【土木計画学研究・講演集 No.17 1995年1月】

ニューラルネットワークの選択行動モデルへの応用可能性*

A Potential for Application of Neural Network to Choice Behavior Model*

清水英範** 平谷浩三***
By Eihan SHIMIZU**, Kozo HIRATANI***

1.はじめに

階層型ニューラルネットワーク（以降、NNと呼ぶ）のバックプロパゲーション（BP）法は、複雑な非線形の入出力関係を精度よく推定する手法として注目され、近年では土木工学の諸分野においても数多くの応用研究が試みられている。

しかし、これら従来の研究には少なからず次のような傾向が見られる。

（1）NN全体の入出力関係を数式として記述すると極めて複雑になる。従って、推定モデルとしてのNNに意味解釈を与えることは困難になり、その利用はブラックボックス的になりがちである。

（2）NNによる高精度な推定は、ニューロン間の結合係数に代表される多数のパラメータによって支えられている。しかし、このことが意外にも無視されがちであり、モデルの再現精度のみの議論に終わることが少なくない。

以上のことは、NNを地域分析モデルに応用する際に特に重要である。換言すれば、これらの問題点や課題への対応方法を一般化できれば、NNは地域分析を行うための強力な道具となる可能性を秘めている。このような観点から、筆者らはNNを地域分析に応用する際の意味解釈とそれに基づくNNの改良方法に関する研究に取り組んでいる。本稿では、立地選択モデルや交通機関分担モデルのような選択行動モデルへの応用に注目し、その代表的モデルである非集計ロジットモデル（以降単にロジットモデルと呼ぶ）との対比からその応用可能性を検討する。

2.ニューラルネットワーク基本型

本研究での議論のベースとする2階層NN（以降NN基本型と呼ぶ）について説明する。

（1）NN

NNでは、情報は入力層から出力層へと順次伝播し、ネットワークの出力値が求まる。その処理過程は以下のように表される。(図1参照）また今回、層数は入力層、出力層の2層としている。

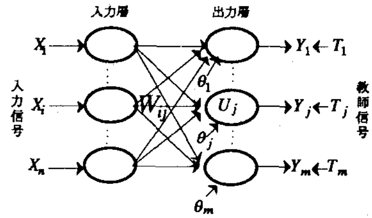

図1 NNモデル（2層）

a）入力層ユニットiは入力信号X_iを入力し、その値をそのまま出力する。

b）出力層ユニットjは入力層ユニットiの出力値X_iに重みW_{ij}（結合係数）を付けた値を入力値とし、その総和に出力層のユニットjのオフセット値θ_jを考慮した値U_j（内部状態）を求め、応答関数により出力値Y_jを算出する。応答関数には以下に示すような単調非減少のシグモイド関数を用いる。

$$\text{入力値の総和：} u_j = \sum_i W_{ij} \cdot X_i \tag{1}$$

$$\text{内部状態：} U_j = u_j - \theta_j \tag{2}$$

$$\text{出力値：} Y_j = \frac{1}{1+\exp(U_j)} \tag{3}$$

（2）BP法

NNでは、出力値と正しい出力（推定したい値：教師信号）との2乗誤差を最小化するように、結合係数及びオフセット値を最急降下法を用いて修正していく過程である。以下に概要を示す。

a）出力層ユニットjの出力値Y_jとそれに対する教師信号T_jとの2乗誤差を最小化するように、結合係数及びオフセット値を修正する。

$$\min_{w_{ij},\theta_j} E = \frac{1}{2}\sum_j (T_j - Y_j)^2 \tag{4}$$

b）ユニットiからユニットjへの結合係数W_{ij}の更新値ΔW_{ij}は、ηを定数として、以下のようになる。

$$\Delta W_{ij} = -\eta \cdot \frac{\partial E}{\partial W_{ij}} = \eta \cdot (T_j - Y_j) \cdot Y_j \cdot (1 - Y_j) \cdot X_i \tag{5}$$

またユニットjにおけるオフセット値θ_jの更新値$\Delta\theta_j$も、同様に求められる。

3.NN基本型とロジットモデルとの構造的関係

ここでは、ある選択主体がいくつかの要因$(X_i; i=1,n)$に基づき、どの選択肢$(j; j=1,m)$を選択するかというモデルを考える。

---

*キーワーズ：ニューラルネットワーク、選択行動モデル、ロジットモデル
**正員、工博、東京大学工学部土木工学科助教授
（東京都文京区本郷7-3-1、TEL03-3812-2111）
***学生員、東京大学大学院工学系研究科土木工学専攻修士課程（同上）

（1）NN基本型による表現

NNは画像分類への応用に関して蓄積が多い。これらの応用では一般に、a）NNの入力層には画素の特性の数のユニットを、また出力層には分類項目の数に対応したユニットを設ける、b）画素の分類項目に関する教師データを用意し、分類項目に合致す場合には1、そうでない場合には0を出力するようにBP法でNNを同定する、という方法を用いる。NNの選択行動モデルへの応用も基本的には分類の問題であり、画像分類への研究成果の還元も考慮して同様の定式化をしておく。

$$P_j = \frac{1}{1+\exp\left\{-(\sum_{i=1}^{n} W_{ij}\cdot X_i - \theta_j)\right\}} \qquad (6)$$

P_j：NNの出力値
W_{ij}：結合係数　　：NNの入力値
θ_j：出力層ユニット j のオフセット値

（2）ロジットモデルによる表現

ロジットモデルは、以下のように定式化できる。

$$P_j = \frac{1}{1+\exp\left\{-\left[V_j - \ln\sum_{k\neq j}^{m}\exp(V_k)\right]\right\}} \qquad (7)$$

P_j：選択肢 j の選択確率
V_j：選択肢 j から享受する効用
m：選択肢の数

また、効用関数には以下のような線形効用関数を仮定する。

$$V_j = \sum_k \theta_{kj}\cdot X_{kj} \qquad (8)$$

X_{kj}：選択肢 j の効用を表現する変数
θ_{kj}：X_{kj} に関する変数パラメータ

式(6)と式(7)を比較するにより、NNとロジットモデルの構造比較を行うと、次の対応関係が成り立つ。

$$\sum_{i=1}^{n} W_{ij}\cdot X_i - \theta_j = V_j - \ln\sum_{k\neq j}^{m}\exp(V_k) \qquad (9)$$

これは、出力層ユニットの内部状態が、ある選択肢効用と、それ以外の選択肢効用の期待値の差に相当していることを表している。[1)]

（3）出力総和制約付きNN[1)]

NNはロジットモデルと違い、出力値の総和が1になる、という構造的特性を持たない。そこでNNとロジットモデルとの意味関連をより明確にするために、NNに、

$$\sum_{j=1}^{m} P_j = 1 \qquad (10)$$

という制約を付けてみる。

このとき、いくつかの仮定のもとに、式(10)が成立する条件は次式であることがわかる。

$$\theta_j = \ln\sum_{k\neq j}^{m}\exp(\sum_{i=1}^{n} W_{ij}\cdot X_i) \qquad (11)$$

式(9),式(11)より、以下のような関係が成立する。

$$\sum_{i=1}^{n} W_{ij}\cdot X_i - \ln\sum_{k\neq j}^{m}\exp(\sum_{i=1}^{n} W_{ij}\cdot X_i) = V_j - \ln\sum_{k\neq j}^{m}\exp(V_k) \qquad (12)$$

この関係式より、次の関係を導くことができる。

$$\sum_{i=1}^{n} W_{ij}\cdot X_i = V_j + C \qquad (13)$$

$$\theta_j = \ln\sum_{k\neq j}^{m}\exp(V_k) + C \qquad C\text{：定数項} \qquad (14)$$

すなわち、

a）$\sum_{i=1}^{n} W_{ij}\cdot X_i$ は選択肢 j から享受する効用

b）θ_j は j 以外の選択肢から享受する効用の期待値に相当していることがわかる。

以降、この出力総和制約付きNNをNN選択モデルと呼ぶことにする。

4．変数特性の観点から見た NN選択モデルとロジットモデルとの比較

（1）変数の分類

効用関数には、選択主体の属性を表す変数（個人特性変数）と選択肢の特性を表す変数（選択肢特性変数）が含まれている。その内、選択肢特性を表す変数には、いずれの選択肢にも共通に含まれるもの（選択肢共通変数）と、特定の選択肢の効用関数に含まれるもの（選択肢固有変数）とがある。

（2）ロジットモデルでの変数

式(8)のような線形効用関数を用いたとき、各種変数及びそのパラメータをまとめたのが表1である。ここでは、選択肢を3種類（A,B,C）とし、説明変数として X_1（選択肢共通変数）、X_2（選択肢固有変数）、X_3（個人特性変数）を考えている。これは多項選択モデルの変数とパラメータに関する一般形と考えてよい。以降、この例に沿って議論を進めることにする。

表1　ロジットモデルでの変数、パラメータ表

$(V_j = \sum_k \theta_{kj}\cdot X_{kj})$

	変数	パラメータ	A	B	C
X_1	X_{1a}	θ_1	X_{1a}	0	0
	X_{1b}	θ_1	0	X_{1b}	0
	X_{1c}	θ_1	0	0	X_{1c}
X_2	X_{2a}	θ_{2a}	X_{2a}	0	0
	X_{2b}	θ_{2b}	0	X_{2b}	0
	X_{2c}	θ_{2c}	0	0	X_{2c}
X_3	X_3	θ_3	X_3	0	0
	X_3	θ_3'	0	X_3	0

また、ロジットモデルで算出される選択確率は選択肢間の効用の差のみに影響されており、効用の絶対的水準には影響を受けないことが知られている。

$$P_j = \frac{\exp(V_j)}{\sum_{j=1}^{m}\exp(V_k)}$$

$$=\frac{1}{\sum_{j=1}^{m}\exp(V_k-V_j)} \tag{15}$$

従って、推定の際に重要なのは、効用の絶対的水準ではなく、効用差V_k-V_jである。一例として、表1の選択肢A、Bに関する効用差は以下のように表現できる。

$$V_a-V_b=\theta_1(X_{1a}-X_{1b})+\theta_{2a}\cdot X_{2a}-\theta_{2b}\cdot X_{2b}+(\theta_3-\theta_3')\cdot X_3 \tag{16}$$

（3）NN選択モデルでの変数

式(13)より、NN選択モデルにおける入力値総和はロジットモデルでの効用項に相当していることがわかる。従って表1と同じ例を用いて、同様な変数表を作成する表2のようになる。但し、見易さの関係上W_{ij}をW_i^jと、以降表記している。

表2　NN選択モデルでの変数、パラメータ表

$$\left(V_j=\sum_{i=1}^{n}W_{ij}\cdot X_i\right)$$

	変数	パラメータ	A	B	C
X_1	X_{1a}	W_{1a}^a	X_{1a}	0	0
	X_{1a}	W_{1a}^b	0	X_{1a}	0
	X_{1a}	W_{1a}^c	0	0	X_{1a}
	X_{1b}	W_{1b}^a	X_{1b}	0	0
	X_{1b}	W_{1b}^b	0	X_{1b}	0
	X_{1b}	W_{1b}^c	0	0	X_{1b}
	X_{1c}	W_{1c}^a	X_{1c}	0	0
	X_{1c}	W_{1c}^b	0	X_{1c}	0
	X_{1c}	W_{1c}^c	0	0	X_{1c}
X_2	X_{2a}	W_{2a}^a	X_{2a}	0	0
	X_{2a}	W_{2a}^b	0	X_{2a}	0
	X_{2a}	W_{2a}^c	0	0	X_{2a}
	X_{2b}	W_{2b}^a	X_{2b}	0	0
	X_{2b}	W_{2b}^b	0	X_{2b}	0
	X_{2b}	W_{2b}^c	0	0	X_{2b}
	X_{2c}	W_{2c}^a	X_{2c}	0	0
	X_{2c}	W_{2c}^b	0	X_{2c}	0
	X_{2c}	W_{2c}^c	0	0	X_{2c}
X_3	X_3	W_3^a	X_3	0	0
	X_3	W_3^b	0	X_3	0
	X_3	W_3^c	0	0	X_3

式(16)と同様に、選択肢A、Bに関する効用差を計算すると以下のようになる。

$$\begin{aligned}V_a-V_b=&(W_{1a}^a-W_{1a}^b)X_{1a}+(W_{1b}^a-W_{1b}^b)X_{1b}+(W_{1c}^a-W_{1c}^b)X_{1c}\\&+(W_{2a}^a-W_{2a}^b)X_{2a}+(W_{2b}^a-W_{2b}^b)X_{2b}+(W_{2c}^a-W_{2c}^b)X_{2c}\\&+(W_3^a-W_3^b)X_3\end{aligned} \tag{17}$$

（4）NN選択モデルとロジットモデルとの比較

式(16)と式(17)とを比較すると、まずV_a-V_bに選択肢Cに関する項目の変数（X_{1c},X_{2c}）が入っていることがわかる。このことはNN選択モデルがロジットモデルとは異なり、選択独立を仮定していない選択行動モデルに相当していることを示している。

次に、各変数について両モデルの関係を検討する。

a）選択肢固有変数（X_2）

選択肢固有変数に関して、NN選択モデルの結合係数とロジットモデルのパラメータの対応関係をまとめると表3になる。

表3　NN選択ﾓﾃﾞﾙとﾛｼﾞｯﾄﾓﾃﾞﾙとの変数、ﾊﾟﾗﾒｰﾀ対応関係1

V_a-V_b	変数	NN選択ﾓﾃﾞﾙ	ﾛｼﾞｯﾄﾓﾃﾞﾙ
X_2	X_{2a}	$W_{2a}^a-W_{2a}^b$	θ_{2a}
	X_{2b}	$W_{2b}^a-W_{2b}^b$	$-\theta_{2b}$
	X_{2c}	$W_{2c}^a-W_{2c}^b$	0

b）個人特性変数（X_3）

個人特性変数に関しても表3と同様にまとめる。

表4　NN選択ﾓﾃﾞﾙとﾛｼﾞｯﾄﾓﾃﾞﾙとの変数、ﾊﾟﾗﾒｰﾀ対応関係2

V_a-V_b	変数	NN選択ﾓﾃﾞﾙ	ﾛｼﾞｯﾄﾓﾃﾞﾙ
X_3	X_3	$W_3^a-W_3^b$	$\theta_3-\theta_3'$

c）選択肢共通変数（X_1）

同様にロジットモデルにおいて選択肢共通変数として扱った変数に対し、どのようなパラメータが付与されるかを整理したものが表5である。

表5　NN選択ﾓﾃﾞﾙとﾛｼﾞｯﾄﾓﾃﾞﾙとの変数、ﾊﾟﾗﾒｰﾀ対応関係3

V_a-V_b	変数	NN選択ﾓﾃﾞﾙ	ﾛｼﾞｯﾄﾓﾃﾞﾙ
X_1	X_{1a}	$W_{1a}^a-W_{1a}^b$	θ_1
	X_{1b}	$W_{1b}^a-W_{1b}^b$	$-\theta_1$
	X_{1c}	$W_{1c}^a-W_{1c}^b$	0

ここで注目するべき点はBP法によって推定されたパラメータが、

$$W_{1a}^a-W_{1a}^b=-(W_{1b}^a-W_{1b}^b) \tag{18}$$

を満足する保証がないことである。

（5）NN選択モデルの改良の視点

表1・表2より、NN選択モデルでのパラメータ数はロジットモデルのそれと比較すると、極端に多くなっていることがわかる。通常のNNでは、オフセット値θ_jもパラメータとして扱うためにパラメータ数はさらに多くなる。これら多数のパラメータによって、NNでの高精度な推定が可能になっているのである。しかし、このNN選択モデルをそのまま選択行動モデルとして利用するのは、2つの意味で

問題が生じる場合がある。

1つは、NN選択モデルではパラメータ数が非常に多いため、サンプル数が十分でないとき予測モデルとしての利用は困難になるということ、もう1つは、NN選択モデルでは選択肢共通変数が扱えず、全て選択肢固有変数として扱われてしまうことである。このことは、立地選択を例にすると分かり易い。ある立地者が、従業地までのｱｸｾｽ手段が全て同じ鉄道であるA,B,Cという3つの地域を選択する際に、それらの所要時間t_a,t_b,t_cは選択肢共通変数として扱うのが一般には合理的であろう。しかしNN選択モデルでは、全て選択肢固有変数として扱われ、同じ鉄道であるのに所要時間の変化が効用に与える影響は選択肢間で異なる、ということになる。

ちなみにこのような問題は、画像分類への応用では生じない。画像分類では、サンプル数の増加は比較的容易であり、また分類項目に関する変数は一般には存在せず、全て画素に関する変数となるからである。

5.目的に応じたNN選択モデルの改良

(1)選択肢共通変数の導入

選択肢共通変数をNN選択モデルで表現するためには、どのような条件が必要なのか考えてみる。

まず、もう一度選択肢共通変数に関するNN選択モデルとロジットモデルとの関係についてまとめる。

表6　NN選択ﾓﾃﾞﾙとﾛｼﾞｯﾄﾓﾃﾞﾙの変数、ﾊﾟﾗﾒｰﾀ対応関係4

	効用差	変数	NN選択ﾓﾃﾞﾙ	ﾛｼﾞｯﾄﾓﾃﾞﾙ
X_1	V_a-V_b	X_{1a}	$W_{1a}^a-W_{1a}^b$	θ_1
		X_{1b}	$W_{1b}^a-W_{1b}^b$	$-\theta_1$
		X_{1c}	$W_{1c}^a-W_{1c}^b$	0
	V_b-V_c	X_{1a}	$W_{1a}^b-W_{1a}^c$	0
		X_{1b}	$W_{1b}^b-W_{1b}^c$	θ_1
		X_{1c}	$W_{1c}^b-W_{1c}^c$	$-\theta_1$
	V_a-V_c	X_{1a}	$W_{1a}^a-W_{1a}^c$	θ_1
		X_{1b}	$W_{1b}^a-W_{1b}^c$	0
		X_{1c}	$W_{1c}^a-W_{1c}^c$	$-\theta_1$

表6から、NN選択モデルで選択肢共通変数を表現するためには、

$$\begin{aligned} &W_{1a}^a-W_{1a}^b=-(W_{1b}^a-W_{1b}^b)\\ &=W_{1b}^b-W_{1b}^c=-(W_{1c}^b-W_{1c}^c)\\ &=W_{1a}^a-W_{1a}^c=-(W_{1c}^a-W_{1c}^c)\end{aligned} \tag{19}$$

という結合係数に関する制約が必要になってくる。

式(19)は、以下のように変形できる。

$$\begin{aligned} W_{1a}^a+W_{1b}^a&=W_{1a}^b+W_{1b}^b\\ W_{1b}^b+W_{1c}^b&=W_{1b}^c+W_{1c}^c\\ W_{1a}^a+W_{1c}^a&=W_{1a}^c+W_{1c}^c\end{aligned} \tag{20}$$

$$\begin{aligned} W_{1c}^a&=W_{1c}^b\\ W_{1a}^b&=W_{1a}^c\end{aligned} \tag{21}$$

(2)BP法の改良によるパラメータ推定

制約式(21)は、表6から、NN選択モデルに関してV_a-V_bのX_{1c}の係数を0、V_b-V_cのX_{1a}の係数を0にする条件であることがわかる。即ち式(21)は、選択独立の仮定と類似しており、その仮定を弱めたものとなっている。(選択独立の仮定を設けた場合には、さらに$W_{1b}^a=W_{1b}^c$という条件が出てくる)

本稿では、BP法のアルゴリズムを簡便に改良することを念頭に置いて、多少強すぎる仮定ではあるが、選択独立の仮定の下でのパラメータ推定方法を示すことにする。

NN選択モデルに選択独立の仮定を設けた場合、以下のような結合係数に関する制約式が導出される。(表3,4,5,6参照)

$$\begin{aligned} W_{1a}^b=W_{1a}^c,W_{1b}^a=W_{1b}^c,W_{1c}^a=W_{1c}^b\\ W_{2a}^b=W_{2a}^c,W_{2b}^a=W_{2b}^c,W_{2c}^a=W_{2c}^b\end{aligned} \tag{22}$$

式(22)を一般的な表現にすると、X_{ip}が選択肢特性要因である時、次のように表現できる。

$$\begin{aligned} W_{ip}^q&=W_{ip}^q \quad (p=q)\\ &=W_{ip} \quad (p\neq q)\end{aligned} \tag{23}$$

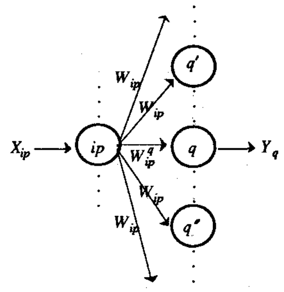

図2　改良NN選択モデル

この改良NN選択モデルは、通常の2層NNモデルの場合と同様に、教師信号T_qと出力値Y_qとの2乗誤差Eから結合係数W_{ip}^qの更新値ΔW_{ip}^qを次式のように計算できる。

$$\begin{aligned} \Delta W_{ip}^q&=\eta\cdot(T_q-Y_q)\cdot Y_q\cdot(1-Y_q)\cdot X_{ip}\\ \Delta W_{ip}&=\eta\cdot\sum_{q\neq p}(Y_q-T_q)\cdot Y_q\cdot(1-Y_q)\cdot X_{ip}\end{aligned} \tag{24}$$

6.今後の課題

一般形である3層NNへの議論の拡張及び、パラメータ推定法が異なることに対する意味解釈等が今後の課題である。

参考文献

1)清水英範,宮城俊彦,森光正:ニューラルネットワークの空間相互作用モデルへの適用可能性,土木計画学研究講演集,No16,pp343～348,1993