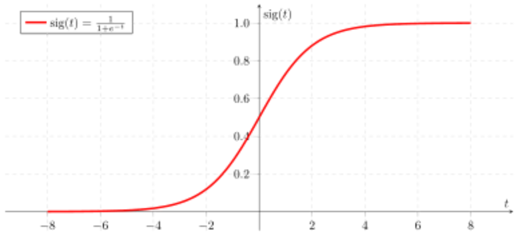
WIKIPEDIA

Sigmoidfunktion

Eine **Sigmoidfunktion**, **Schwanenhalsfunktion**, **Fermifunktion**[1] oder **S-Funktion** ist eine mathematische Funktion mit einem S-förmigen Graphen. Oft wird der Begriff Sigmoidfunktion auf den Spezialfall logistische Funktion bezogen, die durch die Gleichung



Die logistische Kurve

$$\operatorname{sig}(t) = \frac{1}{1+e^{-t}} = \frac{e^t}{1+e^t} = \frac{1}{2} \cdot \left(1 + \tanh \frac{t}{2}\right)$$

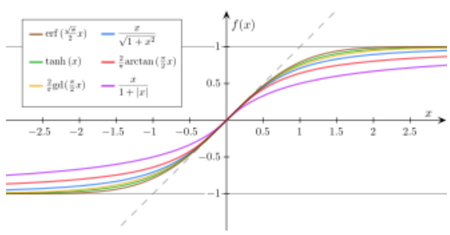
beschrieben wird. Dabei ist e die Eulersche Zahl. Diese spezielle Sigmoidfunktion ist also im Wesentlichen eine skalierte und verschobene Tangens-hyperbolicus-Funktion und hat entsprechende Symmetrien.

Die Umkehrfunktion dieser Funktion ist:

$$\operatorname{sig}^{-1}(y) = -\ln\left(\frac{1}{y} - 1\right) = \ln\left(\frac{y}{1-y}\right) = 2 \cdot \operatorname{artanh}(2 \cdot y - 1).$$

Inhaltsverzeichnis

Sigmoidfunktionen im Allgemeinen



Vergleich einiger Sigmoidfunktionen. Hier sind sie so normiert, dass ihre Grenzwerte −1 bzw. 1 sind und die Steigungen in 0 gleich 1 sind.

Im Allgemeinen ist eine Sigmoidfunktion eine beschränkte und differenzierbare reelle Funktion mit einer durchweg positiven oder durchweg negativen ersten Ableitung und genau einem Wendepunkt.

Die Menge der Sigmoidfunktionen enthält neben der logistischen Funktion den Arkustangens, den Tangens hyperbolicus und die Fehlerfunktion, die sämtlich transzendent sind, sowie auch einfache algebraische Funktionen wie $f(x) = \frac{x}{\sqrt{1+x^2}}$. Das Integral jeder stetigen, positiven Funktion mit einem „Berg“ (genauer: mit genau einem lokalen Maximum und keinem lokalen Minimum, z. B. die gaußsche Glockenkurve) ist ebenfalls eine Sigmoidfunktion. Daher sind viele kumulierte Verteilungsfunktionen sigmoidal.

Sigmoidfunktionen in neuronalen Netzwerken

Sigmoidfunktionen werden oft in künstlichen neuronalen Netzen als Aktivierungsfunktion verwendet, da der Einsatz von differenzierbaren Funktionen die Verwendung von Lernmechanismen, wie etwa dem Backpropagation-Algorithmus, ermöglicht. Als Aktivierungsfunktion eines künstlichen Neurons wird die Sigmoidfunktion auf die Summe der gewichteten Eingabewerte angewendet, um die Ausgabe des Neurons zu erhalten.

Die Sigmoidfunktion wird vor allem aufgrund ihrer einfachen Differenzierbarkeit als Aktivierungsfunktion bevorzugt verwendet, denn für die logistische Funktion gilt:

$$\operatorname{sig}'(t) = \operatorname{sig}(t)\,(1 - \operatorname{sig}(t))\,.$$

Für die Ableitung der Sigmoidfunktion Tangens hyperbolicus gilt:

$$\tanh'(t) = (1 + \tanh(t))\,(1 - \tanh(t)) = 1 - \tanh^2(t).$$

Effiziente Berechnung

Mit Unums vom Typ III lässt sich die oben angegebene logistische Funktion näherungsweise effizient berechnen, indem die Darstellung der Gleitkommazahl-Eingabe elegant genutzt wird.[2]

Siehe auch

- Logistische Verteilung
- Künstliches neuronales Netz
- Populationsdynamik
- Fermi-Dirac-Statistik
- Gompertz-Funktion

Weblinks

- Eric W. Weisstein: *Sigmoid Function*. (https://mathworld.wolfram.com/SigmoidFunction.html) In: *MathWorld* (englisch).

Einzelnachweise

1. *Einzelnes Neuron ::: Neuronale Netze.* (http://www.informatikseite.de/neuro/node16.php) Abgerufen am 4. April 2019.
2. John L. Gustafson, Isaac Yonemoto: *Beating Floating Point at its Own Game: Posit Arithmetic.* (http://www.johngustafson.net/pdfs/BeatingFloatingPoint.pdf) (PDF) 12. Juni 2017, abgerufen am 28. Dezember 2019 (englisch).