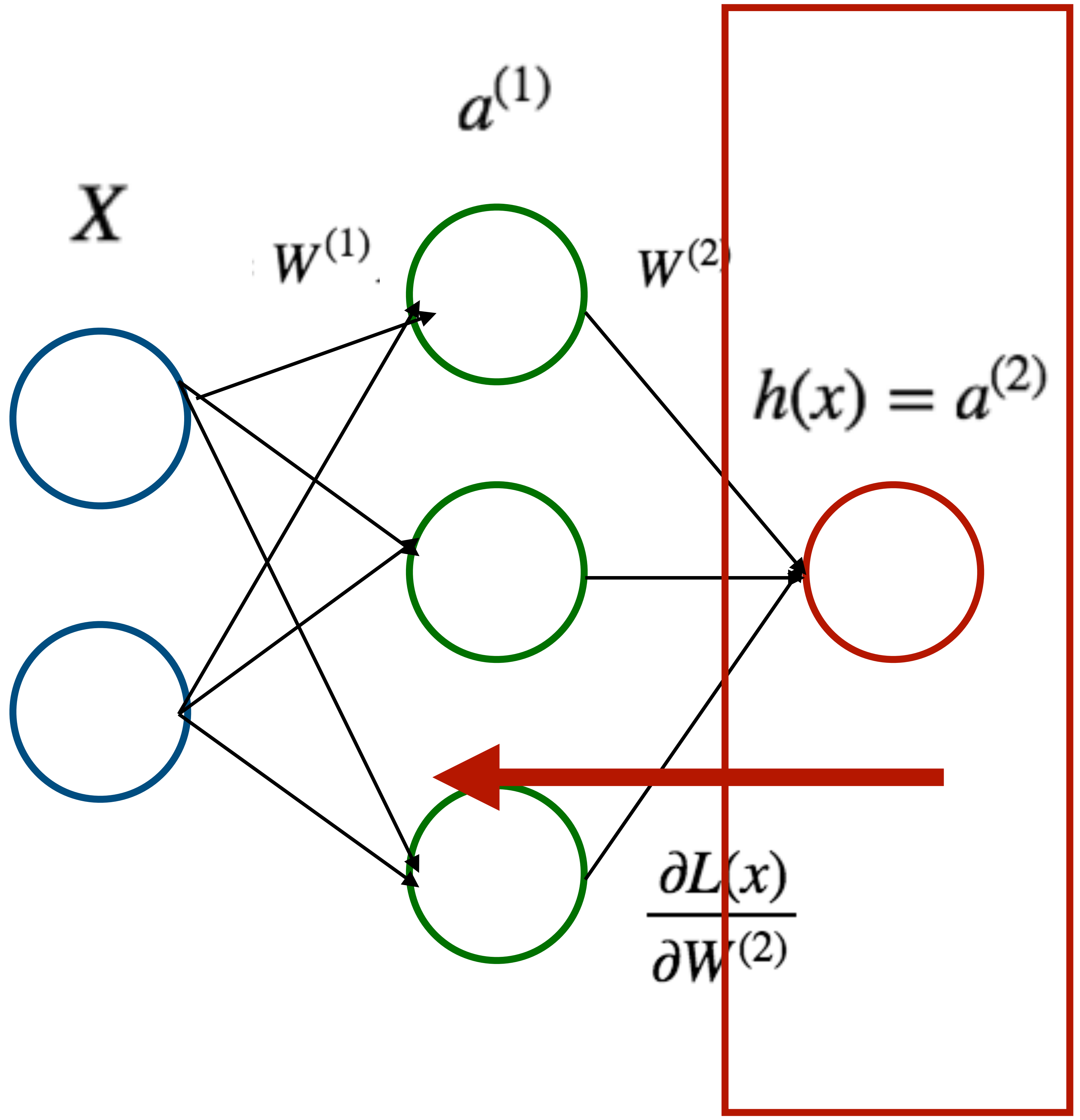
시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial W^{(2)}} &= -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}} \\
&= \frac{-y(1-a^{(2)})}{a^{(2)}(1-a^{(2)})} + \frac{a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\
&= \frac{-y(1-a^{(2)}) + a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})}
\end{aligned}$$



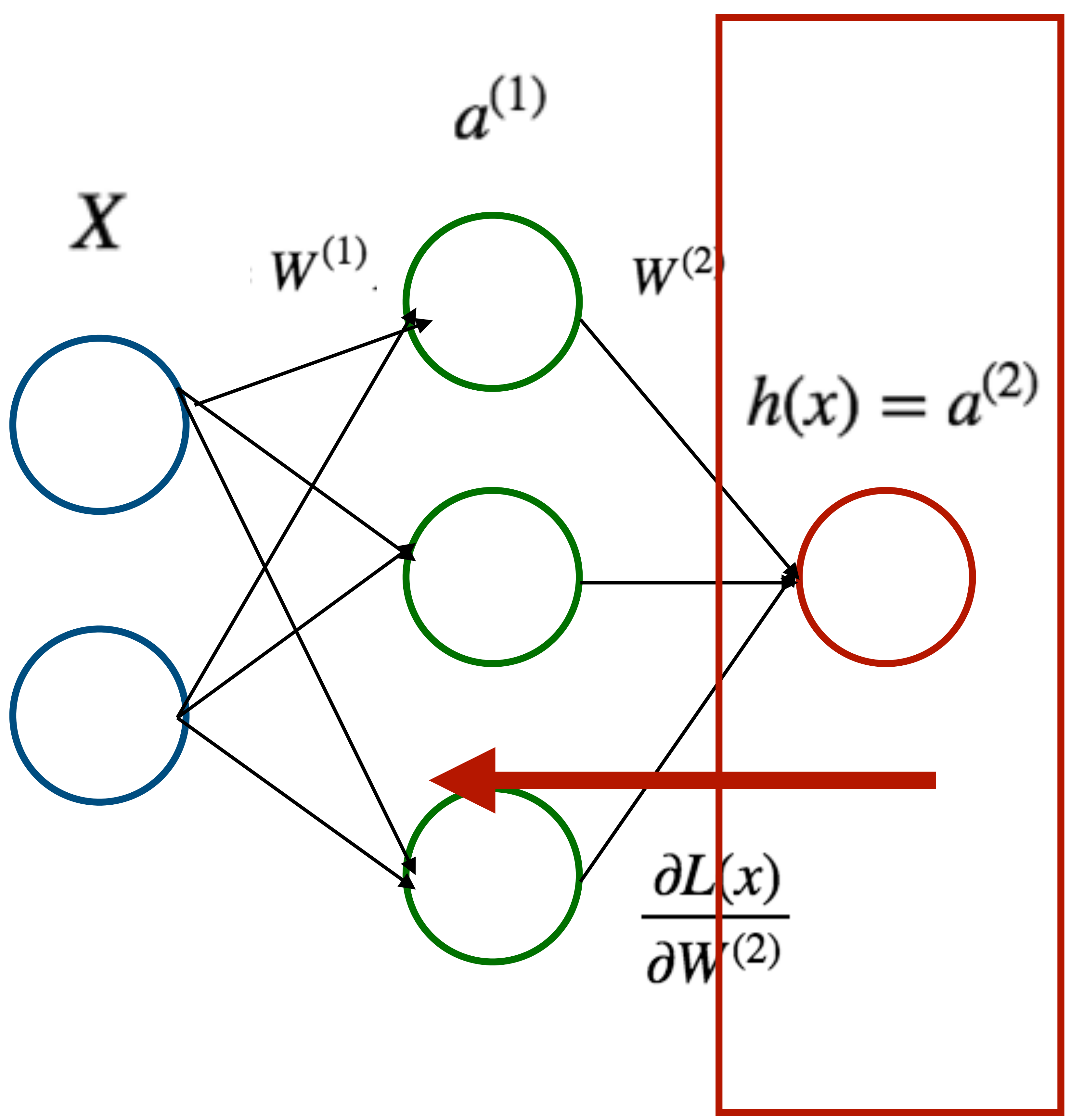
시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial W^{(2)}} &= -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}} \\
&= \frac{-y(1-a^{(2)})}{a^{(2)}(1-a^{(2)})} + \frac{a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\
&= \frac{-y(1-a^{(2)}) + a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\
&= \frac{-y + ya^{(2)} + a^{(2)} - ya^{(2)}}{a^{(2)}(1-a^{(2)})}
\end{aligned}$$



시각화하자면 이 부분

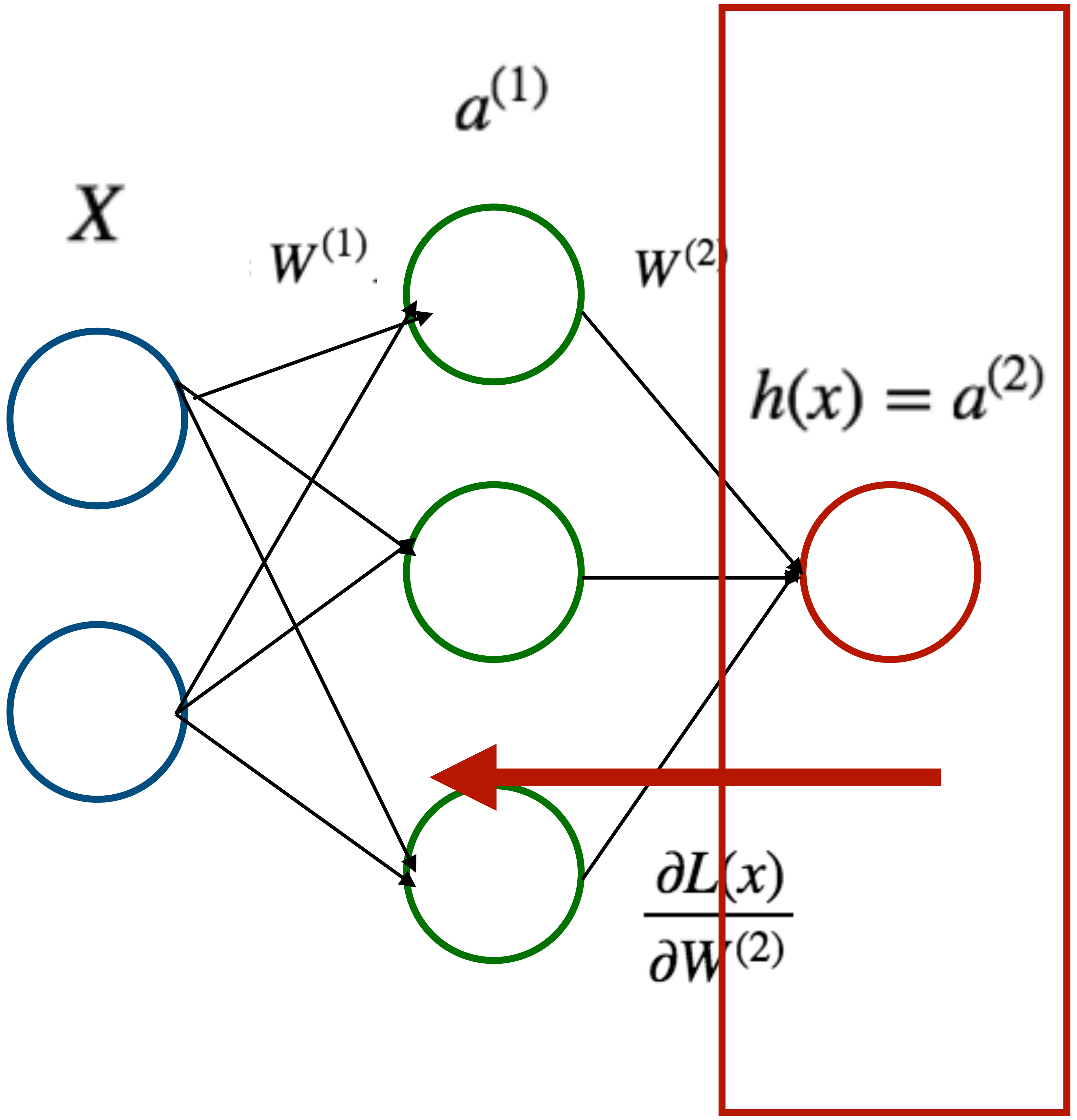
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}
\frac{\partial L(x)}{\partial W^{(2)}} &= -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}} \\
&= \frac{-y(1-a^{(2)})}{a^{(2)}(1-a^{(2)})} + \frac{a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\
&= \frac{-y(1-a^{(2)}) + a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\
&= \frac{-y + \boxed{ya^{(2)}} + a^{(2)} \boxed{- ya^{(2)}}}{a^{(2)}(1-a^{(2)})}
\end{aligned}$$

두 개를 없앨 수 있다



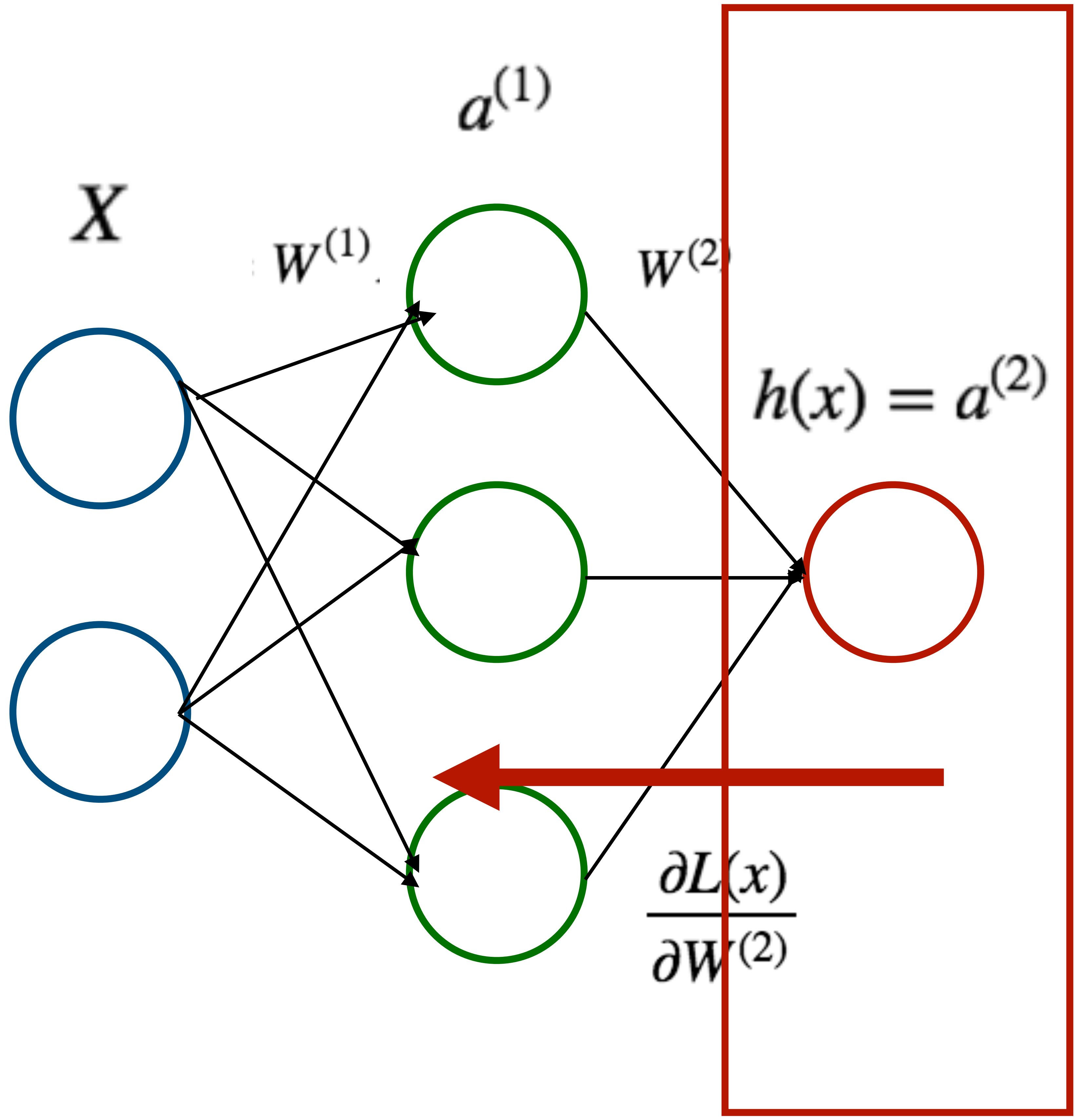
시각화하자면 이 부분

Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \boxed{\frac{\partial L(x)}{\partial a^{(2)}}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\begin{aligned}\frac{\partial L(x)}{\partial W^{(2)}} &= -\frac{y}{a^{(2)}} + \frac{(1-y)}{1-a^{(2)}} \\ &= \frac{-y(1-a^{(2)})}{a^{(2)}(1-a^{(2)})} + \frac{a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\ &= \frac{-y(1-a^{(2)}) + a^{(2)}(1-y)}{a^{(2)}(1-a^{(2)})} \\ &= \frac{-y + ya^{(2)} + a^{(2)} - ya^{(2)}}{a^{(2)}(1-a^{(2)})} \\ &= \frac{a^{(2)} - y}{a^{(2)}(1-a^{(2)})}\end{aligned}$$



시각화하자면 이 부분

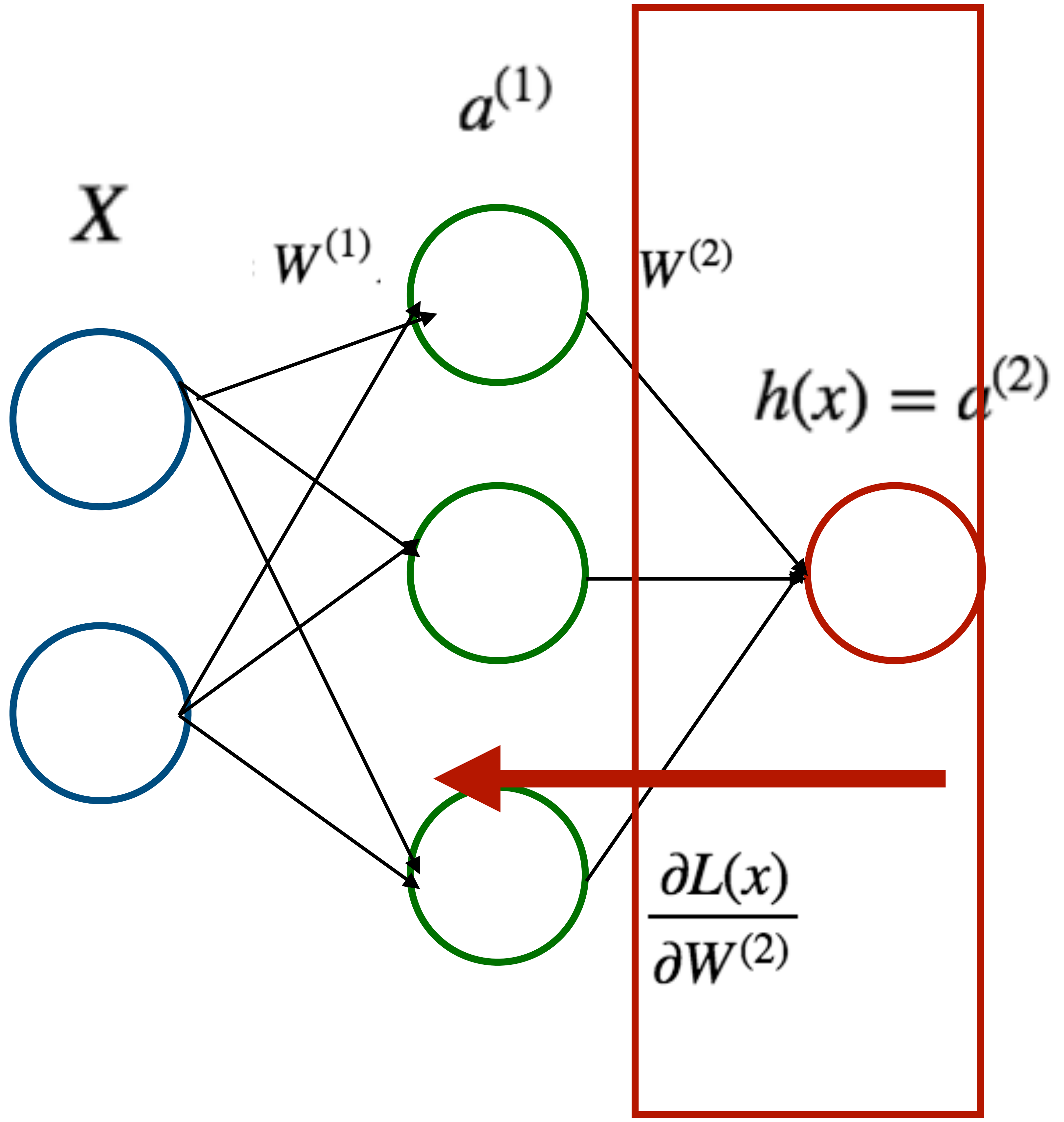
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \boxed{\frac{\partial a^{(2)}}{\partial z^{(2)}}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} =$$



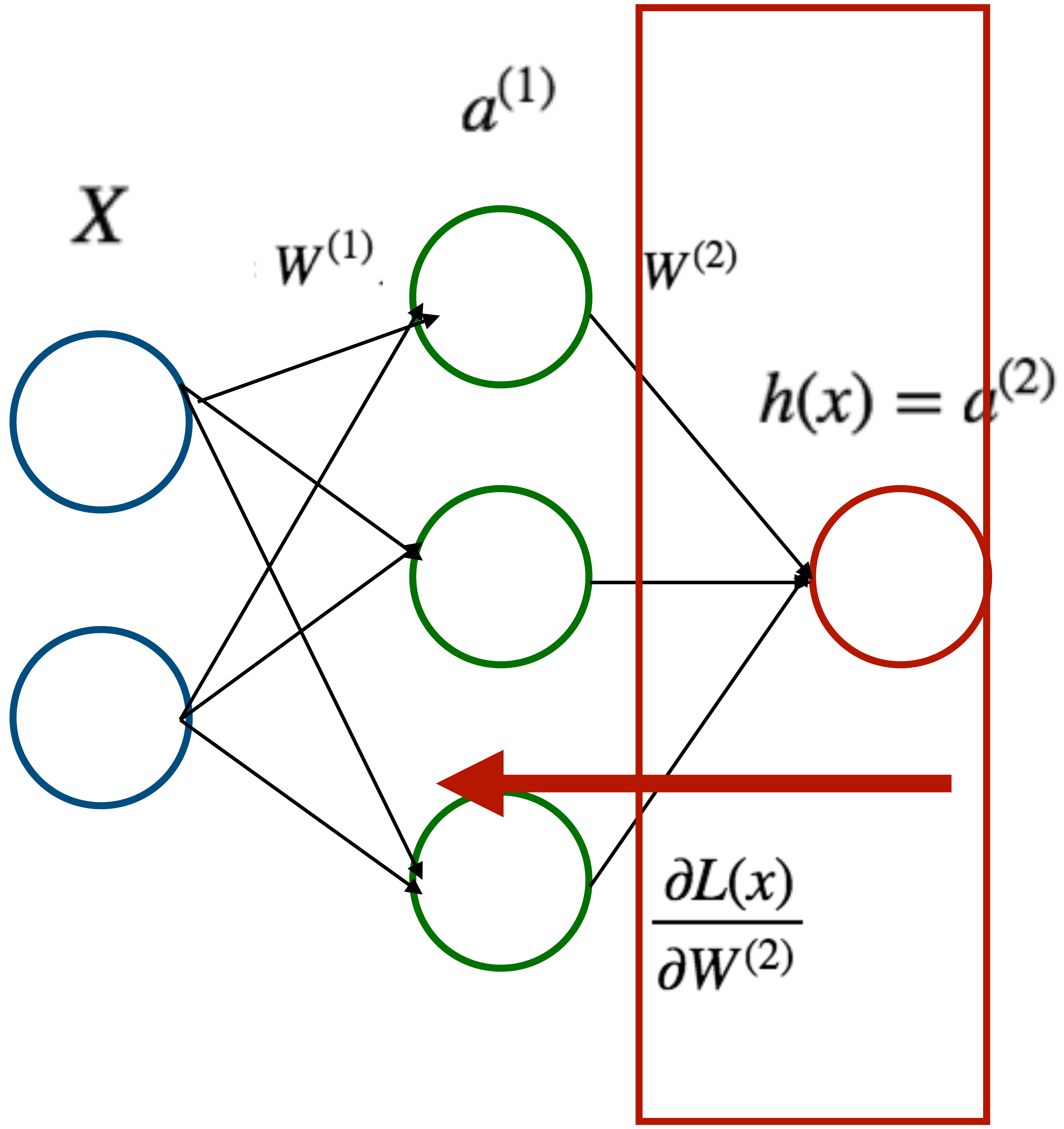
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \boxed{\frac{\partial a^{(2)}}{\partial z^{(2)}}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \frac{\partial}{\partial z^{(2)}} sigmoid\left(z^{(2)}\right)$$



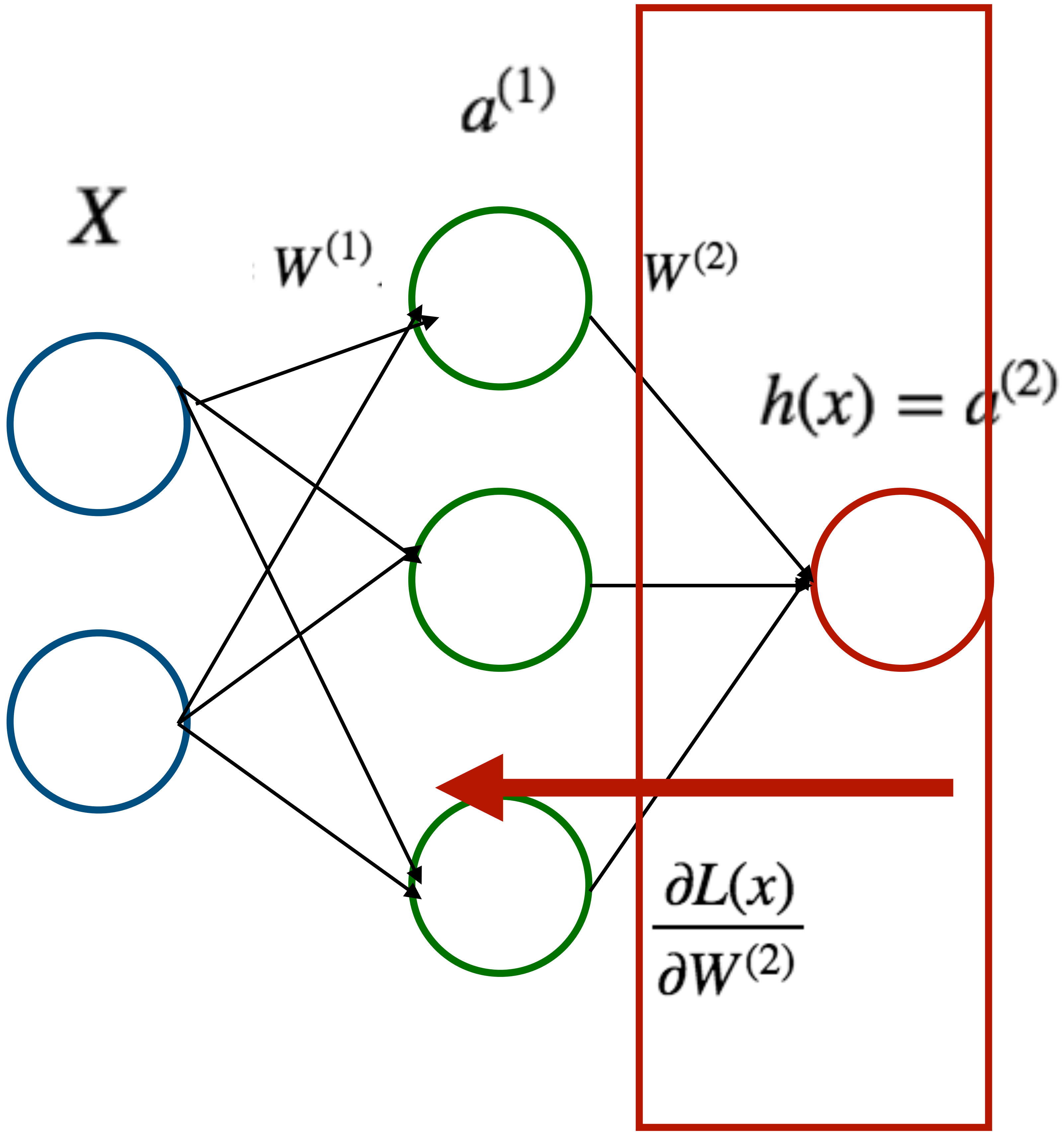
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \boxed{\frac{\partial a^{(2)}}{\partial z^{(2)}}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \frac{\partial}{\partial z^{(2)}} sigmoid\left(z^{(2)}\right)$$
$$= sigmoid\left(z^{(2)}\right)\left(1 - sigmoid\left(z^{(2)}\right)\right)$$



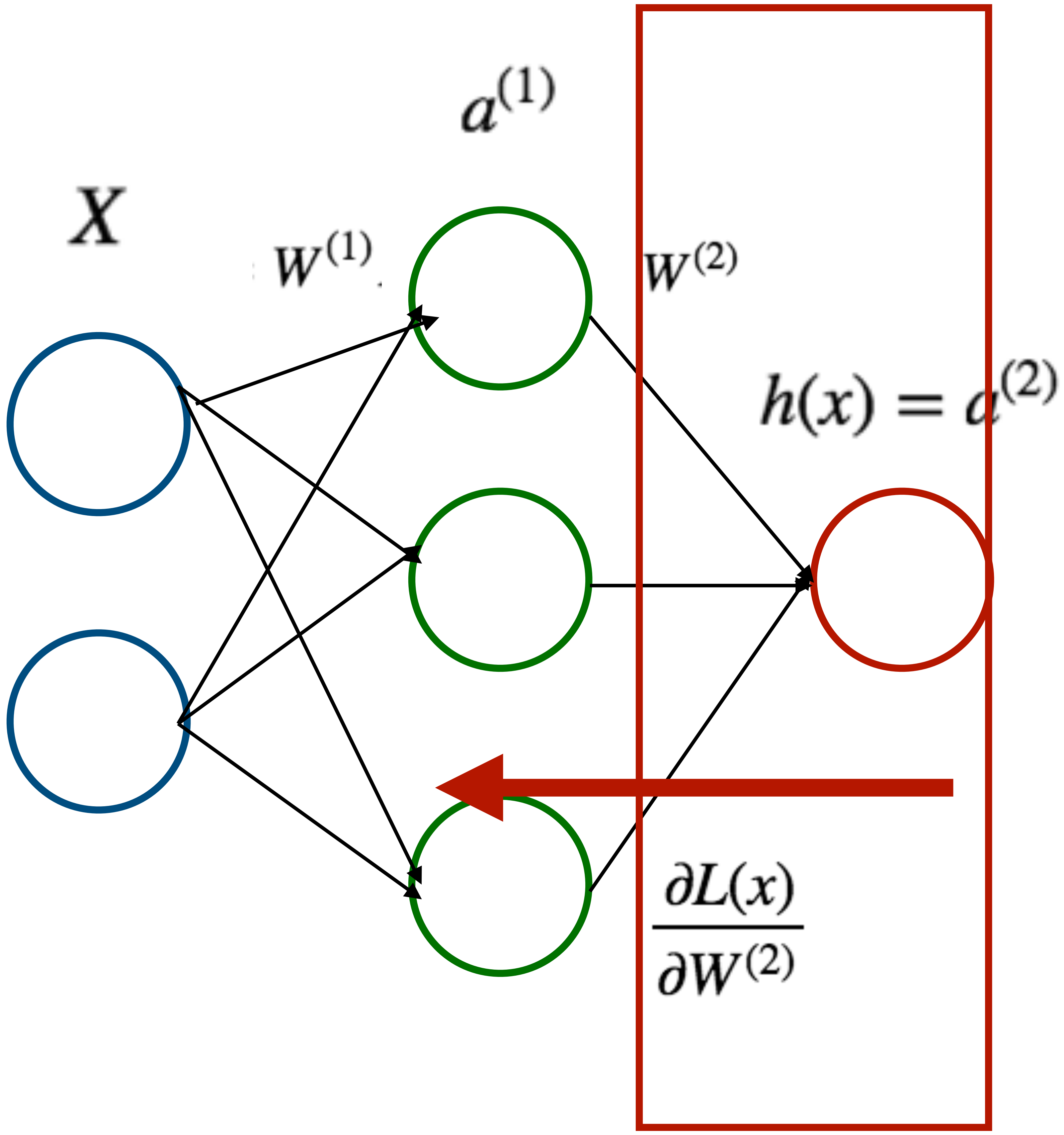
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \boxed{\frac{\partial a^{(2)}}{\partial z^{(2)}}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$a^{(2)} = sigmoid(z^{(2)})$$

$$\begin{aligned}\frac{\partial a^{(2)}}{\partial z^{(2)}} &= \frac{\partial}{\partial z^{(2)}} sigmoid\left(z^{(2)}\right) \\ &= sigmoid\left(z^{(2)}\right)\left(1 - sigmoid\left(z^{(2)}\right)\right) \\ &= a^{(2)}\left(1 - a^{(2)}\right)\end{aligned}$$



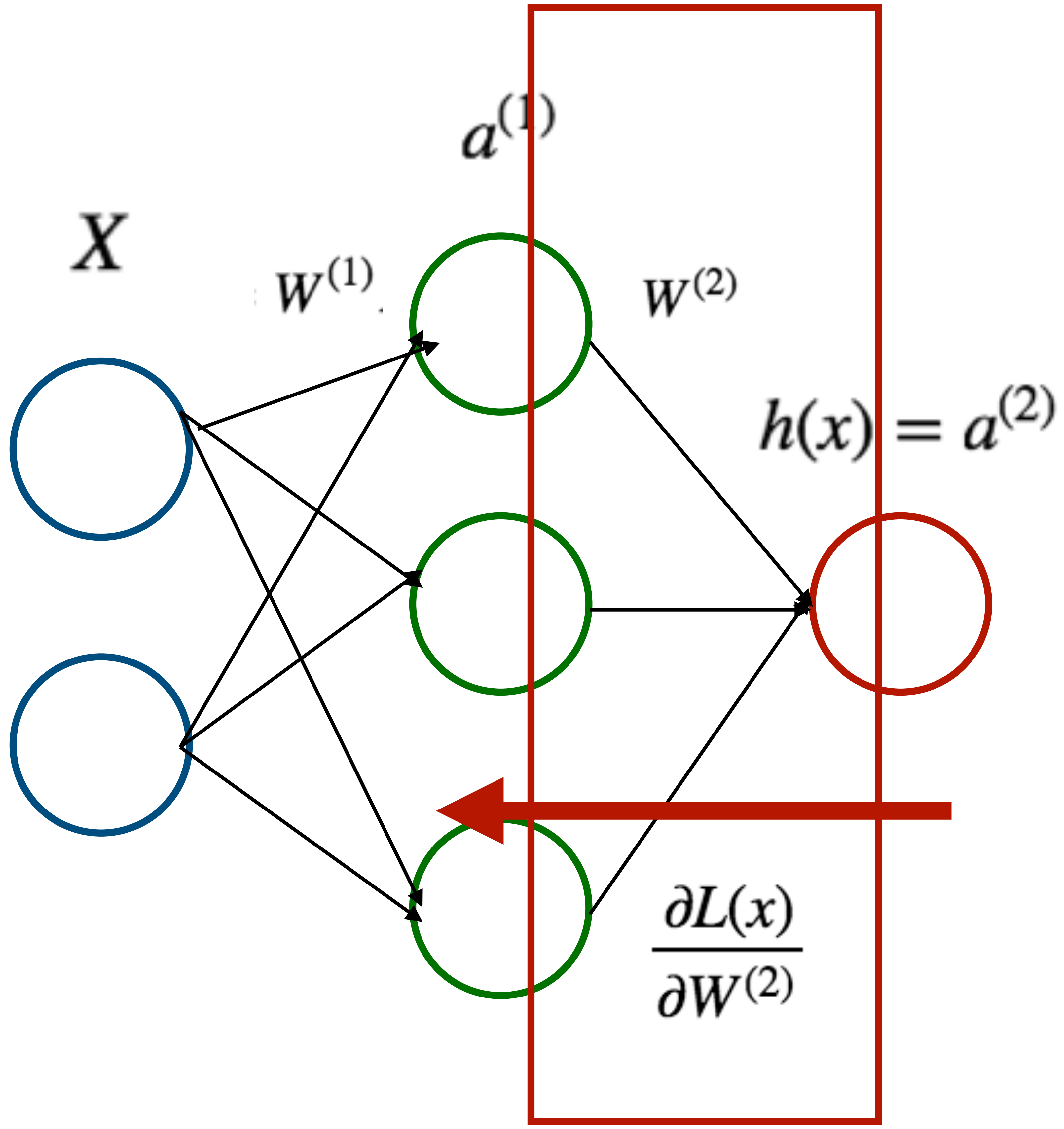
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \boxed{\frac{\partial z^{(2)}}{\partial W^{(2)}}}$$

$$z^{(2)} = W^{(2)} X + b^{(2)}$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} =$$



시각화하자면 이 부분

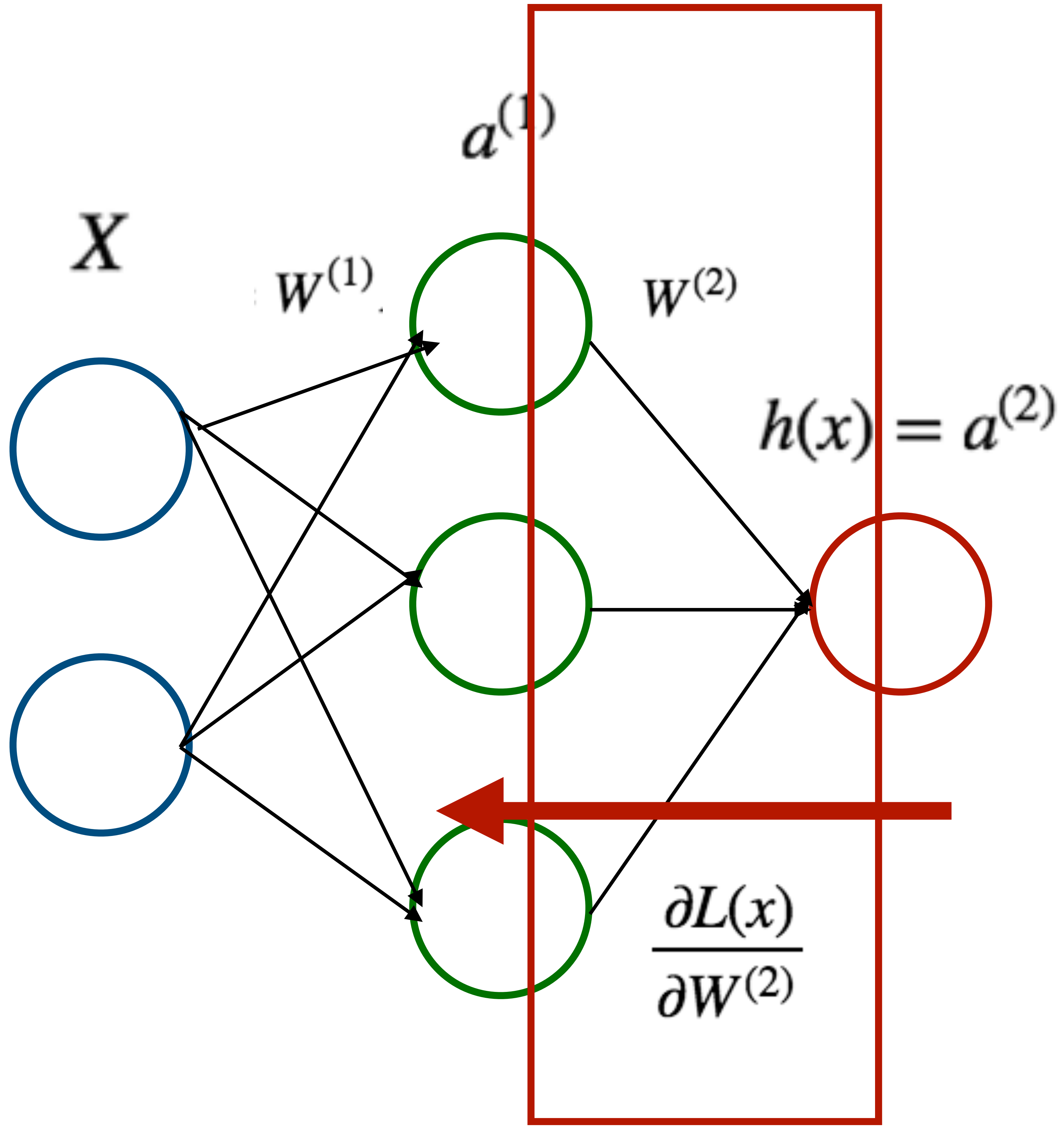
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \boxed{\frac{\partial z^{(2)}}{\partial W^{(2)}}}$$

$$z^{(2)} = W^{(2)} X + b^{(2)}$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = \frac{\partial}{\partial W^{(2)}} \left(W^{(2)} a^{(1)} + b^{(2)} \right)$$



Derivation of Backpropagation

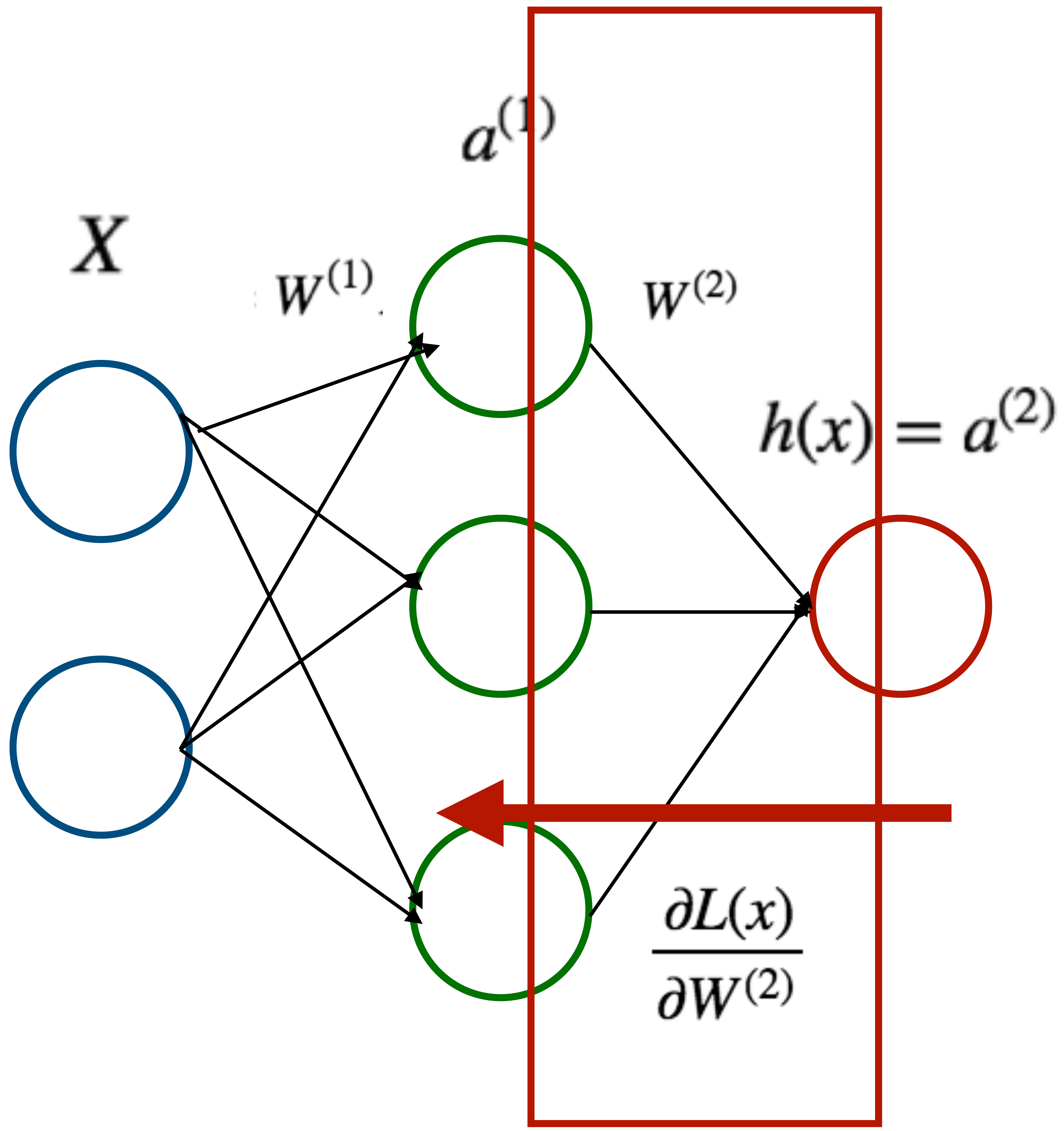
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \boxed{\frac{\partial z^{(2)}}{\partial W^{(2)}}}$$

$$z^{(2)} = W^{(2)} X + b^{(2)}$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = \frac{\partial}{\partial W^{(2)}} \left(W^{(2)} a^{(1)} + b^{(2)} \right)$$

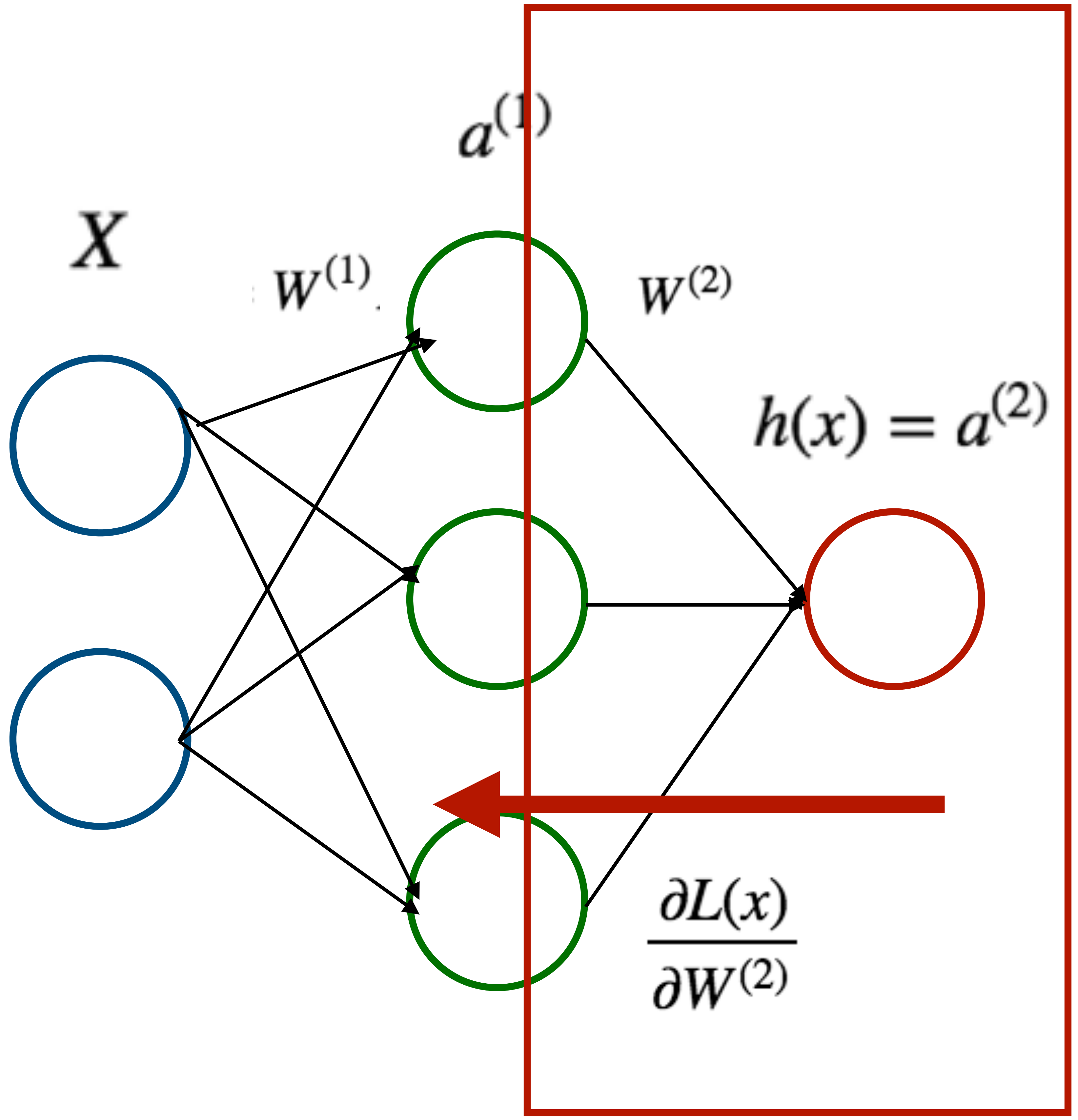
$$= a^{(1)}$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, **chain rule**을 활용해 여러 개로 쪼갤 수 있다
그래서 **Loss Function**의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 **backpropation**의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$



Derivation of Backpropagation

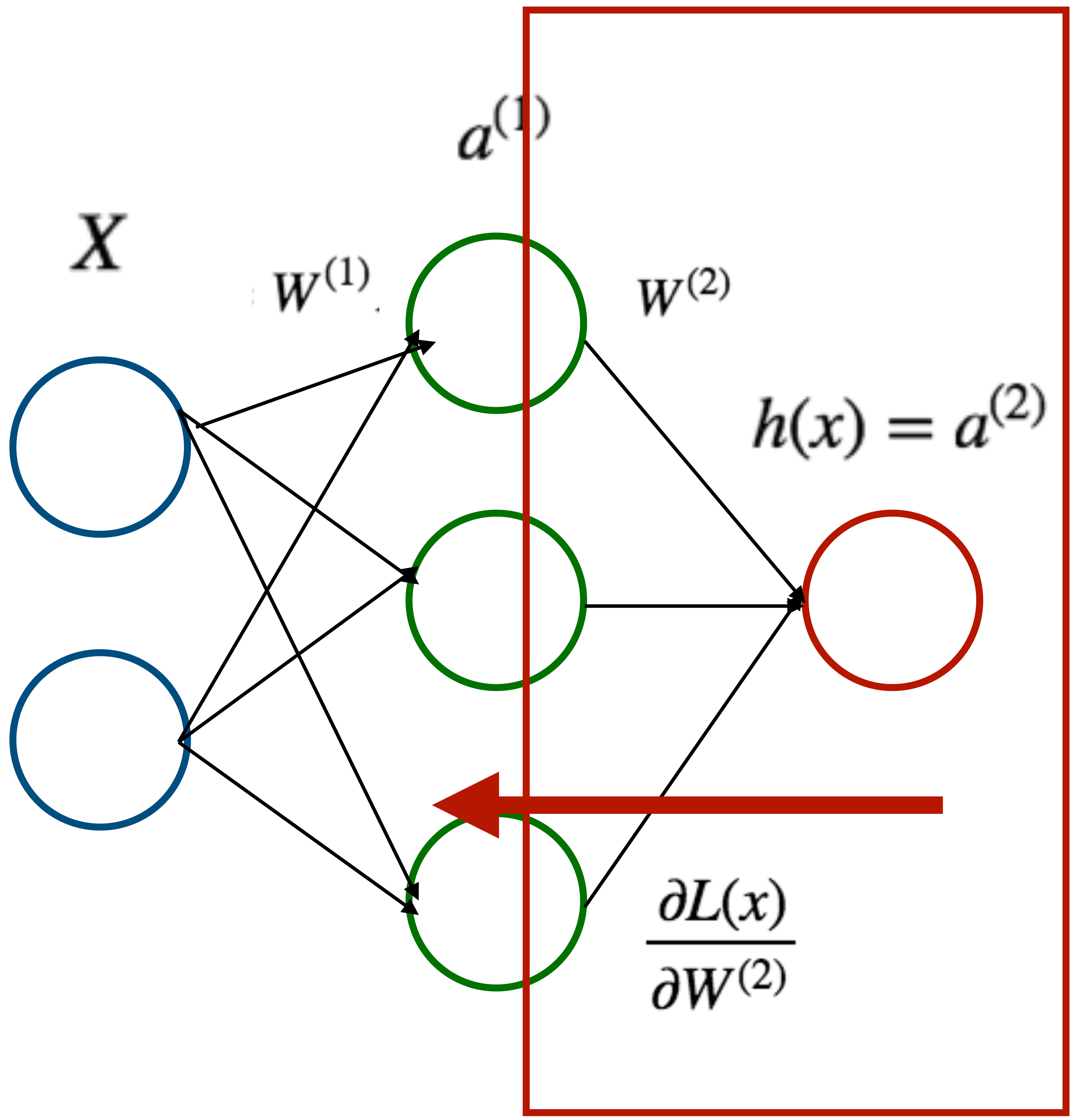
우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = a^{(1)}$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

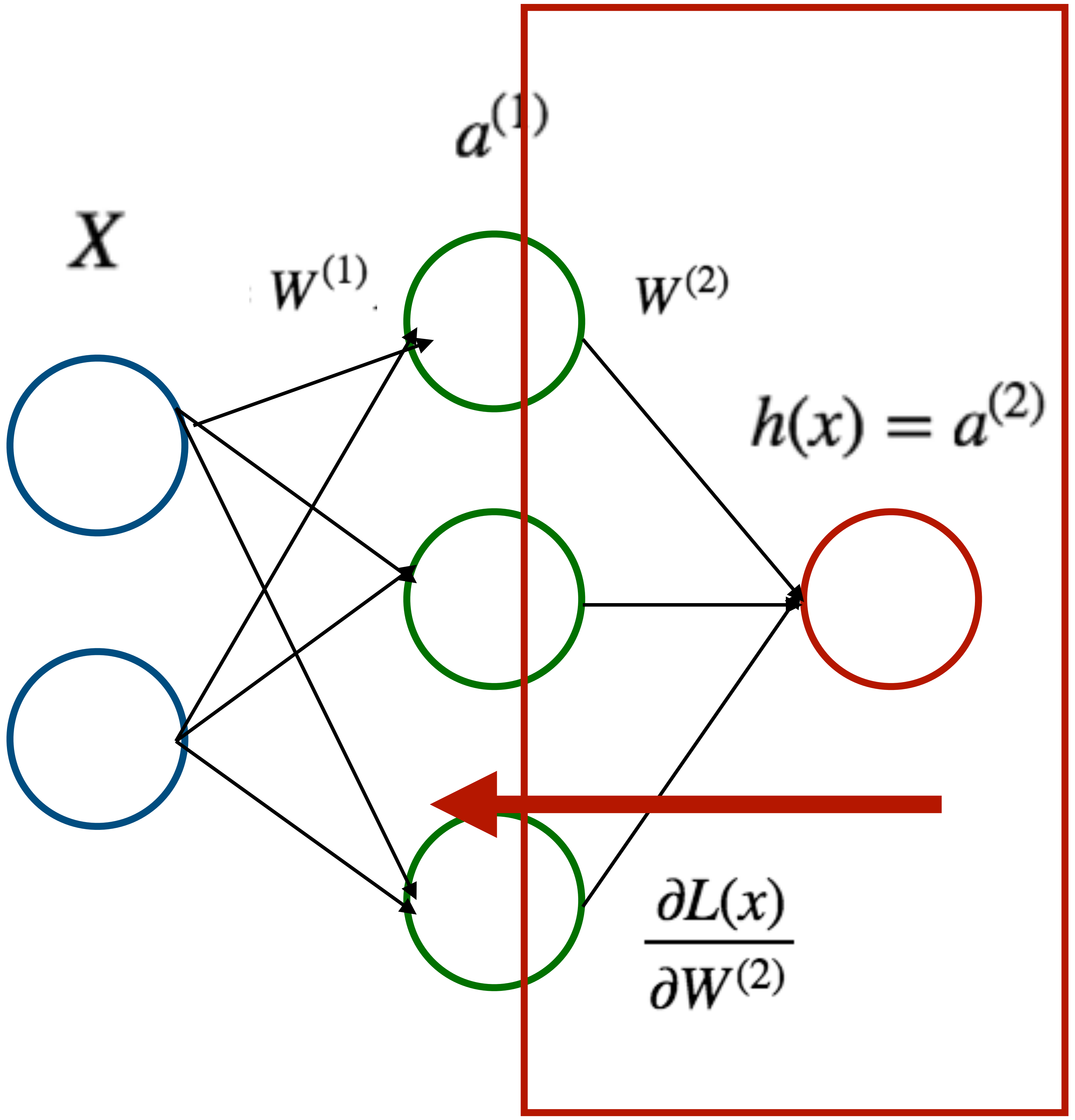
$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = a^{(1)}$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

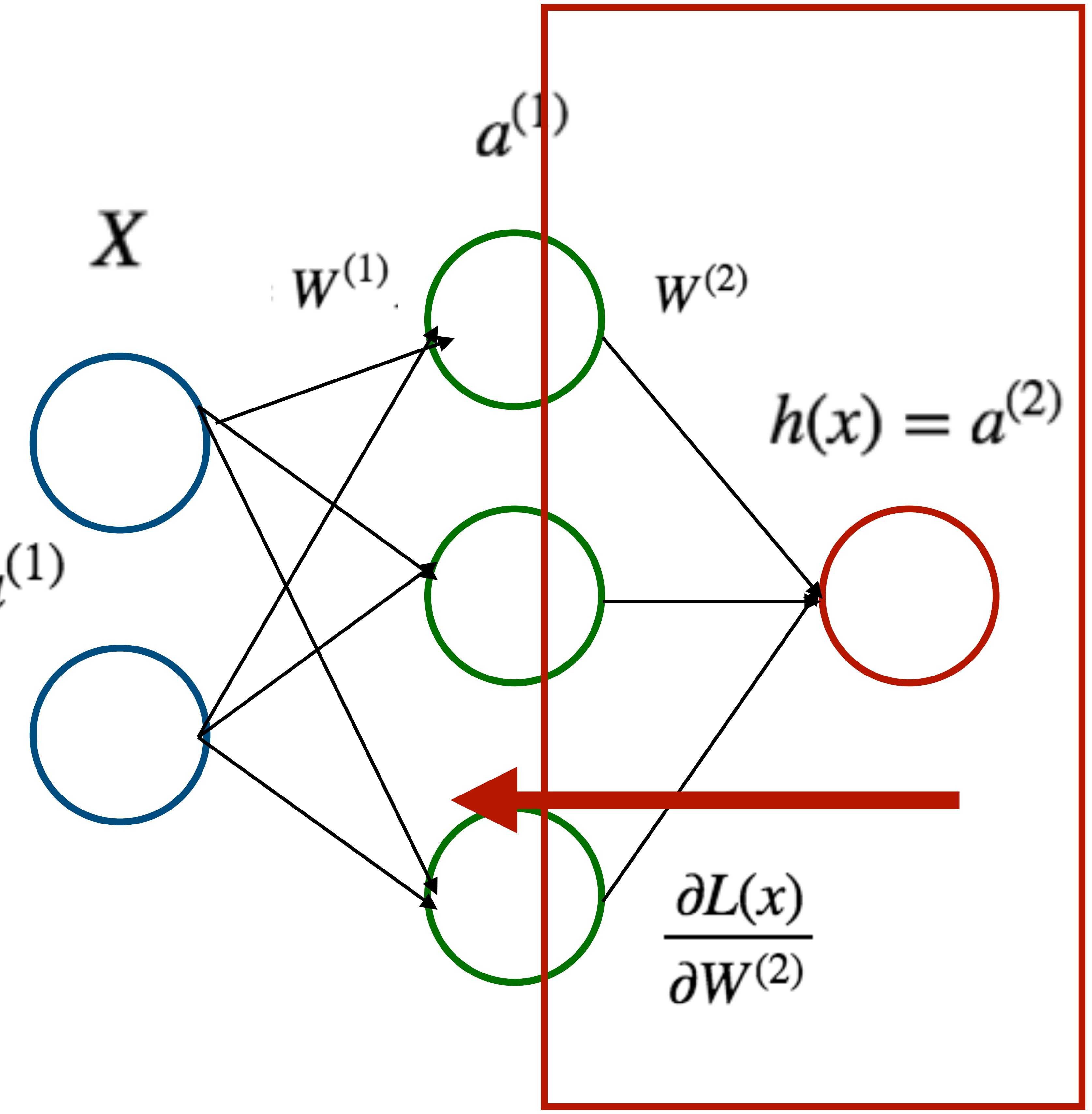
$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = a^{(1)}$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}a^{(2)}(1 - a^{(2)})a^{(1)}$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}$$

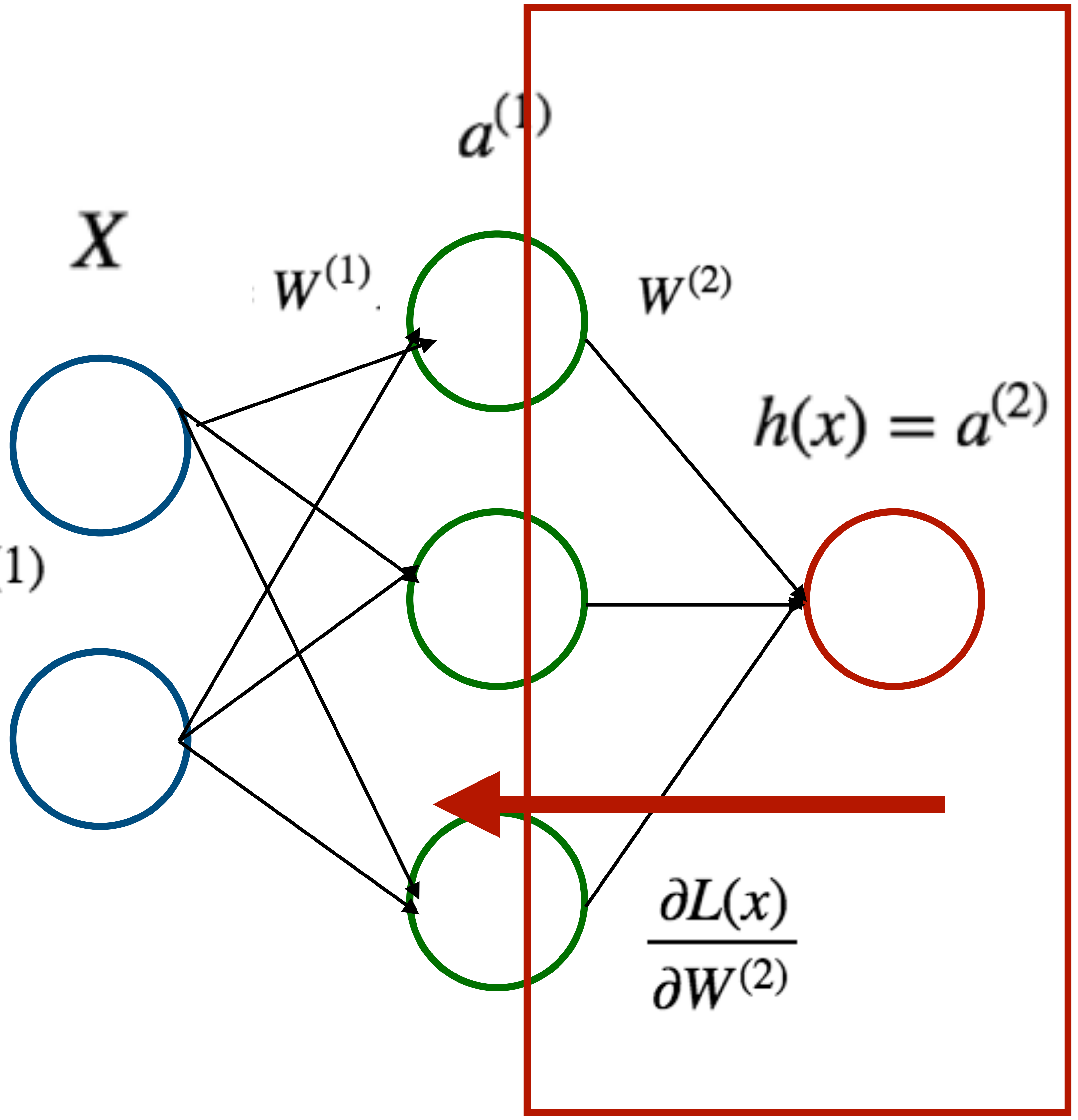
$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = a^{(1)}$$

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$= \frac{a^{(2)} - y}{\boxed{a^{(2)}(1 - a^{(2)})}} \boxed{a^{(2)}(1 - a^{(2)})} a^{(1)}$$

두 개를 없앨 수 있다



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, **chain rule을 활용해 여러 개로 쪼갤 수 있다**

그래서 **Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.**

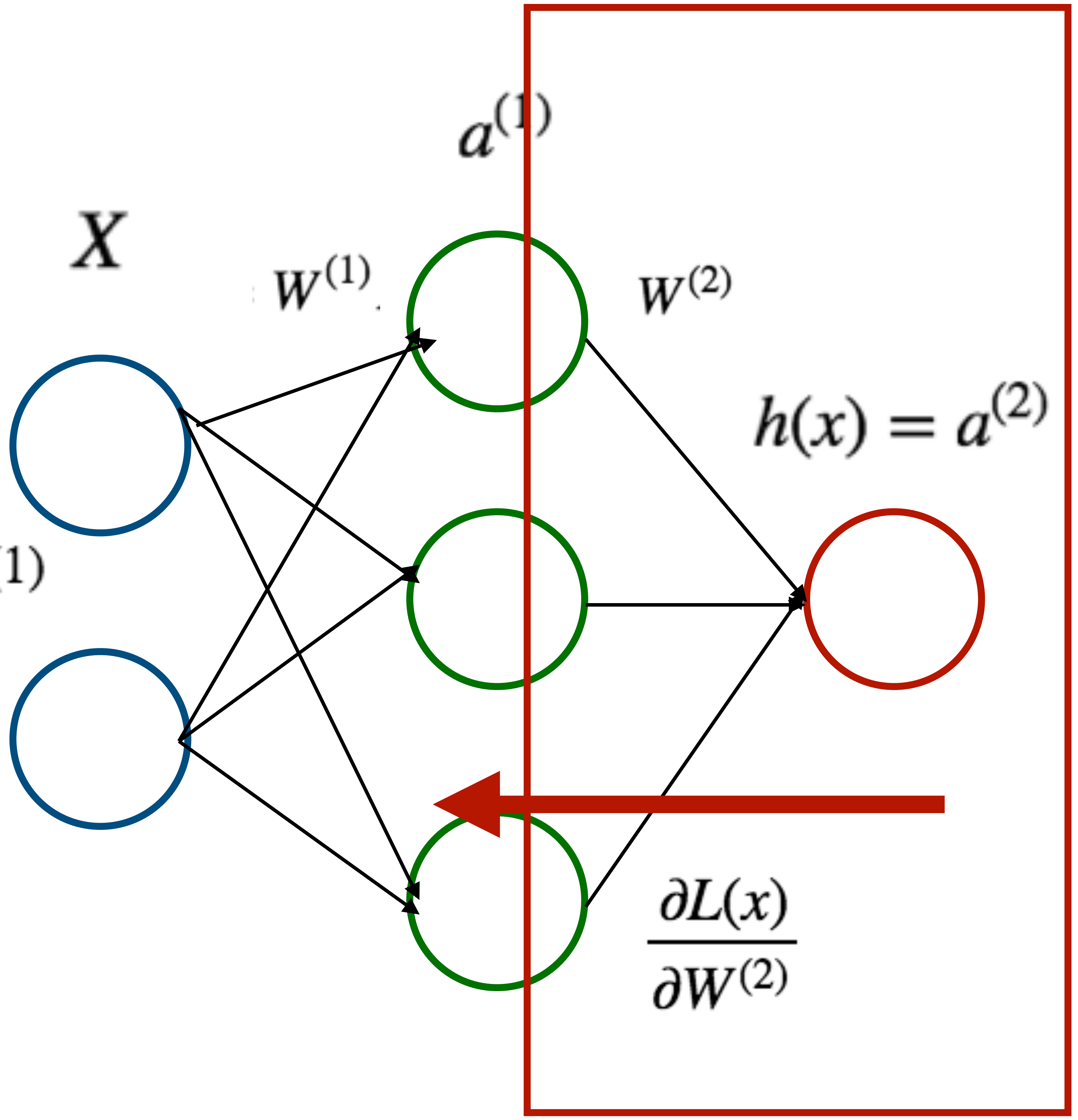
$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$

$$\frac{\partial L(x)}{\partial a^{(2)}} = \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})}$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = a^{(2)}\left(1 - a^{(2)}\right)$$

$$\frac{\partial z^{(2)}}{\partial W^{(2)}} = a^{(1)}$$

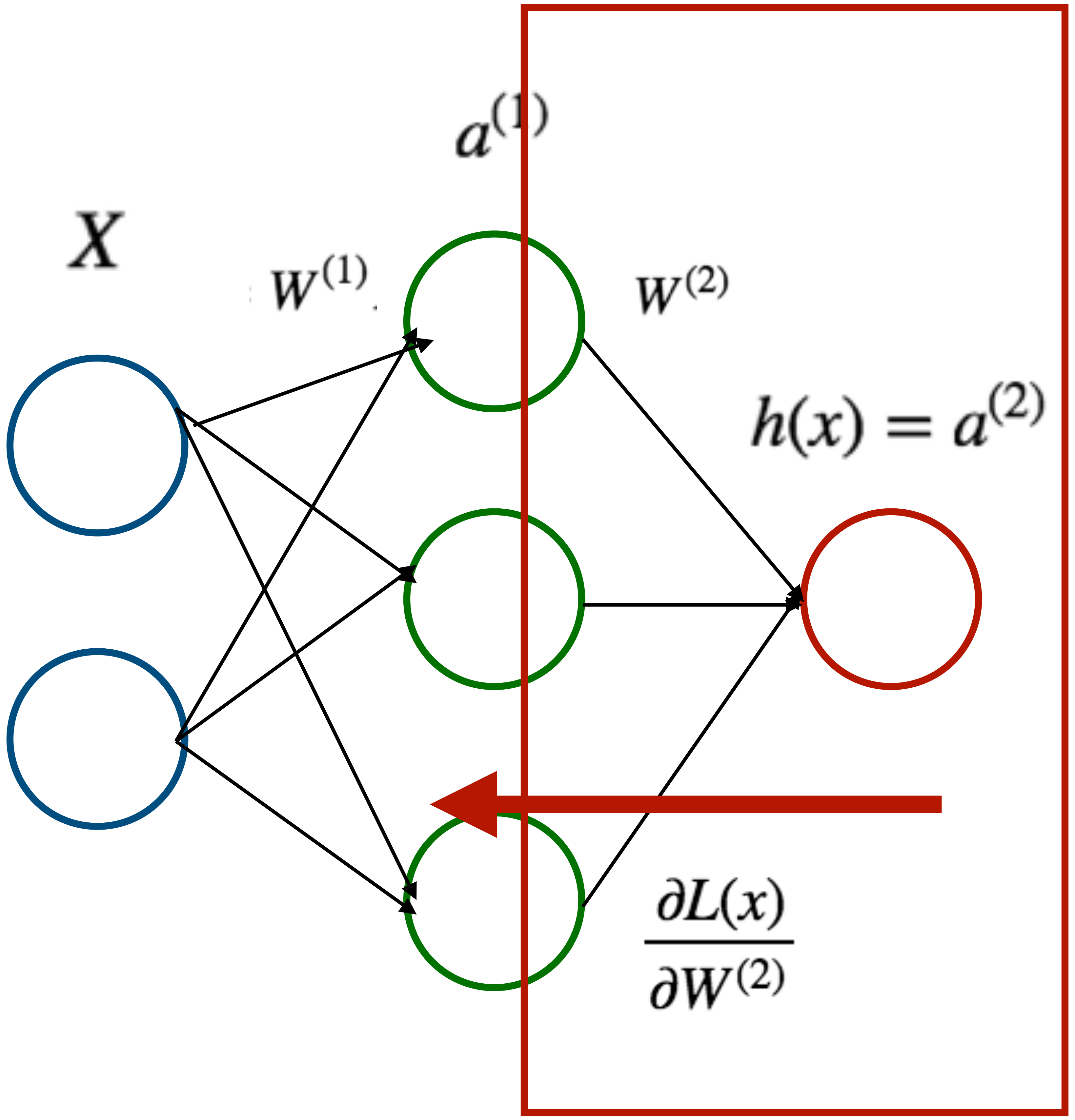
$$\begin{aligned}\frac{\partial L(x)}{\partial W^{(2)}} &= \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}} \\ &= \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})} a^{(2)}(1 - a^{(2)}) a^{(1)} \\ &= (a^{(2)} - y) a^{(1)}\end{aligned}$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(2)}}$$
$$= \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})} a^{(2)}(1 - a^{(2)})a^{(1)}$$
$$= (a^{(2)} - y)a^{(1)}$$



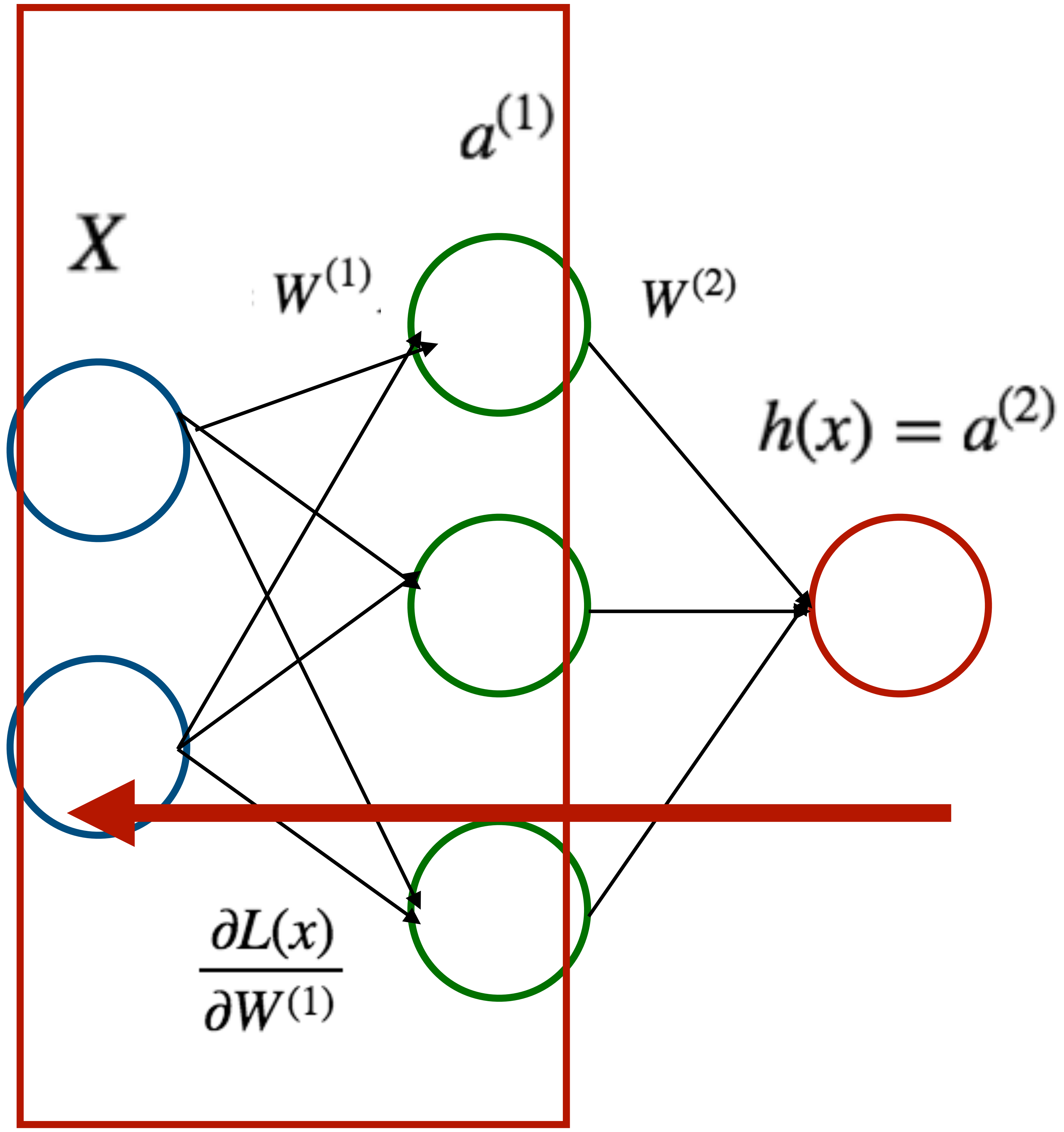
Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

$$\frac{\partial L(x)}{\partial W^{(2)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial W^{(2)}}$$
$$= \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})} a^{(2)}(1 - a^{(2)})a^{(1)}$$
$$= (a^{(2)} - y)a^{(1)}$$

$$\frac{\partial L(x)}{\partial W^{(1)}} = \frac{\partial L(x)}{\partial a^{(2)}} \frac{\partial a^{(2)}}{\partial z^{(2)}} \frac{\partial z^{(2)}}{\partial a^{(1)}} \frac{\partial a^{(1)}}{\partial z^{(1)}} \frac{\partial z^{(1)}}{\partial W^{(1)}}$$

$$= \frac{a^{(2)} - y}{a^{(2)}(1 - a^{(2)})} a^{(2)}(1 - a^{(2)})W^{(2)} a^{(1)}(1 - a^{(1)})X$$
$$= (a^{(2)} - y)W^{(2)} a^{(1)}(1 - a^{(1)})X$$



Derivation of Backpropagation

우리는 아무리 복잡한 편미분 연산이라도, chain rule을 활용해 여러 개로 쪼갤 수 있다
그래서 Loss Function의 편미분을 여러 조각을 내서 하나하나 구한 뒤 이를 곱하는 것이 backpropation의 핵심이다.

결론

$$\frac{\partial L(x)}{\partial W^{(2)}} = (a^{(2)} - y)a^{(1)}$$

$$\frac{\partial L(x)}{\partial W^{(1)}} = (a^{(2)} - y)W^{(2)}a^{(1)}(1 - a^{(1)})X$$

$$\frac{\partial L(x)}{\partial b^{(2)}} = (a^{(2)} - y)$$

$$\frac{\partial L(x)}{\partial b^{(1)}} = (a^{(2)} - y)W^{(2)}a^{(1)}(1 - a^{(1)})$$