

2022年度 統計データ分析コンペティション

総務大臣賞 [大学生・一般の部]

パンデミックは人流をどう変えたか
—地域の特性別—

大古 一聡、西川 直輝

(東京大学大学院情報理工学系研究科)

論文の概要

都道府県別の人流データを用い、移動距離やサービス業割合、旅館数等で重力モデルによる回帰分析を行い、新型コロナウイルス感染症による影響で長距離移動やサービス業・観光業の盛んな地域への移動が減ったことを示した。

論文審査会コメント

テーマが今日的な関心事であること、統計モデルの提案とその実証に関する新たな提案がされており、学術的意義が高く評価できる。政策提言的側面よりは、コロナ下の人流移動変化の現状を極めて良く記述した努力が高く評価される。

パンデミックは人流をどう変えたか ―地域の特性別に―

大古 一聡 ^{*1}・西川 直輝 ^{*1}

^{*1}: 東京大学 (equal contribution)

1. 導入

新型コロナウイルス感染症によって、わが国における人の流れは大きく変化した。これは、緊急事態宣言をはじめとする人流抑制策の結果であり、また、ビジネス需要の減少等が副次的に影響を及ぼしたものである。この変化は極めて大きく、例えば全国の移動人口は、1度目の緊急事態宣言中に2019年同週比で最大61.2%減少し、2022年の平均でも8.7%減少したままである⁽¹⁸⁾。人流の傾向の把握は、感染抑制と経済活動再開の両面で不可欠である。感染抑制の観点からは、繁華街における人流の増加は1週間後の感染者数の増加及び3週間後の実効再生産数の増加と関係があることが指摘されている⁽⁷⁾。同時に、人の集積・移動は経済活動の根幹であり、その変質は、新規出店や広告掲出をはじめとする経済戦略や、消費刺激策や経営支援補助金の配分といった政策立案の前提が、パンデミックの前後で異なることを意味する。

近年、人流を定量的に測る手段として、スマートフォンからのデータ取得が可能になり、人流に関する研究を後押ししてきた。この場合の代表的なデータの作成方法は、スマホアプリから取得したGPSなどの位置情報を匿名化して集約するもので、既存のアンケートベースで作成した人流データセットとの高い相関から、一定の信頼度がある手法と言える。⁽⁵⁾

パンデミック下の人流に注目した研究は大きく2つに分類できる。1つ目は、人流を説明変数として用いて、感染者数を予測し、あるいは感染の空間的な広がりをモデリングするという方向の研究である。2つ目は、人流の変化を目的変数として、その時間的・地域的な傾向を示し、また経済指標等との関係を調べるものである。本論文では、後者を扱う。この方向の先行研究では、Cevik et al. (2022)⁽²⁾が、拡張重力モデルに基づいて行動制限が観光客の移動に与える影響を分析し、Goretta (2021)⁽³⁾はこれを拡張したモデルによって先進国に比べて新興国の観光業が感染症に対して脆弱性を持っていると結論づけている。また、Long & Ren (2022)⁽⁶⁾は、人流の変化傾向は、人口密度、住宅タイプ、地域の経済的困窮度、民族・文化的構成といった指標と関連して地域ごとに異なることを指摘している。これらの研究は、人流の変化の傾向が、地域が個々に持つ性質によって、あるいは地域間の関係性によって、異なって現れることを示唆している。

一方、わが国の人流の変化については、緊急事態宣言の影響を、泰・佐々木・斧田・浅野・鈴木 (2021)⁽¹³⁾が北海道及び東京都に、根元・藤井・佐野・鈴木・雨宮・大澤 (2021)⁽¹²⁾がつくば市に、後藤・田村・田中・塚井 (2022)⁽¹⁰⁾が広島県にそれぞれ注目して定量分析しており、泰らと後藤らは共通して都市部での経済活動の回復の遅れを示唆する結果を得ている。国土交通省 (2022)⁽¹⁶⁾は、全国をブロックに分け、それぞれのブロックで滞在者数の推移の解析と市区町村のクラスタリングを行っている。国土交通省は移動を距離別に分類し、遠距離からの来訪者が減少した一方で同一市区町村内での移動はむしろ増加していることを示した。その他、NTTドコモ社やAgroop社が自社の持つ流動人口データを可視化したものを逐次提供している。また、観光業への影響を調べることを念頭にした人流データの研究も多く、山本 (2021)⁽¹⁵⁾は遠方から京都に来訪する観光客が減少していることを、小西・原嶋・荒川・峯 (2022)⁽⁹⁾は糸島市を事例に施設毎の新型コロナウイルス感染症の影響を、それぞれ明らかにしている。しかし、これらはどれも、地域の持つ特性に関する定量的なデータは用いておらず、恣意的な地域区分の下での、各データの要約に過ぎない。従って、日本国内において人流の変化率に影響を与える、空間的に異なるファクターを探索する研究が急務である。

1.1 貢献

本論文における我々の貢献を以下に示す。

- 市区町村内・都道府県内の人流変化率データを、感染者数の影響を除いた上で都道府県の特徴を用いて回帰した。これにより、移動距離別に人流の変化傾向が異なるという先行研究^(15,16)を裏付けるとともに、そのような傾向が都道府県毎に異なることを示した。その原因として、都道府県の人口・経済状況や産業構造が影響を与えていることを示唆した。
- 重力モデルの限界をその拡張によって克服し、発地・着地都道府県別の人流変化率のグラフデータを用いて、発地・着地それぞれの特徴量が人流に影響を与えていることを明らかにした。従来の研究は、特定地点に滞在する人口の変動に着目して解析を行うものが主であり、例外は滋賀県への流入人口の可視化⁽²⁰⁾などごく部分的なものに留まっていた。従って、これは国内において人流をネットワークとして扱った最初期の研究である。結果は、長距離の移動がより抑制されたことや、経済活動や観光業の盛んな都道府県への移動が減少したことが分かり、従来の結果を異なるモデルから補強することができた。
- 最後に、人流の時系列変化について、時間相関関数を導出し、Graphical Lasso を用いて関係の解析を行った。その結果、人流の変化は全国一律の変化成分が大きく、隣接都道府県間の影響は必ずしも大きくないことが分かった。また、唯一地域的な結びつきが見られたのは東海道・山陽地域だった。

2. 研究方法

2.1 都道府県毎の人流変化率に対する回帰

まず、地域ごとの特性が、市区町村内や都道府県内といった小規模な範囲での人流に与える影響を明らかにする。そのために、地域ごとの特性および各地域における人流の間に線形モデル

$$y_i = \sum_{k=1}^n \beta_k x_{i,k} + \alpha + \varepsilon_i$$

を考える。ここで、 y_i は対象となる人流（市区町村内からの移動/都道府県内からの移動） $x_{i,k}$ は各地域の特徴量、 ε_i は正規分布に従うノイズであるとする。今回、各地域における人流は時系列データとして得ることができる。人流に影響を与える時間変化する変数として新型コロナウイルス感染者数の推移 $z_{i,t}$ があり、これを加えて時間変化する人流 $y_{i,t}$ が以下に従うとする。

$$y_{i,t} = \sum_{k=1}^n \beta_k x_{i,k} + \alpha + \gamma z_{i,t} + \varepsilon_i$$

以上のモデルで二乗誤差最小化によるパラメータ推定を行った。

2.2 都道府県間移動人口流のネットワークデータに対する回帰

次に、緊急事態宣言前後の、都道府県間移動人口の変化率データを用いて、移動人口流のネットワークデータとしての解析を行う。前項では、各都道府県内部の人流に注目し、他都道府県からの流入による人流は考えなかったが、本項では人流をネットワークと捉え、発地・着地の特徴が緊急事態宣言前後の変化にどのように影響を及ぼしているかについて分析する。

通常、人口流をネットワークとして分析する場合、以下の重力モデルを用いた分析が用いられる^(2,3)。(拡張)重力モデルは、地点 i, j 間の人口流 $F_{i,j}$ が、地点 i での人口 p_i 、地点 j での人口 p_j 、 i, j 間の距離 $d_{i,j}$ 、及び各地点のそれ以外の特徴量(所得・産業構造等) $\{x_{i,k}\}_{k=1}^n, \{x_{j,l}\}_{l=1}^n$ を用いて、

$$F_{i,j} = a_0 \frac{p_i^{\alpha_1} p_j^{\alpha_2}}{d_{i,j}^{\alpha_3}} \times \prod_{k=1}^n x_{i,k}^{\alpha_k} \times \prod_{l=1}^n x_{j,l}^{\beta_l} \quad (1)$$

とモデル化される。ここで、 $a_0, \dots, a_3, \alpha_i, \beta_j$ が推定すべきパラメータである。実際、Cevik (2022)⁽²⁾及びGoretti

(2021)⁰はこのモデルに基づき、 $\{x_{i,k}\}_{k=1}^n, \{x_{j,l}\}_{l=1}^n$ の変化が $F_{i,j}$ に影響を及ぼすと仮定している。

しかし、この設定では、人口流の変化について、各都道府県の特徴量が変化する十分長い時間スパンでの分析とならざるを得ない。実際、Goretti (2021)⁹は過去の感染症の中長期的な影響を目的としている。一方、今回の緊急事態宣言前後の変化のモデリングでは、(1)の右辺は変化せず、従って人口流の変化を正しくモデリングすることができない。

そこで、我々は重力モデル(1)を元に、短期的な変化を組み込むための以下の方針を提案する。緊急事態宣言の前後で、パラメータ $a_0, \dots, a_3, \alpha_i, \beta_j$ が、 $a_0 + \delta a_0, \dots, a_3 + \delta a_3, \alpha_i + \delta \alpha_i, \beta_j + \delta \beta_j$ のように変化したとする。このとき、(1)式を元に変形すると、等式

$$\log(1 + \delta F_{i,j}/F_{i,j}) = \delta a_1 \log p_i + \delta a_2 \log p_j - \delta a_3 \log d_{i,j} + \sum_{k=1}^n \delta \alpha_k \log x_{i,k} + \sum_{l=1}^n \delta \beta_l \log x_{j,l} \quad (2)$$

を得る。これより、問題は $\delta a_1, \dots, \delta a_3, \delta \alpha_i, \delta \beta_j$ を推定することに帰着され、特にこれらの符号は、各地域の各特徴が人流の減少に与える影響の正負を表す。

最後に、例えば坂本・川本 (2022)¹¹は大都市から地方への移動に重力モデルを適用しており、この場合発着地 i, j には有向の関係がある。一方、本項の分析では、人流の多くがビジネス客の出張や観光といった、行きと帰りがセットになった移動であると推測される。従って、不必要にモデルが複雑になることを防ぐため、(2)に制約 $\alpha_k = \beta_k$ を課した上で、Elastic Net 回帰を行った。2つの正則化項を合わせたスケールを決めるパラメータ λ は5-fold クロスバリデーションにより、 $\lambda = 0.015$ が、L1 正則化項の割合を決めるパラメータ α は同様に $\alpha = 0.5$ が選択された。

2.3 Graphical Lasso を用いた都道府県別人流傾向の関係可視化

最後に、人流変化における空間的な関係を見るため、Graphical Lasso による可視化を行った。Graphical Lasso では、各観測データ $x \in \mathbb{R}^d$ が、平均ベクトル μ 、分散共分散行列 Σ の正規分布 $N(\mu, \Sigma)$ に従っていると仮定し、 $\Lambda = \Sigma^{-1}$ のスパース推定を行う。 $\Lambda_{i,j} = 0$ は、 x_i, x_j は他の変数を与えた下で条件付き独立を意味するため、 Λ のスパース推定は、変数間のスパースな依存関係を求めていることになる。スパース度を決めるパラメータを λ として、以下の最適化問題を解くことに帰着される。

$$\operatorname{argmax}_{\Lambda, \mu} \frac{1}{2} \operatorname{Tr} \Lambda S - \frac{1}{2} \log \det \Lambda + \lambda \|\Lambda\|_1 \quad (3).$$

ここで、 S は標本共分散行列である。(3)は、座標降下法により解くことができる。スパース度パラメータ λ は5-fold クロスバリデーションにより、 $\lambda = 0.25$ が選択された。

時系列データに対して Graphical Lasso を適用する場合、異なる時点での観測は完全に独立と見なすことはできない。しかし、Chen, Xu, & Wu (2013)¹¹は時間方向の相関が時間経過とともに減衰するという条件を課した下で推定量の漸近収束を示しており、また実用上、各国の金融指標の相関など、時空間データへの適用の例は多い⁸。今回は、2年分のデータを得られていることを考えれば、時間方向の相関は小さいと仮定できるため、Graphical Lasso を採用する一定の妥当性があると言える。

3. 使用データと前処理

説明変数として、はじめに各都道府県の経済活動のスケール、経済状況生活状況、産業構造に関連する9個の説明変数をSSDSE(教育用標準データセット)から取得した(表の(A)から(I))。経済活動のスケールに対応する変数として都道府県別総人口、経済状況に対応する変数として1人あたり県民所得、生活状況に対応する変数として15~64歳人口割合及び一般世帯数に占める単独世帯数の割合を用いた。また、産業構造に対応する

変数としては、全体の就業者に占める第1, 2, 3次産業就業者の割合をそれぞれ算出し、特に昨今のテレワークの普及など「新しい生活様式」への適応を積極的に進めていると考えられる業界として情報通信業従業者数割合を変数として加えた。加えて、旅館営業施設数（ホテル含む）を人口の移動に大きく寄与していると考えられる観光業の規模を示すために用いた。

さらに、時系列で変化する人流に関する制御変数として、該当週に発表された新型コロナ新規感染者数の和を利用した。この値は、NHKにおいて提供されている日次の値を週毎に合計することで計算した⁽²¹⁾。分析 2.2 においては、分析に使用できる期間が短いため、制御変数として新型コロナ新規感染者数に加え、当該期間の気温・降水量の平均を制御変数として用いた。これと(J)は、気象庁より取得した⁽²²⁾。

従属変数として、人流に関するデータ 2 つのデータを用いた。1 つ目は、内閣府ビッグデータチームが V-RESAS において提供している、週別の各都道府県の滞在人口のうち、同一市区町村内に居住する人数の割合、同一都道府県内に居住する人数の割合である⁽¹⁸⁾。このデータを用いることで、市区町村内の比較的小規模な移動を伴う人流と、都道府県内のより大規模な移動を分けて解析することができる。新型コロナ感染拡大当初は人流が急激に変化する時期が多く、データに多くのノイズが乗っていると考えられることから、2021 年の 1 年間のデータを利用した。2 つ目は NTT ドコモがモバイル空間統計において提供する、緊急事態宣言中の 2020 年 6 月における、2 つの都道府県間の移動者数の 2019 年比データである⁽²⁰⁾。表 1 で前者が 1, に、後者が 2 に用いられている。また、V-RESAS から取得した人流の合計データ(発着地を区別しない)を 3. で用いた。

使用する変数の単位と要約統計量を以下の表に示した。アルファベットは説明変数を、数字は従属変数を表す。

表 1: 使用変数の要約統計量

項目	単位	サンプル数	平均	標準偏差	最小値	最大値
(A) 都道府県総人口	人	47	2,683,960	2,767,237	553,407	14,047,590
(B) 15~64歳人口割合	%	47	55.8	2.5	52.2	63.7
(C) 一般世帯数に占める単独世帯数の割合	%	47	34.8	4.1	28.4	50.2
(D) 1人当たり県民所得（平成23年基準）	千円	47	2,983	469	2,349	5,427
(E) 旅館営業施設数（ホテルを含む）	軒	47	1,085	702	344	3,308
(F) 就業者に占める情報通信業従業者の割合	%	47	1.5	1.3	0.5	9.4
(G) 就業者に占める第1次産業就業者の割合	%	47	5.2	2.9	0.4	11.4
(H) 就業者に占める第2次産業就業者の割合	%	47	24.8	5	14.3	33.3
(I) 就業者に占める第3次産業就業者の割合	%	47	70	5.2	62.6	84.4
(J) 都道府県庁間距離（正規化後）	人	1275	608	2,019	0	34,173
(K1) 当該期間の週別発表感染者数（2021年第1週, ..., 51週）	人	2397	0.249	1	0	18
(K2) 当該期間の発表感染者数平均（2020年6月1~14日）	人	47	0.852	3.35	0	22
(L) 平均気温（2019年6月13・14日）	°C	47	20.09	1.6	16.75	25.55
(M) 平均気温（2020年6月15・16日）	°C	47	21.89	1.51	16.75	25.95
(N) 降水量（2019年6月13・14日）	mm	47	17.71	8.79	0	38
(O) 降水量（2020年6月15・16日）	mm	47	3.14	7.75	0	37.75
1 都道府県・週別人流変化率（2021年, 同一市区町村内）	&	2397	4.59	2.73	-1.96	16.86
2 都道府県・週別人流変化率（2021年, 同一都道府県内）	%	2397	-4.92	6.84	-37.82	34.35
3 都道府県間人流増減率	%	1275	-47.71	15.77	-85.798817	72.5274725
3 異なる市区町村からの滞在人口の2019年同週比変化率	%	6486	-12.41	12.1	-72.68	52.33

それぞれのデータは、平均 0, 分散 1 に正規化して用いた。分析 2.1 では(A)-(G), (I), (K1), 1 を、分析 2.2 では(A), (B), (D)-(J), (K2), (L)-(O). 2 を、分析 2.3 では 3 をそれぞれ用いた。比のデータで対数がモデルに出てくる場面では、 $\log(1+x) \approx x$ の近似を用いて、変化率を対数を取ることなく用いた。

4. 分析結果

本章では、それぞれの分析によって得られた結果を示し、その解釈を述べる。

4.1 都道府県毎の人流変化率に対する回帰

表2及び3に同一市区町村内・同一都道府県内の移動人口それぞれに対する回帰モデル(1)で得られた係数を示す。ここから、以下のことが分かる。

都市部に特有の性質と人流変化の関係は移動のスケールによって異なる

都道府県人口(A), 15~65歳人口割合(B), 情報通信業従事者割合(F)は、それぞれ同一市区町村内の移動に対して正の、市区町村を超えた都道府県内の移動に対して負の関係がある。これは、(i)テレワークの導入等により、生活範囲がより自宅の近くにシフトしたことで移動のスケールによって逆向きの変化が生じている、(ii)さらに、都市部の特徴が、双方の向きへの変化をよく説明していること、が分かる。(i)については例えば国土交通省⁽²²⁾のビッグデータを用いた調査によって明らかになっていたが、これを地域の特徴と結び付く可能性を示唆できる点が興味深い。

表2: 同一都道府県・同一市区町村内の移動に対する結果

変数	回帰係数	変数	回帰係数	変数	回帰係数
(A)	0.2557	(D)	0.143	(G)	0.1775
(B)	0.2845	(E)	-0.0676	(I)	0.2454
(C)	-0.2923	(F)	0.0014	(K1)	0.1788

表3: 同一都道府県内の市区町村を越えた移動に対する結果

変数	回帰係数	変数	回帰係数	変数	回帰係数
(A)	-0.1454	(D)	0.1253	(G)	-0.1464
(B)	-0.2074	(E)	-0.0591	(I)	0.1193
(C)	-0.0941	(F)	-0.2554	(K1)	-0.0107

4.2 都道府県間移動人口流のネットワークデータに対する回帰

Elastic Net 回帰の結果、以下の係数ベクトルを得た。

表4: Elastic Net 回帰の結果

変数	回帰係数	変数	回帰係数	変数	回帰係数	変数	回帰係数
(A)	-0.1396	(F)	0	(J)	-0.2762	(N)	0.0372
(B)	0.1387	(G)	0.0706	(K2)	0	(O)	-0.2578
(D)	-0.1636	(H)	-0.0389	(L)	0.2051		
(E)	-0.091	(I)	-0.1514	(M)	0		

制御変数を除いて最も負方向の寄与が大きいのは、移動都道府県間の距離(J)で、次いで1人当たり県民所得(D)、第3次産業従事者の割合(I)、都道府県の人口規模(A)、旅館営業施設数(E)となった。一方、一般世帯数に占める単独世帯数の割合(B)は正の寄与を示し、また都道府県別新型コロナウイルス感染者数(K)は正則化項によって係数が0になった。以下にその解釈を述べる。

長距離の移動がより抑制される

移動都道府県間の距離(J)の回帰係数が最も負方向に大きかったことから、新型コロナウイルス感染症の流行及び緊急事態宣言の発令は長距離の移動をより大きく抑制する影響をもたらしたと言える。観光客の動向については山本 (2021) ⁽¹⁵⁾が京都の観光客の出発地が関西都市部周辺に限定されるようになったことを示しており、

この結果は同様の傾向がビジネス需要を含めた全体として成り立っていることを示唆する。

人口が多く経済活動が盛んな都道府県への移動人口が減少している

1人当たり県民所得(D), 第3次産業従事者の割合(I), 都道府県の人口規模(A)が負方向の寄与を示しており, これは都市部に高い指標であるため, 都市部への移動がより多く減少していることが分かる。

サービス業や観光業の盛んな発着地の移動ほど抑制される

第3次産業従事者の割合(I)及び旅館営業施設数(E)は負方向の寄与を示した。第3次産業に関わる移動は, 対面のコミュニケーションを目的としたものであると考えられる。対面の会議は Web 会議システムに置き換え可能であるが, 一方第1次・第2次産業に関わる移動は例えば物資の輸送等オンラインで代替不可能であることが多く, 第3次産業に関わる移動が多い地域では, 移動が抑制可能であったと考えられる。また, 緊急事態宣言直後では, 観光旅行は基本的に自粛されており, 観光業の盛んな地域では人流のより多くの割合を占める観光客の減少したことが現れていると考えられる。

4.3 都道府県間移動人口流のネットワークデータに対する回帰

図1及び表5には, Graphical Lasso 解析を行った結果を示す。

図1には精度行列の非負成分について, その重みの大きさに対応する太さ・色で辺を引いた都道府県を頂点とするグラフを, 日本地図に重ねたものを示す。このとき, (i)東京・大阪・福岡から全国に満遍なく辺が見られること, 及び(ii)東海道・山陽各県の間で特に精度行列の非負成分が大きいことが見て取れる。

表5には東京都と辺を持つ都道府県について, 対応する辺の重みの絶対値でソートした結果を抜粋して示す。なお, 辺を持つ都道府県は計18道府県であった。ここから以下のことが分かる(i)隣接県だけではなく, 愛知・和歌山・徳島など距離的に離れた地域とも辺を持っている。(ii)神奈川・愛知・大阪・広島・福岡の各東海道・山陽主要都道府県が東京都と辺を持っている。以上は, 図1の全体的な傾向と一致する。

以下, 解釈を述べる。

隣接都府県間の関係が小さく, 大都市を中心にした半径の小さいグラフ

図1及び表5の両方で, 東京・大阪・福岡の各都府県がハブとなり, 半径の小さいグラフが得られていることが分かる。これは, 各都道府県の人流が, 隣接都府県間の関係よりも地理的に離れた大都市の動向と, 関係が強いことを意味している。この一因には, 大都市の感染状況の報道や, 政府の全国に対する行動制限の影響の強さが示唆される。

東海道・山陽の都府県の結びつきの強さ

図1及び表5の両方で, 地域的な結びつきが唯一観察されたのが東海道各都府県である。これは, 経済的な関係が強く, また相互間の交通が発達しており, 人口の移動が地域内で一体的に行われていることを反映していると言える。

5. まとめ

今回の分析では, 以下の知見を得た。1つ目・2つ目の分析からは, まず, 人々の行動範囲が全体として縮小しているという先行研究^(15, 16)を裏付ける結果を得た。1つ目の分析では, 市区町村内と都道府県内というレベルで人流の変化が異なることが分かり, これと整合する結果として2つ目の分析で遠距離の移動ほど抑制されるという結果も得た。さらに, 1つ目・2つ目の分析からは, 都市部への移動ほど抑制されることが分かった。業種別には情報通信業をはじめとするサービス業や, 観光業に関連した人流の落ち込みが示唆された。このように地域の特徴が人流の変化率と関係を持つことを示すことができた。

さらに, 2つ目の分析では, 重力モデルを元に係数の変化を推定するというモデルを立て, 分析を行ってい

る。人流をネットワークとして捉えるモデリングは、国内では先例が少ないが、1つ目の分析と整合する結果を得ることができた。

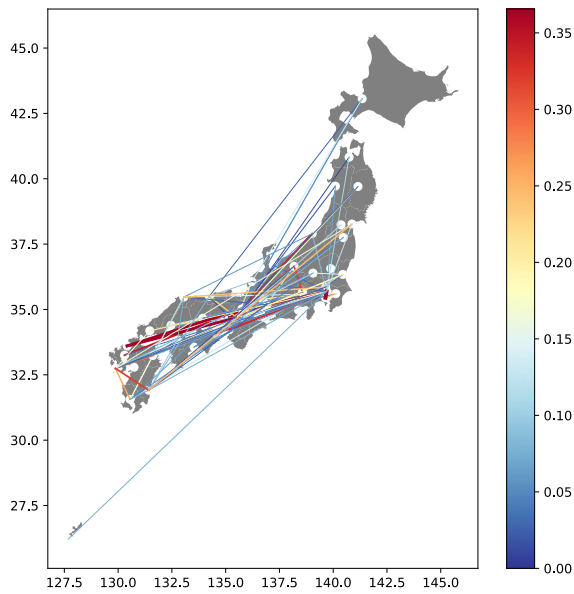


表 5: 精度行列の重みソート (対東京都)

地域1	順位	地域2	係数
東京都	1	神奈川	0.8408
	2	愛知	0.4446
	3	和歌山	0.338
	4	徳島	0.3312
	5	山口	0.2089
	7	大阪	0.1984
	9	新潟	0.0831
	11	沖縄	0.0751
	17	広島	0.0037
	18	福岡	0.0031

図 1: 人流増減の相関

3つ目の分析からは、人流の推移傾向は隣接都道府県内よりも大都市との結びつきが強いことが分かった。Graphical Lasso の辺は必ずしも最寄りの大都市に引かれている訳ではないため、大都市が直接的に影響を与えているというよりむしろ、時系列で見れば人流の傾向は全国的なトレンドが主であると考えるのが自然である。ただし、東海道・山陽地域では、強い相互関係を見ることができた。

ここから導かれる提言としては、人流が、ネットワークとして、大都市をハブとするものから、より小さな地域内で完結する分散型に、移り変わっていることが示唆される。従って、地域内モビリティの充実などが求められていると言える。また、3つ目の分析では、人流の抑制は第一に全国一様な施策が大きな効果を持つことが推測できる。従って、特に東海道・山陽地域では、自治体が一体となって感染症対策を行うことが肝要である。

今後の課題として、ここでは2点言及する。1つ目は、時系列データに対する今回の Graphical Lasso は人流傾向が遅れて伝搬する様子をモデリングできないことである。これについて、全国の傾向ではなく、より高次の傾向を得るのは今後の課題である。2つ目は、国内の人流オープンデータの整備である。特にネットワークフローデータは限定的なものしか得られなかった。Kang et al. (2020)のようなオープンデータセットの整備によって、より精度の高い分析が可能になると期待される。

結論を述べる。本論文では、新型コロナウイルス感染症流行下で人流が変化した様子を、各都道府県が持つ特徴と関連づけて議論した。従来指摘されていた移動距離別の人流の変化の傾向が、各地域の持つ特徴量とも関わっていることが示唆された。これは、線形回帰分析に加えて、重力モデルの係数が変化したと仮定したモデルによっても導かれた。さらに、Graphical Lasso を用いて、人流時系列データの空間的な関係を明らかにした。

参考文献

- (1) Chen, X., Xu, M., & Wu, W. B. Covariance and precision matrix estimation for high-dimensional time series. *The Annals of Statistics*, 41(6), pp. 2994-3021, (2013).
- (2) Cevik, S. Going Viral: A Gravity Model of Infectious Diseases and Tourism Flow. IMF Working Paper 20/112, *International Monetary Fund* (2020).
- (3) Goretti, M., Leigh, L. Y., Babii, A., Cevik, S., Kaendera, S., Muir, D. V., Nadeem, S., Salinas, G. Tourism in the Post-Pandemic World: Economic Challenges and Opportunities for Asia-Pacific and the Western Hemisphere. *International Monetary Fund* (2021).
- (4) Friedman, J., Hastie, T., & Tibshirani, R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3), pp. 432-441 (2008).
- (5) Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J., Kruse, J. Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. *Scientific Data*, 7, 390, (2020).
- (6) Long, J. A., Ren, C. (2022). Associations between mobility and socio-economic indicators vary across the timeline of the Covid-19 pandemic. *Computers, Environment and Urban Systems*, 91, 101710.
- (7) Nakanishi, M., Shibasaki, R., Yamasaki, S., Miyazawa, M., Usami, S., Nishiura, H., & Nishida, A. (2020). On-site Dining in Tokyo During the COVID-19 Pandemic: Time Series Analysis Using Mobile Phone Location Data. *JMIR Mhealth Uhealth*, 9(5):e27342.
- (8) 岩波データサイエンス刊行委員会：“岩波データサイエンス Vol.5”，岩波書店 (2017).
- (9) 小西 宏樹, 原嶋 春輝, 荒川 豊, 峯 恒憲：“観光EBPMに向けた大規模観光人流データ分析～糸島市西部～”，マルチメディア通信と分散処理 (DPS), 2022-DPS-191, 42号, pp.1-8 (2022).
- (10) 後藤 拓, 田村 将太, 田中 貴宏, 塚井 誠人：“新型コロナウイルス感染症による都市内の昼間滞在人口分布の変化 —コロナ収束後の都市のあり方検討に向けて—”，日本建築学会技術報告集, 28 巻, 69 号, pp.906-911 (2022).
- (11) 坂本 大樹, 川本 晃大 “若者の大都市から地方への移動要因を探る —修正重力モデルによる分析—”，「統計」2022年3月号, 日本統計協会 (2022).
- (12) 根本 裕都, 藤井 さやか, 佐野 雅人, 鈴木 勉, 雨宮 護, 大澤 義明：“COVID-19による外出制限が人口流動に及ぼす影響 2020年4月の緊急事態宣言期間中のつくば市の事例分析”，都市計画論文集, 56 巻, 3号, pp.1207-1214 (2021).
- (13) 秦 康範, 佐々木 邦明, 斧田 佳純, 浅野 礼子, 鈴木 俊博：“コロナ禍における緊急事態宣言等が都道府県間流動に与えた影響の検証”，土木学会論文集D3 (土木計画学), 77 巻, 2号, pp.151-159 (2021).
- (14) 宗健：“テレワークに関する実証分析—どのような属性のひとがテレワークをしているのか—”，日本マーケティング学会カンファレンス・プロシーディングス, 9 巻, pp.148-156 (2020).
- (15) 山本 真嗣：“コロナ禍における観光地の人流の変化の考察”，第36回日本観光研究学会全国大会学術論文集, pp.185-188 (2021).
- (16) “令和3年度広域的な人流データ活用による行動変容分析業務報告書 概要編”，国土交通省 不動産・建設経済局, (2022).
- (17) SSDSE: 教育用標準データセット, 独立行政法人統計センター (2022年8月20日取得).
- (18) V-RESAS, 内閣府地方創生推進室ビッグデータチーム (2022年8月20日取得).
- (19) 都道府県をまたぐ移動自粛解除後 (2020年6月19日以降) の県間移動データ, NTTドコモ (2022年8月20日取得).
- (20) “人流ビッグデータを活用した都道府県間人口移動の分析について”，滋賀県 (2020).
<https://www.pref.shiga.lg.jp/kensei/koho/e-shinbun/oshirase/314028.html>
- (21) “特設サイト新型コロナウイルス 都道府県ごとの感染者数の推移”，NHK (2022年8月20日取得).
<https://www3.nhk.or.jp/news/special/coronavirus/>
- (22) “過去の気象データ検索”，気象庁 (2022年8月20日取得).
<https://www.data.jma.go.jp/obd/stats/etrn/>