

IV-30

ニューラルネットワークを用いた  
都市成長過程の分析

室蘭工業大学	学生員	○佐々木	恵一
室蘭工業大学	正員	田村	亨
苫小牧工業高等専門学校	正員	榎谷	有三
室蘭工業大学	正員	斉藤	和夫

1. はじめに

都市成長過程をモデル化した研究は、計量経済学や都市経済学の分析で、これまでも開発されている。特に、交通網整備との関係で分析した研究は1970年代に数多くなされている<sup>1)</sup>。しかし、これらの方法は、都市成長過程の構造が明確な場合には有効な方法であるが、現在のところ理論モデルの枠を出ていない。本研究で扱うニューラルネットワーク（以下NNと呼ぶ）モデルは、外部環境に合うように自己組織化能力を活用して、過去の入出力の結果のみから自立的に学習を行うという、いわば構造をブラックボックス化して解析する方法である。この手法を地域構造分析に用いた先進研究は清水ら<sup>2)</sup>によってなされている。

本研究は、NNモデルを帯広都市圏の都市成長過程分析に用いたものであり、昭和55年、昭和60年、平成2年、平成5年の4時点について住区単位で収集したデータを用いて分析している。本研究は、この適用可能性を検討した基礎分析であるが、目指すところは、交通基盤整備が都市成長過程にどの様に影響していくのかを把握したいと考えている。

2. 本研究で扱うNNモデルの構造<sup>3)</sup>

(1) 多層（階層）モデルについて

多層NNは、人間の神経細胞の人工的なモデルであるユニットが、神経繊維に対応する線で結ばれ（シナプス結合）、網目状のネットワークを形成する。ユニットは、多入力1出力である。出力は重み ( $W_{ij}$ ; 結合係数) を付け、他の入力となる。ユニットは入力の総和にユニット固有の値 ( $\theta_j$ ; オフセット値) を考

慮した値 ( $u_j$ ; 内部状態) を求め、これを応答関数によって出力する。応答関数は、以下に示す準線形の飽和型の応答特性を持つシグモイド関数を用いた。

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (1)$$

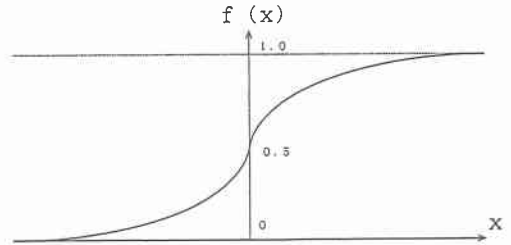


図1 シグモイド関数

本研究では、各ゾーンの商業延べ床 ( $X_1$ )、駅までの直線距離 ( $X_2$ )、公示地価 ( $X_3$ ) を所与として、各ゾーン人口を推定 (教師信号=人口;  $T_n$ ) する分析モデルへの多層NNの適用を試みた。

よって、本研究で扱う具体的な階層は、入力層が1つ、中間層が3つで3-3-1で並んでおり、出力層が1つの3層構造にした (図2)。

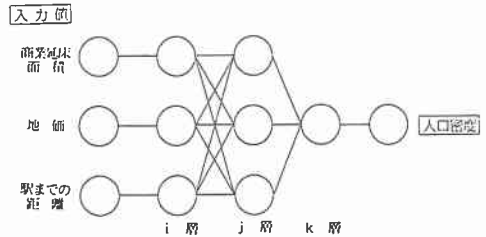


図2 階層NNモデル

Application of NN Model to Urbanization Process

by Keiichi SASAKI, Tohru TAMURA, Yuzou MASUYA, Kazuo SAITO

入力層のユニットは入力信号 $X_{(1,2,\dots)}$ を入力し、その値をそのまま出力する。第一中間層のユニット $i$  ( $i = 1, 2, \dots$ ) は、入力信号をそのまま受ける。第二中間層ユニット $j$ は、第一中間層ユニット $i$ 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、その総和に第二中間層ユニット $j$ の固有の値 $\theta_j$ を引いた値 $U_j$  (内部状況) を求め、応答関数により出力値 $Y_j$ を求める。すなわち、中間層ユニット内で行われている作業を定式化すると、以下のようになる。

$$u_j = \sum_i W_{ij} X_i \quad (2)$$

$$U_j = u_j - \theta_j \quad (3)$$

$$Y_j = f(U_j) \quad (4)$$

第三中間層ユニット $k$ は第二中間層ユニット $j$ からの出力値 $Y_j$ を受け、内部で同様の作業を繰り返す。

$$u_k = \sum_j W_{jk} Y_j \quad (5)$$

$$U_k = u_k - \theta_k \quad (6)$$

$$Y_k = f(U_k) \quad (7)$$

### (2) バックプロパゲーションについて

ニューラルネットワークの自己組織化とは、出力値と正しい出力値 (教師信号) との2乗誤差を最小にするように結合係数とオフセット値を補正しながら最適なネットワークを求めることである。

出力層ユニットの出力値 $Y_n$ とそれに対する教師信号 $T_n$ との2乗誤差は式8となる。

$$E = \frac{\sum_n (T_n - Y_n)^2}{2} \quad (8)$$

第二中間層ユニット $j$ から第三中間層ユニット $k$ への結合係数 $W_{jk}$ の更新値 $\Delta W_{jk}$ 、式9となる。

$$\begin{aligned} \Delta W_{jk} &= -\eta \left( \frac{\delta E}{\delta W_{jk}} \right) \\ &= \eta (T_n - Y_n) Y_n (1 - Y_n) Y_j \quad (9) \end{aligned}$$

( $\eta$ : 一回の更新量を定める定数)

また、第三中間層ユニット $k$ におけるオフセットと値 $\theta_k$ の更新値 $\Delta \theta_k$ は同様に、式10となる。

$$\begin{aligned} \Delta \theta_k &= -\varepsilon \left( \frac{\delta E}{\delta \theta_k} \right) \\ &= \varepsilon (T_n - Y_n) Y_n (1 - Y_n) (-1) \quad (10) \end{aligned}$$

( $\varepsilon$ : 一回の更新量を定める定数)

### 3. ニューラルネットワークの都市成長過程分析モデルへの適用

#### (1) 対象地域と使用データ

対象地域は帯広市であり、市全域を22に分割した住区を単位とした。

使用データは、昭和55年、昭和60年、平成2年、平成5年の4時点についての住区ごとの夜間人口、商業延べ床、公示地価と駅との直線距離である。商業延べ床面積については、町丁目にまとめられた帯広市基礎データをもとに住区単位で集計してデータとした。公示地価は帯広市内にある約30の調査地点をもとに、等地価線図を作成し、これをもとに住区面積との関係から住区ごとの平均公示地価を算出した。

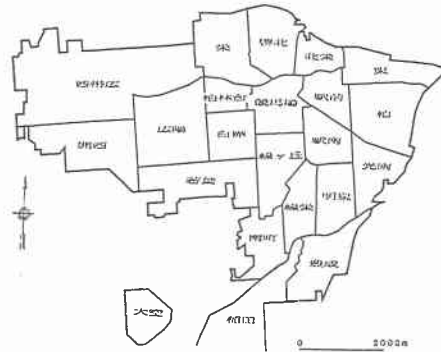


図3 帯広市全域図

#### (2) 分析について

本研究では、4時点22住区のデータを用いて、NNモデルを構築することを目的としていたが、この88個のデータをもとにした分析では、数十万回の学習回数でも集束ができなかったため、どのようなデータ形式によって安定的にNNモデルが構築できるかを検討することにした。このため、3種類の分析を行っている。

①帯広市住居地区 (12住区) の1時点 (昭和55年) のデータによる分析

②多地点 (3地点) 1時点のデータによる分析

③1地点4時点のデータによる分析

分析2) の3住区は、帯広市の旧市街地で、市街化が安定していると考えられる川北地区の3住区 (北栄、啓北、栄) である。

分析における計算打ち切り基準は、学習回数100,000回に達した場合とした。

以下に3つの分析結果をまとめる。

1) 12住区1時点データによる分析

表1 多地点1時点分析 (平成5年)

対象住区	人口密度 人/km <sup>2</sup>	入 力 値			誤 差	
		2814	3.45	25.5	103	
東 南	4892	2814	3.45	25.5	103	
光 明	6498	2387	2.10	43.0	128	6
緑 葉	5610	3624	2.25	37.5	43	7
競 馬	4537	641	2.68	32.5	37	6
緑 葉	3524	2557	2.15	40.0	168	3
白 旗	4541	2606	2.54	31.5	52	7
日 林	5690	1881	3.93	23.4	28	3
北 栄	7886	436	2.93	32.0	21	8
啓 北	5770	4164	2.70	55.5	21	6
栄 北	3939	2852	4.08	38.5	12	5
栄	4136	2992	5.13	30.5	9	1
雨 町	4501	760	4.65	19.5	4	3

・入力値は左側より商業延べ床面積 (m<sup>2</sup>)、駅までの直線距離 (km)、公示地価 (千円) を示す。  
 ・誤差は実測値と予測値の差の絶対値を示す。また、10,000回学習結果である。

平成5年のデータを12住区においてNNモデルにてはめた結果を表-1で表す。表中の人口密度は入力した教師信号であり、次の入力値は左側より商業延べ床面積、駅までの直線距離、公示地価を示す。誤差は実測値と予測の残差の絶対値を示す。この表より以下のことが分かった。

帯広市全域を対象にして分析した結果、誤差に大きなばらつきが現れた。最小では2.1%の誤差(東住区)であるが、最大では47.8%の誤差(競馬場住区)であった。また、学習回数が10,000回に到達した後、誤差の集束がみられなかったため、計算はここで打ち切ることとした。

2) 3住区1時点データによる分析

分析結果を表-2(昭和55年)、表-3(昭和60年)、表-4(平成2年)、表-5(平成5年)に示す。

表2 3地点1時点分析 (昭和55年)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤 差	
北 栄	5770	4164	189	109
		2.70		
		55.5		
啓 北	3939	2852	489	402
		4.08		
		38.5		
栄	4136	2992	295	301
		5.13		
		30.5		

・入力値は上段より商業延べ床面積 (m<sup>2</sup>)、駅までの直線距離 (km)、公示地価 (千円) を示す。  
 ・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

表3 3地点1時点分析 (昭和60年)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤 差	
北 栄	5597	4303	7	49
		2.70		
		80.5		
啓 北	4171	5632	125	118
		4.08		
		33.5		
栄	4488	3520	132	125
		5.13		
		35.5		

・入力値は上段より商業延べ床面積 (m<sup>2</sup>)、駅までの直線距離 (km)、公示地価 (千円) を示す。  
 ・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

表4 3地点1時点分析 (平成2年)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤 差	
北 栄	4423	6318	11	50
		2.70		
		50.0		
啓 北	4208	6223	262	253
		4.08		
		21.0		
栄	4740	5496	273	267
		5.13		
		22.0		

・入力値は上段より商業延べ床面積 (m<sup>2</sup>)、駅までの直線距離 (km)、公示地価 (千円) を示す。  
 ・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

表5 3地点1時点分析 (平成5年)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤 差	
北 栄	4423	6318	180	50
		2.70		
		75.0		
啓 北	4208	7937	320	267
		4.08		
		48.0		
栄	4740	5766	150	230
		5.13		
		45.0		

・入力値は上段より商業延べ床面積 (m<sup>2</sup>)、駅までの直線距離 (km)、公示地価 (千円) を示す。  
 ・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

誤差は最大で10,000回学習後に12.4%、100,000回学習後に10.2%(昭和55年、啓北住区)となっており、またはほぼ安定的に収束している。表中の誤差項において、10万回学習後の方が1万回よりも小さい場合もみられる。全体的な誤差を表す平均誤差は明らかに学習回数を多くすると小さくなるが、それを個別にみると逆転している場合がある。これは1万回から10万回に学習回数を増加させても、収束状況に大きな変化

がないためであろう。また、年次ごとに10万回学習後の平均誤差を比較すると、5.0%、2.0%、4.3%、4.1%となっており、時点による変化がある。

### 3) 1地点4時点のデータによる分析

分析結果を表-6(北栄住区)、表-7(啓北住区)、表-8(栄住区)に示す。

表6 1地点4時点分析(北栄住区)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤差	
昭和55年	5770	4146 2.70 55.5	729	813
昭和60年	5597	4303 5.13 80.5	565	424
平成 2年	4423	6318 2.70 50.0	696	529
平成 5年	4423	6318 2.70 75.0	600	703

・入力値は上段より商業延べ床面積(m<sup>2</sup>)、駅までの直線距離(km)、公示地価(千円)を示す。  
・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

表7 1地点4時点分析(啓北住区)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤差	
昭和55年	3939	2852 4.08 38.5	184	170
昭和60年	4171	5632 4.08 33.5	21	50
平成 2年	4208	6223 4.08 21.0	17	50
平成 5年	4208	7937 4.08 48.0	144	116

・入力値は上段より商業延べ床面積(m<sup>2</sup>)、駅までの直線距離(km)、公示地価(千円)を示す。  
・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

表8 1地点4時点分析(栄住区)

対象年度	人口密度 人/km <sup>2</sup>	入力値	誤差	
昭和55年	4136	2992 5.13 30.5	407	388
昭和60年	4488	3520 5.13 35.5	40	40
平成 2年	4740	5496 5.13 22.0	211	227
平成 5年	4740	5766 5.13 45.0	233	162

・入力値は上段より商業延べ床面積(m<sup>2</sup>)、駅までの直線距離(km)、公示地価(千円)を示す。  
・誤差は実測値と予測値の差の絶対値を示す。また、左側が10,000回学習結果、右が100,000回である。

誤差は最大で10,000回学習後に12.6%、100,000回学習後に14.1%(北栄住区)となっており、またほぼ安

定的に収束している。住区ごとに10万回学習後の平均誤差を比較すると、9.3%、12.2%、4.5%となっており、住区ごとによる都市発展過程が異なることが顕著に示されている。

### 4. おわりに

本研究の分析により分かったことは次のことである。

- ①住居地域について分析した結果でも収束しない。
- ②3住区1時点、及び1住区4時点のデータを利用した場合は、10,000回の学習でも誤差10%前後まで収束する。
- ③3住区1時点の方が1住区4時点よりも収束状況がよい。

今後の課題として、NNモデルの収束を高める方策として以下のことがあげられる。1つは本研究ではモデルを3-3-1の構造のみに限定していたが、他の構造を持つネットワークで適合可能性を検討すべきである。2つめは、入力データの種類についてである。例えば、通勤や買物トリップのODデータの活用が考えられる。3つめは、ゾーンの大きさであり、住区という単位で分析を進めることが良いか否かの検討である。

### 謝辞

本研究においては、ニューラルネットワークのプログラムを、北海道大学土木工学科中辻隆助教授に提供して頂いた。また、データ収集時に帯広市役所、幕別町役場、音更町役場、芽室町役場の方々に多大なご協力をして頂いた。ここに記して感謝いたします。

### <参考文献>

- 1) 例えば肥田野登 「開発プロセスを考慮した地域整備過程に関する研究」 地域学研究 1979年
- 2) 清水英範 「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 1993年3月 pp343~348
- 3) 中野馨、飯沼一元、ニューロンネットグループ、桐谷滋 「ニューロコンピュータ 入門と実習」 技術評論社