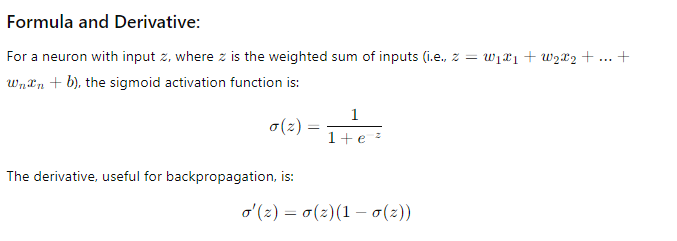
In deep learning, the **sigmoid function** is often used as an activation function for neurons, particularly in the output layer for binary classification problems. It transforms the output of a neuron into a value between 0 and 1, representing probabilities.

**Usage in Deep Learning:**

1. **Logistic Regression**: In binary classification, the sigmoid function is commonly used to map the output of a linear model to a probability, where a value greater than 0.5 is classified as one class, and a value below 0.5 is classified as the other class.
2. **Neural Networks**: The sigmoid can be used as an activation function in hidden layers, though modern architectures typically favor other activation functions (like ReLU or Leaky ReLU) for hidden layers due to performance reasons.

**Limitations in Deep Learning:**

* **Vanishing Gradient Problem**: During backpropagation, the gradients of the sigmoid function become very small for large values of xxx, slowing down learning as the updates to weights become smaller.
* **Non-zero-centered outputs**: The output of the sigmoid function is between 0 and 1, making it harder to optimize certain models because the gradients can push the updates in only one direction.

****