关于插值神经网络的构造性 *

谢庭藩 曹飞龙 **

中国计量学院信息与数学科学系，杭州 310018

摘要 神经网络插值问题是神经网络理论与应用的研究热点与难点之一. 文中研究具有插值性质的前向神经网络的构造与逼近问题. 对于一般的 Sigmoidal 激活函数和 d 维 Euclid 空间中的插值样本，分别构造了精确插值和近似插值的单隐层前向神经网络，研究这两类网络之间的偏差，并分别估计它们对目标函数的逼近误差，指出神经网络插值与一般代数多项式插值之间的本质差异.

关键词 神经网络 精确插值 近似插值 偏差估计

设$\mathbb{R}^d$是 d 维 Euclid 空间，$\left\{X_j\right\}_{j=0}^n$是$\mathbb{R}^d$中的 $n+1$个互异的点，又设$\left\{f_j\right\}_{j=0}^n$是 $n+1$ 个实数，我们称

$$(X_0, f_0), (X_1, f_1), \cdots, (X_n, f_n) \tag{1}$$

为插值样本，$\left\{X_j\right\}_{j=0}^n$为插值结点组. 如果有一个单隐层前向神经网络

$$N_e(X) = \sum_{j=0}^n c_j \phi(W_j \circ X + b_j)$$

满足条件

$$N_e(X_j) = f_j \quad (j = 0, 1, \cdots, n),$$

则说 $N_e(X)$是样本(1)的精确插值网络，这里$W \circ X$表示向量W与X的内积.

一个$\mathbb{R}$上定义的有界函数 σ 被称为 Sigmoidal 函数，如果它满足条件

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

神经网络插值问题是神经网络理论与应用的研究热点与难点之一，至今已有较多研究. 对于激活函数$\phi(x)$为对数 Sigmoidal 函数

$$S(x) = \frac{1}{1+e^{-x}} \tag{2}$$

的情形，精确插值网络存在性的代数证明早已为人所知. 文献[1] 证明了当激活函数ϕ是非减的 Sigmoidal 函数时，精确插值网络是存在的. 文献[2—5] 也研究过精确插值网络的存在性问题. 然而，要具体构造精确插值网络就会出现$(n+1)\times(n+1)$阶矩阵的求逆问题，随着插值节点数 n 的增大，计算量也迅速增大，这是一项非常麻烦的工作. 于是，人们转向寻求近似插值网络的研究，即构造一个网络使之与精确插值网络有任意小的偏差，并由它代替精确插值网络逼近目标函数，文献[6] 曾讨论过这个问题. 最近文献[7] 在激活函数为非减的 Sigmoidal 函数的条件下，给出精确插值网络存在性的代数证明，并且对任意正数 ε，构造了 ε 近似插值网络 $N_a(X)$，即适合条件

$$|N_a(X_j) - f_j| < \varepsilon, \ j = 0, 1, \cdots, n$$

2007-07-24 收稿，2007-09-21 收修改稿

* 国家自然科学基金(批准号：60473034)和浙江省教育厅科研重点基金(批准号：20060543)资助项目

** 通信作者，E-mail：flcao@263.net

的单隐层前向神经网络. 特别地, 他们还就激活函数为对数 Sigmoidal 函数(2), 利用 $S(x)$的可微性构造了在 $C[a, b]$(定义在$[a, b]$上的连续函数空间)中稠密的单隐层前向神经网络集.

如所周知, Sigmoidal 函数是常用的激活函数, 人们在讨论神经网络的稠密性与复杂性时就经常使用一般的 Sigmoidal 函数作为神经网络的激活函数[8—15]. 因此, 自然要问: 在讨论插值神经网络时, 精确网络与近似网络的激活函数可否由一般的 Sigmoidal 函数构成呢? 这是本文将予以肯定回答的第一个问题. 此外, 已有的关于网络插值的研究大多仅停留在存在性的研究上, 很少研究插值网络与近似插值网络对目标函数的逼近误差, 而从应用的角度看, 这又恰是十分重要的问题. 基于这一目的, 我们利用作为刻画逼近阶和描述函数光滑性的重要工具——连续模, 分别建立插值网络和近似网络对目标函数的逼近误差. 同时, 我们将指出神经网络插值与代数多项式插值之间的本质差异. 我们将证明对于 Sigmoidal 激活函数及样本(1), 精确插值网络是存在的, 并且针对一维样本和 Sigmoidal 激活函数构造了近似插值网络, 建立近似插值网络与精确插值网络间的偏差估计. 此外, 我们将分别给出近似插值网络与精确插值网络逼近目标函数的误差估计, 进一步考虑如何将结论推进到多维空间.

1 精确插值网络的存在性

设 $\sigma(x)$是 Sigmoidal 函数, $A>0$, 记

$$\delta(A)=\sup_{x\geqslant A}\max(|1-\sigma(x)|, |\sigma(-x)|).$$

显然 $\delta(A)$是不增函数, 而且

$$\lim_{A\to+\infty}\delta(A)=0.$$

为方便, 本节先讨论一维的情况. 设有实数组 $\{f_j\}_{j=0}^{n}$ 及节点组

$$x_0<x_1<\cdots<x_n. \tag{3}$$

对于样本

$$(x_0, f_0), (x_1, f_1), \cdots, (x_n, f_n) \tag{4}$$

的精确插值网络, 我们先建立如下的

定理 1 设 $\sigma(x)$是 Sigmoidal 函数, 则当

$$\delta(A)<\frac{1}{4n} \tag{5}$$

时, 存在$\{c_j\}_{j=0}^{n}\subset\mathbf{R}$, 使得神经网络

$$N_e(x, A)=\sum_{j=0}^{n-1}c_j\sigma\left(-\frac{2A(x-x_j)}{x_{j+1}-x_j}+A\right)+c_n\sigma\left(-\frac{2A(x-x_n)}{x_{n+1}-x_n}+A\right) \tag{6}$$

是样本(4)的精确插值.

证明 显然, 我们只要证明在条件(5)下关于未知数$\{c_j\}_{j=0}^{n}$的线性方程组

$$N_e(x_i, A)=f_i \quad (i=0, 1, \cdots, n) \tag{7}$$

是有解的. 为此, 简记

$$e_{ij}(A)=\sigma\left(-\frac{2A(x_i-x_j)}{x_{j+1}-x_j}+A\right) \quad (i, j=0, 1, \cdots, n-1),$$
$$e_{in}(A)=\sigma\left(-\frac{2A(x_i-x_n)}{x_n-x_{n-1}}+A\right) \quad (i=0, 1, \cdots, n-1),$$
$$e_{nj}(A)=\sigma\left(-\frac{2A(x_n-x_j)}{x_{j+1}-x_j}+A\right) \quad (j=0, 1, \cdots, n-1),$$
$$e_{nn}(A)=\sigma(A).$$

则由(6)式看到, 方程组(7)的系数行列式为

$$\boldsymbol{D}_n(A)=\begin{vmatrix} e_{00}(A) & e_{01}(A) & \cdots & e_{0n}(A)\\ e_{10}(A) & e_{11}(A) & \cdots & e_{1n}(A)\\ \cdots & \cdots & \cdots & \cdots\\ e_{n0}(A) & e_{n1}(A) & \cdots & e_{nn}(A)\end{vmatrix}.$$

又记

$$d_{ij}(A)=e_{ij}(A)-e_{i+1j}(A) \quad (i, j=0, 1, \cdots, n-1),$$
$$d_{in}(A)=e_{in}(A)-e_{i+1n}(A) \quad (i=0, 1, \cdots, n-1),$$
$$d_{nj}(A)=e_{nj}(A) \quad (j=0, 1, \cdots, n-1),$$
$$d_{nn}(A)=e_{nn}(A).$$

则由行列式的性质知道

$$\boldsymbol{D}_n(A)=\begin{vmatrix} d_{00}(A) & d_{01}(A) & \cdots & d_{0n}(A) \\ d_{10}(A) & d_{11}(A) & \cdots & d_{1n}(A) \\ \cdots & \cdots & \cdots & \cdots \\ d_{n0}(A) & d_{n1}(A) & \cdots & d_{nn}(A) \end{vmatrix}.$$

注意到 $\delta(A)$的定义，不难看出：条件(5)含有当$t\geqslant A$时，成立

$$|\sigma(-t)|<\frac{1}{4n},\quad |1-\sigma(t)|<\frac{1}{4n}.$$

于是，由 $d_{ij}(A)$的定义得到

$$d_{ii}(A)=\sigma(A)-\sigma(-A)=1-(1-\sigma(A))-\sigma(-A)\geqslant 1-\frac{1}{4n}-\frac{1}{4n}=1-\frac{1}{2n}(0\leqslant i\leqslant n-1),$$

$$d_{nn}(A)=\sigma(A)=1-(1-\sigma(A))\geqslant 1-\frac{1}{4n}\geqslant 1-\frac{1}{2n},$$

以及

$$\sum_{j=0,j\neq i}^{n}|d_{ij}|<n\frac{1}{4n}=\frac{1}{4}\quad(i=0,1,\cdots,n).$$

所以，

$$d_{ii}(A)>\sum_{j=0,j\neq i}^{n}|d_{ij}|\quad(i=0,1,\cdots,n).$$

因此，由行列式的对角线严格占优定理[16]推出

$$\boldsymbol{D}_n(A)\neq 0.$$

换言之，方程组(7)是有解的. 定理1证毕.

附注1 定理1证明中所出现的函数

$$\sigma\left(-\frac{2A(x-x_j)}{x_{j+1}-x_j}+A\right)$$

中的系数2是为了使计算简便. 事实上，取任一大于1的正整数不影响定理的结论，只是此时条件(5)应作相应的变动.

附注2 定理1所构造的插值网络的输入权与阈值仅与样本(4)的结点组(3)有关，而输出权$\{c_j\}_{j=0}^{n}$是一个线性方程组的解. 我们证明了这个解在 A 适当大时的存在性，但要具体算出其值需要解方程组，还是比较麻烦的.

附注3 文献[7]在激活函数 $\sigma(x)$为非减的Sigmoidal函数的条件下，证明了精确插值网络的存在性，这里，我们去掉了“非减”的要求，并且证明要比文献[7]简单.

2 插值网络对目标函数的逼近

对于样本(4)，定义

$$N_a(x,A)=\sum_{j=0}^{n-1}(f_j-f_{j+1})\sigma\left(-\frac{2A(x-x_j)}{x_{j+1}-x_j}+A\right)+f_n\sigma\left(-\frac{2A(x-x_n)}{x_n-x_{n-1}}+A\right)$$

作为此样本的近似插值网络. 下面考虑关于样本(4)的近似插值网络 $N_a(x,A)$与精确插值网络 $N_e(x,A)$(即(6))之间的偏差，我们有

定理2 设 $\sigma(x)$是Sigmoidal函数，则当$\delta(A)<\frac{1}{4n}$时，对于样本(4)成立着不等式

$$|N_e(x,A)-N_a(x,A)|\leqslant\frac{(2n+1)\delta(A)\|\sigma\|}{1-(2n+1)\delta(A)}\cdot\left(\sum_{j=0}^{n}|f_j-f_{j+1}|+|f_n|\right),$$

其中

$$\|\sigma\|=\sup_x|\sigma(x)|,\quad i=0,1,\cdots,n.$$

证明 因为 $\delta(A)<\frac{1}{4n}$，所以由定理1知道线性方程组(7)的解是存在的，并记其为$\boldsymbol{C}=(c_0,c_1,\cdots,c_n)$，又记方程组(7)的系数矩阵为$\boldsymbol{M}$，以及$\boldsymbol{f}=(f_0,f_1,\cdots,f_n)$，则方程组(7)可写作

$$\boldsymbol{M}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{f}^{\mathrm{T}}\tag{8}$$

这里$\boldsymbol{C}^{\mathrm{T}}$表示向量$\boldsymbol{C}$的转置. 又定义矩阵

$$U=\begin{bmatrix}1&1&\cdots&1&1\\0&1&\cdots&1&1\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&1&1\\0&0&\cdots&0&1\end{bmatrix},$$

则 U 之逆矩阵

$$U^{-1}=\begin{bmatrix}1&-1&0&\cdots&0&0\\0&1&-1&\cdots&0&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1&-1\\0&0&0&\cdots&0&1\end{bmatrix}.$$

容易看到，若记

$$M-U=(\alpha_{ij})_{i,j=0}^{n,n},$$

则

$$|\alpha_{ij}|\leqslant\delta(A).$$

因此，如记

$$U^{-1}(M-U)=(\beta_{ij})_{i,j=0}^{n,n},$$

那么

$$|\beta_{ij}|\leqslant 2\delta(A),\quad |\beta_{nj}|\leqslant\delta(A)$$
$$(i=0,1,\cdots,n-1,j=0,1,\cdots,n)\tag{9}$$

又记 $F=(f_0-f_1,\ f_1-f_2,\ \cdots,\ f_{n-1}-f_n,\ f_n)$，$\Delta C=C-F$，则有

$$UF^{\mathrm{T}}=f^{\mathrm{T}}.$$

于是(8)式含有

$$(U+(M-U))(F^{\mathrm{T}}+\Delta C^{\mathrm{T}})=f^{\mathrm{T}}.$$

由此不难看到

$$U\Delta C^{\mathrm{T}}=-(M-U)\Delta C^{\mathrm{T}}-(M-U)F^{\mathrm{T}},$$

或者说

$$\Delta C^{\mathrm{T}}=-U^{-1}(M-U)\Delta C^{\mathrm{T}}-U^{-1}(M-U)F^{\mathrm{T}}.$$

上式及(9)式直接算得

$$\sum_{i=0}^{n}|\Delta C_i|\leqslant(2n+1)\delta(A)\sum_{i=0}^{n}|\Delta C_i|+(2n+1)\delta(A)\left(\sum_{i=0}^{n-1}|f_i-f_{i+1}|+|f_n|\right).$$

于是留心到 $\delta(A)<\frac{1}{4n}$，即有

$$\sum_{i=0}^{n}|\Delta C_i|\leqslant\frac{(2n+1)\delta(A)}{1-(2n+1)\delta(A)}\cdot\left(\sum_{i=0}^{n}|f_i-f_{i+1}|+|f_n|\right)\tag{10}$$

显然

$$|N_e(x,A)-N_a(x,A)|\leqslant\sum_{i=0}^{n}|\Delta C_i|\ \|\sigma\|,$$

故(10)式含有

$$|N_e(x,A)-N_a(x,A)|\leqslant\frac{(2n+1)\delta(A)\|\sigma\|}{1-(2n+1)\delta(A)}\cdot\left(\sum_{i=0}^{n}|f_i-f_{i+1}|+|f_n|\right).$$

定理2证明完毕.

下面讨论插值网络对目标函数的逼近. 为此，我们需要连续模的概念. 设 $f(x)$是$[a,b]$上有定义的目标函数，定义

$$\omega(f,\delta)=\sup_{a\leqslant x,\,x+h\leqslant b,\,|h|\leqslant\delta}|f(x)-f(x+h)|$$

为 f 的连续模[17]. 连续模是刻画逼近误差的一个重要尺度，同时也是描述函数光滑性的重要工具.

对于$[a,b]$上有定义的目标函数以及区间$[a,b]$的一个分划

$$a=x_0<x_1<\cdots<x_n=b\tag{11}$$

记样本$\left\{(x_i,f(x_i))\right\}_{i=0}^{n}$的精确插值网络与近似插值网络依次为 $N_e(x,f,A)$与 $N_a(x,f,A)$，下面讨论 $N_a(x,f,A)$与 $f(x)$的偏差.

定理 3 设 $\sigma(x)$ 是 Sigmoidal 函数, $f(x)$ 是 $[a, b]$ 上定义的目标函数, $A>0$, 则在 $[a, b]$ 上有

$$|N_a(x,f,A)-f(x)| \leqslant \delta(A) \cdot \left(\sum_{j=0}^{n-1}|f(x_j)-f(x_{j+1})|+f(b)\right)+\left(\|\sigma\|+1\right)\omega(f,\Delta_n),$$

这里

$$\Delta_n=\max_{0\leqslant j\leqslant n}|x_i-x_{j+1}|.$$

证明 对 $x\in[a, b]$, 存在 $j_0\in\mathrm{N}$, 使得 $x\in[x_{j_0}, x_{j_0+1}]$. 显然

$$\begin{aligned}N_a(x,f,A)-f(x)=&\sum_{j=0}^{j_0-1}(f(x_j)-f(x_{j+1}))\cdot \sigma\left(-\frac{2A(x-x_j)}{x_{j+1}-x_j}+A\right)+\\&\sum_{j=j_0+1}^{n-1}(f(x_j)-f(x_{j+1}))\cdot \left(\sigma\left(-\frac{2A(x-x_j)}{x_{j+1}-x_j}+A\right)-1\right)+\\&f(b)\left(\sigma\left(-\frac{2A(x-x_n)}{x_n-x_{n-1}}+A\right)-1\right)+\\&f(x_{j_0+1})-f(x)+\\&(f(x_{j_0})-f(x_{j_0+1}))\cdot \\&\sigma\left(-\frac{2A(x-x_{j_0})}{x_{j_0+1}-x_{j_0}}+A\right).\end{aligned}$$

因此, 留心到 $\delta(A)$的定义, 就有

$$|N_a(x,f,A)-f(x)|\leqslant \delta(A)\cdot \left(\sum_{j=0}^{n-1}|f(x_j)-f(x_{j+1})|+f(b)\right)+(\|\sigma\|+1)\omega(f,\Delta_n).$$

此即为所求. 定理 3 证毕.

联合定理 2 和定理 3, 注意到 $\|\sigma\|\geqslant 1$, 即得

定理 4 设 $\sigma(x)$ 是 Sigmoidal 函数, $f(x)$ 是 $[a, b]$ 上定义的目标函数, 则当 $\delta(A)<\frac{1}{4n}$时, 在 $[a, b]$ 上有

$$|N_e(x,f,A)-f(x)|\leqslant \frac{(2n+2)\delta(A)\|\sigma\|}{1-(2n+1)\delta(A)}\cdot \left(\sum_{i=0}^{n-1}|f(x_i)-f(x_{i+1})|+f(b)\right)+(\|\sigma\|+1)\omega(f,\Delta_n).$$

特别当分划(11)是$[a, b]$ 的n 等分时, 有

$$\Delta_n=\frac{b-a}{n},\ |f(x_i)-f(x_{i+1})|\leqslant \omega\left(f,\frac{b-a}{n}\right)$$
$$(i=0,1,\cdots,n-1).$$

于是, 我们得到如下的

定理 5 设 $\sigma(x)$ 是 Sigmoidal 函数, $f(x)$ 是 $[a, b]$ 上的目标函数, 则当插值结点$\{x_i\}_{i=0}^{n}$ 为 $[a, b]$ 的 n 等分的分点且 $\delta(A)<\frac{1}{(n+1)^2}$时, 在 $[a, b]$ 上成立着

$$|N_a(x,f,A)-f(x)|\leqslant (\|\sigma\|+2)\omega\left(f,\frac{b-a}{n}\right)+\frac{1}{n+1}|f(b)|,$$

$$|N_e(x,f,A)-f(x)|\leqslant (5\|\sigma\|+1)\omega\left(f,\frac{b-a}{n}\right)+\frac{4\|\sigma\|}{n+1}|f(b)|.$$

附注 4 当插值结点组$\{x_i\}_{i=0}^{n}$为$[a, b]$ 的 n 等分的分点时, 由 $N_e(x, f, A)$与 $N_a(x, f, A)$的定义看到, 这两个网络的输入权是$-\frac{2An}{b-a}$. 因此, 定理 5 给出了具固定权的神经网络的逼近定理[18, 1)].

附注 5 当激活函数为对数 Sigmoidal 函数(2)时, 文献[7] 利用了激活函数的微分性质得到 $N_a(x, f, A)$在连续函数空间中的稠密性定理. 本文的定理 3—5 是在激活函数为一般的 Sigmoidal 函数条件下得到逼近阶估计定理(复杂性定理), 它们自然包含着近似插值网络与精确插值网络在连续的目标函数空间的稠密性定理.

附注 6 定理 5 表明在均匀分布的结点的情况

1) Xie T.F, Cao F.L. The error of approximation of neural networks with fixed weights. (in press)

下，神经网络插值与通常的代数多项式插值的本质差异，因为Lagrange插值多项式不可能对每个连续函数都收敛的，而插值的神经网络对每个连续函数都是一致收敛的.

附注7 我们看到近似插值网络 $N_a(x, f, A)$ 关于 f 是线性的. 于是，在定理5的假设下，我们有如下的不等式：在 $[a, b]$ 上成立着

$$|f(b)+N_a(x, f-f(b), A)-f(x)|\leqslant (\|\sigma\|+2)\omega\left(f, \frac{b-a}{n}\right),$$

$$|f(b)+N_e(x, f-f(b), A)-f(x)|\leqslant (5\|\sigma\|+1)\omega\left(f, \frac{b-a}{n}\right).$$

3 多维空间的插值网络

本节对 $d>1$ 讨论 $\mathbb{R}^d$ 中的神经网络的插值问题. 由于样本(1)中的插值节点组 $\left\{X_j\right\}_{j=0}^{n}$ 是 $\mathbb{R}^d$ 中 $(n+1)$ 个互异的点，所以，根据文献[3]的一个结论，存在点 $\boldsymbol{y}\in\mathbb{R}^d$ 使得

$$\boldsymbol{y}\circ\boldsymbol{X}_j \quad (j=0, 1, \cdots, n)$$

是互异的实数. 将它们按大小排列如下

$$\boldsymbol{y}\circ\boldsymbol{X}_{m_0}<\boldsymbol{y}\circ\boldsymbol{X}_{m_1}<\cdots<\boldsymbol{y}\circ\boldsymbol{X}_{m_n},$$

这里 $(m_0, m_1, \cdots, m_n)$ 是 $(0, 1, \cdots, n)$ 的一个置换. 记 $t_i=\boldsymbol{y}\circ\boldsymbol{X}_{m_i}(i=0, 1, \cdots, n)$，则有

$$t_0<t_1<\cdots<t_n.$$

我们对样本

$$(t_0, f_{m_0}), (t_1, f_{m_1}), \cdots, (t_n, f_{m_n}) \qquad (12)$$

应用定理1，立即看到当 $\delta(A)<\frac{1}{4n}$ 时，存在 $\left\{c_j\right\}_{j=0}^{n}$，使得

$$N_e(t, A)=\sum_{j=0}^{n-1}c_j\sigma\left(-\frac{2A(t-t_j)}{t_{j+1}-t_j}+A\right)+c_n\sigma\left(-\frac{2A(t-t_n)}{t_n-t_{n-1}}+A\right)$$

是样本(12)的精确插值，即有

$$N_e(t_i, A)=f_{m_i} \quad (i=0, 1, \cdots, n).$$

于是网络

$$N_e(\boldsymbol{X}, A)=\sum_{j=0}^{n-1}c_j\sigma\left(-\frac{2A(\boldsymbol{y}\circ\boldsymbol{X}-\boldsymbol{y}\circ\boldsymbol{X}_{m_j})}{\boldsymbol{y}\circ\boldsymbol{X}_{m_{j+1}}-\boldsymbol{y}\circ\boldsymbol{X}_{m_j}}+A\right)+c_n\sigma\left(-\frac{2A(\boldsymbol{y}\circ\boldsymbol{X}-\boldsymbol{y}\circ\boldsymbol{X}_{m_n})}{\boldsymbol{y}\circ\boldsymbol{X}_{m_n}-\boldsymbol{y}\circ\boldsymbol{X}_{m_{n-1}}}+A\right) \qquad (13)$$

是样本(1)的精确插值网络. 因为此时显然有

$$N_e(\boldsymbol{X}_{m_i}, A)=f_{m_i} \quad (i=0, 1, \cdots, n)$$

换言之，有

$$N_e(\boldsymbol{X}_i, A)=f_i \quad (i=0, 1, \cdots, n)$$

我们将上述结论写作

定理6 设 $\sigma(x)$ 是Sigmoidal函数，(1)是 $\mathbb{R}^d$ 中的插值样本，则有 $\boldsymbol{y}\in\mathbb{R}^d$ 及 $(0, 1, \cdots, n)$ 的一个置换 $(m_0, m_1, \cdots, m_n)$，使得

$$\boldsymbol{y}\circ m_0<\boldsymbol{y}\circ m_1<\cdots<\boldsymbol{y}\circ m_n,$$

而且当

$$\delta(A)<\frac{1}{4n}$$

时，存在实数组 $\left\{c_j\right\}_{j=0}^{n}$，使得(13)是样本(1)的精确插值网络.

类似于第三节的讨论，如果定义样本(1)的近似插值网络为

$$N_a(\boldsymbol{X}, A)=\sum_{j=0}^{n-1}(f_{m_j}-f_{m_{j+1}})\circ\sigma\left(-\frac{2A(\boldsymbol{y}\circ\boldsymbol{X}-\boldsymbol{y}\circ\boldsymbol{X}_{m_j})}{\boldsymbol{y}\circ\boldsymbol{X}_{m_{j+1}}-\boldsymbol{y}\circ\boldsymbol{X}_{m_j}}+A\right)+f_{m_n}\sigma\left(-\frac{2A(\boldsymbol{y}\circ\boldsymbol{X}-\boldsymbol{y}\circ\boldsymbol{X}_{m_n})}{\boldsymbol{y}\circ\boldsymbol{X}_{m_{n-1}}-\boldsymbol{y}\circ\boldsymbol{X}_{m_n}}+A\right),$$

那么成立如下的

定理7 设 $\sigma(x)$ 是Sigmoidal函数，则当

$\eth(A) < \frac{1}{4n}$时，成立

$$|N_e(X, A) - N_a(X, A)| \leqslant \frac{(2n+1)\eth(A)\|\sigma\|}{1-(2n+1)\eth(A)} \cdot \left(\sum_{j=0}^{n} |f_{m_j} - f_{m_{j+1}}| + |f_{m_n}|\right).$$

附注 **8** 自然会发问：当样本(1)中的 f_j 取作函数$f(X)$在 X_j 处的值 $f(X_j)$时，相应的插值网络与$f(X)$的偏差该有怎样的刻划？这个问题的回答并不简单，我们将在后续文章中继续讨论.

4 结论

文本首先在激活函数仅为有界的Sigmodial 函数的条件下，证明了精确插值神经网络的存在性. 与已有结果相比，定理所需的条件较为宽松，而采用的证明方法比较简洁并且是构造性的，为构造精确插值网络提供了一种新方法. 其次，给出了近似插值网络与精确插值网络之间以及近似插值网络逼近目标函数偏差的量化估计，从而揭示了网络逼近速度与隐层拓扑结构(如神经元数目)之间的关系. 同时，也指出了神经网络插值与代数插值之间的本质区别以及需进一步研究的问题.

参 考 文 献

1 Ito Y，Saito K. Superposition of linearly independent functions and finite mappings by neural networks. Math Sci，1996，21(1)：27—33

2 Pinkus A. Approximation theory by superposition of sigmoidal and radial basis functions. Adv Appl Math，1992，13(3)：350—373

3 Huang GB，Babri HA. Feedforward neural networks with arbitrary bounded nonlinear activation functions. IEEE Trans Neural Networks，1998，9(1)：224—229

4 Sartori MA，Antsaklis PJ. A simple method to derive bounds on the size and to train multilayer neural networks. IEEE Trans Neural Networks，1991，2(4)：467—471

5 Tamura S，Tateishi M. Capabilities of a four—layered feedforward neural network. IEEE Trans Neural Networks，1997，8(2)：251—255

6 Sontag ED. Feedforward nets for interpolation and classification. J Comp Syst Sci，1992，45(1)：20—48

7 Llanas B，Sainz FJ. Constructive approximate interpolation by neural networks. J Comput Appl Math，2006，188(2)：283—308

8 Chen TP，Chen H，Liu RW. A constructive proof and extension of Cybenko's approximation theorem. In：Computing Science and Statistics. Proceedings of the 22nd Symposium on the Interface. Berlin：Springer，1991，163—168

9 Chen TP，Chen H. Approximation capability to functions of several variables，nonlinear functionals，and operators by radial basis function neural networks. IEEE Trans Neural Networks，1995，6(5)：904—910

10 Chen TP，Chen H，Liu WR. Approximation capability in C(R^n) by multilayer feedforward networks and related problems. IEEE Trans Neural Networks，1995，6(1)：25—30

11 Cybenko G. Approximation by superposition of sigmoidal functions. Mathematics of Control，Signal and Systems，1989，2(2)：303—314

12 Funahashi K. On the approximate realization of continuous mapping by neural networks. Neural Networks，1989，2(1)：183—192

13 陈天平. 神经网络及其在系统识别应用中的逼近问题. 中国科学，A辑，1994，24(1)：1—7

14 曹飞龙，徐宗本. 神经网络的本质逼近阶. 中国科学，E辑，2004，34(4)：361—373

15 Xu ZB，Cao FL. Simultaneous L^p-approximation order for neural networks. Neural Networks，2005，18(11)：914—923

16 程云鹏. 矩阵论. 西安：西北工业大学出版社，2000

17 谢庭藩，周颂平. 实函数逼近论. 杭州：杭州大学出版社，1998

18 Hahm N，Hong BI. Approximation by neural networks with a fixed weight. Computers and Math with Appl，2004，47(11)：1897—1903