

利用者属性や利用時刻に基づく都市内鉄道駅ゾーンの分類

(仙台PTとICAを利用して)

○[土] 加藤 寛道 [土] 嶋原 亮介 [土] 奥村 誠 (東北大学)

[土] 塚井 誠人 (広島大学)

The Classification of Zones with Railway Station, based on the Trip Maker Attributes

○Hiromichi Kato, Ryousuke Shigihara, Makoto Okumura (Tohoku University)

Makoto Tsukai, (Hiroshima University)

It is necessary to know the quality of service that passengers need, in order to plan and manage railway service properly. This paper classifies zones with railway station, based on the trip maker attributes, such as age and departure time. We show that introduction of the zone classification into gravity model improves the fitness of the model to the OD data, observed by the Sendai PT in 2002.

キーワード: パーソントリップ調査, 独立成分分析, パターン抽出

Key Words: Person Trip Survey, Independent Component Analysis, Pattern Extraction

1. はじめに

近年, ライフスタイルの多様化に伴う都市内交通需要の多様化により, 公共交通の利用形態は地域によって大きく異なってきた。事業者が適切な計画・運営を行うためには利用者の特性を把握することが必要である。

本研究は第4回仙台都市圏パーソントリップ調査データに基づき, 鉄道駅が存在するゾーンごとのトリップの年齢構成, 利用時刻構成を集計し, それを独立成分分析を用いて整理することで, ゾーンの性格分類を行う。さらにゾーン間ODの重力モデルにこのゾーン分類を取りこむことにより, 精度が向上することを確認する。

2. 分析手法

2.1 ゾーンの性格分類

(1) データの集計

パーソントリップ(以下PT)調査は都市圏の交通の実態を総合的に把握するために行われる調査で, 「どのような人が」「いつ」「何の目的で」「どこから」「どこへ」「どのような交通手段で」動いたかについて調査し, 1日平日の全ての動きを捉えるものである。本研究では2002年に行われた第4回仙台都市圏PT調査の個票データを用いる。

以下では仙台都市圏全64大ゾーンのうち鉄道駅のある35ゾーンを分析対象とし, 各ゾーンを出発(O)・到着(D)するトリップの年齢・出発時刻をそれぞれ集計する(以下で

は, 年齢に関する分布を例に説明する)。この際, 個人が1日の中で行った全てのトリップを対象に集計を行うと, 例えば通勤トリップと帰宅トリップのように2地域間を往復するトリップが集計され, 居住地・勤務地の区別がつきにくくゾーンの性格分類には適さない。そこでここでは個人が1日の中で最初に行ったトリップのみを集計する。

(2) 独立成分分析(ICA)

集計したデータは, 各ゾーンを発着するトリップの年齢分布を表すが, この分布は例えば通勤や通学, 買物などの目的の異なるトリップの年齢分布パターンがゾーンごとに異なる比率で混ざり合ったものと考えることができる。各トリップ目的に特有の年齢分布パターンを把握することができれば, その比率に従いゾーンの性格分類を行うことが可能である。目的の異なるトリップの年齢分布パターンを $\mathbf{s} = (s_1, \dots, s_j, \dots, s_n)^T$, 各ゾーンにおける s_j の混合比率を行列 $\mathbf{A}$ で表す。各ゾーンを発着するトリップの年齢分布を $\mathbf{x} = (x_1, \dots, x_i, \dots, x_m)^T$ とすれば, これらの関係は式(1)のように表すことができる。

$$\mathbf{x} = \mathbf{A}\mathbf{s} \dots\dots\dots (1)$$

この中で, 既知であるのは $\mathbf{x}$ のみであるため, $\mathbf{A}$ と $\mathbf{s}$ を同時に推定する必要がある。式(1)のように観測データを複数の構成要素が混ざり合ったものと考え, 個々の要素を逆に求めようとする統計的手法として因子分析がある。因子分

析では、得られる因子に正規性を仮定し、2次の誤差を最小にするように計算を行う。しかし、年齢分布パターンには中高生や大学生世代にのみピークを持つようなパターンも考えられるため、その要因に正規性を仮定することには問題がある。

そこで本研究では、因子に正規性を仮定しない独立成分分析 (Independent Component Analysis) (以下ICA) を用いる。推定にあたって用いる仮定は、原信号 $\mathbf{s}$ が各時刻 t において統計的に独立であることと、それぞれ平均0、分散1の分布に従うという仮定である。以下では抽出された原信号 $\mathbf{s}$ を独立成分と呼ぶ。

Hyvarinen ら¹⁾ は原信号の独立性の基準として、尖度を用いる方法を提案した。この方法は「独立な確率変数の和の平均の分布は、和をとる変数の数が増えるに従って正規分布に漸近する」という中心極限定理を踏まえ、尖度を指標として非正規的な分布を探索することで、最も独立な分布を探索するものである。本研究においてもこの FastICA と呼ばれるアルゴリズム²⁾を用い、あるゾーンを発着するトリップの年齢分布 $x_i (i=1, \dots, 35)$ を与えて、年齢分布パターンを表す独立成分 $\mathbf{s}^y$ と、各ゾーンにおける年齢分布パターンの比率を表す混合行列 $\mathbf{A}^y$ を算出する。

(3) ゾーンの性格分類

ICA により導いた $\mathbf{A}^y$ の各行 $\mathbf{a}_i^y (i=1, \dots, 35)$ は各ゾーンにおける年齢分布パターンの比率であり、年齢分布に関するゾーンの性格を表している。発着トリップ数の絶対量の影響を除去するためにここでは $\mathbf{a}_i^y$ を平均0、分散1に基準化した $\mathbf{a}_i^{y*}$ を計算する。

同様の手順で各ゾーンを出発・到着するトリップをその出発時刻について集計し、ICA を適用する。独立成分 $\mathbf{s}'$ はあるゾーンを出発・到着するトリップが出発した時刻の分布パターンを、混合行列 $\mathbf{A}'$ は各ゾーンにおける時刻の分布パターンの混合比率を表すが、これを基準化して $\mathbf{a}_i'^*$ を得る。さらに行列 $(\mathbf{a}_i^{y*} \ \mathbf{a}_i'^*)$ についてクラスター分析を行うことで、年齢・出発時刻に関するゾーンの性格に着目したタイプ分類を行う。

2.2 ゾーンのタイプを考慮した重力モデル

地域間の流動データに対しては、重力モデルを代表とする地域間相互作用モデルの当てはめが多く試みられている³⁾。標準的な重力モデルを式(2)に示す。

$$T_{ij} = \frac{O_i^{\alpha_1} D_j^{\alpha_2}}{d_{ij}^{\beta}} \exp(\lambda \sigma_{ij}) \dots \dots \dots (2)$$

ここで T_{ij} はゾーン i から j へのトリップ総数 (全目的・全交通手段)、 O_i はゾーン i を出発するトリップの総数、 D_j はゾーン j に到着するトリップの総数、 d_{ij} はゾーン i から j の平均所要時間、 σ_{ij} はゾーン内々のトリップを表すダミー変数である。ここで、例えば住宅地と市街地の間のト

リップ数は推定値よりも多く、逆に住宅地と住宅地の間のトリップ数は推定値より少ないなど、実際のトリップ発生の有無にはゾーンの性格が影響すると考えられる。ゾーンのタイプの組み合わせを導入した重力モデルは式(3)のように表せる。

$$T_{ij} = \frac{O_i^{\alpha_1} D_j^{\alpha_2}}{d_{ij}^{\beta}} \prod \exp(\gamma^{kl} \delta_{ij}^{kl}) \exp(\varepsilon^k \phi_j^k) \dots \dots \dots (3)$$

k は出発ゾーンのタイプ、 l は到着ゾーンのタイプであり、 δ_{ij}^{kl} はタイプの組み合わせを表すダミー変数、 ϕ_j^k はタイプ k のゾーン内々のトリップを表すダミー変数である。標準的な重力モデル(2)、ゾーンの性格を考慮した重力モデル(3)それぞれについて対数をとって線形回帰分析を行い、パラメータ $\alpha_1, \alpha_2, \beta, \lambda, \sigma_{ij}, \gamma^{kl}, \varepsilon^k$ を推定する。

3. 分析結果

3.1 ゾーンの性格分類

(1) 年齢構成の分布パターン

ICA により算出された年齢構成の分布パターンを図1に示す。白の棒グラフで示す左半分はゾーンを出発するトリップの年齢構成分布を意味し、黒の棒グラフで示す右半分はゾーンに到着するトリップの年齢構成分布を意味する。紙面の都合により混合係数の図は省略する。

s1 は 20-24 歳を除く 59 歳までの幅広い年齢層で出発・到着が多いことを意味する。このパターンは全てのゾーンにおいて混合係数が正で、平均的に大きな比率をもつことから、各ゾーンを出発・到着するトリップの一般的な年齢構成パターンを表しているといえる。

s2 は 45-79 歳までの中高齢者の出発・到着が多く、また 15-24 歳の年齢層に注目すると出発側では正、到着側では 0 に近い値をとる。このパターンも全てのゾーンで混合係数が正で、大倉、松島町、松陽台などの郊外で比率が大きく、中央や卸町では比率が小さいことから、地方部の特徴を表すパターンである。

s3 は 20-44 歳までの比較的若い労働者世代の出発が多いことを表すパターンである。混合係数も中心市街地と地方部で小さく、市街地周辺の住宅地で大きいことから、単身者世帯の居住を表すパターンであると考えられる。

s4 は 15 歳以上のほぼ全ての世代で出発が少なく、到着が多いゾーンであることを意味する。また混合係数は中央、卸町などの中心部で大きく、その他は負の値をとることから、中心市街地の特徴を表すパターンである。s5 は 15-19 歳の中高生の到着トリップが多く、中学校・高校があるゾーンの特徴を表すパターンである。

s6 も同様に 20-24 歳の大学・専門学校などがあるゾーンの特徴を表すパターンである。

(2) 出発時刻の分布パターン

(1) と同様に ICA により算出された出発時刻の分布パターンを図2に示す。左側の白の棒グラフはゾーンを出発

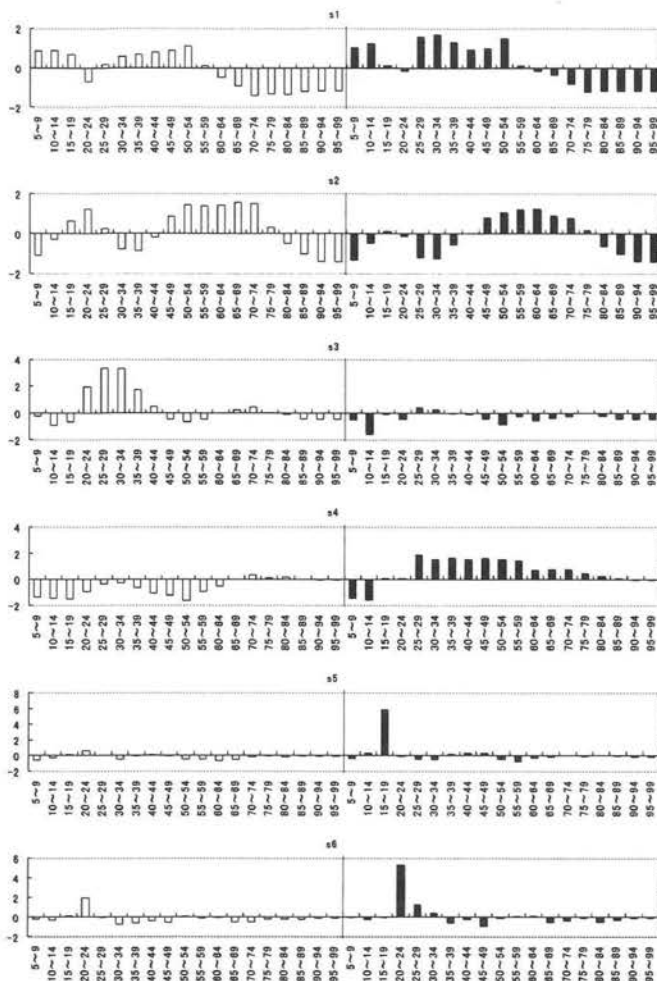


図1 年齢構成の分布パターン

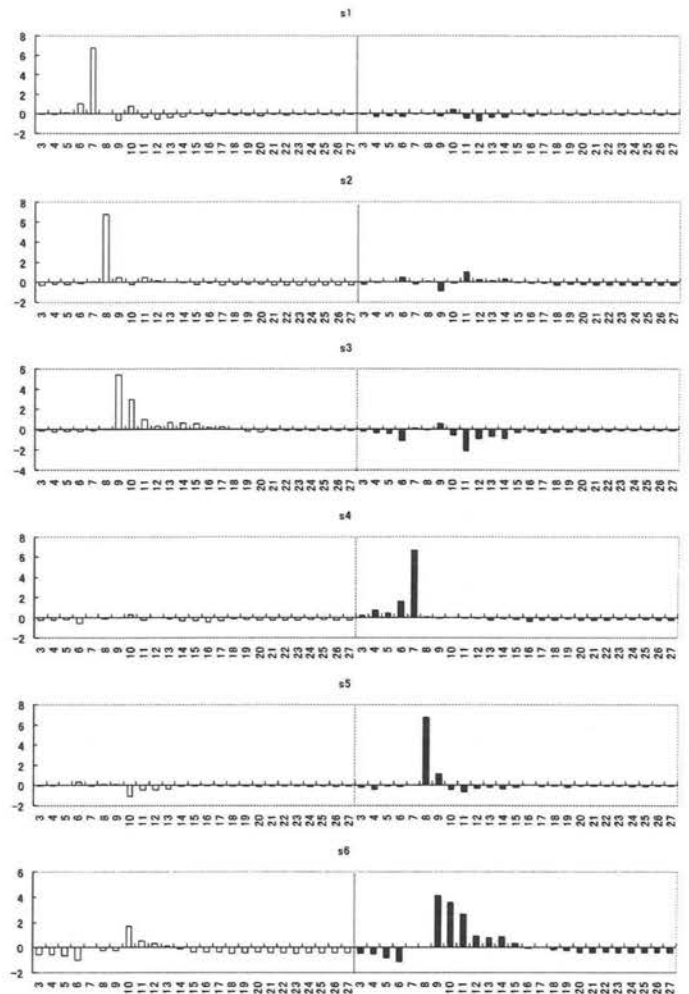


図2 出発時刻の分布パターン

したトリップの出発時刻分布を意味し、右側の黒の棒グラフはゾーンに到着したトリップの出発時刻分布を意味する。紙面の都合上、混合係数の図は省略する。

s1, s2, s3 はそれぞれ7時台、8時台、9時以降に出発するトリップが出発しているというゾーンの特徴を表している。s4, s5, s6 も同様にそれぞれ7時まで、8時台、9時以降に出発するトリップが到着しているというゾーンの特徴を表している。

(3) ゾーンの性格分類

クラスター分析の結果、35ゾーンを5つのグループに分類した。グループごとにパターン分けした地図を図3に示す。図3中の太線は鉄道を表す。また、鉄道駅のないゾーンは対象外であるため白色で示している。各グループの分類と該当するゾーン名を表1にまとめる。

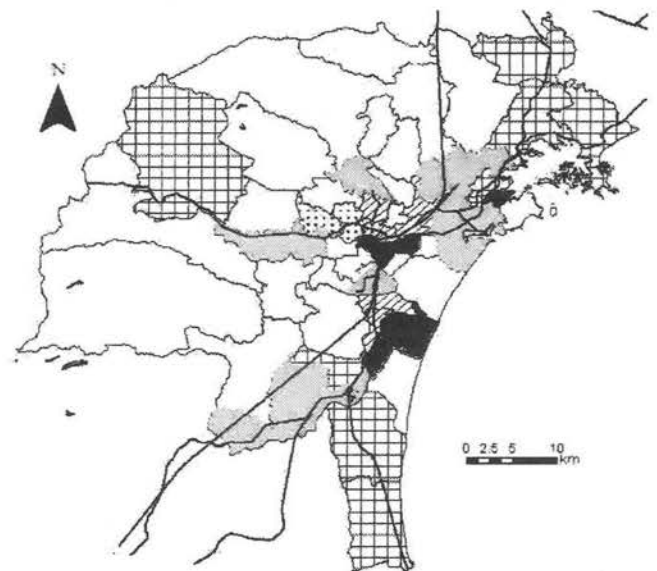


図3 ゾーンの性格分類

表1 グループのラベルと分類

ラベル	パターン	分類	ゾーン名
A	[Pattern A]	中心市街地	中央、卸町、榴岡、関上、尾島町
B	[Pattern B]	周辺市街地	泉中央、高崎、長町、福田町、折立、下馬、愛子、柴田町、大河原町、岩沼、利府町
C	[Pattern C]	近隣住宅地	小田原、七北田、東仙台、岩切、中田、増田
D	[Pattern D]	住宅地(大学)	上杉、柏木、荒巻、貝ヶ森
E	[Pattern E]	郊外住宅地	松陽台、山元町、成瀬町、鹿島台町、亘理町、松島町、松ヶ丘、玉川、大倉

3.2 重力モデルの推定結果

表2に式(2)式(3)の重力モデルのパラメータ推定結果を示す。発生量 O_i 、集中量 D_j はどちらのモデルも大差なく、0.88~0.95のパラメータ値となっており、t値は大きい。平均所要時間 d_{ij} のパラメータは-2よりも若干絶対値が大きいことが共通している。

表2 パラメータの推定結果

説明変数	式(2):標準的な重力モデル			式(3):ゾーンの性格を考慮したモデル		
	非標準化係数	標準化係数	t値	非標準化係数	標準化係数	t値
定数項	-5.452			-5.804		
$\log(O_i)$	0.922	0.332	17.487	0.954	0.344	17.842
$\log(D_j)$	0.883	0.316	16.621	0.916	0.328	16.964
$\log(d_{ij})$	-2.198	-0.513	-23.81	-2.293	-0.535	-25.573
σ	1.744	0.152	7.082	-	-	-
δ^{AE}	-	-	-	0.793	0.076	3.945
δ^{BA}	-	-	-	0.358	0.038	2.022
δ^{BC}	-	-	-	-0.317	-0.037	-1.966
δ^{CC}	-	-	-	-0.629	-0.056	-2.711
δ^{EA}	-	-	-	0.632	0.061	3.176
ϕ^B	-	-	-	1.344	0.066	3.422
ϕ^C	-	-	-	2.148	0.079	3.823
ϕ^D	-	-	-	1.338	0.040	2.115
ϕ^E	-	-	-	2.781	0.125	6.409
調整済みR ² 値	0.595			0.610		

式(2)のモデルでは σ に対するパラメータ λ は正の値となっており、ゾーン間のトリップに比べ内々のトリップが多いことを示している。

一方、式(3)のモデルの結果を見ると、 δ^{AE} 、 δ^{EA} で正の有意なパラメータが得られた。これは中心市街地と郊外住宅地間を往来するトリップが多いことを表している。また δ^{BA} では正、 δ^{BC} では負の有意なパラメータをもつことから、周辺市街地から中心市街地へ向かうトリップは多い一方、周辺市街地から都市近隣の住宅地へのトリップは少ないことがわかる。さらに δ^{CC} では負、 ϕ^C では正の有意なパラメータをもつことから、都市近隣の住宅地で発生する短距離トリップは同様の性格を持つ他のゾーンに向かうことは少なく、そのゾーン内で用事を済ませる傾向にあることがわかる。

全体として、新たに加えたパラメータによって説明力は向上しており、調整済みの決定係数の値も0.595から0.610に向上している。

4. おわりに

本研究は第4回仙台都市圏パーソントリップ調査データに基づき、鉄道駅が存在するゾーンごとの利用者の年齢構成の分布、出発時刻構成を集計し、それをICAを用いて整理することで、目的の異なるトリップの年齢・出発時刻分布パターンを計算した。また、各ゾーンに対するこれらのパターンの比率を基に5つのグループに分類し、グループ間の性格の違いを考慮したゾーン間の重力モデルを推定した。

トリップ目的別の年齢分布パターンとしては、全ゾーンに共通して影響するパターンや都市部・郊外部の特徴を表すパターン、学校があるゾーンの特徴を表すパターン等が得られた。一方、出発時刻の分布パターンは特定

の時間帯に鋭いピークを持つパターンが得られたが、このパターンからトリップの目的を推定することはできなかった。これは、出発時刻が目的よりも目的地までの距離の影響を強く受けるためと推測される。

また重力モデルの推定結果より、発着ゾーン間のタイプの組み合わせがトリップ発生のしやすさに影響していることが確認できた。このゾーンのタイプを適切にとりこむことによって重力モデルの推計精度を向上させることができた。

本研究ではPT調査の利用交通手段は空欄であることが多く、データの代表性に問題がある可能性があると考え、鉄道駅が存在するゾーンを発着する全手段利用トリップを対象に分析を行った。今後は鉄道事業者が所有する実績データを用いて、より詳細な分析を行うことが望まれる。

参 考 文 献

- 1) Hyvarinen, A.: 詳解 独立成分分析, 東京電気大学出版局, 2005.
- 2) 加藤寛道, 木村裕介, 奥村誠, 塚井誠人: 仙台空港における遅れ・欠航の分析, 土木計画学研究・論文集, Vol.25, No.3, 2008.
- 3) 石川義孝, 空間的相互作用モデル—その系譜と体系, 地人書房.