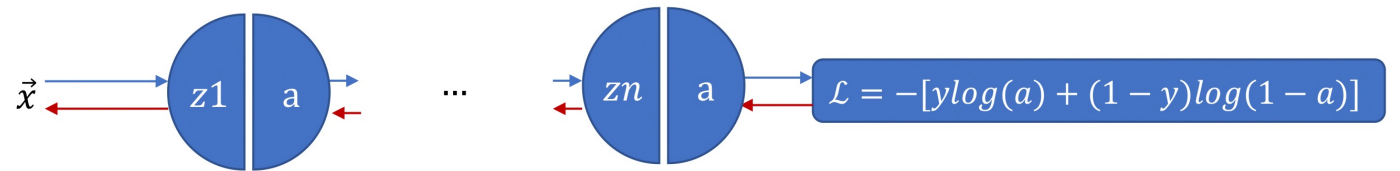


### Question. 7-05

다음과 같이 Tanh Function을 이용하여 Logistic Regression Multi-Layer를 쌓았다고 가정하자.  
Sigmoid Function으로 쌓은 경우와 비교하여 개선된 점과 문제점을 찾고 그 이유를 설명하라.



- 1) Backpropagation 진행시 Tanh Function에서 넘기는 최대값은 1이다.  
따라서 최대 0.25를 넘기는 Sigmoid Function보다 update 되는 양이 커져 연산속도가 빨라진다.  
그러나 레이어가 쌓일수록 1과 같거나 1보다 작은 값들이 계속 공해져  
결국 update되는 양이 매우 작아지는 Vanishing gradient problem이 그대로 남게된다.