Information Journal of Soft Computing 1 (1) : 76-82, 2006

Function Approximation Using Feedforward Networks with Sigmoidal Signals

[1]M. Ramakrishnan, [1]C.M. Velu [1]N. Prabakaran,
[1]K. Ekambaravanan, [1]P. Vivekanandan and [2]P. Thangavelu
[1]Department of Mathematics, Anna University, Chennai, India
[2]Department of Mathematics, Aditanar College, Tiruchendur-628 216, India

Abstract: T. Chen *et al*. and Babri studied Function approximations using sigmoidal and generalized sigmoidal functions. In this study we introduce left sigmoidal and right sigmoidal signals and use them to study function approximations in $C(\bar{R}^n)$.

Key words: Left Sigmoidal, right sigmoidal, function approximation, feedforward networks

INTRODUCTION

Kolmogorov *et al.*[1-4] was pioneer in function approximation theory by neural networks. Since then, many researchers have concentrated on this topic. Funahahsi[5] established that any continuous mapping can be approximately realized by Rumelhart- Hinton-William's multilayer neural networks with at least one hidden layer whose output function is sigmoidal. Hornick *et al.*[6] proved that the standard multi-layer feed forward networks were capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy. Verakurkov[8] derived the estimates of number of hidden units based on the properties of the function being approximated and the accuracy of its approximation. Chen *et al.*[1] investigated the capability of approximating functions in

$$C(\bar{R}^n)$$

by three layer feed forward networks with sigmoidal function in the hidden layer. In this paper, we use the left sigmoidal signals and the right sigmoidal signals for function approximation in

$$C(\bar{R}^n)$$

PRELIMINARIES

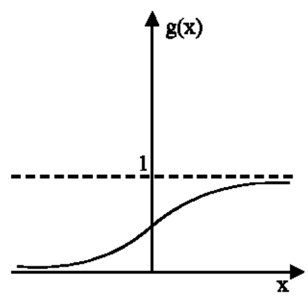
We list certain definitions and results that will be useful in sequel.

Definition 2.1: σ: R → R is called a generalized sigmoidal function,

$$\text{if } \lim_{x\to-\infty}\sigma(x)=0 \quad and \quad \lim_{x\to+\infty}\sigma(x)=1 .$$

The following is an example of a generalized sigmoidal function.

The following theorem holds for the generalized sigmoidal functions.



$$\sigma(x) = \frac{1}{1+e^{-\alpha x}} \ , \alpha > 0$$

Theorem 2.2: If σ(x) is a bounded generalized sigmoidal function and f (x) is a continuous function on (-∞,∞), for which

$$\lim_{x\to-\infty} f(x) = A$$

and

$$\lim_{x\to-\infty} f(x) = B,$$

where A, B are constants, then for any ε>0, there exist N and scalars c_i, y_i, θ_i such that

$$\left| f(x) - \sum_{i=1}^{N} c_i \sigma(y_i x + \theta_i) \right| < \varepsilon$$

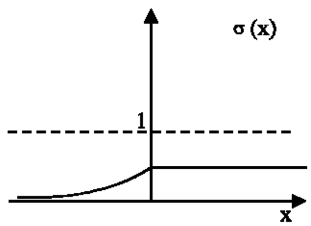
We now introduce the left sigmoidal function and right sigmoidal function.

Definition 2.3: The function σ: R→R is said to be left sigmoidal

Corresponding Author: P. Vivekanandan, Department of Mathematics, Anna University, Chennai-25, India

$$\text{if } \lim_{x \to -\infty} \sigma(x) = 0.$$

The following is an example of a left generalized sigmoidal function.



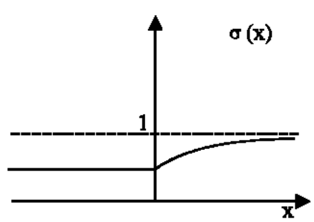
$$\sigma(x) = \begin{cases} e^{\alpha x}, & \text{for } x < 0,\ \alpha > 0 \\ \beta, & \text{for } x \geq 0,\ \beta > 0 \end{cases}$$

Definition 2.4: The function σ: R→ R is said to be right sigmoidal

$$\text{if } \lim_{x \to +\infty} \sigma(x) = 1$$

The following is an example of a right generalized sigmoidal function.

$$\sigma(x) = \begin{cases} 1 + e^{-\alpha x}, & \text{for } x > 0,\ \alpha > 0 \\ \beta, & \text{for } x < 0,\ \beta > 0 \end{cases}$$



Theorem 2.5: Every generalized sigmoidal signal can be decomposed into a left sigmoidal signal and a right sigmoidal signal.

Proof: Let σ: R → R be a generalized sigmoidal signal. $\sigma(0) = \alpha$. Let

$$\sigma_1(x) = \begin{cases} \sigma(x), & \text{if } x < 0 \\ \alpha, & \text{if } x \geq 0 \end{cases} \text{ and } \sigma_2(x) = \begin{cases} \alpha, & \text{if } x < 0 \\ \sigma(x), & \text{if } x \geq 0 \end{cases}.$$

Then it can be proved that σ_1 is a left sigmoidal signal and σ_2 is a right sigmoidal signal.

The following theorem shows that a left sigmoidal signal and right sigmoidal signal can be pasted to form a generalized sigmoidal signal.

Theorem 2.6: If σ_1 is a left sigmoidal signal and σ_2 is a right sigmoidal signal.
such that $\sigma_1(0) = \sigma_2(0)$ and

$$\sigma(x) = \begin{cases} \sigma_1(x), & \text{if } x \leq 0 \\ \sigma_2(x), & \text{if } x \geq 0 \end{cases}$$

then σ is a generalized sigmoidal signal.

Proof: Since σ_1 is a left sigmoidal signal.

$$\lim_{x \to -\infty} \sigma(x) = \lim_{x \to -\infty} \sigma_1(x) = 0$$

Since σ_2 is a right sigmoidal signal

$$\lim_{x \to +\infty} \sigma(x) = \lim_{x \to +\infty} \sigma_2(x) = 1 \cdot$$

This establishes that σ is a generalized sigmoidal signal.

Theorem 2.7: If $\sigma_1(x)$ is a bounded left sigmoidal signal and $\sigma_2(x)$ is a bounded right sigmoidal signal such that $\sigma_1(0) = \sigma_2(0)$ and f(x) is a continuous function on $(-\infty, \infty)$, for which

$$\lim_{x \to -\infty} f(x) = A \text{ and } \lim_{x \to +\infty} f(x) = B$$

where A and B are constants, then for any $\epsilon > 0$, there exist a positive integer N and the scalars c_i, y_i, θ_i such that

$$\left| f(x) - \sum_{i=1}^{N} c_i \sigma(y_i x + \theta_i) \right| < \varepsilon,\ \forall\ x \in (-\infty, \infty)$$

.

where

$$\sigma(x) = \begin{cases} \sigma_1(x), & \text{if } x \leq 0 \\ \sigma_2(x), & \text{if } x \geq 0 \end{cases}$$

Proof. Follows from Theorem 2.6 and Theorem 2.2.

FUNCTION APPROXIMATION

In this section we show that a continuous function can be approximated by a bounded left sigmoidal signal (resp. right sigmoidal signal).

Theorem 3.1: If $\sigma(x)$ is a bounded left sigmoidal function and f(x) is a continuous function in $(-\infty, \infty)$ with

$$\lim_{x\to-\infty} f(x) = A$$

where A is a constant, then for any $\epsilon>0$, there exist N and scalars c_i. y_i, θ_i such that

$$\left| f(x) - \sum_{i=1}^{N} c_i \sigma(y_i x + \theta_i) \right| < \varepsilon, \forall x \in (-\infty, 0)$$

.

Proof: Given that

$$\lim_{x\to-\infty} f(x) = A.$$

For any $\epsilon>0$, we can find M, N>0, such that

$$|f(x) - A| < \frac{\varepsilon}{2} \text{ if } x < -M; \text{and} |f(x_1) - f(x_2)| < \frac{\varepsilon}{4} \text{ if } x > M.$$

If $x_1 \le M$ and $x_2 \le M$ and

$$|x_1 - x_2| \le \frac{1}{N},$$

then divide [-M, 0] into 2MN equal segments and each has length of

$$\frac{1}{N}$$

and

$$-M = x_0 < x_1 < \ldots\ldots < X_{2MN} = 0.$$

$$\text{Let } t_i = \frac{x_i + x_{i+1}}{2}, \quad i = 0, 1, 2, \ldots\ldots., MN - 1.$$

$$\text{Let } g(x) = f(-M) + \sum_{i=1}^{2MN} [f(x_i) - f(x_{i-1})] \sigma(K(x - t_{i-1}))$$

for some K to be determined later. By the definition of Left sigmoidal signals, there exists W>0 such that

$$\text{if } u > W \text{ then } |\sigma(u) - 1| < \frac{1}{MN};$$

$$\text{If } u < -W, \text{ then } |\sigma(u)| < \frac{1}{MN}.$$

Choose K such that $K.\ \frac{1}{2N} > W$.

To prove: $|f(x)-g(x)|<\varepsilon$ for all $x \in (-\infty, 0)$.

Case 1: $x<-M$ and

$$|x - (-M)| < \frac{1}{N},$$

Then

$$|f(x) - f(-M)| < \frac{\varepsilon}{2}$$

Since $g(x) = f(-M) + \sum_{i=1}^{2MN} [f(x_i) - f(x_{i-1})] \sigma(K(x - t_{i-1}))$ ·

$$|g(x) - f(-M)| \le \sum_{i=1}^{2MN} |f(x_i) - f(x_{i-1})| \quad |\sigma(K(x - t_{i-1}))|$$

$$\le \sum_{i=1}^{2MN} \frac{\varepsilon}{4} \frac{1}{MN} = \frac{\varepsilon}{2}.$$

Now

$$|f(x) - g(x)| = |f(x) - f(-M) + f(-M) - g(x)|$$

$$\le |f(x) - f(-M)| + |g(x) - f(-M)|.$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

$= \epsilon$ for all $x \in (-\infty, 0)$.

Case 2: $x \in [x_{k-1}, x_k]$ Then

$$|x - t_{i-1}| \begin{cases} \le \frac{1}{2N} & \text{for } i = k \\ > \frac{1}{2N} & \text{for } i \ne k \end{cases}$$

Further more, If $i<k$, then $K(x-t_{i-1})>W$. Hence

$$|\sigma(K(x - t_{i-1})) - 1)| < \frac{1}{MN};$$

if $i>k$, then $K(x-t_{i-1})<-W$ and hence

$$|\sigma(K(x - t_{i-1}))| < \frac{1}{MN}$$

Consequently we have

$$\left|\begin{matrix} g(x)-f(-M)-[f(x_k)-f(x_{k-1})] \\ \sigma(K(x-t_{k-1}))-\sum_{i=1}^{k-1}[f(x_i)-f(x_{i-1})] \end{matrix}\right|$$

$$\leq \sum_{i=1}^{k-1}\left| f(x_i)-f(x_{i-1})\right| \left|\sigma(K(x-t_{k-1}))-1\right|$$

$$+ \sum_{i=k+1}^{2MN} \left| f(x_i)-f(x_{i-1})\right| \times \left|\sigma\ (K(x-t_{k-1}))\ \right|$$

$$\leq \sum_{i=1}^{k-1}\frac{\varepsilon}{4}\frac{1}{MN}+\sum_{i=k+1}^{2MN}\frac{\varepsilon}{4}\frac{1}{MN}$$

$$< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \epsilon/2 \text{ for all } x \in (0, \infty).$$

It is clear that

$$f(-M)+[f(x_k)-f(x_{k-1})]\sigma(K(x-t_{k-1}))+\sum_{i=1}^{k-1}[f(x_i)-f(x_{i-1})]$$

$$=f(x_{k-1})+[f(x_k)-f(x_{k-1})]\sigma(K(x-t_{k-1}))$$

$$\therefore \left| f(x)-g(x)\right|<\frac{\varepsilon}{2}+\left|f(x_k)-f(x_{k-1})\right| \left| \sigma(K(x-t_{k-1}))\right|$$

$$<\frac{\varepsilon}{2}+\frac{\varepsilon}{2} =\varepsilon .$$

Theorem 3.2. If σ (x) is a bounded right sigmoidal function and f(x) is a continuous function on (-∞,∞) with

$$\lim_{x\to+\infty} f(x) = B$$

where B is a constant, then for any $\epsilon>0$, there exist a positive integer N and scalars c_i, y, θ_i, such that

$$\left|f(x)-\sum_{i=1}^{N}c_i\, s\,(y_i\, x+?_i)\right|<e \quad \text{for all} \quad x\in(0,\infty)\cdot$$

Proof. Since f is continuous on (-∝. ∝).for any $\epsilon>0$, we can find M>0 such that and

$$\left|f(x_1) - f(x_2)\right|<\frac{\varepsilon}{8} \qquad \text{............(*)}$$

whenever $|x_1|\leq M, |x_2|\leq M$. Choose N such that $|x_1|\leq M.|x_2|\leq M$ and

$$|x_1\text{-}x_2|\leq \frac{1}{N}.$$

Divide (0, M) into 2MN equal segments, each has length of and let $0 = x_0<x_1<\ldots\ldots\ldots\ldots<x_{2MN} = M$.
Since

$$\lim_{x\to+\infty} f(x) = B,$$

we choose $x_0.\ x_1 \ldots.. x_{2MN}$ such that

$$\sum_{i=1}^{2MN}\left| f(x_i)-f(x_{i-1})\right|<\frac{\varepsilon}{4} \qquad \text{..........(**).}$$

$$\text{Let } t_i = \frac{(x_i + x_{i+1})}{2} \quad \text{for} \quad i=0,\ 1,\ 2,\ldots\ldots\ldots\ldots, MN-1.$$

$$\text{Define } g(x) = f(M) + \sum_{i=1}^{2MN}(f(x_i)-f(x_{i-1}))\sigma(k(x-t_{i-1})$$

Choose W>0 such that if u>W then

$$\left|\sigma(u)-1\right|<\frac{1}{MN}; \qquad \text{.........(***)}$$

Let K>0 such that

$$K\frac{1}{2N}>W.$$

To prove:

$$\left|f(x)-g(x)\right|<\varepsilon\cdot$$

If x>M then.

$$\left|f(x)-f(M)\right|<\frac{\varepsilon}{2}$$

Now it is easy to see that

$$\left|g(x)-f(M)\right|\leq\sum_{i=1}^{2MN}\left|f(x_i)-f(x_{i-1})\right|\left|\sigma(K(x-t_{i-1})-1)\right|$$

$$+\sum_{i=1}^{2MN}\left| f(x_i)-f(x_{i-1})\right|$$

$$<\frac{\varepsilon}{2}.$$

by using (*), (**), and (***)

$$|f(x)-g(x)| \leq |f(x)-f(M)| + |f(M)-g(x)|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$$

= ϵ for all x $\in$ (0,∞). This proves the Theorem..

LINEAR SUBSPACE

Let $L \subseteq C(\bar{R})$, be the linear subspace of $C(\bar{R})$, where

$$L = \{ f\, C(),: \lim_{x \to -\infty} f(x) = 0 \}.$$

Definition 4.1: g(w, b): $R \to R$ by $g(w, b)(x) = g(wx + b)$ and

$$F(g) = \{g(w, b): (w, b)\, R^2\}.$$

Theorem 4.2: For any g L, F(g) $\subset$ L

Proof: Let m F (g) therefore m = g (w, b), (w, b) R^2

$$\lim_{x \to -\infty} m(x) = \lim_{x \to -\infty} g(w,b)(x)$$

$$= \lim_{x \to -\infty} g(wx+b) = 0 \text{ that proves the Theorem.}$$

$$H = \{f \in C_{(\bar{R})},\ \lim_{x \to +\infty} f(x) = 1\}$$

is the linear subspace of $C_{(\bar{R})}$.

Theorem 4.3: F (g) $\subset$ H, for all g H.

Proof: Analogous to Theorem 4.2.

FUNCTION APPROXIMATION IN ARBITRARY FUNCTIONS

In this section, we use left and right sigmoidal to approximate the arbitrary functions.

Theorem 5.1: Given a bounded non-negative function g(x) in R such that

$$\lim_{x \to -\infty} g(x)$$

exists. Then for any arbitrary mapping f(x) in $C^{(\bar{R})}$, for any ϵ>0 there exist scalars w, b_i and β_i such that

$$\left| \sum_{i=1}^{k} \beta_i g(w_i x + b_i) - f(x) \right| < \varepsilon \text{ for all } x \in R.$$

Proof. Since g(x) is a bounded non negative function, there exists a λ such that.

$$g(x) \leq \lambda \text{ for all } x \in R.$$

This gives that $\frac{g(x)}{\lambda} < 1$ for all x$\in$R.

Let $g_1(x) = \frac{g(x)}{\lambda}$ and therefore $g_1(x) < 1$. This gives that $g_1(x)$ is a sigmoidal function.

Case 1:

$\lim_{x \to -\infty} g(x) = 0$.Therefore $\lim_{x \to -\infty} g_1(x) = 0$ and $g_1(x)$ is a left sigmoidal function.

By Theorem 3.1, we get.

$$\left| \sum_{i=1}^{k} \beta_i^1 g_1(w_i x + b_i) - f(x) \right| < \varepsilon$$

$$\left| \sum_{i=1}^{k} \beta_i^1 \frac{1}{?} g(w_i x + b_i) - f(x) \right| < e \text{ for all } x \in R$$

$$\left| \sum_{i=1}^{k} \beta_i g(w_i x + b_i) - f(x) \right| < e \text{ for all } x \in R. \text{ where } \beta_i = \frac{\beta_i^1}{?}$$

Case 2:

$$\lim_{x \to -\infty} g(x) = N,\ N \neq 0.$$

It is easy to see that $g_2(x) = g(x)-N/\lambda$ is a left sigmoidal function. Then by Theorem 3.1, we get

$$\left| \sum_{i=1}^{k} \beta i^1 g_2(w_i x + b_i) - f(x) \right| < e \text{ for all } x \in R,$$

$$\text{where } k-1 \in N, i = 1, \ldots, k-1$$

Suppose x_0 R and $g(x_0) \neq 0$, then we can choose $w_k = 0$ and $b_k = x_0$ such that

$$g(w_k x + b_k) = g(x_0)$$

Let

$$\beta_k\, g(x_0) = -\sum_{i=1}^{k-1} N \beta_k.$$

$$\left| \sum_{i=1}^{k} \beta_i g(w_i x + b_i) - f(x) \right| = \left| \sum_{i=1}^{k-1} \beta_i g(w_i x + b_i) + \beta_k g(x_0) - f(x) \right|$$

$$= \left|\sum_{i=1}^{k-1}\beta_i\lambda g_1(w_ix+b_i)+\beta_k g(x_0)-f(x)\right|$$

$$= \left|\sum_{i=1}^{k-1}\lambda\beta_i(g_2(w_ix+b_i)+\frac{N}{\lambda})+\beta_k g(x_0)-f(x)\right|$$

$$= \left|\sum_{i=1}^{k-1}\lambda\beta_i\ g_2(w_ix+b_i)+\sum_{i=1}^{k-1}N\beta_i+\beta_k g(x_0)-f(x)\right|$$

$$= \sum_{i=1}^{k-1}\beta_i^1 g_2(w_ix+b_i)-f(x)\Big| <\in \text{ for all x R,}$$

where

$$\beta_i^1 = \lambda\beta_i$$

that completes the proof.

Theorem 5.2: Given a bounded non-negative function g(x) in R such that

$$\lim_{x\to+\infty} g(x)$$

exists and not equal to zero, then for any arbitrary mapping f (x) in

$$C(\overline{R}),$$

Proof: We assume g (x) to be a sigmoidal function (otherwise replace g(x) by

$$\frac{g}{|g|})$$

Case 1:

$\lim_{x\to+\infty} g(x) = 1$. By Theorem 3.2, we get

Case 2:

$$\lim_{x\to+\infty} g(x) = \lambda \neq 1.$$

Let

$$g_1(x) = \frac{g(x)}{\lambda}.$$

Then by Theorem 3.2, we get

$$\left|\sum_{i=1}^{k}\beta_i^1 g_1(w_ix+b_i)-f(x)\right| < \varepsilon\ \forall x\in R.$$

Since $g_1 = \frac{g}{\lambda}$

$$\left|\sum_{i=1}^{k}\beta_i^1\frac{1}{\lambda}g(w_ix+b_i)-f(x)\right| < \varepsilon$$

that implies $\left|\sum_{i=1}^{k}\beta_i g(w_ix+b_i)-f(x)\right| < \varepsilon\ \forall x\in R$

where.

$$\beta_i = \frac{\beta_i^1}{\lambda}$$

The following Lemma is due to Cuang[2].

Lemma 5.3: If g(x) in R satisfies that all linear combinations

$$\sum_{i=1}^{k}\beta_i g(w_ix+b_i) \text{ are dense in } C(\overline{R}),$$

then all linear combinations

$$\sum_{i=1}^{k}\beta_i g(w_ix+b_i)$$

are dense in

$$C(\overline{R^n}),$$

where w_i and w_ix is the inner product of w_i and x.

Theorem 5.4: Given any bounded non-negative function g (x) in R such that

$$\lim_{x\to-\infty} g(x)$$

exists, then all linear combinations

$$\sum_{i=1}^{k}\beta_i g(w_ix+b_i) \text{ are dense in } C(\overline{R^n}) \text{ where } w_i\in \overline{R^n}$$

and wx is the inner product of w and x.

Proof: Follows from Lemma 5.3 and Theorem 5.1.

Theorem 5.5: Given any bounded non-negative function g (x) in R such that

$$\lim_{x\to+\infty} g(x)$$

exists and is not equal to 1. Then all linear combinations

$$\sum_{i=1}^{k}\beta_i\, g(w_i x + b_i) \text{ are dense in } C(\overline{R}^n) \text{ where } w_i \in \overline{R}^n$$

and $w_i x$ is the inner product of w_i and x.

Proof: Follows from Lemma 5.3 and Theorem 5.2.

CONCLUSIONS

Thus we have proved that arbitary continuous real functions with finite limits and (-∞, 0) and (0, ∞) can be approximated by bounded left sigmoidal and right sigmoidal signals respectively. We also approximated the arbitary real function with left and right sigmoidal signals. We further use the left and right sigmoidal signals to prove the approximation theorem due to Chen.

REFERENCES

1. Chen, T.H. Chen and R.W. Liu, 1995. 'Approximation Capability in $C(\overline{R}^n)$ by Multilayer Feedforward Networks and Related Problems', IEEE Transactions on Neural Network, 6: 25-29.
2. Huang, G.B. and H.A. Babri, 1998. Comments in Approximation Capability in $C(\overline{R}^n)$ by Multiplayer Feedforward Networks and Related Problems'. IEEE Transactions on Neural Networks, 9: 714-715.
3. Anonymous, 1998. 'Upper Bounds on the Number of Hidden Neurons in Feedforward Networks with Arbitrary Bounded Nonlinear Activation Functions', IEEE Transactions on Neural Networks, 9: 224-229.
4. Kalmogorov, A.N., 1963. On the representations of continuous functions of many variables by superposition of continuous functions of one variable and addition. Doklady Akademy Nauk SSSR, 114 (1957), 953-956. English Translation: Amer. Math. Soc. Transl., 2: 55-59.
5. Funahashi, K.I., 1989. On the Approximate Realization of Continuous Mapping by Neural Networks, Neural Networks, 2: 183-192.
6. Hornik. K., M. Stinchcombe and H. White, 1989. Multi-layer feed forward networks are Universal approximators, Neural Networks, 2: 359-366.
7. Cybenko, G., 1989. Approximation by superpositions of a sigmoidal functions. Math. Contr. Signals. Syst., 2: 303-314.
8. Verakurkov, A., 1992. Kolmogorv's Theorem a Multilayer Neural Networks, Neural Networks, 5: 501-506.
9. Ito, Y., 1992. Approximation of continuous functions in Rd by linear combinations of shifted Rotations of a sigmoidal function with and without scaling. Neural Networks, 5: 105-115.