

マイクロジオデータを用いた日本全国の家賃形成メカニズムの解明

主査 秋山 祐樹^{*1}

委員 小川 芳樹^{*2}, 秦 桜蘭^{*3}

本稿は日本全国の家賃形成メカニズムの実態を明らかにする主査らによる先行研究の発展的内容として実施した研究成果を報告するものである。まず日本全国の住宅用途の建物約 6,000 万棟の分布を収録した建物マイクロジオデータを整備した。またそれらに約 400 種類に及ぶ家賃形成要素と家賃に関する情報を付与し、日本全国約 1,100 万件の賃貸アパートおよび賃貸マンションの価値を評価出来るデータベースを構築した。さらに同データベースをクラスタリングし、各クラスタに対してスパースモデリング (LASSO 法) を行うことで、クラスタ毎に家賃に対して正あるいは負に寄与する家賃形成要素を明らかにする技術を開発した。その結果、全国的な傾向として物件周辺の人口動態や物件そのものの機能の充実度は家賃形成に正に寄与した一方で、災害リスクは家賃形成にあまり影響を与えないことが明らかになった。

キーワード：1) 家賃, 2) 賃貸物件, 3) マイクロジオデータ, 4) 住宅の性能, 5) 立地環境,
6) 災害リスク, 7) クラスタリング, 8) スパースモデリング, 9) LASSO 法, 10) 居住者

ELUCIDATION OF RENT FORMATION MECHANISM USING MICRO GEODATA

Ch. Yuki Akiyama

Mem. Yoshiki Ogawa, Yinglan Qin

This paper reports the research result carried out as the development research of our previous research to clarify the rent formation mechanism in Japan. First, we developed building micro geodata that recorded the spatial distribution of about 60 million residential buildings throughout Japan. In addition, we developed the database that can evaluate the value of about 11 million rental apartments and condominiums in Japan by giving them information on about 400 types of rent formation elements. Moreover, by clustering the database and performing sparse modeling on each cluster, we developed a method that clarifies rent formation factors that contribute positively or negatively to rent. As a result, as a nationwide trend, population dynamics around the property and the fullness of function of property contributed positively to rent formation. On the other hand, disaster risk does not affect rent formation.

1. はじめに

1.1 主査らによるこれまでの研究

住総研「研究・実践選奨」における受賞論文^{*1)} (以下、主査先行研究) は、日本全国の家賃形成メカニズムの実態把握を目的に、日本全国の住宅用途の建物 1 棟 1 棟の分布を収録した空間的にミクロなデータ (以下、マイクロジオデータ (MGD)) を整備するとともに、それらに様々な家賃形成要素と、郵便番号領域の単位ごとに算定された最低保証家賃に関する情報を付与することで、日本全国の住宅の価値を評価出来るデータベースの構築を実現した。さらに同データを郵便番号領域単位に集計化した後にクラスタリングし、さらに各クラスタに対してスパースモデリングを行うことで、クラスタ毎に最低保証家

賃に対して正あるいは負に寄与する家賃形成要素を明らかにする技術を開発した。その結果、地域によっては住宅の性能や災害リスクが考慮されておらず、すなわち家賃への影響が小さく、住まい手にとっての住宅の価値と市場価値との間に乖離が存在することが確認された。

1.2 主査先行研究における課題

上述する成果が得られた一方で、主査先行研究には、数多くの課題も残された。まずこれまでの研究では、収集できたデータの制約上、戸建て住宅と共同住宅の区別や、住宅の詳細な性能 (間取りや各種附帯設備など) の情報は利用できなかった。また例えば景観や気象条件といったさらに多様な家賃形成要素の収集も課題となった。

^{*1} 東京大学空間情報科学研究センター 助教 博士 (環境学) ^{*2} 東京大学生産技術研究所 特任助教 博士 (環境学) ^{*3} 東京大学大学院新領域創成科学研究科 修士課程

そして MGD の強みを最大限に発揮するためには、非集計（建物単位）の家賃情報を用いた分析が理想的である。主査先行研究では、家賃に関するデータの集計単位が郵便番号単位という制約があった。そのため日本全国の住宅用途の建物 1 棟 1 棟の分布を把握できる MGD を整備したものの、最終的には郵便番号単位で集計化しており、MGD の高精細性を十分に発揮した分析には至らなかった。

主査先行研究は、学術的に斬新かつ挑戦的で、社会的要請も高い内容であると評価された一方で、以上のように数多くの課題が残った。そのため以上の課題の克服を試みることで、主査先行研究をより公益性の高い内容に深化・発展させ、我が国の今後の住宅政策の立案と良好な住環境の実現を支援できる内容にする必要がある。

1.3 関連する既存研究

様々な空間データや各種統計を活用した家賃形成のメカニズム解明に関する研究は、国内外においてこれまでも見られるが^{文²⁾、文³⁾}、いずれも対象地域が限定的であり、本研究のような広域性（日本全国）と詳細性（建物単位）を両立させながら家賃形成メカニズムを把握しようとした研究は前例がない。さらに、家賃形成の要因の解明に潜在的な災害リスクを全国規模で導入した研究は数少なく、全国スケールで災害リスクを考慮した家賃形成メカニズムの解明に取り組んだ事例は、主査先行研究のみである。以上より、主査先行研究の発展的内容となる本研究の学術的価値と新規性は十分に高いと言える。

1.4 本研究の目的

主査先行研究の課題および、その他の既存研究の状況を鑑み、本研究の目的は以上に挙げた課題を克服し、MGD の利点をこれまで以上に活かすことができる空間データの整備を達成するとともに、同データを活用した分析を実施することで、これまでの研究に引き続き、我が国の将来の住宅政策の支援につながる研究成果の実現を目指す。具体的には主査先行研究でも実施した全国的な住宅政策の立案を支援するための基礎的な空間データをより充実したデータに拡充し、我が国の住宅価値形成の現状に課題があることに警鐘を鳴らすとともに、地域ごとに適正な家賃形成を目指すために是正すべき要因を明らかにすることで、これまでの研究よりもより適正な家賃形成の検討につながる成果を得ることを目的とする。なお本研究で分析対象とする物件は、ストック数が多く周辺の環境の影響を受けやすいと考えられる賃貸アパートおよび賃貸マンションとする。

2. マイクロジオデータの開発

図 2-1 に本研究におけるマイクロジオデータの開発方法の流れを示す。研究対象地域は日本全国とする。

2.1 建物マイクロジオデータの整備

まず主査先行研究と同様に、日本全国の住宅の分布を把握できるデータの整備を行った。最新の住宅地図(2016 年)を用いて、日本全国約 6,000 万棟の建物の位置情報（建物ポリゴンの重心経緯度）、面積、階数等を属性として持つポイントデータである「建物マイクロジオデータ（以下「建物 MGD」）」を整備した。

そして以下、2.2 および 2.3 の処理を通して、以下①～④の家賃形成要素（説明変数）と、⑤家賃情報（目的変数）が建物 MGD に付与される。

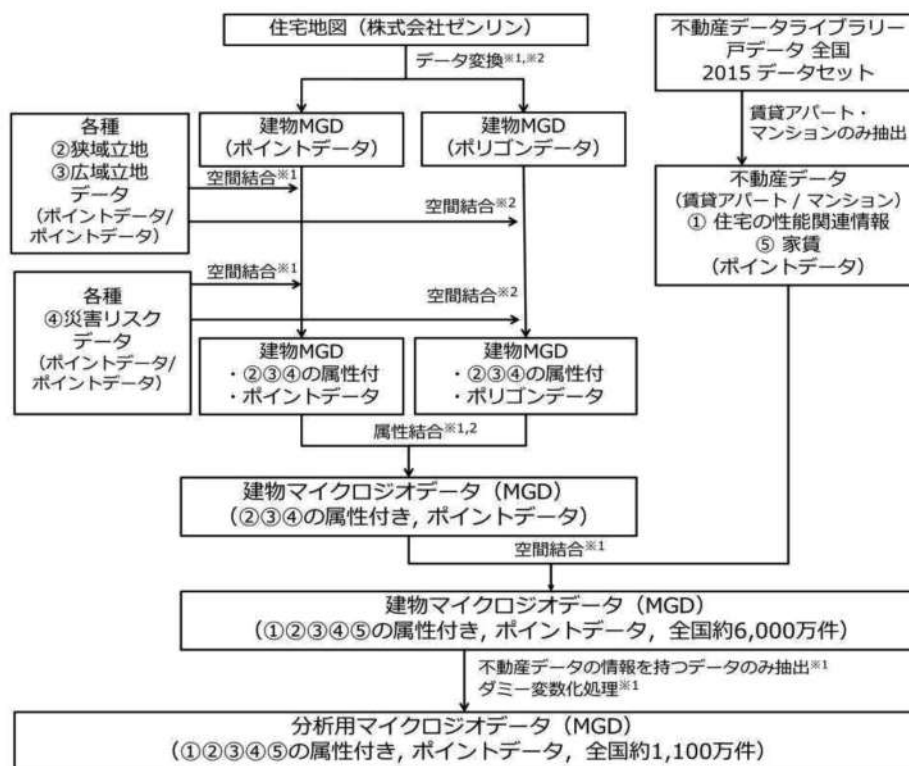
- ① 住宅の性能関連情報:住宅のそのものの性能に関する情報（間取り、面積、所在階、各種付帯設備の有無など）。
- ② 狭域立地関連情報:住宅近隣の立地環境に関する情報。具体的には住宅近隣の利便性に関連する情報（各種生活利便施設や交通機関へのアクセス性など）。
- ③ 広域立地関連情報:より広域の立地環境に関する情報。具体的には住宅が立地する地域の広域的な属性（地域の知名度、近隣大都市へのアクセスの良さ、用途指定など）。
- ④ 災害リスク関連情報:住宅が立地する場所の災害リスクと避難可能性に関する情報。
- ⑤ 家賃:部屋単位の家賃。

2.2 家賃形成要素となる情報の付与

建物 MGD に収録された各建物に、家賃形成要素のうち②狭域立地関連情報、③広域立地関連情報、④災害リスク関連情報となる情報の付与を行った。家賃形成要素がポイントデータの場合は、各建物から最近隣の家賃形成要素までの距離および、徒歩圏（500m 圏内）の家賃形成要素の件数という形で集計を行った。また家賃形成要素がポリゴンデータの場合は、各建物が家賃形成要素のポリゴンデータに含まれるか否か、また最近隣のポリゴンデータまでの距離を計算した。

2.3 家賃情報の付与

続いて⑤家賃情報を建物 MGD に与えた。本研究では東京大学空間情報科学研究センターの共同研究システムに格納された「不動産データライブラリー 戸データ 全国 2013-2017 データセット」(アットホーム株式会社。以下「不動産データ」)を用いて家賃情報の収集を行った。同データはアットホーム株式会社が 2013 年から 2017 年に扱った賃貸物件（賃貸戸建住宅、賃貸共同住宅）の家賃に加えて、①住宅の性能関連情報も保有している。なお本研究では上述したように、賃貸アパートおよび賃貸マンションを用いるものとする。また他の家賃形成要素の作成年との整合性を考慮し、本研究では 2015 年に公開されていた不動産データを用いるものとする。ちなみに 2015 年の不動産データの場合、賃貸アパートは全国で約 450 万件、賃貸マンションは全国で約 650 万件であった。



※1 Rubyによる専用の処理プログラムを開発した。

※2 PythonとArcGISを組み合わせた専用の処理プログラムを開発した。

図 2-1 本研究におけるマイクロジオデータ開発の流れ

また本研究で扱う不動産データは、各物件の住所だけでなく、物件が立地する地点の座標（経度緯度）の情報も保有しているため、各物件の立地をピンポイントに把握できる（図 2-2）。そのため不動産データと、建物 MGD に含まれる建物のうち、最近隣の建物情報と空間結合することにより、建物 MGD の各建物が持つ②～④の情報を不動産データに与えた。すなわちこの処理は両データを建物単位でピンポイントに突合処理したことに等しいといえる。なお 1 つの建物に複数の不動産データが存在する場合は、建物単位ではなく物件単位でデータを整備した。すなわちこのような場合、同一地点に複数の不動産データが分布する形となる。また家賃データは①住宅の性能関連情報も保有するため、以上の処理を行うことで①の情報との統合も同時に達成することができる。なお本研究で利用した家賃データが持つ①の情報は、主査先行研究と比較しても、その種類が大幅に豊富になっており、家賃形成要素をより充実したものとすることができた。

2.4 分析用マイクロジオデータの整備

以上の処理により、日本全国の建物 1 棟 1 棟に多様な家賃形成要素と家賃の情報を与えた新しい建物 MGD が実現した^{注 2)}。また本研究で整備した建物 MGD は物件単位（部屋単位）でデータを扱うことができるため、主査先行研究における郵便番号単位と比べても、大幅に高精細な空間単位で分析することが可能になった。



図 2-2 座標情報付の不動産データ（東京都港区の例）^{注 1)}

続いて 2.3 までは作成した建物 MGD を整理してデータ分析用の MGD に変換する処理を行った。まず 2.3 までは作成した建物 MGD のうち、不動産データを持つデータ、すなわち①住宅の性能関連情報と⑤家賃を持つデータのみを抽出する。これは⑤の情報を持たないデータは目的変数を持たないことを意味しており、分析対象とする意味がないためである。この処理により、賃貸アパートとマンション合わせて、全国約 1,100 万件が分析対象として抽出された。また①～④の家賃形成要素のうち、カテゴリデータについてはダミー変数化を行った。

以上により①～⑤の全ての属性を保有する日本全国の賃貸物件 1 件 1 件単位の分析用 MGD が実現した。表 2-1 に分析用 MGD が保有する属性の一覧を示す。

表 2-1 分析用 MGD が保有する属性の一覧

データの種類		家賃形成要素 (太字斜体は質的変数であるためダミー変数化)	データ 作成年	データ ソース	データ 形式	データ 収録範囲	変数の数
大分類	小分類						
① 住宅の性能 関連情報	①-1. 建物そのものの性能	築年数, 新築中古区分 , 地上階層, 建物構造 , 建物面積, 駐車場有無 , 駐車場料金	2015	アットホーム	ポイント	日本全国	19
	①-2. 物件の性能	所在階, 間取り , バルコニー面積, 各種付帯設備 (141種類) (※1),	2015	アットホーム	ポイント	日本全国	164
② 狭域立地 関連情報	②-1. 生活利便施設	徒歩圏の各種生活利便施設 (23種類) (※2) までの距離, 最近隣の各種生活利便施設 (23種類) までの距離	2015	テレポイント データ (デジタル電話帳)	ポイント	日本全国	55
		最近隣の集客施設までの距離	2014	国土数値情報	ポイント	日本全国	
		最近隣の大規模店舗までの距離, 徒歩圏の大規模店舗数	2018	独自調査	ポイント	日本全国	
		最近隣の市区町村役場までの距離	2015	テレポイント データ	ポイント	日本全国	
		最近隣の商業集積地域までの距離, 最近隣の商業集積地域の店舗数, 徒歩圏の商業集積地域の数, 徒歩圏の商業集積地域の店舗総数	2015	商業集積統計	ポリゴン	日本全国	
	②-2. 交通	最寄駅徒歩(分), バス駅までの時間(分)ダミー , バス駅までの時間(分), 最寄バス停徒歩(分)ダミー , 最寄バス停徒歩(分)	2015	アットホーム	ポイント	日本全国	6
	②-3. 人口	人口, 男性人口, 女性人口, 0~14歳人口, 15~64歳人口, 65歳以上人口, 75歳以上人口, 外国人人口, 世帯人員別世帯数 (1~6人および7人以上), 世帯類型別世帯数 (核家族・高齢者のみ・高齢者単身), 持ち家世帯数, 戸建住宅世帯数, 共同住宅世帯数 (※3)	2015	国勢調査	ポリゴン (500m メッシュ)	日本全国	57
		以上のそれぞれの割合 (※3)					
		平均年齢 (※3)					
		将来推計人口 (2020年・2025年・2030年・2035年・2040年・2045年・2050年) (※3)	2015	国土数値情報	ポリゴン (500m メッシュ)	日本全国	
		以上の将来推計人口の2010年比 (※3)					
	②-4. 生活の安心・安全	最寄りの警察署までの距離, 最寄りの消防署までの距離	2012	国土数値情報	ポイント	日本全国	2
	②-5. 生活環境	最近隣の都市公園までの距離, 最近隣の都市公園の面積, 徒歩圏の都市公園の数, 徒歩圏の都市公園の面積合計	2011	国土数値情報	ポリゴン	日本全国	15
		景観地区指定有無	2014	国土数値情報	ポリゴン	日本全国	
		最近隣の観光資源の種類	2014	国土数値情報	ポイント	日本全国	
③ 広域立地 関連情報	③-1. 立地特性	都市地域指定有無 , 都市地域種別 , 用途地域種別	2011	国土数値情報	ポリゴン	日本全国	18
	③-2. 地域の知名度	最近隣の駅のブランド力, 徒歩圏の駅のブランド力合計 (※4)	2015	独自開発	ポイント	日本全国	3
	③-3. 近隣大都市へのアクセス	最近隣の都道府県庁所在地の主要駅までの距離	2015	テレポイント データ	ポイント	日本全国	1
④ 災害リスク 関連情報	④-1. 被災リスク	浸水想定区域による最大浸水深	2012	国土数値情報	ポリゴン	日本全国	60
		津波浸水想定による最大津波浸水深	2018	国土数値情報	ポリゴン	日本全国	
		東日本大震災による最大津波浸水深	2011	復興支援アーカイブ	ポリゴン	日本全国	
		想定最大津波浸水深 (※6)	2011, 2018	国土数値情報 復興支援アーカイブ	---	日本全国	
		土砂災害警戒区域の現象の種類, 土砂災害警戒区域の指定の種類	2013	国土数値情報	ポリゴン	日本全国 (※5)	
		密集市街地指定有無	2015	国土数値情報	ポリゴン	日本全国	
		低位地帯における浸水深	2015	国土数値情報	ポリゴン	日本全国	
		月別平均降水量, 年平均降水量, 月別平均気温, 年平均気温最高気温, 年平均最低気温, 年平均気温, 月別平均降雪量, 年平均降雪量, 月別平均日照時間, 年平均日照時間	2012	国土数値情報・気象庁	ポリゴン (1kmメッシュ)	日本全国	
		50年超過確率 2 % の計測震度 (最大値・平均値)	2018	防災科学技術研究所	ポリゴン (1kmメッシュ)	日本全国	
	④-2. 避難可能性	最近隣の緊急輸送道路までの距離, 最近隣の緊急輸送道路の種類	2015	国土数値情報	ライン	日本全国	5
		最近隣の避難所までの距離	2012	国土数値情報	ポイント	日本全国	
⑤ 家賃	---	---	2015	アットホーム	ポイント	日本全国	1

変数合計 406

※1 室内洗濯機置場、トランクルーム、追焚機能、エレベータ、バストイレ別、床暖房、ディンブルキーなど各物件の性能に関連する設備

※2 スーパーマーケット、コンビニエンスストア、ドラッグストア、飲食店、総合病院、専門病院、小中学校、郵便局、酒場、金融機関、デパート、ホームセンター、書店、家電量販店、クリーニング店、衣料品店、理美容店、食料品店、ガソリンスタンド、幼稚園・保育園、高等学校、大学

※3 各物件が分布するメッシュとそれに隣接するメッシュが持つ統計情報の合計値より算出した。

※4 各駅の1km圏内の電話帳と住宅地図に収録された企業・店舗・事業所にその駅名が何回出現するかを計算した値。

※5 ただし一部の都道府県では非公開となっている。

※6 「津波浸水想定による最大津波浸水深」と「東日本大震災による最大津波浸水深」の大きい方の値を収録した。

3. 分析用 MGD のクラスタリング

本研究では家賃がどの家賃形成要素の影響を強く受けて決定されるのか、ということを明らかにすることを目的としている。すなわち基本的には表 2-1 に示す家賃形成要素を説明変数、家賃を目的変数とした分析を行い、目的変数である家賃に対して説明力の大きい説明変数を明らかにすればよい。ただし本研究の分析では、分析対象地域が日本全国であり、しかも日本全国約 450 万件の賃貸アパートと約 650 万件の賃貸マンションの情報を含んでいるため、様々なタイプの物件や地域特性を持つ地域が母集団に混在することになる。その上、表 2-1 に示すように説明変数の数も 405 種類と非常に多いため、全国単一のモデルで日本全国の家賃の決定要因を説明することは現実的ではないものと考えられる。

そこで本研究では、家賃形成要素（説明変数）と家賃（目的変数）に基づいて、分析用 MGD を幾つかのクラスター（類型）に分類し、続いてそのクラスターごとに異なるモデルを作成し、クラスターごとに家賃に対して説明力の大きい家賃形成要素を明らかにする方法を採用することで以上の課題に対応した。なお本研究では、不動産データが賃貸アパートと賃貸マンションを含むため、それぞれにおいてクラスタリングを行った後に、分析および結果の考察を行うものとする^{注 3)}。

3.1 クラスタリングの方法

本研究で用いるデータは高次元データであり、計算量も指数関数的に増えていくため、非階層的クラスタリングを用いてクラスタリングを行った。本研究のように大規模なデータに対する非階層的クラスタリングの手法としては k-means 法^{文 4)}や、同手法を改良した k-means++法が広く用いられているが、k-means++法の方が高い分類精度を持つことが報告されているため^{文 5)}本研究においても k-means++法を用いる。

ただし k-means++法はクラスター数を決定してくれないため、あらかじめ最適なクラスター数を決めておく必要がある。そこで本研究ではクラスターからの中心距離を指標にした Gap 統計量を用いてクラスター数を決定した。同手法はクラスター数を 1 から順次増加させていきながら、一様分布から作成されたクラスター内の距離の平均と、元データのクラスター内の距離の平均を比較し、より密集しているクラスター数を採用するという方法である^{文 6)}。図 3-1、と図 3-2 はアパートとマンションそれぞれにおいて表 2-1 の家賃形成要素を全て用いて k-means++法により算出したクラスター数ごとの Gap 統計量をプロットしたものである。Gap 統計量に基づいて、クラスター数に対する距離の傾きに変曲点となるクラスター数を採用した。その結果、賃貸アパートのクラスター数は 5、賃貸マンションのクラスター数は 6 が変曲点として妥当であることが分かった。

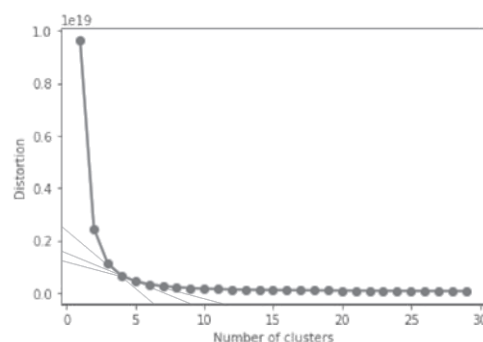


図 3-1 クラスタ数ごとの Gap 統計量（賃貸アパート）

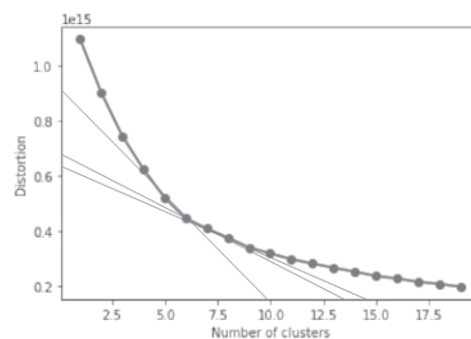


図 3-2 クラスタ数ごとの Gap 統計量（賃貸マンション）

3.2 クラスタリング結果

図 3-3 と図 3-4 に日本全国の賃貸アパートと賃貸マンションを市区町村別に集計し、各クラスターの棟数比率を計算した結果を示す。また表 3-1 と表 3-2 に各クラスターの変数の平均値上位 5 種類と主な分布地域を示す。

3.2.1 賃貸アパートにおける結果

クラスター 0 は主に近畿地方を中心に分布する物件群である。他のクラスターと比べて第二種低層住居専用地域の指定や景観地区指定が上位な点特徴的である。近畿地方では歴史的景観を有する閑静な居住地への評価が他地域よりも高いものと考えられる。クラスター 1 は主に近畿地方と九州地方を中心に分布する物件群である。他のクラスターには見られない徒歩圏の商業集積地域の数が上位に来ている点特徴的である。またクラスター 0 と同様に景観地区指定が上位に入っている。クラスター 2 は主に北海道地方や東北地方などの積雪寒冷地を中心に分布しており、灯油給湯設備や都市ガス暖房の有無など寒さ対策に重視した物件群といえる。そのため北海道・東北地方以外でも、冬季に寒冷となる中山間地域ではこのクラスターに分類される物件が多くなる傾向が見られた。クラスター 3 はクラスター 2 とは対称的に比較的温暖な太平洋側を中心に分布する物件群である。サンルームや太陽光発電システムの有無が上位に入ることからも、温暖な気候という特徴を良く反映していることが分かる。また浸水想定区域指定が上位なことから、太平洋側では浸水想定区域に数多くの賃貸アパートが分布してしまっていること

も示唆している。最後にクラスタ4は比較的全国的に分布するものの、主に東北地方や中四国地方に数多く分布する物件群である。他のクラスタほど特筆すべき特徴

は見られないものの、用途地域指定が行われていない主に中小規模の都市の近郊に数多く分布していた。

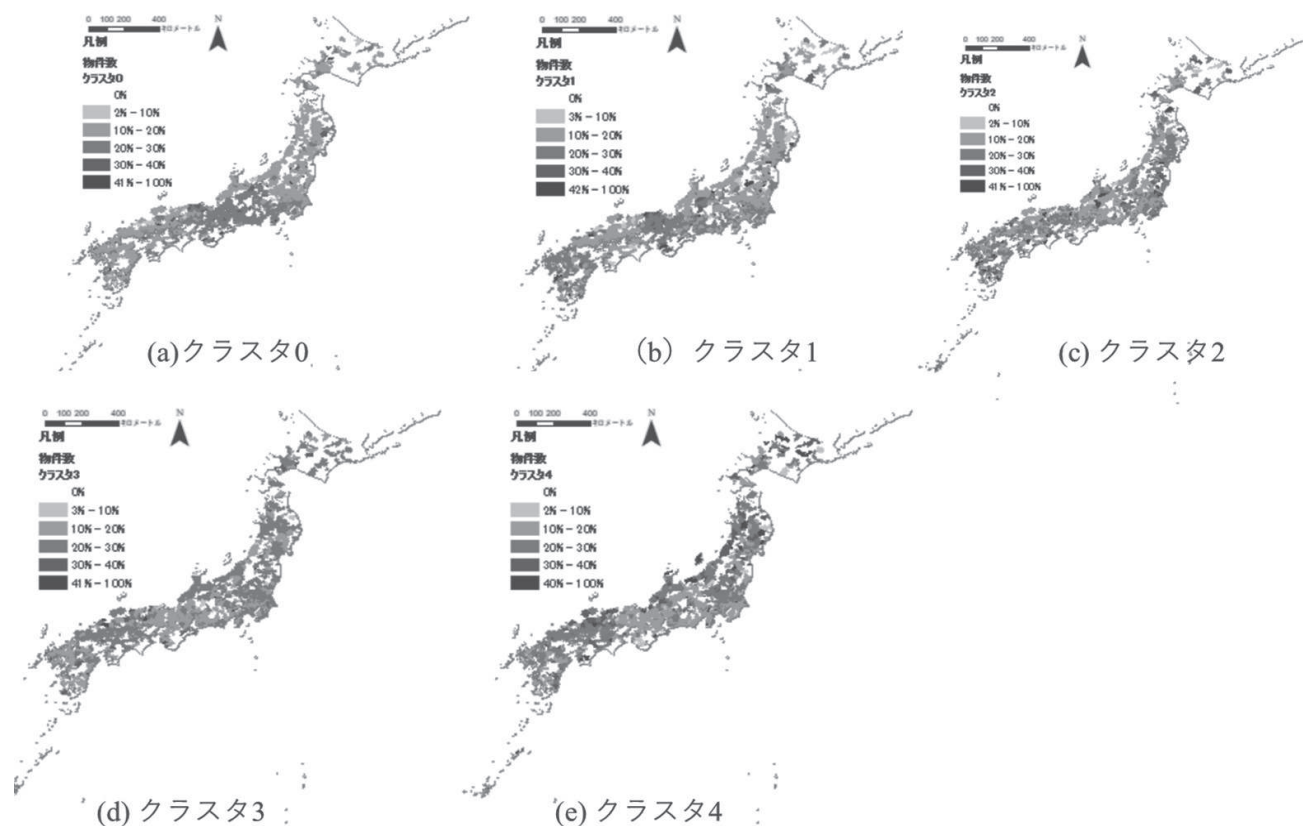


図 3-3 賃貸アパートにおける各クラスタの空間的な分布（構成比）注1)

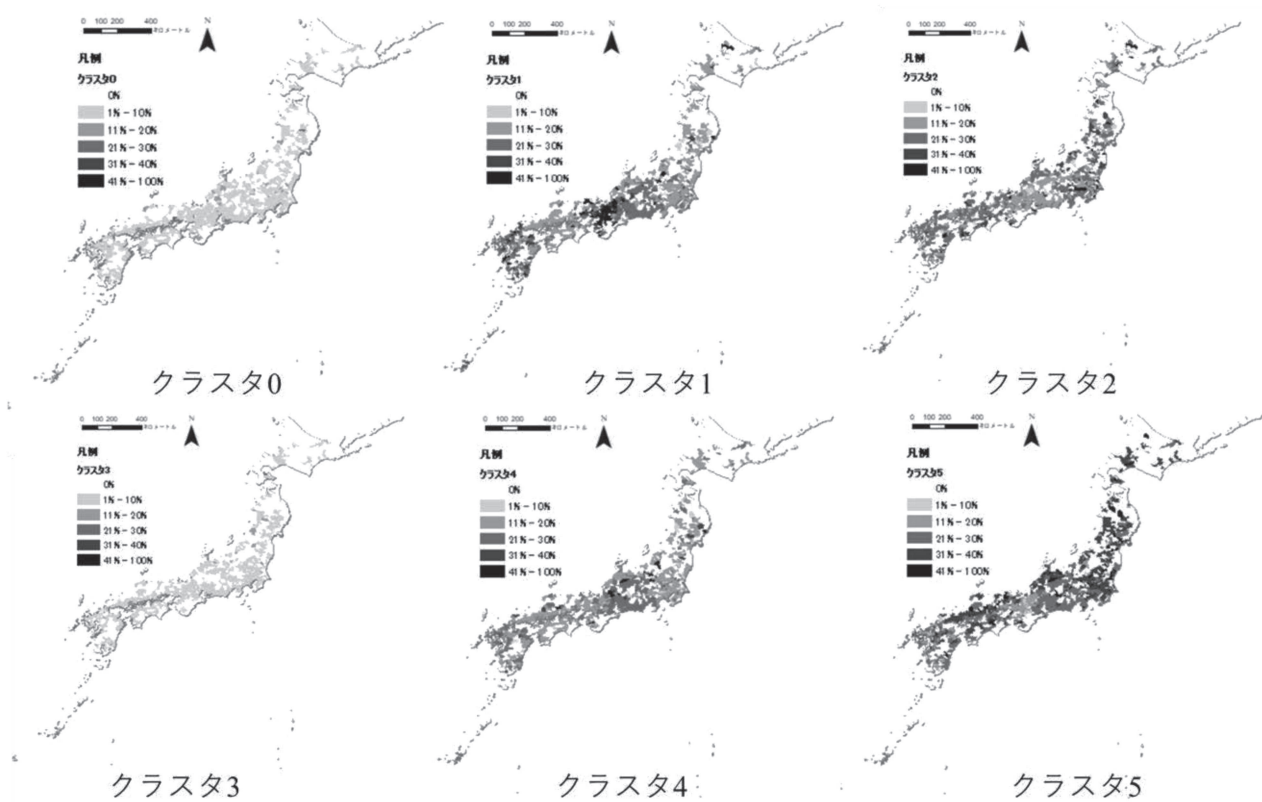


図 3-4 賃貸マンションにおける各クラスタの空間的な分布（構成比）注1)

表 3-1 賃貸アパートにおける各クラスタの変数の平均値上位 5 位と主な分布地域

クラスタ 番号	クラスタの変数の平均値上位5位					件数	平均家賃 [円]	主な 分布地域
	1位	2位	3位	4位	5位			
0	設備：有線放送有	用途地域：第二種低層住居専用地域指定	景観地区指定	設備：共用パーティールーム有	密集市街地指定	971,222	53,006	近畿地方
1	設備：オープンキッチン有	徒歩圏の商業集積地域数	設備：バスオーデオ有	設備：ダウンウォール収納有	景観地区指定	972,349	52,824	近畿・九州地方
2	設備：灯油給湯有	設備：都市ガス暖房有	設備：プロパンガス暖房有	設備：灯油暖房有	設備：都市ガス給湯有	972,812	52,883	北海道・東北地方
3	設備：敷地内AED有	設備：サンルーム有	浸水想定区域最大浸水深	設備：太陽光発電システム有	用途地域：近隣商業地域指定	971,747	52,328	太平洋側
4	用途地域：指定なし	設備：フィットネス施設有	設備：エレベーター2基以上有	市街化区域指定	構造：2X4	971,246	52,177	東北・中国・四国地方

表 3-2 賃貸マンションにおける各クラスタの変数の平均値上位 5 位と主な分布地域

クラスタ 番号	クラスタの変数の平均値上位5位					件数	平均家賃 [円]	主な 分布地域
	1位	2位	3位	4位	5位			
0	徒歩圏駅のブランド力総数	徒歩圏のデパート件数	徒歩圏の大規模店舗件数	徒歩圏の衣料品店件数	徒歩圏の酒場の件数	430,225	103,899	全国
1	景観地区	都市地域指定	用途地域：指定なし	密集市街地指定	建物構造：ALC	1,456,971	61,509	近畿・九州地方
2	密集市街地指定	設備：託児所	用途地域：準工業地域指定	設備：敷地内AED有	設備：MMコンセント有	1,618,224	100,814	東北・関東地方
3	設備：灯油給湯有	設備：灯油有	設備：灯油暖房有	設備：FF暖房有	3月最深積雪	127,440	53,752	全国
4	土砂災害警戒区域（地滑り）	土砂災害特別警戒区域	市街化区域	その他の都市地域	土砂災害警戒区域（土石流）	1,078,261	55,057	東海・近畿地方
5	用途地域：工業専用地域	構造：軽量鉄骨	用途地域：第一種低層住居専用地域	構造：木造	用途地域：第二種低層住居専用地域	1,591,689	74,730	東北・関東・北陸地方

3.2.2 賃貸マンションにおける結果

クラスタ 0 は日本全国に広く分布する物件群である。主に交通の結節点となる大規模な駅周辺に分布し、他のクラスタと比較すると家賃が高い傾向がみられる物件群である。クラスタ 1 は主に近畿地方や九州地方を中心に分布する物件群であり、賃貸アパートのクラスタ 0 やクラスタ 1 と同様、景観地区指定が上位となっている点が特徴的である。クラスタ 2 は主に東北地方と関東地方を中心に分布する物件群である。クラスタ 0 と同程度の家賃となり、全国的には比較的高い家賃のクラスタである。ただしクラスタ 0 ほど駅周辺に分布しているわけではない。クラスタ 3 は日本全国に広く分布する物件群である。ただし灯油給湯設備や灯油暖房設備を持つこと、また 3 月の積雪深が大きいことから、全国の積雪寒冷地に分布する物件群であることが分かる。クラスタ 4 は主に東海地方や近畿地方に数多く見られるが、それ以上に特徴的な点は土砂災害のリスクが大きい斜面地にも分布している点である。平時には日当たりや風通しに恵まれ、良好な景観と自然環境を享受できる一方で、常に土砂災害のリスクを抱える物件群と言える。最後にクラスタ 5 は主に東北地方、関東地方および北陸地方に分布する物件群であり、用途地域や構造からも分かるように郊外の閑静

な住宅街に分布する小規模マンションである。ただし本来住宅が分布することが出来ない工業専用地域が変数に入っている点は、今後さらなる考察が必要であろう^{注 4)}。

以上のように、約 400 次元に及ぶ数百万件という大規模データに対しても、k-means++法を用いることで、それぞれのクラスタの特徴を解釈可能な状態にクラスタリングすることができることが確認された。

4. スパースモデリング（SpM）を用いた特徴抽出

続いてクラスタごとに家賃に対する説明力が高い家賃形成要素を明らかにする方法を検討する。

4.1 重回帰分析の弱点

複数の説明変数に対して目的変数に対する説明力の大きさを比較する最も一般的な方法は、重回帰分析である。ただし重回帰分析では説明変数同士の相関が高くなる多重共線性を考慮する必要がある。また本研究のように説明変数の数が多い場合、複数の説明変数間で多重共線性が発生したり、過学習（オーバーフィッティング）により優れた汎化性能を発揮できなくなったりすることで、有意な結果が得られなくなる場合がある^{文 7)}。

4.2 スパースモデリング（SpM）の導入

そこで主査先行研究では、モデル構築を行う上で必要な説明変数を統計的に選択する「スパースモデリング(以下、SpM)」^{文 8)}を採用することでこれらの課題に対応した。SpM はビッグデータ分析において分析結果から大量のデータに埋もれて見えにくくなってしまいう有為な情報を抽出したり、法則性を導き出したり、断片的なデータを補完して実状を再現することができる手法である^{文 9)}。

これを本研究に適用した場合、未知のデータに対する予測精度を保ちつつ、多重共線性やオーバーフィッティングが起こらないように、データの次元を圧縮しながらモデリング（圧縮センシング）する、すなわち表 2-1 に示す 405 次元にも及ぶ高次元データから説明力が殆ど無い説明変数の係数を敢えて 0 にすることで、より単純な予測式を選ぶことが可能になる。つまり説明力が殆ど無い説明変数が自動的に排除され、人間が認知しうるレベルに結果を落とし込むことが可能になる。その結果、各クラスタでどの説明変数が家賃形成に大きな影響を与えているかを明らかにすることが可能になる。

なお SpM にはいくつかの方法が存在するが、本研究では Least Absolute Shrinkage and Selection Operator 法(以下、LASSO 法)^{文 10)}を用いる^{注 5)}。LASSO 法を 5 種類に分類した賃貸アパートと、6 種類に分類した賃貸マンションの各クラスタに適用し、各クラスタの特徴、すなわち家賃形成に寄与する家賃形成要素を明らかにした。

5. 結果

ここではSpMで得られた分析結果について、賃貸アパートと賃貸マンションの結果の紹介と、結果に対する考察を行った。図5-1および図5-2は賃貸アパートと賃貸マンションのクラスタごとに、どの家賃形成要素が家賃形成に正または負に寄与しているかを示している。横軸に405種類の家賃形成要素が並んでおり、縦軸が家賃への寄与度を示している。また表5-1と表5-2は賃貸アパートと賃貸マンションの各クラスタにおける家賃の説明力が大きい上位5種類の家賃形成要素を示す。さらに図5-3および図5-4は賃貸アパートと賃貸マンションそれぞれの全クラスタによる家賃への寄与度の高い家賃形成要素と累積情報量を示す。以下、賃貸アパートと賃貸マンションそれぞれの結果の考察を行った結果を示す。

5.1 賃貸アパートの結果と考察

全体的な傾向として、まず①住宅の性能関連情報と②狭域立地関連情報が家賃形成に大きく寄与していることが分かった。一方、③広域立地関連情報と④災害リスク関連情報は比較的到家賃に対する影響が小さいことが分上る。また家族世帯が数多く居住する物件で家賃が高くな

かった。また全体的に広い間取りを持ち、周辺に6人以上、周辺の将来人口が減少すると予想されている地域で家賃が低くなる傾向があることが分かった。

続いて各クラスタの特徴を考察する。クラスタ0は周辺の人口構成、各種生活利便施設へのアクセスの良さ、付帯設備として託児所とプロパンガス給湯が家賃に対して大きく正に寄与していた。クラスタ1は概ね周辺の人口構成や、都市地域指定の有無によって家賃が決定されていた。クラスタ2とクラスタ4は人口構成や給湯システムの有無に大きく影響されていた。なおクラスタ2ではプロパンガス給湯、都市ガス給湯、灯油給湯が家賃に正に寄与しているのに対して、クラスタ4ではこれらはすべて負に寄与していた。クラスタ3は敷地内AEDの有無、エレベータの基数など住宅の性能が寄与していた。

5.2 賃貸マンションの結果と考察

全体的な傾向として、①狭域立地関連情報が大きく家賃に寄与していた。特に周辺の人口構成や世帯人別・世帯類型別世帯数の割合に関する変数の寄与が見られた。この結果から、世帯類型（家庭構成）が人々のマンション選択に影響を与えているものと考えられる。

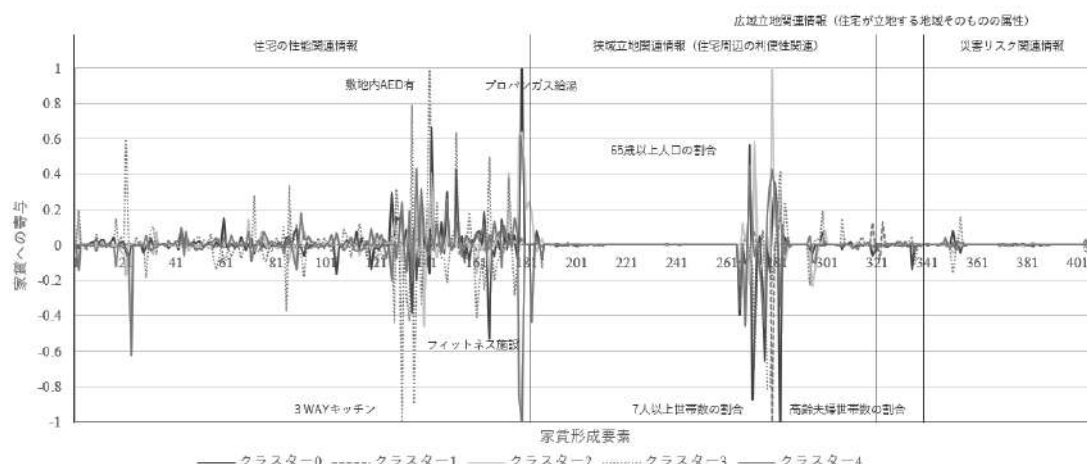


図5-1 各クラスタにおけるそれぞれの家賃形成要素による家賃に対する説明力（賃貸アパートの場合）

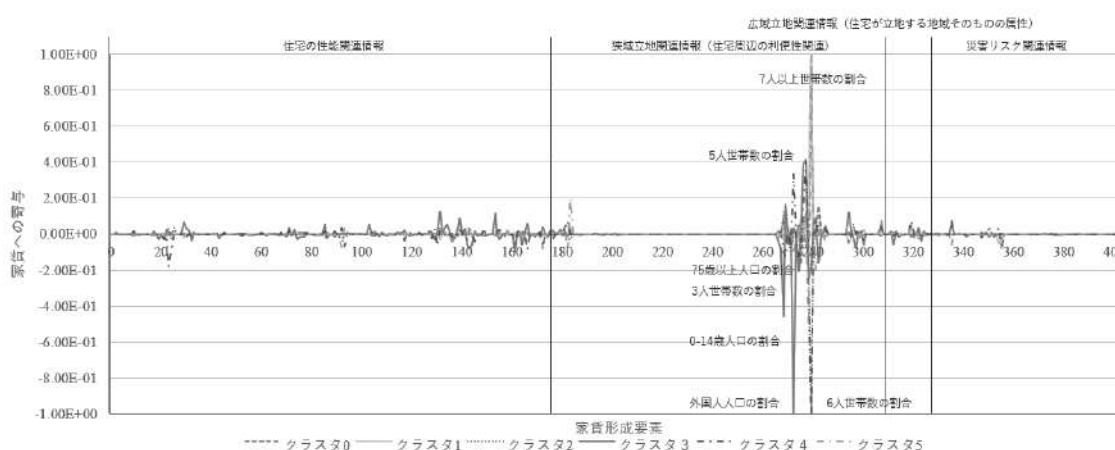


図5-2 各クラスタにおけるそれぞれの家賃形成要素による家賃に対する説明力（賃貸マンションの場合）

表 5-1 各クラスタにおける家賃に対する説明力が高い上位 5 種類の家賃形成要素（賃貸アパートの場合）

寄与	順位	クラスタ番号				
		0	1	2	3	4
正	1	4×4	0～14歳人口の割合	4×4	4×4	4×4
	2	0～14歳人口の割合	4×4	6人世帯数の割合	6人世帯数の割合	3×4
	3	高齢夫婦世帯数の割合	6人世帯数の割合	7人以上世帯数の割合	0～14歳人口の割合	6人世帯数の割合
	4	3×4	高齢夫婦世帯数の割合	高齢夫婦世帯数の割合	高齢夫婦世帯数の割合	4×3
	5	6人世帯数の割合	5×	5×	3×4	5×
負	1	共用シアタールーム	将来推計人口2030年	将来推計人口2035年	共用シアタールーム	共用シアタールーム
	2	将来推計人口2030年	一戸建世帯数の割合	一戸建世帯数の割合	一戸建世帯数の割合	一戸建世帯数の割合
	3	3人世帯数の割合	将来推計人口2035年	3人世帯数の割合	将来推計人口2030年	将来推計人口2030年
	4	将来推計人口2035年	共用シアタールーム	65歳以上人口の割合	将来推計人口2035年	外国人人口の割合
	5	一戸建世帯数の割合	3人世帯数の割合	75歳以上人口の割合	外国人人口の割合	3人世帯数の割合

表 5-2 各クラスタにおける家賃に対する説明力が高い上位 5 種類の家賃形成要素（賃貸マンションの場合）

寄与	順位	クラスタ番号				
		0	1	2	3	4
正	1	6人世帯数の割合	6人世帯数の割合	7人以上世帯数の割合	1～12ヶ所	4×4
	2	0～14歳人口の割合	7人以上世帯数の割合	0～14歳人口の割合	5×	7人以上世帯数の割合
	3	4人世帯数の割合	65歳以上人口の割合	外国人人口の割合	4×3	5×
	4	外国人人口の割合	5人世帯数の割合	5人世帯数の割合	4×4	1～12ヶ所
	5	2人世帯数の割合	4人世帯数の割合	5×	3×4	4×3
負	1	7人以上世帯数の割合	高齢夫婦世帯数の割合	65歳以上人口の割合	7人以上世帯数の割合	外国人人口の割合
	2	核家族世帯数の割合	0～14歳人口の割合	男婦数の割合	男婦数の割合	3人世帯数の割合
	3	65歳以上人口の割合	外国人人口の割合	4人世帯数の割合	5人世帯数の割合	75歳以上人口の割合
	4	男婦数の割合	5×	核家族世帯数の割合	4人世帯数の割合	一戸建世帯数の割合
	5	将来推計人口2020年	4×4	将来推計人口2020年	共用バーテイルーム	将来推計人口2020年

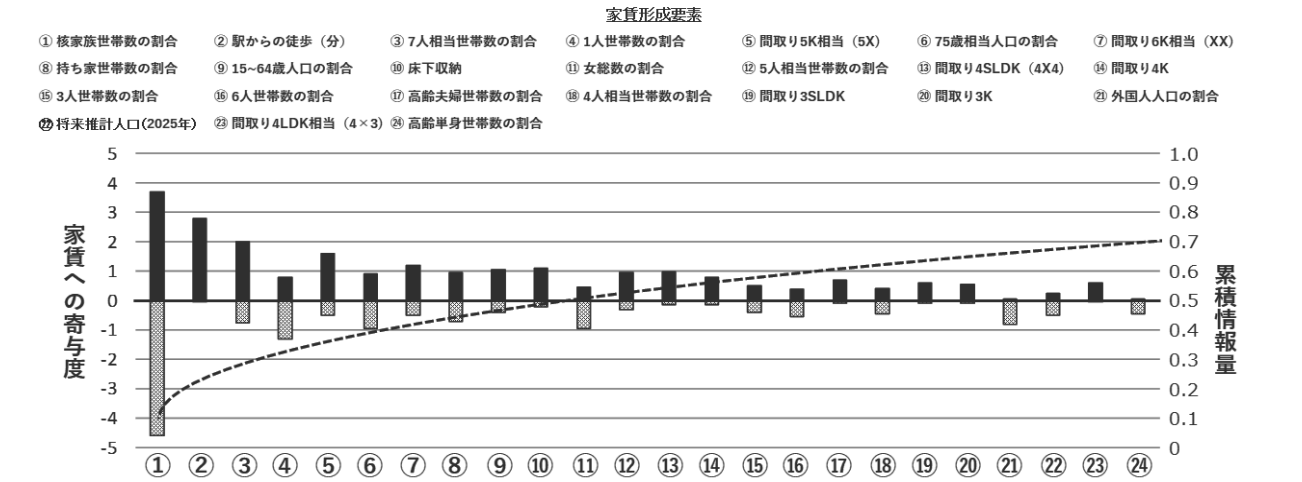


図 5-3 寄与度の高い家賃形成要素（棒グラフ）と累積情報量（破線）（賃貸アパートの場合）

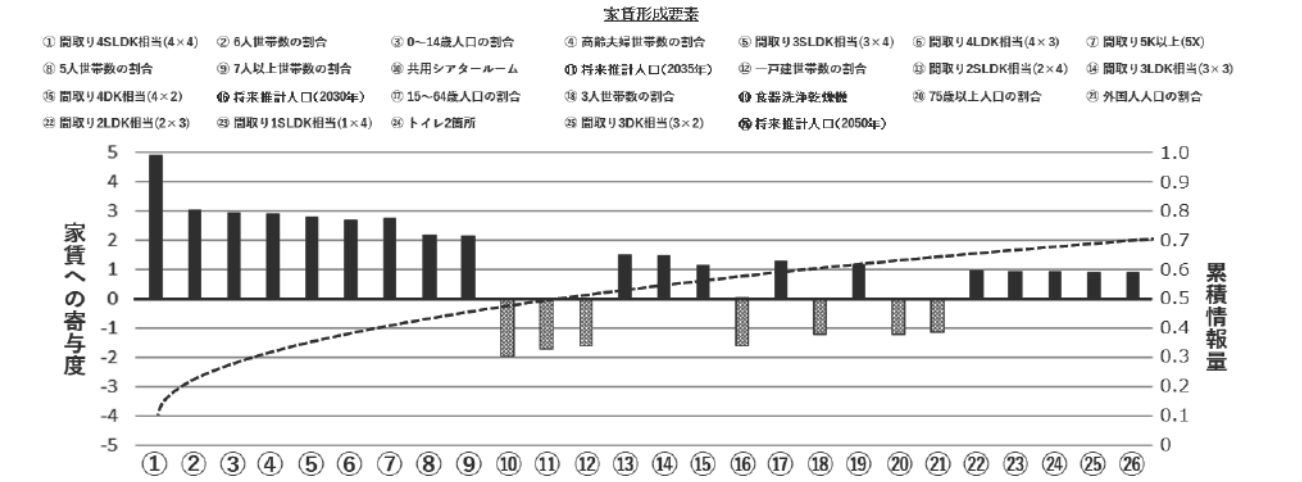


図 5-4 寄与度の高い家賃形成要素（棒グラフ）と累積情報量（破線）（賃貸マンションの場合）

続いて各クラスタの特徴を見ていく。まずクラスタ 0，1，5 は何れも 7 人以上世帯数が正に寄与していた。またそのうちクラスタ 0 は 0-14 歳人口の割合も正に寄与しており，一方 75 歳以上人口の割合は負に寄与していた。クラスタ 1 は土砂災害特別警戒区域，クラスタ 5 は灯油暖房が強く正に寄与していた。ここで特筆すべきはクラスタ 1 の土砂災害特別警戒区域の指定が家賃に正に寄与していることである。土砂災害のリスクよりも，斜面地で享受できる平時の環境の方が優先的に評価されている可能性があると言えよう。一方，クラスタ 2，3，4 では何れも 7 人以上世帯数が負に寄与していた。またクラスタ 3 では外国人人口の割合が大きく負に寄与していた

が，クラスタ 4 では逆に正に寄与していた。

以上のように賃貸マンションの家賃はマンション周辺の居住者の属性に大いに影響を受ける可能性があることが分かった。今後，マンション周辺の人口・世帯数やその構成が，住まいやその地域のコミュニティに与える影響を定量的に分析することで，この結果に対して更に深い理解を得ることが出来るものと期待できる。

5.3 分析結果のまとめ

SpM によりクラスタごとに家賃に対して正または負に寄与する家賃形成要素が明らかになった(図 5-1, 図 5-2)。また寄与度の高い順に全体の家賃形成要素に対する累積

情報量（各家賃形成要素の寄与率が全体に占める割合の累積）を計算することで、情報量の観点からも家賃形成要素に対する考察を行うことができた（図 5-3、図 5-4）。

まず上位の家賃形成要素に着目すると、その多くは 7 人以上の世帯の割合や 0-14 歳人口の割合など物件周辺の人口動態に関する要素および、間取りに関する要素であった。この結果から、周辺に比較的世帯人員が大きい世帯や子供が多い地域で、面積や部屋数が大きい物件において家賃が高くなる傾向があることが分かった。この結果は常識的な家賃形成と一致していると言えよう。

さらに以上で挙げたもの以外の家賃形成要素も見ていく。まず人口動態関連の家賃形成要素は、基本的に将来人口の現象幅が大きく、高齢者人口、外国人人口が多いほど負に寄与する傾向が見られた。また建物の付帯設備は、トイレ 2 基やエレベータ 2 基、サウナなどの付加価値的な要素が家賃に正に寄与していることが明らかになった。また情報量の観点から見ても、人口関連と建物の付帯設備に関する上位 50 位の家賃形成要素で 80% 以上の累積情報量を持っており、他の約 350 要素は家賃形成に大きく影響していないことが明らかになった。情報量の増加率に着目してみても、情報量が対数関数的に増えていることから、上位の構成要素が潜在的な家賃形成に大きく影響を与えていることが分かる結果となった。

一方、居住者が洪水や地震などの自然災害のリスクに晒されにくい住宅選択に繋がる家賃形成を実現するためには、自然災害に関するリスク要因（浸水想定区域や津波想定区域、土砂災害警戒区域等の指定有無）や、都市計画区域・用途地域の指定状況も家賃形成に寄与していることが重要である。しかし、本研究の成果からはそれらは家賃形成に対して大きな寄与は示さなかった。この結果から、居住者の自然災害への安全性を考慮すると、自然災害のリスクの小ささが家賃により強く正に寄与するべきである。しかし今日の日本では、災害のリスクが高い地域だとしても、多くの居住者にとって利便性が高く支持される地域であれば、災害に対して安全な地域と同水準（場合によってはそれ以上）の家賃が設定され、物件が持つ機能などが災害の危険性よりも重視されて物件の選択が行われていることを示唆している。なおこの結果は著者既存研究でも類似した結果が得られたが、本研究で物件単位の分析にしたことで、改めて確認された。

また生活利便性（生活利便施設へのアクセスの良さ）も、家賃への寄与は小さかった。ただしこれは生活利便性が家賃形成要素にならないという意味ではなく、先述した人口動態関連の要素に生活利便性に関する要素が従属する形となったため、結果的に寄与が小さくなったものと考えられる。

以上のことから、家賃形成に大きく影響するのは、賃貸アパート、賃貸マンションともに物件周辺の人口動態

および建物機能に関連する要素であることが分かった。また同じ性能をもつ物件の場合、人口が多い都市部では家賃が高く、一方、同じ地域内では高付加価値な機能が付随する物件ほど家賃が高くなる傾向があることが分かった。ただし、人口が多くても高齢化が進んでいる場合や、将来人口が減少する可能性が高い地域、また外国人が多く住む地域では人口の多さが負に寄与することもあることが分かった。さらに自然災害のリスクは家賃形成にあまり影響を与えていないことも明らかになった。

6. おわりに

本研究では主査先行研究の発展的研究として取り組んだ。まず日本全国の家賃形成の実態明らかにするために、日本全国の住宅用途の建物 1 棟 1 棟の分布が把握できる建物 MGD を整備し、続いてそれらに家賃形成要素と家賃に関する情報を付与することで、住まい手にとっての住宅の価値を評価出来るデータベースを構築した。また同データベースを 405 種類の家賃形成要素（説明変数）および家賃（目的変数）に基づいてクラスタリングし、さらに各クラスターでスパースモデリングを実施することで、クラスター毎に家賃に対して正あるいは負に寄与する家賃形成要素を明らかにすることが可能になった。

その結果、地域による違いはあるものの、全国的な傾向としては、賃貸アパート、賃貸マンションともに物件周辺の人口動態や建物そのものの機能に関する家賃形成要素が家賃形成に大きな影響を与えていることが分かった。また高齢化率の高さや外国人人口の多さは家賃に負に寄与する傾向にあることも分かった。さらに主査先行研究では住宅の性能の高さは必ずしも家賃に正に寄与するものではないという結果が得られたが、本研究ではこれを否定し、付帯設備が充実することで家賃を高める効果があることが示された。一方、主査先行研究と同様に、本研究でも自然災害のリスクが家賃形成にあまり影響を与えていないことが分かった。これは住まい手にとって災害に強い物件を選択したいという住まい手側の価値と市場価値との間に乖離が見られる状況であり、今後は正されていくべき点であると言えよう。

以上のように本研究は主査先行研究の課題を数多く克服し、研究目的の達成にさらに近づくことができた。しかしながら、本研究をさらに有用な内容にしていくためには、なお数多くの課題が残されている。まず主査先行研究でも指摘したように、家賃形成に影響を与えると考えられる家賃形成要素のさらなる充実が必要である。例えば犯罪発生に関する情報や、空き家の分布情報などはデータとしては存在していることを確認しているため^{注 7)}、将来的に導入することを検討したい。また景観や居住者のライフスタイル^{注 8)}といった定性的な家賃形成要素の導入も検討する必要がある可能性がある。さらに従

来の統計から得られる定住人口だけではなく、時々刻々と変化する時間別の人口とその変化も家賃形成に大きな要因になるものと予想される。すでに主査らが実施した研究では、携帯電話の移動履歴から得られる人流ビッグデータを用いて推定した時間別人口と家賃との間に関係性が見られることが示されており^{文 13)}、加えて類似の研究は国内外においても殆ど見られないため、学術的にも斬新な挑戦と言える。今後はこうしたデータを用いた家賃形成要素の拡張も検討したい^{注 9)}。

また分析手法の改善も必要である。今回のような高次元データを取り扱う場合、潜在的な家賃形成構造を容易に理解できるモデル選択が好ましい。しかし、本研究の手法 (LASSO 法) は 1 つ 1 つの要素にたしてペナルティ計算を行うため、選択される変数も多様となり結果的に考察が難しくなった。そのため今後は高次元データの考察をより容易にするために、類似した家賃形成要素同士をグループ化することで、グループごとに解釈が可能なモデル手法 (Group LASSO 法) などの利用を検討したい。

さらに、本研究の海外展開も視野に入れている。主査先行研究では大韓民国の不動産に関する統計・データの整備状況を調査し、大韓民国では我が国よりも空間的にも時間的にも高精細な MGD や、統計情報、不動産関連データが整備されており、我々の研究を適用出来る可能であることが分かった。本研究ではさらに中華人民共和国 (上海市) と、アメリカ合衆国 (ニューヨーク市) の調査を実施した^{注 10)}。その結果、何れの国も日本のように全国的に均一な性能の不動産関連のデータや統計は整備されていないものの、上海市やニューヨーク市といった都市単位の場合、日本と同程度、あるいは日本以上に高精細かつ高頻度に整備・更新されているデータも存在していることが明らかとなった。ただし中国の場合、国外へのデータの持ち出しが制限されている場合があるため、中国で同様の研究を展開することを考える場合は、中国国内の大学・研究機関等との連携が不可欠となる。今後も引き続き海外調査を進めていきたいと考えている。

加えて、本研究を実施する上で大きな課題となったのが分析に要する計算時間である^{注 11)}。そのため今後は GPU ベースの高速計算環境の導入や、計算アルゴリズムの改良が必要になるものと考えている。

なお本研究ではデータから得られる情報に基づいて、家賃形成に正、あるいは負に寄与する家賃形成要素を定量的に探索、特定することが実現したが、このような結果となる要因が住まい手による価値観やニーズによるものなのか、不動産提供者や行政による何らかのコントロールによるものなのか、マスメディアから発信される情報の過不足によるものなのか、あるいはこれらが複合することによるのかなどは、本稿執筆の時点では明らかにはなっていない。この点を明らかにする方法の開発と実

施は重要な課題であり、今後実施を検討していきたい。

以上のように本研究はまだ数多くの課題は残っているものの、主査既存研究で挙げられた課題を数多く克服し、学術的により有意義な成果を達成することができた。主査らは本研究の手法およびデータの発展性・拡張性の高さや、本研究課題の社会的要請の高さは、我が国の今後の住宅政策の立案を大いに支援できるものと期待している。我が国の良好な住環境の実現に向けて、本研究の更なる発展を続けていきたい。

<謝辞>

本研究を実施するにあたり、様々な空間情報および統計情報を東京大学空間情報科学研究センターの共同研究利用システムより、共同研究 No. 884 の一環として提供を受けました。加えて中華人民共和国 (上海市) およびアメリカ合衆国 (ニューヨーク市) の空間情報に関する聞き取り調査でご協力頂いた、New York University Shanghai の ChengHe Guan 先生と同研究室の研究メンバーの皆様に、ここに記して謝意を表します。

<注>

- 1) 以下の URL において図 2-2、図 3-3、図 3-4 のカラー版と、Web GIS による図 3-3 および図 3-4 の結果の閲覧が可能。
<http://akiyama-lab.jp/yuki/jusoken/jusoken2020.html>
- 2) 主査先行研究の家賃形成要素 (説明変数) は 82 種類であったが、本研究では 405 種類 (表 2-1 参照) と大幅に増加した。
- 3) なお k-means++法でクラスタリングを行う際には、説明変数間の単位が異なるため、全ての変数を標準化した。これは Gap 統計量の計算や、スパースモデリングを行う際にも同様である。
- 4) 本研究では用途地域の判定に国土数値情報の用途地域データを用いた。国土数値情報の用途地域データは最終更新年が 2011 年であるため、2015 年の物件の空間分布の実態とは必ずしも一致していない可能性もあることも原因であると考えられる。
- 5) LASSO 法では推定したいパラメータを ax 、対象データを y とした時、 x の 1 次ノルムを、非スパース性のペナルティ項 (λ) として付加して最小化 (L1 正則化) させることで解析し、スパース解 ax を得る。 λ はスパース性の制御パラメータであり、 $\lambda = 0$ で通常の重回帰分析と一致する。LASSO の L1 正則化では、適切な規範で λ を選択し、説明変数を刈り込むことでデータの特徴構造を適切に抽出したスパース解を得ることが可能である。また本研究ではモデルのパラメータチューニングに 5-fold Cross Validation による平均二乗誤差根 (RMSD: Root-Mean-Square Deviation) を評価指標に用いて実施した。
- 6) 間取りに関する情報のうち、「3Xn」 (n は 1 から 4 が入る) となっているものは X に SLDK が入り、n はそれらが何種類入るかを意味している。例えば「4X2」なら 4DK や 4LK など、「3X4」なら 3SLDK となる。
- 7) いくつかの自治体に対するヒアリングを行った結果、犯罪発

生地点（あるいは町丁目単位程度の領域単位）の情報は各都道府県警が保有していることを確認した。また空き家の分布情報は自治体によっては独自に調査を行っており、建物単位のピンポイントなデータを保有している場合がある。ただし何れのデータも全国規模で収集することは極めて困難なため、推定値の活用も検討する必要がある。なお空き家の分布推定は、Akiyama et al. (2020)^{文 11)}による自治体全域を対象にピンポイントに推定した例や、石河ほか (2016)^{文 12)}による日本全国を 3 次メッシュ単位で推定した例などがある。

8) 例えば町丁目単位の居住者のライフスタイルを推定したデータである Chomonicx3.0 (株式会社ゼンリンジオインテリジェンス) は、町丁目を都市化の程度と富裕度に応じて 36 通りに分類している。

9) 当初は携帯電話の移動履歴に基づく人流ビッグデータを用いた物件周辺の時間別人口を家賃形成要素に盛り込む予定であったが、データ保有企業との研究利用のための調整が研究期間中に完了できなかったこと、またもしデータを購入する場合、非常に高額となるため、本稿作成時点での導入は断念した。

10) New York University Shanghai を訪問し (2019 年 10 月 14 ~17 日)、中華人民共和国 (上海市) およびアメリカ合衆国 (ニューヨーク市) の空間情報データや不動産関連データの整備状況についてヒアリング調査を実施した。

11) 例えば全国約 450 万件の賃貸アパートのクラスタリング (kmeans++法) は、6 コアの CPU と 256GB のメモリを搭載したサーバ環境を利用しても約 24 時間の計算時間を要した。また SpM の計算も、6 クラスタ (賃貸マンション) の計算に同じ計算環境下で約 12 時間の計算時間を要した。

<参考文献>

- 1) 秋山祐樹・小川芳樹, マイクロジオデータを用いた家賃形成メカニズムの研究—住まい手からみた住宅の価値と市場価値の乖離に着目して—, 住総研 研究論文集・実践研究報告集, 44, 1-12, 2018.
- 2) 崔廷敏・浅見泰司, 賃貸住宅居住者の満足度評価に見られる潜在的評価構造. 都市住宅学, 42, 86-97, 2003.
- 3) Lazauskaitė, D., Burinskienė, M. and Podvezko, V., Subjectively and objectively integrated assessment of the quality indices of the suburban residential environment, International Journal of Strategic Property Management 2015, 19, 297-308, 2015.
- 4) MacQueen, J. B. : Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281-297, University of California Press, 1967. 1
- 5) Arthur, D. and Vassilvitskii, S.: K-means++: the advantages of careful seeding, Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms.

pp. 1027-1035, ACM, 2007. 1

6) 志津綾香, 松田眞一: クラスタ分析におけるクラスター数自動決定法の比較, アカデミア情報理工学編, Vol. 11, pp. 17-34, 南山大学, 2011. 3

7) Nguyen, N. and Cripps, A.: Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. Journal of Real Estate Research, Vol. 22, No. 3, pp. 313-336, 2001

8) Hastie, T., Tibshirani, R. and Wainwright, M.: Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC, 2015. 8

9) 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. 2

10) Tibshirani, R., Walther, G. and Hastie, T.: Estimating the number of clusters in a data set via the gap statistic, Series B Statistical Methodology, Vol. 63, No. 2, pp. 411-423, Royal statistical society, 2001. 1

11) Akiyama, Y., Ueda, A., Ouchi, K., Ito, N., Ono, Y., Takaoka, H. and Hisadomi, K.: Estimating the Spatial Distribution of Vacant Houses using Public Municipal Data. Geospatial Technologies for Local and Regional Development, pp. 165-183, 2020. 3

12) 石河 正寛, 松橋 啓介, 有賀 敏典, 金森 有子, 栗島 英明: 空家の地域内分布に関する現況及び将来推計—世帯数と住宅数の差分に着目して—, 都市計画論文集, Vol. 51, No. 3, 833-838, 2016. 10

13) Qin, Y., Akiyama, Y., Ogawa, Y., Shibasaki, R. and Sakuramachi, R., Study of the relationship between rental house prices and people congestion of urban areas based on mobile phone GPS data, Accepted Short Papers and Posters from the 22nd AGILE Conference on Geo-information Science, #80, 2019. 6