WIKIPEDIA

Sigmoid function

A **sigmoid function** is a mathematical function having a characteristic "S"-shaped curve or **sigmoid curve**. A common example of a sigmoid function is the logistic function shown in the first figure and defined by the formula:[1]

$$S(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1} = 1 - S(-x).$$

Other standard sigmoid functions are given in the Examples section.

Special cases of the sigmoid function include the Gompertz curve (used in modeling systems that saturate at large values of x) and the ogee curve (used in the spillway of some dams). Sigmoid functions have domain of all real numbers, with return (response) value commonly monotonically increasing but could be decreasing. Sigmoid functions most often show a return value (y axis) in the range 0 to 1. Another commonly used range is from −1 to 1.

A wide variety of sigmoid functions including the logistic and hyperbolic tangent functions have been used as the activation function of artificial neurons. Sigmoid curves are also common in statistics as cumulative distribution functions (which go from 0 to 1), such as the integrals of the logistic density, the normal density, and Student's t probability density functions. The logistic sigmoid function is invertible, and its inverse is the logit function.
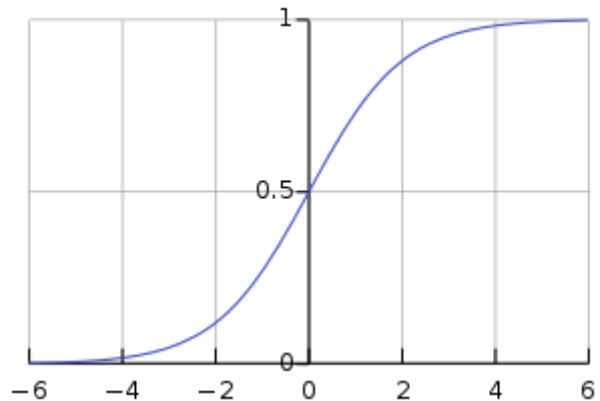
The logistic curve
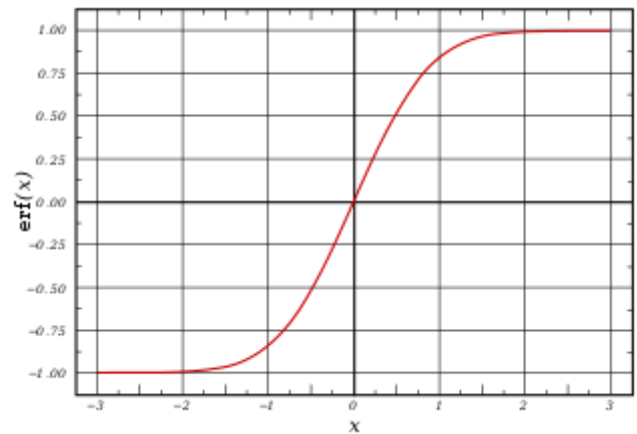
Plot of the error function

Contents

Definition

Properties

Examples

Applications

See also

References

Definition

A sigmoid function is a [bounded](), [differentiable](), real function that is defined for all real input values and has a non-negative derivative at each point[1] and exactly one [inflection point](). A sigmoid "function" and a sigmoid "curve" refer to the same object.

Properties

In general, a sigmoid function is [monotonic](), and has a first [derivative]() which is [bell shaped](). Conversely, the [integral]() of any continuous, non-negative, bell-shaped function (with one local maximum and no local minimum, unless degenerate) will be sigmoidal. Thus the [cumulative distribution functions]() for many common [probability distributions]() are sigmoidal. One such example is the [error function](), which is related to the cumulative distribution function of a [normal distribution]().

A sigmoid function is constrained by a pair of [horizontal asymptotes]() as $x \to \pm\infty$.

A sigmoid function is [convex]() for values less than 0, and it is [concave]() for values greater than 0.

Examples

- [Logistic function]()

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

- [Hyperbolic tangent]() (shifted and scaled version of the logistic function, above)

$$f(x) = \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- [Arctangent function]()

$$f(x) = \arctan x$$

- [Gudermannian function]()

$$f(x) = \mathrm{gd}(x) = \int_0^x \frac{dt}{\cosh t} = 2\arctan\left(\tanh\left(\frac{x}{2}\right)\right)$$

- [Error function]()

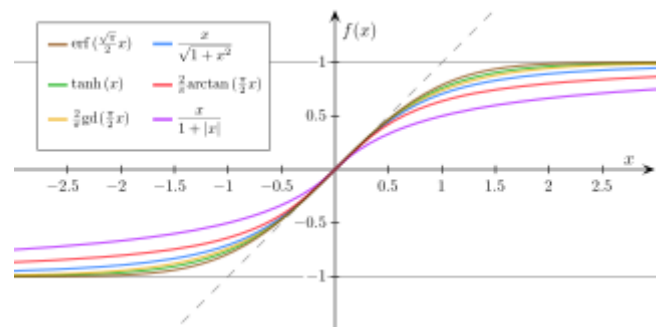$$f(x) = \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

- [Generalised logistic function]()

$$f(x) = \left(1 + e^{-x}\right)^{-\alpha}, \quad \alpha > 0$$

- [Smoothstep]() function

$$f(x) = \begin{cases} \left(\int_0^1 \left(1 - u^2\right)^N du\right)^{-1} \int_0^x \left(1 - u^2\right)^N\, du, & |x| \le 1 \\[2ex] \mathrm{sgn}(x) & |x| \ge 1 \end{cases} \quad N \ge 1$$

- Some [algebraic functions](), for example

Some sigmoid functions compared. In the drawing all functions are normalized in such a way that their slope at the origin is 1.

$$f(x) = \frac{x}{\sqrt{1+x^2}}$$

- and in a more general form

$$f(x) = \frac{x}{(1+|x|^k)^{1/k}} \text{ (see } [2])$$

Applications

Many natural processes, such as those of complex system learning curves, exhibit a progression from small beginnings that accelerates and approaches a climax over time. When a specific mathematical model is lacking, a sigmoid function is often used.[4]

The van Genuchten–Gupta model is based on an inverted S-curve and applied to the response of crop yield to soil salinity.

Examples of the application of the logistic S-curve to the response of crop yield (wheat) to both the soil salinity and depth to water table in the soil are shown in logistic function#In agriculture: modeling crop response.
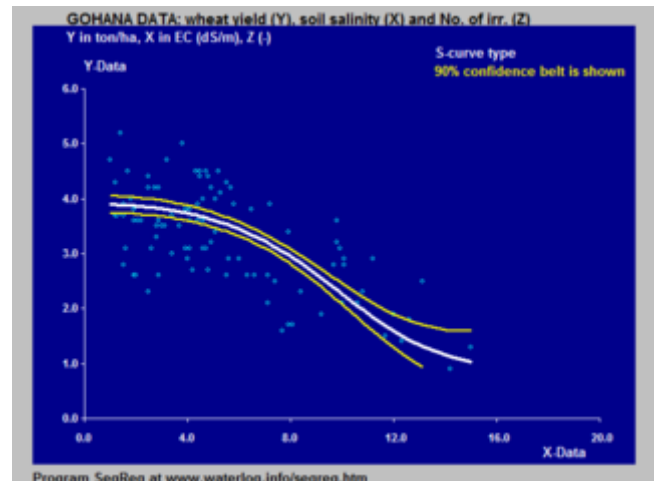
Inverted logistic S-curve to model the relation between wheat yield and soil salinity. [3]

In artificial neural networks, sometimes non-smooth functions are used instead for efficiency; these are known as hard sigmoids.

In audio signal processing, sigmoid functions are used as waveshaper transfer functions to emulate the sound of analog circuitry clipping.[5]

In biochemistry and pharmacology, the Hill equation and Hill–Langmuir equation are sigmoid functions.

In computer graphics and real-time rendering, some of the sigmoid functions are used to blend colors or geometry between two values, smoothly and without visible seams or discontinuities.

Titration curves between strong acids and strong bases have a sigmoid shape due to the logarithmic nature of the pH scale.

See also

- Heaviside step function
- Logistic regression
- Logit
- Softplus function
- Soboleva modified hyperbolic tangent
- Softmax function
- Swish function
- Weibull distribution

- Fermi–Dirac statistics

References

1. Han, Jun; Morag, Claudio (1995). "The influence of the sigmoid function parameters on the speed of backpropagation learning" (https://archive.org/details/fromnaturaltoart1995inte/page/1 95). In Mira, José; Sandoval, Francisco (eds.). *From Natural to Artificial Neural Computation*. Lecture Notes in Computer Science. **930**. pp. 195–201 (https://archive.org/details/fromnaturalto art1995inte/page/195). doi:10.1007/3-540-59497-3_175 (https://doi.org/10.1007%2F3-540-594 97-3_175). ISBN 978-3-540-59497-0.
2. Dunning, AJ; Kensler, Jennifer; Coudeville, Laurent; Bailleux, Fabrice (28 Dec 2015). "Some extensions in continuous methods for immunological correlates of protection" (https://www.ncbi. nlm.nih.gov/pmc/articles/PMC4692073). *BMC Medical Research Methodology*. **15** (107). doi:10.1186/s12874-015-0096-9 (https://doi.org/10.1186%2Fs12874-015-0096-9). PMC 4692073 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4692073). PMID 26707389 (http s://pubmed.ncbi.nlm.nih.gov/26707389).
3. Software to fit an S-curve to a data set [1] (https://www.waterlog.info/sigmoid.htm)
4. Gibbs, M.N. (Nov 2000). "Variational Gaussian process classifiers" (https://semanticscholar.org/ paper/66d429b63e0b8e329c565766289b4189c9398174). *IEEE Transactions on Neural Networks*. **11** (6): 1458–1464. doi:10.1109/72.883477 (https://doi.org/10.1109%2F72.883477). PMID 18249869 (https://pubmed.ncbi.nlm.nih.gov/18249869). S2CID 14456885 (https://api.se manticscholar.org/CorpusID:14456885).
5. Smith, Julius O. (2010). *Physical Audio Signal Processing* (https://ccrma.stanford.edu/~jos/pas p/Soft_Clipping.html) (2010 ed.). W3K Publishing. ISBN 978-0-9745607-2-4. Retrieved 28 March 2020.

- Mitchell, Tom M. (1997). *Machine Learning*. WCB–McGraw–Hill. ISBN 978-0-07-042807-2.. In particular see "Chapter 4: Artificial Neural Networks" (in particular pp. 96–97) where Mitchell uses the word "logistic function" and the "sigmoid function" synonymously – this function he also calls the "squashing function" – and the sigmoid (aka logistic) function is used to compress the outputs of the "neurons" in multi-layer neural nets.
- Humphrys, Mark. "Continuous output, the sigmoid function" (http://www.computing.dcu.ie/~hum phrys/Notes/Neural/sigmoid.html). Properties of the sigmoid, including how it can shift along axes and how its domain may be transformed.