

Softmax関数

softmax (ソフトマックス) 関数は、次の式で定義される関数です。

$$y_i = \frac{e^{x_i}}{\sum_{k=1}^n e^{x_k}} \quad (i = 1, 2, \dots, n)$$

なんだかややこしい関数ですが、一言でいえば「 n 個のごちゃごちゃなデータ $\{x_1, x_2, \dots, x_n\}$ を、 n 個全ての数の合計が 100% (1.0) となったデータ $\{y_1, y_2, \dots, y_n\}$ に変換する関数」です。以下では、この式に出てくる Σ や e の意味を一つ一つ確認していき、この式を理解していきます。

♣ 2 の小数乗？ マイナス乗？

softmax 関数を理解するために、まず累乗について確認しましょう。中学校では累乗に関しては次のように習うと思います。

「 a を n 回掛け算するとき、 $a \times a \times \dots \times a = a^n$ と表す」

具体的には、

$$2^3 = 2 \times 2 \times 2 = 8$$

$$(-2)^3 = (-2) \times (-2) \times (-2) = -8$$

といった具合です。この累乗の考え方は間違っていないのですが、次のような式が出てきたときに困ってしまいます。

$$2^0, \quad 2^{-3}, \quad 2^{0.5}$$

2 を「0 回」かける、2 を「-3 回」かける、2 を「0.5 回」かけた値と言われても、さっぱり分かりません。これらの値を考える際には、この「何回かける」という考え方を捨てて、次のようにして規則性を見出すのが良いでしょう。

$$\begin{array}{l} 2^3 = 8 \\ 2^2 = 4 \\ 2^1 = 2 \\ 2^0 = 1 \end{array} \begin{array}{l} \left. \begin{array}{l} \times \frac{1}{2} \\ \times \frac{1}{2} \\ \times \frac{1}{2} \end{array} \right\} \end{array}$$

3 乗、2 乗、1 乗と小さくしていくと、値は $\frac{1}{2}$ 倍されていくことが分かります。この規則性を用いると、0 乗は

$$2^0 = 1$$