ニューラルネットを利用したシフトバリアントデータの解析

橋　本　文　彦

1．序論
2．ニューラルネットの基本構造
3．シフト・バリアントなデータからのモデル構築
4．因子分析への応用
5．今後の展望

1. 序　論

近年，様々な分野においてニューラルネットワークを用いたデータ解析が行われるようになってきている。旧来のデータ解析の手法が，専ら代数的な処理に依るものであるのに対して，ニューラルネットワークによるデータ解析は，観測値として与えられたデータを，ネットワークの内部に分散された情報として表現する。しこうして，そのネットワークの持つエントロピーを減少させる（エネルギーを減少させる）ことで解を得るという，数学的にはきわめて単純な処理を行うものである。ニューラルネットは，特に多パラメタ問題の最適解を探索する場合に，かなりよい近似値を与えるものとして利用できることが明らかになってきている。

ニューラルネットワークに関する研究には主に二通りのアプローチがある。そのひとつは，人工知能の理論的な研究分野からのアプローチであり，ここでは，人間の脳の神経細胞（ニューロン）の構造にヒントを得て，その分散情報処理の過程が如何に人間の情報処理過程と似ているかという点を大きな特徴と考えており，将来的には人間が行うと同様の精神活動が行えるようになると考えてのアプローチである。また，いまひとつは，広大な解空間上に多数のパラメタを持つような，計算機の実時間内で代数的，あるいは反復的に解を得ることが困難であるような問題に対して，実用的な計算時間での「近似解」を得るための手法として，人間の情報処理活動との類似は棚上げした形でニューラルネットを利用しようとするアプローチである。

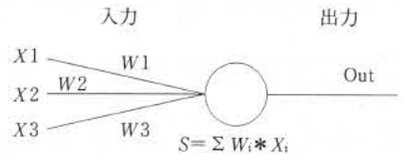
このような二つのアプローチのうち，本論では後者の立場，すなわちニューラルネットワークを統計解析のための近似的解法の一道具だてとして用いることで，これまでは問題が複雑になりすぎて，解を得ることができなかったような問題に対して，近似的ではあるが，十分に実用になる解を与えることを目標として，ニューラルネットワークの応用の可能性を探ることにしたい。

〔キー・ワード〕
ニューラルネット，ボルツマンマシン，シフト・バリアント，最小二乗法，因子分析

2. ニューラルネットの基本構造

ニューラルネットは，いくつかのニューロンが接続されたものであり，一つのニューロンは，一つのユニットと，他のユニットに接続するためのラインよりなる。（図1）

図1



一つのニューロンに対して，複数の入力がなされることがあるが，ラインに入力された情報は，各ラインに固有の重み（ウェイト）を持ち，その重み付けがなされた上でユニットに入力される。

重み付けされた複数の入力を受け取ったユニットは，それらの入力を総合する。具体的には，ライン i からの入力 x_i に，w_i なる重み付けがなされているならば，ユニットへの総合された入力（S）は，

$$S=\Sigma\ x_i * w_i$$

で与えられる。

入力の総和を与えられたニューロンは，以下のような出力を行う。

$$\mathrm{Out}=1\cdots\cdots(S>=h\text{の時})$$

$$\mathrm{Out}=0\cdots\cdots(S<h\text{の時})$$

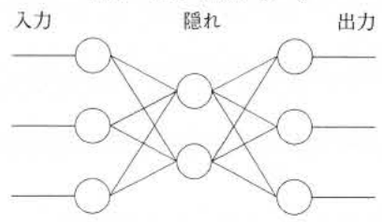
（ここで，h はあらかじめ定められた閾値）

出力ラインが複数ある時には，ユニットは，すべての出力ラインに対して同じ値を出力する。

このようなニューロンが複数結合されたものが，ニューラルネットワークである。

ニューラルネットの形態には，大きく二種類がある。そのひとつは，パーセプトロンと，その発展系であるPDP（誤差逆伝搬型ニューラルネット）であり，この形式は（図2）のように，入力された信号は常にラインを一方向に流れ，出力ラインから出力される。出力された反応は外界からの教師信号と照合され教師信号と出力信号との差が小さくなるように，ニューラルネット内部の重み付けが，出力層に近い層から入力層へと逆方向に向けて改訂される（これが「誤差逆伝搬」である）。

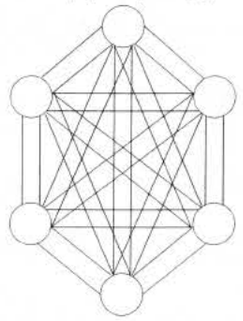
図2　PDPネットワーク



ニューラルネットの今一つの形態は，ホップフィールドモデルに代表される，相互結合型の

ニューラルネットである（図3）。この形式は，基本的にはすべてのユニットが，他のすべてのユニットとラインで結合されており，入力のみのラインや，出力のみのラインを持たない。この形式のネットワークでは，あらかじめ問題にあわせてネットワークの設計がなされた上で，ニューラルネット内部のエネルギー量を次のように定義し，このエネルギーが減少するように，ユニットを相互に結ぶラインの重み付けを変更してゆき，最終的に収束させるというものであり，外部からの教師信号を必要としないという点で，前記のPDP型のニューラルネットと異なる。

図3　ポップフィールドネットワーク
2本のラインの内，1本が入力，1本が出力



相互結合型のニューラルネットの内部エネルギーは次の様に定義される。

$$E=-(1/2)\,\Sigma\, W_{ij}{*}x_i{*}x_j-\Sigma h_i{*}x_i$$

すなわち，相互結合型のニューラルネットワークでは，解くべき問題（最適解を求めるべき問題）を，ある方程式の値を最小にするという問題に置き換え，この方程式の最小化問題がエネルギーの最小化問題と等値になるように，ニューラルネットのユニット数や結合状態を定めて設計すればよいのである。

本研究で用いたネットワークは上記の内，相互結合型のネットワークである。ニューラルネットによる多パラメタの最適化では，解空間が変数に対して比較的なだらかに変化する際には多くの場合，近似的にほぼ最適な解を与えるが，解空間が複雑に入り組んだ形をしている場合には，エネルギーが局所最小解に落ち込み，真の解とは異なる値を与える可能性がある。このような局所解への落ち込みを回避する方法としてボルツマンマシンのモデルがある。ボルツマンマシンは，ニューラルネットの内部エネルギーを熱エントロピーとして捉える。つまり，エネルギーの減少による解の収束は，熱エントロピーの減少に相当し，解が局所解に落ち込むことを防ぐためには，温度を急激に下げるのではなく，意図的にエントロピーを与えることで，大域的な最小値を見いだそうとするものである。

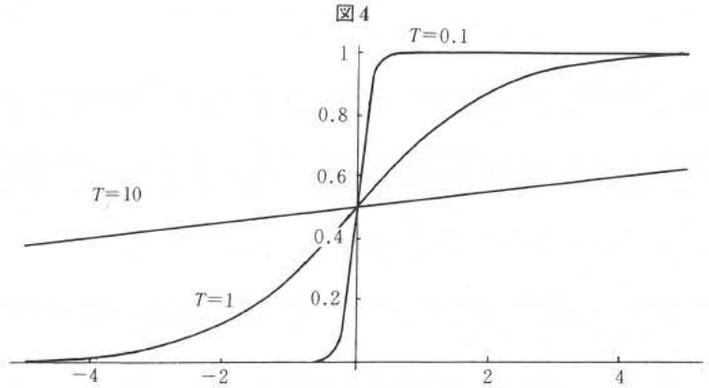
すなわち，通常のホップフィールドモデルでのニューロンがある閾値でオール・オア・ナッシングで1または0を出力するのに対して，ボルツマンマシンでは次のような確率値に従って1または0を出力することになる。

$$\mathrm{Prob}\ (\mathrm{Out}=1)=\mathrm{Exp}\ (S/T)/(1+\mathrm{Exp}\ (S/T))$$

ここで，Sは各ラインからの入力の総和，Tは「温度」と呼ぶべき量で，これが大きい程，エントロピーは大きく，収束を遅らせる役割を果たし，小さいほどエントロピーが小さく，収束を早める役割を果たす。（図4）

ボルツマンマシンでは，無限に長い時間をかけて温度を下げていった場合には，代数的な方法で得られる真の解に収束することが知られている。しかしながら，収束までに無限の時間をかけることは不可能であるので，いかなる割合で温度を下げていくかという点が問題として残されている。

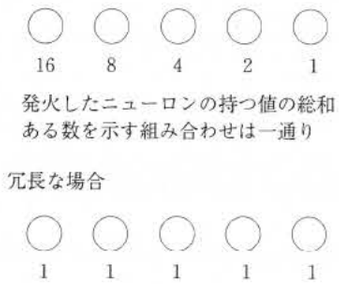
図4



本研究では，以下のような設計をボルツマンマシンに与えることで，収束時間と解の精度をできるだけ自由に扱えるように改良を行った。

☆　通常は，回路上で各ニューロン素子に与える意味付けとして，2進数で与えることで，冗長さを廃するが，本研究で用いたニューラルネットでは，逆に特定の変数に対して冗長さを持たせることで，高い精度の望まれる変数に対して，収束しやすくなるようにした。逆に，さほど精度が必要とされないような変数に対しては，冗長さを廃し，低い精度でも収束とみなすような回路とした。（図5）

図5



以上の準備のもと，通常の代数的解法では，実時間内に解くことがきわめて困難な統計的問題に対して，ニューラルネットによる解の有効性を以下に探った。

3. シフトバリアントなデータからのモデル構築

観測されたデータが何らかの変動を含むと考えられる場合に，この変動の性質（変動を与える分布関数）が既知であれば，そこからモデル推定を行うことは，基本的な考え方の枠組み自体はそれほど難しいものではない。しかしながら，多くの場合はこのような変動の性質そのものについてはほとんど知識がないために，もしくは，数学的な構成の複雑さをさけるために，一般的な形として観測データのすべての位置で同様の分散を持つガウス分布による変動を前提としてモデル推定を行うことが多い。あるいは，このような前提をいかすために，むしろモデ

ルを構成する軸の方をログ変換などによって修正した上で改めて均質の分散を持つガウス分布を当てはめるといった手法も用いられる。

このうち，変動の性質についての知識が乏しいことを理由に，上記のような仮定の下にモデル推定を行うことは仕方のないことであるが，変動に対する何らかの知識が得られている場合にまで，変動を無視したり，あるいは軸側を変換することで，このように単純化された仮定を維持する必要が本当にあるのだろうか？「真の構造は単純である」とは多くの科学者の信念であるが，明らかに伝達関数側が軸上の位置に依存して変化する場面は実際にみられることである。

最小二乗による直線回帰のような単純な推定の場合には，たとえば Y 軸方向での変動が X 軸上での位置に依存して関数そのものが変化する（これをシフト・バリアントと呼ぶ）としても計算機の実用的時間で解くことができる可能性がある。しかし，推定すべきモデルが複雑な構成を持つようになるほど，また変数が増えるほどシフト・バリアントな変動関数に対して計算すべき量は爆発的に増大することになり，通常の代数的あるいは反復的な解法によってはモデルを得ることができない。

図6

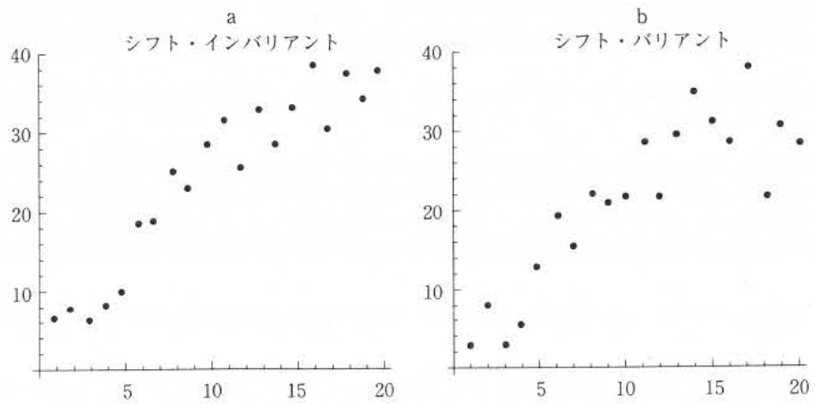


図6aにシフト・バリアントでないデータ（シフト・インバリアントと呼ぶ）のグラフ，図6bにシフト・バリアントなデータを含むグラフを示す。

図6aでは，X 軸方向でのシフトに対して，Y 軸方向に発生する誤差の分散が常に一定である。これに対して，図6bでは，X 軸上での値が大きくなるほど，Y 軸上での誤差の分散が大きくなる。

このグラフは，モデルにノイズを加えることで，理論的に作成したもので，ノイズを加える前の真のモデルは，

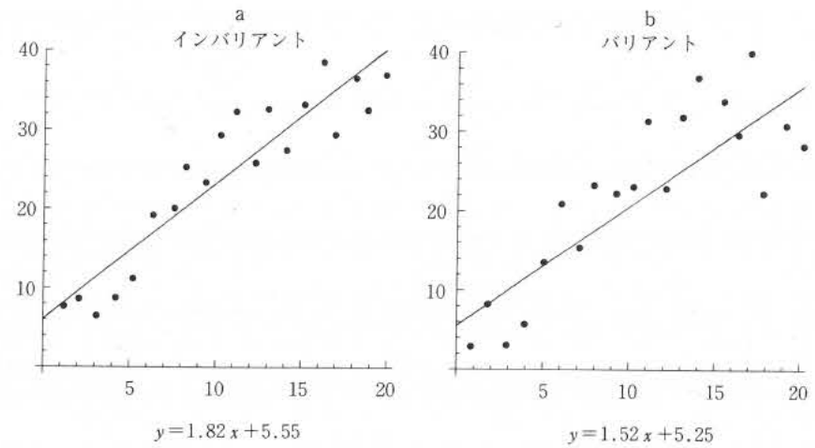
$$Y=2*X+4$$

である。

データ全体に対する誤差の分散は，両者で等しくしてある。

以上のように作成されたデータに対して，通常の最小二乗法で与えられた解が図7aと図7bである。シフト・バリアントなデータの方は，X 軸上の大きな値での誤差（この場合には Y 軸方向を下向きに引っ張っている）に惑わされて，真のデータよりもかなりはずれた値になっている。

図7

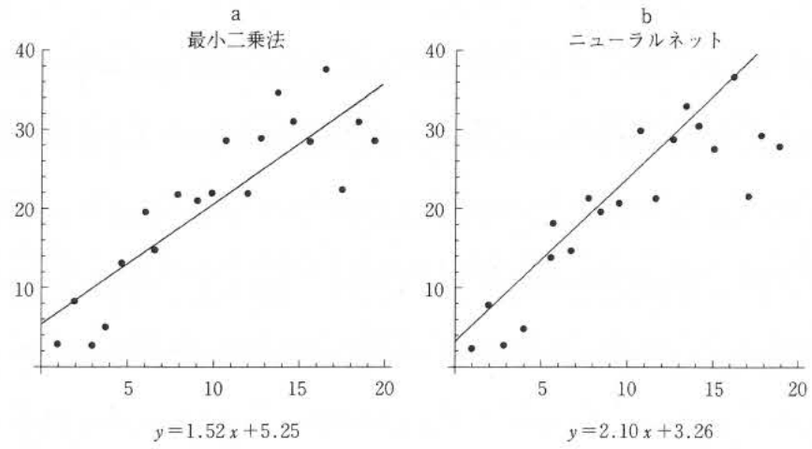


前章の末尾で触れたように，本研究では，通常のボルツマンマシンに対して，次のような改良を施したニューラルネットを用いて，上記のデータに対する解を求めた。

〔改良点〕

求めるべきモデルを $Y=a*X+b$ で表すと，a および b を意味するユニットとして，ともに

図8



整数部10ビットで［0, 10］を表現する事で，冗長度を付加した。一方，整数部については二進数で表現することが冗長さを取り除いた。

これによって，少なくとも整数部は，直ちに収束し，求解に要する時間の殆どは，小数点以下の値の精度向上に用いられることになる。

図8aと図8bに，以上で作成したニューラルネットによって得られた結果を示した。図8aは，先に示した図7bと同じもので，通常の最小二乗法によって得られた結果であり，図8bが改良型ボルツマンマシンのニューラルネットによって得られた結果である。

プロットされた観測値に対して引かれた直線としては，一見すると通常の最小二乗法を用いた図8bの方がフィットしているように見えるが，実際には，X軸方向で値の大きな領域は，より大きな誤差を含んでいるので，ニューラルネットによって得られた解の方がはるかに真の値に近い解を求めていることが解る。

4. 因子分析への応用

前章で示した程度の，シフト・バリアントなデータならば，旧来の代数的な解法でも，労力と時間をかければ，近似解ではなく，厳密な解を求めることができるが，より複雑な構造を持ったデータに対しては，代数的解法では求解が困難な場合が殆どである。

何らかの変量を観測・測定する場合，各変量の測定点によって誤差の大きさが変動するということは現実の場面で多く経験することである。その変動が，リニアな場合には，代数的解法が有効な場合もあるが，例えば，測定レンジの中央付近では誤差が小さいが，レンジの両端付近で誤差が大きくなるというような場合には，ニューラルネットを用いることがきわめて有効であるように思われる。

次に示すように，実験的に因子分析モデルのデータを作成した。（表1）

表1　実験データ

因子負荷量		変量1	変量2	変量3	変量4	変量5
	因　子　1	1	0.7	0.5	0.3	0
	因　子　2	0	0.3	0.5	0.7	1
誤　　差	理　論　値	0	0	0	0	0
	インバリアント	5	10	15	10	5
	バリアント	5	10	15	10	5
因子得点		平均	標準偏差			
	因　子　1	60	15			
	因　子　2	40	9			

求解に用いたニューラルネットは，前章で用いたものと同様の構造であり，ユニットの数のみがかなり多くなっている。すなわち，ユニットとして，各変量と因子毎に，整数部10ビットで値［0, 10］を，小数部4ビットで，小数点以下二桁目までを意味するようにネットを作成した。

通常の因子分析によって求めた解と，ニューラルネットを用いて求めた解とを表2で比較した。

表2　計算された因子得点と元の因子得点
（標準化されたもの）との二乗誤差を求めた。

分析結果		残　差
通常の因子分析	誤差なしモデル	0.719726
	インバリアント誤差	17.58055
	バリアント誤差	60.0297
ニューラルネット	バリアント誤差	49.02223

＊ニューラルネットを用いたのは因子負荷量の計算部分まで

表の値は，求められた因子得点ともとの因子得点（標準化されたもの）との二乗誤差を求めたものである。

通常の因子分析は，シフト・インバリアントなデータに対しては，予想以上に良い値を示したが，シフト・バリアントなデータでは，かなり大きな誤差を含む結果を得ることとなった。

一方，ニューラルネットによる結果では，シフト・インバリアントなデータに対する因子分析には劣るものの，通常の因子分析に比較してよりよい近似値を解として求めることができた。

5. 今後の展望

ニューラルネットを用いたデータ解析には，様々な手法があり，本研究で示した回路も，その一つである。しかしながら，様々な統計解析に対して，最適なニューラルネットの回路を構築するには，現在のところ，試行錯誤による以外に方法はない。特定の問題に対して，理論的に最適なニューラルネットの回路を機械的に設計するための指針が必要である。

また，ニューラルネットの動作原理自体が単純であるにも関わらず，通常のコンピュータで計算をシミュレートするには，未だ長い時間を要する（勿論，代数的に厳密な解を求めるよりもはるかに速いが）。

問題がより大規模なものになったときに，果たして，ニューラルネットの設計方針が現在のままでうまく動作するのかという点が，今後考慮されなくてはならない問題であり，また，これらの結果得られた解が，近似値としてもどの程度信頼できるのかという点について，より明確に数学的な根拠付けを行っていく必要がある。

参考文献

Churchland, P. M. (1979), *Scientific Realism and Plasticity of Mind*, New York : Cambridge University Press.

Churchland, P. M. (1989), *A Neurocomputational Perspective*, Cambridge, MA : MIT Press.

Domany, E. (Eds.) (1994), *Models of Neural Networks.*

橋本文彦（1995),「日本行動計量学会第23回講演抄録」70-71。

Hopfield, J. J. and D. W. Tank, (1985), "Neural Computation of Decisions in Optimization Problems", *Biological Cybernetics*, 52, 141-152.

McCulloch, W. S. and W. H. Pitts, (1943), "A Logical Calculus of the Ideas Immanent in Neural Nets", *Bull. Math. Biophys.*, 5, 115-133.

松岡清利（編）(1992),『ニューロコンピューティング』, 朝倉書店。

上坂吉則 (1993),『ニューロコンピューティングの数学的基礎』, 近代科学社。