# SoftMax and Cross Entropy

Both those two methods serve the classification purpose

## SoftMax

### Why we need it

The softmax function, also known as the normalized exponential function. There are three purpose to apply softmax function

(1) restrict all outputs within the interval [0,1]

(2) all outputs will add up to 1

(3) new results remain in the same order as the original (monotonous)

### Definition

So, it is defended as

### Examples

## Cross Entropy

### Why we need it

When we evaluate the results for classification, we find the squared error is too strict. Usually, we only require the highest value is the right label. For instance, we have a, b, c three possible outputs, if b is the ground truth, we only care about whether the value of b is the largest. If the output of b is 0.6, it does not matter whether the value of a and c is (0.2,0.2) or (0,0.4). But from squared error estimation, the first situation is preferable.

To overcome the bias from norm, we apply Cross Entropy

### Definition