発生重回帰モデルにおける説明変数選択法に関する研究

北海道大学 工学部 正員 ○ 佐藤 馨一
建設省中国地建 〃 桐越 信
静岡県土木部 〃 岩立 忠夫

1. はじめに

発生交通量の推計は交通計画の最も基本的なプロセスであり、種々の手法，モデルが研究されてきた。表-1は昭和42年から昭和50年までに我国で実施されたパーソントリップの実施箇所と，そのときに採用された発生交通量の推計モデルや経済指標を取りまとめたものである。表-1からも明らかなように，発生交通量の推計法としては大部分の都市圏において重回帰モデル法が採用されており、説明変数は広島都市圏を除いた他の都市圏すべてにおいて人口関連指標が用いられている。

表1　全国各都市圏の発生交通量推計法と採用された経済指標

都市圏名	推計方法	採用された経済指標
広島	原単位	土地利用関連
東京	重回帰モデル	人口関連
京阪神	重回帰モデル	人口関連
中京	重回帰モデル	人口関連
岡山県南	重回帰モデル	人口関連
静岡清水	重回帰モデル	人口関連
札幌	重回帰モデル	人口関連
仙台	重回帰モデル	人口関連
北部九州	重回帰モデル	人口関連
熊本	重回帰モデル	人口関連
富山高岡	重回帰モデル	人口関連
金沢	原単位	人口関連
香川	重回帰モデル	人口関連
長崎	重回帰モデル	人口関連
鹿児島	重回帰モデル	人口関連

その採用理由を整理すると以下のようになる。

〈重回帰モデル法の採用理由〉

i) 重相関係数が一般に高く，現状を良く説明していると考えられる。
ii) 重回帰モデルの偏回帰係数が容易に推計できる。
iii) 原単位法のようにゾーン別のバラツキが大きくならない。
iv) マルコフチェーン法やシステムダイナミックス法に比較してモデルの構築が容易である。

〈人口関連指標の採用理由〉

i) 人間の行動原理から考えて交通の発生を人口関連指標を用いて説明することは妥当であり、相関性も高い。
ii) 人口関連指標と土地利用関連指標とを比較すると、現況値，将来値とも人口関連指標の方が入手しやすい。

重回帰モデル法による発生交通量の推計は，以上の理由により今後とも多用されていくものと考えられる。とくに最近はプログラムパッケージの整備により、重回帰モデル法の適用はごく容易になり，それとともに種々の問題が顕在化しはじめた。その代表的なものは説明変数相互間に高い相関関係を有するために生じる多重共線性の問題であり，また説明変数の選択基準に関する問題である。本文はこれらの問題について交通計画学の立場から考察を試みたものであり，種々の説明変数の選択基準についてその特徴を明らかにするとともに、実験的回帰分析法による寄与率が説明変数の選択基準として有用であることを提言するものである。

2. 説明変数の選択基準

重回帰モデルを用いて発生交通量推計モデルを構築するとき、最も重要でかつ困難な課題は説明変数の選択問題である。説明変数が先験的に決定されることはほとんどないと言って過言ではない。一般にはいくつかの候補変数が与えられ、試行錯誤の末に最良と思われる変数の組合せを選択し、最終的な重回帰モデルを構築している。試行錯誤の過程においては、主観的判断と客観的判断が入り乱れる。このこと自体に問題はないが、主観的判断の取り入れ方と、客観的判断のルールを確立する必要がある。本節においては重回帰モデルの説明変数の選択に用いられる種々の基準について簡単に紹介する。

1) ステップワイズ法：〔RSS基準〕[1) 2)]

ステップワイズ法は重回帰モデルに使用する説明変数を一定の規則にしたがって，逐次選択して行く方法であり、①変数増加法，②変数減少法，③変数増減法，④変数減増法などの手法が提案されている。いずれの手法も全変動に対する残差平方和（Residual Sum of Squares）の比を基準としているため，本研究ではステップワイズ法による説明変数の選択基準をRSS基準と呼ぶことにした。ステップワイズ法は前のステップで導入された最良の変数であっても，後のステップではその時に導入されている他の変数の関係で役に立たないことがあるという点を考慮している。このため偏F値の基準をつくり，各ステップで採用された説明変数の影響度を判断しなければならない（詳しくは文献1), 2)を参照のこと）。

2) 予測平方和基準：〔PSS基準〕[3) 4)]

PSSは（Prediction Sum of Squares）の略であり，予測平方和と訳されている説明変数の選択基準である．PSSの計算は次のように行なう。

今，n組のデータがあるとする．このα番目の観測値y_αを推計するとき，n組のデータ全部を使わずに、α番目のデータだけを除いた（$n-1$）組のデータを用いて回帰分析を行ない、パラメータを決定する．すなわちα番目の対象の予測値$\hat{y}_\alpha$は（2.2）式によって求められる。

$$\hat{y}_\alpha = \hat{a}_{0\alpha} + \sum_{i=1}^{p} \hat{a}_{i\alpha} x_{\alpha i} \qquad (2.2)$$

ここで，$\hat{a}_{0\alpha}, \cdots \hat{a}_{i\alpha}$：（$y_\alpha, x_{\alpha 1}, x_{\alpha 2} \cdots x_{\alpha p}$）のペアを除いた（$n-1$）組のデータから推計したパラメータ値．PSSは（2.3）式によって計算する。

$$PSS = \sum_{\alpha=1}^{n} (y_\alpha - \hat{y}_\alpha)^2 \qquad (2.3)$$

（2.3）式によって計算されたPSSには次のような性質がある．

i) PSSは同じ変数の組合せにおいて，RSS（残差平方和）より大きな値となる。両者の関係を期待値をもって示すと（2.4）式のようになる。

$$E(PSS) \geqq \left(\frac{n}{n-p-1}\right)^2 \cdot E(RSS) \qquad (2.4)$$

ここで，n：データ数，p：説明変数の数

ii) PSSには極値が存在する．この性質を利用して，種々の変数の組合わせの中でPSSを最小にする重回帰モデルを構築することができる。

3) 赤池氏の情報量規準：〔AIC〕[5) 6)]

AICとは，現実に観測されるデータを用いて統計的モデルの評価を行うための規準であり，1971年に文部省統計数理研究所の赤池氏が（2.5）式のように定義したものである。

$$AIC = (-2)\,ln(\text{最大尤度}) + 2\,(\text{パラメータ数}) \qquad (2.5)$$

（2.5）式の右辺第1項は，モデルの適合度の良否をはかる量であり、第2項はパラメータの増加に対するペナルティーとなっている。したがって、「AICの小さいモデルほど望ましい」ことになり、「データへのあてはまりが良く，なるべくパラメータ数が少ない」モデルが構築されることになる．いくつかの可能なモデルが与えられたとき，AICを最小にするようなモデルを選択する方式を，MAIC（Mimimum A.I.C.）方式という。この方式はモデルの尤度関数が定義されているかぎり，どのような統計的問題にも適用することが可能であるという長所を有している。しかし、常にモデルの分布型を特定化しなければならないという欠点のあることも指摘されている[6)].

さて，AICによる重回帰モデルの評価法を以下に示そう。

目的変数の観測値をy_i，その予測値$\hat{y}_i$が（2.6）式の重回帰モデルで推計されたとする。

$$\hat{y}_i = \hat{a}_0 + \sum_{i=1}^{k} \hat{a}_i x_{ji} \qquad (2.6)$$

y_i と $\hat{y}_i$ との残差を ε_i とすると ε_i は（2.7）式のようになる。

$$\varepsilon_i = y_i - \hat{y}_i \qquad (2.7)$$

（2.7）式で示した ε_i が，$N(0, \sigma^2)$ という正規分布に従うとすると，

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right) \qquad (2.8)$$

（2.8）式の対数尤度を計算すると，

$$\begin{aligned} \ln(\theta) &= \sum_{i=1}^{n} \ln f(\varepsilon_i) \\ &= \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi \cdot \sigma^2}} \exp\left(-\frac{\varepsilon_i^2}{2\sigma^2}\right)\right) \\ &= \sum_{i=1}^{n} \left\{-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{\varepsilon_i^2}{2\sigma^2}\right\} \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{2 \cdot \sigma^2} \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{n}{2} \qquad (2.9) \end{aligned}$$

（2.9）式を（2.5）式へ代入すると

$$AIC = n\ln\sigma^2 + 2(K) + (n\ln 2\pi + n)$$

ここで，$(n\ln 2\pi + n) = const$ であるから，

$$AIC = n\ln\sigma^2 + 2(K) + C \qquad (2.10)$$

それゆえ，重回帰モデルの比較は（2.11）式の大小によって行なえばよい。

$$w = n\ln\sigma^2 + 2K \qquad (2.11)$$

3．実験的回帰分析法とERSS基準[7)8)9)]

発生重回帰モデルは発生交通量を説明変数の線形式で推計しようとするものであり，パラメータ（偏回帰係数）は一般に最小二乗法を用いて決定される。しかし，最小二乗法を用いてパラメータを推計する際に次のような問題が発生する．

表2　道央都市圏における指標の単相関係数

	夜間人口	就業人口	第1次就業人口	第2次就業人口	第3次就業人口	従業人口	第1次従業人口	第2次従業人口	第3次従業人口	児童・生徒学生数
夜間人口	1									
就業人口	0.989	1								
第1次就業人口	-0.002	-0.193	1							
第2次就業人口	0.911	0.901	-0.192	1						
第3次就業人口	0.967	0.984	-0.292	0.824	1					
従業人口	0.237	0.301	-0.203	0.171	0.347	1				
第1次従業人口	0.084	0.099	0.437	0.085	0.048	0.345	1			
第2次従業人口	0.448	0.511	-0.237	0.477	0.513	0.873	0.290	1		
第3次従業人口	0.189	0.252	-0.196	0.107	0.305	0.995	0.336	0.823	1	
児童・生徒学生数	0.769	0.764	-0.240	0.573	0.801	0.242	0.037	0.380	0.209	1

i）採用される説明変数の組み合せによって，一部のパラメータの符号が変化することがある。また，符号が変化しないまでも，値が著しく変動する。これは最小二乗法を適用する際に，パラメータが区間（−∞，∞）のすべての値をとり得ることを仮定しているためであり，説明変数間の相関性が高いためである。すなわち，重回帰モデルに採用される説明変数は原則として互いに独立であることが必要条件となっているにもかかわらず，発生重回帰モデルに数多く採用されている人口関連指標は，表2からも明らかなように互いに強い相関性を有している。このため統計学でいう重共線問題が発生し，多くの計画者，研究者がこの問題の解決に苦慮している。

ii）最小二乗法を用いて推計したパラメータが説明変数を取り入れていく段階ごとに大きく変動するため，目的変数に対する各説明変数の影響度を知ることができない。また，説明変数の増加と共に決定係数も増加するので，決定係数を高めることに注目するあまり，必要以上の説明変数を採用しがちな傾向が生れる。

本研究で取り上げた実験的回帰分析法は、このような最小二乗法のもつ欠点を克服したものであり、次のような利点を有している。

i) 現実感覚にあたったパラメータの範囲を設定し、その範囲内で目的変数を良く説明できるパラメータを決定することができる。符号の逆転は当然のことが発生しない。

ii) 説明変数を加えていく段階でもパラメータの変動は小さく、説明変数への影響度を知ることができる。

iii) 決定係数は変数を加えていくある段階で最大に達し、以後は変数の増加と共に減少する。実験的回帰分析法によるERSS基準はこの点に注目したものである。すなわち、重回帰モデルの決定係数は一般に(3・1)式によって計算される。

$$r^2 = 1 - S_e / S_{yy} \qquad (3 \cdot 1)$$

ここで、r^2：決定係数，S_e：残差平方和，S_{yy}：全変動

従来の重回帰モデルは残差平方和S_eを減少させる形で変数の追加が行なわれるため，説明変数の増加によってr^2は次第に高くなっていく。これに対して実験的回帰分析法では，パラメータの値をあらかじめ限定しているので、変数の増加につれてr^2が増加しつづけることはない。つまり，パラメータのとりうる範囲をあらかじめ限定しているためその回帰空間は歪められており，この影響が変数を加えていくある段階で変数増加による残差平方和の減少効果を上回るので，決定係数は減少していくことになる。本文では実験的回帰分析法で得られた決定係数r^2をとくにERSSと呼び，説明変数を選択するための基準の一つとした。

さて，実験的回帰分析法の詳しいアルゴリズムは文献[6),7)]に譲るとして，その基本概念を簡単に説明すると次のようになる。

〈前提条件〉

1) 与えられた観測データに対して妥当な任意の関数形$f(a_1, a_2, \cdots a_\ell ; x_1, x_2 \cdots x_k)$を仮定できる。

2) 仮定された関数に含まれる未知パラメータに関して，固有技術などの知識からパラメータの範囲を設定できる。

以上の前提のもとで，あらかじめ与えられたパラメータの範囲内で残差平方和

$$S_e = \sum_{i=1}^{n} \{ y - f(a_1, a_2, \cdots a_\ell ; x_{i1}, x_{i2} \cdots x_{ik}) \}^2$$

を十分小さくするような$\hat{a}_1, \hat{a}_2, \cdots \hat{a}_\ell$を逐次計算法によって求めようとするのが実験的回帰分析法の基本概念である。

このとき，未知のパラメータが1個の場合には，あらかじめ与えられた範囲内で，残差平方和を最小にするパラメータ値は容易に求めることができる。しかし，未知のパラメータが多数ある場合には，逐次近似の計算を合理的に行なうことが必要になってくる。直交表を用いて逐次近似計算を行うのはまさにこのためであり、実験的回帰分析法は結局のところ，直交表を利用した多変数逐次近似推計法と定義することができよう。

4. 多次元正規乱数データによる各種選択基準の評価

本研究で取り上げた4種類の選択基準（ERSS基準，RSS基準，PSS基準，AIC）について，その特徴を明らかにするために多次元正規乱数データを作成してシミュレーション分析を行なった。シミュレーションに用いたデータは次のように作成した。

〔ステップ1〕；説明変数と目的変数との間に真の関係を設定する

真の重回帰モデル；$y = 100 + x_1 + x_2$

〔ステップ2〕；説明変数x_1，x_2と相関関係（$r = 0.80$）を有するx_3，x_4の説明変数を考え，仮の重回帰モデルを構築する。

仮の重回帰モデル；$y' = 100 + x_1 + x_2 + x_3 + x_4$

〔ステップ3〕；かく乱項ε_iの分散($V(\varepsilon) = \sigma_\varepsilon^2 = 10^2$)を仮定し、$N(0, \sigma_\varepsilon^2)$に従う$n$個の乱数を発生させる。

〔ステップ4〕: シミュレーションに用いる目的変数yの値を次式から計算する。

$$y_i' = y_i + \varepsilon_i = 100 + x_{1i} + x_{2i} + x_{3i} + x_{4i} + \varepsilon_i$$

以上のステップに従って一組の多変量データを作成した後，説明変数を種々組み合せた重回帰モデルを構築し、説明変数の選択基準をもとめた。

表3 模擬データによる説明変数の選択結果（I）

データの作成条件

i）真の関係：$y = 100 + x_1 + x_2$
ii）かく乱項ε_iの分布：$N(0, 10^2)$
iii）予測誤差：W＝0％

説明変数の数	用いた説明変数の組合せ	ERSS (%)	RSS×10^5	PSS×10^5	AIC
1	(x_1)	86.57	0.4265	0.4424	611.58
	(x_2)	87.98	0.3818	0.3963	600.51
	(x_3)	71.49	0.9056	0.9452	686.88
	(x_4)	71.36	0.9101	0.9514	687.38
2	(x_1, x_2)	97.54*	0.0781	0.0827*	444.82*
	(x_1, x_3)	88.20	0.3538	0.3802	597.31
	(x_1, x_4)	89.16	0.3298	0.3481	588.91
	(x_2, x_3)	90.31	0.2949	0.3135	577.73
	(x_2, x_4)	90.01	0.3101	0.3270	582.72
	(x_3, x_4)	79.92	0.6378	0.6785	654.85
3	(x_1, x_2, x_3)	97.15	0.0779	0.0849	447.61
	(x_1, x_2, x_4)	97.23	0.0770*	0.0832	446.53
	(x_1, x_3, x_4)	88.78	0.3073	0.3323	584.88
	(x_2, x_3, x_4)	91.48	0.2689	0.2917	571.53
4	(x_1, x_2, x_3, x_4)	94.63	0.0771	0.0858	449.52

*：各基準のうちで最も望ましい説明変数の組合せ

表4 模擬データによる説明変数の選択結果（II）

データの作成条件

i）真の関係：$y = 100 + x_1 + x_2 + x_3$
ii）かく乱項εの分布：$N(0, 20^2)$
iii）予測誤差：W＝0％

説明変数の数	用いた説明変数の組合せ	ERSS (%)	RSS×10^5	PSS×10^5	AIC
1	(x_1)	78.54	1.1546	1.1956	711.17
	(x_2)	79.99	1.2099	1.2555	715.85
	(x_3)	78.14	1.0781	1.1227	704.32
	(x_4)	70.04	1.8807	1.9596	759.96
2	(x_1, x_2)	91.75	0.5446	0.5816	639.05
	(x_1, x_3)	91.39	0.5679	0.6022	643.24
	(x_1, x_4)	86.68	0.8769	0.9222	686.68
	(x_2, x_3)	92.14	0.5161	0.5502	633.68
	(x_2, x_4)	85.64	0.9492	1.0017	694.61
	(x_3, x_4)	87.66	0.8039	0.8546	677.99
3	(x_1, x_2, x_3)	95.28*	0.3116	0.3395*	586.25*
	(x_1, x_2, x_4)	91.86	0.5143	0.5597	636.36
	(x_1, x_3, x_4)	92.16	0.5077	0.5495	635.07
	(x_2, x_3, x_4)	92.41	0.4816	0.5238	629.79
4	(x_1, x_2, x_3, x_4)	94.61	0.3080*	0.3431	588.16

*：各基準のうちで最も望ましい説明変数の組合せ

表3は真の関係を〔$y = 100 + x_1 + x_2$〕，かく乱項の分布を$N(0, 10^2)$，説明変数間の相関係数を$r = 0.80$としたときの、各種選択基準の値を示したものである。また，表4は真の関係を〔$y = 100 + x_1 + x_2 + x_3$〕かく乱項の分布を$N(0, 20^2)$，説明変数間の相関係数を$r = 0.80$としたときの結果を示したものである。

表3，表4に示した基準値のうち，実験的回帰分析法による決定係数〔ERSS〕は大きな値ほど望ましく，その他の基準値は小さい値の方が良い結果となっている。パラメータ（偏回帰係数）の推計は、ERSSについては実験的回帰分析法を用いており，他の基準値の場合には最小二乗法を採用した。表3，表4から次のことが明らかにされた。

1) 実験的回帰分析法による決定係数〔ERSS〕は、予測平方和〔PSS〕，赤池氏の情報量規準〔AIC〕とともに、真の関係に一致した説明変数の組み合せを選び出すことができた。PSSやAICは説明変数の選択基準として理論的に確立された基準でありERSSがこれらの基準と同等の判別能力を有していることが分った。
2) ステップワイズ法による選択基準〔RSS〕に従うと，その最良モデルはERSS，PSS，AICに比べて常に説明変数の数が多くなっている。その数は説明変数の取り入れ，取り去りの基準となる偏F値の大きさによって左右される。本研究では参考文献(1)の指摘をふまえ，$F = 2.0$として計算を行った。
3) ステップワイズ法は取り込む説明変数の数を指定したとき，最良の組合せを捜し出すことができる。

この特徴を生かし，ステップワイズ法と実験的回帰分析法とを組み合せた発生重回帰モデルの構築法を次節において提言する。

5. 実験的回帰分析法による発生重回帰モデルの構築法

図1はステップワイズ法と実験的回帰分析法とを組合せて発生重回帰モデルを構築するプロセスを示したものである。図1のフローチャートは大きく次の3ステップに分割することができる。

〔ステップ1〕ステップワイズ法による説明変数の組合せの決定

発生交通量を説明すると思われる変数（n個）に対してステップワイズ法を適用し、説明変数が1個，2個，… $n-1$個の場合の最適組合せをもとめる。

〔ステップⅡ〕：実験的回帰分析法によるパラメータの推計

説明変数が1個，2個，…，$n-1$個の重回帰モデルについて実験的回帰分析法を適用し，パラメータを推計するとともに，そのERSSを求める。

〔ステップⅢ〕：ERSSによる最良モデルの決定

実験的回帰分析法による最良モデルは，オⅡステップで計算したERSSが最も高い重回帰モデルとする。このモデルは「現状データの説明力が高く，しかもパラメータの数が少ない（これは予測時の誤差発生を抑制する）」という特長を有している。

図1に示したフローチャートに従い，道央都市圏パーソントリップ調査のデータを用いて発生重回帰モデルを構築した例を次に示そう。

目的変数は全目的発生交通量とし，説明変数として表2に示した10個の人口関連指標を採用した。ステップワイズ法により説明変数が1個，2個，3個，4個の場合の最良組合せを求め，そのパラメータを最小二乗法で推計した。表5はその結果を示したものである。表5からも明らかなように，RSSは説明変数が多くなるほど減少してくる。また，説明変数が3個の場合にはx_3（就業人口総数）のパラメータがマイナスとなり，説明変数が4個のときはx_3，x_4（オ2次従業人口）のパラメータがマイナスになっている。これは説明変数間の相関性が高いことを意味しており，発生重回帰モデルの論理性が損なわれていることを示している。なお，ステップワイズ法では説明変数が4個のときの重回帰モデルを最良としている。

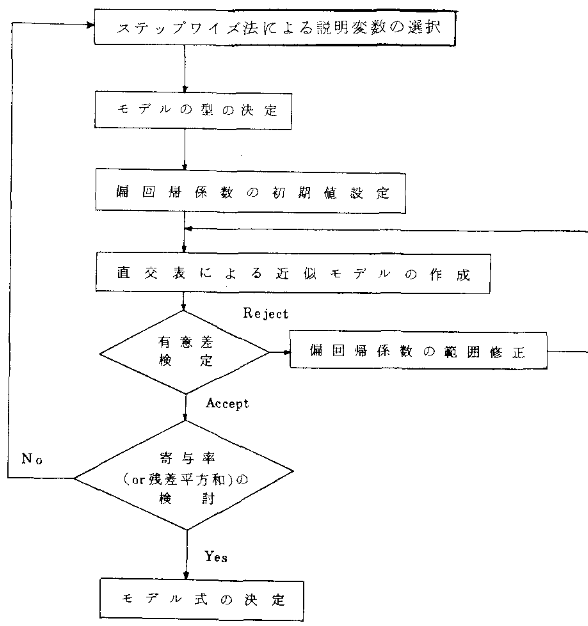

図1 実験的回帰分析法による偏回帰係数の推計プロセス

表5 ステップ・ワイズ法による発生重回帰モデルの構築

説明変数の数	重回帰モデル	RSS
1 個	$y = 19,438.60 + 3.234\,x_1$	0.2356×10^{11}
2	$y = 379.96 + 2.984\,x_1 + 1.220\,x_2$	0.5557×10^{10}
3	$y = -283.55 + 3.224\,x_1 + 4.942\,x_2 - 8.124\,x_3$	0.2381×10^{10}
4	$y^* = 66.57 + 3.547\,x_1 + 4.703\,x_2 - 7.328\,x_3 - 2.549\,x_4$	0.1877×10^{10}

＊：最良モデル

表6 実験的回帰分析法による発生重回帰モデルの構築

説明変数の数	重回帰モデル	ERSS(%)
1	$y = 19,565.16 + 3.129\,x_1$	88.01
2	$y^* = 280.09 + 3.000\,x_1 + 1.219\,x_2$	97.17
3	$y = 675.78 + 2.969\,x_1 + 1.094\,x_2 + 0.250\,x_3$	97.07
4	$y = 813.98 + 2.688\,x_1 + 0.781\,x_2 + 0.688\,x_3 + 1.938\,x_4$	96.07

＊：最良モデル

ただし，表5，表6とも，x_1：従業人口総数（人）　x_3：就業人口総数（人）
x_2：夜間人口（人）　x_4：第2次従業人口（人）

表6は実験的回帰分析法による発生重回帰モデルの構築例を示したものである。この表からも明らかなように，最小二乗法推計でマイナスとなったパラメータは実験的回帰分析法ではプラスの符号を維持しており，発生重回帰モデルとしての論理性は確保されていることがわかる。一方，ERSSの値は符号の逆転したモデルから減少しており，実験的回帰分析法では説明変数が2個の場合のモデルを最良モデルと判定している。

実験的回帰分析法で求めたパラメータは，最小二乗法推計値に比べて変動が少なく，しかも説明変数が1個や2個の場合には，最小二乗法推計値とほぼ同等の値を有することがわかった。それゆえ，実験的回帰分析法で推計したパラメータは，各説明変数が目的変数に及ぼす影響度を示しているものと考えられる。しかし，実験的回帰分析法における定数項は交通計画学的な意味をもたず，統計的には誤差項にすぎないことに注意する必要がある。

6. おわりに

実験的回帰分析法は重回帰モデルにおける説明変数の選択や、パラメータの推計に優れた機能を発揮し、発生重回帰モデルの構築にきわめて有用な手法であることが明らかにされた。重回帰モデルは多変量解析法の中でもとくに多く用いられるモデルであり、その長所，短所についてはすでに十分な研究が行なわれているといっても過言ではない。しかし一方では、多重共線性問題や予測誤差問題等に残された課題の多いことも事実である。[10][11] 実験的回帰分析法はパラメータに関して非線形な関数や、未知数の個数がデータ数より多い場合にも適用することができ，その工学的価値は今後ますます注目されてくるものと考えられる[7]。したがって今後は次の点に関してさらに研究を進めて行く必要がある。

1) 実験的回帰分析法における決定係数（ERSS）の統計学意味の確立，とくに極値の存在を理論的に明らかにすること。
2) 実験的回帰分析法における収束条件の吟味，とくに直交表の選択と収束判定値Fの決定を合理的に行なう必要がある[12]。
3) 予測誤差を考慮したときの実験的回帰分析法の適用

謝辞　本研究は北大工学部交通計画学講座で一貫して行なってきた発生交通量推計モデルの研究をとりまとめたものである。本研究を進めるにあたり、貴重な御意見，御指導を賜った五十嵐日出夫教授に厚くお礼を申し上げるとともに，幾多の課題を解決してきた白石悟君（運輸省），五十嵐力君（労働省）に対して感謝の意を表する次第である。

参考文献

1)、奥野忠一，芳賀敏郎ほか；多変量解析法，日科技連，1971年
2)、N.R.DRAPER：Applied Regression Analysis：　中村慶一訳，応用回帰分析，森北出版，1968年
3)、奥野忠一，芳賀敏郎ほか；続多変量解析法，日科技連 1976年
4)、白石悟、山形耕一，佐藤馨一；PSS基準による重回帰モデルの変数選択法に関する研究，才32回土木学会年次学術講演会概要集才IV部門，1977年
5)、赤池弘次：情報量規準とは何か，数理科学（特集情報量規準），サイエンス社，1976年3月
6)、佐和隆光：回帰分析，朝倉書店，1979年
7)、田口玄一，横山巽子；ビジネスデータの分析，丸善，1975年
8)、佐藤馨一，五十嵐力；実験的回帰分析法による発生交通量推計モデルの構築，才34回土木学会年次学術講演会概要集才IV部門，1979年
9)、佐藤馨一，五十嵐日出夫；実験的回帰分析法による発生重回帰モデルの構築，北海道大学大型計算機センターニュース，Vol.12，No.5，1980年
10)、山形耕一，桐越信：交通需要予測プロセスの予測精度について，才1回土木計画学研究発表会講演集，1979年
11)、桐越信，佐藤馨一；予測誤差基準による発生重回帰モデルの評価方法に関する研究，才35回土木学会年次学術講演会概要集才IV部門，1980年
12)、横山巽子：実験的回帰分析－事前情報を考慮したパラメータの逐次近似推定法－，応用統計学，Vol.8，No.3，1979年