

IV-70

立地モデルへの応用を前提としたニューラルネットワークの改良

東京大学工学部 正会員 清水英範

岐阜大学工学部 正会員 宮城俊彦

岐阜県庁 正会員 森 光正

1. はじめに

階層型ニューラルネットワーク(以後、NN)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強力な手法である。近年では、NN関係の専門書等を通してプログラムも流通しており、多変量解析の種々の手法やロジットモデルなどと同様に身近な実用的道具となりつつある。

しかし、NNの利用に際しては、これまでの応用研究等を通して次のような問題点や課題が指摘されていることを念頭におく必要がある。

① NN全体の入出力関係を数式として記述するとき極めて複雑になる。したがって、推定モデルとしてのNNに意味解釈を与えることが困難となり、その利用はブラックボックス的になる。

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、このことが意外と無視されがちである。推定モデルとしてのNNの汎化性の検討は課題である。

以上のことは、立地モデルのような地域分析モデルに適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を改善するようにNNを改良することができれば、NNは立地モデルなどを構築するための強力な道具となる可能性を秘めている。

筆者らは、以上の観点から、簡単な立地モデルを例にNNに意味解釈を与え、それを踏まえたNNの発展的改良を試みている。本稿では、これまでの検討の成果の概要について示す。

2. NNによる立地モデルの基本型¹⁾

(1) 想定する立地モデル

本研究では、各ゾーンの従業者数(X_i)が与えられた時に、交通条件等の立地条件を考慮して各ゾーンの人口(T_j)を推定する居住立地モデルを考える。

(2) NNの基本型

ネットワークは、入力層と出力層のみの2層構造と

し、各々のユニット数はゾーンの数と一致させる。また、出力範囲は0から対象地域の総人口(N)までとし、1セットの入力データ(従業者数)と教師データ(人口)によりBP法による学習を行うものとする。

この2層NNによる立地モデルは以下のように定式化される。

$$Y_j = \frac{N}{1 + \exp\left\{-\left(\sum_i W_{ij} X_i - \theta_j\right)\right\}} \quad (1)$$

ここで、 Y_j : NNの出力(j ゾーンの人口推定値)

$\sum_i W_{ij} X_i - \theta_j$: ユニットの内部状態

W_{ij} : 結合係数, θ_j : オフセット

このとき、学習は以下のようになされ、

$$\min. E = \frac{1}{2} (T_j - Y_j)^2 \quad (2)$$

例えば、 W_{ij} の修正量 ΔW_{ij} は以下になる。

$$\begin{aligned} \Delta W_{ij} &= -\eta \frac{\partial E}{\partial W_{ij}} \\ &= \eta (T_j - Y_j) Y_j (1 - Y_j/N) X_j \end{aligned} \quad (3)$$

ここで、 η : 1回の修正量を決めるパラメータ

さて、一般的なNNの適用では、中間層を入れた3層以上の構造とし、またモデルのすべての説明変数を入力層への入力変数とする。本研究で上記のような簡単な構造のNNを基本型としたのは以下のような理由による。

2階層にした理由

① NNの一般的な入出力関係はロジスティック関数の合成関数で記述されるため、ロジットモデルとの対比の上で意味解釈が与えられそうである。このとき、構造が簡単な2階層の方が比較対照が容易である。

② 2階層でかつ各ユニットにゾーンを対応させることによって、結合係数は交通条件等のゾーン間の相互作用を表すパラメータとして解釈できそうである。

入力層への入力を従業者数のみにした理由

①予測モデルとして必要最小限の変数である従業者数のみを入力変数とすることにより、構造を簡略化でき比較が容易になる。

②NNの意味解釈が可能になれば、交通条件等の立地条件は結合係数を関数化することによって取り込むことができそうである。

3. NN基本型の改良

上記の検討を踏まえ、ロジットモデルとの比較からNN基本型の改良を試みる。

(1) 改良の方針

①NNとロジットモデルがモデル構造的に最も異なるのは、NNでは出力の総和に対して以下のような制約を持たないことである。

$$\sum_j (Y_j/N) = 1 \quad (4)$$

すなわち、NNではコントロールトータルが与えられるような場合でもそれを明示的に考慮できない。そこで、ロジットモデルとの比較を明確にすることも踏まえて、式(4)を満足するようにNNを改良する。

②上記のNN基本型の最大の問題点は、予測に際して立地条件を考慮できないことである。そこで、結合係数を交通条件等の立地条件の関数とし、これを教師データから同定できるように改良する。

結合係数の関数化は、意味解釈を明確にした上でパラメータ数を調整できるということを意味し、AICなどをを用いた汎化性の検討にも柔軟に対応できる。

(2) 出力に総量制約を与えたNN

いくつかの仮定のもとに、式(4)を満足する必要十分条件は次式であることが分かった。

$$\theta_j = \ln \sum_{k \neq j} \exp \left(\sum_i W_{ik} X_i \right) \quad (5)$$

これにより、ロジットモデルとの対応が明確になった。すなわち、 $\sum_i W_{ij} X_i$ がjゾーンへの立地効用を表すと解釈できる。

NN基本型は以下のように改良されたことになる。

$$Y_j = \frac{N}{1 + \exp \left\{ - \left(\sum_i W_{ij} X_i - \ln \left[\sum_{k \neq j} \exp \left(\sum_i W_{ik} X_i \right) \right] \right) \right\}} \quad (6)$$

このとき、 W_{ij} の修正量 ΔW_{ij} は以下ようになる。

$$\Delta W_{ij} = \eta (T_j - Y_j) Y_j (1 - Y_j / N) X_j \quad (7)$$

式(3)から分かるようにBP法のアルゴリズムをそのまま用いることができる。

(3) 結合係数を関数化したNN

いま、結合係数が交通条件(t_{ij})と、その他の立地条件(A_j)によって以下のように定義したとしよう。

$$W_{ij} = \alpha A_j \beta t_{ij}^\gamma \quad (8)$$

ここで、 α, β, γ : パラメータ

すなわち、NN基本型は式(6)、(8)により、 α, β, γ の3つのパラメータしかもたないNNへと改良されたことになる。

このとき、例えばパラメータ $\Delta \alpha$ の修正量は、以下のようになる。

$$\Delta \alpha = -\eta \frac{\partial E}{\partial \alpha} \\ = \eta \sum_j \delta_j \left\{ \sum_i M_{ij} \frac{\sum_{k \neq j} [(\sum_i M_{ik}) \exp(\sum_i \alpha M_{ik})]}{\sum_{k \neq j} \exp(\alpha M_{ik})} \right\} \quad (9)$$

ここで、 $\delta_j = (T_j - Y_j) Y_j (1 - Y_j / N)$

$$M_{ij} = \alpha A_j \beta t_{ij}^\gamma X_j$$

4. おわりに

本研究では、ロジットモデルとの対比から、NNの立地モデルへの応用に意味解釈を与えた。また、この結果を踏まえ、①出力の総量制約を与えたNN、②結合係数を関数化したNN、という2つの改良NNを提示し、その学習が通常のBP法のプログラムの若干の修正によって行えることを示した。これらの改良は、立地モデルのみならず広く応用可能なものであり、またその有用性も高いと考える。

(参考文献)

1) 清水、宮城、森：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16、pp.343-348、1993。