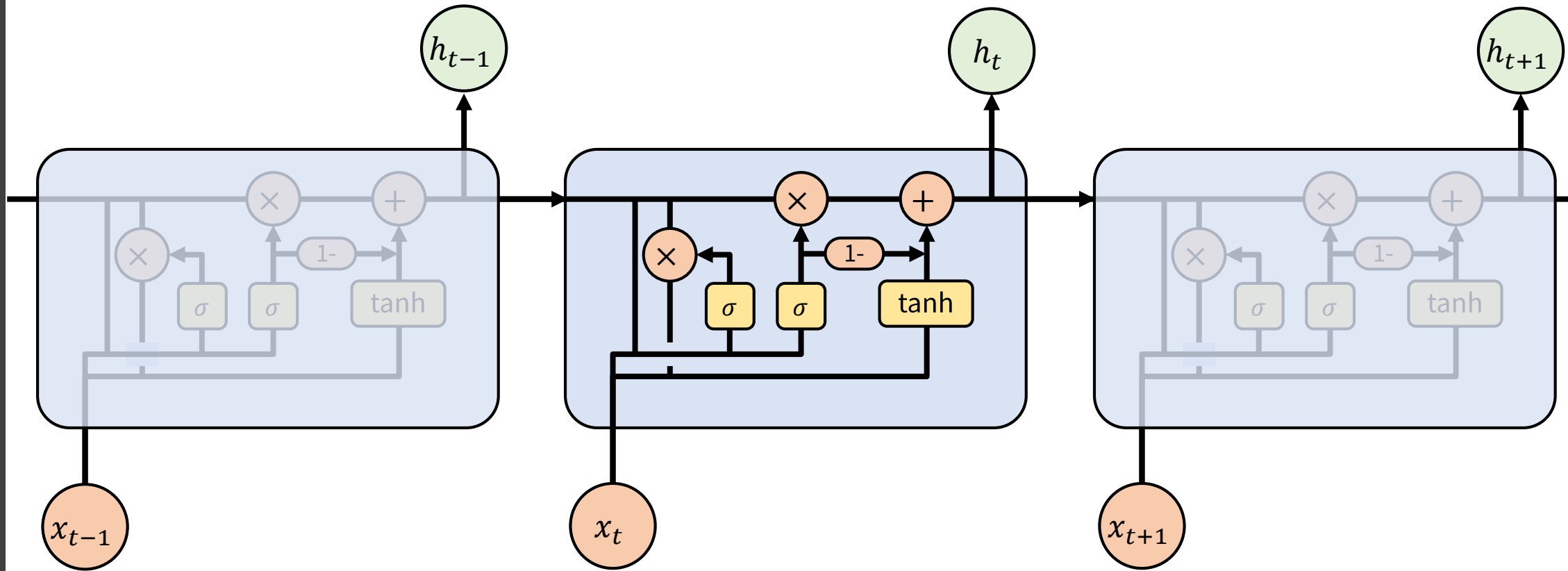
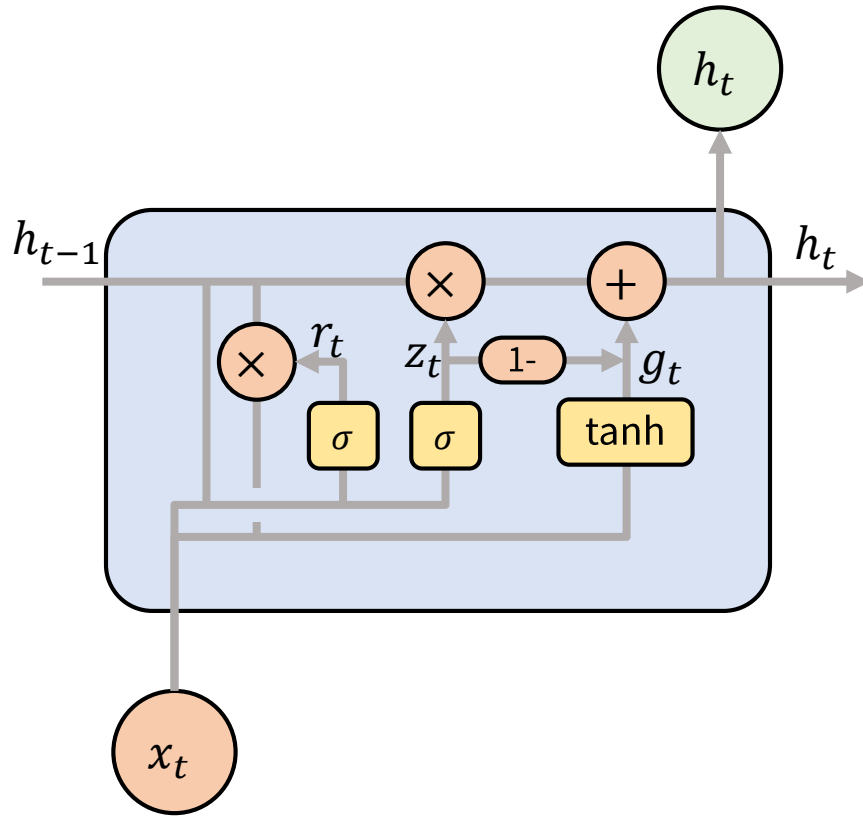


# GRU



LSTM보다 조금 더 간단한 GRU도 마찬가지로 수식을 통해서 살펴보자.

# GRU 수식



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

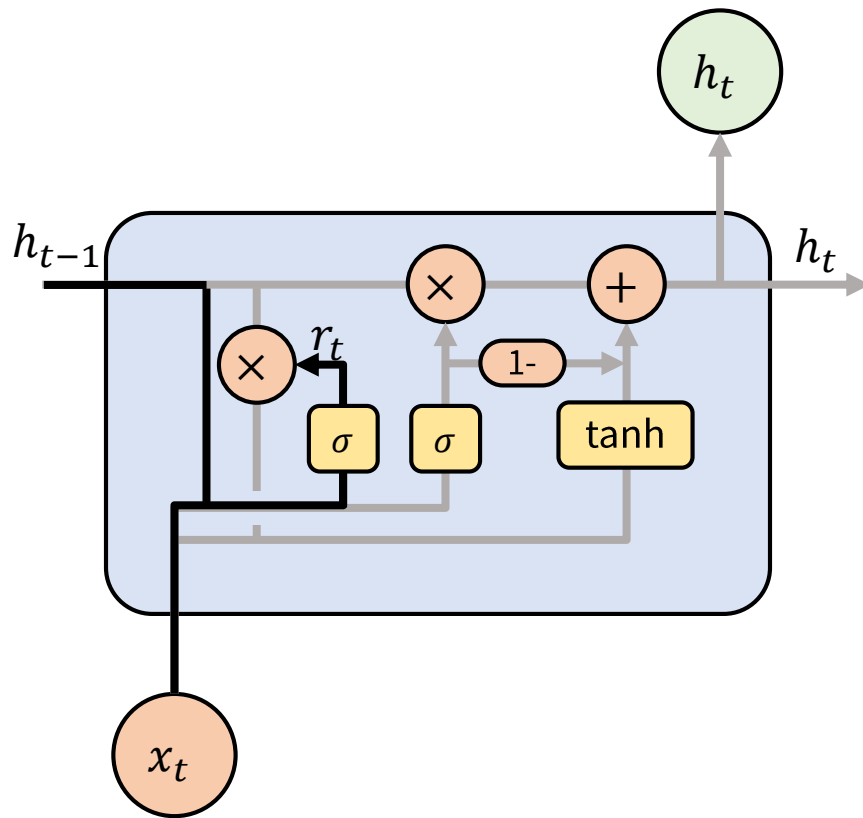
$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

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

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

LSTM보다 조금 더 간단한 GRU도 마찬가지로 수식을 통해서 살펴보자.

# Reset gate



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

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

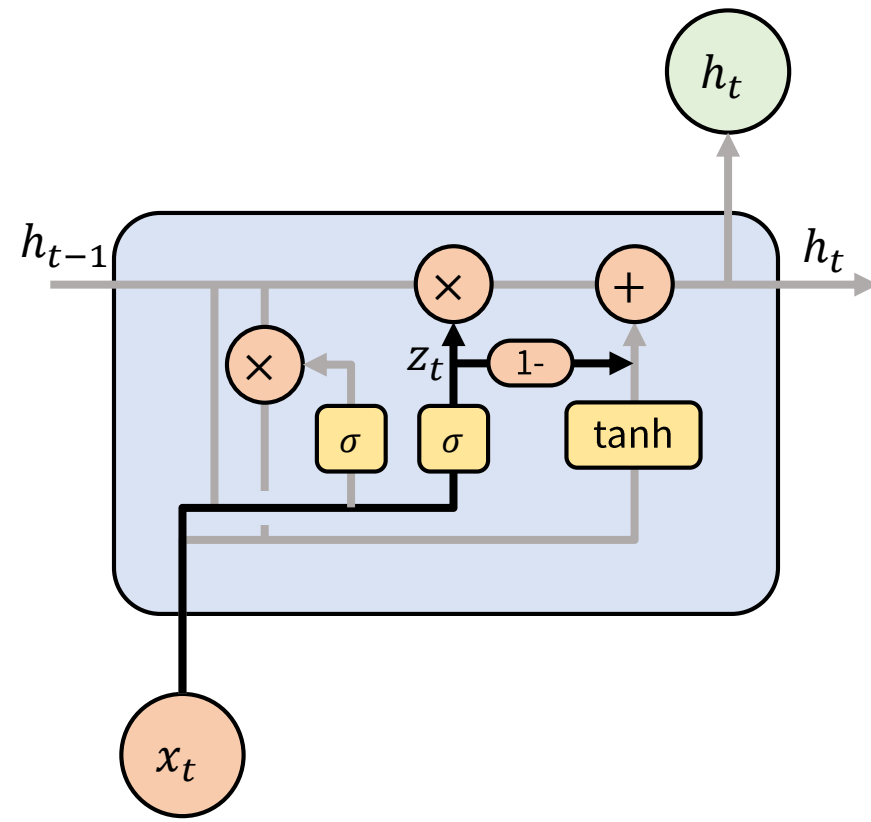
$$g_t = \tanh(W_{xg}x_t + W_{hg}(r_t \odot h_{t-1}) + b_g)$$

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

Reset gate는 Hidden state 중 어떤 특징을 reset할지 결정한다.

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

# Forget gate



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

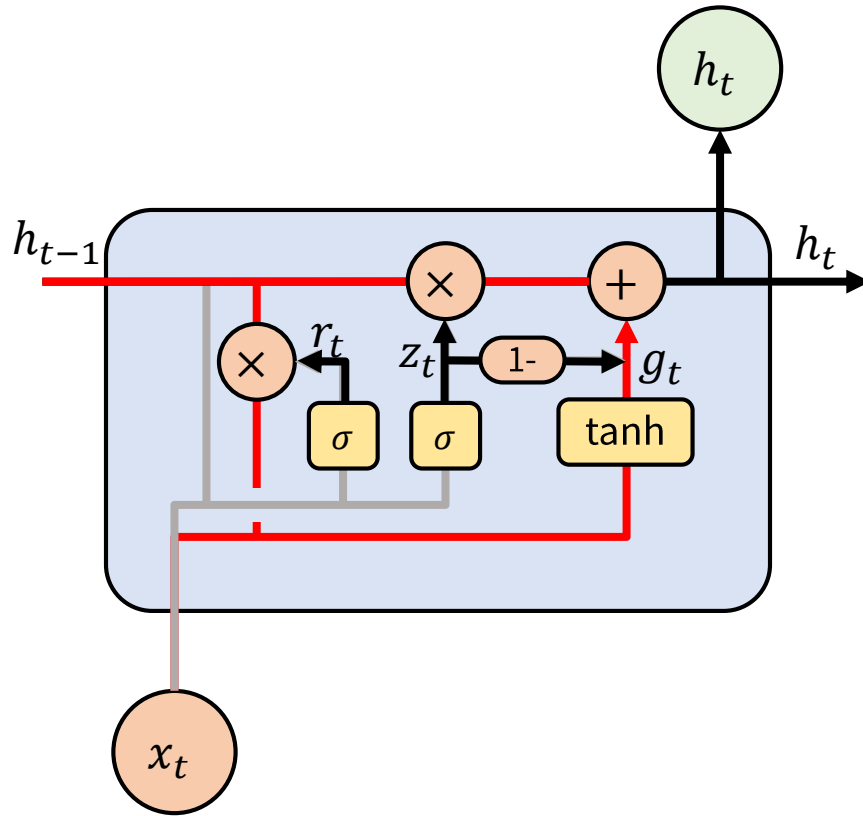
$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

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

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

Forget gate는 LSTM의 Forget gate와 Output gate를 겸한다.

# Hidden state



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

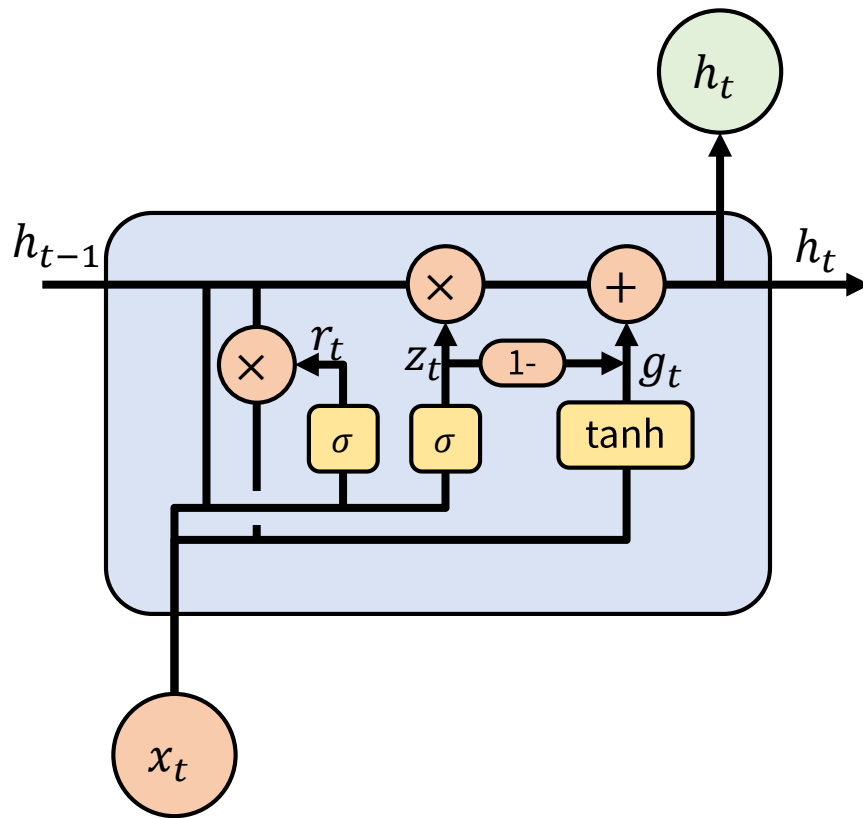
$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

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

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

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

# GRU Overview



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

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

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

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

LSTM보다 조금 더 간단한 GRU도 마찬가지로 수식을 통해서 살펴보자.