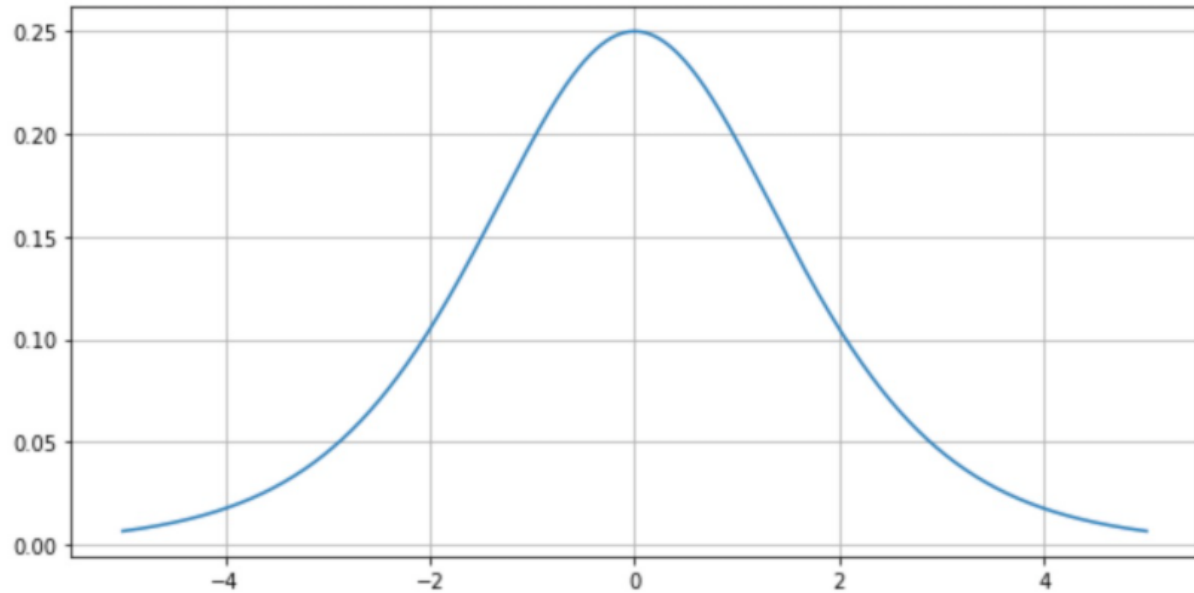


7.

$$\frac{d}{dx} \text{sigmoid}(x) = \text{sigmoid}(x)(1 - \text{sigmoid}(x))$$

다음은 sigmoid 미분결과(도함수)로 그래프를 그린 결과입니다.



그래프의 미분계수를 보면 최댓값은 0.25입니다. deep learning에서 학습을 위하여 역전파(Backpropagation)를 계산하는 과정에서 activation function의 미분 값을 곱하는 과정이 포함됩니다. sigmoid를 활성화 함수로 사용하는 경우 은닉층의 깊이가 깊다면 오차를 계산이 어렵다는 문제가 발생합니다. 이것이 딥러닝에 sigmoid를 사용할 때 “**vanishing gradient problem**”이 발생하는 이유입니다.