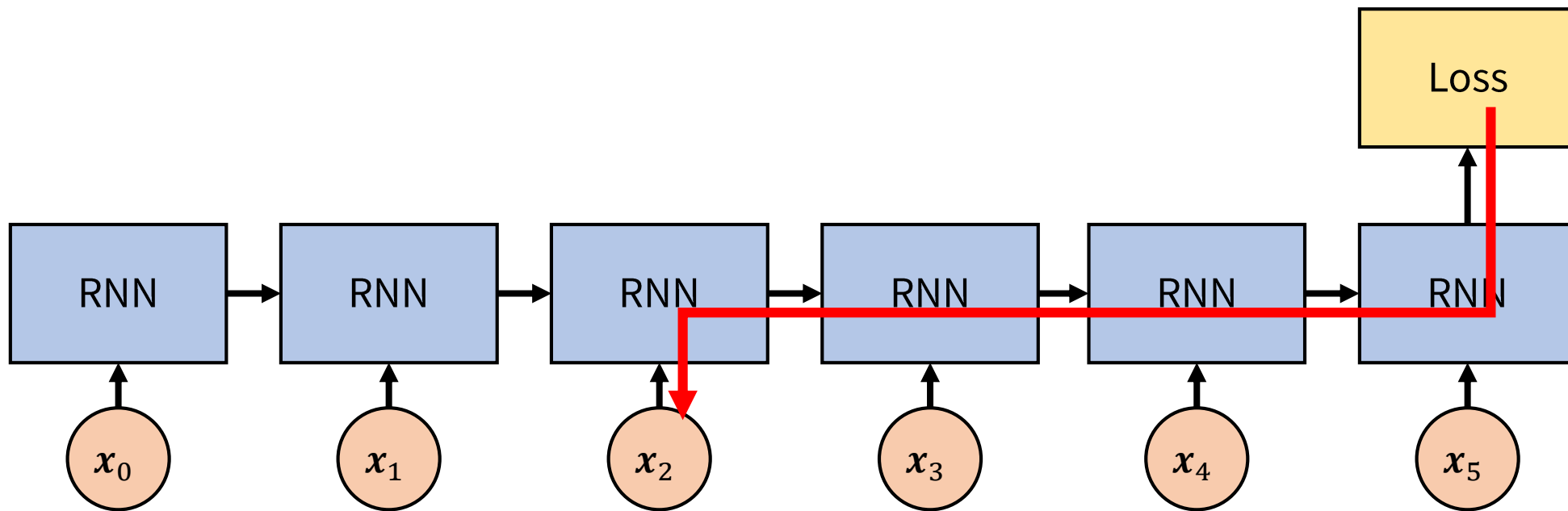


Chapter 06. 순환 신경망(RNN)

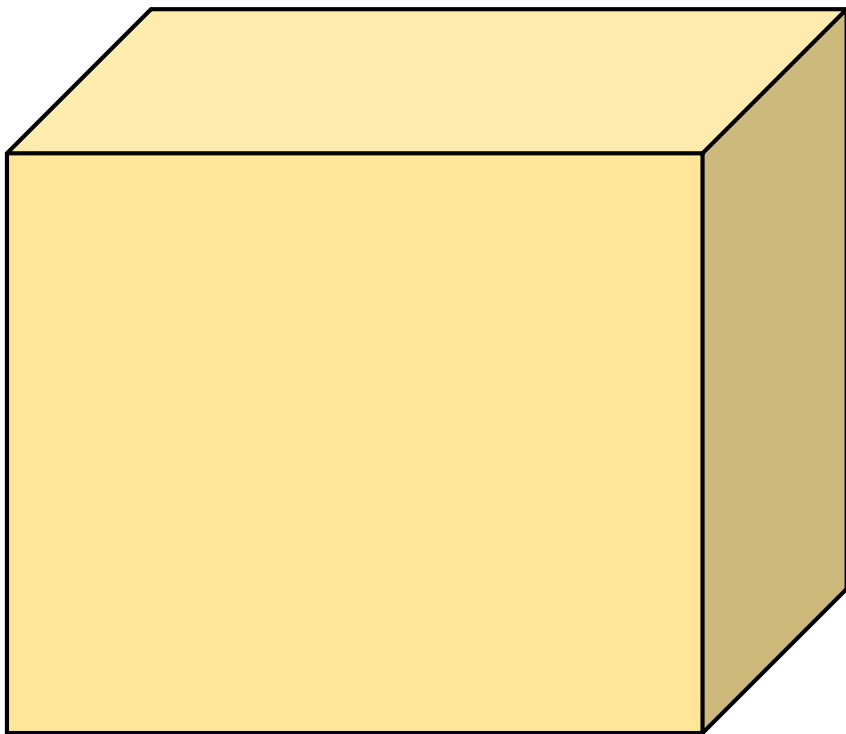
# STEP2. 순환 신경망의 학습법

# BPTT

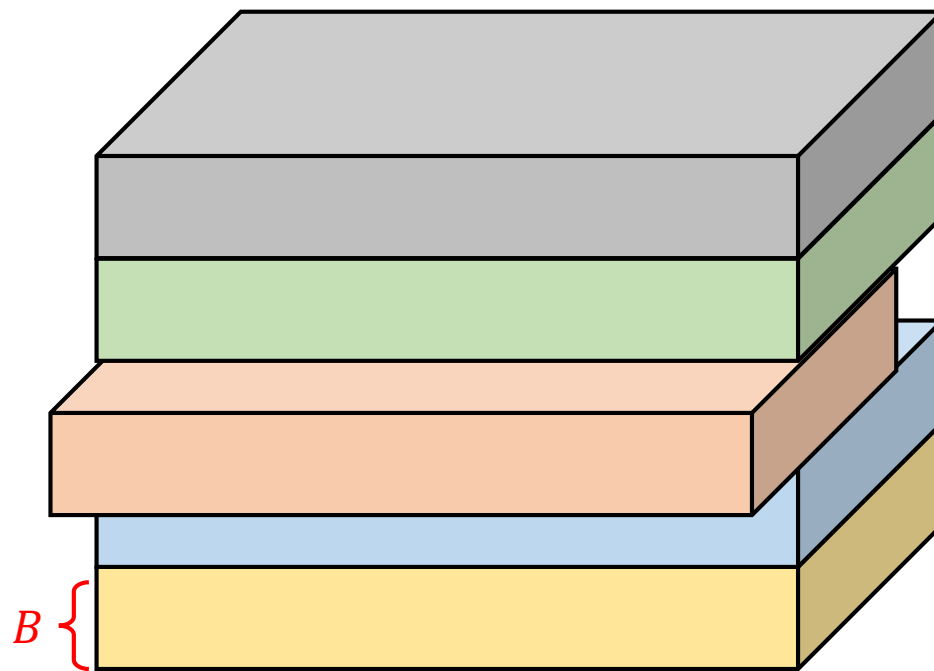


시간펼침 역전파(Backpropagation through time; BPTT)를 개념적으로 배웠다.  
이것을 실제로 학습에 사용하기 위해 어떻게 해야 하는지 자세히 알아보자.

# BPTT 데이터 입력

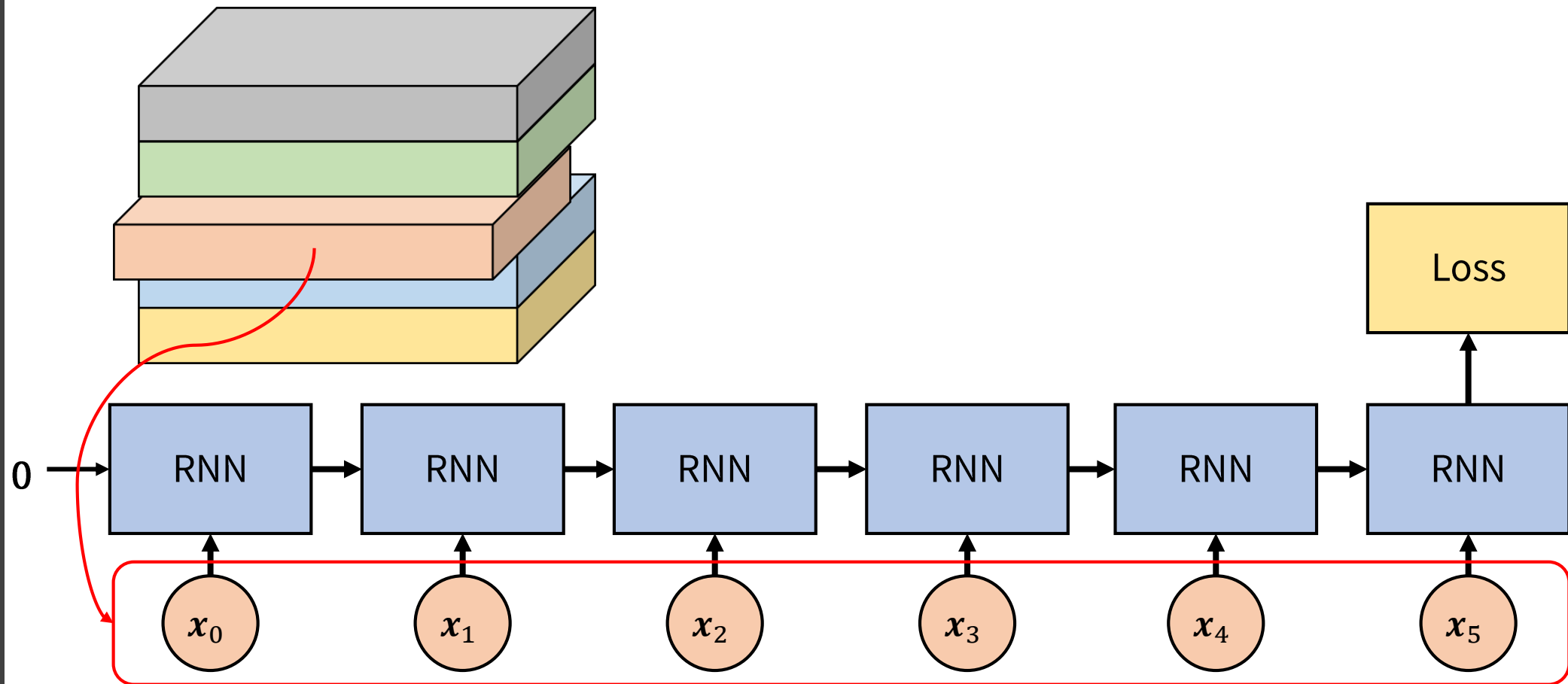


학습 데이터 입력 ( $N \times L \times I$ )



배치로 나눈 학습 데이터 입력( $B \times L \times I$ )

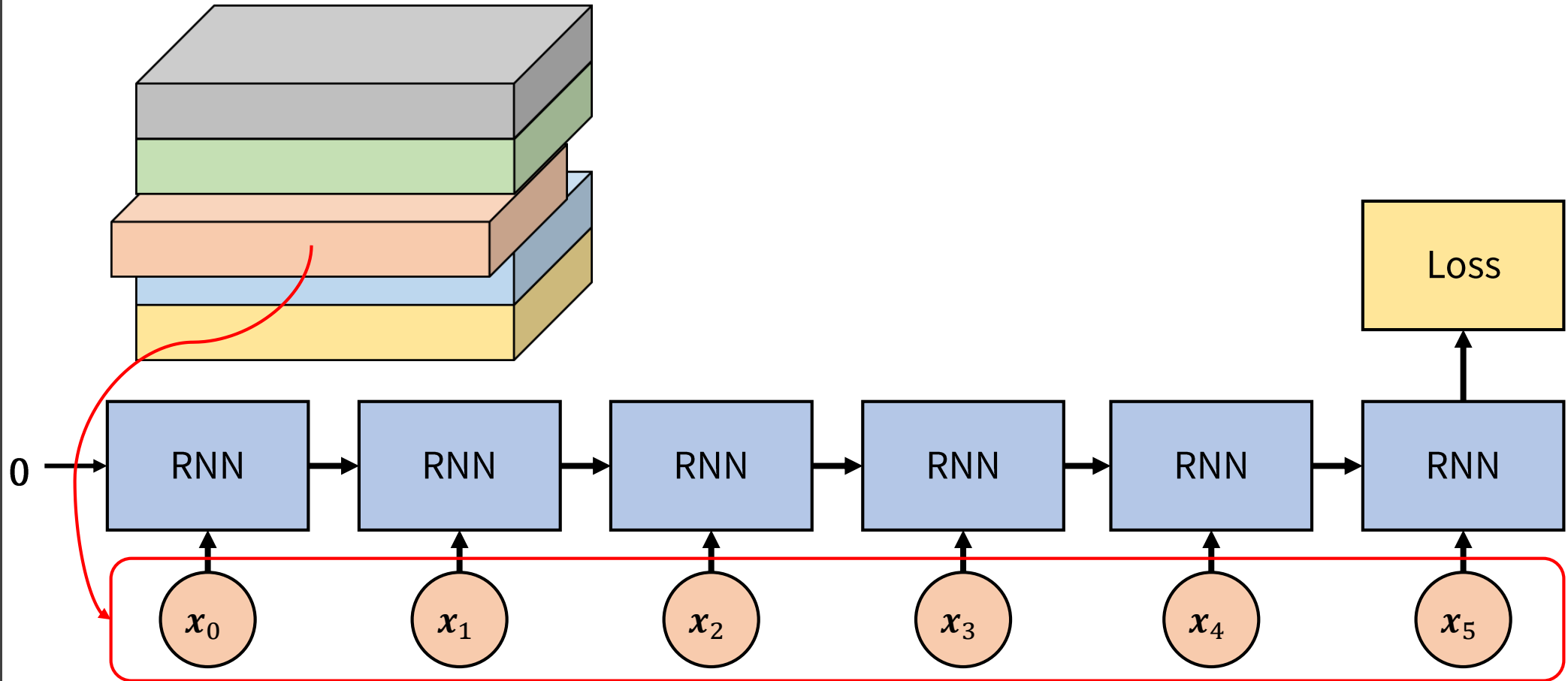
# BPTT의 배치 학습법



시간적으로 펼치면서 계산해야 하는 점을 제외하면, 보통의 역전파와 동일하다.

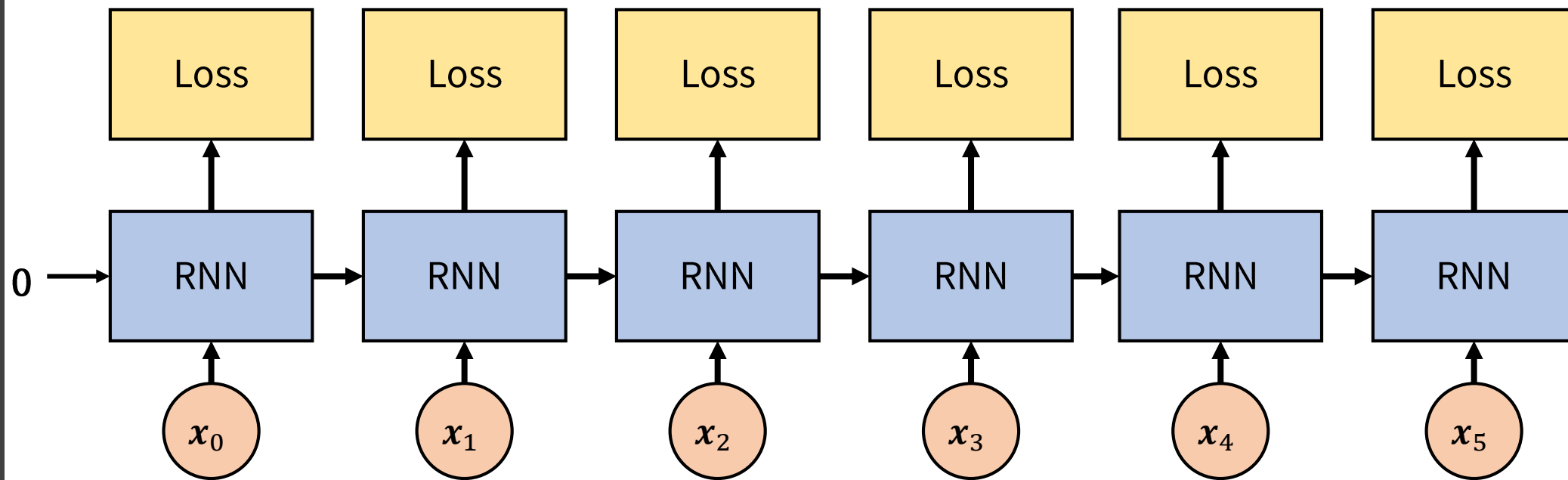
단, 시간적으로 펼칠 때 **역전파를 위한 추가적인 메모리**가 필요하다.

# BPTT의 문제점



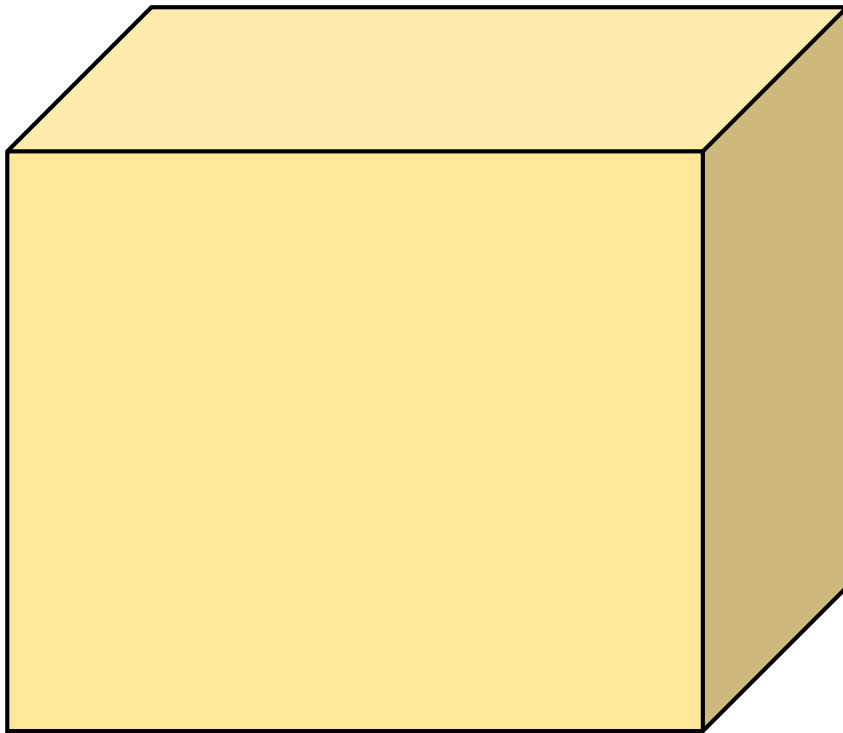
순차 데이터의 길이  $L$ 이 매우 클 경우, 시간 펼침이 늘어나면서 **필요 메모리가  $L$ 배 증가**한다.  
 $B$ 개의 샘플을 동시에 계산하므로, 얇은 신경망에 비해 훨씬 큰 메모리 필요.

# 다중 입력, 다중 출력

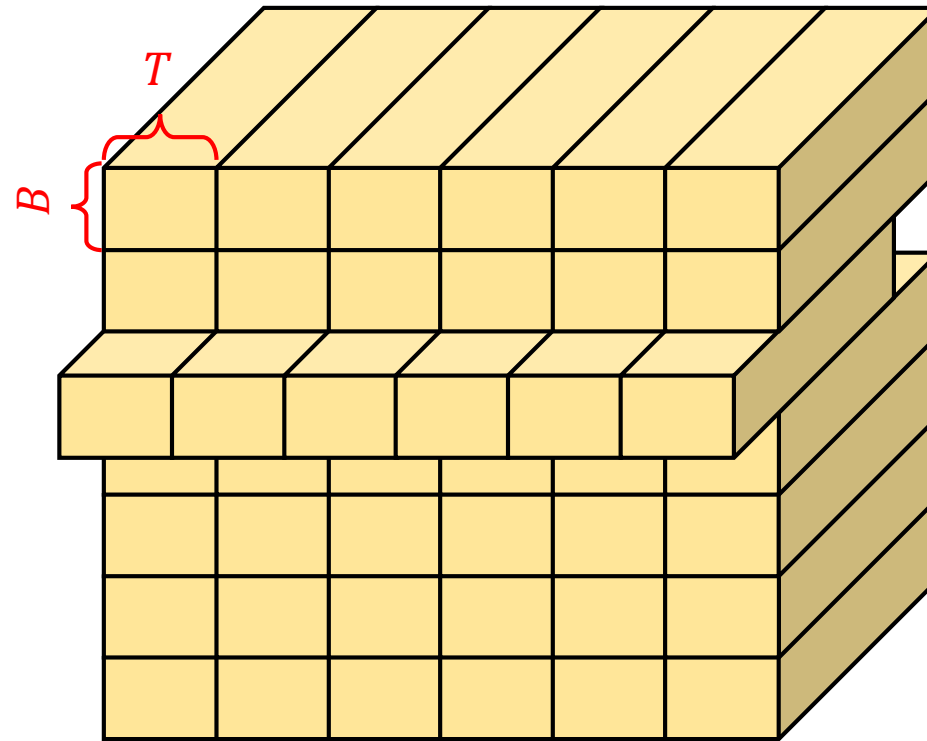


다중 입력, 다중 출력의 경우 Truncated BPTT를 이용해 BPTT에서 발생하는 메모리 문제를 해결할 수 있다.

# Truncated BPTT 데이터 입력



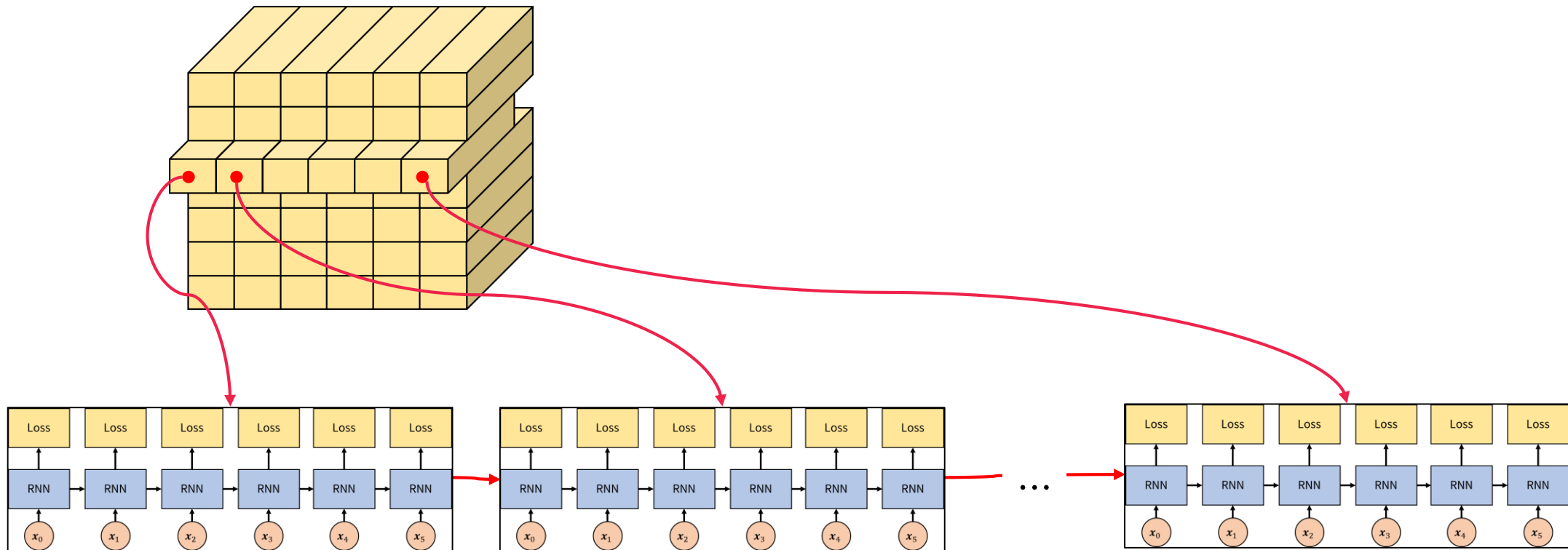
학습 데이터 입력 ( $N \times L \times I$ )



Truncated BPTT 학습 데이터 입력 ( $B \times T \times I$ )

순차 데이터의 길이를 일정한  $T$  길이로 잘라서 배치를 나누듯이 한번에 계산하는 크기를 줄인다.

# Truncated BPTT



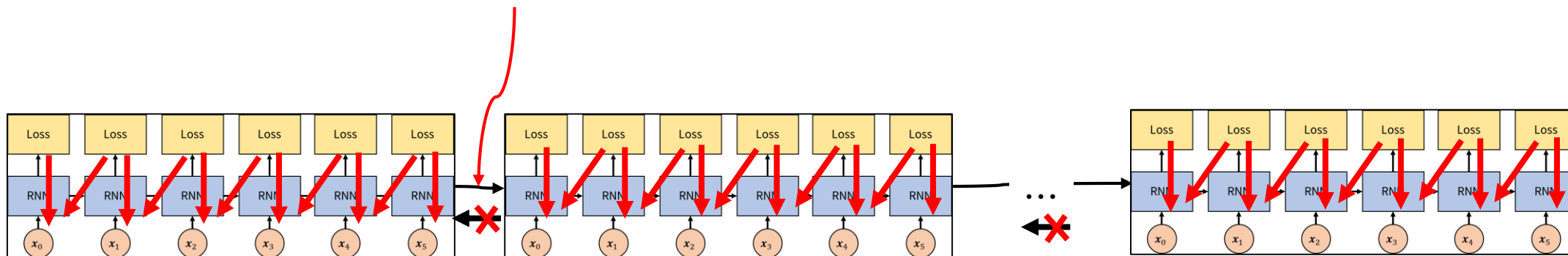
Truncated BPTT는 길이  $L$ 의 입력을 길이  $T$ 로 쪼개어 순서대로 학습한다.

한번에 역전파 하는 길이가 제한되므로 메모리 사용이 줄어든다.



# Truncated BPTT의역전파 흐름

Hidden state, Cell state 전달



길이  $T$ 로 쪼개진 Truncation 사이에서는 기울기 역전파가 이루어지지 않는다.

즉, Time step이  $T$  이상 떨어진 입-출력 관계는 학습되지 않는다.