

グラフの形状をみると、 $x = 0$ の部分を基点として、点対称なS字型のカーブとなっていることが分かります。sigmoid 関数の名前は、このグラフの形状に由来しています。sigmoid という英単語を辞書で調べてみると、「S字形、S字の」という意味になっています。つまり、sigmoid 関数は「S字の関数」ということでグラフの形状をそのまま表現したネーミングとなっています。

### ♣ より一般的な sigmoid 関数

sigmoid 関数をより一般化した関数として、logistic (ロジスティック) 関数が知られています。

$$f(x) = \frac{L}{1 + e^{-\alpha(x-x_0)}}$$

この関数に含まれる各パラメータは、次の働きをしています。

- $L$ : 関数  $f(x)$  の最大値を決めるパラメータ。関数  $f(x)$  は 0 から  $L$  の範囲の値を取る。
- $x_0$ : S 字カーブをしている曲線の真ん中の点<sup>18</sup>の位置を決めるパラメータ。
- $\alpha$ : 最小値から最大値に至る増加の緩急をコントロールするパラメータ<sup>19</sup>。

最初の sigmoid 関数の式は、これらのパラメータを、

$$L = 1, \quad x_0 = 0, \quad \alpha = 1$$

と設定したものになっています。なお、この sigmoid 関数は特に「標準 sigmoid 関数」と呼ばれます。一般には、 $\alpha$  が 1 以外の場合も含めて sigmoid 関数と呼ばれることがあります<sup>20</sup>。

### ♣ softmax 関数との関係

入力データが 2 個  $\{x_1, x_2\}$  の場合、softmax 関数の出力値  $y_1$  は次のように変形することができます<sup>21</sup>。

$$y_1 = \frac{e^{x_1}}{e^{x_1} + e^{x_2}} = \frac{1}{1 + e^{-(x_1-x_2)}}$$

この形は sigmoid 関数とよく似ています。つまり、 $n = 2$  の softmax 関数は sigmoid 関数となり、逆に sigmoid 関数を  $n = 2$  以外の場合まで一般化したものが softmax 関数だと解釈することができます。このことから、機械学習の分類問題において、2 値分類では sigmoid 関数、多値分類 (3 値以上) では softmax 関数が、出力層の活性化関数として用いられることが多くなっています。

---

<sup>18</sup>正確には「変曲点」と言います。

<sup>19</sup> $\alpha$  値が大きくなるほど急激な増加となり、step 関数のような形状に近づいていきます。

<sup>20</sup>ただしニューラルネットワークの文脈では、sigmoid 関数は通常、標準 sigmoid 関数のことを指します。

<sup>21</sup>分母と分子を両方  $e^{x_1}$  で割っています。