

Part. 08

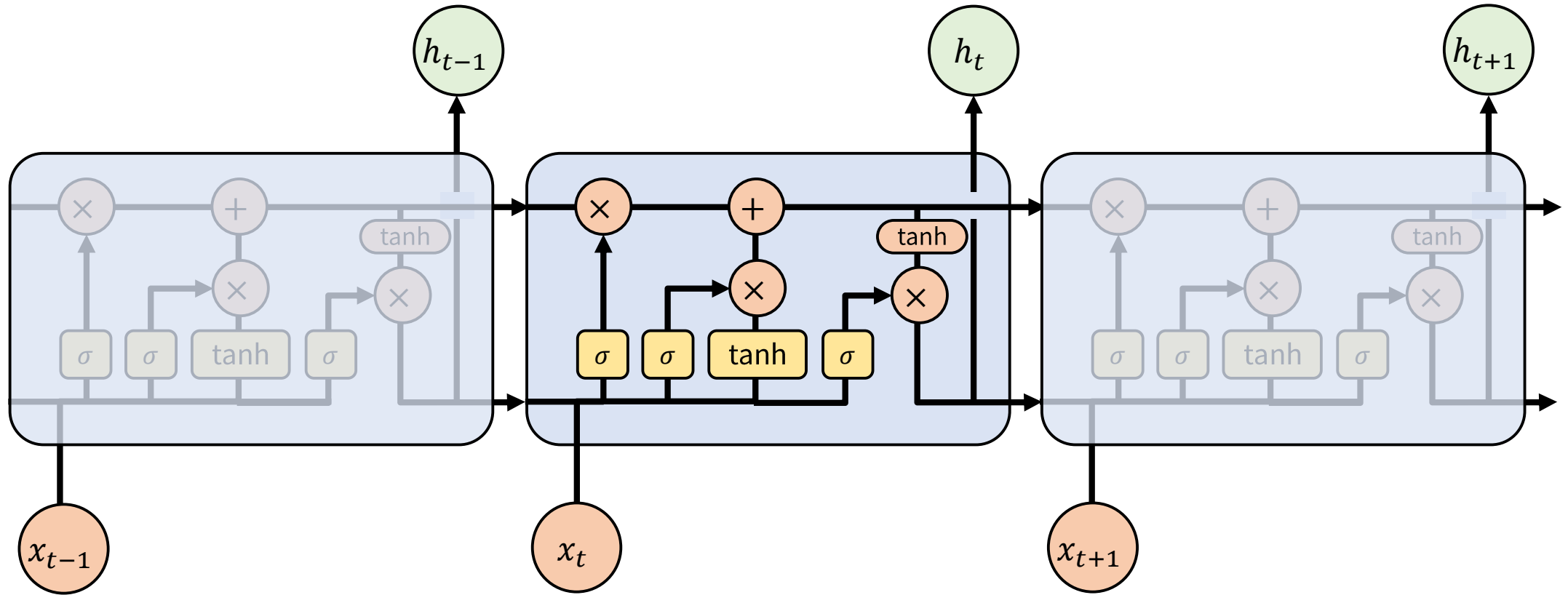
Recurrent Neural Network

# | Gated Recurrent Unit

FASTCAMPUS  
ONLINE

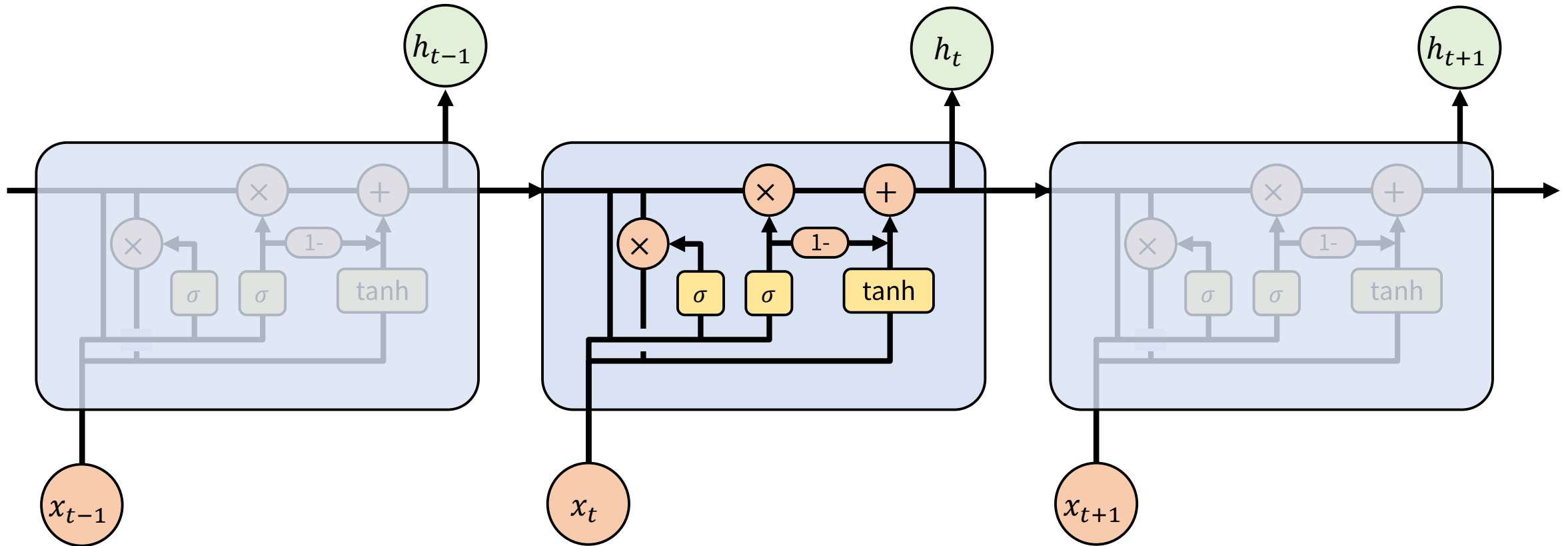
강사. 신제용

# I LSTM (Long Short-Term Memory)



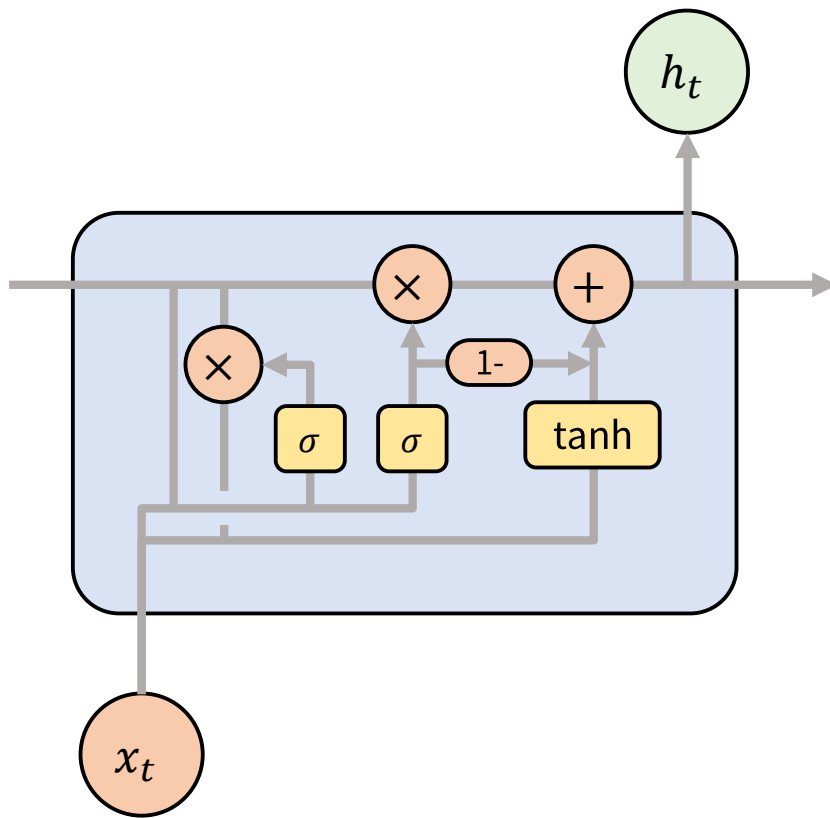
LSTM은 전반적으로 잘 구성되어 있으나, 학습 매개변수가 많고 학습 노하우가 중요한 편이다.

# I GRU (Gated Recurrent Unit)



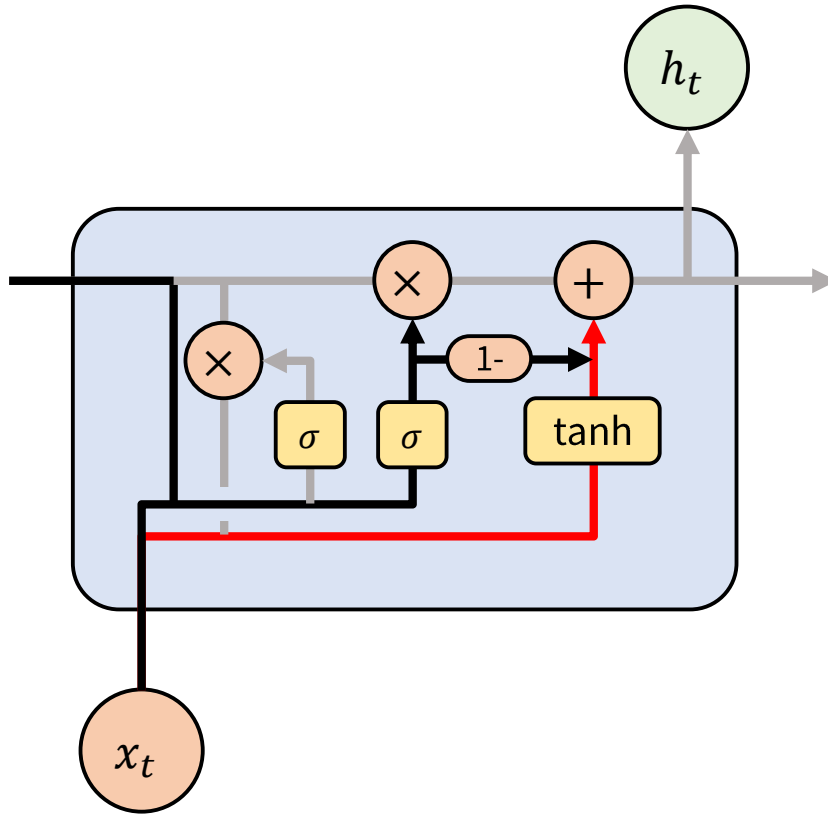
LSTM을 간소화한 버전이라고 할 수 있는 GRU. 마찬가지로 차근차근 보면 쉽게 의미를 알 수 있다.

# I GRU의 구조



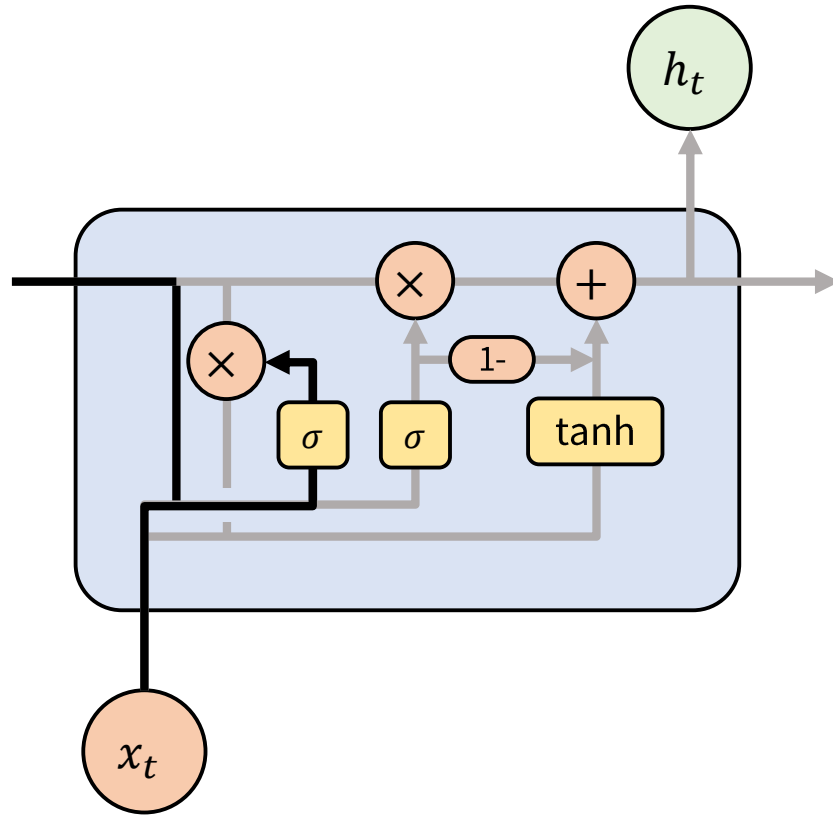
- Cell State가 없고, Hidden State만 존재한다.
- Forget Gate와 Input Gate를 결합하였다.
- Reset Gate를 추가하였다.

# I Forget Gate & Input Gate



- LSTM과 동일한 Forget Gate를 사용한다.
- Forget Gate를 1에서 빼서 Input Gate로 사용하였다.

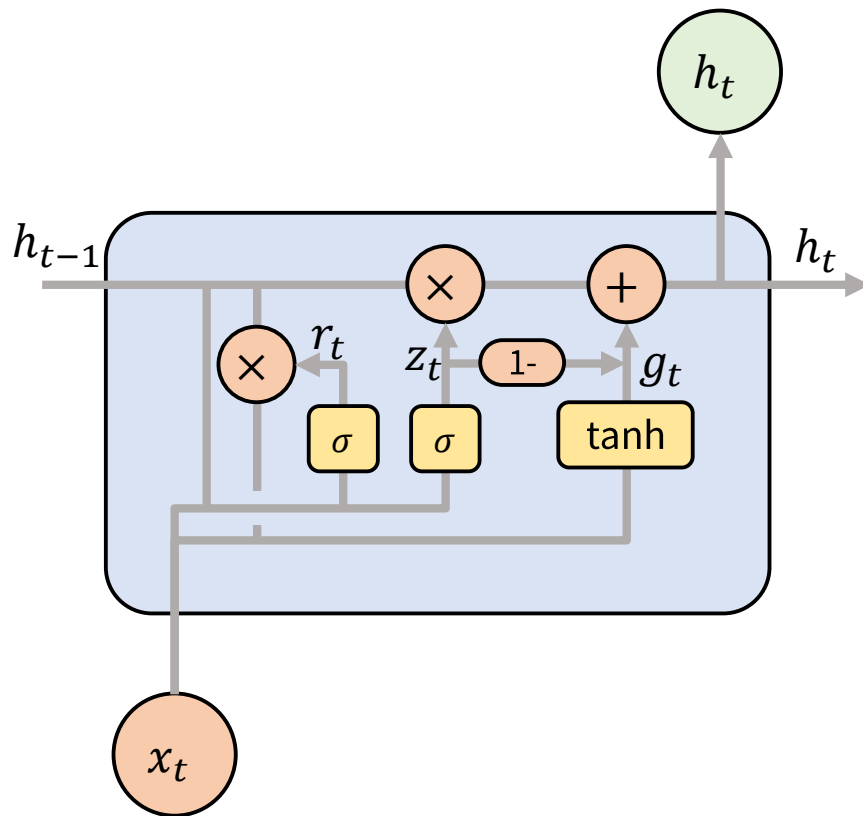
# I Reset Gate



## Reset Gate

- Sigmoid 활성화 함수로, 0~1의 값을 가진다.
- 이전 Hidden state를 얼마나 사용할지 정하는 역할
- 0에 가까운 값이 되면 'Reset'이 된다. (ex. 새 문장의 시작)

# I GRU 수식



$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

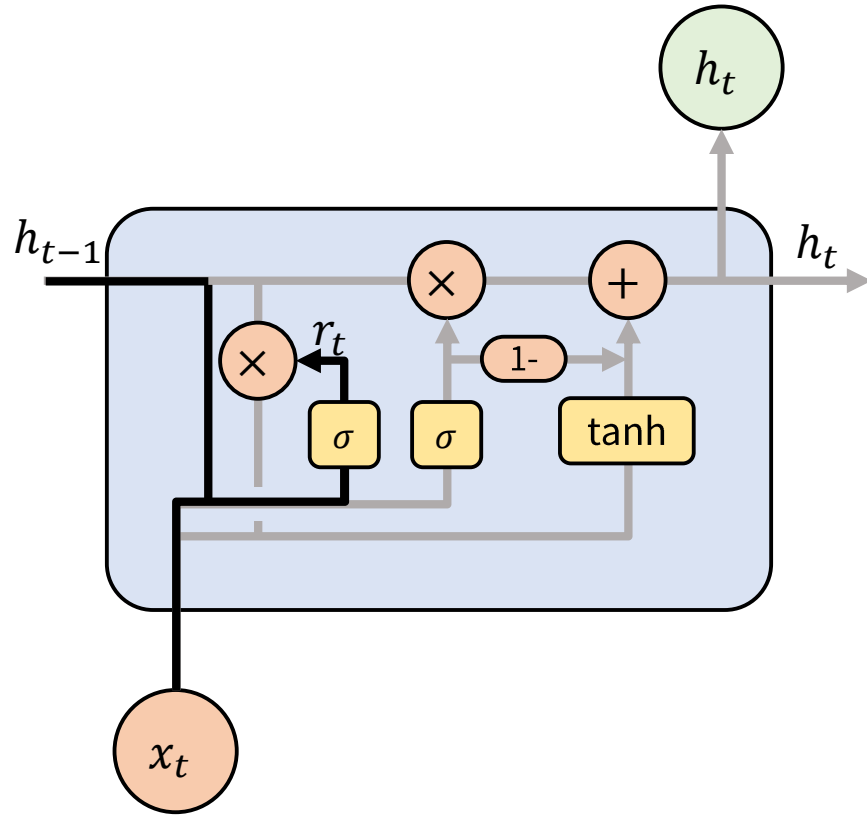
$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

$$g_t = \tanh(W_{xg}x_t + W_{hg}(r_t \odot h_{t-1}) + b_g)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t)g_t$$

GRU는 수식도 한번 도전해 보자. 그리 어렵지 않다!

# I Reset gate



$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

$$g_t = \tanh(W_{xg}x_t + W_{hg}(r_t \odot h_{t-1}) + b_g)$$

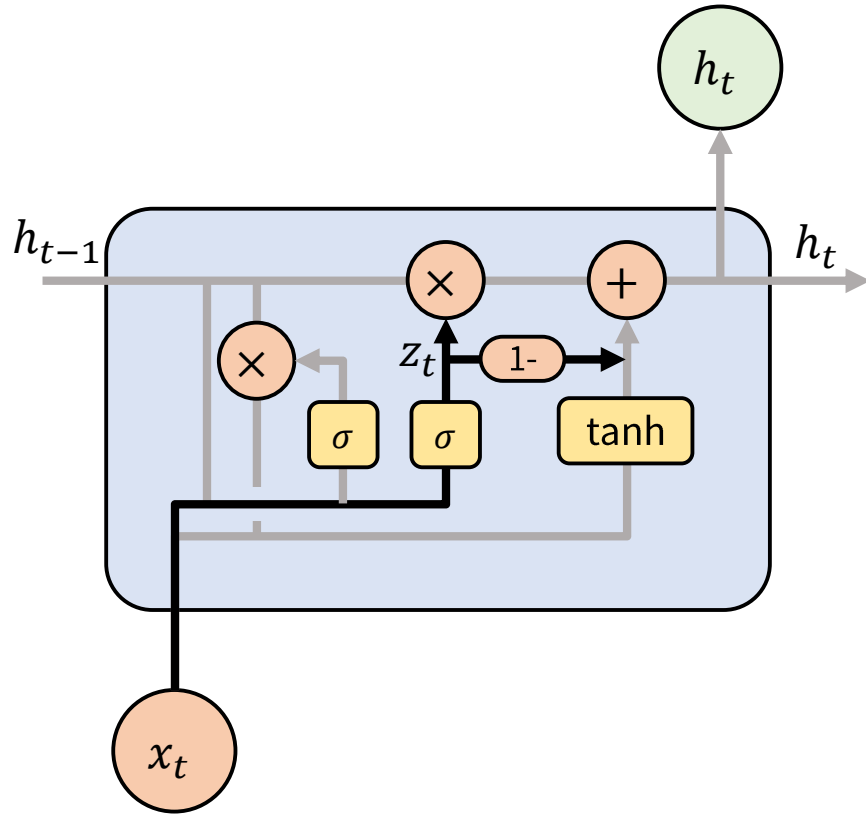
$$h_t = z_t \odot h_{t-1} + (1 - z_t)g_t$$

Reset gate는 Hidden state 중 어떤 특징을 reset할지 결정한다.

Reset 된 특징은 **현재 time step부터 Fully-connected layer 입력에서 제외된다.**



# I Forget gate



$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

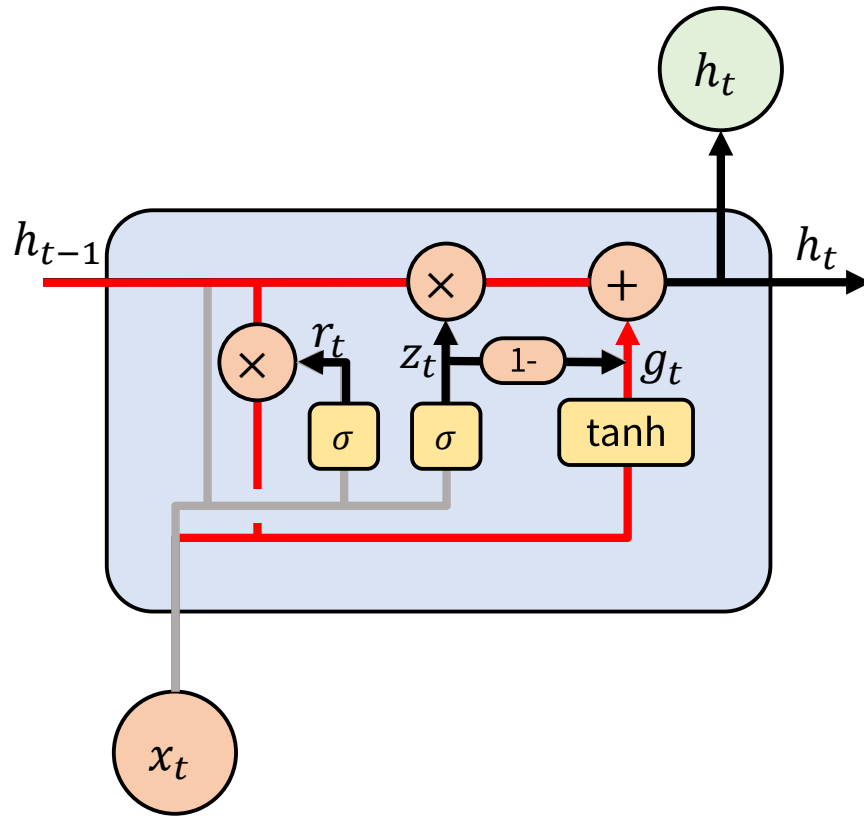
$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

$$g_t = \tanh(W_{xg}x_t + W_{hg}(r_t \odot h_{t-1}) + b_g)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t)g_t$$

Forget gate는 LSTM의 Forget gate와 Output gate를 겸한다.

# I Hidden state



$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

$$g_t = \tanh(W_{xg}x_t + W_{hg}(r_t \odot h_{t-1}) + b_g)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t)g_t$$

Reset gate, Forget gate를 모두 적용하여 Hidden state를 계산한다.  
LSTM의 Cell state와 Hidden state 역할을 모두 겸하고 있다.

# I GRU Overview

