# ヘシアン行列を用いた階層型ニューラルネットワークの評価

## － 認知神経心理学のシミュレーションツールとして活用するために －

**○浅川伸一（あさかわ　しんいち）**
**東京女子大学・情報処理センター**

**(要旨)**　ニューラルネットワークの認知神経心理学的な応用を目指した研究では，多くの場合パラメータ過剰により不良設定問題になっていることを指摘する．不良設定問題の場合，誤差関数の２階微分行列（ヘシアン行列）は正則行列とは限らない．数値実験から，多くのモデルでは，実際に最適解でない可能性があることがわかった．このような不良設定問題の場合，収束性の判定，一般化可能性，訓練データの持つ構造の獲得，などが満たされていない可能性がある．不良設定問題を回避するためにパラメータを減らす努力がなされるべきである．

Key words: **ニューラルネットワーク，ヘシアン行列，不良設定問題**

### 1.　問題点の指摘

伝統的にニューラルネットワークの神経心理学的応用の研究分野では，学習成立後に任意のユニット間の結合をゼロにする，または，任意のユニットを除去することによって損傷が表現されてきた．これらの研究では，損傷の程度を NN の切断，または除去の割合として表現し，損傷の程度によって課題成績の変化する割合を検討している．ニューラルネットワークは訓練データが固定されたときのパラメータ推定問題とみなすことができる．ところが，今までに提案されたモデルでは，パラメータ数（ネットワークの結合数）に比べてデータ量が極端に少ない．しかも，ほとんどの研究でニューラルネットワークの挙動が体系的に調べられていない．一般に推定すべきパラメータ数が多ければデータ近似の意味では精度が向上するが，一方でパラメータ過剰により，解くべき問題は不良設定問題となってしまう．このような不良設定問題では，

・最適値に収束することが保証されない
・モデルの一般化可能性が低い
・データ構造の獲得に失敗している

等の問題が指摘できる．ところが，ほとんどの文献で収束性についての議論が行なわれていない．シミュレーション結果が不適解であるのならば，神経心理学的現象のシミュレーションとしての損傷実験は意味をなさず，従って神経心理的モデルとして不適切であると言わざるを得ない．推定すべきパラメータ数の数倍程度の入力データ次元数がなければ一般化は困難である．モデルが一般化可能性を持つためには，データの持つ次元数／パラメータ数の比が 3 から 5 程度であることが望ましいといわれている．この比を基準と考えれば，過去のモデルのほとんどは一般性が保証されず，モデルがデータ構造を学習したとは考えられない．

以下では，上記のようなパラメータ過剰による不良設定問題の場合における収束性の判定に誤差の２次微分行列（ヘシアン行列）を用いる方法を述べている．

### 2.　ヘシアン行列による収束判定

標準的な誤差逆伝播学習則，すなわち勾配降下法においては，最小化すべき目標関数として平均２乗誤差関数 $E$ が用いられている．学習すべきデータセットが固定されていれば $E$ はニューラルネットワーク内の全結合係数の関数と見なすことができるので $E(w)$ と表記できる．このとき $E$ を各パラメータについて 2 階微分したヘシアン行列が $E$ の形状についての情報を与える．解くべき問題が良設定問題であれば誤差関数の変化は Taylor 展開を用いて

$$\delta E(w|x)=\left(\frac{\delta E}{\delta w}\right)^t\delta w+\frac{1}{2}\delta w^t H\delta w+O(|\delta w|^3) \tag{1}$$

のように近似することができる．ここで添字 t は行列の転置を表し，H はヘシアン行列である．学習が成立したと仮定した場合，もしくは，局所最小点で学習が進まなくなった状況では，誤差関数 $E$ の 1 次微分はゼロとみ

**連絡先：〒**167-8585 **杉並区善福寺** 2-6-1　Tel. 03-5382-6746 e-mail: asakawa@twcu.ac.jp

なすことができる．また，誤差関数が 2 乗の形で定義されていれば，(1) 式第 3 項は 3 次以上の項なので無視できる．このとき学習が収束しているとは

$$\frac{\partial E(w)}{\partial w}=0 \qquad (2)$$

を意味し，かつ，ヘシアン行列の対角要素がすべて正でなければならない．

$$\frac{\partial^2 E(w)}{\partial w^2}\geq 0 \qquad (3)$$

そこで，ヘシアン行列の対角要素がすべて正であることを基準として提案する．この基準は，誤差関数の形状が凹であることを意味する．

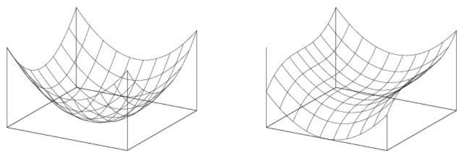

図 1 ヘシアン行列の対角要素がすべて正であるときの誤差関数の形状（左）と一つ以上の対角要素が負である場合の形状（右）の概念図

上記の基準がどの程度満たされているかを排他的論理和 XOR 問題を用いて検討した．ここではすべての場合で 3 層のネットワークとし，入力層ユニットから中間層ユニットへと，中間層ユニットから出力層ユニットへの結合は全結合とした．入力層ユニット数を 2，出力層ユニット数 1 に固定し，中間層のユニット数を 2 から 10 まで変化させた．このとき推定すべきパラメータ数は中間層ユニット数によって決まり，中間層ユニット数を n とすればパラメータ数は (2+1)×n　+ (n+1)×1 となる．学習は通常の誤差逆伝播則とした．図 2 は，中間層数を 2 から 10 まで変化させたときの収束基準に達した割合を示している．図 2 からパラメータ数が 17 個（中間層のユニット数は 4）以上では必ず収束していることがわかる．すなわちパラメータ数が多ければ訓練データは任意の精度で近似できることが確認できる．一方，図 3 に収束した場合，ヘシアン行列の対角要素に一つ以上の負の値が含まれる割合を示した．図 3 から，パラメータ数が 9 のとき（中間層のユニット数は 2 のとき）には全ての場合でヘシアン行列の対角要素が正であったが，パラメータ数の増加に伴いヘシアン行列の対角要素に負の

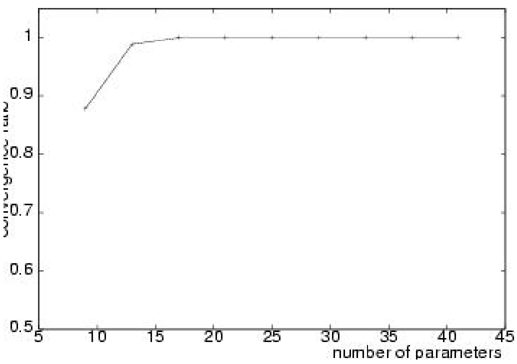


図 2　各パラメータ数において 1000 回数値実験を繰り返したとき，収束基準に達した割合

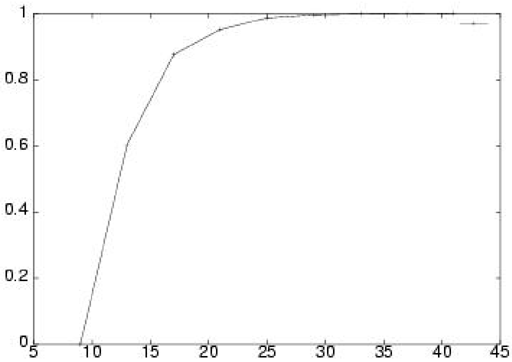


図 3　各パラメータ数において 1000 回数値実験を繰り返したとき，収束基準に達したパラメータにおいてヘシアン行列の対角要素に負の値が出現する割合

値が現れる確率は急速に上昇することがわかる．ヘシアン行列の対角要素に負の値が現れることは局所的に誤差関数の形状が凸に近いことを表しており，学習が収束したとしても，パラメータ過剰による偽収束である可能性がある．

以上のように不良設定問題では，推定すべきパラメータ数が多くなればなるほど，不適解に収束する割合が高く，適解に収束したとみなすことはできない．このような場合にニューラルネットワークモデルを認知神経心理学のシミュレーションツールとして用いることは危険であると言わざるを得ない．

不良設定問題の解決策は，正則化項の付加による拘束条件，または適切な前処理を施すなどしてパラメータ数を減らすことであり，推定すべきパラメータ数を訓練データの持つ次元数より小さくする努力がなされるべきである．