ニューラルネットの入出力関係に関する一考察

逢坂一正・前田曜也
岡山理科大学工学部機械工学科

（1990年９月30日　受理）

1.　はじめに

機械の知能的制御法の一つとして，ニューラルネットを制御装置に用いる方法が注目されている。ニューラルネットの基本的機能といえるニューロコンピューティングに関する研究はかなり報告[1),2)]されている。R. P. Lippmann は，幾種類かのニューラルネットモデルにおける計算アルゴリズムを簡潔にまとめて報告[3)]している。しかし，ニューラルネットを一つのシステムと考えたときの入出力関係を表す関数形，その関数の性質などは明確にされてない。前報[4)]では，カスケード結合された３層ニューラルネットを入出力関係の関数発生器とみなし，それに教師値付きバックプロゲパーション法による学習を行なわせたとき，関数精度におよぼす中間層ニューロン数，教師値数，学習回数などの影響について報告した。しかし，教師値の出力に観測雑音が伴う場合には，ニューラルネットは単なる関数発生器というよりシステム同定器とみなすのが妥当であると思われる。つまり，ニューラルネットの学習過程は，ニューロン結合荷重の同定過程とみなすことができる。また，教師値を追加していく学習は逐次同定の一種と考えられる。

そこで本報告では，ニューラルネットで形成される関数の精度におよぼす観測雑音などの影響を明らかにするため，教師値の出力に白色観測雑音が重畳しているシステムの入出力関係を簡単な多項式で近似する関数の精度と前報の方法[4)]によりニューラルネットで形成される非線形関数の精度を数値シミュレーションにより比較，検討する。

2.　ニューラルネットの入出力関係

一般のニューラルネットは多入力多出力の関数発生器とみなせるが，本報告では一般性を失わないので，図１に示す２入力１出力に限定した場合を取り上げる。ここで，u_1, u_2

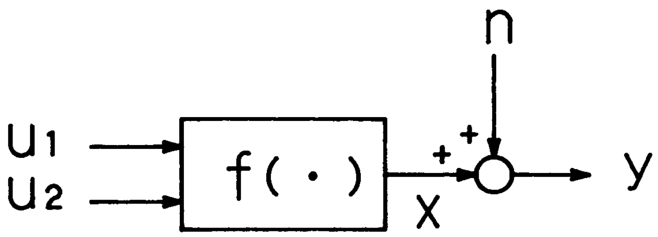

図１　同定システム

は入力，n は白色観測雑音，y は出力，f（•）は入力から出力への変換を表す関数である。ニューラルネットでは，一般に f（•）はシグモイド関数で合成される非線形関数としてモデル化される。

サンプル番号 i の入出力に対して，f（•）のモデルとして次の多項式近似の線形構造をもつ同定モデルを考える。

$$y^i = u^i a + n^i \tag{1}$$

ここで，a は p×1 の未知係数ベクトルで，$a = [a_0, a_1, a_2, a_3, a_4, a_5, \cdots, a_{p-1}]^T$，$u^i$ は入力 u_1^i，u_2^i で構成される 1×p の入力ベクトルで，$u^i = [1, u_1^i, u_2^i\ (u_1^i)^2, (u_2^i)^2, u_1^i\ u_2^i, (u_1^i)^3, \cdot\cdot\cdot]$ である。ただし，$(\cdot)^T$ は（•）の転置を表す。また白色観測雑音n は、平均値 $E\{n\}=0$，分散 $\mathrm{var}\{n\}=\sigma^2$ であるとする。(1)式を$i=1, 2, \cdot\cdot\cdot, q$ とおいて得られる q 個の方程式をまとめると次式となる。

$$\mathbf{y} = \mathbf{Ua} + \mathbf{n} \tag{2}$$

ここで，$\mathbf{y}$ は $q\times 1$ の出力ベクトルで $\mathbf{y} = [y^1, y^2, \cdot\cdot\cdot, y^q]^T$，$\mathbf{U}$ は $q\times p$ の入力行列で $\mathbf{U} = [\mathbf{u}^1, \mathbf{u}^2, \cdot\cdot\cdot, \mathbf{u}^q]^T$，$\mathbf{n}$ は $q\times 1$ の観測雑音ベクトルで $\mathbf{n} = [n^1, n^2\ \cdot\cdot\cdot, n^q]^T$ である。(2)式のもとで入出力の組 $\{\mathbf{U}, \mathbf{y}\}$ が得られたとして，

$$J = (\mathbf{y} - \mathbf{Ua})^T (\mathbf{y} - \mathbf{Ua}) \tag{3}$$

を最小にする未知係数ベクトル $\mathbf{a}$ の推定値 $\hat{\mathbf{a}}$ は，$p \leqq q$ のとき次式のように求まる（最小2乗推定[5]）。

$$\hat{\mathbf{a}} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{y} \tag{4}$$

ただし，$(\cdot)^{-1}$ は（•）の逆行列である。ニューラルネットで形成される関数を(4)式の $\hat{\mathbf{a}}$ を用いて(1)式で近似する場合，入出力のサンプル数を追加していく同定法は，ニューラルネットにおける教師値の増加で関数精度の向上をはかる学習法に対応する。

3.　数値シミュレーション

ここでは，前報の方法によりニューラルネットで形成される関数の精度と本同定モデルによる関数(1)式の精度について，教師値の追加，観測雑音の分散の大きさなどの影響を数値シミュレーションにより調べる。

図1に対応する数値シミュレーション用3層ニューラルネットを図2に示す。中間層のニューロン数 m は形成すべき関数を考慮して決めるべきだが，ここでは関数精度を調べるために使われる見本関数を次の2次関数とするので，十分な精度が出ると思われる $m=10$

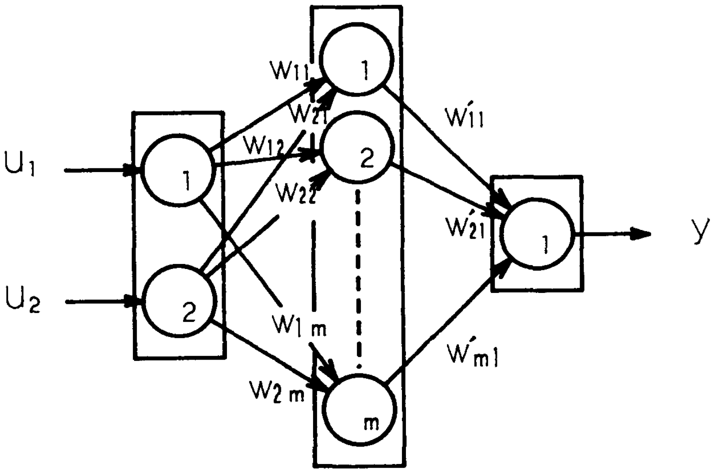

図2　3層ニューラルネット

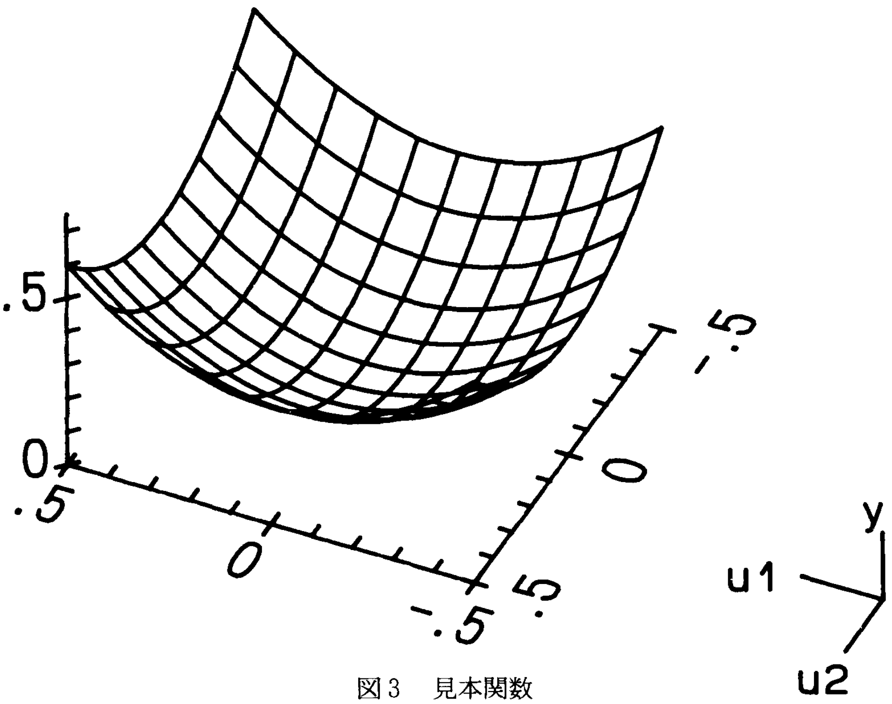

図3　見本関数

を採用[4]する。

$$y = u_1^2 + u_2^2 + 0.1 \tag{5}$$

ただし，$-0.5 \leqq u_1^2$，$u_2^2 \leqq 0.5$ の矩形領域とする。見本関数はどのような関数でもかまわないが，関数の簡単さと図示の容易さを考慮して(5)式を採用した。図3にその見本関数を示す。観測雑音は一様乱数を発生させて所要の平均値，分散になるように変形して用いる。

関数精度を判定する誤差eを次式のように決める。

$$e = \sum_{i=1}^{100} (y^i - \hat{y}^i)^2 \tag{6}$$

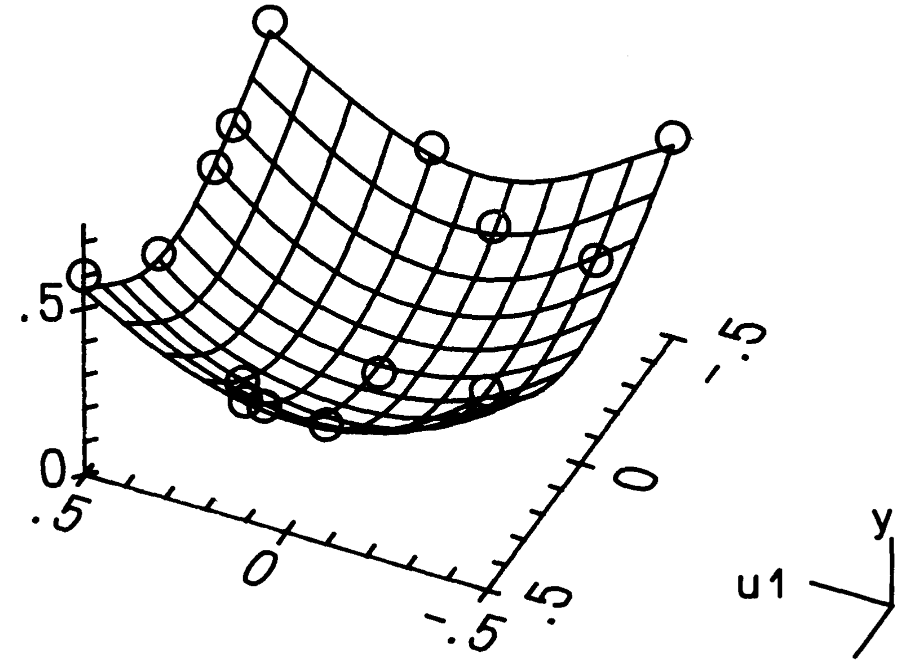

図4　ニューラルネットで形成された関数

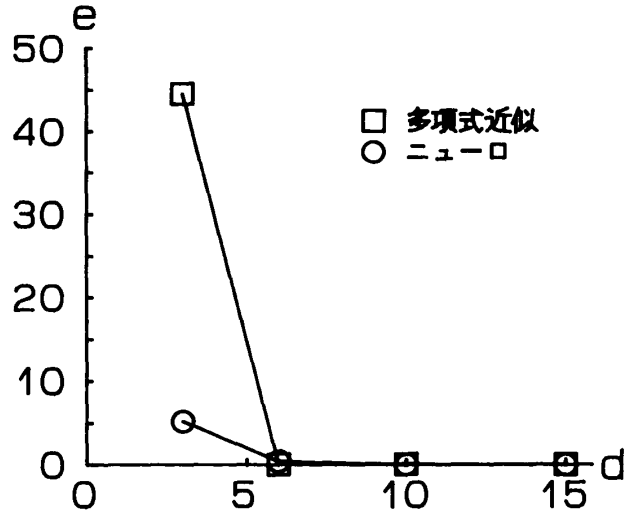

図5　教師値数と誤差の関係（ＳＮ比；２％）

ここで，y^i は入力領域内の u_1 ，u_2 軸を原点を中心にそれぞれ10等分する100個の格子点における見本関数(5)式の y の値，$\hat{y}^i$ は同じ格子点におけるニューラルネットで形成された関数，または同定モデル(1)式の y^i の値である。

観測雑音がない場合，ニューラルネットにより15個の教師値を用いて前報の方法で形成された関数を図４に示す。ここで，○印は教師値である。図３とほぼ完全に一致していて，この関数の誤差は $e\fallingdotseq 0.01$ であるので，十分な精度をもつ関数といえる。

逐次学習過程におけるニューラルネットと同定モデルで形成される関数精度について述べる。同定モデルの多項式近似(1)式は，次数の低いほうから１次式，２次式，３次式，４

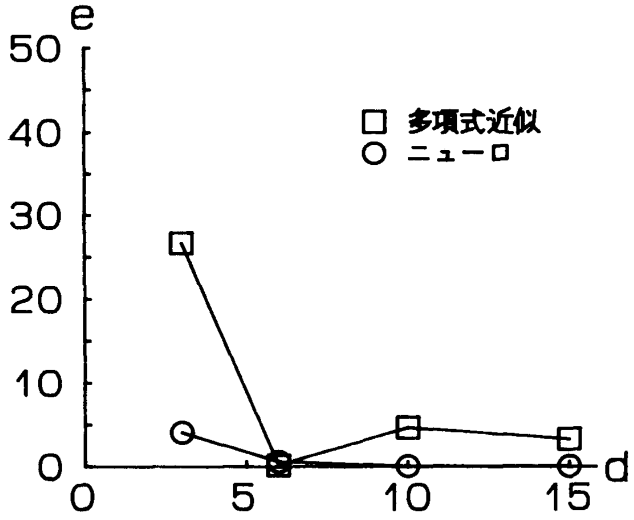

図6　教師値数と誤差の関係（*SN*比；10%）

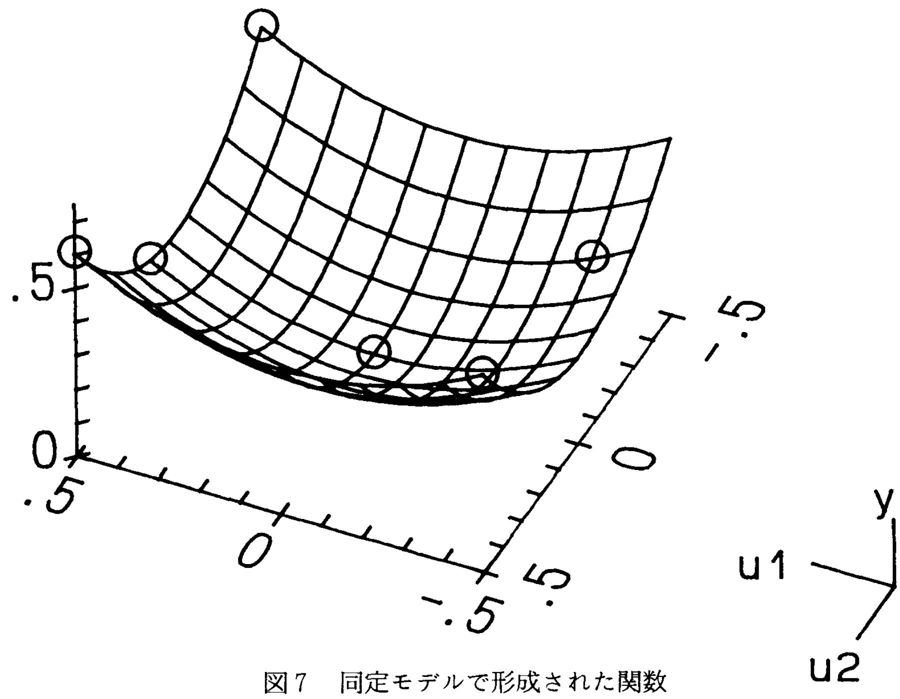

図7　同定モデルで形成された関数

次式の順に同定していく。そのとき，同定に最小必要な教師値数（係数ベクトルの要素数）はそれぞれ3個，6個，10個，15個である。したがって，ニューラルネットに対する教師値数も同じ数にして学習させる。

ニューラルネットと同定モデルで形成される関数の教師値数 d と誤差 e の関係を図5に示す。この結果は，観測雑音の標準偏差と信号の代表値の比（σ / y_r)，(以後，簡単にSN比という）が2%の場合である。ただし，信号の代表値 y_r は，(5)式の y の最小値と最大値の算術平均をとり，$y_r = 0.35$である。ニューラルネットと同定モデルの両方とも教師値の数が増加すると誤差は減少（ある意味の学習）し，教師値数が3でニューラルネッ

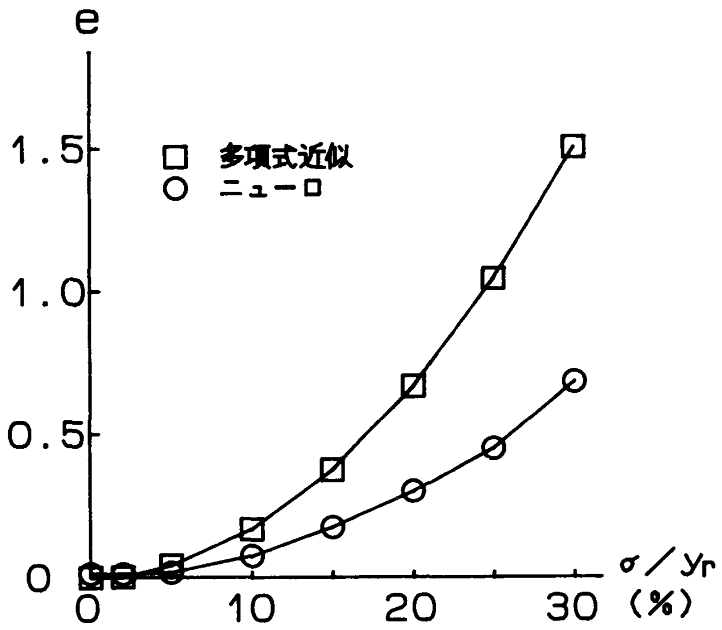

図8　SN比と誤差の関係

トのほうが同定モデルより精度が良いが，6以上で両方ともほぼ見本関数に一致する。この場合の同定モデルで求まった2次近似と3次近似の係数ベクトルは，それぞれ次のとうりである。

2次近似；$\hat{\mathbf{a}}=$ [0.103, 0.001, 0.073, 0.965, 1.021, 0.030]^T

3次近似；$\hat{\mathbf{a}}=$ [0.103, 1.432, −0.164, 1.083, 0.910, 0.010, −0.080, 0.004, 0.004, 0.020]^T

SN比が10%の場合の教師値数 d と誤差 e との関係を図6に示す。この場合もニューラルネットでは図5の結果と同じことがいえる。同定モデルでは見本関数と同じ次数の2次式近似が最良であり，同定モデルの近似式の次数があらかじめ分かっていれば精度の良い同定が可能であることを示している。この場合の同定モデルで求まった2次近似と3次近似の係数ベクトルは,それぞれ次のとうりである。

2次近似；$\hat{a}=$ [0.130, 0.002, 0.073, 0.652, 1.213, 0.030]^T

3次近似；$\hat{a}=$ [0.130, 1.434, −0.164, 1.832, 0.101, 0,098, −8.009, 0.410, 0.400, 2.143]^T

同定された関数の一例として，この2次近似の係数ベクトル$\hat{\mathbf{a}}$による(1)式を図7に示す。図3の見本関数とよく似ているように見えるかもしいが，よく見るとかなりずれているのが分かる。

次にニューラルネットと同定モデルにおいて，SN比と誤差の関係を図8に示す。教師値数は，同定モデルが6個，ニューラルネットが10個である。どのSN比でもニューラルネットの誤差が同定モデルの誤差の約半分であり，ニューラルネットで形成される関数精度のほうが同定モデルの関数精度より良い。しかし，ニューラルネット，同定モデルの

両方ともＳＮ比の増加につれて誤差が急激に増加し，関数精度は急激に悪くなる。

4. おわりに

ニューラルネットと多項式近似の同定モデルにおける入出力関係を表す関数の精度におよぼす観測雑音のＳＮ比と教師値数の影響について数値シミュレーションにより調べた。得られた結果は,つぎのようにまとめられる。

(1) 簡単な２次式の見本関数の形成では，ニューラルネットのほうが同定モデルにより精度の良い関数が形成できる。

(2) 観測雑音のＳＮ比が２％以下の範囲では，ニューラルネットと同定モデルの両方とも関数の逐次学習が成立する。

(3) 観測雑音のＳＮ比が増加すると，ニューラルネット，同定モデルの両方とも関数精度が急激に悪くなる。

終わりに，日頃ご指導いただいている大阪府立大学の小野敏郎教授に感謝します。

参考文献

1） J. J. Hopfield : Neurons wich graded response have collective computational properties like those of two-state neurons, Proc. of the National Academy of Science U.S.A., vol. 81, pp. 3088 - 3092, 1984

2） D. W. Tank, J. J. Hopfield : Simple Neural Optimization Networks : An A／D Converter, Signal Decision Circuit, and a Linear Programming Circuit, IEEE Trans. on Circuits and Systems, vol.CAS-33, no. 5, pp. 533-541, 1986

3） R. P. Lippmann : An introduction to Computing with Neural Nets : IEEE ASSP Magazine, pp. 4-22, April 1987

4） 逢坂一正：コントローラーとしてのニューラルネットワークの学習，岡山理科大学紀要，第25号A, pp. 273-282, 1990

5） 中溝高好：信号解析とシステム固定，コロナ社， 1989

A Study of Function Generated from Neural Network

Kazumasa OHSAKA and Teruya MAEDA

Department of Mechanical Engineering,
Okayama University of Science,
1-1, *Ridaicho, Okayama* 700, *Japan*

(Received September 30, 1986)

In this paper, the precision of function which expresses the relation between input and output of a system and is generated from a neural network is compared with that of the function generated from an identification model which is expressed b y a polynominal, and the effect of observation noise on the precision is investi-g ated by numerical simulation.

The results obtained are summarized as follows.

(1) In the generation of a sample function expressed by a square polinominal,the precision of function generated from the neural network is better than that of function generated from the identification model.

(2) In the case of less than 2% of signal to noise ratio, a learning process which means the decreasing of the error between the sample function and generated functions with the increasing of the number of teaching values is realized in both neural network and identification model.

(3) The precision of functions generated from both neural network and identification model grows worse rapidly with the increasing of signal to noise ratio.