

10주차(1/3)

다층 신경망 모델링

파이썬으로 배우는 기계학습

한동대학교
김영섭 교수

다층 신경망 모델링

- 학습 목표

- 미분의 연쇄법칙을 학습한다.
- 오차함수의 행렬 표기에서 미분하는 방법을 학습한다.
- 다층 인공 신경망의 행렬 모델을 학습한다.

- 학습 내용

- 미분의 연쇄법칙
- 오차함수의 행렬 미분
- 다층 인공 신경망 행렬 모델

연쇄법칙: 연쇄법칙이란?

- “연쇄”

연쇄법칙: 연쇄법칙이란?

- “연쇄”

$$F(x) = f(g(x))$$

연쇄법칙: 연쇄법칙이란?

- “연쇄”

$$F(x) = f(g(x))$$

$$F'(x) = f'(g(x))g'(x)$$

연쇄법칙: 연쇄법칙이란?

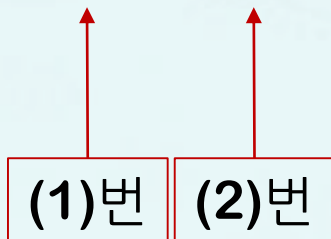
- “연쇄”

$$F(x) = f(g(x))$$

$$F'(x) = f'(g(x))g'(x)$$

$y = f(u), u = g(x)$ 일 때,

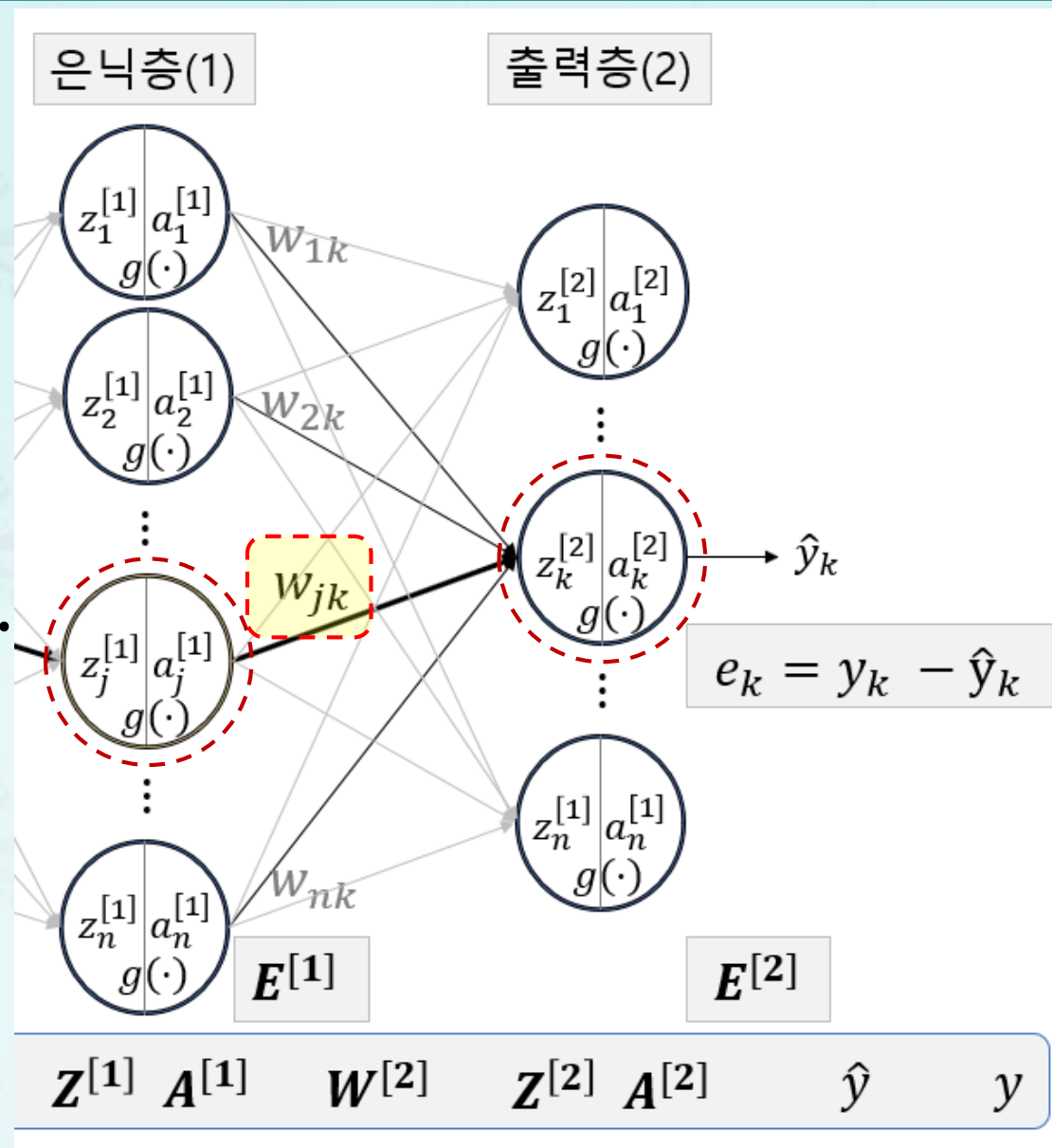
$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 3단계

$$\begin{aligned}
 \frac{\partial E}{\partial w_{jk}} &= -(y_k - \hat{y}_k) \cdot \frac{\partial}{\partial w_{jk}} g(z_k) \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \frac{\partial z_k}{\partial w_{jk}} \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \frac{\partial}{\partial w_{jk}} \left(\sum_j w_{jk} \cdot \right. \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j
 \end{aligned}$$

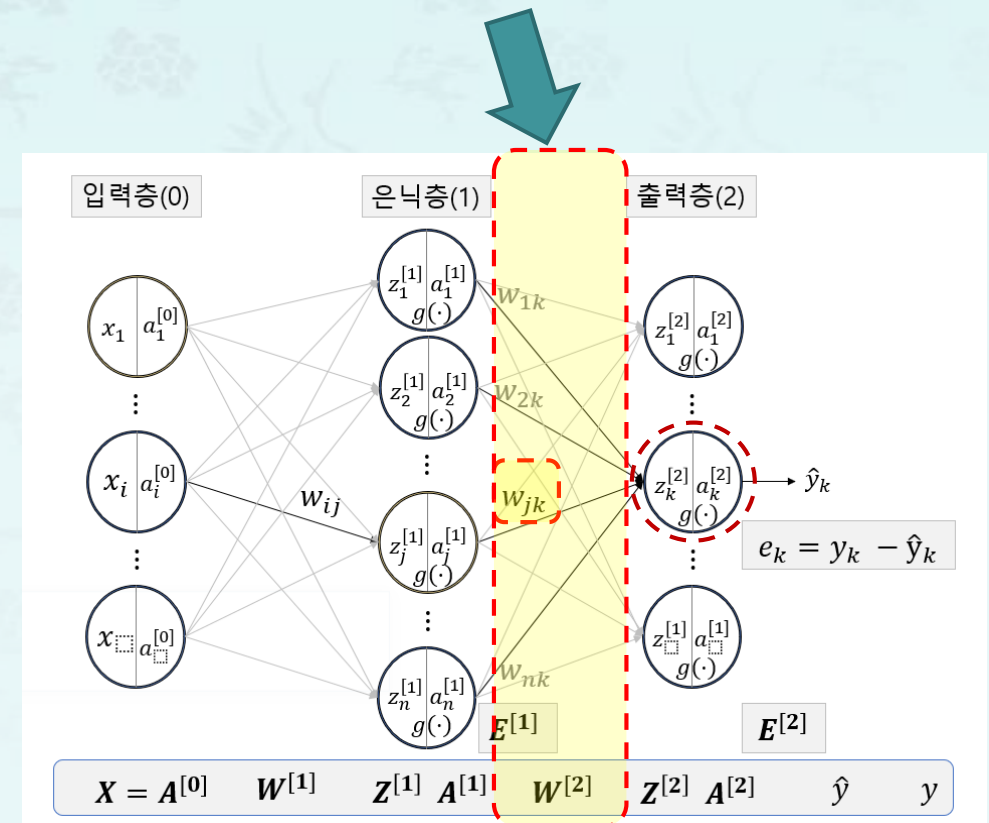


오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 3단계

$$\begin{aligned}
 \frac{\partial E}{\partial w_{jk}} &= -(y_k - \hat{y}_k) \cdot \frac{\partial}{\partial w_{jk}} g(z_k) \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \frac{\partial z_k}{\partial w_{jk}} \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \frac{\partial}{\partial w_{jk}} \left(\sum_j w_{jk} \cdot a_j \right) \\
 &= -(y_k - \hat{y}_k) \cdot g'(z_k) \cdot a_j
 \end{aligned}$$

$$\frac{\partial E}{\partial W^{[2]}} = -E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계

$$\begin{aligned} W^{[2]} &:= W^{[2]} - \alpha \Delta W^{[2]} \\ &= W^{[2]} - \alpha \frac{\partial E}{\partial W^{[2]}} \end{aligned}$$

■ 2단계

$$\begin{aligned} w_{jk}^{[2]} &:= w_{jk}^{[2]} - \alpha \Delta w_{jk}^{[2]} \\ &= w_{jk}^{[2]} - \alpha \frac{\partial E}{\partial w_{jk}^{[2]}} \end{aligned}$$

■ 3단계

$$\frac{\partial E}{\partial w_{jk}} = -(y_k - \hat{y}_k) \cdot \frac{\partial}{\partial w_{jk}} g(z_k)$$

■ 4단계

$$\Delta w_{jk}^{[2]} = \frac{\partial E}{\partial w_{jk}} = \boxed{-(y_k - \hat{y}_k)} \cdot \boxed{g'(z_k)} \cdot \boxed{a_j}$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}} = \boxed{-E^{[2]}} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$

■ 결론

$$\begin{aligned} W^{[2]} &:= W^{[2]} - \alpha \Delta W^{[2]} \\ &= W^{[2]} - \alpha \frac{\partial E}{\partial W^{[2]}} \\ &= \boxed{W^{[2]} + E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1]T}} \end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

- 1단계

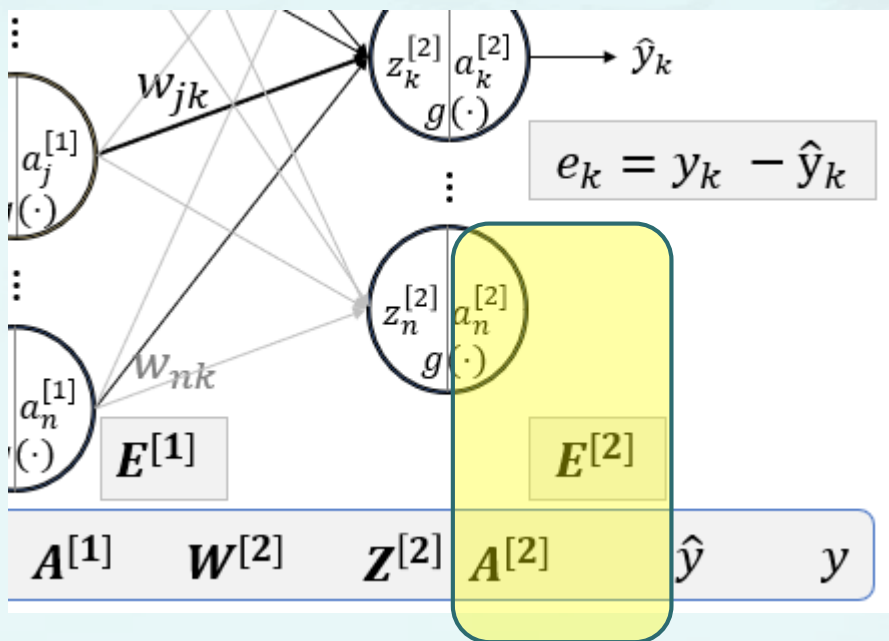
$$\begin{aligned} W^{[2]} &:= W^{[2]} - \alpha \Delta W^{[2]} \\ &= W^{[2]} - \alpha \frac{\partial E}{\partial W^{[2]}} \end{aligned}$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

- 1단계

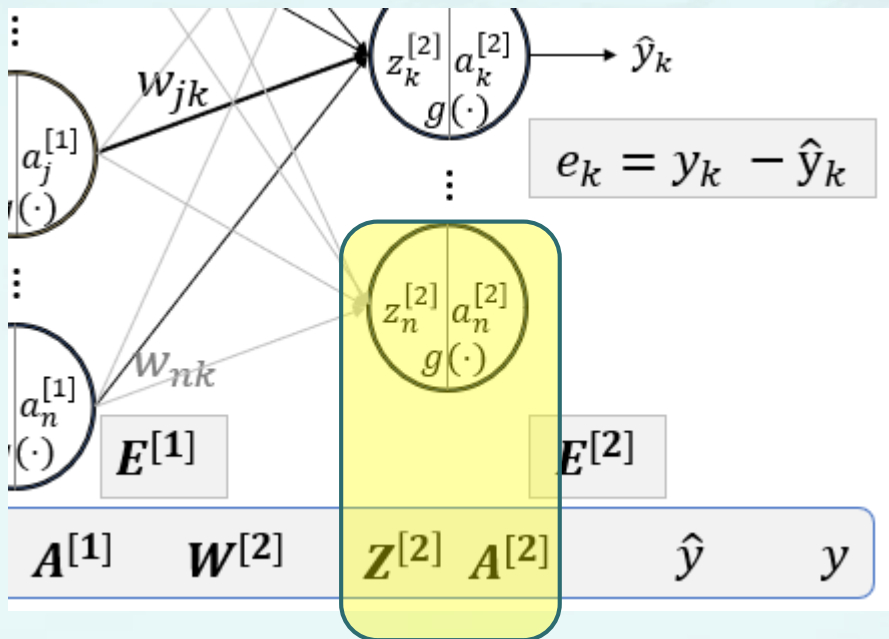
$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}}\end{aligned}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

- 1단계

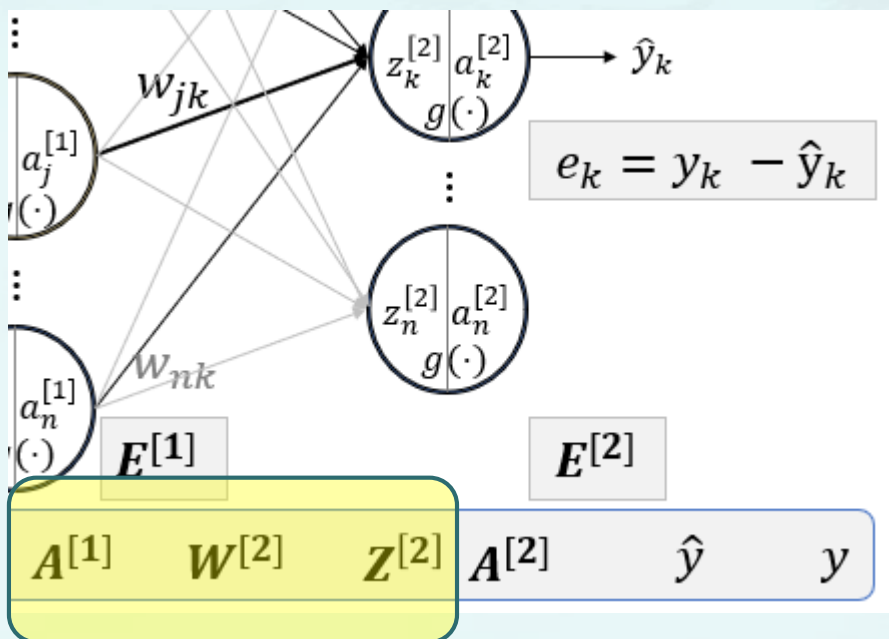
$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}\end{aligned}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

- 1단계

$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}\end{aligned}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

- 1단계

$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \boxed{\frac{\partial E}{\partial A^{[2]}}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}\end{aligned}$$

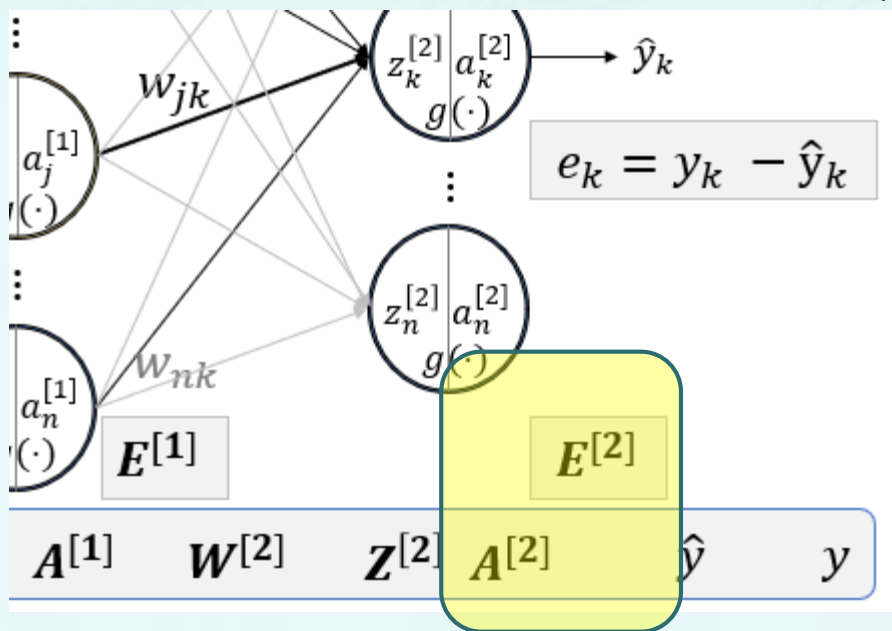
오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계

$$E = \frac{1}{2} (A^{[2]} - Y)^2$$

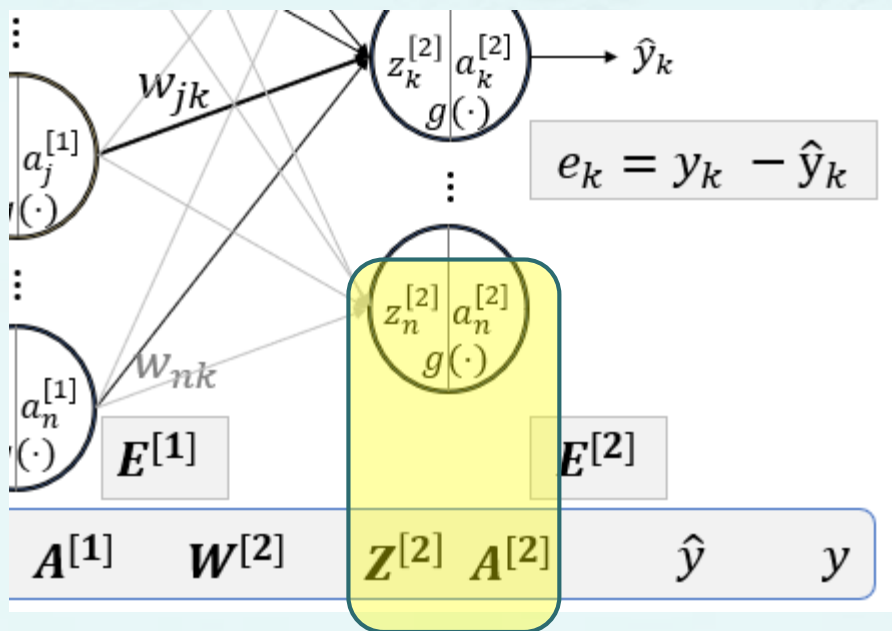
$$E^{[2]} := (A^{[2]} - Y)$$

$$\begin{aligned} \Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \end{aligned}$$



오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

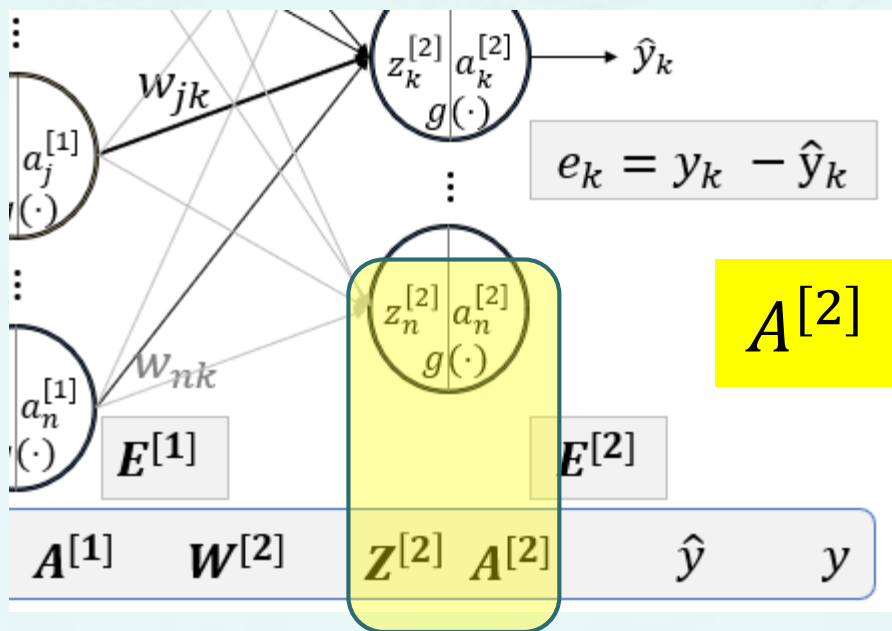
■ 1단계



$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}\end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계



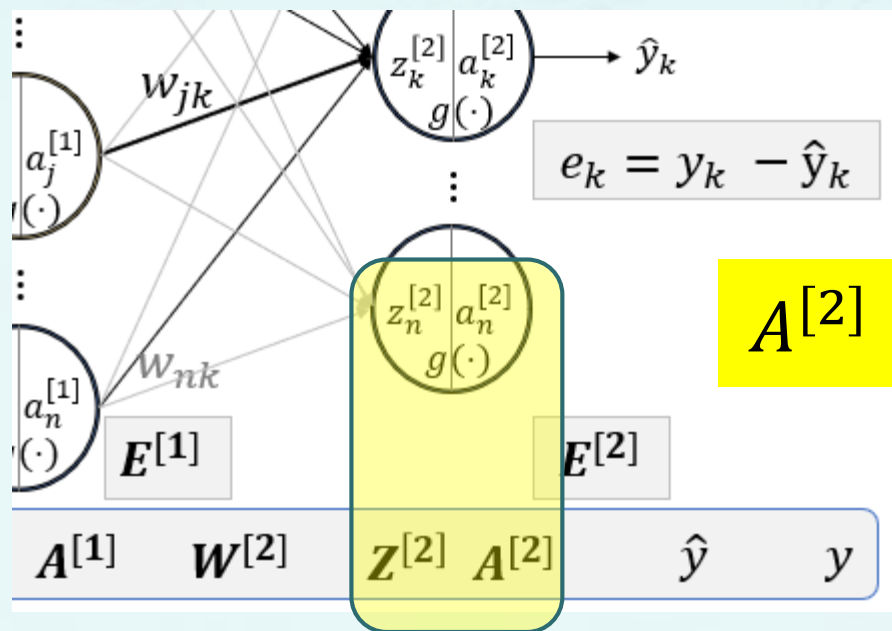
$$A^{[2]} = g(Z^{[2]})$$

$$\begin{aligned} \Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계

$$E^{[2]} \cdot \sigma(Z^{[2]})(1 - \sigma(Z^{[2]})) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$



$$A^{[2]} = g(Z^{[2]})$$

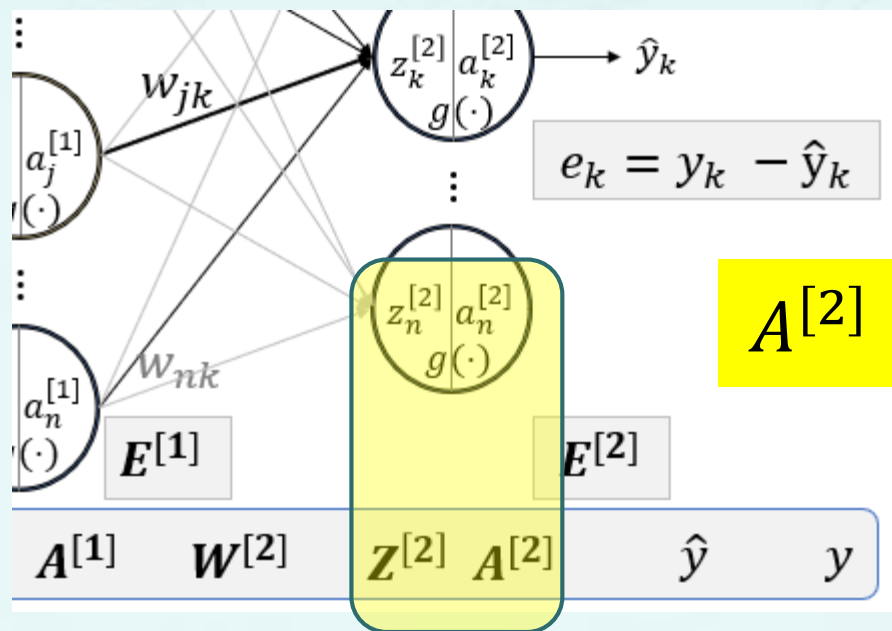
$$\begin{aligned} \Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계

$$E^{[2]} \cdot \sigma(Z^{[2]})(1 - \sigma(Z^{[2]})) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

$$= E^{[2]} \cdot A^{[2]}(1 - A^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$



$$A^{[2]} = g(Z^{[2]})$$

$$\Delta W^{[2]} = \frac{\partial E}{\partial W^{[2]}}$$

$$= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}}$$

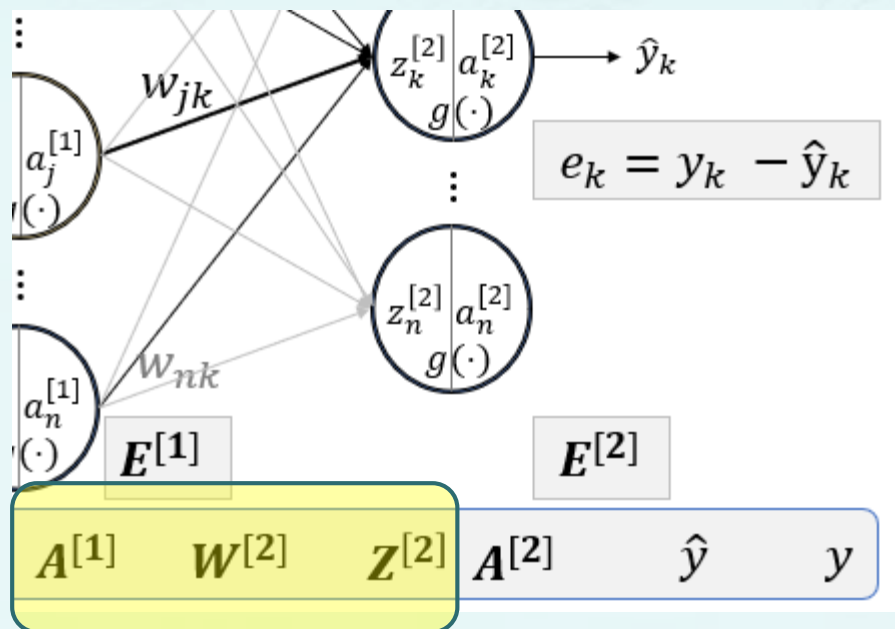
$$= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

$$= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

$$= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

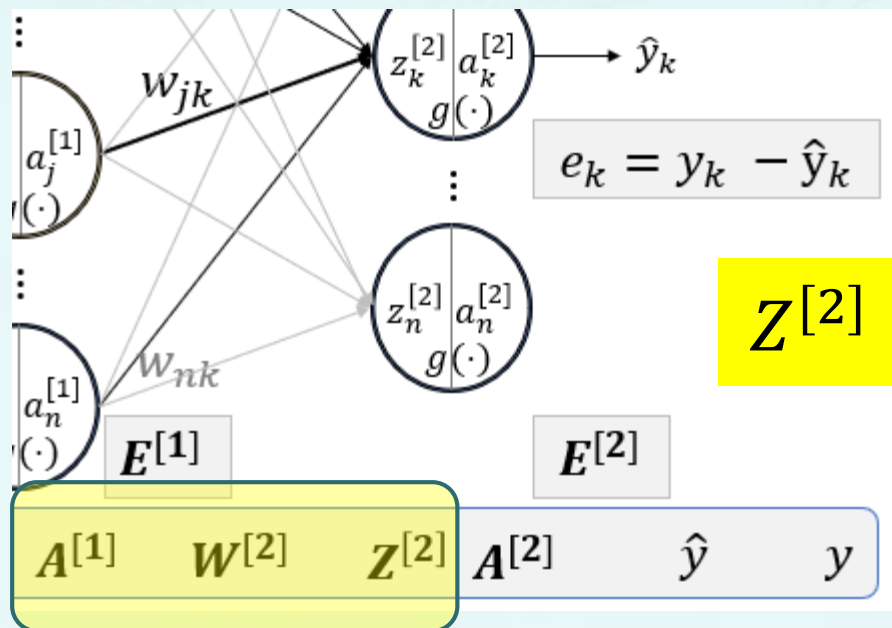
■ 1단계



$$\begin{aligned}
 \Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\
 &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\
 &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\
 &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\
 &= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}}
 \end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

■ 1단계



$$Z^{[2]} = W^{[2]} A^{[1]}$$

$$\begin{aligned} \Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\ &= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\ &= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \end{aligned}$$

오차함수의 행렬 미분: $W^{[2]}$ 의 오차함수 미분

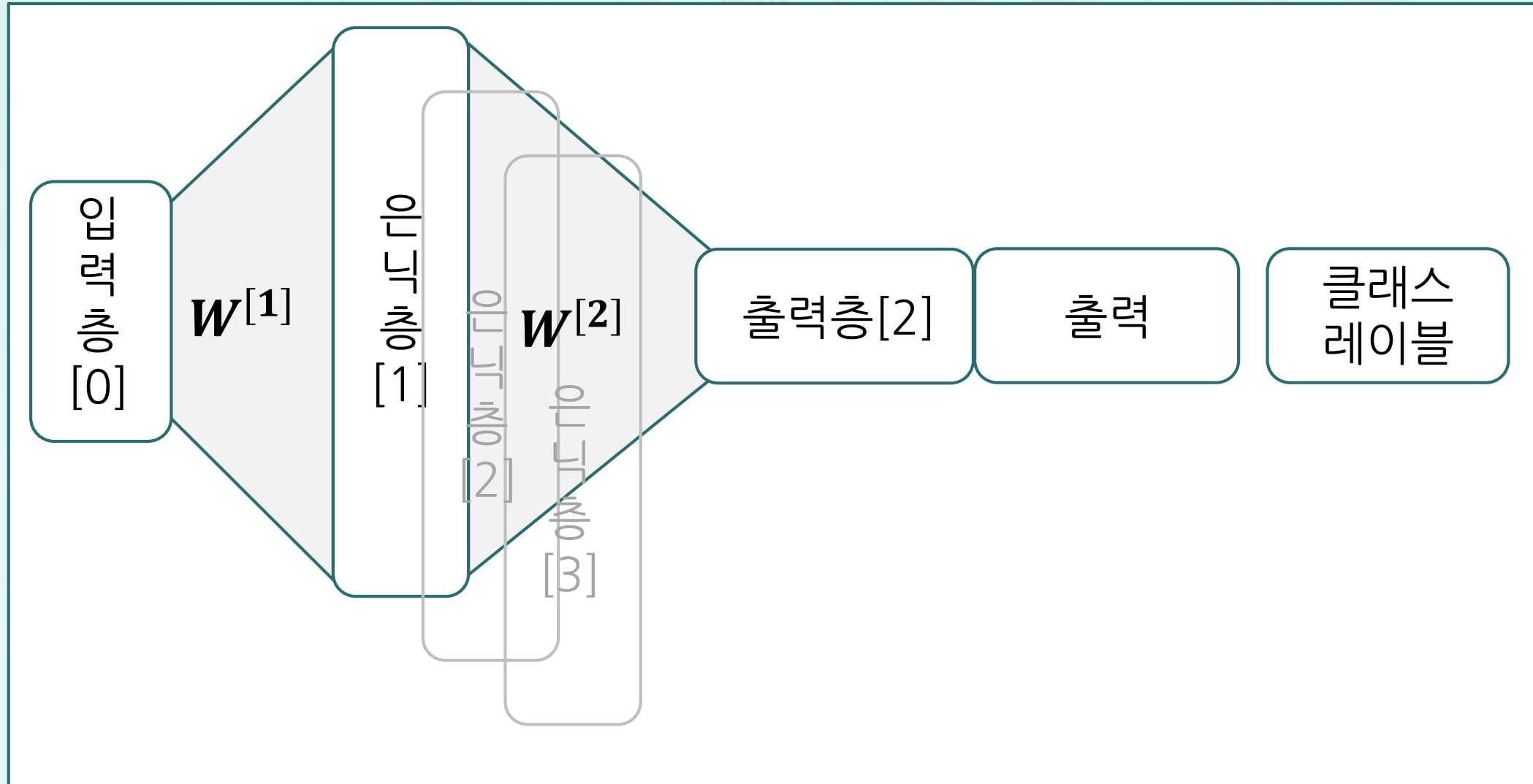
- 1단계

$$\begin{aligned}\Delta W^{[2]} &= \frac{\partial E}{\partial W^{[2]}} \\&= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial W^{[2]}} \\&= \frac{\partial E}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\&= E^{[2]} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}} \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\&= E^{[2]} \cdot g'(Z^{[2]}) \cdot \frac{\partial Z^{[2]}}{\partial W^{[2]}} \\&= E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1] \cdot T}\end{aligned}$$

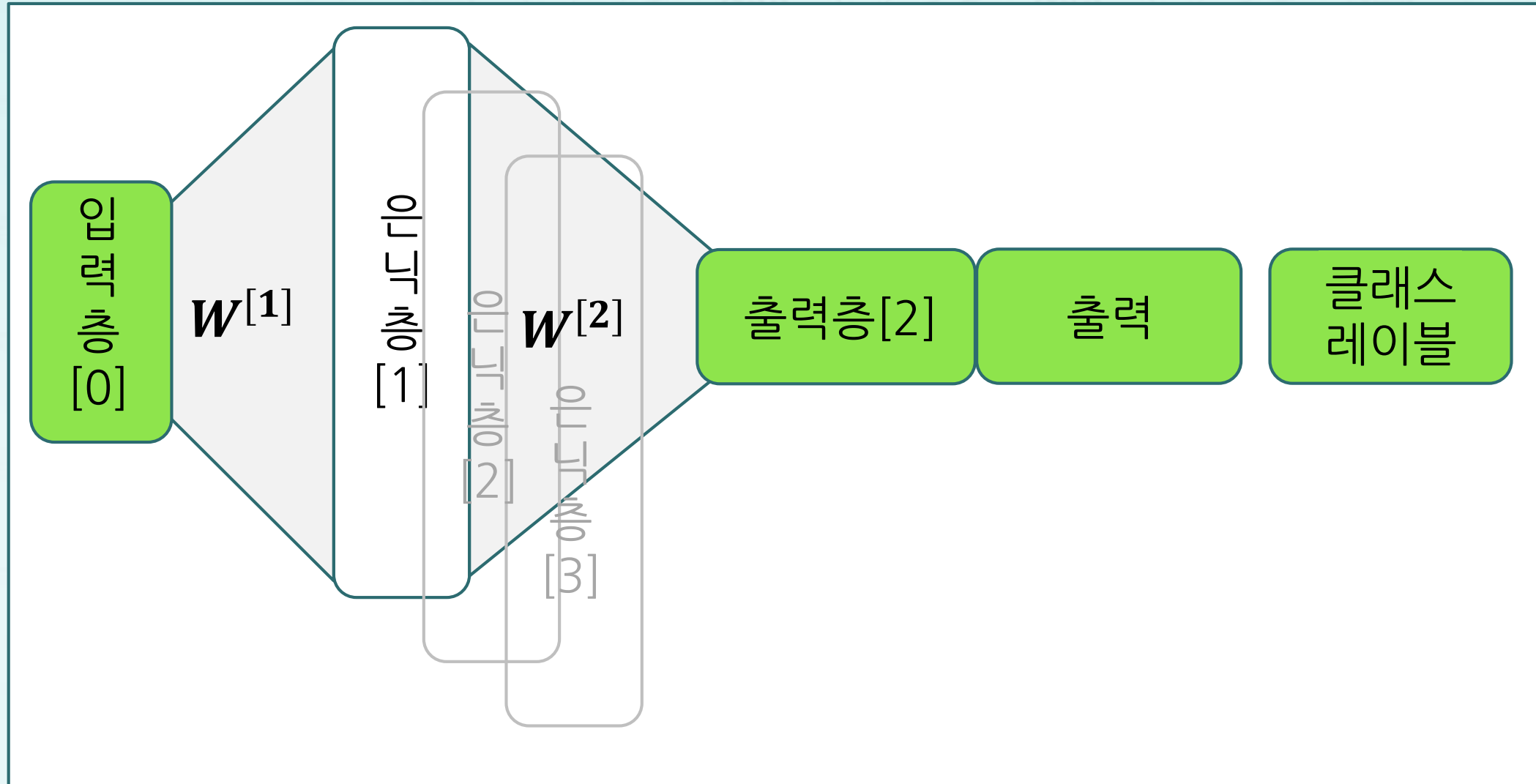
다층 인공신경망 행렬 모델:

- 다층 신경망의 구조

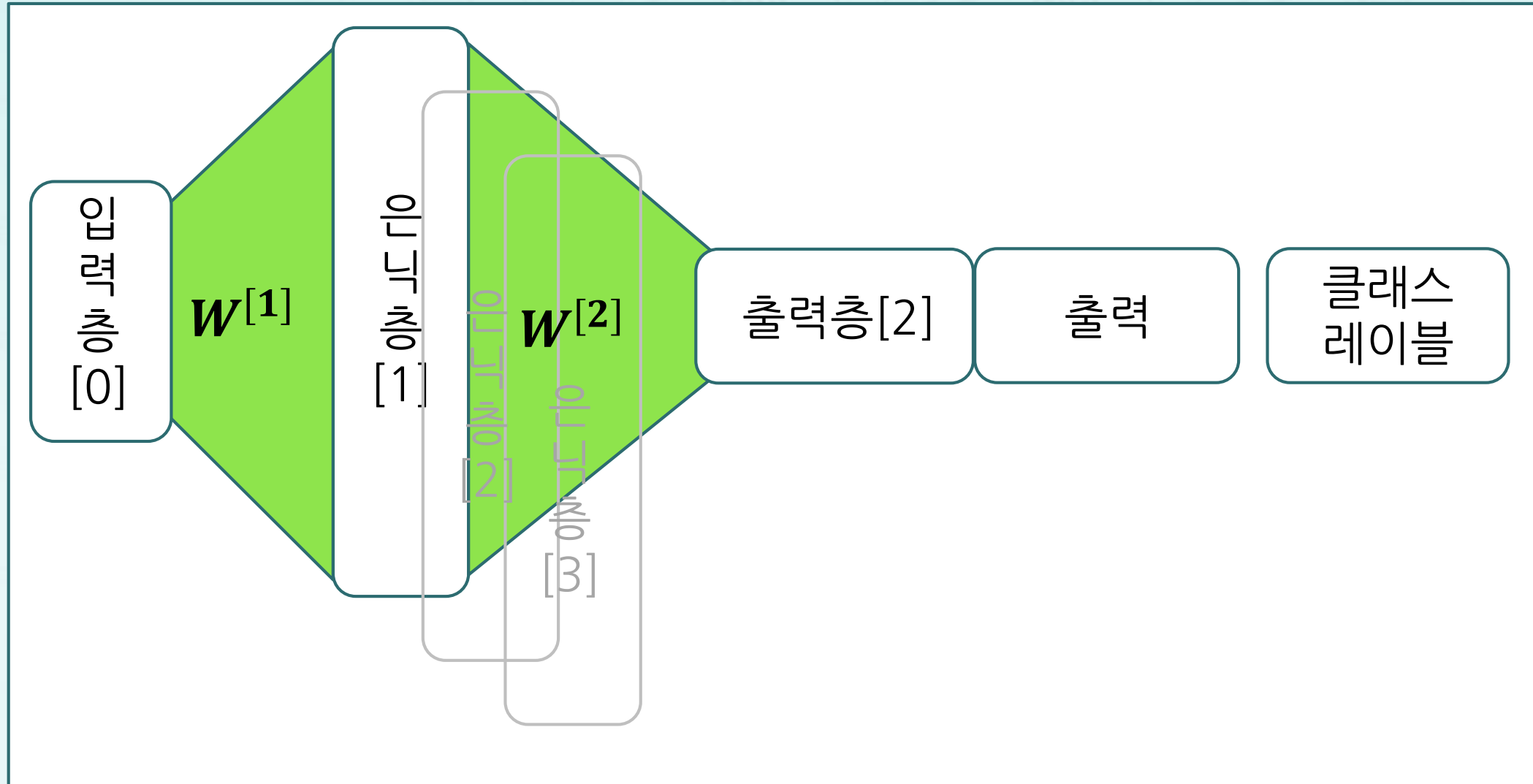
다층 인공신경망 행렬 모델:



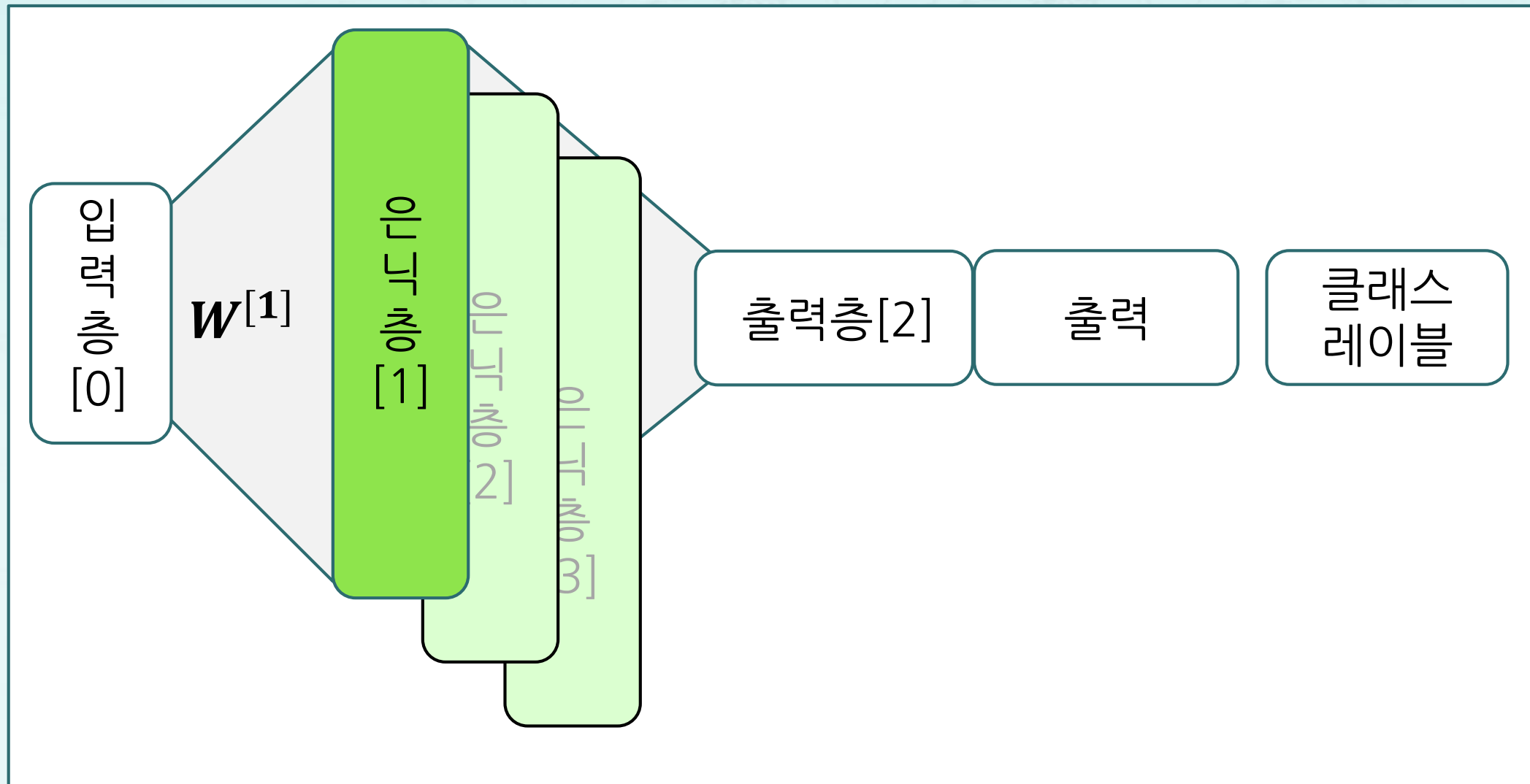
다층 인공신경망 행렬 모델: 입력층, 출력층



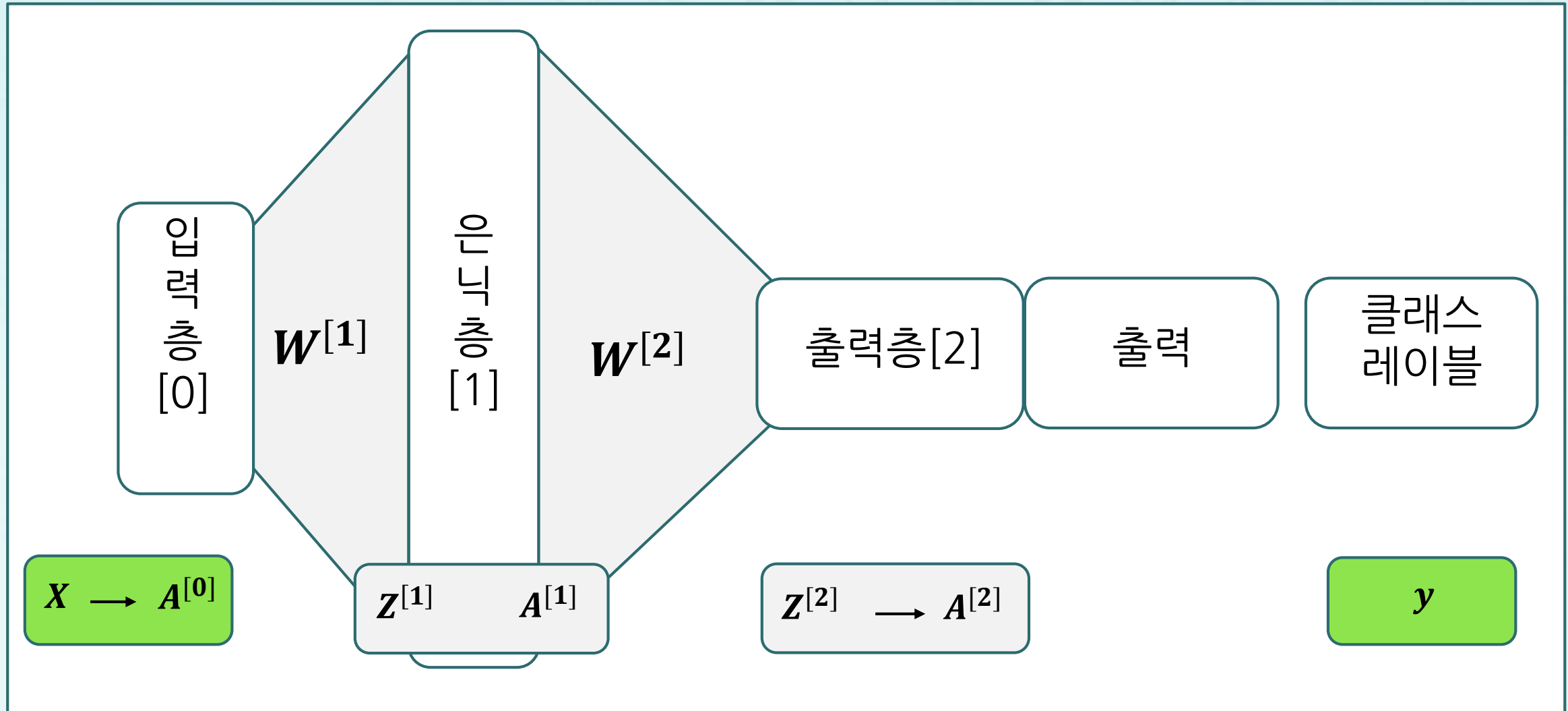
다층 인공신경망 행렬 모델: 가중치



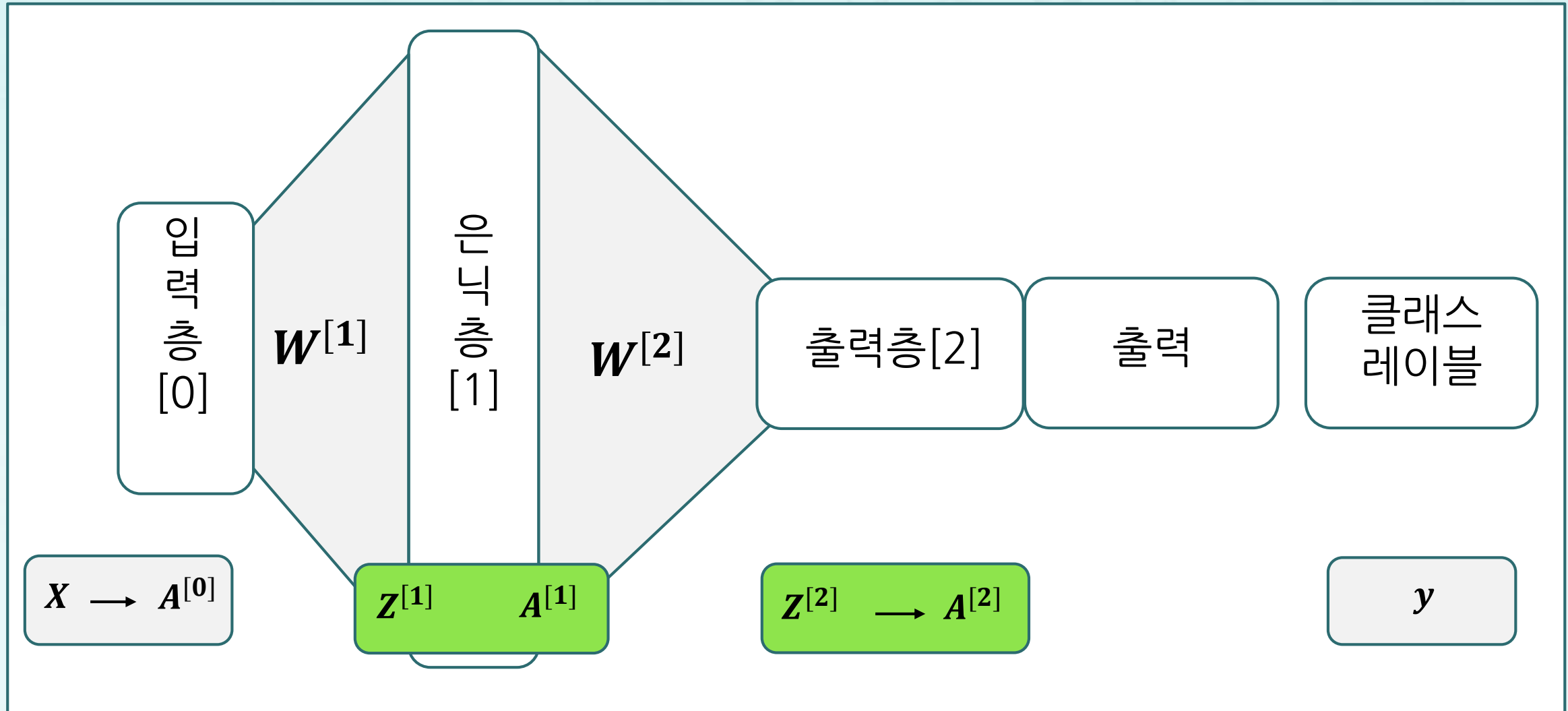
다층 인공신경망 행렬 모델: 은닉층



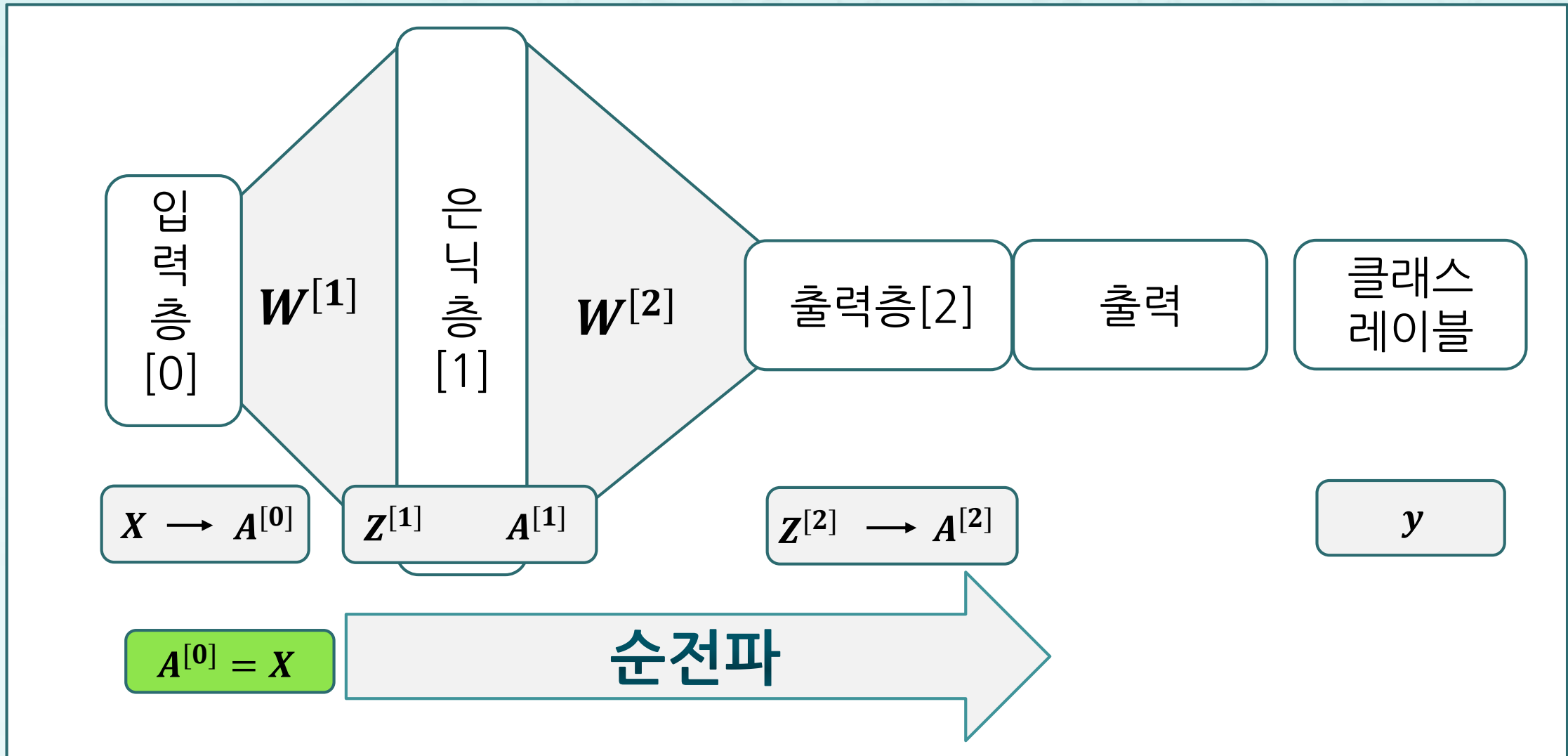
다층 인공신경망 행렬 모델: 입력과 출력



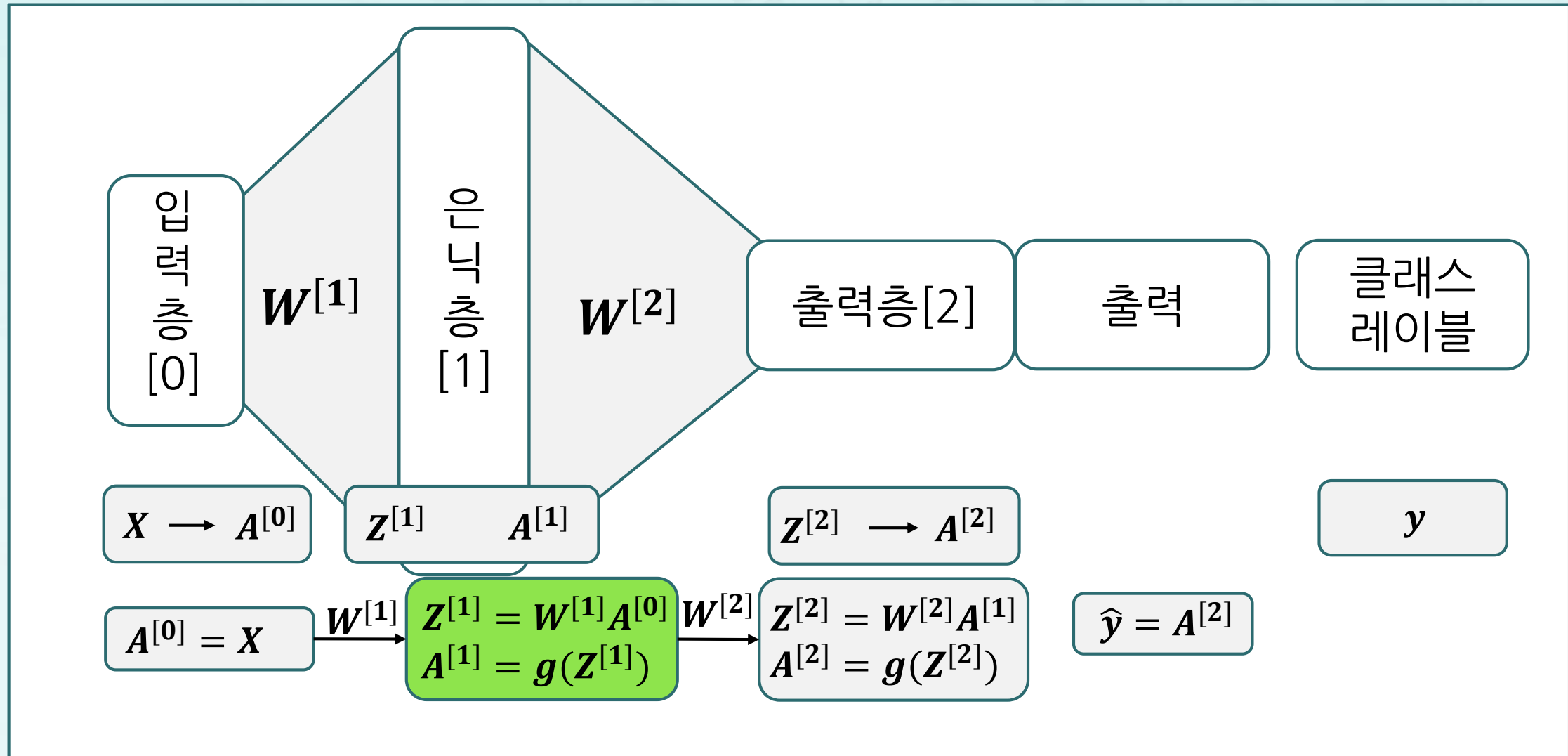
다층 인공신경망 행렬 모델: 입력과 출력



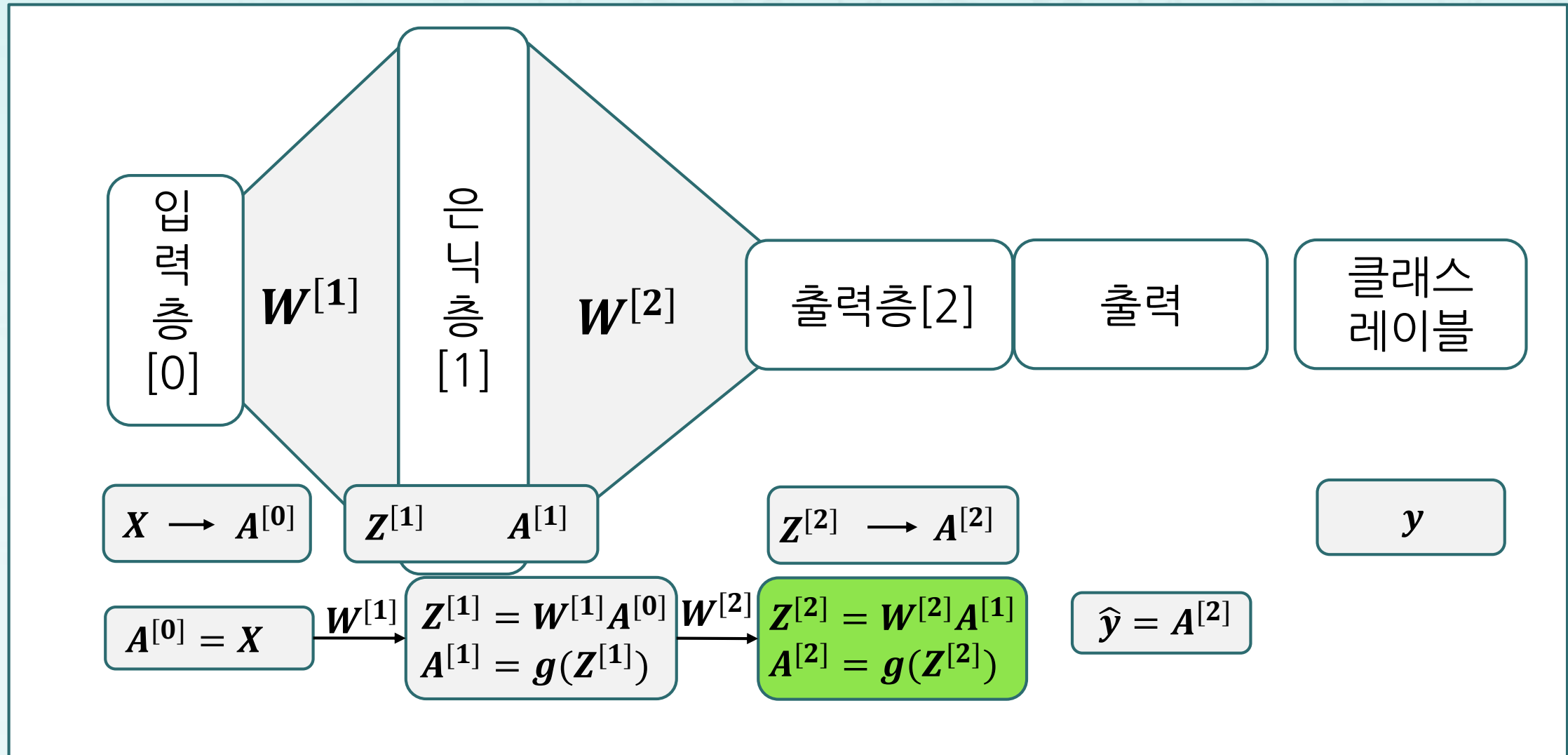
다층 인공신경망 행렬 모델: 순전파



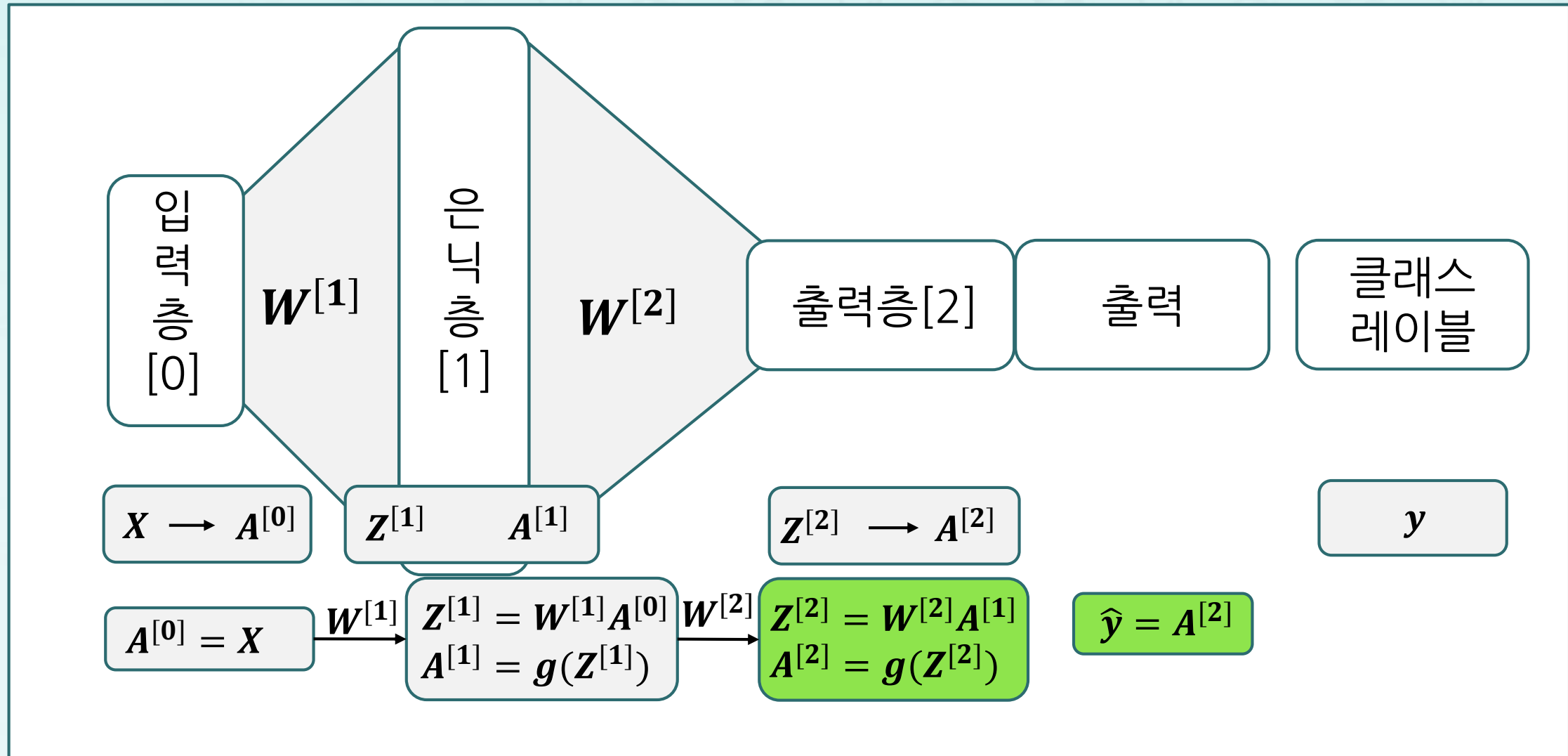
다층 인공신경망 행렬 모델: 순전파



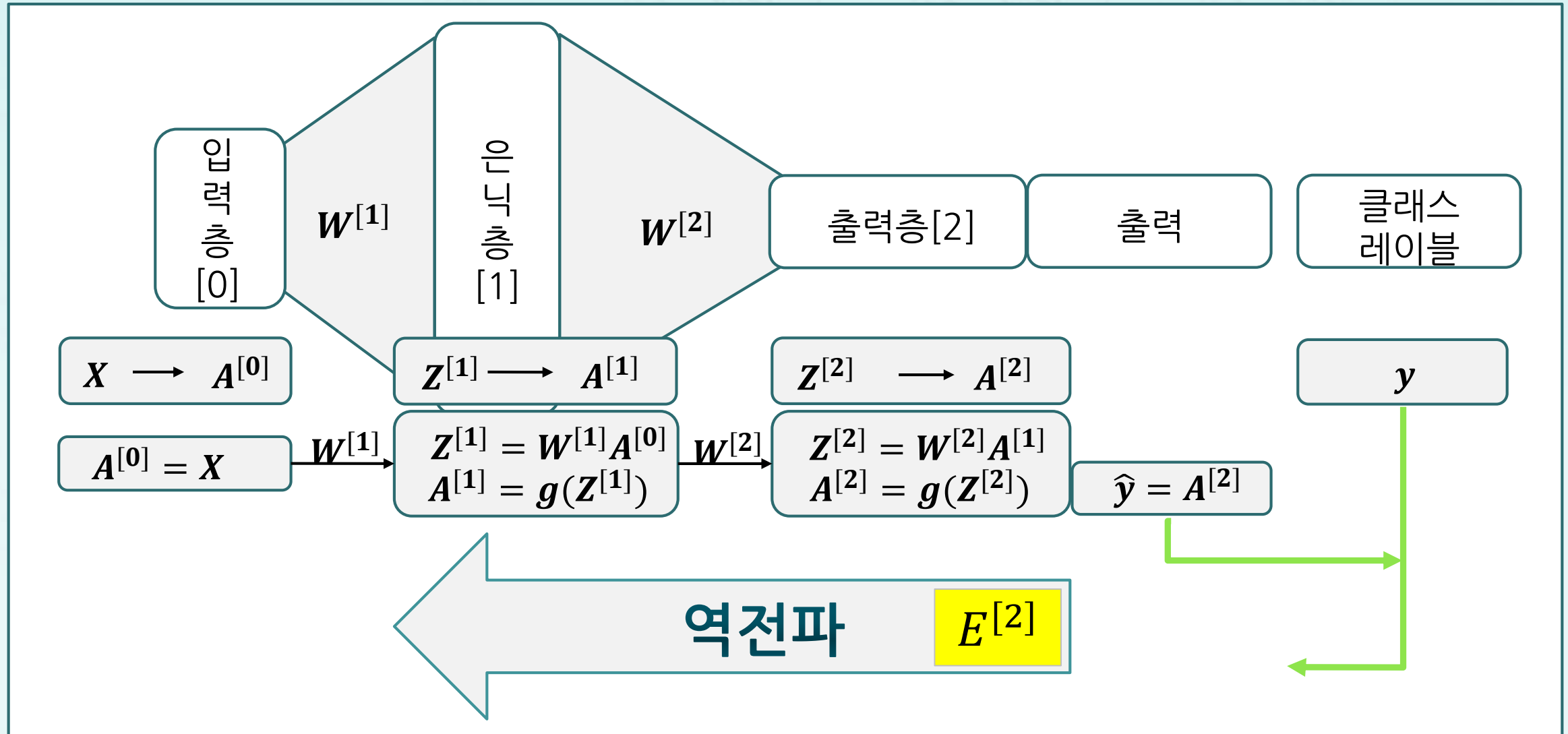
다층 인공신경망 행렬 모델: 순전파



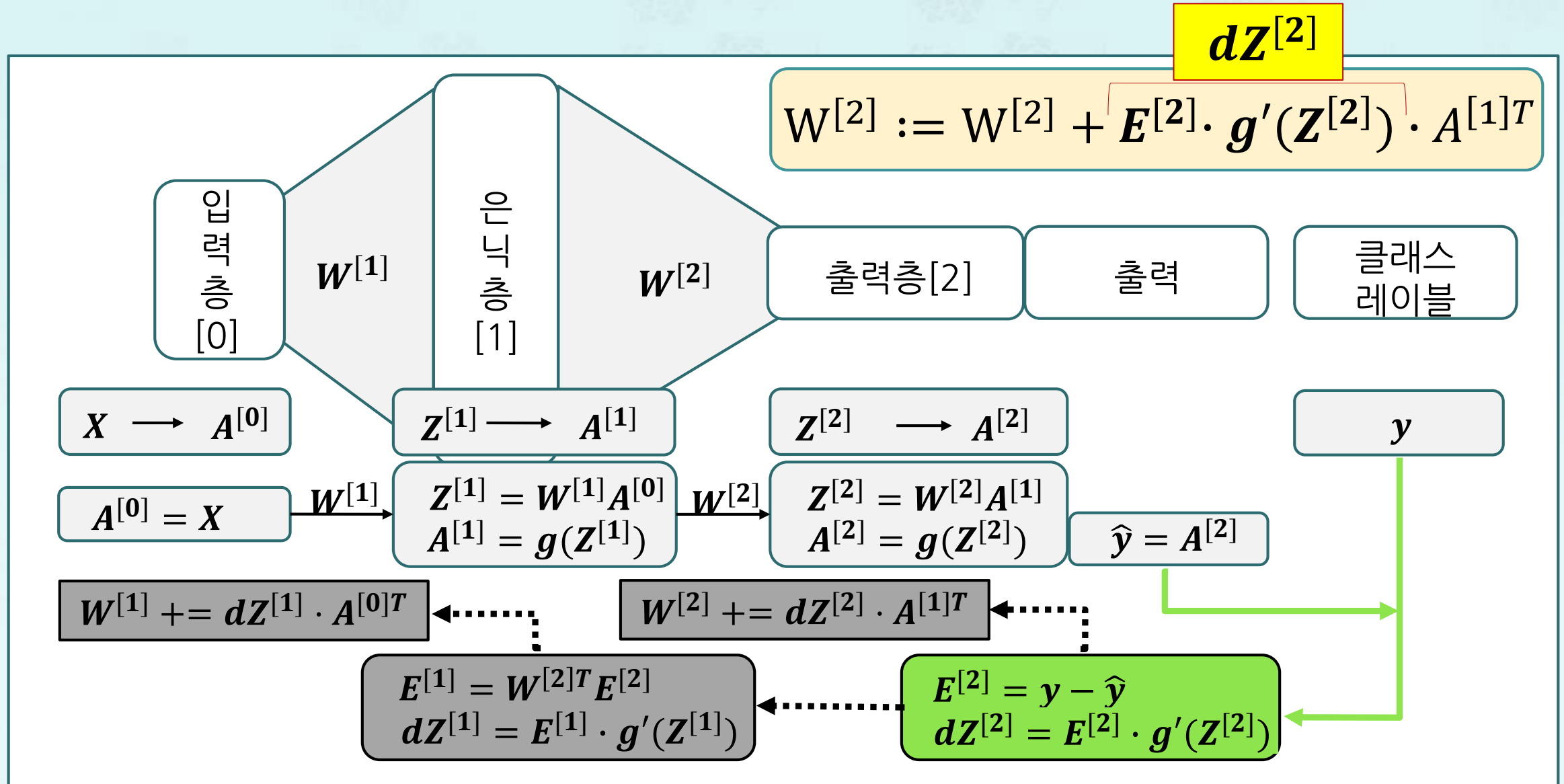
다층 인공신경망 행렬 모델: 순전파



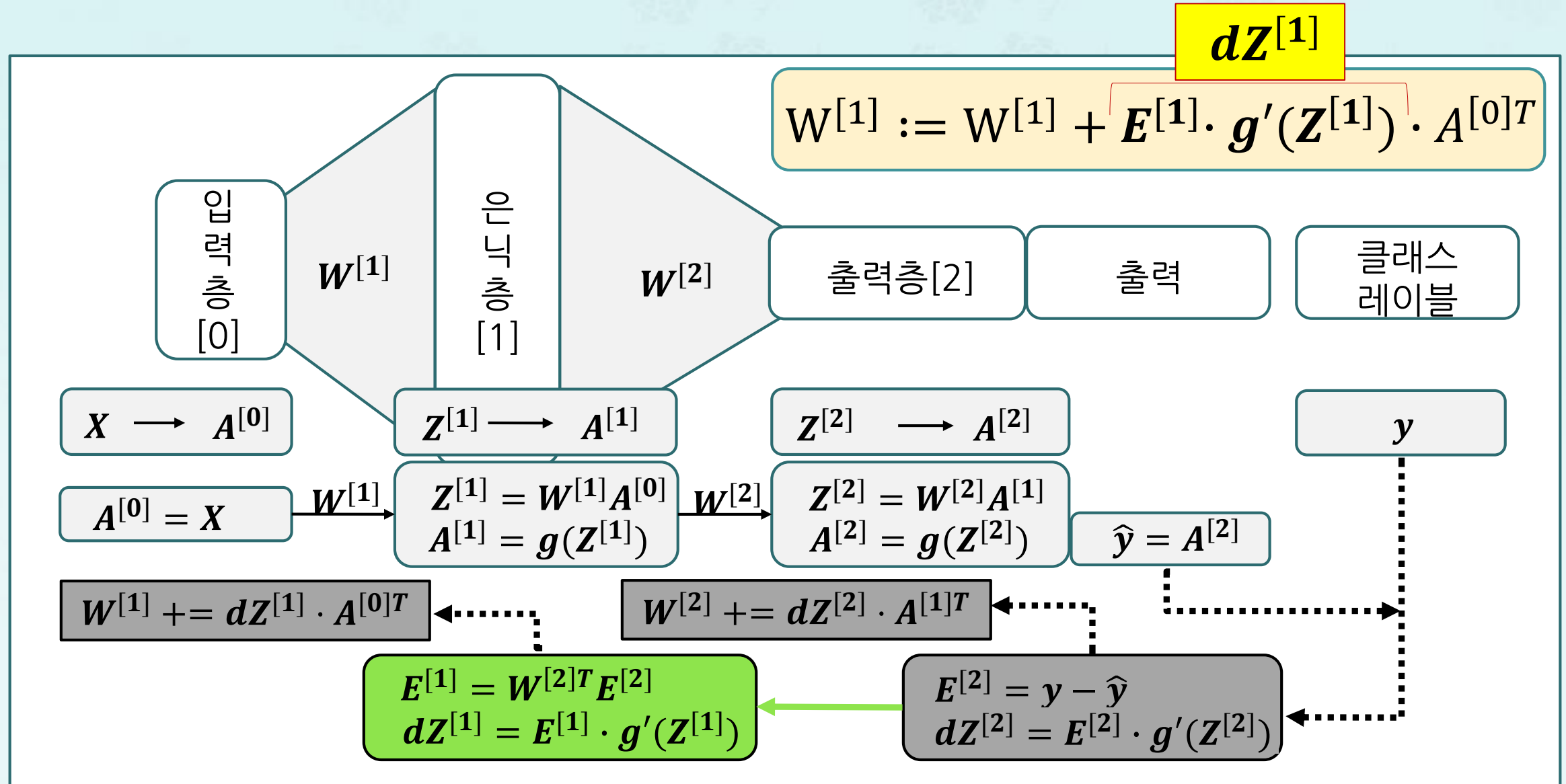
다층 인공신경망 행렬 모델: 역전파



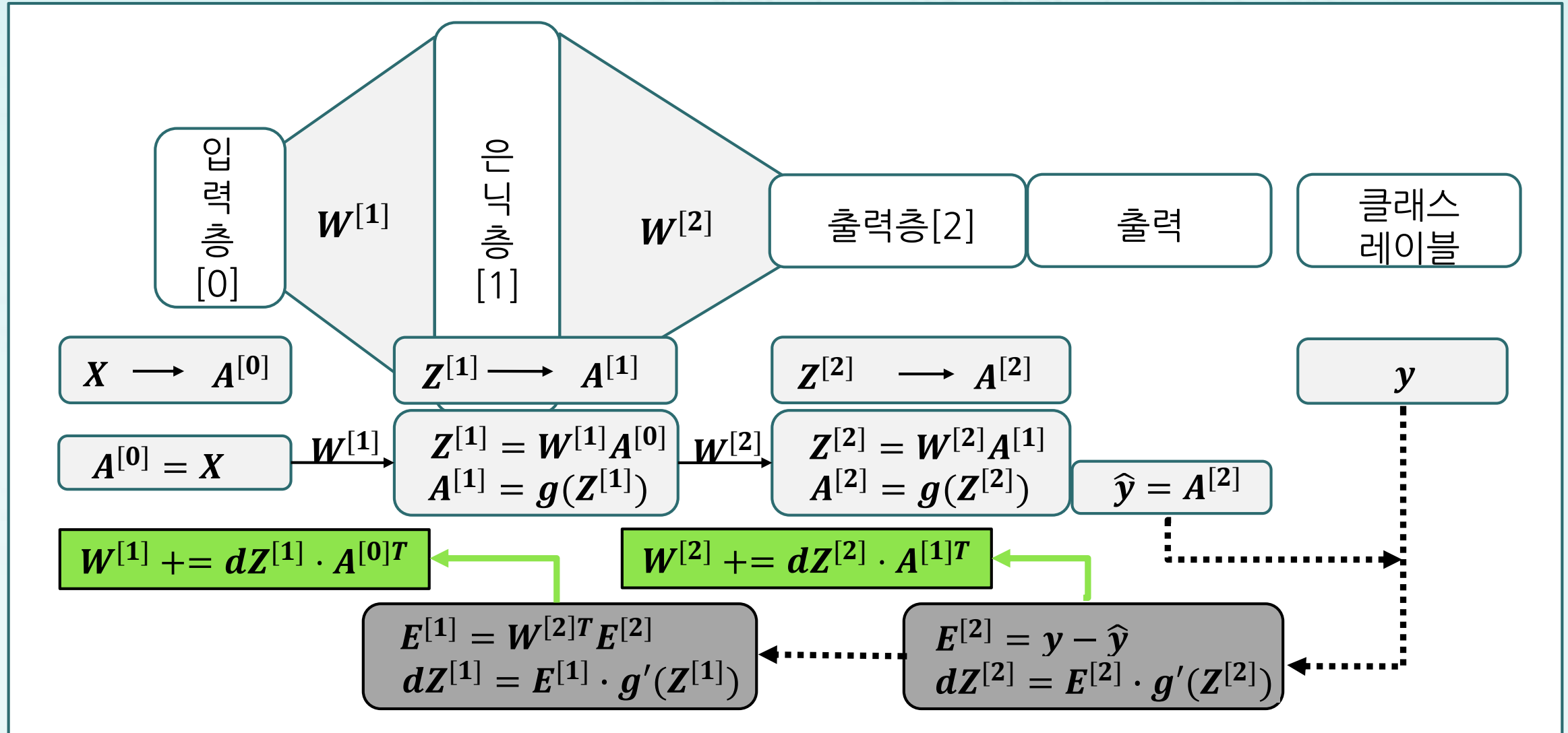
다층 인공신경망 행렬 모델: 역전파



다층 인공신경망 행렬 모델: 역전파



다층 인공신경망 행렬 모델: 역전파



역전파 2: 역전파의 가중치 조정

- 최종 결과

$$W^{[2]} := W^{[2]} - \alpha \Delta W^{[2]}$$

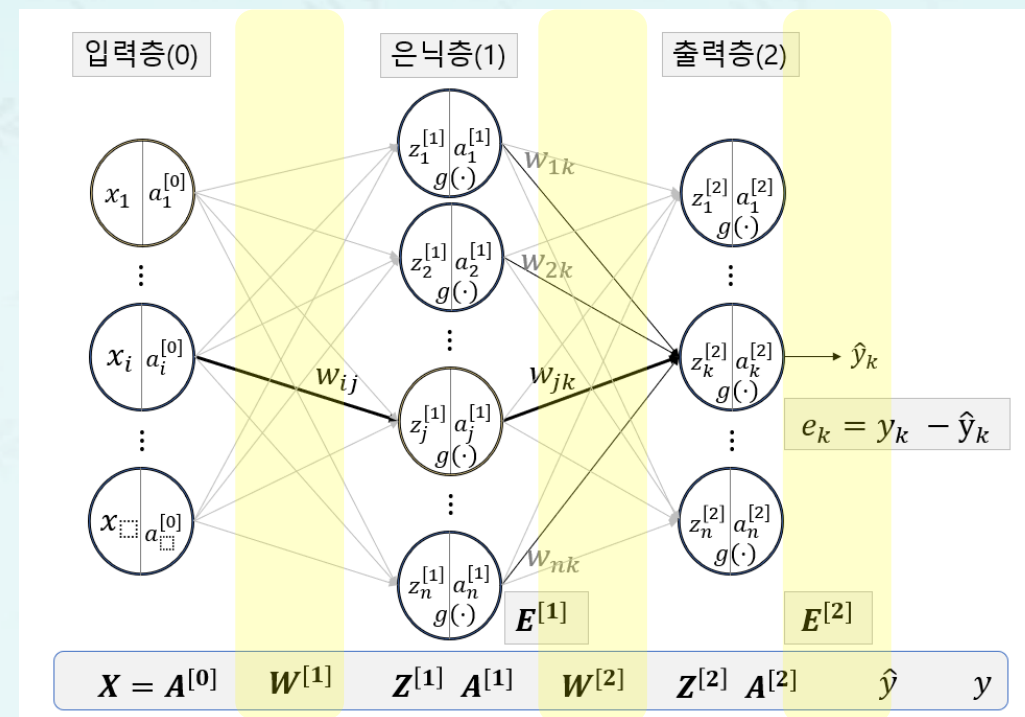
$$= W^{[2]} - \alpha \frac{\partial E}{\partial W^{[2]}}$$

$$= W^{[2]} + E^{[2]} \cdot g'(Z^{[2]}) \cdot A^{[1]T}$$

$$W^{[1]} := W^{[1]} - \alpha \Delta W^{[1]}$$

$$= W^{[1]} - \alpha \frac{\partial E}{\partial W^{[1]}}$$

$$= W^{[1]} + E^{[1]} \cdot g'(Z^{[1]}) \cdot A^{[0]T}$$



다층 신경망 모델링

- 학습 정리
 - 미분에서 연쇄법칙이 무엇인지 이해하기
 - 오차함수의 행렬로 미분하기
 - 다층 인공 신경망의 행렬 모델을 학습하기
- 차시 예고
 - **10-2** 로지스틱 회귀

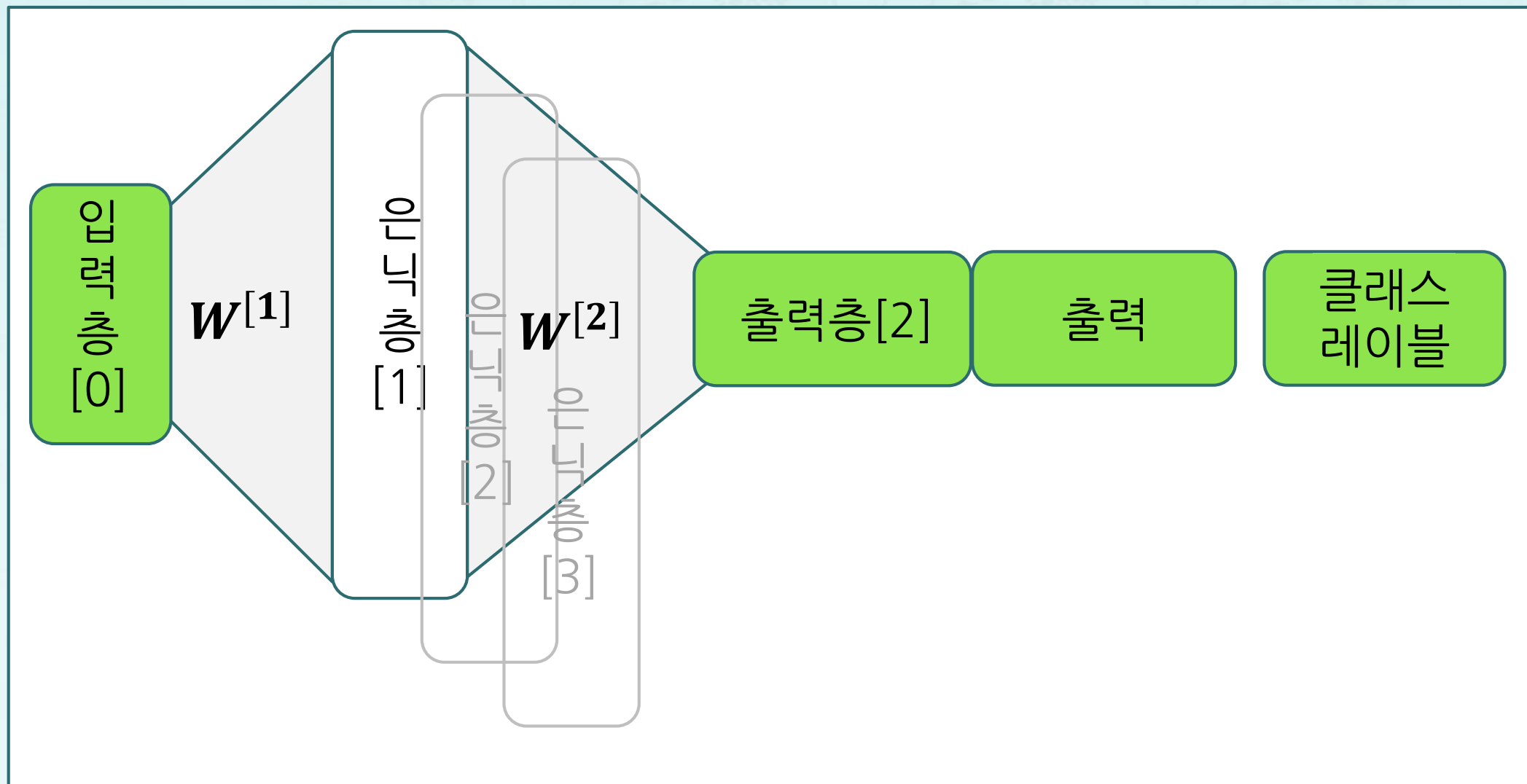
10주차(1/3)

다층 신경망 모델링

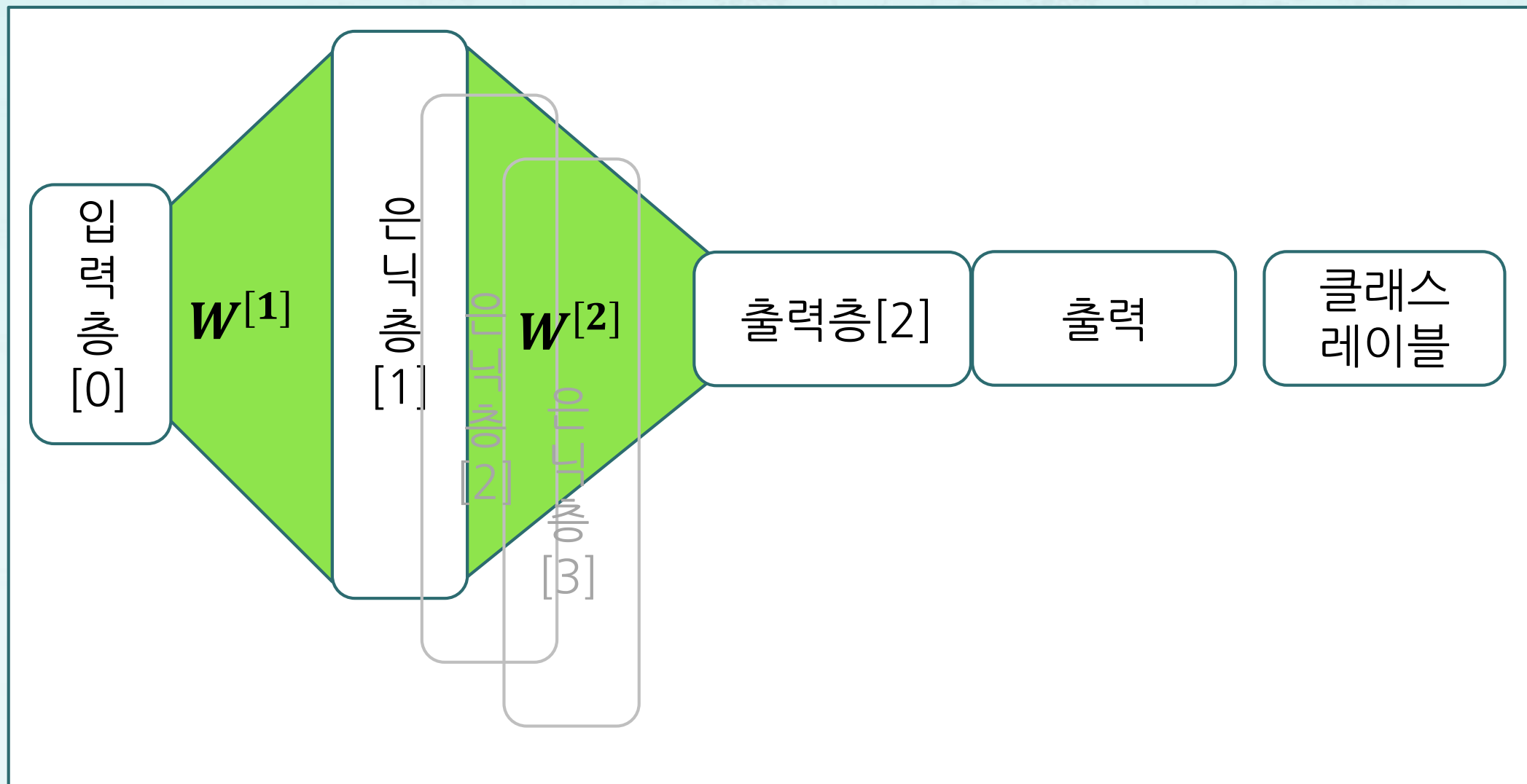
파이썬으로 배우는 기계학습

한동대학교
김영섭 교수

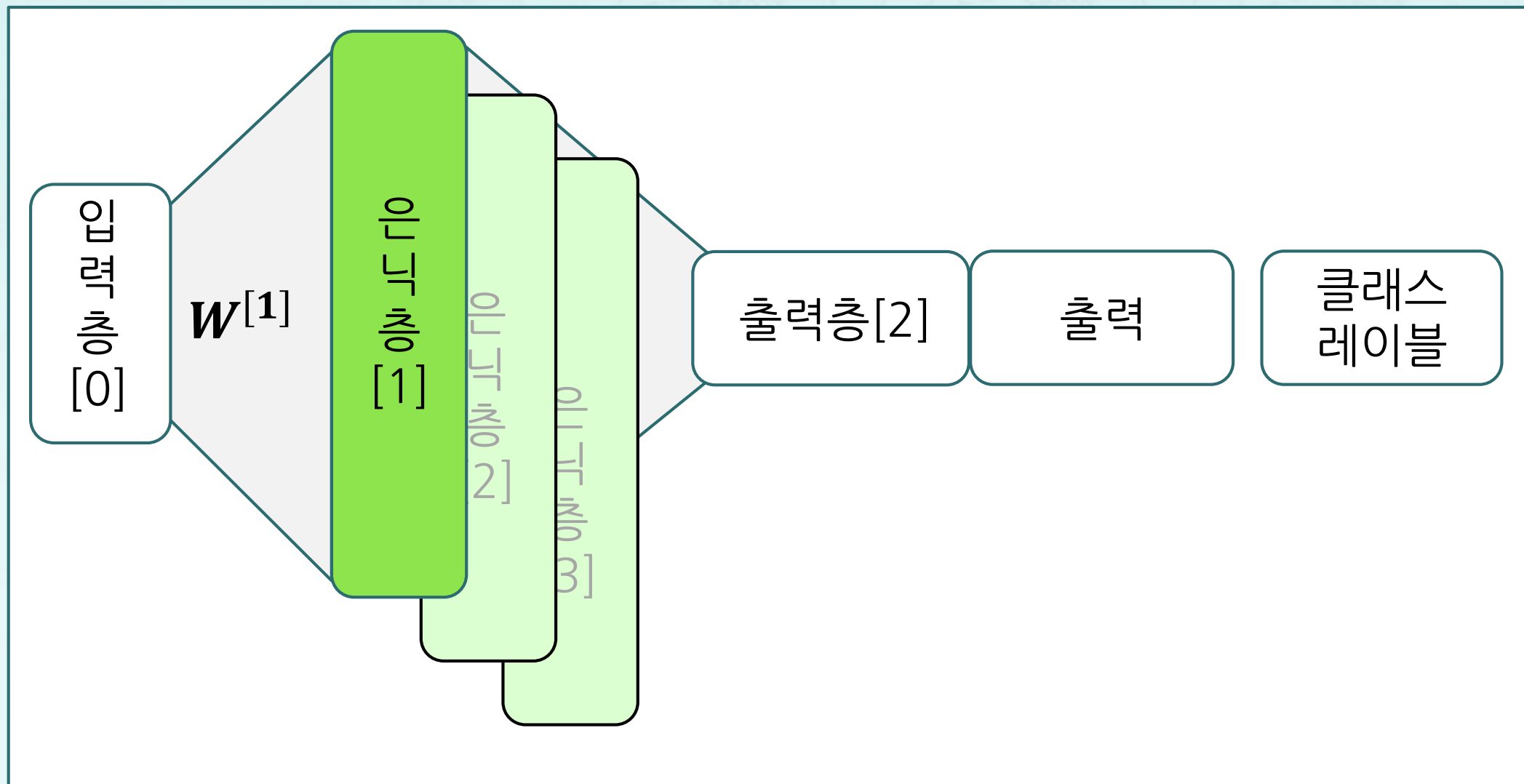
- 다층 신경망의 구조

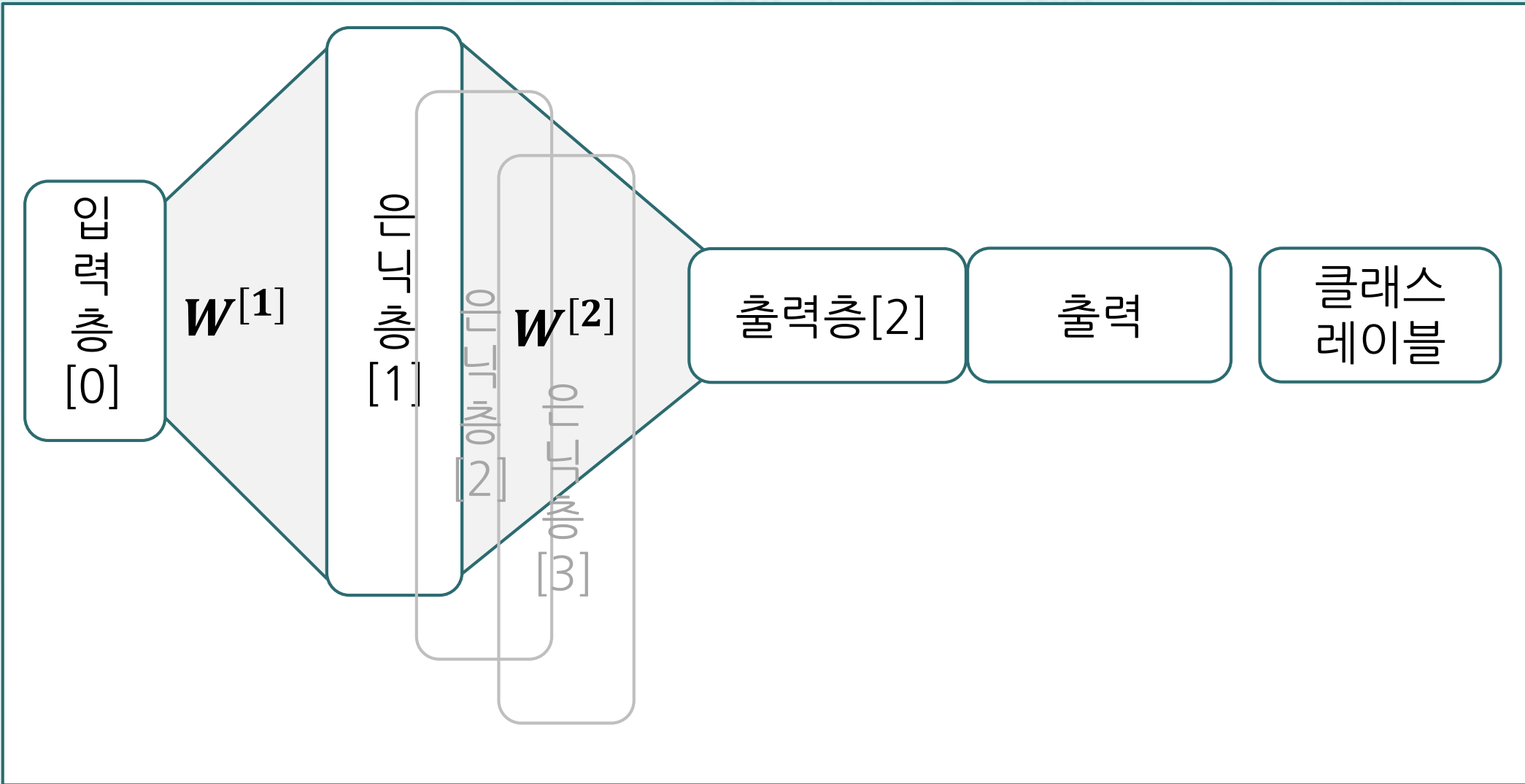


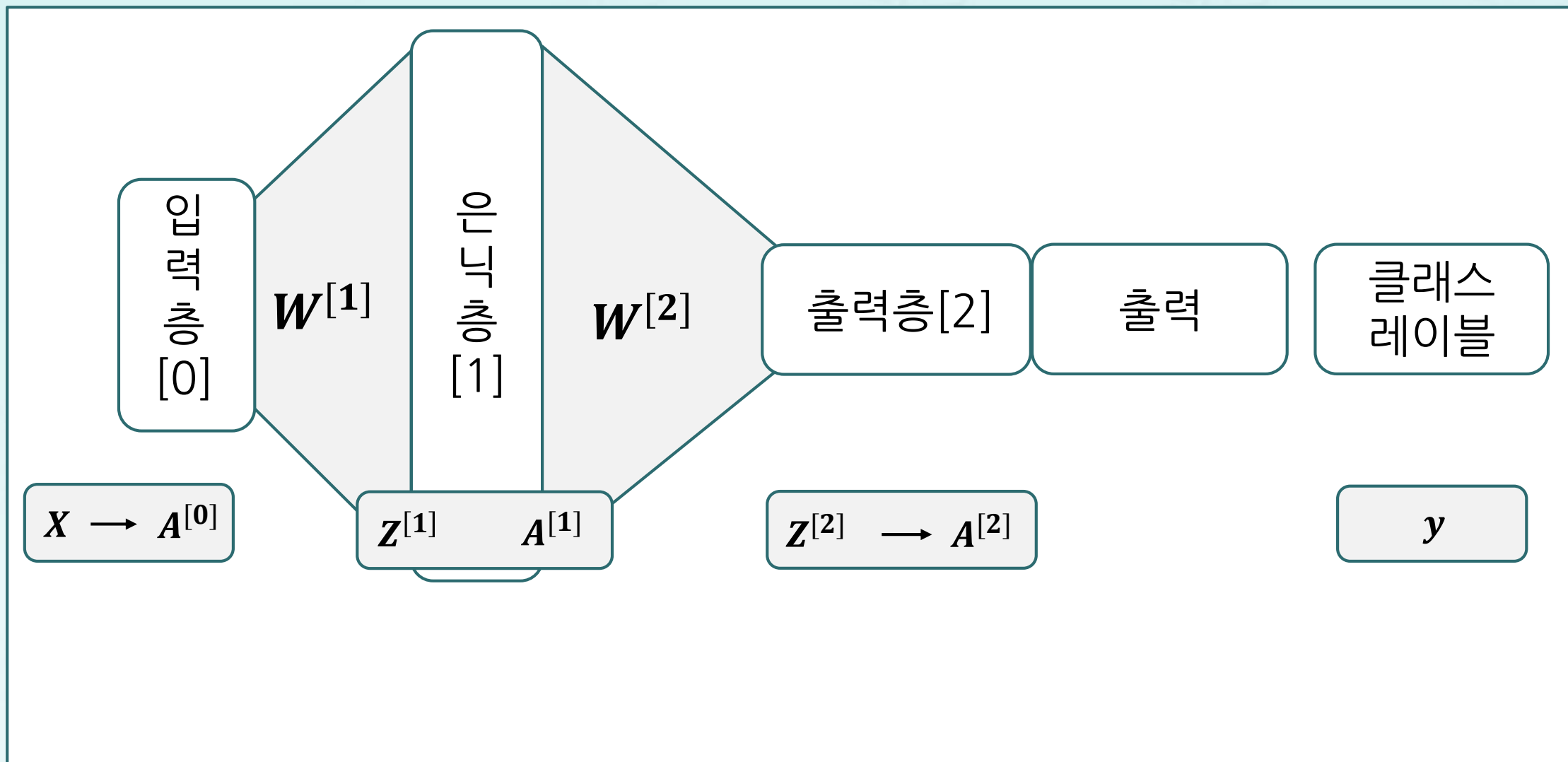
- 다층 신경망의 구조



- 다층 신경망의 구조







■ 역전파 모델링

