

Kanazawa University,
Faculty of Economics and Management

Discussion Paper Series

No. 053

スパース非負値行列因子分解を用いた
COVID-19流行期の県間旅行行動の変容分析

金沢大学大学院 人間社会環境研究科 原田 魁成
金沢大学 地球社会基盤学系 山口 裕通
金沢大学 経済学経営学系 寒河江 雅彦

Sagae.masahiko@gmail.com

December 2020



金沢大学経済学経営学系
〒920-1192 金沢市角間町

Faculty of Economics and Management,
Kanazawa University

Kakumamachi, Kanazawa-shi, Ishikawa, 920-1192, Japan

<http://econ.w3.knazawa-u.ac.jp/DP/>

スパース非負値行列因子分解を用いた COVID-19 流行期の県間旅行行動の変容分析

原田 魁成¹・山口 裕通²・寒河江 雅彦³

¹非会員 金沢大学 人間社会環境研究科 人間社会環境学専攻 (〒 920-1192 金沢市 角間町)

E-mail: h.01kaisei25@gmail.com

²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

³非会員 金沢大学 教授 人間社会学域 経済学経営学系 (〒 920-1192 金沢市 角間町)

E-mail: sagae.masahiko@gmail.com

COVID-19 下で、人々の都道府県間旅行はどのように変化したのだろうか？この疑問に答えるには、人口分布の推定精度が高く、かつ広範囲・長期の時間変動が入手できる携帯電話位置情報データが有効である。本稿では、モバイル空間統計による都道府県単位の居住地-旅行先表データを用いて上述の疑問に答えていく。このとき、膨大な次元のデータに対して、スパース非負値行列因子分解を適用することで、時間変化の特徴を簡潔に分解・記述するアプローチをとっている。その結果、首都圏での通勤の抑制効果、週末の外出控え、長期休暇時の帰省の抑制および、居住地の都道府県内滞在の時系列変化を定量的に明らかにすることに成功した。さらに、残差情報から、COVID-19 による行動変容には、通常時の時間変動から説明できない、空間的な偏在があったことを明らかにした。

Key Words: *Sparse Non-negative Matrix Factorization, change of behavior pattern, mobile phone location data, COVID-19*

1. はじめに

新型コロナウイルス感染症 (COVID-19) への対応として、我が国の政府及び都道府県は、緊急事態宣言などに連動させて行動自粛要請を実施した。その影響により、2020 年の 2 月以降に、人々の行動が大きく変化していることは報道等を通じて明らかである。さらに、長距離旅行は「都道府県を跨ぐ移動の自粛要請」の情報が発信されたために、より大きく旅行行動の変化が起こっていると想定される。

それでは、A) 緊急事態宣言とその前後の間に、どれほどの量の行動変化があったのであろうか？、B) その変化は、行動の内容 (旅行目的) ごとに、どれほど異なるものだったのだろうか？、C) この時の行動の変化は、「通常時の都道府県を跨ぐ行動の変化パターン」で単純に説明できるものだろうか？これらの疑問に回答することは、国や自治体からの大規模な行動自粛要請による人々の行動変化の記録として重要であるとともに、将来の施策決定のための重要なエビデンスになろう。

長距離旅行行動を対象として上述の疑問に答えるためには、携帯電話位置情報データなどの大規模な位置情報記録データが有用である。長距離旅行行動はほとんどの人にとって低頻度の行動であり、大規模なサンプル数がなければその増減を精度よく把握することは困

難であるため、膨大なサンプル数のデータを持つ携帯電話位置情報が最も適しているといえる。一方で、位置情報の記録データでしかないために、旅行目的などの人の意思の情報が分からないという欠点がある。旅行目的などは人の意思決定に大きな影響を与える情報であり、COVID-19 状況下における人々の行動を記録・記述するにあたって非常に重要な情報であるが、携帯電話位置情報データの情報をそのまま集計するだけでは、推測することは困難である。その結果として、上述の B) と C) の疑問に回答することができない。

このような欠点に対して、位置情報記録データを用いて、人の意思の情報である「旅行目的」を推測する手法はいくつか提案されている^{1),2)}。その中でも、Yamaguchi and Nakayama (2020)³⁾ では、1461 日分の時系列情報を持つ集計データに対して非負値行列 (テンソル) 因子分解を適用することで、新幹線開業に対する感度の異なる旅行グループに分解する手法を提案している。その結果として、ほぼ既存のモデルで扱われてきた「旅行目的」と整合が取れ、対応する旅行グループに分解されることを確認している。

そこで、本研究では、COVID-19 による影響下で、「都道府県間旅行がどのように変化したのか？」に定量的に答えることを目的として、Yamaguchi and Nakayama (2020)³⁾ と同様のアプローチで COVID-19 下における

携帯電話位置情報データを解析していく。具体的には、携帯電話位置情報データの一つであるモバイル空間統計^①によって入手できる、都道府県ごとの居住地・滞在先ペア毎の推計人口分布データから、各行を2284日(2014.3.1-2020.5.31)の時点・各列を47都道府県毎の居住地滞在人数、他県からの来訪者、他県への旅行者とする 47×3 種類の滞在空間情報とした行列 \mathbf{Y} を用いていく。そして、この情報に含まれるCOVID-19による影響を、非負値行列因子分解によって行動内容ごとに分解しつつ、明らかにしていく。

非負値行列因子分解とは、非負値のデータを加法的な構成成分に分解することを目的とした多変量解析手法であり、Lee and Seung (1999)⁴⁾はデータ行列 \mathbf{Y} を非負の2つの行列、基底行列 \mathbf{H} と表現行列 \mathbf{U} の積で近似する手法を提案した。(以降、本論文のデータの構造に合わせて、 \mathbf{H} を時系列パターン、 \mathbf{U} を空間パターンと呼ぶ。)彼らは、具体例として顔画像データに対する同モデルの適用例が挙げられており、元データが顔のパーツで構成された基底行列と0~1の範囲で定義された表現行列に分解できることを示している。そして、交通や社会基盤関連データに対しても、Yao et al. (2015)⁶⁾、細江ら (2018)⁵⁾などで応用されている。

非負値行列因子分解に罰則項を付与することで明瞭な特徴を導出する手法も提案されている。Hoyer(2002)⁷⁾は分解される行列のうち、空間パターンに L_1 ノルムを付与したモデルを提案しており、Kim and Park(2007)⁸⁾は分解される行列のうち、時系列パターンと空間パターン両方に L_2 ノルムを付与したモデルを提案している。これらのモデルはCichocki et al. (2007)⁹⁾によって総括され、体系化されている。罰則項は、時系列パターン及び空間パターンの片側または両側に L_p ノルムを付与する様々なバリエーションが組み合わせ的に考えることができる。本稿では、より解釈が容易な分解を実現するために、日数が多く情報量の多い時系列パターンに L_1 ノルムを付与したモデルを構築し、使用データに適用する。

2. データと基礎集計

(1) 多時点の居住地-旅行先表データ

本稿では、携帯電話運用情報を用いた人口分布の統計である、モバイル空間統計¹⁰⁾による居住地-旅行先表データを用いる。このデータは、NTTドコモが提供している8千万台もの携帯電話の運用データから作成されたものであり、最頻で1時間ごとの日本全体の人口分布を把握することが可能な集計データである(ただし本稿では、最も移動先に滞在している可能性が高い

と推測される、昼間の13時台のみを用いる)。その運用データに含まれる旅行先情報と居住地情報を用いて作成される、以下のような2284日分の都道府県間居住地-旅行先表を対象に分析を行う：

$$\mathbf{Q} = \begin{pmatrix} q_{1,1,d_1} & \cdots & q_{i,j,d_1} & \cdots \\ \vdots & & \vdots & \\ q_{1,1,d_t} & \cdots & q_{i,j,d_t} & \cdots \\ \vdots & & \vdots & \end{pmatrix} \quad (1)$$

$\forall(d_t \in D).$

ここで $D = [2014.3.1, \dots, 2020.5.31]$ 、 $i, j \in Z$ はそれぞれ居住地、旅行先に対応し、 Z は47都道府県の集合である。つまり、 q_{i,j,d_t} の値は、 d_t 日の居住地の都道府県 i の人の中で都道府県 j に滞在している人数を、携帯電話運用データから推計したものである。ただし、 $i \neq j$ のとき、特定の居住地からある都道府県へ流出する人数、あるいは流入する人数は、いずれも居住地に滞在する人数より小さいため、そのままでは次元縮約をして行列データから特徴を抽出するスパース非負値行列因子分解に適用するとうまくこれらの特徴が抽出されない可能性がある。そこで地域間の旅行が行われた属性に対し、居住地 i を固定し旅行先 j に関して総和をとる属性と旅行先 j を固定し居住地 i に関して総和をとる属性に2分し、それぞれ「 $q_{i,\bullet}$ 」、「 $q_{\bullet,i}$ 」と表すことにする。これらの処理によって使用データは以下のように表せる。

$$\mathbf{Y} = \begin{pmatrix} \cdots & q_{i,i,d_1} & \cdots & q_{i,\bullet,d_1} & \cdots & q_{\bullet,i,d_1} & \cdots \\ \vdots & \vdots & & \vdots & & \vdots & \\ \cdots & q_{i,i,d_t} & \cdots & q_{i,\bullet,d_t} & \cdots & q_{\bullet,i,d_t} & \cdots \\ \vdots & \vdots & & \vdots & & \vdots & \end{pmatrix}$$

$\forall(d_t \in D).$

(2)

ただし $q_{i,\bullet} = \sum_{j,i \neq j} q_{i,j}$ であり、 $q_{\bullet,i} = \sum_{j,i \neq j} q_{j,i}$ である。さらに $[q_{i,i}, q_{i,\bullet}, q_{\bullet,i}]$ は、自地域滞在、他地域への流出、他地域からの流入に対応し、いずれも各都道府県の[47,47,47]パターン存在する。よってこれらの処理を施した2284時点(日) \times 141空間種類を本稿の使用データ行列 \mathbf{Y} とした。

(2) 居住県外滞在者の時間推移

上述のデータから、 $\sum_{i=1}^{47} q_{i,\bullet,d_t}$ で、時点 d_t において居住県外に滞在している人数の総数を算出することができる。まず、この数値の時間変動から、「緊急事態宣言とその前後の間に、どれほどの行動変化があったのであろうか？」という問いに答えていこう。

¹ 「モバイル空間統計」は株式会社NTTドコモの登録商標です。

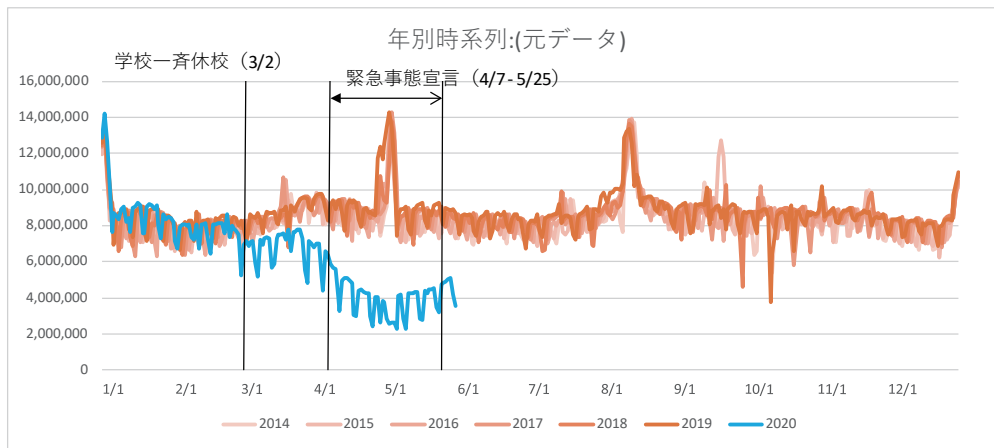


図-1 居住県外滞在者数の時間推移

図-1は、モバイル空間統計において推定された、2014年3月1日から2020年5月31日までの「居住県外滞在人数」の推移を示している。まず、2019年までのCOVID-19の影響を受けていない期間の時間推移を見てみると、おおむね6年ともに共通の時間変動であることが確認できる。具体的には、約800万人程度の人が13時台に県外滞在做しており、年間に3回のシーズンにおいてその量が大きくなる。それは、年末年始、5月初旬（ゴールデンウィーク）と8月中旬（お盆）であり、この期間には、通常時の1.5倍以上、最大で140万人も人が居住している都道府県の外に滞在している。なお、9月末にある年にだけ確認できる、居住県外滞在人数の増加は、2015年のシルバーウィークである。この年には祝日の配置によってカレンダー上で5連休（2015年9月19日から9月23日）になっており、この年に限定して5月のゴールデンウィークに匹敵するほどの、大きな県外滞在人数の増加が観測されている。

次に、COVID-19の影響を受けた2020年の推移（青線）を見ていこう。1月と2月は、ほとんど2019年以前と同じ範囲を変動しているが、3月から大幅に減少していることが確認できる。この図から、5月末までのCOVID-19による居住県外での滞在量の推移は、以下の3フェーズに分けることができる：

1) 2020年の3月に入ってから、居住県外での滞在量が大幅に減少した。これは、およそ学校の一斉休校やイベント自粛の検討要請など、政府が活動の抑制を発信し始めた時期と一致する。この時の減少量は、通常時より約100万人（約12%）程度少なく、この水準が1か月程度継続している。

2) 3月末から、5月初旬（ゴールデンウィーク）にかけて、'徐々に' 県外滞在人数が減少し続けた。その結果、ゴールデンウィークにあったピーク時には、居住県外の滞在人数は約200万人まで減少した。これは、通

常時の約800万人の1/4であり、ゴールデンウィーク期間において期待される1,400万人のたった1/7程度に過ぎない。なお、この期間において政府による緊急事態宣言が発令されている（7都府県：4/7、全国：4/16）が、その時点において急激に県外滞在者数が減少するといった傾向は図-1では確認できない。また、この居住県外滞在人数の低下は、7都府県で先行して緊急事態が宣言された日より2週間も早く始まっていた。以上のことから、居住県外への旅行行動の減少は、政府の緊急事態宣言などの要請のタイミングに急激に進んだわけではなく、1日当たりの感染者数の増加や「自粛要請」の内容が強化されるに従って、'徐々に' 進んだことがわかる。

3) 5月のゴールデンウィークから、5月末にかけて若干の回復傾向が確認できる。短期間かつ緩やかな変動であるため、より明確なトレンドを把握するためにはより長期のデータを見る必要があるが、この回復が開始した時点は、緊急事態宣言の解除（39都府県：5/14、関西圏：5/21、首都圏・北海道：5/25）があった後である。なお、この時点では「県外への移動自粛」は継続した状態であり、全国的に県外への移動自粛要請が解除されたのは6月19日であり、その影響を見ていくには6月以降のより長期のデータを確認する必要がある。

3. 居住地-旅行先表データのスパース非負値行列因子分解

本節では、都道府県単位の時系列データを非負値行列因子分解することで、COVID-19による都道府県を跨ぐ滞在行动の変化を、以下の2点に着目して明らかにする。一つ目は、「行動の内容」ごとに分解することで、COVID-19対応における変化は、旅行目的など行動内容ごとに明らかにする。その上で、COVID-19対

表-1 ハイパーパラメータ μ の変化と時系列・空間パターンのスパース度の変化

μ	10^2	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}
H スパース度	1.2%	6.4%	12.0%	15.4%	25.2%	27.1%	29.4%	28.3%	27.8%
U スパース度	9.9%	19.6%	12.6%	27.9%	13.2%	12.1%	14.9%	13.2%	12.6%
RSS(10^{13})	2.5	3.5	3.7	5.8	8.2	8.6	9.9	8.3	8.6

※スパース度は H,U それぞれの全要素数に占める 0 の割合と定義している。

応における変化には、通常時の時間変動データから導出される「移動の行動内容」の全国的な増減のみで説明しきれない、地域ごとの変化の偏在性を確認していく。

本稿の非負値行列因子分解では、より「移動の行動内容」を解釈しやすい分解結果を得られることが期待される、時系列パターンに L_1 ノルム罰則項を付与したモデルを使用する。ここで、2284 日分時系列データ $\times 141$ (47 都道府県 $\times 3$ パターン) の行列 $\mathbf{Y} = [\mathbf{x}_1, \dots, \mathbf{x}_{141}] \in \mathbb{R}_+^{2284 \times 141}$ を、2 つの非負値行列 $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_k] \in \mathbb{R}_+^{2284 \times k}$, $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{141}] \in \mathbb{R}_+^{k \times 141}$ に分解するとき、スパース非負値行列因子分解は、残差行列 $\mathbf{Y} - \mathbf{H}\mathbf{U}$ のフロベニウスノルムと時系列パターン行列 \mathbf{H} の L_1 ノルム罰則項の和である

$$L = \|\mathbf{Y} - \mathbf{H}\mathbf{U}\|_F^2 + \mu \|\mathbf{H}\|_1 \quad (3)$$

を最小化する問題となる。ただし μ はハイパーパラメータを表す。

このとき、上記の式を最小化する行列 \mathbf{H}, \mathbf{U} は、任意の初期値を設定したうえで以下の更新式を収束するまで繰り返し適用することによって、高速に導出することができる⁹⁾。

$$\mathbf{H}_{i,j} \leftarrow \mathbf{H}_{i,j} \frac{(\mathbf{Y}\mathbf{U}^T)_{i,j}}{(\mathbf{H}\mathbf{U}\mathbf{U}^T)_{i,j} + \mu(\mathbf{1}_{2284}\mathbf{1}_k^T)_{i,j}} \quad (4)$$

$$\mathbf{U}_{i,j} \leftarrow \mathbf{U}_{i,j} \frac{(\mathbf{H}^T\mathbf{Y})_{i,j}}{(\mathbf{H}^T\mathbf{H}\mathbf{U})_{i,j}} \quad (5)$$

ただし、 $\mathbf{1}_k$ は要素が全て 1 の $k \times 1$ ベクトルである。また上記式は行列の要素ごとの演算である。

本稿の計算では、式 (4)・式 (5) の更新式を用いて行列分解を実施した。なお、計算においては各 500 回ずつ更新を行い、いずれの式も更新変化率が 10^{-3} を下回った時を停止条件とした。ハイパーパラメータ μ の設定として、本研究では $10^1 \sim 10^{10}$ の範囲で探索的に行い実施した。表-1 は、ハイパーパラメータごとの、 \mathbf{H} , \mathbf{U} の要素のスパース率 (要素の中でゼロである割合) を示したものである。基本的には、ハイパーパラメータ μ の値が大きくなればなるほど、罰則項が強く作用するために、 \mathbf{H} のスパース率は大きくなる傾向にあり、スパース率が大きいほど分解された移動の時間変動成分が、特定の時期に偏るため、より解釈が容易となる。一方で、罰則項の影響を大きくすると、残差 RSS も大き

くなるため、ある程度のスパース化の効果が確認できる中で最小の μ の値を設定することが望ましいと考える。そこで、 \mathbf{H} のスパース率が大きく増加する値である最小値 $\mu = 10^6$ を設定した。

また、事前に決定するパラメータとして基底数 k も外生的に決める必要がある。その決定方法として情報量基準の一種である AIC や MDL を利用した決定方法があるが、本稿では結果の解釈しやすさを優先して、一般的に利用される都市間旅行データにおける旅行目的 (業務、観光、私用・帰省) を用いて解釈できる基底数 $k = 5$ を選択した。

この基底数に限定することによって、代表的な旅行目的ごとの全国的な移動量変化を、時系列パターン行列 \mathbf{H} から捉えることができる。そして、「この分類による全国的な移動量変化」として説明できないような変化を、残差行列 $\mathbf{Y} - \mathbf{H}\mathbf{U}$ を用いて捉えることができる。

4. スパース非負値行列因子分解からわかる行動変容パターン

(1) 分解した行動内容ごとの COVID-19 の影響

本節では、時系列パターン \mathbf{H} と空間パターン \mathbf{U} の 2 つの非負値行列を、スパース非負値行列因子分解によって算出した結果を確認していこう。これらは、スパース非負値行列因子分解によって抽出された、行動内容ごとの時系列変動情報である。

本節では、時系列パターンと空間パターンの 2 つの非負値行列を、スパース非負値行列因子分解によって算出した結果を示す。図-2 から図-6 は、それぞれ時系列パターン \mathbf{H} と空間パターン \mathbf{U} ごとの成分の値を、基底ごとに示したものである。

なお、図-2 から図-6 には、それぞれ 3 種類の情報を掲載している。(a) は時点 d_t において、基底成分 k によって表される居住者数及び旅行者数 ($h_{d_t,k}$) を示している。つまり、行列 \mathbf{H} の 1 列分の情報を、年ごとに分けて折れ線グラフで示したものである。つぎに、(b) も (a) と同じで、時系列パターン行列 \mathbf{H} の情報を示したものであり、全期間を通じた曜日ごとの平均値を示している。この図から、(a) の中で把握しにくい 7 日ご

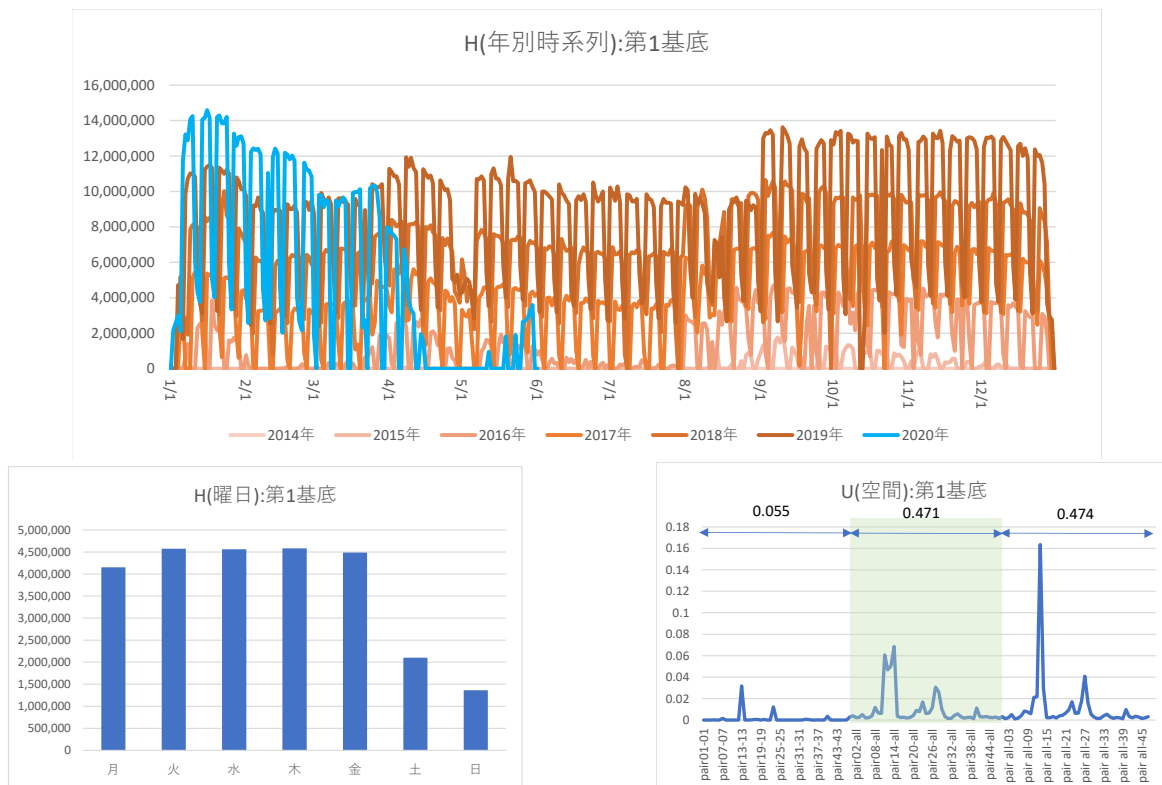


図-2 第1基底 (上 (a):年別時系列パターン, 左下 (b):曜日別時系列パターン, 右下 (c):空間パターン)

との時間変動を把握することができる。他方、(c)は基底成分 k における空間パターン i の割合を示している。横軸の空間パターンは、左から、居住地内の滞在人数 (47 都道府県)、特定の居住都道府県からの流出人数 (47 都道府県)、特定の都道府県への他県からの流入人数 (47 都道府県) の順に並べている。ここで都道府県はコード順に並べてある。また、式 (2) における $q_{i,\bullet}$ と $q_{\bullet,i}$ はそれぞれ図中において便宜上「 i -all」, 「all- i 」と表している。

a) 業務・通勤目的変動の時系列推移

図-2の第1基底と図-3の第2基底は、それぞれ類似性が高い成分であるため、まとめて結果を確認していこう。まず、COVID-19の影響を受けていない期間の各基底の時間変動を図-2(a)と図-3(a)から見ていこう。図-2(a)では2014年から2019年にかけて経年的に増加している。一方で、図-3(b)では2014年から2019年にかけて経年的に減少し続けていることが確認できる。そして、この時間変動のうちで7日間周期の変動のみに着目した図-2(b)と図-3(b)をみると、どちらも共通して平日に突出したパターンであることがわかる。これらの基底の空間パターンを示した図-2(c)と図-3(c)からも、かなり類似性が高いことが確認される。具体的には、流入量(最も右の領域)では少数の都道府県(大きいピークから順に、東京都・大阪府・愛知県)に集中する一方で、流出量(緑色の領域)は日本全

国に微量の値を取り、とくに大都市圏周辺で大きい傾向にある。これらは図-2(c)と図-3(c)のどちらでも共通の傾向である。唯一の差異は、第1基底(図-2(c))では、居住都道府県内の滞在量(最も左の領域)において東京と愛知県の居住都県内滞在が比較的大きい値をとっている点である。

以上の時間変動の特徴を踏まえると、平日に3大都市圏へ旅行する行動であることから、この第1・2基底は業務・通勤行動が多くを占める成分であることが推測される。そして、この平日の代表的な行動では、その量の構成が2014年から2019年にかけて徐々に第2基底中心から第1基底中心に移ってきており、東京と愛知県の居住都県内滞在の量が増加する変化があったことを示している。

つぎに、これらの業務・通勤目的のCOVID-19による行動変化を2020年の情報から見ていこう。まず、図-3(a)に示される第2基底では、2020年は継続してほぼゼロである。この基底の時間変動では、COVID-19の影響がない2019年末からほぼゼロで推移しており、COVID-19に関係なく経年的に減少し続けた結果、近年はほとんど見られなくなった成分であることが確認できる。

図-2(a)における、COVID-19の期間における推移を見ていこう。2020年3月以降の推移を見ていくと、1) まず、2020年3月上旬(学校の休校が発表された

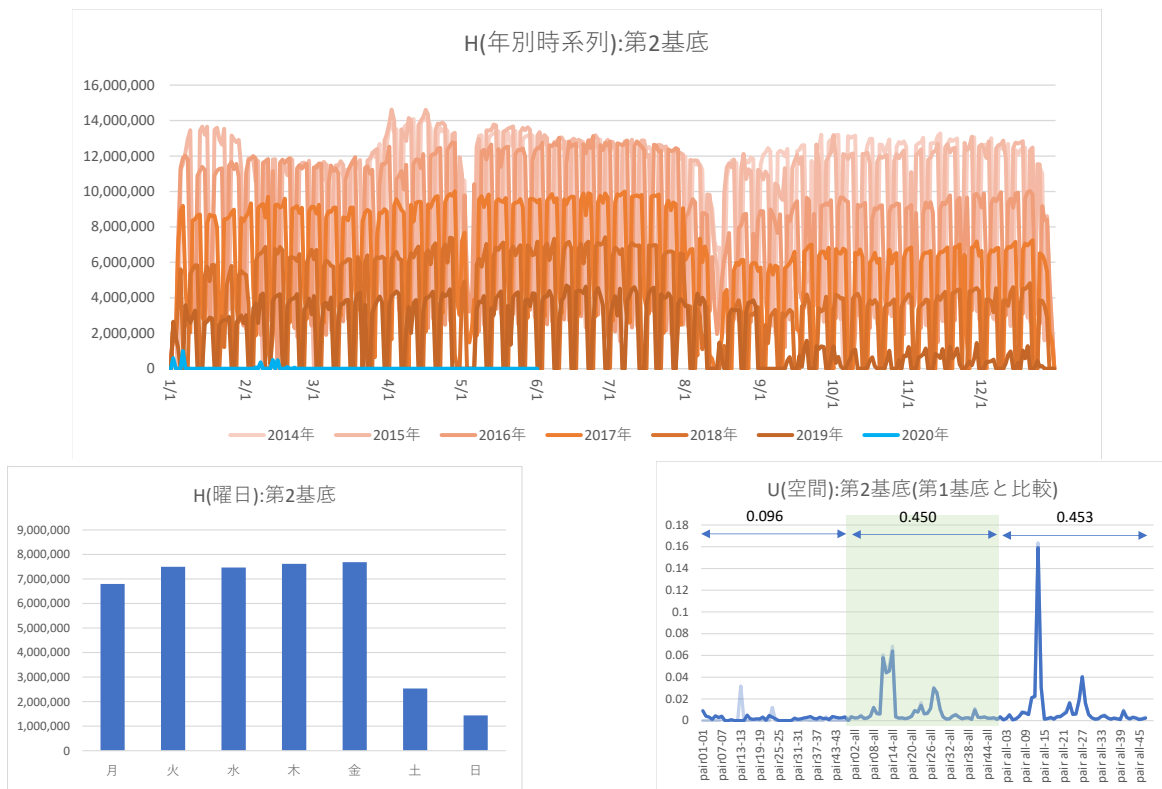


図-3 第2基底 (上 (a):年別時系列パターン, 左下 (b):曜日別時系列パターン, 右下 (c):空間パターン (第1基底比較))

タイミング)に1,200万人から1,000万人²に減少し、1ヵ月程度その人数が継続している。このような変化は、2019年以前には見られず、COVID-19による影響であろうと推測できる。2) 7都府県で先行して緊急事態宣言が発令される、およそ2週間前からさらに急激に減少を始める。そして、全国が緊急事態宣言下にあった、4/16から5/7までの期間は、この通常時の平日にある大都市圏へ集中するパターンの旅行はほぼゼロで推移していた。3) 緊急事態宣言が段階的に解除されるにしたがって、徐々に回復するものの、5月末の時点で通常時に期待される量(2019年の後半の水準が約1,300万人)の1/4にも満たない量である。

以上から、通常時の平日に強く確認された、(業務・通勤目的と思われる)大都市へ集中する旅行行動は、緊急事態宣言中はほぼゼロになるなど、大幅に減少していたことが確認できる。

b) 居住都道府県における滞在量変動の時系列推移

図-4から、第3基底の情報を見ていこう。まず、図-4(c)から、この成分のうちの95%程度が居住地県内での滞在行動であり、その都道府県内での大小関係はおおよそ人口規模に比例する形である。つまり、「居住都道府県での滞在行動」を示す成分であるといえる。

² なお、この数値は単位が人であるが、居住県外に滞在している人は「流入人数」と「流出人数」でダブルカウントされるために、実際の行動人数の2倍近い数値となる点は注意が必要である。

この通常時(2019年まで)の時間変動を、図-4(a)から見ていくと、普段は安定した居住者の推移を示すが、特にGWや盆休み、正月休み期にこの滞在者が大きく減少する、つまり県外への旅行行動をとる人が多くなることが確認できる。また、図-4(b)の曜日ごとの変動をみると、金曜日・土曜日・日曜日で平均値が小さく、これらの週末に比較的県外への旅行者数が多いことが確認できる。

COVID-19による行動変化を見ていくと、おおよそ図-1で確認できた県外旅行者と反対の特徴を示すことが分かる。つまり、学校の一斉休校が開始されたタイミングでの増加した後、さらに3月下旬からゴールデンウィークにかけて徐々に増加し続ける。そして、緊急事態宣言の解除後に若干の回復傾向が確認できるという推移である。これは、COVID-19の状況下において、居住県外への移動を自粛して都道府県内での滞在量が増加していることを示すものである。

c) 長期休暇時の大都市から地方への行動の時系列推移

図-5から、第4基底の情報を確認していく。まず、COVID-19の影響を受けていない期間の各基底の時間変動を図-5(a)から見ていこう。図-5(a)と(b)から、この基底は通常時の休日にも確認できるものの、GWや夏休み(お盆の期間がピーク)・年末年始の3期間に集中して発生している行動であることが分かる。つぎに、図-5(c)からこの行動パターンの空間的な特徴を

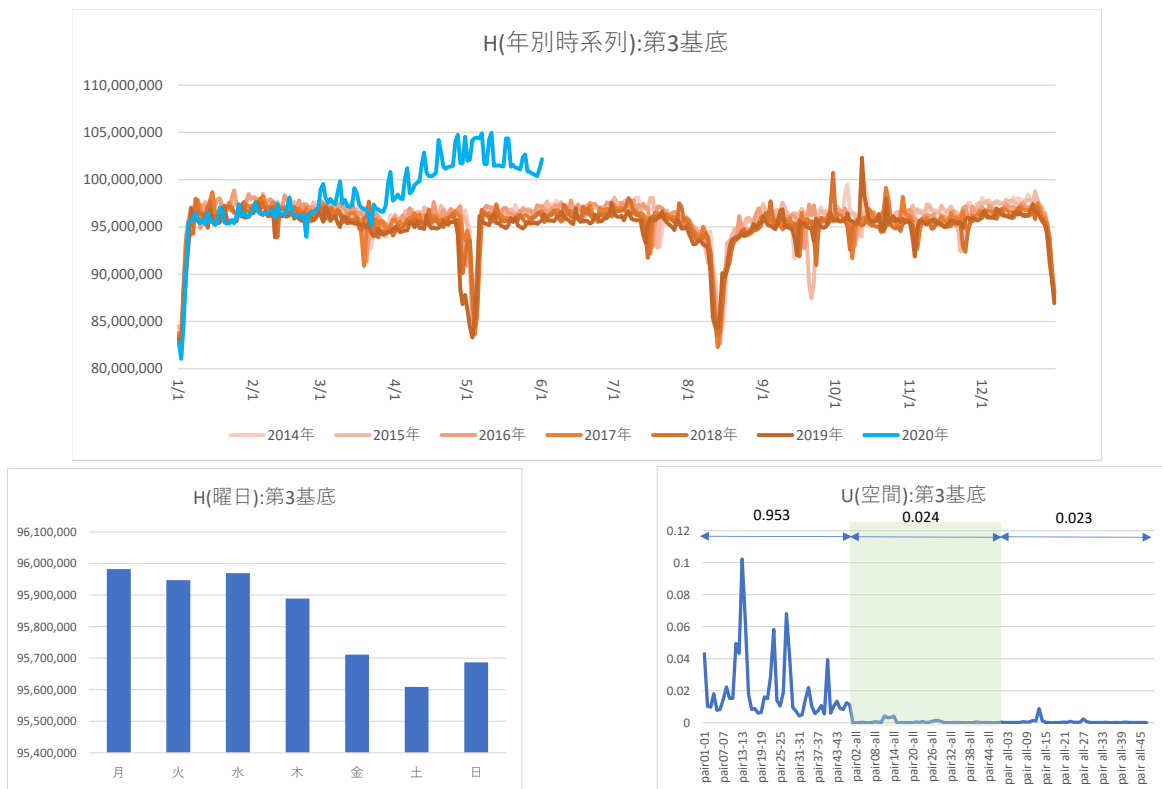


図-4 第3基底 (上 (a):年別時系列パターン, 左下 (b):曜日別時系列パターン, 右下 (c):空間パターン)

見ていこう。まず流入量（最も右の領域）は、地方部も含めて（東京都・大阪府を除いて）日本全国でほぼ満遍なくあることが確認できる。一方で、流出量（真ん中の緑の領域）は、東京都・大阪府・愛知県とその隣県といった大都市部がほとんどを占める。つまり、大都市から地方部への旅行が多くを占める行動であることがわかる。以上の内容から、第4基底は「長期休暇時における大都市から地方への行動」に対応し、多くを帰省行動が占めるものと推測できる。なお、図-5(c)では、この成分には北海道居住者が北海道に滞在する行動も多くあることが確認できる。これは、北海道居住者が、他の県の人と比較して長期休暇の時期に道内に滞在する割合が大きいことを示している。

COVID-19による行動変化を見ていくと、2020年では数回の微量の移動が確認できるものの、ほぼ一貫してこの成分がゼロであることが確認できる。とくに、2014年から2019年までの傾向を踏まえると膨大な量の移動が予定されていたと思われるゴールデンウィークにおいても完全にゼロであり、極めて異質な年であったことが確認できる。

d) 週末の外出行動の時系列推移

図-6から、第5基底の情報を確認していく。まず、COVID-19の影響を受けていない期間の各基底の時間変動を見ていこう。図-6(a)から、この基底は2019年までの間にはそれほど大きな季節変動がないことが分

かる。さらに、図-6(b)から、とくに曜日ごとの変動が大きく、平日と比較して土曜日と日曜日に特に量が多い行動であることが分かる。つぎに、図-6(c)からこの行動パターンの空間的な特徴を見ていこう。まず、流出量（真ん中の緑の領域）をみると、地方部においても一定の値をとるものの、東京都・大阪府・愛知県といった大都市部で大きいことが分かる。つぎに、流入量（最も右の領域）をみると、こちらも地方部全体で一定の量を占めるが、関東・関西地方の各県において比較的大きな値をとることが分かる。これは、特に東京からの流出量が多いことを踏まえると、東京から神奈川や千葉、埼玉など比較的近県への旅行行動が多いことを示唆していると考えられる。以上の特徴を踏まえると、この行動は「週末の外出行動」であり、大都市から近隣県への旅行が多くを占めるものと推測できる。

この行動のCOVID-19による量の変化を見ていくと、以下のような時系列推移が確認できる：1) この成分では、3月上旬の学校の一斉休校などによる行動量の減少は、土日ではあまり確認できない。一方で、通常時ではほぼゼロである、平日の行動量が徐々に増加する傾向にある。この成分の空間パターンの特徴を踏まえると、大都市居住者が平日に郊外に滞在する量が増えたことを意味する。これは、COVID-19状況下においてのみ見られる特殊な変動である。仮説にすぎないが、普段は東京都や大阪に住んでいた人が、テレワーク等が推

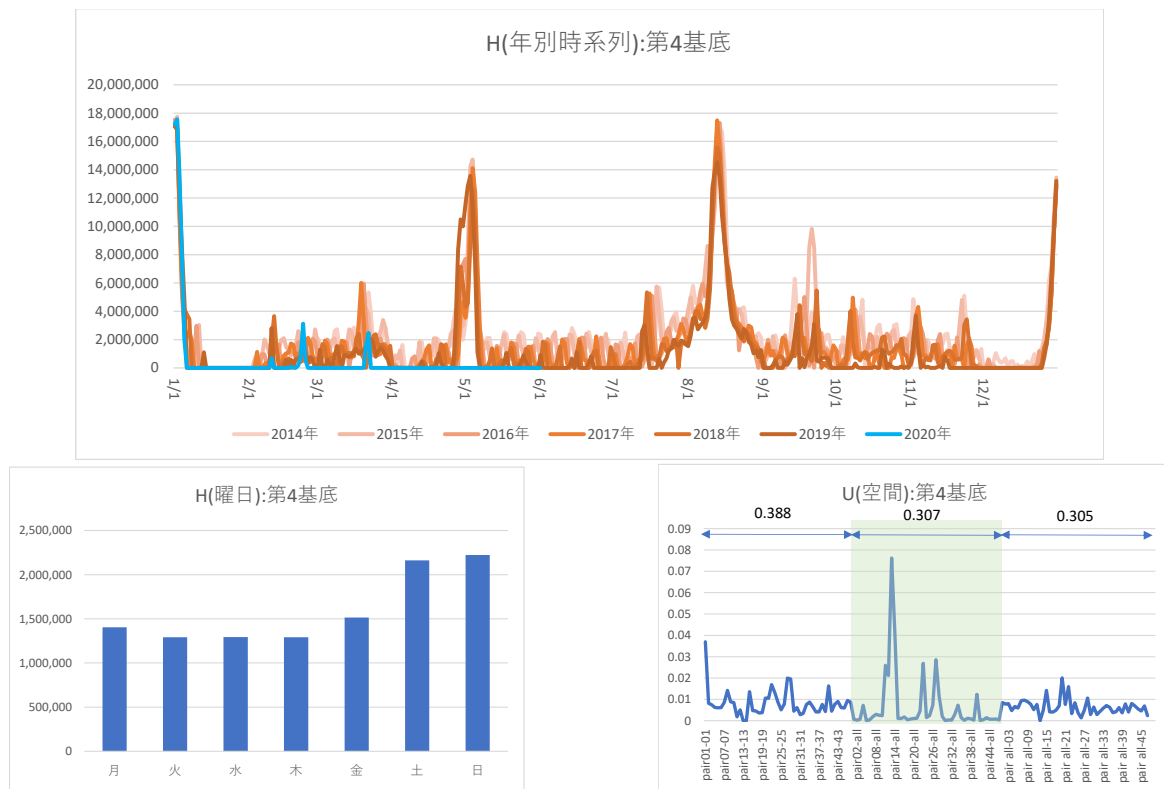


図-5 第4基底 (上(a):年別時系列パターン, 左下(b):曜日別時系列パターン, 右下(c):空間パターン)

進されるにしたがって平日の滞在場所を郊外の近隣県に変更した可能性がある。

2) 4月上旬の緊急事態が宣言される前後になると、土日のこの成分の行動量が大幅に減少し、ゴールデンウィークの期間にはほぼゼロとなる。一方で、平日の滞在量は増加したままで推移しており、大都市居住者の郊外部滞在量は、土日と平日で反転するような現象が起こっている。

3) 5月後半の緊急事態宣言解除後には、(日数が少ないために明確なトレンドを確認することは難しいが、)土日の行動量について回復の兆しが見られる。また、平日の滞在量は3月から4月にかけて増加したままの水準で維持している。

(2) 残差の時空間分布

次に、 $Y - HU$ で算出される残差行列の時空間分布を見ていこう。ここでの残差は、上述したスパース非負値行列因子分解の5つの基底では十分に特徴が抽出されなかった要素であり、これを見ることで元データに対する近似の損失分を補った分析が出来る。つまり、COVID-19等による滞在行动の変化が、上記で述べた5つの行動分類(基底)ごとに日本全体で一律に起こった場合は、それぞれの時間変動成分の大小で説明可能であるため、残差は大きく変わることはない。一方で、5つの行動内容(業務・通勤行動、県内滞在、長期休暇

の大都市から地方への行動、週末の外出行動)の空間分布と大きく乖離する変化が起こっていた場合には、非負値行列因子分解で導出した基底ではその変化を説明することができず、その時点の残差の値が大きくなる。以降では、このような残差の特徴を踏まえて、その大きさを確認するとともに、「COVID-19下ではどのような特殊な行動変化があったのか？」をその空間分布から明らかにしていく。

図-7は、元データと罰則付き非負値行列因子分解における残差の2乗和を、日付ごとに算出したものである。図-7より、2020年ではCOVID流行期と重なる2020.3.29(日)(外出自粛要請後初めての週末)以降において残差が非常に大きく、例年とは極めて異質な行動パターンであったことがわかる。COVID-19の影響がない時点のなかで、次に大きい残差が確認されたのは、2019.10.12(土)と2018.9.4(火)である。これらは、記録的暴風が日本に接近した日であり、前者は台風19号が静岡県や関東甲信越・東北地方に接近し、人的被害の他、土砂災害、家屋の崩壊、浸水、断水、停電等の被害をもたらした。そして、交通面を見ると関東地方を中心に大規模な計画運休が実施された日であり、北陸新幹線がこの日より被災により長期間の運休を余儀なくされたタイミングである。後者は、台風21号が近畿地方に接近し、大規模停電や関西国際空港の浸水に伴う閉鎖等の被害をもたらした時期である。残差と

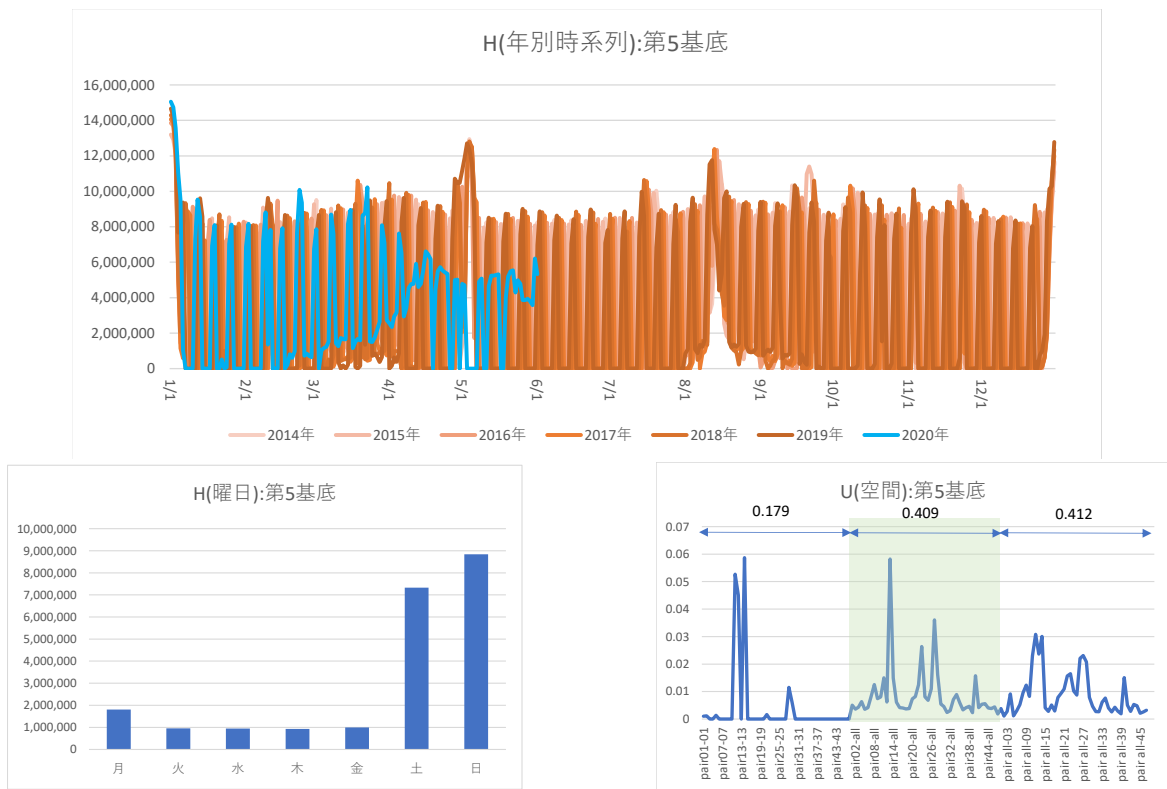


図-6 第5基底 (上 (a):年別時系列パターン, 左下 (b):曜日別時系列パターン, 右下 (c):空間パターン)

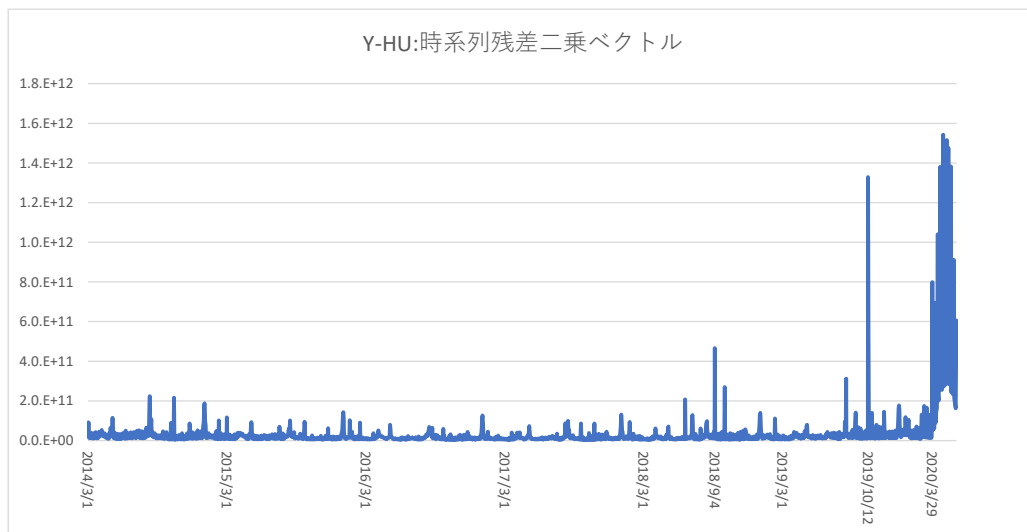


図-7 元データと罰則付き非負値行列因子分解における時系列残差ベクトル 2284 日分

して特徴が抽出されたことは、これらの時期には、平常時の行動パターンとは空間的に大きく異なる行動パターンで行動変化が起きていたことを意味している。

これらの時期における、残差の空間分布を用いて、「どのような特殊な行動パターンがあったのか？」に答えていこう。図-8と図-9は、それぞれ2020.3.1(日) - 2020.5.31(日)のCOVID-19の影響が大きいと予想される大きな残差が確認された期間と、2019.10.12(土)の2019年台風19号による大きな残差が確認された期

間の、平均的な残差の空間パターンの内訳を表している。図-8のCOVID-19時期の残差をみると、東京をはじめとする関東地域の居住都県内滞在量が大きく正の残差を示しており、そのほかの地域で負の値をとっている。つまり、COVID-19による居住都道府県内の滞在量増加は、平常時の行動内容パターンでは説明できないほど、関東地域の各県で特に大きい(行動量の変化が偏在していた)ことを示している。

一方で、2019年の台風19号が接近していた時の残差

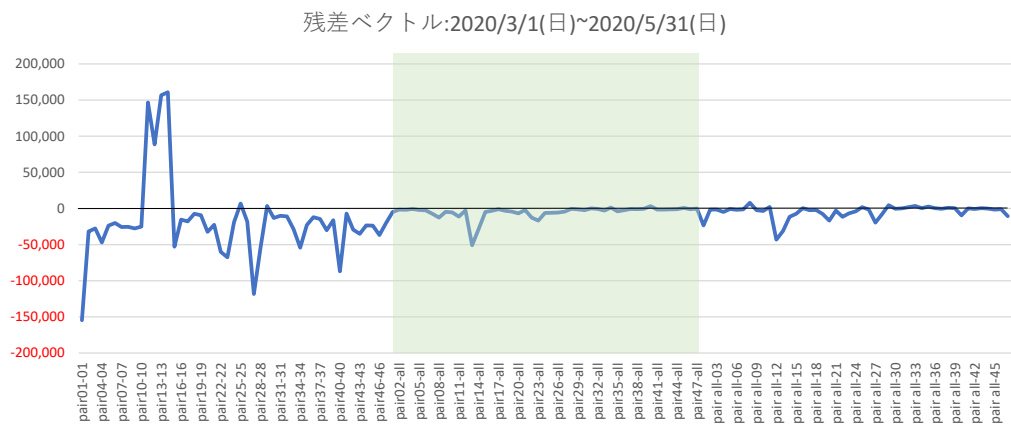


図-8 COVID 流行時の時系列残差ベクトル:2020.3.1(日)-2020.5.31(日)

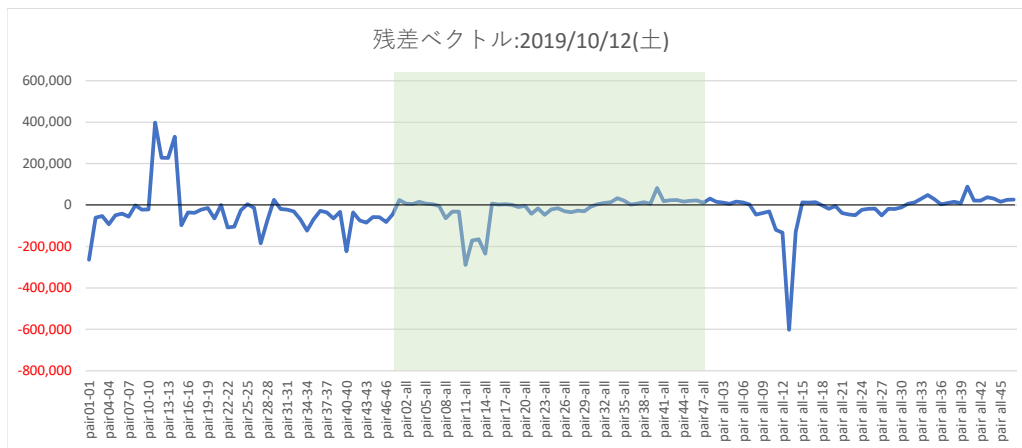


図-9 関東への記録的暴風接近時の時系列残差ベクトル:2019.10.12(土)

ベクトルを示した図-9を見ると、こちらでも東京をはじめとする関東地域の自地域内滞在パターンが正の特徴を示しており、さらに東京への流入量が顕著に負の値をとっている。この台風では、計画運休をはじめとする台風による影響は、とくに関東地方に偏在していた。そのことを反映して、関東地方を中心に残差が大きい結果であった。COVID-19の影響についても、台風19号とは数量オーダーは半分程度ではあるものの、とくに感染者数が多かった東京都を中心に、関東地域に行動変容が偏在していた傾向が確認できた。

5. おわりに

本研究では携帯電話の位置情報データ（モバイル空間統計）から入手できる長期間の居住地-旅行先データを用いて、2020年の5月末までのCOVID-19による居住県内および県間における移動行動の推移を確認した。とくに、スパース非負値行列因子分解を用いることで、定点に絞らず、データの特徴から客観的に分解を行い、全国の行動パターンの特徴量を抽出すること

で、COVID-19による都道府県間を跨ぐ行動変容をサンプルにできた点を強調したい。得られた成果は、以下のとおりである。

まず、時系列データの基礎集計を通じて、「緊急事態宣言とその前後の間にどれほどの量の行動変化があったのか？」という疑問に対して、以下のような時系列推移を明らかにした：[2020年3月] 学校の全国一斉休校が開始された3月2日前後で居住県外での滞在数が12%程度減少し、1ヵ月程度はそのまま推移していた。[2020年4月] さらに、緊急事態宣言が開始される数週間前から徐々に行動量が減少し、ピークとなるゴールデンウィークには通常時の1/4、ゴールデンウィーク期間において期待される量のたった1/7程度まで居住県外での滞在人数が減少した。[2020年5月] 緊急事態宣言が解除された5月末には若干の回復傾向がみられるものの、この時点では引き続き大幅に低下した状態であった。

つぎに、より細かい都道府県単位での傾向を日本全体で簡潔に記述するために、次元縮約の観点からスパース非負値行列因子分解を用いた。その結果、日本国内で

の都道府県間を跨ぐ旅行パターンとして、「平日の業務・通勤行動」「居住都道府県での滞在行動」「長期休暇時の大都市から地方部への移動行動」「週末の外出行動」の4パターンが抽出され、それぞれの行動のCOVID-19による影響を明らかにした。具体的には、「平日の業務・通勤行動」は、おおむね居住県外の総数と同様の推移をとり、とくに緊急事態宣言状態ではこのパターンがゼロになっていた。「長期休暇時の大都市から地方部への行動」は、2020年の年末年始以降、ほぼゼロのまま推移し続けていた。「週末の外出行動」は、3月末からゴールデンウィークにかけて減少し続けた。一方で、2020年の3月以降には平日にこのパターンが見られるようになった。これは、平日の滞在場所（勤務場所）を大都市から（これまでは週末の旅行先であった）近隣県に旅行させた行動が無視できない程度に増加したことが原因と推測されるものの、特定するためにはより詳細な調査が必要である。

さらに、残差の情報を確認すると、COVID-19では、「普段の都道府県をまたぐ行動変化パターン」の減少或いは増加で単純に説明できないような行動変化があったことが明らかになった。このような特殊な変動は、隣接地域のみ限定して行動変容が起こる、台風など同様のパターンである。その内訳をみると、とくに東京都を中心とする関東地方において、居住県内に滞在するという行動変容が集中していることが分かった。これは、移動の自粛要請などは、全国的に実施されていたものが多いものの、感染者数が多かった関東地方でとくに大きな行動変容が起こっていたことを示している。

謝辞：本研究は、科学研究費補助金 16K00043, 16H02790, 17H03321, 20H02270, 北陸銀行若手研究者助成金の支援を受けた研究活動による成果の一部です。

参考文献

- 1) Alexander, L., Jiang, S., Murga, M., González, M.C.: *Origin-destination trips by purpose and time of day inferred from mobile phone data*, Transportation Research Part C: Emerging Technologies, Vol. 58, pp.240250, 2015.
- 2) Alsger, A., Tavassoli, A., Mesbah, M., Ferreira, L., Hickman, M.: *Public transport trip purpose inference using smart card fare data*, Transportation Research Part C: Emerging Technologies, Vol. 87, pp.123137, 2018.
- 3) Yamaguchi, H., Nakayama, S.: *Detection of base travel groups with different sensitivities to new high-speed rail services: Non-negative tensor decomposition approach*, Transport policy, Vol. 97, pp.3746, 2020.
- 4) D, Lee., H, S, Seung.: *Learning the parts of objects by non-negative matrix factorization*, Nature 401, 788-791, 1999
- 5) 細野美欧, 桑野将司, 谷本圭志: 非負値テンソル因子

分解を用いた交通系 IC カードデータからの移動パターンの抽出に関する研究, 都市計画論文集 Vol.53, No.3, pp.1320-1326, 2018.

- 6) Yao, D., Yu, C., Ding, Q.: *Human mobility synthesis using matrix and tensor factorizations*, Information Fusion, Vol.23, pp.25-32, 2015.
- 7) P,O,Hoyer.: *Non-negative sparse coding*, In Neural Networks for Signal Processing XII (Proc. IEEE Workshop on Neural Networks for Signal Processing), 557565, Martigny, Switzerland, 2002
- 8) H,kim., H,park.: *Sparse non-negative matrix factorization via alternating non-negative-constrained least squares for microarray data analysis*, Bioinformatics, 23(12):1495-1502, 2007
- 9) Cichocki, A., Zdunek, R., Amari, S.: *Hierarchical ALS Algorithms for Nonnegative Matrix and 3D Tensor Factorization*, In: Davies M.E., James C.J., Abdallah S.A., Plumbley M.D. (eds) Independent Component Analysis and Signal Separation. ICA 2007. Lecture Notes in Computer Science, vol 4666. Springer, Berlin, Heidelberg, 2007
- 10) NTT ドコモ: モバイル空間統計に関する情報, (https://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/, last access: 2020/09/23).

(Sep. 30, 2020 受付)

THE TIME-SERIES ANALYSIS OF INTER-PREFECTURE TRAVEL UNDER THE COVID-19 BY SPARSE NON-NEGATIVE MATRIX FACTORIZATION

Kaisei HARADA, Hiromichi YAMAGUCHI and Masahiko SAGAE

The mobile phone location data enable us to obtain the accurate and temporally detailed inter-prefecture travel information. The non-negative matrix factorization can extract interpretable features by dimensionally reduction into two matrices, a time-series pattern and a spatial pattern. Furthermore, by adding a penalty term of L_1 norm to the model, it is possible to extract more distinct features. In this paper, we extract the characteristics of intra-residence stay and inter-regional migration by comparing demographic patterns before COVID-19 (normal time) and during the COVID-19 epidemic using 2284 days (2014.3.1, \dots , 2020.5.31) of residence-migration destination table data per each prefecture. In particular, we analyzed the effect of reducing commuting to the Tokyo metropolitan area, reducing the number of weekend outings and returns to the city during long vacations, and increasing the number of people who stay in their homes.