

# 순환 신경망

Recurrent neural network, RNN

# 목차

- I. 시퀀스(sequence) 데이터
- II. 시퀀스 모델링을 위한 RNN

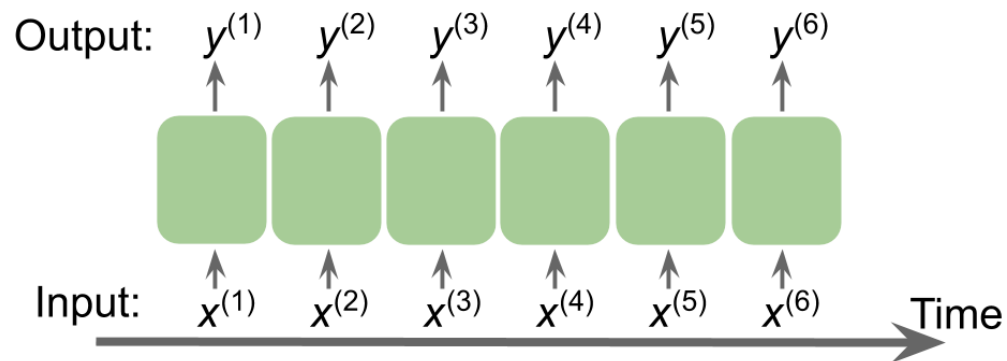
# I. 시퀀스 sequence

- 시퀀스란 순서에 따라 정렬된 데이터이다.
- (I, am, eating), (am, I, eating) : 같은 집합, 다른 시퀀스
- 데이터의 나열 순서가 데이터의 의미나 해석에 영향을 미침
  - RNN은 순서 정보도 반영하는 시퀀스 모델링을 위해 고안됨. 학습 과정에서 과거 정보가 현재 출력에 영향을 줌(순서가 독립적이지 않음)
  - GARCH 모형도 시퀀스 모델, 전날의 변동성이 오늘의 변동성에 영향을 줌.
  - 시계열 데이터는 순서(index)가 시간인 시퀀스이다. time series  $\subset$  sequence

# I. 시퀀스 sequence

## ▪ 시퀀스 표현

- 이 장에서는 시퀀스를  $\langle \mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^T \rangle$  로 표현한다.
- 위 첨자는 샘플의 순서, T는 시퀀스 길이
- 각 샘플 포인트  $\mathbf{x}^t$ (입력특성)는 특정시간 t에 속한다.

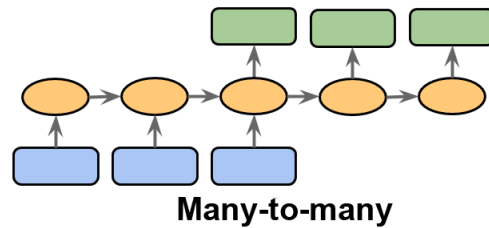
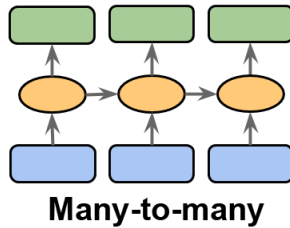
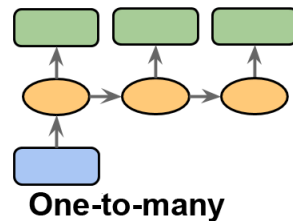
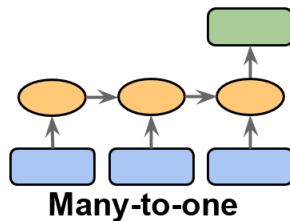


$$X = \begin{bmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \\ \mathbf{x}^{(3)} \\ \mathbf{x}^{(4)} \\ \mathbf{x}^{(5)} \end{bmatrix} = \begin{bmatrix} 22.1 & 45.0 \\ 22.5 & 46.2 \\ 23.0 & 47.8 \\ 22.8 & 48.0 \\ 22.3 & 46.5 \end{bmatrix}$$

# I. 시퀀스 sequence

## ▪ 시퀀스 모델링 종류

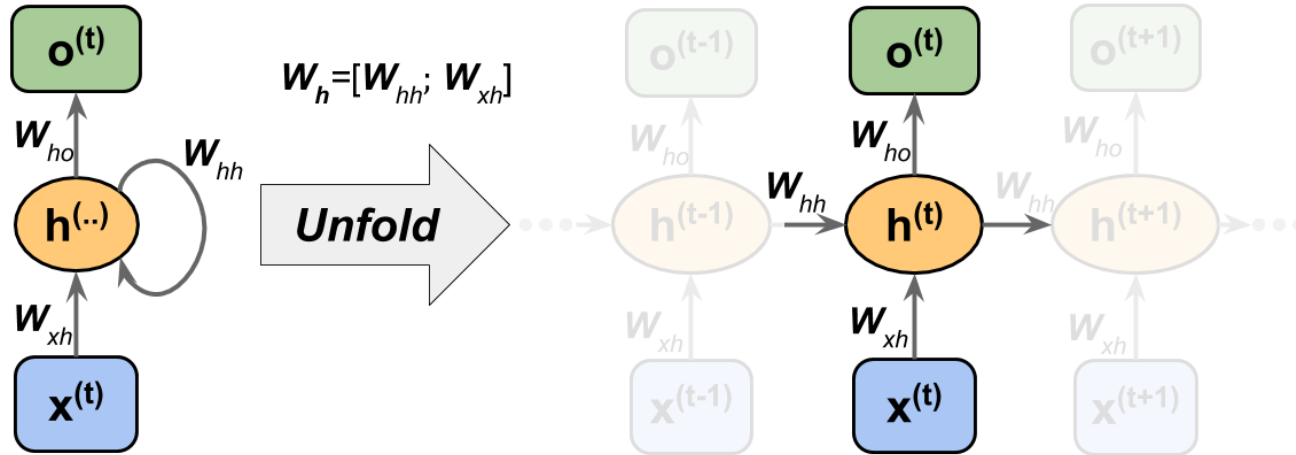
- 가장 널리 사용하는 시퀀스 모델링 작업의 종류 - 3 가지



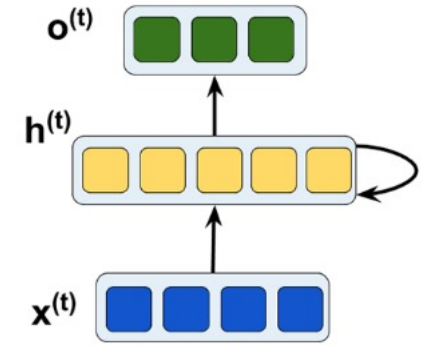
- 다대일 **Many to one** : (입력 = 시퀀스, 출력 = 고정크기 벡터(스칼라))  
예) 영화 리뷰 → 긍정 or 부정
- 일대다 **One to many** : (입력 = 일반적, 출력 = 시퀀스)  
예) 이미지 → 이미지 내용 요약 문장  
\* Vision Transformer (ViT)는 이미지를 패치 단위로 나눠 시퀀스처럼 처리
- 다대다 **Many to many** : (입력 = 시퀀스, 출력 = 시퀀스) 동기과 지연으로 나뉨  
예) 각 프레임을 레이블링, 번역(문장을 먼저 다 읽고 해석)

## II. 시퀀스 모델링을 위한 RNN

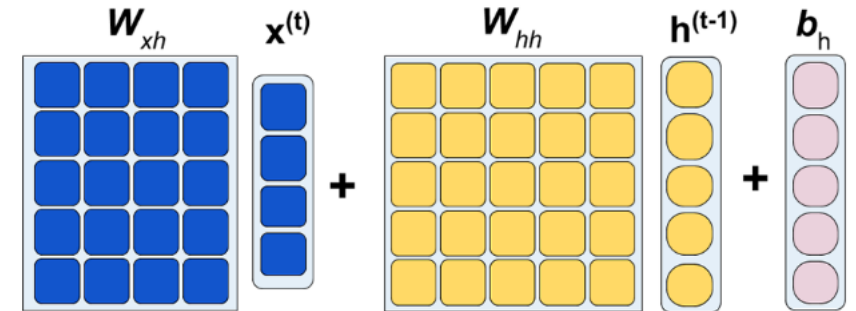
### ▪ 단일 층 RNN의 구조



$W_{ho}$   
 $W_{hh}$   
 $W_{xh}$  이 세 가중치 행렬은 특정시간  $t$  에 종속되지 않고 동일함

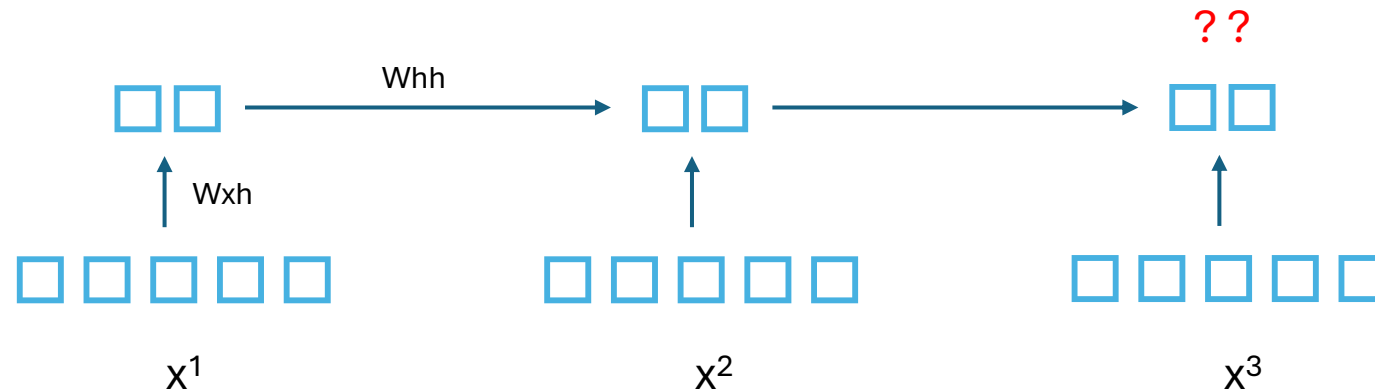


Formulation 1:  $h^{(t)} = \sigma_h( W_{xh}x^{(t)} + W_{hh}h^{(t-1)} + b_h )$



## II. 시퀀스 모델링을 위한 RNN

- **예제**\_입력특성이 5차원, 시퀀스 길이가 3인 단일 층 RNN의 출력



예제코드

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{tensor}([ [1., 1., 1., 1., 1.], [2., 2., 2., 2., 2.], [3., 3., 3., 3., 3.] ])$$

예 : 하나의 주식에 대해 5가지  
입력특성을 3일간 관찰한 시퀀스