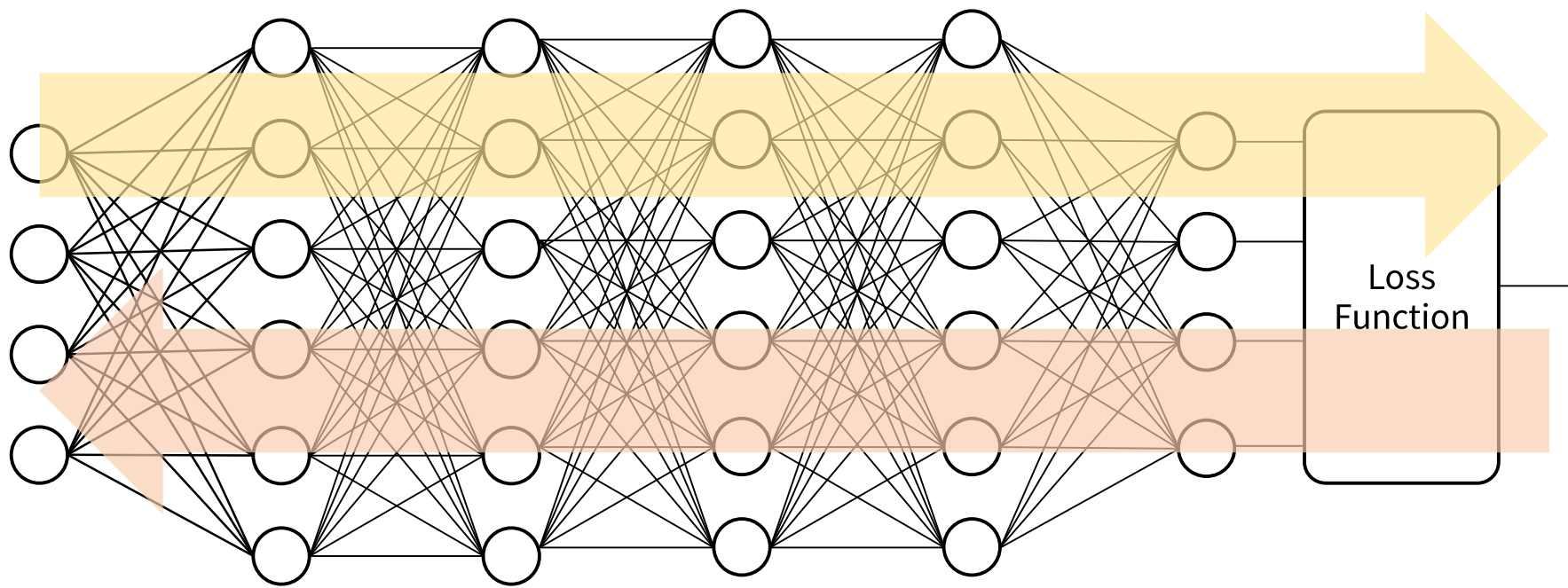


Chapter 04. 쉽게 배우는 역전파 학습법

# STEP2.

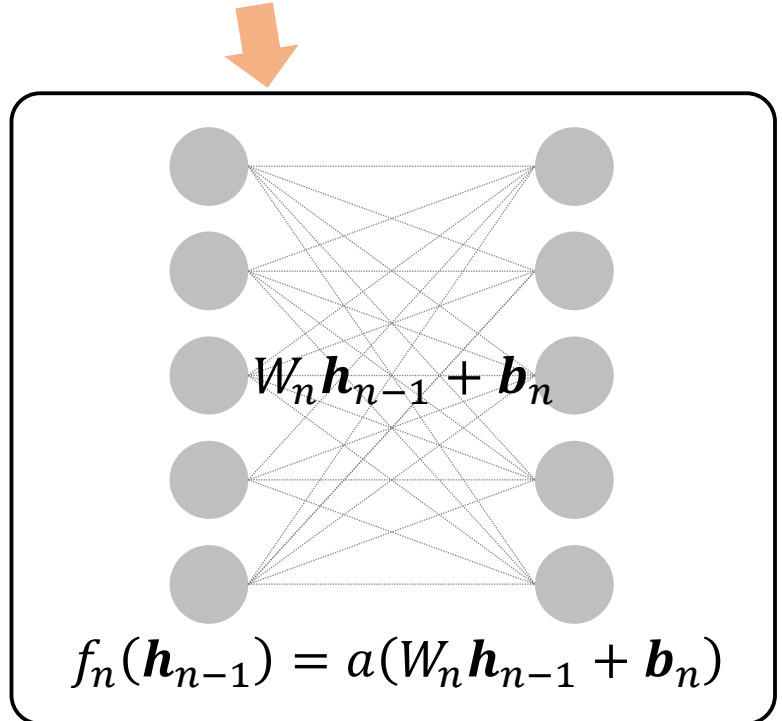
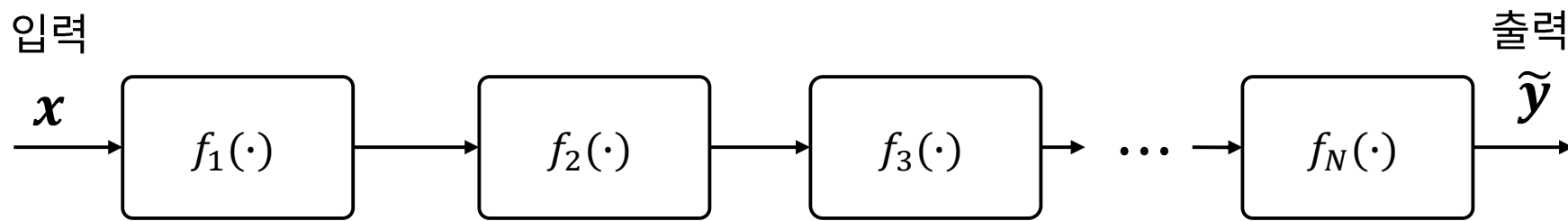
## 역전파 학습법의 수식적 이해

# 역전파 학습법 Back-Propagation



역전파 학습법에서 가장 핵심적인 미분 계산을 수식적으로 자세히 뜯어보고 이해해보자.

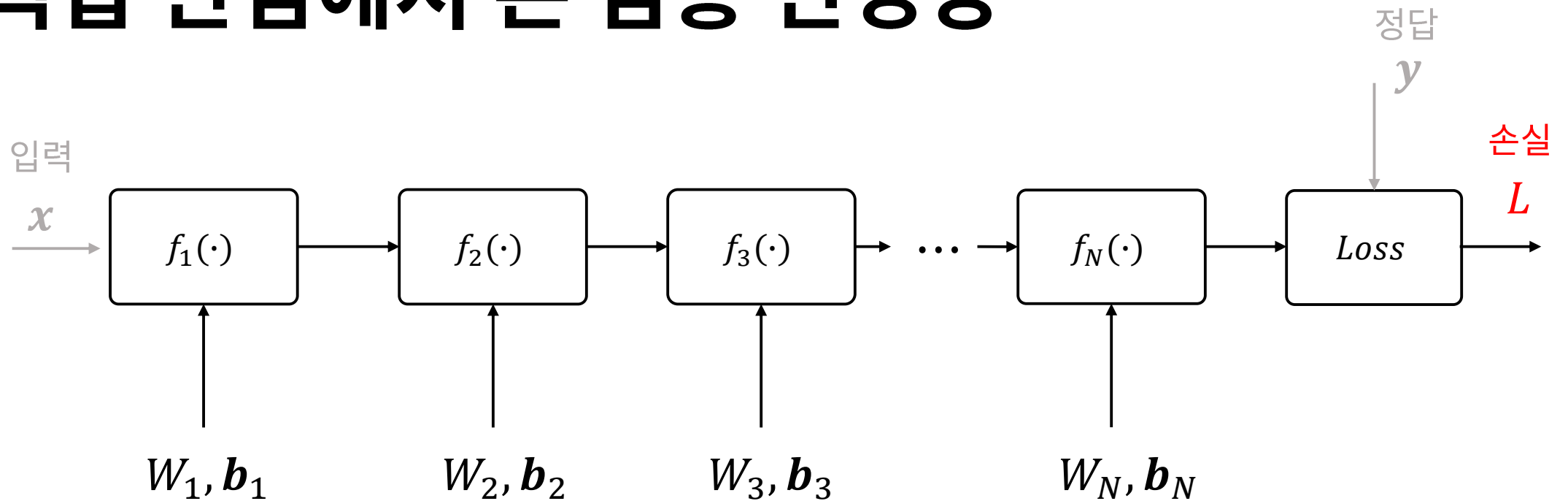
# 합성 함수로서의 심층 신경망



$$\underline{\tilde{y} = f_N(f_{N-1}(\cdots f_2(f_1(x))))}$$

심층 신경망의 각 Layer를 하나의 함수로 본다면, **신경망을 합성 함수로 표현**할 수 있다.

# 학습 관점에서 본 심층 신경망

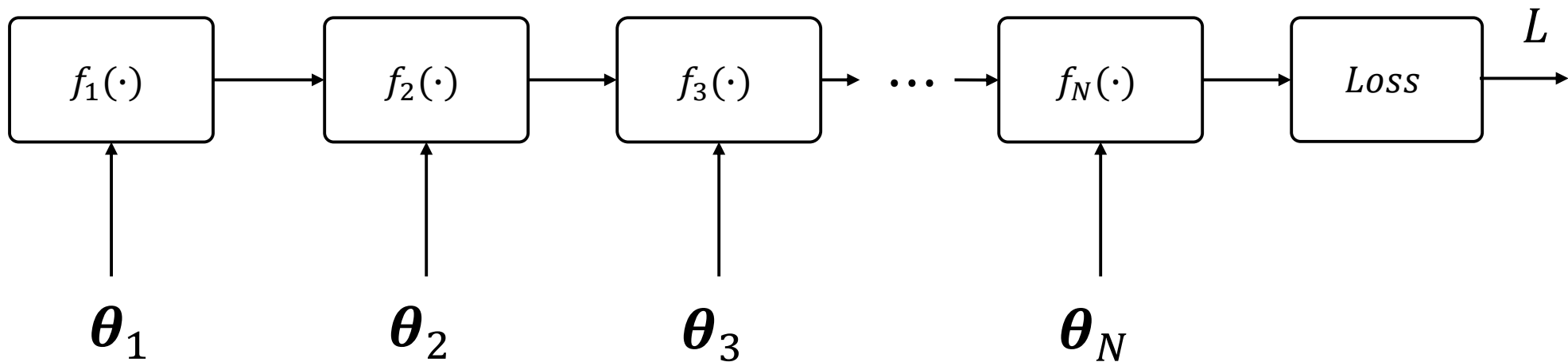


$$f_n(\mathbf{h}_{n-1}; W_n, \mathbf{b}_n) = a(W_n \mathbf{h}_{n-1} + \mathbf{b}_n)$$

이미 손실을 구했다면, 데이터셋의 입력과 출력은 학습 과정에서 중요하지 않다.

**손실을 최소화하는 파라미터만** 찾으면 되기 때문!

# 심층신경망의 연쇄 법칙



$$f_n(\theta_n) = a(W_n \mathbf{h}_{n-1} + \mathbf{b}_n)$$

$$\theta_n = [\text{vec}(W_n)^T | \mathbf{b}_n^T]^T$$

$$\frac{\partial L}{\partial \theta_n} = \frac{\partial L}{\partial \mathbf{h}_N} \prod_{i=n+1}^N \frac{\partial f_i(\mathbf{h}_{i-1})}{\partial \mathbf{h}_{i-1}} \frac{\partial f_n(\theta_n)}{\partial \theta_n}$$

미분하고자 하는 **경로 사이에 있는 모든 미분 값을 곱하면** 원하는 미분을 구할 수 있다.