Ⅳ-38　実験的回帰分析法による発生交通量推計モデルの構築

北海道大学　正員　○佐藤　馨一
労働省　〃　五十嵐　力
北海道大学　〃　山形　耕一

1．はじめに、　交通計画において用いられるモデルは、現実の問題を解決するために構築するものであって、モデルと現象との関係を網羅的に記述するものではなく、現象を抽象化した上でその論理構造を明確化し、さらに操作性や一致性でも優れたものでなければならない。しかしながら実際のモデル・ビルディングにおいては、一致性を重じるために論理性に矛盾が生じたり、論理性や一致性を重視するあまり操作性において種々の障害が認められるケースが多い。たとえば、発生交通量の推計モデルとして最も一般的に用いられる重回帰モデルを考えると、説明変数を加えていく過程において偏回帰係数の符号が変化したり、寄与率を高めることに注目するあまり、必要以上に多くの変数を選択しがちな傾向は従来から指摘されてきたとおりである。

本研究は重回帰モデルのもつこのような問題点を解決するために、実験的回帰分析法を用いて偏回帰係数の推計を行なったものであり、実験的回帰分析法の有する意義と課題について検討を加えたものである。

2．実験的回帰分析法によるモデル構築のプロセス[1]

実験的回帰分析法は重回帰モデルの偏回帰係数を実験的にシミュレーションによって求める方法であり、その基本概念は次のようになっている。いま、P個の説明変数からなる重回帰モデルを(1)式のように記述する。

$$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_Px_P \qquad (1)$$

重回帰モデル式においては、偏回帰係数$a_0, a_1, \cdots a_P$を最小二乗法を用いて推定しているが、これは偏回帰係数が区間$(-\infty, \infty)$のすべての値をとり得るということを前提条件としている。一方、実験的回帰分析法の場合には、偏回帰係数のとりうる値をあらかじめ設定し、その範囲内で残差平方和を最小にする偏回帰係数を求めようとするものである。図・1は実験的回帰分析法によって偏回帰係数を推定する一般的プロセスを示したものであり、説明変数が4個ある重回帰モデルの場合には以下のとおりとなる。

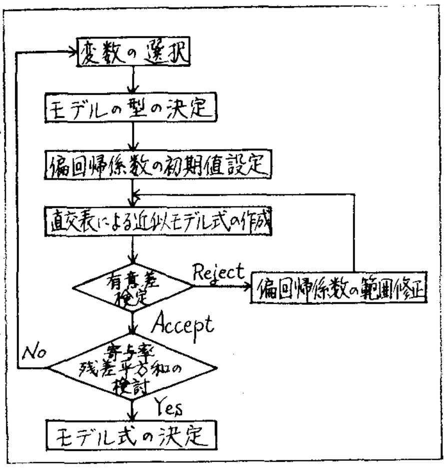

図1　モデル決定のプロセス

i) 偏回帰係数の初期値設定 ：　このステップにおいては、偏回帰係数a_Pの範囲を現実感覚に即して限定し、その範囲を3等分して表・1を作成する。

ii) 直交表による近似モデル式の作成 ：　表・1で設定した初期値を表・2に示すL9の直交表に割り付ける。このとき、(2)式で示すような9種類の近似モデル式が作成される。

$$\left.\begin{array}{ll} No1 & \hat{y} = a^0_{11}x_1 + a^0_{21}x_2 + a^0_{31}x_3 + a^0_{41}x_4 \\ No2 & \hat{y} = a^0_{11}x_1 + a^0_{22}x_2 + a^0_{32}x_3 + a^0_{42}x_4 \\ \vdots & \\ No9 & \hat{y} = a^0_{13}x_1 + a^0_{23}x_2 + a^0_{32}x_3 + a^0_{41}x_4 \end{array}\right\} \quad (2)$$

iii) 残差ε，変動Sの計算 ：　外的基準yの実際値と、(2)式から求められる$\hat{y}$との差をとり、残差εを計算する。　$\varepsilon_{ij} = (y_{ij} - \hat{y}_{ij})$　　(3)

次に残差の総和Tと残差の変動Sを(4)式，(5)式から計算する。

表1　偏回帰係数の初期値

パラメータ＼水準	1	2	3
a_1	a^0_{11}	a^0_{12}	a^0_{13}
a_2	a^0_{21}	a^0_{22}	a^0_{23}
a_3	a^0_{31}	a^0_{32}	a^0_{33}
a_4	a^0_{41}	a^0_{42}	a^0_{43}

表2　L9 直交表

No.＼列番	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

iv) 変動Sを各偏回帰係数について水準別に計算する。たとえば、偏回帰係数a_1について計算式を示すと(6)式のようになる。ただし、(6)式中に示したSの添字は直交表の組合せNoと対応している。

v) 有意差検定 ： 水準別Sの和を用いて、水準間の有意差検定を行なう。a_{11}°とa_{12}°の有意差検定は、(7)式から分散比F_0を計算し、F_0が2〜4以上の場合を有意差ありと判定する。ただし、(7)式中のV_eは最小二乗解の誤差分散である。

$$T_j = \sum_{i=1}^{n} \varepsilon_{ij} \tag{4}$$

$$S_j = \sum_{i=1}^{n} (\varepsilon_{ij} - \bar{\varepsilon}_j) = \sum_{i=1}^{n} \varepsilon_{ij}^2 - T_j^2/n \tag{5}$$

ただし i：サンプル番号　j：直交表の組合せ番号

$$\left.\begin{aligned} S(a_{11}^\circ) &= a_{11}\text{のSの和} = S_1+S_2+S_3 \\ S(a_{12}^\circ) &= a_{12}\text{のSの和} = S_4+S_5+S_6 \\ S(a_{13}^\circ) &= a_{13}\text{のSの和} = S_7+S_8+S_9 \end{aligned}\right\} \tag{6}$$

$$F_0 = |S(a_{11}^\circ) - S(a_{12}^\circ)| / 3V_e \tag{7}$$

vi) モデル式の決定 ： 有意差の状況により偏回帰係数の範囲を修正し、上記のステップを繰返す。偏回帰係数の水準別の有意差が認められなくなった時点で計算を打切り、モデル式を決定する。

3. 実験的回帰分析法の適用と意義

実験的回帰分析法の特徴を明らかにするため、昭和46年，52年度の自動車OD調査結果表（北海道）を用いて発生交通量推定モデルの構築を行なった。表・3は昭和46年度のOD表をもとに、ステップ・ワイズ法により変数選択を行ない、最小二乗法により偏回帰係数を推定したモデルを示したものである。また、表・4は実験的回帰分析法を適用したときの発生交通量推計モデルである。これらの表から次のことがわかった。

表3　最小二乗法による結果（昭和46年）（全目的）

step	モデル式	寄与率
1	$y = 31656.5 + 3.41x_1$	92.19
2	$y = 28825.1 + 6.50x_1 - 5.92x_2$	94.75
3	$y = 6935.7 + 6.49x_1 - 5.89x_2 + 14.05x_3$	94.99
4	$y = 5023.2 + 6.62x_1 - 6.66x_2 + 14.41x_3 + 0.21x_4$	95.05
5	$y = 12221.5 + 7.10x_1 - 6.42x_2 + 15.34x_3 + 0.25x_4 - 0.22x_5$	95.07

表4　実験的回帰分析による結果（昭和46年）（全目的）

step	モデル式	寄与率
1	$y = 31464.1 + 3.42x_1$	92.18
2	$y = -592.3 + 3.48x_1 + 2.45x_2$	93.08
3	$y = -5231.7 + 3.48x_1 + 2.45x_2 + 3.00x_3$	92.98
4	$y = -9499.0 + 3.19x_1 + 2.50x_2 + 3.05x_3 + 0.32x_4$	92.54
5	$y = -220853.3 + 3.01x_1 + 2.40x_2 + 3.00x_3 + 0.30x_4 + 1.40x_5$	86.08

y：全目的発生交通量
x_1：第3次従業者人口
x_2：第2次従業者人口
x_3：第1次従業者人口
x_4：工業出荷額
x_5：夜間人口

表5　最良モデル式の適合性比較

年度	昭和46年	
手法	最小二乗法	実験的回帰分析
寄与率	94.99(%)	93.08(%)
残差平方和	1.44×10^{11}	1.98×10^{11}
相関係数	0.9746	0.9642

1) 表・3からも明らかなように、最小二乗法を用いた場合には第3ステップにおいてx_2の偏回帰係数が負となる。しかしながら、実験的回帰分析法によって推定された偏回帰係数は負とはならず、モデルの論理性が保たれている。

2) 最小二乗法を用いて構築された発生交通量推定モデルは、説明変数が多くなるにつれ寄与率も一様に増加する。しかしながら、実験的回帰分析法を用いて構築したモデルは、説明変数の増加は第2ステップまでであり、それ以降は説明変数が増加するにつれて寄与率も低下する。この傾向昭和52年度のデータについても認められステップ・ワイズ法を用いて説明変数の選択を行なうのに比較して、実験的回帰分析法による説明変数の選択は非常に容易である。

3) ステップ・ワイズ法による回帰式は第3ステップのものであり、実験的回帰分析法の場合には寄与率が最大になる第2ステップの回帰式が最良となる。これらの回帰式の適合性を調べてみると、実験的回帰分析法によるモデルは最小二乗法で構築されたモデルに比べて多少寄与率は小さくなっているが、ほぼ同等の適合性を有していると判断される。

参考文献　1) 田口玄一、横山巽子：ビジネスデータの分析－手法と実例－，丸善，1975