

# 非負値テンソル因子分解を用いた 旅行需要の北陸新幹線開業に対する感度の計測

山口 裕通<sup>1</sup>・三宅 寛之<sup>2</sup>・中山 晶一郎<sup>3</sup>

<sup>1</sup>正会員 金沢大学助教 理工研究域地球社会基盤学系 (〒 920-1192 金沢市角間町)  
E-mail: hyamaguchi@se.kanazawa-u.ac.jp

<sup>2</sup>学生会員 金沢大学大学院 自然科学研究科 (〒 920-1192 金沢市角間町)

<sup>3</sup>正会員 金沢大学教授 理工研究域地球社会基盤学系 (〒 920-1192 金沢市角間町)  
E-mail: nakayama@staff.kanazawa-u.ac.jp

近年活用が進みつつある携帯電話位置情報データを用いることで、広範囲かつ膨大なサンプルの人口分布情報を、長期かつ時間解像度も高く把握することが可能である。一方で、交通サービスに対する感度を推計し将来予測するために重要である、旅行目的などの「人の意思・考え」の情報が含まれないという欠点がある。この課題に対して、本研究では、膨大な空間・時間情報を用いて旅行目的に対応する行動分類情報を推計することを試みる。具体的には、非負値テンソル因子分解を用いて、北陸新幹線に対する感度が異なるグループを抽出した。その結果、新幹線開業前後においてその増加比が大きく異なる、いくつかの旅行パターンを分類することができた。

**Key Words:** mobile phone data, demand model, non-negative tensor factorization, high speed rail

## 1. はじめに

新幹線の開業は、より容易に長距離移動を可能にすることができるため、長距離行動の構造を顕著に変化させることが期待される。このような、新幹線の整備による要距離旅行構造の変化は、国内における広域の交流行動の活発化を意味しており、新幹線整備において実現すべき主要な効果である。そのため、新幹線整備の計画やその受け入れ準備を検討するためには、長距離旅行パターンへの効果に着目して、それを予測する交通需要予測が不可欠である。そのため長距離旅行需要予測モデルが多く提案されてきている<sup>1) 2)</sup>。

これらのモデルでは、新幹線整備などの交通サービス変化に対する需要の変化を精度よく推定するために、感度を示すパラメータは3つのグループ（観光・業務・私用）ごとに別に推定されてきた。しかし、このグループは、「直感的に妥当な項目」としてアンケート調査しやすいものを適用しているに過ぎず、「新幹線整備に対する感度が異なるグループ」としての妥当性はほとんど検討されていない。つまり、代表的な3パターンのほかにも感度の異なるグループがある場合、従来モデルのアプローチでは、精度よく長距離旅行行動の変化を予測することは困難である。

一方で、携帯電話位置情報をはじめとするパッシブ型の位置情報ビッグデータは、大量のサンプル（携帯電話ユーザ）の位置情報を、広範囲かつ高頻度で取得

した信頼性の高いデータであり、国レベルでの長距離旅行の実態を俯瞰的かつ高頻度に把握することが可能である。すでに、Ahas et al. (2007, 2008)<sup>3) 4)</sup> や室井ら (2015)<sup>5)</sup>, Janzen et al. (2016)<sup>6)</sup>, 鈴木ら (2019)<sup>8)</sup> をはじめとして、観光旅行・長距離旅行の行動分析に活用され始めている。これらのデータを活用することで、新幹線開業に対する「旅行需要の感度（弾力性）」をより詳細に把握できる可能性がある。

しかし、これまでの研究では、従来調査を代替に向けた方法論が主に検討されてきた。これらは、携帯電話位置情報の「空間範囲の広さと空間解像度の高さ」に着目して、旅行目的などの情報を付与しながら<sup>6) 7) 8)</sup>、従来型の長距離旅行分析・需要予測に必要なデータを取得しようとするものである。これにたいして、携帯電話位置情報には、面的な移動情報を高頻度で取得しているという「時間的に高い解像度」・膨大なサンプル数があることによる「制度の良い空間・属性の構成情報」という特徴がある。これらの情報を適切に分解することで、従来型調査で設定されてきた「旅行目的分類」よりさらに細かい分類ごとに、新幹線開業に対する感度の情報を入手できる可能性がある。

そこで本研究は、より高精度な需要予測方法の開発を目指して、アンケート調査による分類を用いずに、より細かいグループの新幹線に対する感度を分類・推定する方法論を検討・提案することを目的とする。具体的には、NTTドコモが提供する携帯電話位置情報を用いた

人口分布統計データであるモバイル空間統計に、非負値テンソル因子分解（以下 NTF : Non-negative tensor factorization）<sup>9)</sup> を拡張したモデルを適用して、新幹線に対する感度が異なるグループを抽出を試みる。

そして、近年あった大きな交通サービスレベルの変化として、2015年3月14日の北陸新幹線の延伸開業がある。この新幹線の整備によって、東京駅-金沢駅間をそれまでの所要時間から1時間20分も大幅に短縮し、2時間30分で同区間を移動することが可能になった。山本(2016)<sup>10)</sup>の研究ではNTTドコモが提供する携帯電話位置情報を用いた人口統計データであるモバイル空間統計を用いて石川県への訪問客数を分析した結果、北陸新幹線の開業前後で石川県への観光客数が増加したことを確認しているものの、将来予測を可能とするような、グループごとの感度の推計はできていない状態である。さらに、近い将来には福井県への延伸が決定しており、この延伸に際する需要予測はマーケティングなどの施策を開業前に講じるためには非常に重要であり、石川県で観測された旅行パターン構造の変化を参考にできる可能性が高い。そこで、本研究では個々の北陸新幹線を対象として、従来の3種類の旅行目的の感度だけでなく、季節や曜日を含む長期の時間変動や個人属性（性別や年齢）・より細かい空間の情報といった携帯電話位置情報から得られる膨大な情報を駆使しながら、新幹線開業に対する感度の推定を実施していく。

## 2. モバイル空間統計によるテンソルデータ

本研究では、「モバイル空間統計」を用いて分析を行う。このデータは、NTTドコモ携帯電話の基地局との通信で得られる位置情報データから作成された1時間ごとの人口分布データであり、性別・年齢・居住地・滞在地の情報も有している。

本研究では、北陸新幹線整備の効果を分析するために、( $d$ : 2014年3月1日~2018年2月28日の1,461日の時点)  $\times$  ( $c$ : PM1時台の石川県19市町村の滞在地)  $\times$  ( $a$ : 性別(男女)  $\times$  年齢(20s~70sの10歳ごと)の合計12分類の個人属性) ごとの北陸3県(富山県・石川県・福井県)外居住者の石川県滞在人数  $y_{d,c,a}$  を分析する。具体的には、「3階テンソル」として解釈してテンソル分解に適用していく。なお、PM1時台の滞在推計人口を分析する理由は、(日帰り・宿泊を含む)長距離旅行が最も観測される時間と想定されるためである。清家ら(2010)<sup>11)</sup>の研究ではモバイル空間統計を用いた1日の人口推移によれば昼間人口、特に13時台の滞在人口が最も多いことを示しており、それと同じ時間を本研究でも適用していく。

## 3. 長距離旅行パターン分解のためのNTFモデル

### (1) 非負値テンソル因子分解

本研究では非負値テンソル因子分解のうちの一つの手法であるCP分解を取り扱う。CP分解とは所与のテンソルを1階テンソル(ベクトル)の外積の和に分解する方法である。ここで、人口分布データ  $y_{(d,c,a)}$  の分解は以下式のように書くことができる:

$$y_{d,c,a} = \sum_{k \in (1,;K)} w_{k,d} h_{k,c} n_{k,a} + e_{d,c,a}, \quad (1)$$

$$\forall (d, c, a) \in (D \times C \times A).$$

$$w_{k,d} > 0, \quad \forall (k, d) \in [(1,;K) \times D],$$

$$h_{k,c} > 0, \quad \forall (k, c) \in [(1,;K) \times C], \quad (2)$$

$$n_{k,a} > 0, \quad \forall (k, a) \in [(1,;K) \times A],$$

$$\sum_{c \in C} h_{k,c} = 1, \quad \forall k \in (1,;K), \quad (3)$$

$$\sum_{a \in A} n_{k,a} = 1, \quad \forall k \in (1,;K)$$

ここで、 $D, C, A$  は、それぞれ分析対象とする日付  $d$  の集合(1,461日分)、石川県内市町村  $c$  の集合(19市町村)、属性情報  $a$  の集合(2性別  $\times$  6年齢階層)に相当する。そして、 $w_{k,d}, h_{k,c}, n_{k,a}$  が分解後の変数であり、 $k$  は基底番号、 $K$  が基底数、 $e_{d,c,a}$  が残差(誤差)に相当する。

まず、式(1)が分解の基本式であり、旅行数  $y_{d,c,a}$  を3種類の変数  $w_{k,d}, h_{k,c}, n_{k,a}$  に分解している。ここで、分解後の3つの変数は、それぞれテンソルにおける3種類の情報に対応し、基底  $k$  とそれぞれ1種類の情報の組み合わせごとに定義している。つまり、 $w_{k,d}$  は  $1,461 \times K$  個、 $h_{k,c}$  は  $19 \times K$  個、 $n_{k,a}$  は  $12 \times K$  個の変数であり、ここでは  $1,461 \times 19 \times 12$  個のデータを、 $(1,461 + 19 + 12) \times K$  個に圧縮することに相当する。

これを、概念的に示したものが図-1である。この図のように、式(1)の分解は、基底  $k$  毎に各変数の情報をベクトル化した、時間変動  $w_k$ ・旅行先構成  $h_k$ ・属性構成  $n_k$  の3つのベクトルの外積で得られる3階テンソルを、基底の数だけ和算した形に分解するものとして解釈できる。

そして、式(2)が、すべての分解後の数値に対する非負制約であり、“非負値”テンソル因子分解の最大の特徴である。さらに、式(3)のように、 $h_{k,c}, n_{k,a}$  には、各基底ごとに成分の和がゼロになる、つまり旅行先構成ベクトル  $h_k$  と属性構成ベクトル  $n_k$  は、それぞれ成分の和が1であるという制約をつける。この結果、 $h_{k,c}, n_{k,a}$  は、それぞれ基底  $k$  における、旅行先  $c$  あるいは属性  $a$  の“構成割合”を意味する。そして、 $h_k, n_k$  のベクトルの成分が非負かつ和がそれぞれ1となる割合の情報で

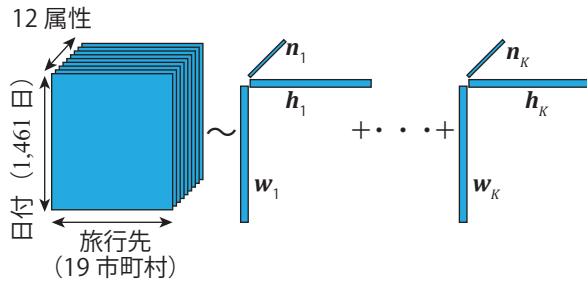


図-1 非負値テンソル因子分解 CP モデルの図解

あるため、 $w_{k,d}$  は  $y_{d,c,a}$  は同じ単位 (人) となる。つまり、 $w_{k,d}$  は、 $d$  日に基底  $k$  の旅行パターンで石川県に訪れる人数であり、その旅行先と属性の内訳が  $h_k, n_k$  で定義される。

そのうえで、非負値テンソル因子分解では ( $w_{k,d}, h_{k,c}, n_{k,a}$  の全成分を並べて行列化した,)  $W, H, N$  を、ノルム (本研究ではフロベニウスを用いていく) が最小になるように、推定する。これを数式で表記すると、以下ようになる:

$$\begin{aligned} \min_{W, H, N} \sum_{d \in D} \sum_{c \in C} \sum_{a \in A} (e_{d,c,a})^2, \\ \text{s.t.} \quad \text{eq.(1), eq.(2), eq.(3)}. \end{aligned} \quad (4)$$

## (2) 新幹線整備に対する感度が異なるグループ推定のアイデア

本研究では、非負制約を加えた NTF をベースとしたモデルを活用して、北陸新幹線開業に対して感度が異なるグループごとの情報に分解することを目指す。このアプローチには 2 点の特徴がある。

まず、1 点目は、「非負値」への分解を目指すポイントである。各成分が非負値と仮定する、最も大きい理由は、もともとのデータが複数の旅行目的等のグループの人数 (非負値) の和算によって成り立っており、現実の加法性や実現象と一致している点である。NTF で分析するデータであるモバイル空間統計は滞在人口推計を表すデータであり、人が毎日・任意の時間にどこにいるのかを属性込みで把握できる。対象地域に滞在している人は、何かしらの目的をもって滞在していることに留意すると、各滞在人口データは観光目的の人数+業務目的の人数+私用目的の人数+…のような足し算で表現され、それぞれが異なる新幹線整備に対する異なる感度を持つと解釈できる。NTF における分解は、このような構成にほぼ合致しており、基底 = 感度が異なるグループ (旅行目的グループ) として解釈できる可能性がある。

一方で、主成分分析や独立成分分析といった手法を、同様に適用したケースについて考えていこう。これら

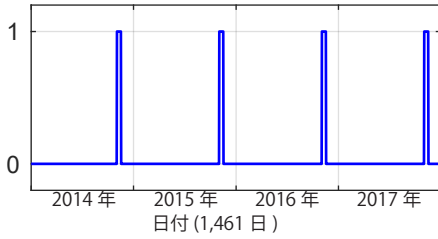
のモデルではグループの足し算で表現せず、「足し引き」でデータを表現する形になる。このような場合は、Lee and Seung (1999)<sup>9)</sup> が顔の画像情報の分解で示しているように、加法的に構成される各パーツを抽出するような結果を得られない可能性が高い。このとき、非負値の分解であれば、パーツ以外の部分はなるべく抽出されない (要素がゼロとなる) スパースな要素を導出する特徴があり、画像の分解などではおおむね良好にパーツを抽出することが可能であることを示している。本研究で用いる長距離旅行数のテンソルデータにおいても、これと同様に総旅行数を構成する「パーツ (グループ)」を、基底として時間変動・旅行先・属性の規則性から導出することで、旅行目的に相当するような北陸新幹線開業に対する感度が異なるグループを位置情報データのみから導出できると期待される。

つぎに、2 点目としては、旅行数情報を 3 種類 (日付単位の時間変動  $w_{k,d}$ ・市町村単位の旅行先割合  $h_{k,c}$ ・性別年齢の情報  $n_{k,a}$ ) に分割したテンソルを用いる点である。この 3 種類の情報のうちでは、新幹線整備の前後の変化は、すべて日付単位の時間変動  $w_{k,d}$  で表現されることとなる。つまり、ほかの二つの情報は新幹線整備前と後で共通である。そのうえで、本研究では、後述するように時間変動情報のなかに、新幹線整備前・後で共通するような周期パターンを導入していく。このことによって、「業務目的の移動は、平日に 30 代~60 代の男性による大都市や工業地への移動が大半を占める」といった、旅行目的に対応する時間的な周期性・属性・旅行先の特徴を示す情報が得られると推測される。そのうえで、それぞれの新幹線整備に対する感度を分析できるようにモデルを工夫していく。

また、テンソルの形にしなくても、列方向の情報として属性 × 旅行先の組み合わせを全て並べた「行列」を考えると、3 階テンソルではなく行列の形で分解することも可能である。しかし、将来予測において場所を移転することを考えると、本研究のようにテンソルのまま分解するアプローチが望ましいと考える。なぜなら、グループごとの属性の組み合わせの情報や時間の周期的な変動の情報は、別の地域であっても同じ構成である (空間移転性がある) 可能性が高いと推測される。つまり、旅行先情報のベクトルだけ、予測対象地域の来訪者数の変動に合わせて改めて推定することができれば、従来のデータから観測された属性・時間変動で定義されるグループごとの感度を、外挿して予測するアプローチが考えられる。本研究では、石川県における北陸新幹線が開業したときのデータしか存在せず、外挿の妥当性は確認できないが、将来予測を念頭にしたテンソルアプローチを適用して感度が異なるグループの推計を試みる。



新幹線開業前後に共通



新幹線開業後のみ

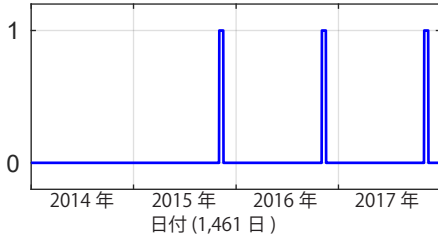


図-2 周期成分ダミー項  $m_{d,p}$  の例 (1 年周期成分)

### (3) 時間変動情報の周期パターンへの分解

ここでは、現在の 1,461 日分の時系列情報  $\mathbf{W}$  について、いくつかの周期的な成分にまとめる方法を説明していく。具体的には、1 年間の中で週周期・月周期・年周期 (季節変動) と、カレンダーで定義される大型連休や祝日の情報をもとに、時系列変動成分を更に分解していく。このように、時間変動情報  $\mathbf{W}$  を、新幹線開業前後で共通する周期成分と新幹線開業前後の変化に分解することで、より明確に各基底 (グループ) の旅行の特徴と新幹線開業に対するグループごとの需要の感度を解釈することが可能となる。

この分解のために、周期パターンを表現するダミー変数行列  $\mathbf{M} = |m_{d,p}|_{(D \times P)}$  を考える。これは周期成分  $p$  に日付  $d$  が該当すれば 1, それ以外は 0 を取る成分から構成される行列である。そして、各周期ごとの新幹線開業による効果を検出するために、全期間を通じた周期成分だけでなく、開業前はすべて 0 で開業後のみに周期成分が発現ような周期成分  $p$  も用いていく (図-2)。そして、本研究では、月成分・週成分・曜日成分・代表的な連休成分などの 50 種類の周期 (開業前後で合計 100 種類) の成分に時系列情報を分解していく (表-1)。

この行列  $\mathbf{M}$  の係数を  $x_{p,k} > 0$  (基底 (グループ)  $k$  における、周期成分  $p$  に対応する重み) とすると、式 (1) は以下の式のように書き換えられる：

$$y_{d,c,a} = \sum_{k \in (1,;K)} \sum_{p \in (1,;100)} m_{d,p} x_{p,k} h_{k,c} n_{k,a} + e_{d,c,a}, \quad \forall (d, c, a) \in (D \times C \times A). \quad (5)$$

そして、この時の未知変数は  $x_{p,k} h_{k,c} n_{k,a}$  であり、これ

表-1 分解に用いた周期成分一覧

周期成分名	成分の数
<b>周期成分</b>	
1 年周期	24 (2 × 12 カ月分)
7 日周期 (曜日)	7+1 (「祝日」)
<b>特別な日付</b>	
年末年始	5 日分 (12/30 × 1/3)
お盆	4 日分 (8/13 × 8/16)
3 連休	3
4,5 連休	4
7 日周期 (曜日)	7+1 (「祝日」)
小松基地航空祭	2 (当日と前日)

らを式 (4) で示したフロベニウスノルム最小化で推定していく。

### (4) 基底数の決定方法

NTF において重要な課題が基底数の決定である。基底数  $K$  を増やしていけばいくほどに、推定したモデルと持っているデータとの残差は小さくなっていくが、その分だけ過学習を起こしてしまうリスクが高まる。過学習とは現在のデータに対する適応度は高いが、未知のデータに対しての適応度が低いことを示す。逆にグループ数を少なくすれば残差が大きくなり、現在のデータに対する適応度が低くなり、結果として予測精度も低下してしまう可能性が高い。そのため、元のデータを表現できるのであれば、多くの基底数ではなく最小限の基底数で表現すべきという、オッカムの剃刀の概念に基づいて基底数を決める必要がある。

この残差と過学習のトレードオフ関係のバランスを取るためにモデル選択基準という考え方がある。どの基準を使うのかについては様々な議論があり、やや研究者の恣意的な側面で決められているところがあるが、データと将来的な (未知の) データとのバランスを汎化誤差の概念で定義された情報量基準で行うモデル選択は、試行錯誤した探索的な決定よりも客観的であろう。ここでいう探索的とは基底数を変えて複数回計算し、説明がつくグループに分類ができていない基底数を選択することを指す。

モデル選択基準はいくつか存在し、代表的なもので AIC (赤池情報量基準) と MDL (記述長最小化原理) が挙げられる。それぞれのモデル選択基準の算出式は以下のとおりである：

$$\text{AIC} = -2 \ln(L) + 2\theta, \quad (6)$$

$$\text{MDL} = -2 \ln(L) + \theta \ln(N).$$

ここで  $\ln(L)$  は対数尤度関数を表し、 $\theta$  はパラメータ数

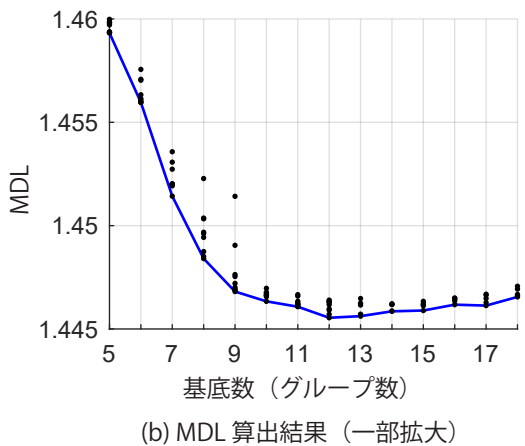
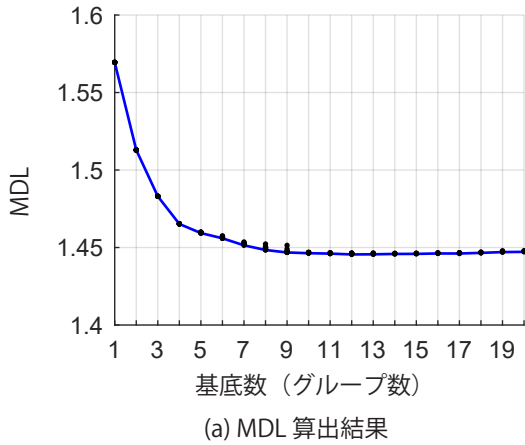


図-3 MDL 推定結果と基底数の関係

表し、 $N$  はサンプル数を表す。式の第 1 項は今持ちうるデータに対するモデルの当てはまりの良さを示し、第 2 項は過学習を防ぐための罰則項を示している。この 2 つの基準式の大きな違いは第 2 項にある。MDL には第 2 項にサンプル数が含まれており、AIC には含まれていない。そのためサンプル数が膨大な場合、AIC ではより複雑なモデルを選択する傾向にある。また、一般的に AIC はサンプル数の少ないときに有効とされており、本研究では膨大なサンプル数を扱うため、MDL を採用して基底数の決定を行う。

#### 4. NTF モデルによる分解結果

##### (1) 基底数の決定

本研究では MDL 基準を用いて基底数を決定する。図-3 は基底数  $K$  を、1~18 に設定したうえで、初期値を乱数で発生させて Lee and Seung のアルゴリズム<sup>12)</sup> で分解を実施することを、150 回繰り返して MDL を算出した結果である。基本的には、非負制約を含む本モデルの推定は、解が一意であることは保証されてい

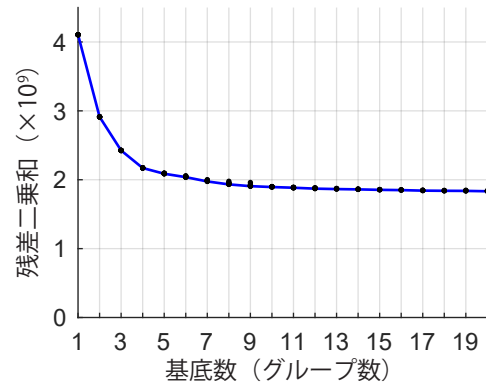


図-4 残差二乗和と基底数の関係

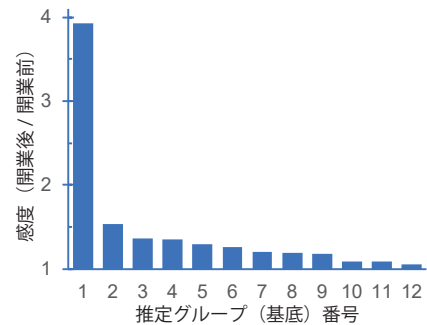


図-5 グループごとの北陸新幹線開業前後比 (感度)

い。そのため、同じ基底数  $K$  を設定していても図-3 に示すように、MDL の算出結果は異なる値をとる。この図の青い線は、150 回繰り返した中での最小値をつないだものである。この MDL の最小値と  $K$  との関係を見ていくと、基底数  $K = 12$  の時に MDL が最小の値となることが分かった。

あわせて、基底数  $K$  と残差二乗和 (フロベニウスノルム) との関係を示した図-4 を見ると、 $K = 1$  の残差と比較して、 $K = 12$  ではおおよそ半分になっていることが分かる。さらに、これ以上基底数を増加させても、残差はそれほど小さく成ら解ことが分かる。そこで、本研究では最小値をとった  $K = 12$  を採用し、その分解結果を確認していく。

##### (2) NTF による分解結果

###### a) 新幹線整備に対する感度の概要

本研究ではグループ (基底) 数を 12 個として人口分布のテンソルデータを分解した。その結果として得られたグループの特徴と、新幹線整備に対する感度 (日平均滞在者数の新幹線開業前後比) をまとめたものが図-5 である。

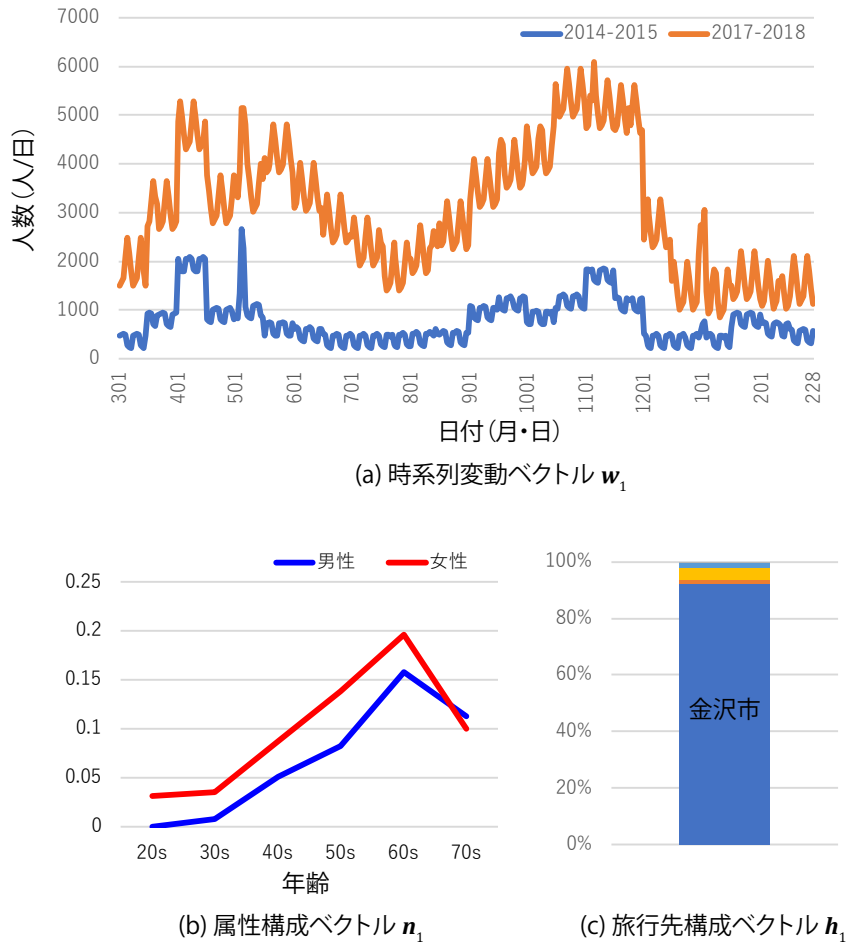


図-6 グループ 1 の推定結果

なお、NTF から得られるグループには特に順序の情報はないので、本稿では新幹線整備に対する感度が最も大きいグループを「グループ 1」と称し、以降その順番で「グループ 2」、「グループ 3」、... と表記していく。

図-5 をみると、分解されたグループごとに新幹線整備前後比（感度）が大きく異なることが分かる。感度が最も大きいグループ 1 では 3.93 倍であった一方で、最も小さいグループ 12 では 1.05 倍に過ぎない。以降では、代表的な 3 グループをピックアップしてその特徴を確認していく。

**b) グループ 1 の特徴**

まずグループ 1 の各要素の推定結果は図-6 に示すとおりである。ここでは、(a) に時間変動ベクトル  $w_{(k=1)}$  を、(b) に属性構成ベクトル  $a_{(k=1)}$ 、(c) に旅行先構成ベクトル  $c_{(k=1)}$  を示している。

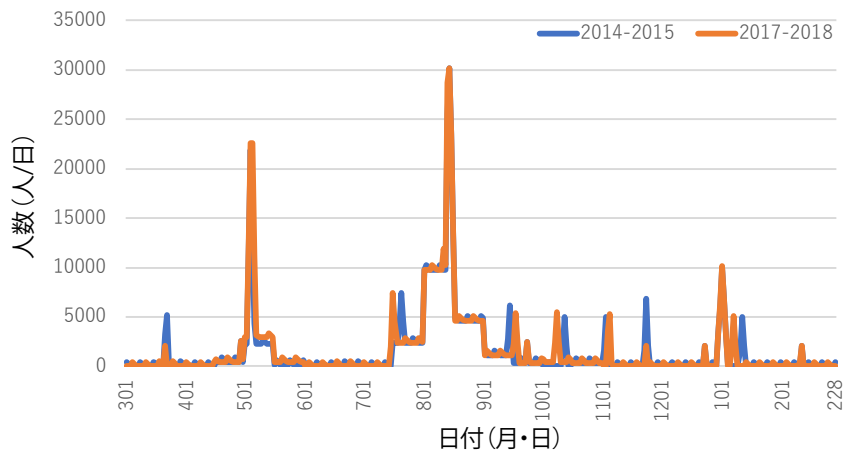
これらの図から、このグループの特徴として、以下の 5 点の特徴があげられる：まず、開業前の旅行者数（図-6 (a) の紺色線）を見ると、(1) 平均的な来訪人数は、1,000（人/日）程度である。(2) とくに GW と桜の時期である 3 月終わりから 4 月始めにかけて、最も来

訪人数が多く平均の 3 倍以上となる。(3) 普段の変動を見ると、休日 > 平日の傾向がある。以上のような時間変動の特徴から、このグループ 1 はおおむね「観光旅行」に相当するものと推測できる。

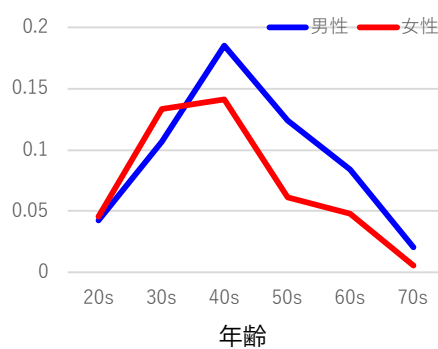
つぎに、図-6 (b) から属性構成をみると、(4) 男女ともに 40 歳代以上を幅広く含まれるグループであるが、全体的に男性より女性の方が多く、年齢ではとくに 60 歳代が高い割合を占めていることがわかる。図-6 (c) から旅行先構成をみると、(5) 旅行先としては金沢が 90% を占め、七尾市・輪島市・珠洲市で残りの 10

このようなグループにおける、開業後の推移（図-6 (a) の橙色線）をみると、開業前と比較してほぼすべての時点で大きく増加していることが確認できる。

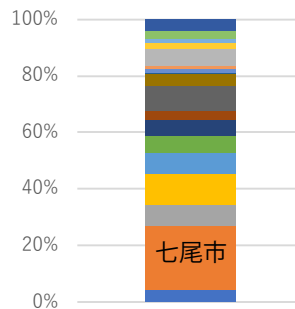
以上から、グループ 1 は、観光目的で（GW をはじめとする休日に多く）金沢市を訪れる 50 代・60 代が中心のグループであり、このようなタイプの旅行で特に新幹線開業による変化が大きかった（感度が高い）ことがわかる。



(a) 時系列変動ベクトル  $w_{12}$



(b) 属性構成ベクトル  $n_{12}$



(c) 旅行先構成ベクトル  $h_{12}$

図-7 グループ 12 の推定結果

### c) グループ 12 の特徴

つぎに、新幹線開業による感度が最も小さかった、グループ 12 の特徴を図-7 からみていこう。

この図から、このグループの特徴として、以下の 5 点の特徴があげられる：まず、開業前の旅行者数（図-7 (a) の紺色線）を見ると、(1) 普段の来訪人数はほとんどゼロであるが、特定の時期に限って 1 日当たり数千～数万人もの来訪者数がある。(2) 最大となる時期はお盆であり、石川県全体で 30,000 (人/日) もの人が来訪している。(3) 次に多いのが GW であり、20,000 (人/日) 以上の人々が来訪している。以上のような時間変動の特徴から、このグループ 12 はおおむね「帰省旅行」に相当するものと推測できる。

つぎに、図-7 (b) から属性構成をみると、(4) 男女ともにおよそ同じ割合で、年齢が 30 歳代と 40 歳代の比較的若い世代で多い傾向にある。図-7 (c) から旅行先構成をみると、(5) 滞在地割合は金沢市が 4%ほどに対して、他の地域（ほとんどのグループで 0%に近い割合を占めている市町村）が 95%以上を占めている。最も多く占めるのは七尾市である。

このようなグループにおける、開業後の推移（図-7 (a) の橙色線）をみると、開業前と比較してほとんど変わっておらず、新幹線整備に対する感度がほとんどないことが分かる。

以上から、グループ 12 は、GW・お盆の比較的若い世代による金沢市以外への（おもに帰省などが中心と思われる）旅行であり、新幹線整備にもかかわらず、ほとんど増加しなかったことが分かる。

なお、「帰省旅行」については、ここで示した時期以外に年末年始にも多くの人々が旅行することが想定されるが、その分はグループ 11 として別のグループとして推定されている。こちらのグループの詳細は、旅行先グループの内訳で金沢市が比較的多い以外は、新幹線整備に対する感度も含めて、このグループ 12 とかなり類似した特徴を持つ。

### d) グループ 8 の特徴

つぎに、新幹線開業による感度が中程度であった、グループ 8 の特徴を図-8 からみていこう。

この図から、このグループの特徴として、以下の 5 点の特徴があげられる：まず、開業前の旅行者数（図-

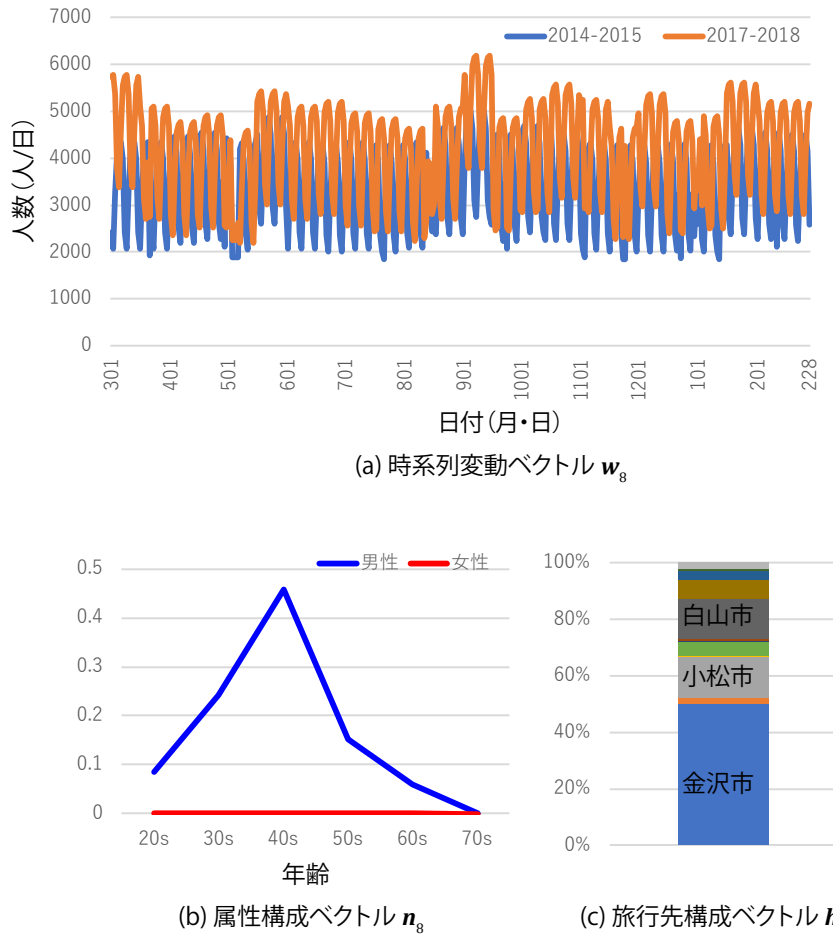


図-8 グループ 8 の推定結果

8 (a) の紺色線)を見ると、(1) 普段の来訪人数は、2,000~5,000 (人/日) 程度である。(2) 季節的な周期はほとんどなく、1週間周期で変動するのみである。(3) 1週間周期の変動を見ると、圧倒的に平日>休日の傾向がある。以上のような時間変動の特徴から、このグループ 8 はおおむね「平日の業務旅行」に相当するものと推測できる。

つぎに、図-8 (b) から属性構成をみると、(4) 女性がほとんどゼロであり、ほとんどを 30 歳代・40 歳代・50 歳代の男性が占める。図-8 (c) から旅行先構成をみると、(5) 滞在地割合は金沢市が 50%ほどである特徴に加えて、小松や白山といった有名な工業地区のある市町村で合わせて 20 これらの特徴からも、このグループが「業務目的」に相当する情報であることが確認できる。

このようなグループにおける、開業後の推移(図-8 (a) の橙色線)をみると、開業前と同じ時間変動の傾向のまま全時点を通じて 1.2 倍に増加していることが確認できる。

以上から、グループ 8 は、平日に 30-50 歳代男性が金沢市・小松市・白山市を訪れる(おもに業務と思われ

る)旅行であり、新幹線整備によって 1,2 倍に増えたことがわかる。観光と比較すると小さいが、このような業務タイプの旅行も新幹線開業前後で大きく増えていることが確認できた。

#### e) 全 12 グループの特徴まとめ

以上までの考察と同様に、全 12 グループの特徴をまとめた表が、表-2 である。

この表から、分解された旅行パターン・グループには、以下の 4 点の特徴があることが分かる。(1) NTF アプローチで分解した結果、12 種類の新幹線に対する感度が異なるグループが抽出されており、それぞれの分類はおおよそ従来まで用いられてきた 3 目的をより細かく分類した形として解釈可能である。(2) 本分析の結果では、業務目的と帰省目的がそれぞれ 2 グループずつと、観光と考えられるものが 8 グループに分解された。(3) 「帰省」目的と考えられる 2 種類は、共通して新幹線に対する感度が非常に小さい特徴がある。(4) 「観光」と分類した目的が、最も感度の差異が大きく、金沢市内に休日に来訪する 50 代・60 代(女性)の旅行パターンは、新幹線開業に対する感度が大きかったの



表-2 全 12 グループ推定結果の特徴まとめ

グループ 番号	目的 (推測)	季節変動	曜日変動	属性構成	旅行先	新幹線増加 2018/2014	新幹線増加 (人/日)
1	観光	-	休日 > 平日	50s60s	金沢	×3.93	+2,285
2	観光	-	休日 > 平日	20s30s	金沢	×1.53	+840
3	航空祭	-	休日 > 平日	30s-50s 男	小松	×1.36	+111
4	観光	GW	休日 > 平日	30s40s	金沢	×1.35	+550
5	観光	-	休日 > 平日	30s-50s 男	金沢	×1.29	+580
6	観光	-	休日 > 平日	30s 女	金沢, 加賀	×1.26	+556
7	業務	-	平日 > 休日	30s-50s 男	金沢	×1.20	+348
8	業務	-	平日 > 休日	30s-50s 男	金沢, 小松	×1.19	+670
9	観光	-	休日 > 平日	20s-50s 男	金沢近郊	×1.18	+150
10	観光	-	休日 > 平日	70s	金沢, 加賀	×1.09	+215
11	帰省	お盆・新年	休日 > 平日	-	-	×1.09	+165
12	帰省	GW・お盆	休日 > 平日	30s40s	金沢以外	×1.05	+82

に対して、比較的若年層や男性については感度が小さいことがわかる。

## 5. おわりに

本研究では、新幹線整備に対する旅行行動の誘発需要を、精度よく把握する手法の開発を目指して、アンケート調査による「旅行目的分類」を用いずに、より細かい新幹線整備に対する感度が異なるグループを分類・推定する方法論を提案した。具体的には、携帯電話位置情報データから得られる、長期の時間・空間・属性情報のテンソルデータを、非負値テンソル因子分解にさらに時間変動情報を周期ごとに分解するモデルを用いるアプローチを提案した。

そして、提案した手法を、北陸新幹線整備の前後を含むモバイル空間統計データに適用した結果、MDL 基準によってグループ数は 12 個に決定し、3 種類よりも多くの感度を持つグループが抽出されることがわかった。最も感度の高いグループの開業前の時系列変動は、「GW などを中心とする平日に金沢市に訪れる、50-60 歳代が多くを占める（観光と推測できる）旅行グループ」であり、このグループでは新幹線開業前後で 4 倍近くも増加していることが分かった。一方で、「GW・お盆・年末年始に石川県に訪れる、比較的若年の人々による（帰省行動と思われる）旅行グループ」は、新幹線が開業してもその量はほとんど変わらず、数パーセントの増加にすぎないことがわかった。

今後の課題としては、分解されて得られた「旅行グループ（基底）」情報の特徴と解釈の妥当性の確認があげられる。現時点では、周期パターンの特徴からお

およそ旅行目的を推測しているが、非常に根拠に乏しい状態である。そこで、旅行目的も把握することができ全国幹線旅客純流動調査などの情報と比較しながら、本研究で得られた旅行行動分類の妥当性・意味をより丁寧に考察することが望ましいであろう。そのうえで、将来予測への適用を考えるためには、推定した感度の時間・空間的に移転可能な形のモデルに発展させていく必要がある。本研究では、石川県のみを対象としていたために、所要時間変化の情報などを利用していなかったが、感度に相当する情報を所要時間変化に対する弾力性にするようなモデル改良が必要であろう。そのうえで、感度情報の時間・空間的な移転性を確認するために、ほかの新幹線開業効果などと本研究で得られた感度情報を比較・検討していく必要がある。

謝辞： 本研究は、科学研究費補助金・若手研究(B)17K14736、文部科学省卓越研究員事業として支援を受けた研究による成果の一部です。このスペースをお借りして、支援に感謝申し上げます。

## 参考文献

- 1) Yijing, L. and Arash, A. and Sepehr, G.: Natinal Travel Demand Model for the U.S.: Applications to Statewide Modeling, Fuel Price and HSR, *TRB 2019 Annual Meeting*, 2019.
- 2) Fu, X. and Tae, H.O.: An Analysis of Travel Demand in Japan's Intercity Market Empirical Estimation and Policy Simulation, *Journal of Transport Economics and Policy*, Vol. 48, No. 1, pp.898-910, 2014.
- 3) Ahas, R., Aasa, A., Mark, Ü., Pae, T. and Kull, A.: Seasonal tourism spaces in Estonia: Case study with mobile positioning data, *Tourism Management*, Vol. 28, No. 3, pp.898-910, 2007.
- 4) Ahas, R., Aasa, A., Roose, A., Mark, Ü. and Silm, S.: Evaluating passive mobile positioning data for

- tourism surveys: An Estonian case study, *Tourism Management*, Vol. 29, No.3, pp. 469–486, 2008.
- 5) 室井寿明, 磯野文暁, 鈴木俊博: モバイル・ビッグデータを用いた都市間旅客交通への活用に関する研究, 土木計画学研究・講演集, Vol. 51, 2015.
  - 6) Janzen, M., Vanhoof, M., Axhausen, K. and Smoreda, Z.: Estimating Long-Distance Travel Demand with Mobile Phone Billing Data, *In 16th Swiss Transport Research Conference*, 2016.
  - 7) Alexander, L., Jiang, S., Murga, M. and González, M. C.: Origindestination trips by purpose and time of day inferred from mobile phone data, *Transportation Research Part C: Emerging Technologies*, Vol. 58, pp. 240–250, 2015.
  - 8) 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, forthcoming, 2019.
  - 9) Lee, D.D., and Seung, H.S.: Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol. 401, pp. 788–791, 1999.
  - 10) 山本真嗣: 携帯電話ユーザーの位置情報を活用した旅行者動態の把握: 北陸新幹線開業効果の検討, 日本観光研究学会全国大会学術論文集, Vol.53, 77-84, 2016.
  - 11) 清家剛, 三牧浩也: 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, 2010.
  - 12) Lee, D.D., and Seung, H.S.: Algorithms for nonnegative matrix factorization, *in Adv. NIPS*, pp.135-138, 2000.

(2019. 3. 10 受付)

---

## ANALYSIS OF TRAVEL PATTERN CHANGE BY HOKURIKU HIGH SPEED RAILWAY BY NON-NEGATIVE TENSOR FACTORIZATION

Hiromichi YAMAGUCHI, Hiroyuki MIYAKE and Shoichiro NAKAYAMA

Most conventional long distance traveling models are based on three classifications of travel purpose (sensitivity) by net flow survey, but in order to make future demand prediction more accurate, season, day of the week, individual attributes (gender and age). Is handled and a model that reflects finer sensitivity difference is required. Therefore, in this research, we propose a method called nonnegative tensor factorization using mobile space statistical data, which is one of mobile position information data that continuously acquires movement information of a large number of people with high accuracy Apply the extended model and ex-tract groups with different sensitivities to Shinkansen. In addition, it applies to the data of Ishikawa prefecture where the Hoku-riku Shinkansen opened, estimates the sensitivity of each group, applying the result to the Fukui prefecture and examining the re-sult of the demand forecast.