ON A GENERALIZED SIGMOID FUNCTION AND ITS PROPERTIES

K. NANTOMAH¹ *, C. A. OKPOTI² AND S. NASIRU³

¹ Department of Mathematics, Faculty of Mathematical Sciences, University for Development Studies, Navrongo Campus, P. O. Box 24, Navrongo, UE/R, Ghana

² Department of Mathematics, University of Education, Winneba, Ghana

³ Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies, Navrongo Campus, P. O. Box 24, Navrongo, UE/R, Ghana

*Correspondence: knantomah@uds.edu.gh

ABSTRACT. Motivated by generalized forms of the exponential and hyperbolic functions, we introduce a new generalization of the sigmoid function and further study some of its analytical properties.

1. INTRODUCTION

The sigmoid function, which is otherwise known as the standard logistic function is defined for \( z \in \mathbb{R} = (-\infty, \infty) \) as

\[
S(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}
\]

(1)

\[
= \frac{1}{2} + \frac{1}{2} \tanh \left( \frac{z}{2} \right).
\]

(2)

It is monotone increasing and its first derivative is bell-shaped. See the plots below.

This special function is applied in many scientific disciplines such as machine learning, artificial neural networks, computer graphics, medicine, probability and statistics, biology, ecology, population dynamics, demography, and mathematical psychology. See for instance [1], [2], [3], [4], [5], [6], [8], [10], [11], [12], [13], and the several related references therein. Due to the important roles of this function, recent works have focused on investigating its analytical properties. In [9], the authors studied properties of the function such as starlikeness and convexity in a unit disc. In [14], the author studied among other things

Keywords: Generalized exponential function; generalized hyperbolic functions; generalized sigmoid function; generalized softplus function; inequality.
some properties such as subadditivity, log-concavity, monotonicity and geometric concavity of the function. Thereafter, in [17], the authors introduced a generalization of the function and by adopting the techniques of [14], they proved analogous properties for the generalized function. They also considered some statistical properties of the generalized function.

Motivated by generalized forms of the exponential and hyperbolic functions, the goal of this paper is to introduce a new generalization of the sigmoid function and to further study some of its analytical properties.

**Definition 1.1 ([16])**. Let \( \phi : \mathbb{I} \subseteq \mathbb{R}^+ = (0, \infty) \to \mathbb{R}^+ \). Then \( \phi \) is said to be geometrically convex on \( \mathbb{I} \) if

\[
\phi(x^ry^{1-r}) \leq [\phi(x)]^r [\phi(y)]^{1-r}
\]

for all \( x, y \in \mathbb{I} \) and \( r \in [0, 1] \). If the inequality in (3) is reversed, then \( \phi \) is said to be geometrically concave on \( \mathbb{I} \).

**Lemma 1.2 ([15])**. Let \( \phi : \mathbb{I} \subseteq \mathbb{R}^+ \to \mathbb{R}^+ \) be a differentiable function. Then the following statements are equivalent.

(a) The function \( \phi \) is geometrically convex (concave).

(b) The function \( \frac{\phi'(x)}{\phi(x)} \) is increasing (decreasing).

2. **Generalized Sigmoid Function and its Properties**

In this section, we introduce a new generalization of the sigmoid function and further study some of its properties. We begin with the following auxiliary definitions.
Definition 2.1. Let $a > 1$. Then the generalized exponential function is defined as

$$a^z = \sum_{n=0}^{\infty} \frac{(\ln a)^n z^n}{n!},$$

for all $z \in \mathbb{R}$.

From this definition, we define the generalized hyperbolic cosine, hyperbolic sine and hyperbolic tangent functions as follows.

Definition 2.2. Let $a > 1$. Then the generalized hyperbolic cosine, hyperbolic sine and hyperbolic tangent functions are respectively defined as

$$\cosh_a(z) = \frac{a^z + a^{-z}}{2},$$

$$\sinh_a(z) = \frac{a^z - a^{-z}}{2},$$

$$\tanh_a(z) = \frac{\sinh_a(z)}{\cosh_a(z)} = \frac{a^z - a^{-z}}{a^z + a^{-z}} = 1 - \frac{2}{1 + a^{2z}},$$

for all $z \in \mathbb{R}$.

These generalized functions satisfy the following identities.

$$\cosh_a(z) + \sinh_a(z) = a^z,$$

$$\cosh_a(z) - \sinh_a(z) = a^{-z},$$

$$(\cosh_a(z))' = (\ln a) \sinh_a(z),$$

$$(\sinh_a(z))' = (\ln a) \cosh_a(z),$$

$$(\tanh_a(z))' = \frac{\ln a}{\cosh^2_a(z)},$$

$$(\cosh_a(z))'' + (\sinh_a(z))'' = (\ln a)^2 a^z,$$

$$(\cosh_a(z))'' - (\sinh_a(z))'' = (\ln a)^2 a^{-z},$$

$$\cosh^2_a(z) + \sinh^2_a(z) = \cosh_a(2z),$$

$$\cosh^2_a(z) - \sinh^2_a(z) = 1.$$
**Definition 2.3.** Let $a > 1$. Then the generalized sigmoid function is defined as

\[
S_a(z) = \frac{a^z}{1 + a^z},
\]

\[
= \frac{1}{2} + \frac{1}{2} \tanh \left( \frac{z}{2} \right),
\]

for all $z \in \mathbb{R}$. Clearly, when $a = e$, then $S_a(z)$ reduces to $S(z)$.

Direct computations reveals that this special function satisfies the following properties.

\[
S'_a(z) = (\ln a) \frac{a^z}{(1 + a^z)^2} = (\ln a) S_a(z) (1 - S_a(z)),
\]

\[
S''_a(z) = (\ln a)^2 \frac{a^z(1 - a^z)}{(1 + a^z)^3} = (\ln a)^2 S_a(z) (1 - S_a(z))(1 - 2S_a(z)),
\]

\[
(\ln S_a(z))'' + (\ln a) S'_a(z) = 0,
\]

\[
S_a(z) = 1 - S_a(-z),
\]

\[
S'_a(z) = (\ln a) S_a(z) S_a(-z),
\]

\[
\lim_{z \to \infty} S_a(z) = 1,
\]

\[
\lim_{z \to 0} S_a(z) = \frac{1}{2},
\]

\[
\lim_{z \to -\infty} S_a(z) = 0,
\]

\[
\lim_{z \to \pm \infty} S'_a(z) = 0,
\]

\[
\lim_{z \to 0} S'_a(z) = \frac{\ln a}{4},
\]

\[
\int S_a(z) \, dz = \frac{\ln(1 + a^z)}{\ln a} + k,
\]

where $k$ is a constant of integration. The function $\frac{\ln(1+a^z)}{\ln a}$ is a generalization of the softplus function which was first defined in [7].

**Remark 2.4.** Figure 3 and Figure 4 respectively show the plots of the generalized sigmoid function and its derivative for some particular values of $a$. It is evident that the shapes of these functions depend on the values of $a$. The maximum value of $S'_a(z)$ is $\frac{\ln a}{4}$ for any $a > 1$.

**Remark 2.5.** The function $S_a(z)$ is a cumulative distribution function in the sense that:

(a) $\lim_{z \to -\infty} S_a(z) = 0$,
(b) $\lim_{z \to \infty} S_a(z) = 1$,
(c) $S_a(z)$ is monotone increasing,
(d) $S_a(z)$ is right-continuous.

The function $S_a'(z)$ is the associated probability density function.

Theorem 2.6. The generalized sigmoid function satisfies the inequality

$$S_a(w + z) < S_a(w) + S_a(z),$$

for all $w, z \in \mathbb{R}$. That is, $S_a(z)$ is subadditive on $\mathbb{R}$.

Proof. Let $\psi(t) = \frac{t}{1+t}$ where $t > 0$. Then $\psi(t)$ is strictly increasing for all $t > 0$. Also, let $w, z \in \mathbb{R}$. Then,

$$S_a(w) + S_a(z) = \frac{a^w}{1 + a^w} + \frac{a^z}{1 + a^z} = \frac{a^w + a^z + 2a^{w+z}}{1 + a^w + a^z + a^{w+z}} \geq \frac{a^w + a^z + a^{w+z}}{1 + a^w + a^z + a^{w+z}} \geq \psi(a^w + a^z + a^{w+z}) \geq \psi(a^{w+z}) = S_a(w + z),$$

which gives the desired result. \hfill \Box

Theorem 2.7. The generalized sigmoid function satisfies the inequality

$$S_a(\delta w + (1 - \delta)z) \geq [S_a(w)]^\delta [S_a(z)]^{1-\delta},$$

for all $w, z \in \mathbb{R}$ and $\delta \in [0, 1]$. That is, $S_a(z)$ is logarithmically concave on $\mathbb{R}$.

Proof. It suffices to show that $(\ln S_a(z))'' \leq 0$ for all $z \in \mathbb{R}$. Let $z \in \mathbb{R}$. Then

$$(\ln S_a(z))'' = \left(\frac{S_a'(z)}{S_a(z)}\right)' = \frac{S_a''(z)S_a(z) - (S_a'(z))^2}{[S_a(z)]^2} = -\frac{(\ln a)^2 a^z}{(1 + a^z)^2} \leq 0,$$
which concludes the proof. □

**Remark 2.8.** Theorem 2.7 is equivalent to the following statements.

(a) \( S_a''(z)S_a(z) \leq (S_a'(z))^2 \) for all \( z \in \mathbb{R} \).

(b) \( \frac{S_a'(z)}{S_a(z)} \) is decreasing for all \( z \in \mathbb{R} \).

**Corollary 2.9.** The inequality

\[
S_a(x) - S_a(x-1) \geq S_a(x+1) - S_a(x),
\]

is satisfied for all \( x \in \mathbb{R} \).

**Proof.** By letting \( w = x - 1, z = x + 1 \) and \( \delta = \frac{1}{2} \) in Theorem 2.7, we obtain

\[
[S_a(x)]^2 \geq S_a(x-1)S_a(x+1),
\]

which yields inequality (32). □

**Theorem 2.10.** The generalized sigmoid function satisfies the inequality

\[
\left( \frac{a}{1+a} \right)^{1-\lambda} \leq \frac{[S_a(z+1)]^\lambda}{S_a(\lambda z + 1)} \leq 1,
\]

for \( z \in (0, \infty) \) and \( \lambda \geq 1 \). The inequality is reversed if \( 0 < \lambda \leq 1 \).

**Proof.** For \( z \in (0, \infty) \) and \( \lambda \geq 1 \), let \( \beta(z) = \frac{[S_a(z+1)]^\lambda}{S_a(\lambda z + 1)} \) and \( h(z) = \ln \beta(z) \). Then

\[
h'(z) = \lambda \left[ \frac{S_a'(z+1)}{S_a(z+1)} - \frac{S_a'(\lambda z + 1)}{S_a(\lambda z + 1)} \right] > 0,
\]

since \( \frac{S_a'(z)}{S_a(z)} \) is decreasing. Thus, \( \beta(z) \) is increasing. Then for \( z \in (0, \infty) \), we have

\[
f(0) \leq \beta(z) \leq \lim_{z \to \infty} \beta(z) = 1,
\]

which yields inequality (33). The proof for the case where \( 0 < \lambda \leq 1 \), follows a similar procedure. Hence we omit the details. □

**Theorem 2.11.** For \( \alpha \in [0, 1] \), the inequalities

\[
\frac{1 - \alpha}{2} \leq S_a(\alpha z) - \alpha S_a(z) \leq 1 - \alpha, \quad z \in (0, \infty),
\]

\[
0 \leq S_a(\alpha z) - \alpha S_a(z) \leq \frac{1 - \alpha}{2}, \quad z \in (-\infty, 0),
\]

are satisfied.

**Proof.** Note that

\[
(S_a'(z))^2 = S_a''(z) = (\ln a)^2 a^2(1-a^2)/(1+a^2)^3 > 0
\]
respectively if \( z \in (-\infty, 0) \) or \( z \in (0, \infty) \). Thus, \( S'_a(z) \) is increasing for \( z \in (-\infty, 0) \) and decreasing for \( z \in (0, \infty) \). Let \( \alpha \in [0, 1] \) and \( \phi(z) = S_a(\alpha z) - \alpha S_a(z) \). Now suppose that \( z \in (0, \infty) \). Then \( \alpha z < z \) and then

\[
\phi'(z) = \alpha [S'_a(\alpha z) - S'_a(z)] > 0,
\]

since \( S'_a(z) \) is decreasing for \( z \in (0, \infty) \). Hence \( \phi(z) \) is increasing for \( z \in (0, \infty) \). Therefore

\[
\frac{1 - \alpha}{2} = \phi(0) \leq \lim_{z \to \infty} \phi(z) = 1 - \alpha,
\]

which gives (34). Next suppose that \( z \in (-\infty, 0) \). Then \( \alpha z > z \) and

\[
\phi'(z) = \alpha [S'_a(\alpha z) - S'_a(z)] > 0,
\]

since \( S'_a(z) \) is increasing for \( z \in (-\infty, 0) \). Hence \( \phi(z) \) is increasing for \( z \in (-\infty, 0) \) as well. Consequently

\[
0 = \lim_{z \to -\infty} \phi(z) \leq \phi(z) \leq \phi(0) = \frac{1 - \alpha}{2},
\]

which gives (35).

**Remark 2.12.** Inequalities (34) and (35) imply that \( S_a(\alpha z) \geq \alpha S_a(z) \) for all \( z \in (-\infty, \infty) \) and \( \alpha \in [0, 1] \). In other words, the function \( S_a(z) \) is \(+\infty\)-star-shaped.

**Theorem 2.13.** The inequality

\[
(\ln a)S_a(x) < \frac{\ln(1 + a^z) - \ln(1 + a^x)}{z - x} < (\ln a)S_a(z),
\]

holds for \( 0 \leq x < z \).

**Proof.** Let \( 0 \leq x < z \) and consider the function \( \gamma(t) = \ln(1 + a^t) \) on the interval \((x, z)\). By the mean value theorem, there exist a \( c \in (x, z) \) such that

\[
\frac{\ln(1 + a^z) - \ln(1 + a^x)}{z - x} = \gamma'(c).
\]

Since \( \gamma'(t) = (\ln a)S_a(t) \) is increasing, then for \( c \in (x, z) \), we have

\[
\gamma'(x) < \gamma'(c) < \gamma'(z),
\]

which is equivalent to (36).

**Remark 2.14.** If \( x = 0 \) in Theorem 2.13, then we obtain

\[
\ln 2 + \frac{(\ln a)}{2} z < \ln(1 + a^z) < \ln 2 + (\ln a)zS_a(z), \quad z > 0,
\]

which gives an estimate for the generalized softplus function in terms of the generalized sigmoid function.
Theorem 2.15. Let \( r \geq 1 \). Then

\[
\left( \frac{S_a(x)}{S_a(y)} \right)^r \leq \frac{S_a(rx)}{S_a(ry)},
\]

if \( x \leq y \) and

\[
\left( \frac{1}{2} \right)^{r-1} \leq \frac{[S_a(x)]^r}{S_a(rx)} \leq \frac{1 + ar}{(1 + a)^r},
\]

if \( x \in (0,1) \). Equality holds when \( r = 1 \).

Proof. Let \( r \geq 1 \), \( \theta(x) = \frac{[S_a(x)]^r}{S_a(rx)} \), and \( g(x) = \ln \theta(x) \). Then

\[
g'(x) = r \left[ \frac{S_a'(x)}{S_a(x)} - \frac{S_a'(rx)}{S_a(rx)} \right] > 0.
\]

Thus, \( \theta(x) \) is increasing. Then for \( x \leq y \), we have

\[ \theta(x) \leq \theta(y), \]

which yields (37). Similarly for \( x \in (0,1) \), we have

\[ \left( \frac{1}{2} \right)^{r-1} = \lim_{x \to 0} \theta(x) \leq \theta(x) \leq \lim_{x \to 1} \theta(x) = \frac{1 + ar}{(1 + a)^r}, \]

which yields (38). This completes the proof. \( \square \)

Remark 2.16. If \( 0 < r < 1 \) in Theorem 2.15, then we obtain the reverse cases of inequalities (37) and (38).

Theorem 2.17. The inequality

\[
S_a'(x^\lambda y^{1-\lambda}) \geq [S_a'(x)]^\lambda [S_a'(y)]^{1-\lambda},
\]

holds for \( x, y \in \mathbb{R}^+ \) and \( \lambda \in [0,1] \). That is, \( S_a'(z) \) is geometrically concave on \( \mathbb{R}^+ \).

Proof. Let \( z \in \mathbb{R}^+ \). Then it follows from (19) and (20) that

\[
\frac{S_a''(z)}{S_a'(z)} = (\ln a) \frac{1 - az}{1 + az},
\]

and by direct computation, we obtain

\[
\left\{ \frac{z S_a''(z)}{S_a'(z)} \right\}' = (\ln a) \left[ \frac{1 - az}{1 + az} - 2(\ln a) \frac{za^2}{(1 + az)^2} \right] < 0.
\]

In view of Lemma 1.2, we conclude that \( S_a'(z) \) is geometrically concave on \( \mathbb{R}^+ \) and this is equivalent to inequality (39). \( \square \)
Theorem 2.18. The inequalities

\begin{align}
1 < & \frac{S_a(z + 1)}{S_a(z)} < a, \\
0 < & \frac{S_a'(z)}{S_a(z)} < \ln a,
\end{align}

hold for all \( z \in \mathbb{R} \).

Proof. Let \( \Delta(z) = \frac{S_a(z+1)}{S_a(z)} \) for all \( z \in \mathbb{R} \). Then

\[
\lim_{z \to -\infty} \Delta(z) = \lim_{z \to -\infty} \frac{a + a^{z+1}}{1 + a^{z+1}} = a
\]

and

\[
\lim_{z \to \infty} \Delta(z) = \lim_{z \to \infty} \frac{a + a^{z+1}}{1 + a^{z+1}} = 1.
\]

Also,

\[
\frac{\Delta'(z)}{\Delta(z)} = \frac{S_a'(z + 1)}{S_a(z + 1)} - \frac{S_a'(z)}{S_a(z)} < 0,
\]

since \( \frac{S_a'(z)}{S_a(z)} \) is decreasing. Thus, \( \Delta(z) \) is decreasing. Hence for all \( z \in \mathbb{R} \), we have

\[
1 = \lim_{z \to \infty} \Delta(z) < \Delta(z) < \lim_{z \to -\infty} \Delta(z) = a,
\]

which gives inequality (40). Next we have

\[
\lim_{z \to -\infty} \frac{S_a'(z)}{S_a(z)} = \lim_{z \to -\infty} \frac{\ln a}{1 + a^z} = \ln a,
\]

and

\[
\lim_{z \to \infty} \frac{S_a'(z)}{S_a(z)} = \lim_{z \to \infty} \frac{\ln a}{1 + a^z} = 0.
\]

Then by the monotonicity of \( \frac{S_a'(z)}{S_a(z)} \), we obtain

\[
\lim_{z \to \infty} \frac{S_a'(z)}{S_a(z)} < \frac{S_a'(z)}{S_a(z)} < \lim_{z \to -\infty} \frac{S_a'(z)}{S_a(z)},
\]

which gives inequality (41). \( \square \)

Theorem 2.19. The inequalities

\begin{align}
\frac{1}{a} < & \frac{S_a'(z + 1)}{S_a'(z)} < a, \\
-\ln a < & \frac{S_a''(z)}{S_a'(z)} < \ln a,
\end{align}

hold for all \( z \in \mathbb{R} \).
Proof. Recall that $\frac{S''(z)}{S'(z)} = (\ln a)\frac{1-a^z}{1+a^z}$ for all $z \in \mathbb{R}$. Then

$$\left\{ \frac{S''(z)}{S'(z)} \right\}' = -2(\ln a)^2 \frac{a^z}{(1+a^z)^2} < 0,$$

which implies that $\frac{S''(z)}{S'(z)}$ is decreasing. Also,

$$\lim_{z \to -\infty} \frac{S''(z)}{S'(z)} = \ln a \quad \text{and} \quad \lim_{z \to \infty} \frac{S''(z)}{S'(z)} = -\ln a.$$

Hence for $z \in \mathbb{R}$, we have

$$\lim_{z \to -\infty} \frac{S''(z)}{S'(z)} < \frac{S''(z)}{S'(z)} < \lim_{z \to \infty} \frac{S''(z)}{S'(z)},$$

which gives inequality (43). Next, let $D(z) = \frac{S'(z+1)}{S'(z)}$ for all $z \in \mathbb{R}$. Then

$$\lim_{z \to -\infty} D(z) = \lim_{z \to -\infty} a \left( \frac{1+a^z}{1+a^{z+1}} \right)^2 = a$$

and

$$\lim_{z \to \infty} D(z) = \lim_{z \to \infty} a \left( \frac{1+a^z}{1+a^{z+1}} \right)^2 = \frac{1}{a}.$$

In addition, we obtain

$$\frac{D'(z)}{D(z)} = \frac{S''(z+1)}{S'(z+1)} - \frac{S''(z)}{S'(z)} < 0.$$

Thus, $D(z)$ is decreasing. Hence for all $z \in \mathbb{R}$, we have

$$\frac{1}{a} = \lim_{z \to \infty} D(z) < D(z) < \lim_{z \to -\infty} D(z) = a,$$

which gives inequality (42). \hfill \Box

**Theorem 2.20.** The inequalities

(44) \quad \frac{a}{1+a} \leq \frac{S_a(x+1)S_a(y+1)}{S_a(x+y+1)} < 1,

holds for all $x, y \in [0, \infty)$.

Proof. Let $\lambda(x, y) = \frac{S_a(x+1)S_a(y+1)}{S_a(x+y+1)}$ for all $x, y \in [0, \infty)$. Further let $h(x, y) = \ln \lambda(x, y)$. Then

$$\frac{\partial}{\partial x} h(x, y) = \frac{S_a(x+1)}{S_a(x+1)} - \frac{S_a(x+y+1)}{S_a(x+y+1)} > 0.$$

Thus, $\lambda(x, y)$ is increasing. Hence

$$\frac{a}{1+a} = \lim_{x \to 0} \lambda(x, y) \leq \lambda(x, y) < \lim_{x \to \infty} \lambda(x, y) = S_a(y+1) < 1,$$

which gives inequality (44). \hfill \Box
3. Conclusion

In this work, we have introduced a new generalization of the sigmoid function which is frequently applied in machine learning, artificial neural networks, computer graphics, medicine, probability and statistics, biology, ecology, population dynamics, demography, and mathematical psychology. It is our fervent hope that, this new generalization and the properties established thereafter, will find useful applications in these scientific disciplines. We also anticipate that for some particular values of $a$, this generalized function may be applied in modeling datasets with heavy tails.

References