

モバイル空間統計を用いた 都市間流動の特徴に関する分析

菅原 優志¹・塚井 誠人²

¹学生会員 広島大学 大学院工学研究科(〒739-8527 東広島市鏡山 1-4-1)
E-mail:D194510@hiroshima-u.ac.jp

²正会員 広島大学准教授 大学院工学研究科(〒739-8527 東広島市鏡山 1-4-1)
E-mail:mtukai@hiroshima-u.ac.jp

情報通信技術の発展によって、従来は困難であった、短いスパンでの都市間流動調査が可能になった。新たな都市間流動データとして着目されるモバイル空間統計はデータの速報性、収集量、蓄積において優れており、同データを用いた都市間流動の研究が進められている。しかし、同データは、過去のデータとの計測特性の違いがあるため、調査の精度や信頼性に課題が多く指摘される。本研究では、既存の調査による都市間流動の OD 表を記述する基底を非負値行列因子分解によって定め、その基底によってモバイルデータの OD 表を記述することで、モバイルデータを既存調査と共通する部分と相違する部分に分離する。これによって、既存調査と新規調査の比較、新規調査に現れる固有部分の検証を試みる。

Key Words: *inter-regional, travel demand, NMF, liner programming*

1. はじめに

2017 年の旅行・観光消費動向調査¹⁾によると日本人の国内延べ旅行者数は約 6 億 4,000 万人にのぼった。人口規模の大きな都市間については、北海道新幹線、北陸新幹線の新函館、金沢までの開業、各地の高速道路整備などにより、高速・大容量の旅客交通サービスが整備されている。新たな交通ネットワークで結ばれた都市間では、所要時間の短縮によって、流動量の増加が期待される。都市間ネットワーク整備による需要予測や整備の事業評価を行うために、より詳細かつ精緻な流動量の分析が求められる。

従来、都市間旅客流動の分析には、日本全国の純流動を調査対象とした、国土交通省が実施している全国幹線旅客純流動調査(以下、NPTS)による純流動データが用いられてきた。NPTS によって蓄積されたデータは、たとえば、人口減少下で必要最小限の幹線交通ネットワークを維持する方策や、長期的な視点からの幹線交通ネットワーク計画に活用される。これまで NPTS の純流動データは 5 年に 1 度の秋に行われるアンケート調査によって収集、分析されていた。しかし、情報通信技術の発達によって、携帯電話やカーナビゲーションシステムの移動記録から、人や自動車の移動実態を把握できる交通関連ビッグデータが開発されている。これらのビッグデータ

は、アンケート調査のような、特定の調査日の移動実態だけでなく、常時移動実態を把握できる。交通関連ビッグデータのうち、個人が所有する携帯電話位置情報を用いたモバイル空間統計(以下、MOBI データ)は、速報性に優れ、多時点での収集が容易である。この利点を活かし、都市間純流動の季節変動や幹線交通機関の開業前後の純流動の分析²⁾などが進められている。MOBI データの課題として、収集される旅客純流動の信頼性や、従来の調査との計測特性の違いが指摘されることが多い。本研究では、過去の NPTS データと MOBI データを比較するために、NPTS の純流動によって記述できる MOBI の純流動について考察を行う。

2. 既往研究

MOBI データはその利点を活かし、既存の調査を補完、代替することを目的とした分析に用いられることが多い。たとえば、都市内の人口分布を把握するために従来は、国勢調査や PT 調査が活用されてきた。ただし、調査間隔は国勢調査は 5 年、PT 調査は 10 年と長いほか、短時間での人の動きや、細かな集計単位での情報入手は困難であるという課題がある。清家ら³⁾はモバイル空間統計を用いて、概ね 1-2 万人程度の都市内のコミュニティエ

リアを対象に、時間単位で滞在人口の推計を行った。MOBI データによる推計人口と、従来調査で公表されている人口とを比較して、MOBI データの信頼性について検証し、都市内流動における MOBI データの信頼性の高さを明らかにした。また MOBI データを用いる際の留意点について言及した。菊池ら⁴⁾は、2015 年の PT 調査結果から交通量を推定した。このとき MOBI データの観測値を用いて補正を行うことで、補正前に比べて補正後の推定値は、PT 調査の実績値との相関が大きくなり再現性が向上した。

都市間流動に MOBI データを活用した研究として、高橋ら²⁾や鈴木ら⁹⁾の研究があげられる。高橋ら²⁾は北海道新幹線の開業効果を明らかにするために、開業前後の来訪者の変化を考察した。MOBI データを利用することで、平日別、昼間、夜間別など細かな時間単位での分析が可能となった。課題として、個人属性の影響や来訪目的を考慮できなかったことを指摘した。ただし、この研究ではモバイル空間統計の信頼性については検討されていない。鈴木ら⁹⁾は MOBI データの課題である旅行目的別の流動量を推計するために、NPTS の目的別宿泊先選択比率と MOBI データの目的別宿泊先選択比率が等しいと仮定して、MOBI データの全流動量と目的別流動量の和が等しくなるように旅行目的別の流動量を推計した。この研究では、MOBI データのサンプル数を既存調査に対する利点としてあげ、既存の都市間交通調査と MOBI のようなビッグデータとを適切に融合することで、都市間流動の全体像を適切に把握できると期待を示している。

MOBI データは、都市内流動においては、従来調査の補完、代替を目的とし、MOBI データの信頼性検証や MOBI データを用いた補正方法の検討が行われている。一方、都市間流動においては、MOBI データの利点に着目してその活用を図る研究は多いが、データの信頼性や既存調査との比較などはあまり行われていない。

本研究では、MOBI データと従来データとの対応を明確にすることを目的とする。これにより、MOBI の観測特性をわかり易く示すとともに、同データの信頼性を示す指標を得ることを目的とする。具体的には、MOBI データによる OD 表と従来調査の NPTS による OD 表に関して、後者の特徴によって前者の特徴が記述できる部分と記述できない部分に分離することで、OD 別に MOBI の NPTS に対する適合度を考察する。OD 表の記述には非負値行列因子分解(以下、NMF)を用いる。

3. NPTS 基底による MOBI データの記述

(1) 非負値行列因子分解(NMF)

m 行 \times n 列サイズの非負値行列を X とする。NMF は、

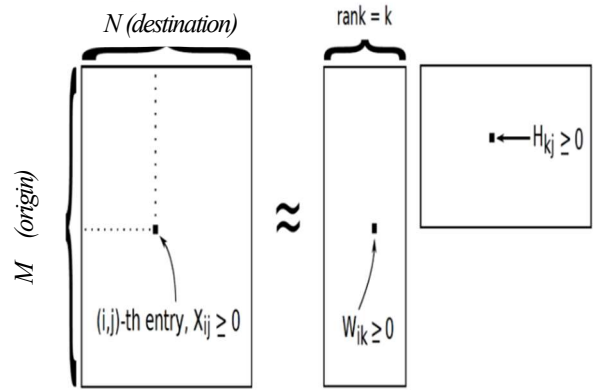


図 1 非負値行列因子分解(NMF)による行列分解

行列 X を $m \times k$ の非負値行列 W と $k \times n$ の非負値行列 H の積の形に近似的に分解する(図 1)。図中の k は NMF の基底数である。Paatero ら¹⁰⁾は、 X を近似する行列の積 $W \times H$ の最適解を導出するために、以下の最適化問題の極小値を推定した。

$$\min_{W, H \geq 0} [D(X, WH)] + R[W, H] \quad (1)$$

式(1)で、 D は損失関数、 R は任意の正則関数である。近似精度は損失関数 D によって測定され、以下の 2 種類が広く用いられる。式(2)は Euclid 距離に基づく D の計算方法であり、 X_{ij} と $(WH)_{ij}$ の距離は、右辺から計算される左辺の期待値に対して対称の 2 乗和で定義される。式(3)は一般化 Kullback-Leibler 距離であり、 X_{ij} と $(W \times H)_{ij}$ の距離は、期待値に対して、非対称である。

$$D(X, WH) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n (X_{ij} - (WH)_{ij})^2 \quad (2)$$

$$D(X, WH) = \sum_{i=1}^m \sum_{j=1}^n X_{ij} [\log(\frac{X_{ij}}{(WH)_{ij}}) - X_{ij} - (WH)_{ij}] \quad (3)$$

本研究では、式(1)を極小化する行列 W 、 H を、式(3)の目的関数の下で最小化する。以下実際のアルゴリズムは式(4)、(5)で表される更新式に基づいて推定する。

$$H_{kj} \leftarrow H_{kj} \frac{\sum_l W_{lk} X_{lj}}{\sum_l (WH)_{lj}} \quad (4)$$

$$W_{ik} \leftarrow W_{ik} \frac{\sum_l H_{kl} X_{il}}{\sum_l (WH)_{il}} \quad (5)$$

NMF の実際の計算にはフリーの統計計算ソフト R の NMF パッケージを用いた。

(2) 共通部分、相違部分の分離

MOBI の純流動を NPTS の純流動と共通部分と相違部分に分離する。純流動の特徴を数値化する手法として NMF を用いる。NMF を用いることで、OD 表は任意のパターン数によって 2 つの行列 W 、 H の積と近似できる。このとき行列 W は基底行列、 H は重み行列である。本研

究では、行方向に m 個の出発地、列方向に n 個の目的地、要素に旅客純流動量をとった非負値行列の NMF を行う。行列 W の第 k 列はパターン k における出発地の特徴、行列 H の第 k 行は k における目的地選択の重みを表している。

ここで、NPTS、国勢調査の OD 表から NMF によって得られる基底行列 W^N 、 W^C と任意の非負値行列 H^{MN} 、 H^{MC} を用いると、MOBI の OD 表 X^M は任意の行列を A とし、式(6)と表現できる。

$$X^M \approx W^N \times H^{MN} + W^C \times H^{MC} + A \quad (6)$$

式(6)において、基底行列 W と重み行列 H で表せられる部分は、NPTS あるいは国勢調査と同じ基底によって記述できる MOBI の純流動量である。任意の行列 A は NPTS と同じ基底では記述できない MOBI 固有の部分である。MOBI の純流動において、NPTS と同じように表現できる部分と、表現できない部分を明示するために、行列 A も非負値行列として求める。そこで次の制約条件(式(7))を満足する非負値行列 H^M の要素 $h_{k,j}$ を求める。

$$\begin{aligned} & \max_h \sum_{k=1, j=1}^{K, J} (h_{k,j}^{MN} + h_{k,j}^{MC}) \\ \text{st. } & \sum_{k=1}^K (w_{ik}^N h_{k,j}^{MN} + w_{ik}^C h_{k,j}^{MC}) \leq x_{i,j} \\ & h_{k,j} \geq 0, \forall_k, \forall_j \end{aligned} \quad (7)$$

一方で、調査の共通部分と相違部分の分離、旅行目的の判別には、複数調査による OD 表を結合させて NMF を行い、推定したパターンの類似性やパターンの特徴を考察する手法もある。後述の手法では、NMF の性質上複数ある OD 表のそれぞれの合計量を一致させる必要がある。しかし、のちの図-2 に示すように、NPTS、国勢調査と MOBI の OD 表は NPTS の方が流動量の多い OD と MOBI の方が多い OD とが混在している。そのため、MOBI の方が全体として NPTS の 2 倍に当たる流動を観測したため、MOBI 全体を半分にし 2 つの OD 表それぞれの合計量を一致させると、個々の OD における NPTS と MOBI の観測量の比率が歪む。これに対して前述の手法は、推定した重み行列 H^M がそれぞれの合計量の比率を考慮できるため、無理に合計量を一致させる必要性がないという利点がある。

4. 純流動データの概要

都市間流動の既存調査である NPTS は、旅行中の旅行者に対してアンケートを配布、回収する調査形式である。ある出発地 O から、同一都道府県外を目的地 D としたサンプルを対象とする。またサンプルを抽出する条件は、

年齢を 15~79 歳まで、(ただしその不明者を除く)、旅行目的が仕事、観光あるいは私的とした。調査によって得られたサンプルは、輸送統計の輸送量実績値とサンプル数と乗継特性にもとづいて拡大処理される。

MOBI データは株式会社 NTT ドコモ(以下、ドコモ)契約者の携帯電話位置情報を随時収集し、携帯電話契約者の移動に関するデータを得る調査形式である。得られたサンプルは、契約者の個人属性が、契約者の居住地での国勢調査の個人属性構成と合致するように拡大処理される。本研究で用いた MOBI データは人口流動統計と呼ばれるデータである。人口流動統計とは、ある 1 日の中で出発地から目的地へ何人移動したかを推計する統計である。移動人口の推計は、携帯電話の移動を特定条件によって 1 トリップ(流動量が 1)と定義する。その定義とは、一定時間ごとに基地局に届く携帯電話の位置情報の変化から、基準となる距離をもとに移動と滞留を判定し、滞留から移動、さらに滞留へと切り替わることである。

第 6 回 NPTS は 2015 年秋に調査が実施されているため、2015 年 10 月 26 日から 30 日の 5 日間をデータ収集期間とした。先述の定義によって 207 生活圏単位で、24 時間毎の OD 間流動量を 5 日間集計し、記録したデータからその平均を、1 日当たりの MOBI の純流動量とする。

なお、NPTS は都道府県を越える移動であっても、通勤・通学を目的とした流動は調査の対象外としている。一方 MOBI データは移動目的を判別できないため、調査日に行われた全ての移動を記録している。そこで、2015 年の国勢調査によって公表されている、常在地別の従業地・通学地集計データを用いて、NPTS がカバーしていない流動の特性を把握する。常在地から従業地への集計を通勤、従業地から常在地への集計を退勤とし、通勤と退勤の合計を 1 日当たりの常在地と従業地間の純流動量とする。

NPTS と MOBI データは、両調査の最も大きな違いとして考えられる、通勤・通学流動を国勢調査で補完したあとでも、調査形式、対象者、データ処理方法が異なることで、流動量が異なる、目的地の選択比率が異なるなどの差異が存在する。図-2 に式(8)で定義する NPTS、国勢調査と MOBI データの差異を示す。 y_{ij}^M は MOBI の純流動量、 y_{ij}^N 、 y_{ij}^C は NPTS、国勢調査の純流動量である。 P_{ij} の 50%ile 値は 0.96、最大値は 12.4 であった。図-2 において、縦軸は出発地を表し上から順に、横軸は目的地を表し左から順に都道府県を北から並べている。図-2 の左上は北海道、東北地方内々、右下は九州地方内々の OD を表している。図-2 中の白は、MOBI データの観測が 0、もしくは都道府県、大都市圏内々の OD を示している。

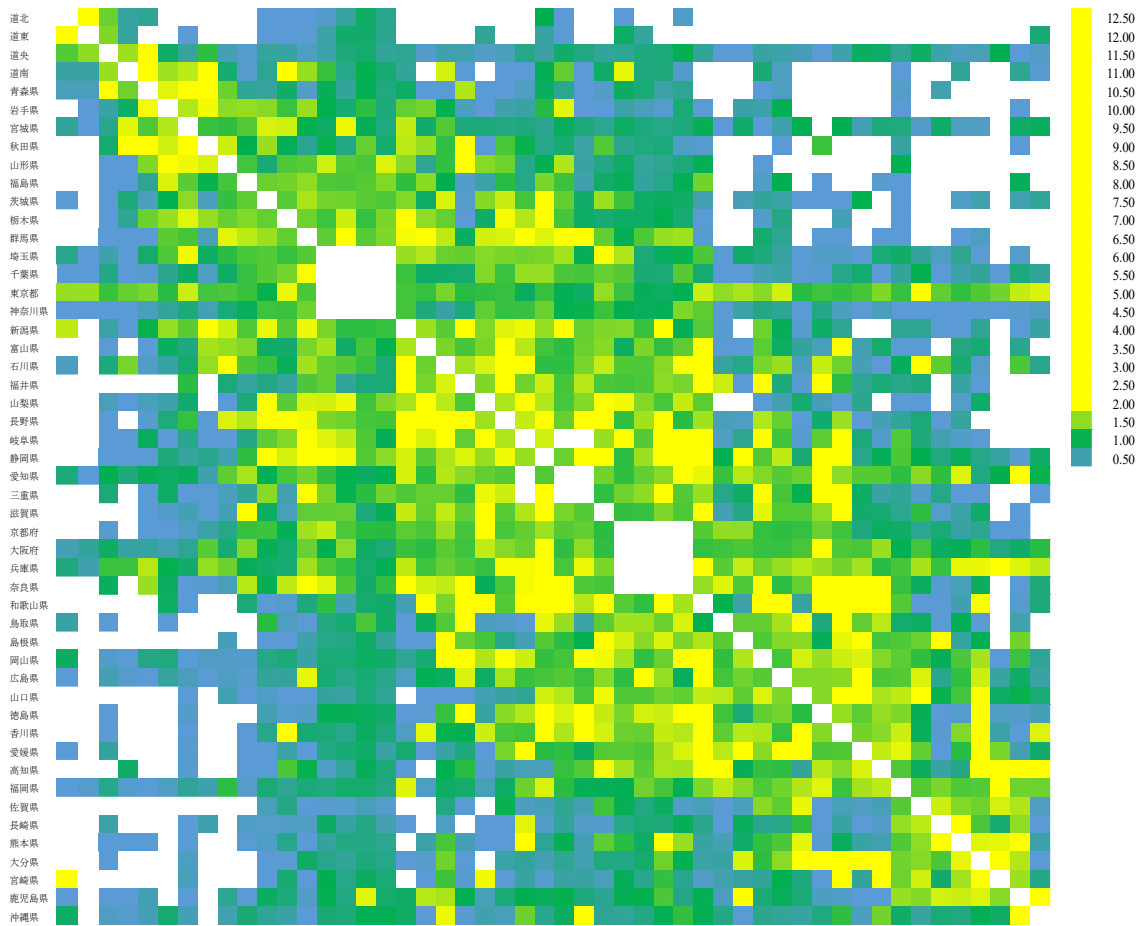


図-2 MOBI, 既存統計の数量比較

$$P_{ij} = \frac{y_{ij}^M}{y_{ij}^N + y_{ij}^C} \quad (8)$$

MOBIがNPTS, 国勢調査より過大なOD間の特徴として、東北、関東の地方内々、地方間、関東、中部、近畿の地方内々、地方間、近畿、中国の地方内々、地方間があげられる。MOBIがNPTS, 国勢より過小なOD間の特徴として、北海道、東北と九州の地方間、関東と北海道、近畿以西の地方間、北海道、九州発近畿があげられる。遠距離OD間ではNPTS, 国勢の方が多く観測され、近距離OD間ではMOBIが多く観測される傾向にある。

図-2 から NPTS, 国勢調査と MOBI の数量差を比較することもできる。しかし、日比野ら⁸⁾は複数主体による統計調査の調査結果の違いは、その違いの原因や、どのように補正し組み合わせるかといったことを明らかにする必要性を述べている。そこで、NPTS, 国勢調査と MOBI の違いの原因を明らかにするために、両データの比較を行う。そのために、MOBI データを既存調査の結果によって記述できる部分(共通部分)と記述できない部分(相違部分)に分離する。

5. 純流動データへの適用

(1) 南関東を出発地とする流動の分析

南関東(埼玉、千葉、東京、神奈川)を出発地とし、同一都県以外を目的地とする NPTS と国勢調査に基づく OD 表に対して、NMF 分析を行う。パターン数 k は NMF パッケージによって示される評価尺度と解釈性を考慮し、どちらも 2 とする。図-3~図-6 に双方の NMF 結果を示す。KN1 は東京、神奈川を主な出発地とし北海道、山梨、静岡、大阪、福岡への流動に当てはまるパターンである。KN2 は埼玉、千葉を主な出発地とし北関東を目的地とする短距離流動のパターンである。通勤のパターンはどちらも北関東への短距離流動パターンである。

NPTS と国勢調査の NMF 分析によって得た 4 パターンを基底行列とし、MOBI OD 表の重み行列を求める。これにより、MOBI と既存の 2 調査との共通部分と相違部分に分離する。また、NPTS と共通する部分と国勢調査と共通する部分とを区別し、旅行目的が不明の MOBI の流動を、通勤以外と通勤にも分離する。

前節によって得た基底行列から式(6)、制約条件(7)を満たす行列 H を推定し、MOBI と 2 調査に共通する流動量を算出する。それらの結果を図 7-9 に示す。

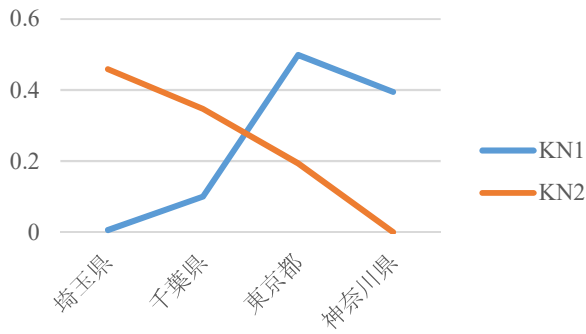


図-3 NPTS 行列W

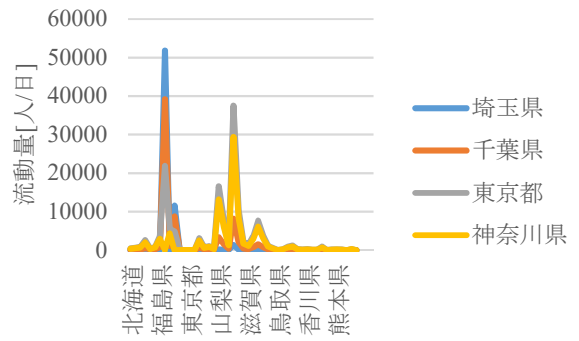


図-7 MOBI NPTS と共通する部分

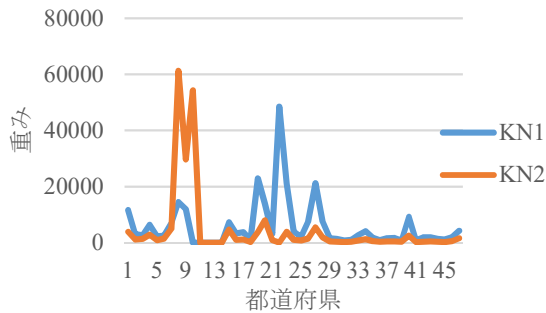


図-4 NPTS 行列H

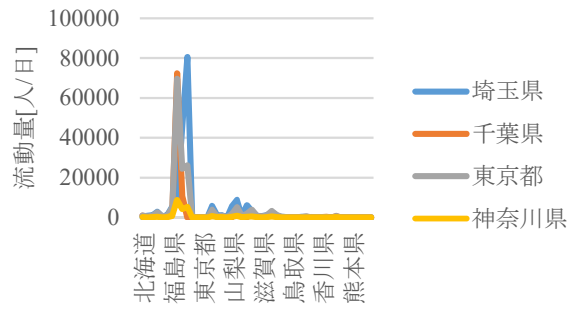


図-8 MOBI 国勢調査と共通する部分

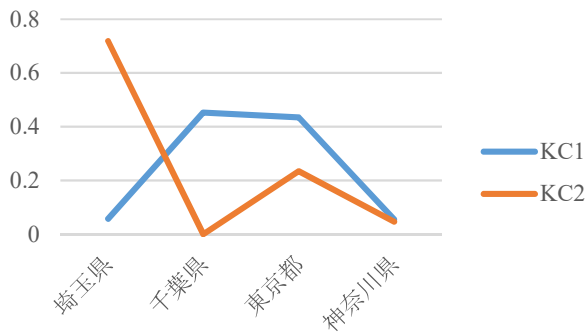


図-5 国勢調査(通勤) 行列W

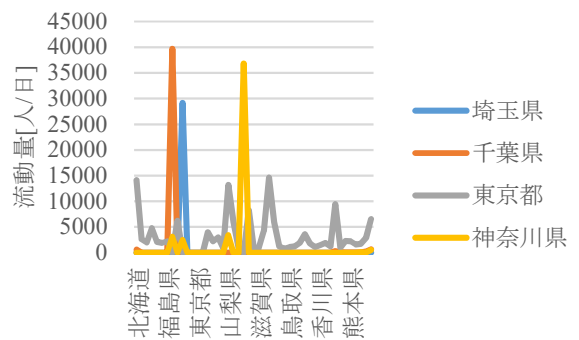


図-9 MOBI 行列A

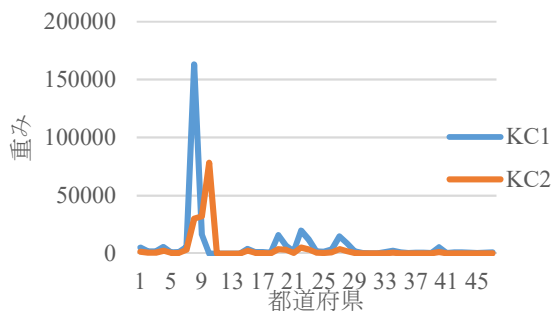


図-6 国勢調査(通勤) 行列H

NPTS, 国勢調査と共通する部分は、それぞれの調査の NMF 分析から考察できた流動が抽出されていることがわかる。相違部分である行列 A から求めた流動量を見ると、埼玉、千葉、神奈川を出発地とする流動は、それぞれ群馬、茨城、静岡への流動が突出しており、MOBI による観測の特徴だと考えられる。東京都はほかの 3 県

と異なり、全国に向け流動が発生していることが特徴である。

図-10は MOBI で観測した流動量のうち、NPTS と共通、国勢調査と共通する部分、どちらも異なる部分(残差)として算出した流動量が占める割合を示す。神奈川を出発地とする流動の多くは NPTS と MOBI で共通する流動が多く、埼玉の場合は国勢調査と共通する部分が多い。千葉の場合は、九州への流動では残差が占める割合が多い。それ以外への流動では、国勢調査と共通する部分の方が NPTS との共通部分より多い。東京は他の 3 県と異なり、MOBI と NPTS, 国勢調査に共通する部分が少ない。その中でも北関東、東海は共通部分が多く、北関東では国勢調査、東海では NPTS との共通部分がもう一方の共通部分よりも多い。

NPTS, 国勢調査との共通部分をそれぞれ算出することによって、MOBI の流動を通勤以外と通勤という目的

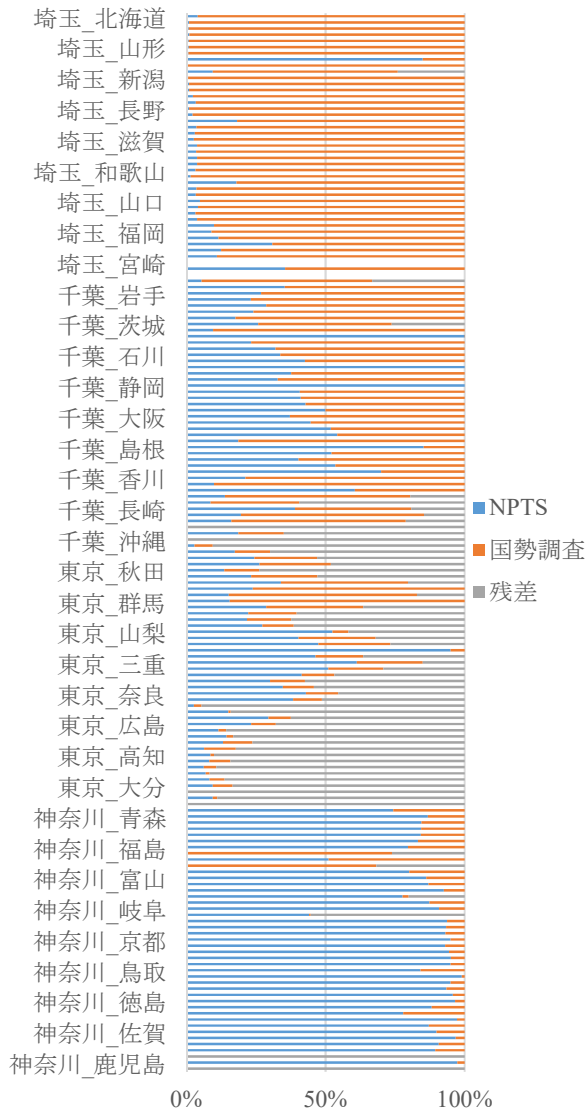


図-10 MOBI 流動量に占める共通部と相違部(関東発の流動)

別に分離を試みた。しかし、埼玉発の流動は遠方でも通勤が多く、神奈川発の流動では通勤以外がほとんどを占めるなど、実態に即していない部分が多い。この原因として埼玉の場合は、NPTS、国勢調査ともに短距離流動を表す基底への当てはまりが大きい。神奈川の場合は国勢調査の基底への当てはまりが小さいことが考えられる。今後は基底の求め方について改良を行う。

(2) その他の地域を出発地とする流動分析

OD 表を北海道・東北、北関東・山梨、南関東、北陸、東海、近畿、中国・四国、九州・沖縄ごとに集計し、NPTS、国勢調査の基底をそれぞれの地方ごとに定める。前節と同様に、2調査の基底にもとづき、MOBIと2調査の共通部分と相違部分の流動量を算出する。図-11にある出発地から同一都道府県外への流動量のうち共通部分と相違部分の占める割合を示す。

九州を出発地とする流動は、NPTS、国勢調査との共

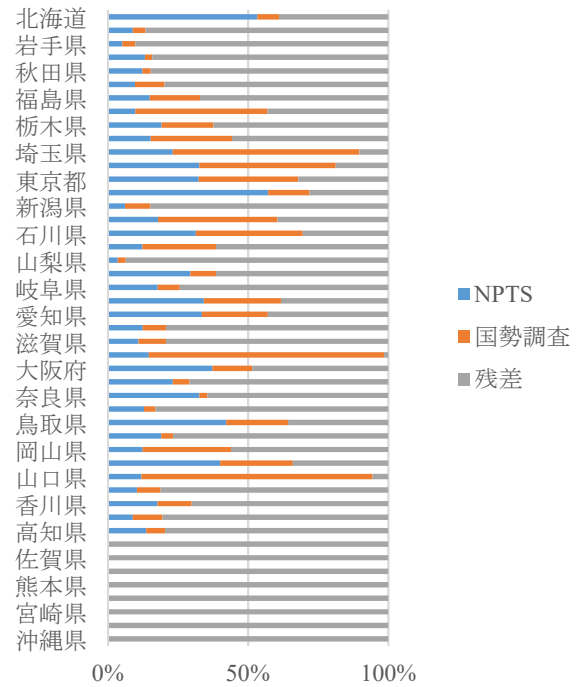


図-11 MOBI 流動量に占める共通部と相違部(発地別の流動)

通部分を推定することができなかった。これは、MOBIの観測量が0となるODが多く、式(7)の制約条件から重み行列の要素が0と推定されるためであった。今後0を多く含む、すなわち欠損が多いデータへの対処を検討する必要がある。

6. まとめ

MOBI データは、NPTS などの既存のアンケート調査に比べて、短期間のうちに複数回データを収集できるため、新たな都市間流動調査の媒体として期待されている。しかし、調査方法の違いから、収集する純流動データの対象が異なるため、過去の純流動データと比較をすること、データの精度検証が不十分なことが懸念される。

本研究では、2015 年の秋を対象に MOBI の流動を NPTS と国勢調査の NMF 結果にもとづく基底によって分解を行った。特に、NPTS と国勢調査それぞれに共通した部分を推定することで、MOBI では把握できなかった、旅行目的別の流動量を推定することを試みた。

既存の調査と共通する部分と相違する部分の割合を推計することで、比較が困難な MOBI データの調査特性を明らかにできる可能性が示せた。

課題として、目的変数としている OD 表に 0 が多くと制約条件から行列 H^M の多くの要素が 0 となるといったことが残る。

今後の発展として、NPTS の旅行目的別 OD 表から基

底を定め、MOBI の流動量を目的別に推計できる可能性がある。しかし、MOBI を記述する基底数が増えることは、多次元の行列をより少ない次元で表現できる NMF の利点を活かさないため、方法論の検討が必要である。

票データ及び都市間交通データを用いた国内宿泊観光行動の時系列分析, 土木学会論文集 D3, Vol.69, No.5, pp.I_533-I_543, 2013

参考文献

- 1) 観光庁：旅行・観光消費動向調査, 2017
- 2) 高橋央亘, 浅田拓海, 有村幹治：北海道新幹線開業前後における函館エリア来訪者数の比較分析, 土木学会論文集 D3, Vol.74, No.5, pp.I_827-I_835, 2018
- 3) 清家剛, 三牧浩也, 原裕介, 小田原亨, 永田智大, 寺田雅之：まちづくり分野におけるモバイル空間統計の活用可能性に係る研究, 都市計画論文集, Vol.46, No.3, pp.451-456, 2011
- 4) 菊池雅彦, 岩館慶多, 羽藤英二, 茂木渉, 森尾淳：全国 PT 調査データと携帯電話基地局データを用いた地方都市での OD 表の実務的推計, 土木学会論文集 D3, Vol.74, No.5, pp.I_677-I_691, 2018
- 5) 鈴木新, 山口裕通, 福田大輔：データ融合による日別・旅行目的別都市間旅客流動量推計, 運輸政策研究, Vol.21, pp.48-59, 2019
- 6) Paatero, P., Tapper, U. (1994). Positive matrix factorization: A non - negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2), pp.111-126.
- 7) 川上陸, Han-Dirk Schmocker：モバイル空間統計のデータ特性を考慮した OD 推計手法：京都観光地間流動におけるケーススタディ, 第 59 回土木計画学研究発表会・講演集, 114, 2019
- 8) 日比野直彦, 佐藤真理子, 森地茂：複数の観光統計の個

(2019. 10. 4 受付)

A CHARACTERISTICS OF INTER-REGIONAL TRAVEL FLOW USING MOBILE PHONE DATA

Yushi SUGAHARA, Makoto TSUKAI

The development of information and communication technology has made it possible to conduct inter-regional travel surveys in a short span, which was difficult in the past. Mobile phone data, which are attracting attention as new inter-regional travel data, are excellent in data promptness, collection amount, and accumulation, and research on inter-regional travel using the data is underway. However, it is difficult to confirm the compatibility with past data, and many problems are pointed out in the accuracy and reliability of the survey. In this study, the base describing the OD table of inter-regional travel from existing surveys is determined by non-negative matrix factorization, and the mobile data is shared with the existing surveys by describing the OD table of mobile data by that base. Is separated into different parts. It is hoped that this will enable comparison between existing and new surveys, and verification of specific parts that appear in new surveys.