
原著論文

大都市から地方への移住と社会経済的要因の関連

—Elastic net回帰を用いたポアソン重力モデルによる分析—

The Relationship of Socioeconomic Factors and Migration from Large Cities to Rural Areas: Poisson Gravity Model Analysis with Elastic Net Regression

キーワード：

地方創生, 地域間人口移動, ポアソン重力モデル, Elastic net回帰

keyword：

Regional revitalization, interregional migration, Poisson gravity model, elastic net regression.

経済産業研究所 コンサルティングフェロー・東京大学大学院 情報理工学系研究科 協力研究員 荒川清晟
Consulting Fellow, Research Institute of Economy, Trade and Industry; Cooperative Researcher,
Graduate School of Information Science and Technology, The University of Tokyo Kiyooki ARAKAWA

東京大学大学院 情報理工学系研究科 客員研究員 野寄修平
Visiting Researcher, Graduate School of Information Science and Technology, The University of Tokyo
Shuhei S NOYORI

要約

地方から都市圏への人口流出が続く中、移住誘致のための取り組みが全国の市区町村で実施されている。その一方で、日本国内の人口移動に関する研究は、市区町村単位で解析したもの、大都市から地方への人口移動流を対象としたものが少なく、これらの人口移動に関連する社会・経済的な要因は未だ明らかになっていない。本研究では、これを明らかにするため、大都市から地方への人口移動量と社会・経済的な要因（所得、就業機会、アメニティなど、移動先と移動元合わせて53変数）の関連を市区町村単位で網羅的に分析した。一般に、回帰モデルはモデルに使用できる変数の数の制限や多重共線性の

原稿受付：2022年9月25日

掲載決定：2023年2月9日

問題があり、多くの変数を使用することは難しい。これらの問題に対処するため、本研究では、モデル内に取り込める説明変数の数の制限が弱く、かつ相関の強い変数の回帰係数の絶対値が近い値をとる特徴（Grouping効果）をもつElastic net回帰を使用した。この結果、重力モデルの基本変数である距離、移動元の総人口、移動先の総人口を含む20の変数が移動人口と関連のあるものとして選択された。今後は年齢や子供の有無など属性別の分析、複数年次のパネルデータを用いた各要因の人口移動量への影響評価が必要である。

Abstract

Efforts are underway to attract interregional immigration by every local government throughout Japan to deal with the population outflow from rural areas to urban areas. However, few studies on population migration in Japan have analyzed migration at the municipal level or those from large cities to rural areas so far. In this study, to clarify what kind of socioeconomic factors are associated with migration from large cities to rural areas, we analyzed the relationship between migration from large cities to rural areas and socioeconomic factors (53 variables in total, including income, employment opportunities, amenities, etc.). In general, it is difficult to use many variables in regression problems due to the limitation of the number of variables that can be used in a model and the problem of multicollinearity. To address these problems, we used elastic net regression, which has a weak restriction on the number of explanatory variables and a feature in which the absolute values of the regression coefficients become close if they are strongly correlated (Grouping effect). As a result, 20 variables were selected as relevant to migration, including the basic variables of the gravity model: distance between the source and destination cities, populations at the source and destination cities. Future study includes analyses of the relationship between personal characteristics such as age and whether they are raising children on migration and migration, as well as evaluation of the impact of each factor on migration using panel data.

1 はじめに

日本の総人口が2008年をピークに減少に転じ、少子高齢化が深刻な状況にある中、東京圏への人口の一極集中が大きな問題となっている（閣議決定、2020）。東京圏に人口が集中すると、人口が流出している地域の魅力・活力を低下させ、更なる人口流出を招くことになり、特に中山間地域や農山漁村等では、不可欠な生活サービスの維持・確保が困難となる（閣議決定、2019）。そこで、政府は東京圏への人口の一極集中の是正を図るため、「まち・ひと・しごと創生総合戦略」の中で、「地方への新しいひとの流れをつくる」という基本目標を設定し（閣議決定、2015）、地方移住を推進している。

東京圏から地方への移住に限らず、人口の移動要因はこれまでに複数の観点から研究されてきた。人口移動要因は、当初は所得格差や就業機会に着目していたが、周辺的生活環境も考慮するようになり、対象とする人口移動流についても、大都市から地方への移動流と地方から大都市への移動流を分けた分析が求められるようになった（青木・稲村、1997）。しかし、これらの研究は、データの制約上都道府県単位の分析が多く、市区町村単位の分析はあまり行われていない（荒川、2018）。

人口移動要因の分析手法は、人口集団とそれを取り巻く地域条件との関連に着目するマクロな分析、個人の移動理由に着目するミクロな分析に大別される（日本人口学会、2002、606頁）。マクロな分析は、ミクロな分析に比べてデータを収集しやすく、網羅的な分析が可能である。しかし、マクロな分析で用いられることが多い重力モデルを用いた回帰分析では、多重共線性の問題を回避するため、相関の強い複数の変数を同時に扱うことができない（伊藤、2003；荒川、2018）。これまでの研究では、相関の強い変数は、片方の変数のみを残して分析したり（荒川、2018）、それぞれ

別に回帰式を作成して分析したり（伊藤、2003）するなどの対策をとっているが、説明変数の選択に研究者の恣意性が入る余地がある、同時に考慮できる説明変数が少なくなるという問題があった。

本研究の目的は、大都市から地方への人口移動に関連する社会・経済的な要因を市区町村単位で網羅的に分析すること、さらにその際に生じる多重共線性の問題を解決するため、重力モデルに対してElastic net回帰を適用することである。本論文は次の通りに構成される。第2章では、関連する先行研究を整理するとともに、本研究の位置付けを示す。第3章では、本研究で使用するデータについて説明し、第4章では本研究で使用するモデルと分析手法を説明する。第5章では、分析結果を示し、第6章でそれらを踏まえた考察を述べ、第7章で、まとめとして本研究の結論について述べる。

2 関連する先行研究

人口移動要因に関する研究には、人口集団とそれを取り巻く地域条件との関連に着目するマクロな手法と、個人の移動理由に着目するミクロな手法がある（日本人口学会、2002、606頁）。これらの方法を用いて、様々な人口移動流を対象とした分析が行われているが、Cadwallader（1996、334-337頁）によると、人口移動要因は、主に所得格差、就業機会、教育、行政サービス、アメニティ、年齢に分類される。以下、本研究に関連するマクロな手法を用いた先行研究について、国内の人口移動を対象に要因別に整理する。

所得格差は、人々は自己の効用を最大にするため、賃金あるいは所得の高い地域へ移住するという経済学の伝統的モデルを基にしている（伊藤、2001；伊藤、2003）。これまでの研究では、各都道府県間・各地方間の人口移動を対象とした分析で、高齢者以外ではこのモデルが成立するものの、高齢者では一部逆の現象が観察されている。

転入先の所得水準については、全年齢・各年齢層で所得が高い地域への転入者が多く（伊藤，2003；張ら，2016；當麻，2016），転出元の所得水準について、高齢者以外（特に15-24歳）は転出元の所得が低いほど転出者が多いが、高齢者は転出元の所得が高いほど転出者が多かった（伊藤，2003）。

就業機会は、賃金の地域的差異よりも就業機会の地域的差異が労働人口を移動させるという考えを基にしている（青木・稲村，1997）。當麻（2016）は、都道府県間の人口移動を対象に経済社会的要因、地域アメニティとの関連を分析し、有効求人倍率の高い地域への転入者が多いことを明らかにした。

教育は、15歳-34歳を対象とした場合、大学・短大数の多い地域への転入者が多いが、幼稚園保育園数が多い地域への転入者数は少ないという報告がある（張ら，2016）。また、大学進学に伴って移住する者を対象とした場合、教員数の少ない地域へ、科研費配分額が高い地域から低い地域への転入者が多いとの研究もある（田村，2017）。さらに、大都市から地方への市区町村間人口移動を対象に、教育・子育てのまちという地域イメージが高い大都市は、転出者が多いという報告がある（荒川，2018）。

行政サービスは、都道府県間の人口移動を対象に、知事演説から解析した政策プライオリティとの関連を分析し、経済的・社会的・文化的政策プライオリティが人口移動に影響を与えることが明らかにされている（尾崎，2013）。また、各地方間の人口移動を対象に、公共的なサービス水準を含む社会環境アメニティが人口移動に影響を与えること（伊藤，2006a）、高齢者を対象とした場合、面積当たり医師数が影響を与えることが明らかにされている（伊藤，2011）。

アメニティは、地方間の人口移動を対象に、寒冷地からの転出者が多く、積雪の少ない地域への転入者が多いと報告されている（伊藤，2003）。

また、地方間の人口移動を対象に、主に35歳以上で、温暖な地域への転入者が多いことが報告されている（伊藤，2006b）。

年齢については、年齢層別に分析が行われており、伊藤（2003）、伊藤（2011）、張ら（2016）、田村（2017）などで、年齢層ごとの人口移動量を被説明変数とした分析が行われている。その結果、前述の通り、所得水準に関して高齢者のみ転出元の所得が高いほど転出者が多いとの報告や（伊藤，2003）、教育に関して15歳-34歳の若年層では大学・短大数の多い地域への転入者が多いが、幼稚園保育園数が多い地域への転入者数は少ないという報告など（張ら，2016）、世代間で異なる傾向が見られている。

これまでの研究の問題として、（1）人口移動を都道府県単位で解析していること、（2）大都市から地方への人口移動流を対象とした研究が少ないこと、（3）相関の強い説明変数を同時に扱えていないことの3点が挙げられる。（1）人口移動を都道府県単位で分析していることについて、これは国内の市区町村単位の人口移動統計が近年まで公開されていなかったことに起因している。そのため、市区町村単位の人口移動を分析している研究は、荒川（2018）のみと限られている。Uターンの人口移動が、都道府県単位ではなく市町村単位で生じていることや（江崎ら，2000）、「まち・ひと・しごと創生総合戦略」を始めとする移住定住関連施策が市区町村単位で作成されていることを考えると、市区町村単位の分析が求められる。（2）大都市から地方への人口移動流を対象とした研究が少ないことについて、これまでの多くの研究では、都市の属性を無視して、地方から大都市への集中移動、大都市から地方への分散移動をまとめて分析していた。地方から大都市への集中移動、大都市から地方への分散移動は、移動理由や移動先の都市に求めることが異なると考えられ、両者を区別して分析する必要がある（青木・稲村，1997）。（3）相関の強い説明変

数を同時に扱っていないことについて、多重共線性の問題から、回帰モデルの中で相関の強い説明変数を同時に扱ってこなかった（伊藤，2003；當麻，2016；荒川，2018）。そのため、変数選択の恣意性や、同時に考慮できる変数が減るという問題が生じている。

本研究では、(1)、(2)の問題への対応として大都市から地方への分散移動に注目し、政令指定都市と東京都特別区部から、大都市圏に属さない地域への人口移動を市区町村単位で分析する。先行研究で使用されている社会経済変数を可能な限り網羅的に使用した上で、(3)の問題への対処として、相関の強い変数を同時に扱うことのできるElastic net回帰を用い、それぞれの変数が大都市から地方への市区町村単位の人口移動にどのように関連しているかを明らかにする。相関の強い変数を同時にモデルに組み込むことで、これまでの研究では他の変数の影響を考慮せずに関連を評価していたり、因子分析や主成分分析で次元削減していたために変数単位ではなく概念単位でしか関連を評価できなかつたりしたところを、他の変数の影響を調整した上で、変数単位で人口移動との関連を評価できるようになる。一般に、社会経済変数には相関の強いものが多いため、この手法は人口移動に限らず社会情報学分野一般に応用可能であると考えられる。

3 使用したデータ

3.1 移動元と移動先の地域の定義

本研究では、移動元を政令指定都市と東京都特別区部、移動先を大都市圏に属さない市町村とする。大都市圏に属する市町村は、e-Stat政府の統計窓口（以下e-Stat）に掲載されている平成27年国勢調査（総務省）での大都市圏構成市町村の定義に依った（総務省統計局，2013）。なお、大都市圏に属する市町村への転入は、大都市圏中心市への通勤など中心市の影響が大きいと考えられる

ため、移動先の地域から除外した。

3.2 使用する変数

表-1に使用した変数を示す。被説明変数として用いる、移動元の地域 (i) から移動先の地域 (j) への人口移動量 M_{ij} のデータは、住民基本台帳人口移動報告（総務省）を用いた。

重力モデルの基本となる3つの変数について、人口 (P_i, P_j) は、平成27年国勢調査の人口を用い、地域間の距離 (d_{ij}) については、国土地理院ホームページから各市区町村役場の緯度経度を取得し、国土地理院測量計算サイト⁽¹⁾を用いて距離を算出した。社会経済変数については、先行研究で使用されていた変数のうち、e-Statの都道府県・市区町村のすがた（社会・人口統計体系）から取得することができるものを用いた。また、気象データは、国土数値情報（国土交通省）より各市区町村役場のメッシュデータを取得し、地価は、都道府県基準地価を取得し、各市区町村の観測地点の平均値（住宅地）を利用した。大学数は、全国学校データ（教育ソリューション協会）のものを利



図-1. 解析対象の市区町村の役所/役場所在地。解析対象とならなかった市町村役所/役場所在地のうち一部の離島は表示していない。

地図データ： <https://github.com/dataofjapan>

表-1. 使用した変数

変数名	単位	出典
総移動人口	人	住民基本台帳人口移動報告 (総務省)
距離	km	国土地理院
総人口	人	国勢調査 人口推計 (総務省)
可住地面積1 km ² 当たり人口密度	人	国勢調査 全国都道府市区町村別面積調 (国土地理院) 農林業センサス・世界農林業センサス (農林水産省)
総面積 (北方地域及び竹島を除く)	ha	国勢調査 全国都道府市区町村別面積調
可住地面積割合	%	国勢調査 全国都道府市区町村別面積調 農林業センサス・世界農林業センサス
課税対象所得 (納税義務者1人当たり)	千円	市町村税課税状況等の調 (総務省)
幼稚園数	園 *	学校基本調査 (文部科学省)
小学校数	校 *	学校基本調査
中学校数	校 *	学校基本調査
高等学校数	校 *	学校基本調査
第1次産業就業者比率	%	国勢調査
第2次産業就業者比率	%	国勢調査
第3次産業就業者比率	%	国勢調査
完全失業率	%	国勢調査
公民館数	館 *	社会教育調査 (文部科学省)
図書館数	館 *	社会教育調査
一般病院数	施設 *	医療施設調査 (厚生労働省)
老人福祉施設数	所 *	社会福祉施設等調査 (厚生労働省)
保育所等数	所 *	社会福祉施設等調査
都道府県基準地価	円	国土数値情報 都道府県地価調査データ (国土交通省)
大学 (本部) 数	校 *	全国学校データ (教育ソリューション協会)
年降水量	mm	国土数値情報 平年値メッシュデータ (国土交通省)
年間最高気温	K	国土数値情報 平年値メッシュデータ
年間最低気温	K	国土数値情報 平年値メッシュデータ
年間平均気温	K	国土数値情報 平年値メッシュデータ
年最深積雪	cm	国土数値情報 平年値メッシュデータ
年合計日照時間	時間	国土数値情報 平年値メッシュデータ

被説明変数1変数, 説明変数53変数。

* (変数×総人口/100,000) により総人口10万人あたりに統一した。総移動人口, 距離は1つの移動元, 移動先の組につき1つずつ, その他の26変数については移動元, 移動先それぞれに1つずつ。

用した。その際に, 先行研究で用いられているデータと同一のデータがない場合は, 筆者らにより適切と思われるデータを用いた⁽²⁾。その結果, 53の説明変数を用いた。なお, 気温のデータは国土数値情報メッシュデータで公開されている摂氏のデータを絶対温度に変換して用いた。

2015年度の全1,741市区町村から, 移動元から移動先への人口移動数が10人未満の市区町村の組⁽³⁾, 気候のデータが得られなかった東京都小笠原村を除外した結果残った747市区町村 (移動元43, 移動先704), 4,071組の人口移動データを解析した (図-1)。移動人口は平均42人, 中央値20人, 第1四分位13人, 第3四分位37人, 移動

元/移動先の市区町村役所/役場間の距離は平均405 km, 中央値240 km, 第1四分位114 km, 第3四分位589 kmであった。表-2に, 年齢層, 性別の総移動人口を示す。男性が全体の56.9%,

表-2. 年齢層・性別の人口移動量

	男性	女性	総数
総移動人口	97,070 (100.0)	73,673 (100.0)	170,743 (100.0)
0~9歳	9,420 (9.7)	8,983 (12.2)	18,403 (10.8)
10~19歳	5,892 (6.1)	3,651 (5.0)	9,543 (5.6)
20~29歳	32,555 (33.5)	26,323 (35.7)	58,878 (34.5)
30~39歳	21,049 (21.7)	17,620 (23.9)	38,669 (22.6)
40~49歳	13,555 (14.0)	7,368 (10.0)	20,923 (12.3)
50~59歳	8,242 (8.5)	4,025 (5.5)	12,267 (7.2)
60歳以上	6,326 (6.5)	5,677 (7.7)	12,003 (7.0)
不詳/その他	31 (0.0)	26 (0.0)	57 (0.0)

実際の値 (割合) で記載。

女性が43.1%と男性がやや多く、いずれの年齢層でも男性が女性よりも多かった。また、40、50歳代は男性が女性のおよそ2倍と性別に大きく偏りがあった。年齢の分布は、20、30歳代が全体の半数を占める一方で60歳以上は全体の10%未満であった。

4 使用するモデルと分析手法

4.1 ポアソン重力モデル

本研究では、人口移動を社会の空間的相互作用として扱うため、人口移動要因のマクロな分析で用いられるポアソン重力モデルを用いた (Flowerdew, 1982; 矢野, 1991)。ポアソン重力モデルの基礎となっている重力モデルは、人口移動量 (M_{ij}) が移動元、移動先双方の人口規模 (P_i, P_j) と地域間の距離 (d_{ij}) に対して

$$M_{ij} = G \frac{P_i^{a_1} P_j^{a_1}}{d_{ij}^e} \quad (1)$$

の関係にあるとし (Ravenstein, 1885)、移動元、移動先の人口が多いほど、また地域間の距離が小さいほど、人口移動量が多いと考える (ただし G, a_1, a_2, e は各変数の寄与を定める定数)。

重力モデルは人口移動を分析する際に人口と距離のみ考慮しているが、これに人口と距離以外の説明変数を追加したものが (2) 式に示す修正重力モデルである (V_{i1}, \dots, V_{im} は移動元の変数, V_{j1}, \dots, V_{jm} は移動先の変数, $b_1, \dots, b_n, c_1, \dots, c_m$ は各変数の寄与を定める定数) (Greenwood and Hunt, 2003)。

$$M_{ij} = G \frac{P_i^{a_1} P_j^{a_1} V_{i1}^{b_1} \dots V_{im}^{b_m} V_{j1}^{c_1} \dots V_{jm}^{c_m}}{d_{ij}^e} \quad (2)$$

(2) 式の両辺を対数変換することで、線形回帰モデルとして各変数の回帰係数を推定できる。こ

こで、(2) 式左辺の人口移動量が非負の整数であることを利用し、ポアソン分布に基づく一般化線形モデルとして回帰係数を推定する手法がポアソン重力モデルである (Flowerdew, 1982)。ポアソン重力モデルは (3), (4) 式により表現される。

$$Pr(M_{ij} = k) = \frac{\lambda_{ij}^k \cdot \exp(-\lambda_{ij})}{k!} \quad (3)$$

$$\lambda_{ij} = \exp(\ln G + a_1 \ln P_i + a_2 \ln P_j - e \ln d_{ij} + \sum_{l=1}^n b_l \ln V_{il} + \sum_{l=1}^m c_l \ln V_{jl}) \quad (4)$$

本研究では、ポアソン重力モデルを用いて、説明変数として移動元の人口 P_i 、移動先の人口 P_j 、移動元と移動先の距離 d_{ij} に加え社会経済変数 $V_{i1}, \dots, V_{in}, V_{j1}, \dots, V_{jm}$ を投入し、後述のElastic net回帰によって人口移動数と関連のある変数を選択した。

4.2 Elastic net回帰

一般に、社会経済変数には互いの相関が強いため、多重共線性の問題から線形回帰モデルに同時に投入できない変数の組が複数存在する。本研究では正則化回帰の手法の1つであるElastic net回帰 (Zou and Hastie, 2005) を用いて、相関が強い変数を同時に扱うことを可能にした。

Elastic net回帰は回帰係数を推定する際に係数のL1ノルムとL2ノルムを用いて正則化する手法である。L1ノルムのみを正則化項としたLasso回帰では、被説明変数との関連の小さい説明変数の係数がゼロになることで変数選択ができるものの、相関が強い説明変数の組が存在する場合には、どれか1つだけが選択され、他は係数がゼロになり選択されないという問題、選択される変数の数に上限があるという問題がある。Elastic net回帰は正則化項としてL1ノルムとL2ノルムの和を用いることでこの問題を解決しており、相

関が強い説明変数の組が存在する場合にはそれらの説明変数の係数の絶対値がほぼ同じになるという特徴（Grouping効果）、サンプル数を超える数の説明変数を選択できるという特徴を持つ。本研究では、前者の特徴を活用し、相関が強い変数を同時に扱った。

Elastic net回帰は、(5)式に示した目的関数を最小化する回帰係数 $\beta = [\beta_0, \beta_1, \dots, \beta_m]$ を求めることにより、与えられたデータに対する最適なモデルを推定する（ y_i , \hat{y}_i はそれぞれ被説明変数の観測値と推定値）。(5)式の第2項が正則化項であり、 λ は正則化項の影響の大きさを制御するパラメータ（ $\lambda > 0$ ）、 α はL1ノルムとL2ノルムの相対的な寄与を調整するパラメータ（ $0 \leq \alpha \leq 1$ ）である。本研究では、10分割交差検証法によって逸脱度の最小値を与えるときの λ , α を用いてモデルを決定した。

$$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^m \left(\frac{1-\alpha}{2} \hat{\beta}_j^2 + \alpha |\hat{\beta}_j| \right) \quad (5)$$

Elastic net回帰は回帰係数の信頼区間を直接計算できないため、回帰係数の95%信頼区間の算出には、Elastic net回帰と同様、回帰係数の大きさを用いた正則化回帰手法であるLasso回帰を用いた先行研究（Caster, 2007；林ら, 2017）に倣いブートストラップ法を採用した。重複を許したりサンプリングにより全データと同数のブートストラップ標本を作成し、Elastic net回帰による回帰係数の推定を、予備的検討から1,000回繰り返して、各説明変数についてパーセントイル法により95%信頼区間を算出した。95%信頼区間がゼロを含まない場合に、その説明変数は被説明変数を説明するものとして選択した。ポアソン重力モデルによる回帰係数の推定の際は、説明変数を対数変換したのち、標準化したものを用いた。説明変数には値がゼロのものも含まれていたため、

全ての値に1を足してから対数変換を行なった。

統計解析にはR version 4.0.2を使用し、Elastic net回帰、ブートストラップ法によるリサンプリングにはそれぞれglmnet version 4.1.3, boot version 1.3.28を用いた。

5 結果

使用した変数の移動元、移動先別の基本統計量と変数間の相関（対数変換後のPearson積率相関係数）をそれぞれ表-3, 表-4に示す。i_総人口とi_図書館数は絶対値で0.8以上の強い相関がみられた。

Elastic net回帰により選択された説明変数の回帰係数の95%信頼区間を図-2に示す。すべての変数を同等に扱い、変数の選択に制約を課していなかったが、重力モデルで使用される距離、移動元の総人口、移動先の総人口は、総移動人口と関連のあるものとして選択された。係数の符号も距離は負、総人口は正と、重力モデルで使用されているものと一致した。

社会経済変数については、移動元の変数ではi_公民館数, i_図書館数, i_一般病院数, i_年降水量, i_年合計日照時間は係数の符号が正, i_中学校数, i_第2次産業就業者比率, i_年間最高気温は係数の符号が負であった。また、移動先の変数ではj_総面積, j_中学校数, j_年間最低気温は係数の符号が正, j_課税対象所得, j_第1次産業就業者比率, j_第2次産業就業者比率, j_公民館数, j_年降水量, j_年最深積雪は係数の符号が負であった。

6 考察

本研究は、大都市から地方への人口移動に関連する社会・経済的な要因の市区町村単位でのポアソン重力モデルを用いた分析に、Elastic net回帰を適用した変数選択を試みた。その結果、重力モデルの基本変数である距離、移動元の総人口、移

表-3. 使用した変数の基本統計量

a) 移動元 (n=43)

	平均	標準偏差	中央値	第1四分位数	第3四分位数
i_総人口	855,115	736,842	717,082	330,888	1,027,020
i_可住地面積 1 km2当たり人口密度	10,897	6,479	11,943	4,823	16,554
i_総面積(北方地域及び竹島を除く)	29,371	39,977	5,805	1,930	41,390
i_可住地面積割合	84	25	100	74	100
i_課税対象所得(納税義務者1人当たり)	4,188	1,490	3,619	3,352	4,181
i_幼稚園数	8	3	8	6	9
i_小学校数	11	2	11	9	12
i_中学校数	6	3	6	5	7
i_高等学校数	4	5	3	3	4
i_第1次産業就業者比率(対就業者)	1	1	0	0	1
i_第2次産業就業者比率(対就業者)	17	5	17	13	20
i_第3次産業就業者比率(対就業者)	72	4	72	70	74
i_完全失業率	4	1	4	4	4
i_公民館数	1	3	0	0	1
i_図書館数	2	2	2	1	3
i_一般病院数	5	4	5	3	7
i_老人福祉施設数	2	2	2	1	3
i_保育所等数	16	3	16	13	18
i_都道府県基準地価	393,034	421,835	287,100	120,140	498,758
i_大学(本部)数	1	4	1	0	1
i_年降水量	1,517	221	1,515	1,432	1,565
i_年間最高気温	293	1	293	293	293
i_年間最低気温	285	1	285	285	286
i_年間平均気温	289	1	289	289	289
i_年最深積雪	5	14	3	2	4
i_年合計日照時間	1,888	98	1,866	1,844	1,905

b) 移動先 (n=704)

	平均	標準偏差	中央値	第1四分位数	第3四分位数
j_総人口	54,637	79,482	29,677	13,704	55,550
j_可住地面積 1 km2当たり人口密度	623	706	440	247	734
j_総面積(北方地域及び竹島を除く)	33,461	29,273	24,034	11,983	46,922
j_可住地面積割合	40	23	34	21	51
j_課税対象所得(納税義務者1人当たり)	2,684	338	2,642	2,487	2,831
j_幼稚園数	11	12	9	5	15
j_小学校数	34	25	27	19	40
j_中学校数	18	17	13	9	20
j_高等学校数	7	8	6	4	8
j_第1次産業就業者比率(対就業者)	12	9	9	5	17
j_第2次産業就業者比率(対就業者)	24	8	25	18	30
j_第3次産業就業者比率(対就業者)	62	8	61	56	67
j_完全失業率	4	1	4	3	5
j_公民館数	34	59	18	5	39
j_図書館数	6	6	4	2	7
j_一般病院数	9	7	7	4	11
j_老人福祉施設数	8	9	6	3	11
j_保育所等数	30	14	28	20	37
j_都道府県基準地価	21,144	16,141	17,318	10,665	27,562
j_大学(本部)数	0	1	0	0	0
j_年降水量	1,631	520	1,527	1,213	2,023
j_年間最高気温	291	4	292	289	294
j_年間最低気温	282	5	283	279	285
j_年間平均気温	286	4	287	284	289
j_年最深積雪	30	40	14	0	45
j_年合計日照時間	1,800	203	1,814	1,620	1,955

表-4. Elastic net回帰により選択された説明変数間の相関係数

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)	(P)	(Q)	(R)	(S)	(T)	
(A) 距離	1.000																				
(B) i_総人口	0.018	1.000																			
(C) i_中学校数	-0.106	-0.263	1.000																		
(D) i_第2次産業就業者比率	-0.064	0.475	0.031	1.000																	
(E) i_公民館数	-0.109	0.034	0.117	0.109	1.000																
(F) i_図書館数	0.009	-0.837	0.268	-0.357	-0.229	1.000															
(G) i_一般病院数	-0.065	0.060	0.457	0.033	0.106	-0.184	1.000														
(H) i_年降水量	-0.005	-0.168	0.077	0.142	-0.018	0.207	-0.329	1.000													
(I) i_年間最高気温	-0.023	-0.113	0.026	0.214	0.004	0.242	-0.201	0.554	1.000												
(J) i_年合計日照時間	-0.035	0.278	0.004	0.539	-0.027	-0.172	-0.025	0.167	0.604	1.000											
(K) j_総人口	0.324	-0.407	0.057	-0.102	-0.054	0.399	-0.175	0.182	0.221	-0.018	1.000										
(L) j_総面積	0.076	-0.139	0.014	-0.081	-0.032	0.121	-0.022	-0.022	-0.075	-0.112	0.391	1.000									
(M) j_課税対象所得	-0.045	-0.273	0.032	-0.091	-0.115	0.269	-0.114	0.035	0.011	-0.051	0.492	-0.007	1.000								
(N) j_中学校数	-0.075	0.236	-0.024	0.024	0.025	-0.250	0.188	-0.180	-0.210	-0.035	-0.576	0.227	-0.459	1.000							
(O) j_第1次産業就業者比率	-0.234	0.269	-0.042	0.039	0.068	-0.279	0.151	-0.144	-0.188	-0.027	-0.626	0.160	-0.544	0.605	1.000						
(P) j_第2次産業就業者比率	-0.431	0.084	-0.005	0.123	0.055	-0.042	-0.101	0.043	0.071	0.049	-0.007	0.060	0.139	-0.135	-0.034	1.000					
(Q) j_公民館数	-0.218	0.117	0.006	0.070	0.038	-0.100	0.044	-0.019	0.059	0.067	-0.211	0.135	-0.214	0.242	0.300	0.155	1.000				
(R) j_年降水量	0.141	0.056	0.040	0.087	0.027	-0.022	0.029	0.165	0.240	0.159	-0.080	-0.139	-0.231	0.090	-0.079	-0.169	0.057	1.000			
(S) j_年間最低気温	0.308	-0.037	0.052	0.079	0.033	0.065	0.019	0.165	0.352	0.211	0.138	-0.457	-0.049	-0.141	-0.276	-0.244	-0.107	0.568	1.000		
(T) j_年最深積雪	-0.146	-0.014	-0.060	-0.086	-0.063	0.007	-0.101	-0.120	-0.310	-0.231	-0.056	0.357	0.000	0.039	0.098	0.159	0.095	-0.298	-0.722	1.000	

対数変換後の相関係数。絶対値が大きいものはセルの背景色を濃くしている。

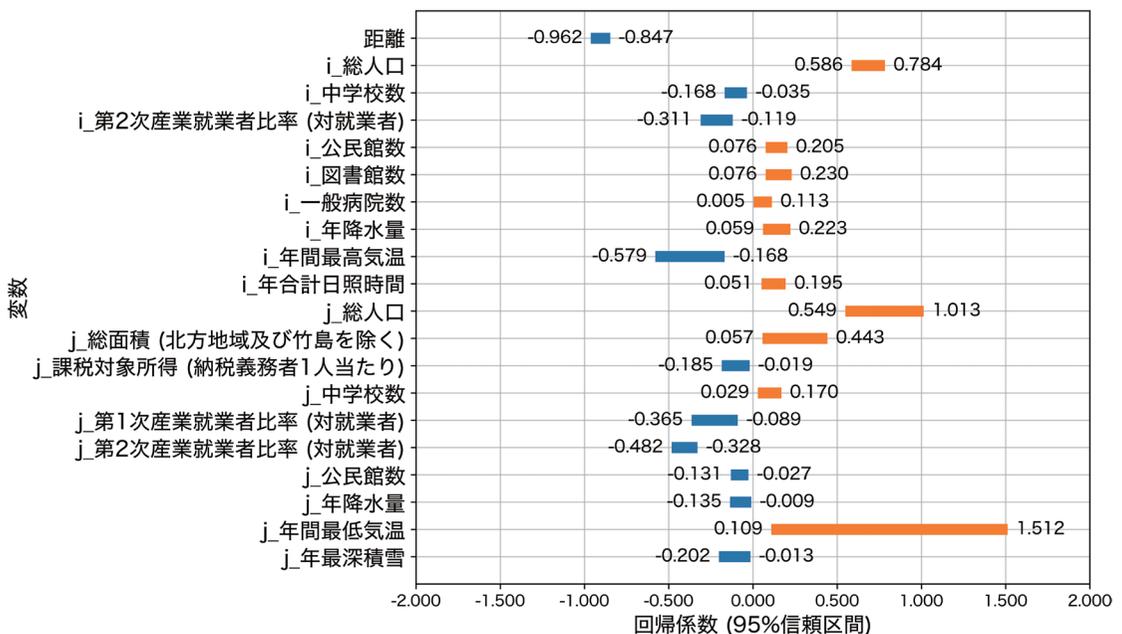


図-2. Elastic net回帰により選択された変数の回帰係数。人口移動量と正の関連を持つものは橙色、負の関連を持つものは青色で描画した。棒グラフの左右の数字はそれぞれ回帰係数の95%信頼区間の下限と上限。

動先の総人口を含む20の変数が移動人口と関連のあるものとして選択された。これらの説明変数には相関が強い変数の組も存在したが、これらを1つの回帰式で扱うことができたのはElastic net回帰の利点である。本研究ではElastic net回帰ですべての変数を同等に扱い、変数の選択に制約を課さなかったが、重力モデルの基本変数が回帰モデルに必ず組み込まれるように、基本変数の係数

を正則化項に組み込まず、係数が縮小しないようにするといった応用も可能である。以下、選択された変数に関して、人口移動との関連について考察する。

重力モデルの基本変数について、係数の符号は先行研究と一致しており、大都市から地方への人口移動では様々な変数を入れて調整しても、距離と総人口が人口移動と関連することが改めて確認

された。基本変数の中でも距離の係数の絶対値は移動元の総人口のものよりも大きく、地方の市町村による人口誘致の視点では、遠くの大都市よりも近くの大都市をターゲットにする方が有効である可能性がある。

人口移動量と正の関連があった移動元の変数の中で、 i 公民館数に関して、1990年以降の規制緩和・地方分権等の行政改革が本格化し、公民館は統廃合やコミュニティセンター等への用途変更が行われているが（文野，2017）、この状況は大都市に顕著である（佐藤，2010）。社会教育に力を入れると定住を促進できるという研究もあり（原，2016）、改革に積極的な大都市ほど社会教育法に基づく公民館数が少ないため、公民館が多いほど転出者が多いという結果になったと考えられる。今後は公民館のみならずコミュニティセンター等の施設を含めた解析が求められる。また、 i 図書館数は、人口100万人当たりの図書館数が多い地域の方が若者の転出が多いという先行研究の結果と一致した（後河，2019）。図書館が整備されていることは社会教育に対する公的投資が多いと考えられ、教育投資が多い都道府県ほど若者は転出しやすいこと（後河，2019）、教育・子育てのまちという地域イメージのある地域ほど、教育・子育てに関心のある人々が住んでいるため、より良い環境を求めて移住していることから（荒川，2018）、転出者が多いと考えられる。 i 一般病院数について、医療・福祉環境が充実しているほど20-24歳の純転入率が低いという研究や（林・林，2017）、健康・医療のまちという地域イメージのある地域ほど、転出者が多いという研究の結果と変数の関連の向きが一致した（荒川，2018）。 i 年降水量について、10~24歳の人口移動について、湿潤な地域ほど純移動率が少ないという先行研究の結果と変数の関連の向きが一致した（伊藤，2006b）。 i 年合計日照時間について、先行研究の結果と変数の関連の向きが異なっていた。都市化の進んでいる地域ほど、日照時間は短

くなる傾向にあるため（Landsberg，1981，258頁）、日照時間が長い、つまり都市化の進んでいない地域からの転出者が多いものと考えられる。

人口移動量と負の関連があった移動元の変数の中で、 i 中学校数に関して、中学校に通う子どもを育てている、もしくは今後子育てをする者が比較的多いと考えられる20歳から49歳が総移動人口の約70%を占めており（表-2）、中学校が多い地域から移住する者は少ないと考えられる。また、 i 第2次産業就業者比率に関して、工場は一度立地すれば移転が少ないため（當麻，2016）、第2次産業就業者は移動が難しく、第2次産業就業者比率の高い大都市からは移住する者が少ないと考えられる。 i 年間最高気温に関しては、気候が寒冷であるほど転出が多いという先行研究の結果と変数の関連の向きが一致した（伊藤，2003）。

人口移動量と正の関連があった移動先の変数の中で、 j 総面積は、松本・松下（2002）の先行研究と変数の関連の向きが一致し、居住可能な地域が広い場所に移住する者が多いと考えられる。また、 j 中学校数に関しては、中学生を含む10~19歳が解析対象の人口移動に占める割合は5.6%と中学生自体の割合は小さいものの、前述の通り、今後中学生を育てる可能性のある層の割合は大きく、教育環境を重視しているものと考えられる。 j 年間最低気温に関しては、温暖な地域であるほど転入者が多いという先行研究の結果と変数の関連の向きが一致した（伊藤，2006b）。

人口移動量と負の関連があった移動先の変数の中で、 j 課税対象所得に関しては、所得の高い地域への人口移動が多いという多くの先行研究と変数の関連の向きが異なっていた（伊藤，2003；田村，2017）。 j 課税対象所得の係数の符号が負であった理由として、大都市から地方への人口移動を対象とした場合、所得の高い地域ではなく、生活コストの低いところへ移住することが多い可能性が考えられる。平成26年全国消費実態調査

(総務省)での消費支出と平成26年の課税対象所得(納税義務者1人当たり)のPearsonの積率相関係数は0.551, このうち食料, 住居, 被服及び履物, 教育, 教養娯楽との相関係数はそれぞれ0.627, 0.716, 0.719, 0.739, 0.753であり, 消費支出と課税対象所得(納税義務者1人当たり)の相関は高く, 消費支出の多寡が大都市から地方への人口移動と関連する可能性がある。さらに, 地方への移住に伴う所得の減少は許容されやすいことも理由の1つと考えられる。一般社団法人移住・交流推進機構(2017)によると, 地方移住時の所得に関して, どの世帯年収層も「変化なしが好ましい」と答える者が最も多かった(39.2%)が, 現在の世帯年収が500万円以上700万円未満の層では「5%程減少まで許容できる」が, 700万円以上の層では「10%程減少まで許容できる」という回答がそれに続いて多かった。j_第1次産業就業者比率, j_第2次産業就業者比率に関して, 第1次産業就業者は農林水産業の性質上, 特定の場所でのみ生産活動が可能であり, 第2次産業についても前述のとおり, 第1次, 第2次産業就業者は第3次産業就業者と比較して移動が困難である。そのため, 第1次産業就業者比率, 第2次産業就業者比率の高い地方に移住する者は少ないと考えられる。j_公民館数に関しては, 文教施設ストックが高い地域ほど転入者が少ないという先行研究の結果と変数の関連の向きが一致した(張ら, 2016)。1985年以前に設立された公民館は全体の約6割であり, 1981年の耐震基準改正前に建築された公民館のうち, 耐震性なし及び耐震診断未実施の公民館は23.9%である(文部科学省生涯学習政策局社会教育課, 2017)。そのため, 老朽化した公民館が地域のマイナスの魅力となってしまう可能性がある。また, 公民館の耐震化等未実施は, 財政状況の悪さを反映しているとも考えられ, 地方公共団体の財政状況も含めた検証が今後必要である。j_年降水量に関して, 前述した通り, 10~24歳の人口移動で, 湿潤な地域は

ど純移動率が少ないという先行研究の結果と変数の関連の向きが一致した(伊藤, 2006b)。また, j_年最深積雪に関して, 積雪の少ない地域への転入者が多いという先行研究の結果と変数の関連の向きが一致した(伊藤, 2003)。

7 まとめ

本研究では, 日本国内の市区町村間の人口移動について, 政令指定都市と東京都特別区部から, 大都市圏に属する周辺市町村を除く市町村への人口移動を対象に分析した。これまでの多くの研究では, 都市の属性を無視し, 地方から大都市への集中移動, 大都市から地方への分散移動をまとめて, かつ都道府県単位で分析しているという問題があった。本研究は対象を大都市から地方への分散移動に絞り, 市区町村単位で分析したことで, 人口移動及び移住施策の実態により即した解析ができた。また, ポアソン重力モデルにElastic net回帰を適用して関連の強い複数の社会経済変数を同時に扱い, 研究者による変数選択の恣意性を減らした上で, 各変数と大都市から地方への人口移動にどのように関連しているかを明らかにした。

本研究の限界は, 解析対象の人口移動・市区町村がある程度以上の規模に限定されていることである。本研究では, データの制約上, 移動元から移動先への人口移動数が10人未満の市区町村の組を除外した。そのため, 解析対象の人口移動が限定されており, 小規模な市区町村ほど欠損となりやすいセレクションバイアスが生じている可能性がある。また, 今後の課題は, 大きく以下の3点である。1点目は, 年齢や子供の有無などの属性別に分析を行うことである。本研究では全年齢をひとまとめにして解析したが, 年齢は人口移動に関連する要因とされており(Cadwallader, 1996, 334-337頁), また, 本研究の解析結果としても中学校数や図書館数といった, 子育て中あるいは今後子育てをする世代が特に注目すると考えられ

る教育に関連する変数が選択された。年齢や子供の有無などの属性別に分析することで、属性ごとに人口移動量と関連する要因を抽出可能であり、移動先である地方公共団体が移住誘致施策を立案する上でも有用な情報を提供できると考えられる。2点目は、複数年次のデータを用いたパネルデータ分析を行うことである。本研究は横断的な分析を実施しており、説明変数と人口移動量との因果関係は不明である。本研究で抽出された変数を含むパネルデータ分析をすることで、完全な因果の証明にはならないものの、大都市から地方への人口移動に影響を与える要因として補強をすることができる。3点目に、分析手法の課題として、重力モデルへのサンプルの階層性に対応した回帰係数推定手法の導入が挙げられる。(ポアソン)重力モデルに使用される各サンプルは移動元と移動先の2方向に対してネストされており、本来ならばサンプル間の相関を考慮したマルチレベル分析が必要である。筆者の知る限りでこの点に対応した人口移動分野の研究はこれまでになく、なおかつ本研究で採用した手法も、サンプル間の相関を考慮できていない。サンプルが2方向に対してネストされている、重力モデルへのマルチレベル分析導入は今後の課題である。

謝辞

本研究は、2020年6月に経済産業研究所ディスカッション・ペーパー「大都市から地方への移住における社会経済的要因の影響—Elastic net回帰を用いたポアソン重力モデルによる分析—」として公開した原稿を修正し、再構成したものである。本研究の原案に対して、経済産業研究所ディスカッション・ペーパー検討会の方々から多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。

注

(1) 国土地理院測量計算サイト：[https://vldb.](https://vldb.gsi.go.jp/sokuchi/surveycalc/main.html)

[gsi.go.jp/sokuchi/surveycalc/main.html](https://vldb.gsi.go.jp/sokuchi/surveycalc/main.html)

- (2) 例えば、先行研究で県民所得の変数が使用されている場合、e-Statには市民所得のデータが掲載されていないため、課税対象所得を用いた。
- (3) 政府統計では10人未満（ゼロを含む）のデータは実際の数値が公開されていないため除外した。

参考文献

- 青木俊明・稲村肇 (1997) 「人口移動研究の展開と今後の展望」、『土木計画学研究・論文集』, 14, pp.213-224.
- 荒川清晟 (2018) 「大都市から地方への移住における地域イメージの影響に関する研究」、『東京大学大学院情報学環紀要 情報学研究』, 95, pp.57-71.
- 伊藤薫 (2001) 「戦後日本の人口移動に対する所得格差説の説明力と今後の課題」、『地域と社会』, (4), pp.9-38.
- (2003) 「バブル経済期の男女・年齢別人口移動—1990年国勢調査人口移動集計結果を利用して—」、『地域学研究』, 33(3), pp.85-102.
- (2006a) 「長距離人口移動に対する社会環境アメニティの作用—1970年から2000年の国勢調査人口移動集計結果を利用して—」、『Review of economics and information studies』, 7 (1-2), pp.21-49.
- (2006b) 「戦後日本の国内長距離人口移動の決定因の変化—純移動率に対する所得・気候の作用—」、『地域学研究』, 36(1), pp.85-99.
- (2011) 「高齢者の長距離人口移動の決定因の変化—1960年国勢調査から2000年—」、『地域学研究』, 41(1), pp.179-194.
- 一般社団法人 移住・交流推進機構 (JOIN) (2017) 「若者の移住に関する調査」, <<https://www.iju->

- join.jp/material/files/group/1/JOIN_report_201710.pdf), 参照2022年9月4日.
- 後河正浩(2019)「若者の地域間移動の傾向と要因: 都道府県データでみる大学進学・初職就職時の地域間移動」, 『京都産業大学経済学レビュー』, 6, pp.1-42.
- 江崎雄治・荒井良雄・川口太郎(2000)「地方圏出身者の還流移動 長野県および宮崎県出身者の事例」, 『人文地理』, 52(2), pp.190-203.
- 尾崎雅彦(2013)「地域間人口移動に対する地域別政策プライオリティの影響—テキストマイニングによる政策プライオリティの定量的計測—」, 『RIETI Policy Discussion Paper Series』, 13-J-072.
- 閣議決定(2015)「まち・ひと・しごと創生基本方針2015—ローカル・アベノミクスの実現に向けて—」, 2015年6月30日決定, <<https://www.chisou.go.jp/sousei/info/pdf/20150630siryou3.pdf>>, 参照2022年9月4日.
- (2019)「第2期『まち・ひと・しごと創生総合戦略』」, 2019年12月20日決定, <<https://www.chisou.go.jp/sousei/info/pdf/r1-12-20-senryaku.pdf>>, 参照2022年9月4日.
- (2020)「第2期『まち・ひと・しごと創生総合戦略』(2020改訂版)」, 2020年12月21日決定, <<https://www.chisou.go.jp/sousei/info/pdf/r02-12-21-senryaku2020.pdf>>, 参照2022年9月4日.
- 佐藤智子(2010)「中核市における教育行政組織機構の再編と公民館の位置づけ: 愛知県豊田市を事例として」, 『東京大学大学院教育学研究科紀要』, 49, pp.149-159.
- 総務省統計局(2013)「平成25年住宅・土地統計調査用語の解説《地域》」, <<http://www.stat.go.jp/data/jyutaku/2013/1-1.htm>>, 参照2022年9月4日.
- 田村一軌(2017)「大学進学にとまなう都道府県間人口移動の定量分析—修正重力モデルによる分析—」, 『AGI Working Paper Series』, 3, pp.1-17.
- 張峻屹・瀬谷創・兼重仁・力石真(2016)「都道府県間人口移動の影響要因の経年的分析—空間的文脈依存性をもつ選択モデルに基づく分析—」, 『地域学研究』, 71(3), pp.118-132.
- 當麻雅章(2016)「人口移動要因としての地域アメニティ近接性」, 『大阪大学経済学』, 66(3), pp.1-23.
- 日本人口学会編(2002)『人口大辞典』培風館.
- 林裕志・平松達雄・小出大介・田中勝弥・大江和彦(2017)「電子カルテデータベースからのLASSOロジスティック回帰による医薬品副作用シグナルの検出: ケース・コントロール研究」, 『薬剤疫学』, 21(2), pp.51-62.
- 林勇貴・林宜嗣(2017)「地域競争力と地域間人口移動」, 『経済学論究』, 71(3), pp.59-81.
- 原俊彦(2016)「人口減少に対応した社会教育の在り方—北海道・道南ブロックの事例—」, 『札幌市立大学研究論文集』, 10(1), pp.67-76.
- 文野克成(2017)「公民館に関する研究: その来歴と現代的役割に関する考察」, 博士論文, 法政大学, pp.1-132.
- 松本茂・松下敬一郎(2002)「住宅環境が居住地の選定に与える影響: Poisson Gravity Modelによるティプー仮説の検証」, 『關西大學経済論集』, 51(4), pp.455-471.
- 文部科学省生涯学習政策局社会教育課(2017)「公民館の現状と課題」, <https://www.soumu.go.jp/main_content/000513104.pdf>, 参照2022年9月4日.
- 矢野桂司(1991)「空間的相互作用モデルの精緻化に関する研究」, 『都市計画論文集』, 26, pp.517-522.
- Cadwallader, M. (1996), *Urban Geography: An Analytical Approach*, New Jersey, Prentice-hall.
- Caster, O. (2007), *Mining the WHO Drug Safety*

- Database Using Lasso Logistic Regression, *UUDM Project Report*, 16.
- Flowerdew, R. and Aitkin, M. (1982), A Method of Fitting the Gravity Model Based on the Poisson Distribution, *Journal of Regional Science*, 22(2), pp.191-202.
- Greenwood, M.J. and Hunt, G. (2003), The Early History of Migration Research, *International Regional Science Review*, 26 (1), pp.3-37.
- Landsberg, H.E. (1981) *The Urban Climate*, Academic Press, New York.
- Ravenstein, E.G. (1885), The laws of migration, *Journal of the Statistical Society of London*, 48(2), pp.167-235.
- Zou, H. and Hastie, T. (2005), Regularization and Variable Selection via the Elastic Net, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 67(2), pp.301-320.