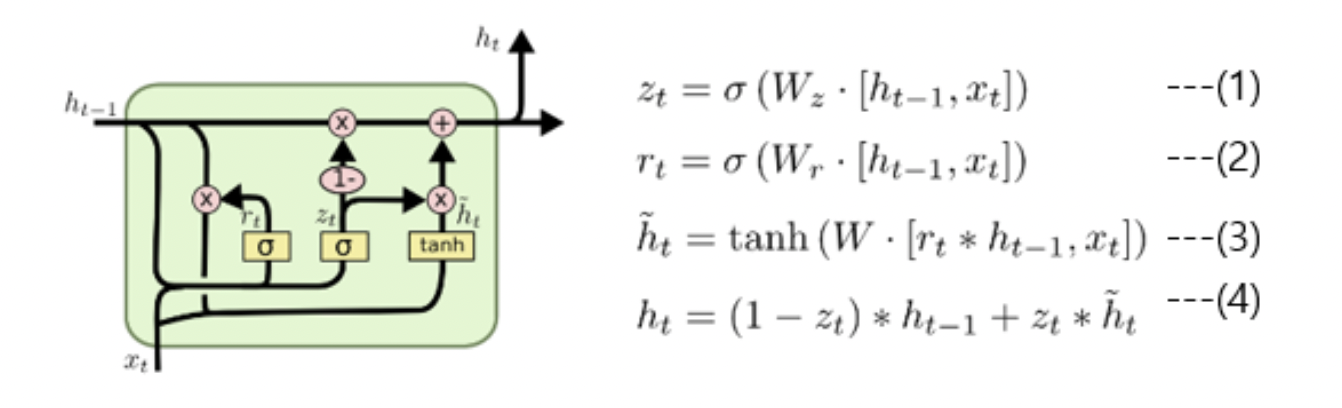
* GRU (Gated Recurrent Unit, 게이트 순환 유닛)

LSTM의 경우 forget gate, input gate, output gate 3개의 gate가 존재했던 반면 GRU에서는 reset gate, update gate 2개의 gate만이 존재한다.



* Reset Gate

reset gate는 이전 시점의 hidden state와 현 시점의 x를 sigmoid function을 적용해 구하는 방식으로, 위 그림의 공식 (2)를 통해 구할 수 있다. 이전 hidden state의 값을 얼마나 활용할 것인지에 대한 정보로 활용할 수 있다.

* Update Gate

update gate는 LSTM의 input, forget gate와 비슷한 역할을 하며, 과거와 현재의 정보를 각각 얼마나 반영할 것인지에 대한 비율을 구한다. 식 (1)을 통해 구한 z는 현재 정보를 얼마나 사용할지 반영하며, (1-z)는 과거 정보를 얼마나 사용할지를 반영한다. 그래서 각각의 역할을 LSTM의 input gate, forget gate의 역할로 볼 수 있고, 최종적으로 식 (4)를 통해 현 시점의 출력값 hidden state를 구할 수 있다.

* 장점
  + LSTM에 비해 구조가 간단하며, 학습할 parameter의 수가 적다
  + LSTM에 비해 학습 속도가 빠르다
  + 마지막 출력값에 활성화함수를 적용하지 않는다
* 단점
  + 성능 면에서 LSTM에 비해 우월한 것은 아니다