用 Sigmoidal 函数的叠合逼近 Hilbert 空间中的连续泛函

陈 天 平

(复旦大学数学系，上海 200433)

关键词 神经网络、Sigmoidal 函数、Hilbert 空间、紧集、连续泛函

近来，Cybenko[1] 证明了下述

定理 A 设 $\sigma(x)$ 是一个连续的 Sigmoidal函数，则下述形式

$$\sum_{j=1}^{N} \alpha_j \sigma(x \cdot y_j + \theta_j)$$

的函数全体在 $C(I^n)$ 中是稠密的，其中 $y_j \in \mathbb{R}^n, x \in I^n, x \cdot y$ 是 x 与 y 的内积，α_j, θ_j 分别为实数，$I^n = [0,1]^n$.

上述结果不仅回答了关于用单个隐层的前馈神经网络 (Feedforward neural network) 的表示问题，而且在数学上也是很有意义的。早年，Kolmogorov 解决了 Hilbert 第13问题，证明了 $C(I^n)$ 上的连续函数可用有限个单变量函数的有限次复合及叠合精确表示。而定理 A 表明，$C(I^n)$ 上的连续函数可用单个 Sigmoidal 函数的有限次叠合以任意精度逼近。

但在文献 [2] 中，有些证明不够确切，证明也不是构造性的，因此，我们给出了一个构造性的证明并指出 $\sigma(x)$ 的连续性是不必要的，起本质作用的是它的有界性[2]。我们给出了

定理 B 设 $\sigma(x)$ 是一个有界的 Sigmoidal 函数，则

$$\sum_{j=1}^{N} \alpha_j \sigma(x \cdot y_j + \theta_j)$$

在 $C(I^n)$ 中稠密。

我们还指出[2]，如用 B-样条、Schoenberg Cardinal 样条以及 Wavelet 来代替Sigmoidal 函数，定理 B 也成立。

本文的目的是讨论 Hilbert 空间中紧集上的连续函数用 Sigmoidal 函数叠合的逼近问题。

定义 若 $\sigma(x)$ 是 $\mathbb{R} \to \mathbb{R}$ 的一个函数，且满足

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

则称 $\sigma(x)$ 是一个 Sigmoidal 函数。

设 H 是一个 Hilbert 空间。当 $x, y \in H$ 时，用 $\langle x, y\rangle$ 表示它们的内积 $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$ 表示 x 的模。

1991—10—30 收稿，1991—12—17 收修改稿

定理 1 设 $\sigma(x)$ 是一个有界 Sigmoidal 函数，H 为一个 Hilbert 空间,$U\subseteq H$ 是 H 中一个紧集。f 为定义在 U 上的一个连续泛函,则在 U 中存在可列个元素 $\{x_1,x_2\cdots x_n,\cdots\}$. 使 $\forall \varepsilon>0$，存在正整数 M,N，使

$$\left|\sum_{i=1}^{N}\alpha_i\sigma(\langle x,u_i\rangle+\theta_i)-f(x)\right|<\varepsilon$$

对一切 $x\in U$ 成立,其中 $\alpha_i,\theta_i\in \mathrm{R}$，$u_i\in \mathrm{Span}\{x_1,\cdots,x_M\}$，$i=1,\cdots,N$。

首先给出几个引理。

引理 1 设 U 是 R^n 中一个紧集，$\sigma(x)$ 是一个有界 Sigmoidal 函数,则

$$\sum_{i=1}^{N}\alpha_i\sigma(x\cdot y_i+\theta_i)$$

在 $C(U)$ 中稠密,其中 $x\in U$，$y_i\in\mathrm{R}^n$，$\alpha_i,\theta_i\in\mathrm{R}$。

证 首先，我们可将 U 上的连续函数延拓成一个包含 U 的 n 维立方体上的一个连续函数,然后通过平移压缩可将其归结为定理 B。

引理 2 设 U 是 H 中的一个紧集,则在 U 中存在可列个元素 $x_1,x_2,\cdots$ 在 U 中稠密。

引理 3 在引理 2 的假设下,存在一组可列的就范正交向量 $y_1,y_2,\cdots,y_n\cdots$，使 $y_i\in \mathrm{Span}\{x_1,\cdots,x_i\},i=1,\cdots$,且 U 是 $\widetilde{H}_1=\mathrm{Span}\{y_1,\cdots,\}$ 中的一个紧集。

引理 4 U 为 $\widetilde{H}$ 中一个紧集的充要条件是 U 在 $\widetilde{H}_1$ 中是有界闭的以及 $\forall\varepsilon>0$，存在正整数 M 和 $A>0$，使

$$\left(\sum_{i=M+1}^{\infty}|\langle x,y_i\rangle|^2\right)^{1/2}<\varepsilon$$

对一切 $x\in U$ 成立。同时，

$$U^M=\left\{x^M:x^M=\sum_{i=1}^{M}\langle x,y_i\rangle y_i,x\in U\right\}$$

是 $H^M=\mathrm{Span}\{y_1,\cdots,y_M\}$ 中的一个紧集。

引理 2—4 可在一般泛函分析书中找到。

引理 5[3] 设 f 是定义在 U 上的一个连续泛函,则它可延拓成 $\widetilde{H}_1$ 上的一个连续泛函。

引理 5 由著名的 Urilson 定理或 Tietze 定理得知。

定理 1 的证明 首先,把 f 延拓成 $\widetilde{H}_1$ 上的一个连续泛函，因此一开始就可假定 f 的定义域为 $\widetilde{H}_1$。

令 $\widetilde{U}$ 是 U 及一切 $U^M,M=1,2,\cdots$的并集,则由引理 4 可证 $\widetilde{U}$ 是 $\widetilde{H}^1$ 中的一个紧集。

由于 $\tilde{f}$ 是 $\widetilde{U}$ 上一个连续泛函，$\forall\varepsilon>0$，存在 $\delta>0$,使 $\forall x',x''\in\widetilde{U}$，当 $\|x'-x''\|<\delta$ 时，

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

又由引理 3，对 $\delta>0$，存在正整数 M，使

$$\left\|x-\sum_{i=1}^{M}\langle x,y_i\rangle y_i\right\|<\delta$$

对一切 $x\in U$ 成立。从而对一切 $x\in U$，成立着

$$\left|f(x)-f\left(\sum_{i=1}^{M}\langle x,y_i\rangle y_i\right)\right|<\frac{\varepsilon}{2}$$

或写成

$$|f(x)-f(x^M)|<\frac{\varepsilon}{2}. \tag{*}$$

这里 $x^M\in U^M$，而 U^M 是 $\mathrm{Span}\{y_1,\cdots,y_M\}$ 中的一个紧集. 利用引理 1，存在正整数 N，以及 $u_j\in\mathrm{Span}\{y_1,\cdots,y_M\}$，$\alpha_j,\theta_j\in\mathrm{R}$，$j=1,\cdots,N$ 使

$$\left|f(x^M)-\sum_{j=1}^{N}\alpha_j\sigma(\langle x^M,u_j\rangle+\theta_j)\right|<\frac{\varepsilon}{2} \tag{**}$$

对一切 $x^M\in U^M$ 成立.

由 U^M，x^M的定义，$\forall x\in U$，$u_j\in\mathrm{Span}\{y_1,\cdots,y_M\}$，$\langle x,u_j\rangle=\langle x^M,u_j\rangle$，从而结合(*)及(**)式，$\forall x\in U$，

$$\left|f(x)-\sum_{j=1}^{N}\alpha_j\sigma(\langle x,u_j\rangle+\theta_j)\right|<\varepsilon.$$

因为 $u_j\in\mathrm{Span}\{y_1,\cdots,y_M\}\subseteq\mathrm{Span}\{x_1,\cdots,x_M\}$. 故得证.

我们可以把上述结果推广到定义在距离空间中的连续泛函.

设 U 是距离空间 X 中的一个紧集，由于 U 是紧的(本身是个可分度量空间). 由 Urilson 定理，存在一个同胚映射 T，把 U 映成 Hilbert 空间 l_2 中一个紧集. 因此，成立着

定理 2 设 U 是距离空间 X 中一个紧集，f 为 U 上一个连续泛函，则存在 U 中可列个元素 $x_1,x_2,\cdots,x_n,\cdots$，使 $\forall\varepsilon>0$，存在正整数 N,M，$\alpha_j,d_{i,j},\theta_j\in\mathrm{R}$，$j=1,\cdots,N$，$i=1,\cdots,M$，成立着

$$\left|f(x)-\sum_{j=1}^{N}\alpha_j\sigma\left(\sum_{i=1}^{M}d_{i,j}\langle Tx,Tx_i\rangle+\theta_j\right)\right|<\varepsilon$$

对一切 $x\in U$ 成立，其中 T 是把 U 映入 l_2 中的同胚映射.

注 1 正如文献[2]中指出的，定理中的 Sigmoidal 函数可用 B-样条，Wavelet，Schoenberg Cardinal 样条以及其它函数代替，结论依然成立.

注 2 本文结果表明，任何 Hilbert 空间紧集上的连续函数可用单个特定一元函数的叠合及复合以任意精度来逼近. 同时也表明用一个隐层的前馈神经网络可逼近 Hilbert 空间紧集上的连续函数.

参 考 文 献

[1] Cybenko G., *Mathematics of Control, Signals and Systems*, 2 (1989), 4: 303—314.

[2] Chen Tianping, Chen Hong, Liu Ruey wen, A constructive proof of Cybenko's theorem and its extension (to appear in *IEEE Neural Networks*).

[3] Александров, Д. С., *Введение в общую Теорию Иномесмв в функций*, 1948.