Glossary

Softmax

Marcel Turcotte

2025-07-10

The standard (unit) **softmax** function is a mathematical function that converts a vector of real numbers into a probability distribution, where the components of the vector are exponentiated and then normalized by dividing by the sum of all exponentiated components. This ensures that the output values are in the range $(0, 1)$ and sum up to 1, making them suitable for representing probabilities. Specifically, $\sigma : \mathbb{R}^K \to (0, 1)^K$, for $K > 1$ and $\mathbf{z} = (z_1, \dots, z_K) \in \mathbb{R}^K$.

$$\sigma(\mathbf{z}) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

```python
import numpy as np

def softmax(z):
  return np.exp(z) / np.sum(np.exp(z))

s = softmax([1.47, -0.39, 0.22])
print(s, sum(s))
```

```
[0.69339596 0.10794277 0.19866127] 1.0
```