

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—

主査 秋山 祐樹^{*1}

委員 小川 芳樹^{*2}

本研究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ために、日本全国の住宅用途の建物1棟1棟の分布を収録したマイクロジオデータを整備した。またそれらに様々な家賃形成要素と最低保証家賃に関する情報を付与し、日本全国の住宅の価値を評価出来るデータベースを構築した。さらに同データを集計化後にクラスタリングし、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自動的に明らかにする技術が実現した。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、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確認された。

キーワード：1) 家賃，2) マイクロジオデータ，3) 住宅の性能，4) 狭域立地，5) 広域立地，
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) 居住者，10) 市場価値

A STUDY FOR THE FORMATION MECHANISM OF HOUSE RENT USING MICRO GEODATA

- Focusing on the mismatch between house value for residents and market value -

Ch. Yuki Akiyama

Mem. Yoshiki Ogawa

In order to clarify the formation mechanism of house rent throughout Japan, this study developed the micro geodata (MGD) that contains the spatial distribution of all residential buildings in Japan with their various rent formation factors and information on minimum guaranteed rent. In addition, by clustering the MGD after spatial aggregation and performing the sparse modeling on each cluster, a methodology has been realized that automatically reveals rent formation factors that contribute positively or negatively to minimum guaranteed rent. The result showed that depending on the region, we confirmed mismatches between house values for residents and market values.

1. はじめに

1.1 本研究の背景

我が国では今日に至るまで「市場における不動産の使用価値（家賃や住宅価格）」が、本来住宅が持っているはずの「住まい手から見た価値」と比較して低く、あるいは高く設定されてきた。例えば一般的な戸建住宅や共同住宅の価格がサラリーマン世帯の平均年収に占める割合は、首都圏や大阪、名古屋などの大都市では平均年収が高いにも関わらず、他の地方都市と比べると大きくなっており、その状況は長らく横ばいである^{文1)}。すなわち東京をはじめとする大都市と、その他の都市の住宅の性能には大きな差が見られないにも関わらず、大都市圏の住宅は大都市圏に分布している、というだけで極端に価値が高い状況が続いているとされている。

また近年では「リセールバリュー」という考え方が生まれており、住宅を「住み潰す」のではなく、「資産」と

して捉えていく動きが出てきている。しかしリセールバリューの多くの部分はその住宅が都市中心部か郊外のどちらに立地しているかという「広域立地」で決まるとされており、住宅周辺の利便性を左右する「狭域立地」や、「住宅の性能」の影響は小さいとされている^{文2)}。このように住宅の価値は住宅の性能や近隣空間の利便性という、そこに住む住まい手にとって重要であるはずの価値はあまり考慮されず、「都市か郊外か」で決定される場合がある。加えて「地名のブランドイメージ」がこの傾向をさらに加速させている側面もあると言われている^{文3)}。

以上のような状況から、今日の住宅価値決定のメカニズムは、居住者にとってそこに居住する上で重要な住宅の性能や、その周辺の利便性といった「狭域立地」よりも、都市からの距離や地名のブランドなどの「広域立地」が支配的である可能性がある。このような状況を放置すれば、地方都市や郊外の住宅価値が低下する一方で、大

^{*1} 東京大学空間情報科学研究センター 助教 ^{*2} 東京大学生産技術研究所 特任研究員

都市、特に東京での住宅価値の高止まり、あるいは更なる高騰に歯止めがかからなくなることが懸念される。またリセールバリューの低下による中古住宅市場の停滞を招いている可能性もあり、その結果ライフステージの変化に合わせた住みかえが困難になり、住生活の改善・向上が困難になっているものと考えられる。

さらに近年では「レジリエンス」という概念を取り入れた人命を守り抜ける災害に強靱な住まいづくり、地域づくりも求められている^{文4)}。先の東日本大震災では過去の地震や津波で大きな被害を受けた地域に立地していた住宅が大きな被害を被っており^{文5)}、本来はこのような防災上危険な地域での居住は抑制されるべきであった。住宅が立地する地域の災害リスクが十分に考慮されずに住宅価値が決定されている現状は、レジリエンスな住まいと地域を実現する上で今後は正されていくべきである。

1.2 本研究の目的

以上の現状と課題を鑑み、本研究では今日の家賃形成の実態を、様々な統計データやマイクロな空間データ（以下「マイクロジオデータ」^{文6)}）を駆使して定量的に明らかにする。そしてその結果を持って今日の家賃形成の実態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るべき要因の解明に向けた基礎的な空間データを整備し、我が国の将来的な住生活の向上に貢献できる成果の実現を目指すものである。

本研究ではまず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マイクロジオデータを整備し、続いてそれらに以下の「家賃形成要素」に関する情報を付与することで、住まい手にとっての価値を評価出来るデータベースを構築する。

- ① 住宅の性能：住宅の構造や築年代
- ② 狭域立地：住宅周辺の立地環境に関する情報（生活便利施設や駅、主要道路等へのアクセス性）
- ③ 広域立地：都市域からの距離やブランド力の高い地名を持つ地域からの距離
- ④ 災害リスク：地震や津波、洪水、土砂崩れなどの自然災害を被るリスクと避難所へのアクセス性

さらに同データベースに一般社団法人 移住・住みかえ支援機構から提供された「全国郵便番号別最低保証家賃査定データ」から得られる、⑤地域ごとの最低保証家賃情報を与えて比較することで、地域ごと（郵便番号単位）の家賃形成のメカニズム、すなわちどの家賃形成要素が住宅価格を決定するのに支配的であるか、ということを明らかにする。その上で地域ごとにどの家賃形成要素に関する是正策を講じていくべきか、またその要因の大小に地域的な偏りはあるのかなどを明らかにし、我が国の今後の住宅政策、すなわち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の整備を目指す

とともに、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ことを目標とする。

1.3 本研究の新規性

空間データや統計を活用した家賃形成のメカニズムに関する研究は国内外においてこれまでに数多く見られる^{文7), 文8), 文9)}。しかしいずれも対象地域が限定的であり、日本全国というマクロなスケールに対して、建物1棟1棟というマイクロジオデータを大規模に運用して、家賃形成のメカニズムの実態把握を試みた研究は前例がない。また家賃形成の要因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皆無である。1.2で挙げた「我が国の将来的な住生活の向上」を目指すためには、レジリエントな住まいと地域形成に繋がる全国的な定量的評価が求められていると考えられる。本研究はこれらの課題に対して、日本全国のマイクロジオデータと、最新のデータサイエンスを駆使して挑戦する、学術的に極めて意欲的なだけでなく、社会的にも実用性、緊急性が高く将来の拡張性も高い研究であると言えよう。

2. 建物マイクロジオデータ開発

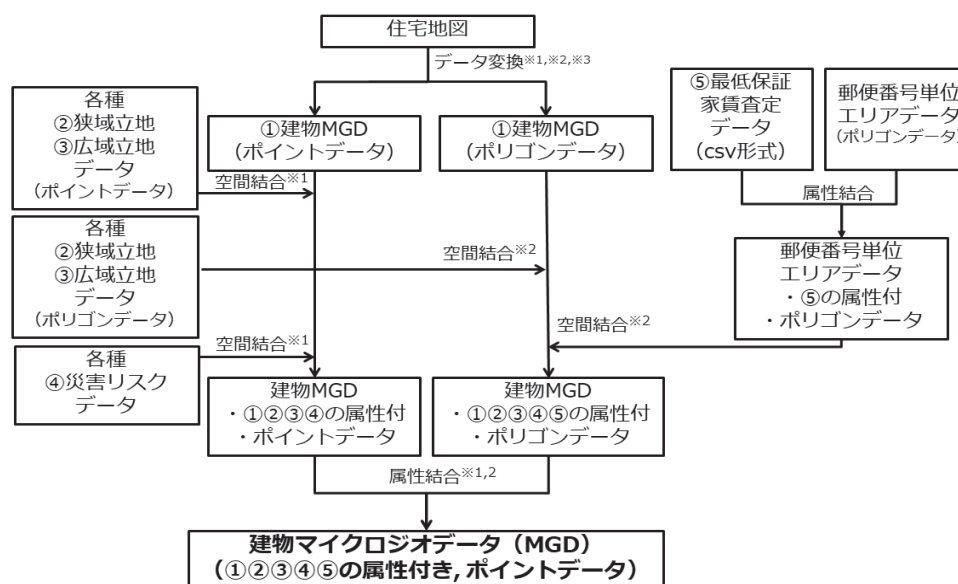
図 2-1 に本研究におけるマイクロジオデータの開発方法の流れを示す。なお対象地域は日本全国とする。

2.1 建物分布情報の取得

日本全国の住宅の家賃形成を決定する家賃形成要素を明らかにするためには、まず日本全国の住宅用途の建物の分布を把握できるデータが必要である。そこで最新の住宅地図（2013～2014年）から日本全国約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。さらにこのデータに①住宅の性能に関する情報を与えるために、小川（2016）^{文10)}の手法を適用することで、建物ごとの構造（木造/非木造）や築年代^{注1)}の推定を行った。さらにAkiyama et al. (2013)^{文11)}の手法で整備した建物ごとの居住者データである「マイクロ人口統計」から、2010年の建物ごとの年齢性別人口の推定値を与えた。以上により建物ごとに①住宅の性能に当たる構造、築年代、年齢性別人口の情報が与えられた。

2.2 狭域立地および広域立地情報の付与

続いて建物MGDに格納された建物に、表2-1に示す②狭域立地と③広域立地に関する情報を与えた。何れのデータも建物MGDと空間結合を行うため、②狭域立地と③広域立地に関するデータを全て空間結合しやすいデータ形式に変換した。また全国のデータを高速かつ並列的に空間結合処理が行えるように専用のデータ処理プログラ



※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。
 ※3 秋山ほか（2013）を用いて、①住宅の性能である建物構造や築年代を与えた。

図 2-1 本研究におけるマイクロジオデータ開発の流れ

ムを開発し、全国約 6,000 万棟の建物 1 棟 1 棟全てに表 2-1 の属性を与えた。

2.3 災害リスク情報の付与

さらに建物 MGD に格納された建物に、表 2-1 に示す④災害リスクに関する情報を与えた。建物ごとの最大倒壊率と最大出火率は、防災科学技術研究所で公開されている確率的地震動予測地図の全シナリオを格納し、各シナリオにおける建物ごとの倒壊率と出火率を推定した情報を保有する「震災ビッグデータ」（秋山ほか，2013^{※10）}より情報を取得し建物ごとに付与した。

2.4 最低保証家賃の付与

最後に建物 MGD に格納された建物に、⑤地域ごとの最低保証家賃情報を与えた。最低保証家賃に関する情報は一般社団法人 移住・住みかえ支援機構から提供を受けた「全国郵便番号別最低保証家賃査定データ」により収集した。同データからは、郵便番号単位でその地域の賃貸物件（戸建住宅・共同住宅）の一括査定の査定結果としての最低保証家賃（3 万円，5 万円，7 万円の 3 区分）の情報を得ることができる。ただし本研究で用いる建物 MGD には建物ごとの立地する場所の郵便番号の情報を持っていないため、このままでは全国郵便番号別最低保証家賃査定データの属性を建物に与えることが出来ない。

そこでまず株式会社ゼンリンジオインテリジェンスから提供を受けた日本全国の郵便番号の領域を表現するポリゴンデータである「郵便番号地図データ」と、建物 MGD を空間結合し、全ての建物に郵便番号の情報を与えた。この段階で建物 MGD に格納された全ての建物に郵便

番号が与えられた。続いてその郵便番号に基いて各建物に最低保証家賃情報を付与した。以上により日本全国の建物 1 棟 1 棟に様々な家賃形成要素と最低保証家賃の情報を格納した建物 MGD が実現した（図 2-2）。

3. 地域のクラスタリング

3.1 郵便番号単位による家賃形成要素の集計化

本研究では最低保証家賃がどのような家賃形成要素により決定されるのか、ということ明らかにすることを目指している。そこで基本的には家賃形成要素を説明変数、複数の最低保証家賃を目的変数とした分析を行い、目的変数である最低保証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良い。ただし本研究では最低保証家賃が郵便番号単位で与えられるため、家賃形成要素もその単位に合わせて集計化する必要がある。すなわち建物 MGD の建物ごとに与えられた家賃形成要素を郵便番号単位に集計化する必要がある。そこで最低保証家賃の情報を持つ全国 29,043 箇所の郵便番号単位の領域において、表 3-1 に示すように建物単位に与えられた家賃形成要素（説明変数）を郵便番号単位に集計化した。

3.2 地域のクラスタリング

以上のデータ整理により、郵便番号単位ごとに家賃形成要素を説明変数、最低保証家賃を目的変数とする回帰分析（重回帰分析）が理論的には可能になった。ただし本研究の分析では、分析対象地域が日本全国となるため、様々な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 3-1 に示すように説明変数の数も非常に多くなるため（82 種類）、単一のモデルで日本全国の最

表 2-1 建物 MGD が持つ家賃形成要素

データの種類		内容	データ作成年	データソース	データ収録範囲
大分類	小分類				
①住宅の性能	各種建物性能情報	構造推定値（木造/非木造）	2014	住宅地図・住宅土地統計等	日本全国
		築年代推定値（～81年/82～2000年/2000年～）	2014	住宅地図・住宅土地統計等	日本全国
		建物面積	2014	住宅地図	日本全国
		建物階数	2014	住宅地図 ※6	日本全国
		建物の居住者（年齢・性別）情報	2010	マイクロ人口統計 ※7	日本全国
②狭域立地情報 （住宅周辺の立地環境に関する情報）	生活利便施設 ※1	各種生活利便施設の分布 （スーパーマーケット、大規模小売店舗、コンビニエンスストア、ドラッグストア、各種飲食店、総合病院、一般診療所、小学校、中学校、郵便局それぞれの分布）	2015	テレポイントデータ ※8	日本全国
		市区町村役場・支所の分布	2014	国土数値情報	日本全国
		商店街・商業集積地域の分布	2014	商業集積統計 ※8	日本全国
	交通 ※1	主要道路（国道・都道府県道）の分布	2005	国土数値情報	日本全国
		鉄道路駅の分布	2015	国土数値情報	日本全国
		バス停の分布	2010	国土数値情報	日本全国
	人口情報	建物周辺（徒歩圏）の年齢性別人口	2010	マイクロ人口統計 ※7	日本全国
	生活の安心・安全 ※2	警察署の分布	2012	国土数値情報	日本全国
		消防署の分布	2012	国土数値情報	日本全国
		都市公園の分布	2010	国土数値情報	日本全国
	生活環境 ※1	景観計画区域の分布	2014	国土数値情報	日本全国
		観光資源の分布	2014	国土数値情報	日本全国
③広域立地情報 （住宅が立地する地域の属性に関する情報）	立地特性	用途地域の指定有無	2011	国土数値情報	日本全国
	地域の知名度	地名のブランド力分布 ※3	2015	テレポイントデータ ※8 住宅地図 ※8	日本全国
	大都市へのアクセス性	都道府県庁所在地までの距離 ※4	2015	国土数値情報	日本全国
		三大都市圏までの距離 ※4	2015	国土数値情報	日本全国
④災害リスク情報	被災リスク情報	建物最大倒壊率（推定） ※5	2014	震災ビッグデータ ※9	日本全国
		建物最大出火率（推定） ※5	2014	震災ビッグデータ ※9	日本全国
		浸水想定区域の指定有無	2011	国土数値情報	日本全国
		津波最大浸水深の指定有無	2015	国土数値情報	日本全国 ※12
		土砂災害警戒区域の指定有無	2015	国土数値情報	日本全国 ※12
	避難可能性 ※2	避難施設の分布	2012	国土数値情報	日本全国
⑤最低保証家賃	最低保証家賃関係情報	郵便番号	2016	郵便番号地図データ ※10	日本全国
		郵便番号単位の最低保証家賃	2016	全国郵便番号別最低保証家賃査定データ ※11	日本全国 ※13

- ※1 何れも最近隣施設までの距離、徒歩圏（500m圏内）の施設数を算出した。 ※2 何れも最近隣施設までの距離を算出した。
 ※3 国土数値情報から得られた駅から10km圏内を対象に、テレポイントデータと住宅地図にその駅名が何件含まれるか、という値を計算したものの。
 ※4 各都道府県庁の代表駅までの距離（三大都市圏の場合は、東京駅、大阪駅、名古屋駅）
 ※5 防災科学技術研究所より公開されている確率的地震動予測地図から得られる最大の被害値を採用した。
 ※6 ただし2階以下は属性が含まれていないため、全て2階建とした。 ※7 主査ほかにより開発された世帯の分布と居住者の年齢性別情報を推定出来るポイントデータ。
 ※8 東京大学空間情報科学研究センター 共同研究利用システム（JoRAS）より提供。
 ※9 主査ほかにより開発された日本全国の建物の分布と建物ごとの様々な属性、また災害危険性に関する属性を持つポイントデータ。
 ※10 株式会社ゼンリンジオインテリジェンスより提供 ※11 一般社団法人 移住・住みかえ支援機構より提供
 ※12 ただし一部の都道府県では自治体からの要望により非公開となっているためデータが存在しない。
 ※13 ただし全国の都道府県庁所在地およびそれに類する規模の都市のみを対象とする。



図 2-2 様々な家賃形成要素が空間結合された建物 MGD と属性情報（東京都世田谷区の例）

低保証家賃の決定要因を説明することは殆ど出来ないものと考えられる。そこで本研究では、最低保証家賃と家賃形成要素に基づいて郵便番号単位の領域を幾つかのクラスタ（類型）に分類し、続いてそのクラスタごとに異なるモデルを作成し、クラスタごとに最低保証家賃に対して説明力の大きい説明変数を明らかにする方法を採用することにした。そのためには最低保証家賃と家賃形成要素に基づいて、郵便番号単位を幾つかのグループに分類（クラスタリング）する必要がある。

多次元データのクラスタリングには様々な手法が知られている。まず古賀ほか（2013）^{文12)}でも用いられている、階層的クラスタリングがある。ただし同手法は本研究のように分類対象となるサンプル数や次元数が大きいデータの分類には適していないとされている。

一方、サンプル数や次元数が大きいデータの分類には k-means 法^{文13)}や、同手法を改良した k-means++法^{文14)}が広く用いられている。ただしこれらの手法は事前にいくつかのクラスタに分類するかを決める必要があるため、分類前に最適な分類数を決定しておく必要がある。

この課題を解決する手法としては、x-means 法^{文15)}が知られている。x-means 法は k-means 法を拡張した分類手法であり、データに応じて最適なクラスタ数を推定出来ることが特徴である。すなわち事前にクラスタ数を決める必要がない。しかし志津ほか（2011）^{文16)}は x-means 法について「目で見て明確に判断できるようなデータは分けることができるが、そうでない場合は信頼するのに不安が残る。」と言及している。秋山ほか（2015）^{文17)}による研究でもサンプル数、次元数ともに大きいデータのクラスタリングの際に x-means を適用すると、分類結果が安定しないことが明らかとなっている。そのため本研究で扱うデータのような、もはや人間の認知を越えた目視で判断不能なレベルのビッグデータに対して x-means 法を適用することには適切ではないと言える。

そこで本研究では特にサンプル数や次元数が大きいデータのクラスタ数の決定に広く利用されている「Gap 統計量」を用いてクラスタ数を決定した。同手法はクラスタ数を 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内の距離の平均と、元データのクラスタ内の距離の平均を比較し、より密集しているクラスタ数を採用するという方法である^{文18)}。図 3-1 は表 3-1 の家賃形成要素を全て用いて算出したクラスタ数ごとの Gap 統計量である。図 3-1 における最適なクラスタ数とは、クラスタ数を増加させていく中で Gap 統計量の増加するペースが低下する場所となる。図 3-1 でそれに該当するクラスタ数は 6, 34, 42 であった。しかしクラスタ数があまりにも多い場合、結果の解釈が困難となる上に、母集団が非常に小さいクラスタが多数出現することが予想されるため、本研究では 6 と 34 を採用し、k-means++

法による k=6 と k=34 のクラスタリングを実施した^{注2)}。

図 3-2 に日本全国の郵便番号単位を 6 種類のクラスタに分類した結果を示す。また図 3-3 は図 3-2 の関東地方を拡大した結果である。さらに表 3-2 に 6 分類した各クラスタの変数の平均値上位 5 種類と主な分布地域を示す。

クラスタ 0 は主要都市中心部に分布する地域であり、徒歩圏の商業集積地域の店舗数の平均値が最も影響を及ぼしている変数となっていることが特徴である。これらの地域は他の地域よりも最低保証家賃が特に高い。また周辺の生活利便性が高く、高層化が進んだ地域であった。

クラスタ 1 と 4 は何れも郊外に広く分布していた。ただしクラスタ 1 は地方都市の中心市街地とその周辺にも分布しており、徒歩圏の商業集積地域の店舗数の平均値や都道府県庁所在地までの距離が重要な変数となっていた。一方、クラスタ 4 はクラスタ 1 よりも郊外に分布しており、全郵便番号単位の地域（29, 043 地域）のうち 80.53%（23, 388 件）を占めており、ほぼ全ての地方都市の郊外に分布する、我が国において典型的な郊外型の居住地といえる。変数を見ても木造率や古い建物の割合が高い住宅地や郊外の集落であることが分かる。

クラスタ 2 はニュータウンなどの新興住宅地や、近年開発が進みつつある地域が該当する。例えば近年開発が進んでいるつくばエクスプレス沿線の地域や、大阪府の国際文化公園都市（彩都）、仙台市の錦ヶ丘ニュータウンなどが該当する。生活利便施設までの距離の大きさ、すなわちアクセスの悪さが主な変数であった。

クラスタ 3 はクラスタ 4 よりもさらに郊外の人口がまばらな地域（中山間地域、漁村、リゾート地など）が該当していた。例えば愛知県豊田市北東部や佐賀県佐賀市北部の中山間地域、志賀島などが該当した。基本的には様々な生活利便施設へのアクセスが不利な地域であった。

そしてクラスタ 5 は大都市の臨海部や、大規模小売店舗周辺に掲載された共同住宅群が該当した。例えば東京都港区台場や横浜市西区みなとみらいなどが該当した。

以上のように 6 種類の分類の場合、概ね都市部から郊外に向かってクラスタが変化していくとともに、ニュータウン（クラスタ 2）や、大都市臨海部（クラスタ 5）という特徴的なクラスタも形成された。

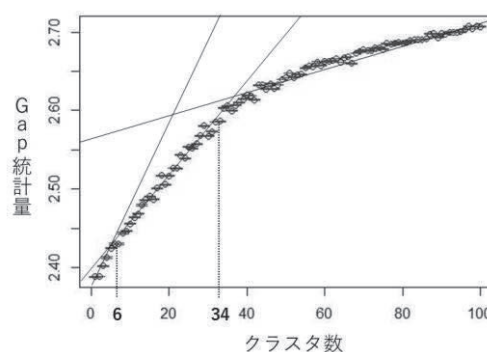


図 3-1 クラスタ数ごとの Gap 統計量

表 3-1 本研究で用いた郵便番号単位ごとの家賃形成要素

データの種類		建物建物単位に与えられた 家賃形成要素から得られる変数	データ タイプ	郵便番号単位での データ集計方法	集計後の 変数の数
大分類	小分類				
①住宅の性能	各種建物 性能情報	構造推定値（木造/非木造）	カテゴリ	木造の割合	1
		築年代推定値（～81年/82～2000年/2000年～）	カテゴリ	～81年/82～2000年/2000年～の割合	3
		建物面積	数値	相加平均	1
		建物階数	数値	相加平均	1
		建物容積	数値	相加平均	1
②狭域立地関連 情報（住宅周辺 の利便性に関する 情報）	生活利便 施設	以下の各種生活利便施設（10種類）について ・最近隣の施設までの距離 ・徒歩圏※1の施設数 各種生活利便施設： スーパーマーケット、大規模小売店舗、コンビニエンスストア、ドラッグストア、各種飲食店、総合病院、一般診療所、小学校、中学校、郵便局	数値	相加平均	20
		再近隣の市区町村役場・支所までの距離	数値	相加平均	1
		再近隣の商業集積地域までの距離 / 店舗数	数値	相加平均	2
		徒歩圏の商業集積地域の数 / それらの店舗総数	数値	相加平均	2
		再近隣の主要道路までの距離	数値	相加平均	1
	交通	再近隣の駅 / バス停までの距離	数値	相加平均	2
		徒歩圏の駅 / バス停の件数	数値	相加平均	2
	人口	徒歩圏の平均人口（推定値）	数値	相加平均	1
		徒歩圏の平均高齢化率 / 平均子供人口率 / 平均若年人口率（それぞれ推定値）※2	数値	相加平均	3
	生活環境	再近隣の警察署 / 消防署までの距離	数値	相加平均	2
		再近隣の都市公園までの距離	数値	相加平均	1
		再近隣の都市公園の面積	数値	相加平均	1
		徒歩圏の都市公園の件数	数値	相加平均	1
		徒歩圏の都市公園の合計面積	数値	相加平均	1
		景観地区指定の有無	数値	景観地区内の建物の割合	1
		準景観地区指定の有無	数値	準景観地区内の建物の割合	1
		再近隣の観光資源までの距離	数値	相加平均	1
		再近隣の観光資源の種類	カテゴリ	観光資源の種類別割合（6種類）	6
		徒歩圏の観光資源の件数	数値	相加平均	1
		郵便番号単位内の建物棟数	数値	---	1
③広域立地関連 情報（住宅が立 地する地域の属 性に関する情報）	立地特性 地域の 知名度	用途地域の指定有無	カテゴリ	用途地域の種類別（工業専用地域を除く11種類）	11
		再近隣の駅のブランド力	数値	相加平均	1
		徒歩圏の駅のブランド力合計	数値	相加平均	1
	大都市への アクセス性	関東大都市圏（東京駅）までの距離	数値	相加平均	1
		京阪神大都市圏（大阪駅）までの距離	数値	相加平均	1
		中京大都市圏（名古屋駅）までの距離	数値	相加平均	1
		都道府県庁所在地（代表駅）までの距離	数値	相加平均	1
④災害リスク情報	被災リスク 情報	建物最大倒壊率（推定）	数値	相加平均	1
		建物最大出火率（推定）	数値	相加平均	1
		浸水想定区域における浸水深	数値	相加平均	1
		津波最大浸水深	数値	相加平均	1
		土砂災害警戒区域 / 同特別警戒区域の内外判定	数値	土砂災害（特別）警戒区域内の建物の割合	2
	避難可能性	再近隣の避難施設までの距離	数値	相加平均	1
⑤最低保証家賃	最低保証家賃 関係情報	郵便番号	数値	---	1
		郵便番号単位の最低保証家賃	数値	---	1
家賃形成要素数（①+②+③+④） 合計					82

※1 各住宅から500m圏内を徒歩圏とする。他の家賃形成要素も同様。

※2 高齢化率：65歳以上人口の割合 子供人口率：15歳未満人口の割合 若年人口率：16～29歳人口の割合

同様に 34 分類を行った結果を表 3-3 に示す。概ね 6 分類のそれぞれの分類がさらに細分化されたが、ある地域にしか出現しない特徴的なクラスタが見られた。例えばクラスタ 5（山口県の太平洋側のみ）、クラスタ 6（長野県内のみ）、クラスタ 13（札幌市の郊外のみ）、クラスタ 15（札幌市の郊外を除く地域のみ）などが挙げられる。また都市近郊は説明力の高い要素の違いによって複数のクラスタに分類された。中心市街地も業務地区が繁華街かといった都市機能によって、また駅周辺が臨海部かといった地理的特性によって分類された。

4. スパースモデリングによるクラスタごとの特徴抽出

続いてクラスタごとに「どの家賃形成要素が最低保証家賃を説明しているのか」ということを明らかにする方法を検討する必要がある。すなわち複数の家賃形成要素（説明変数）のうち、どの要素が他の要素よりもより強

く最低保証家賃（目的変数）を説明しているか、ということ を明らかにするということになる。

複数の説明変数間の目的変数に対する説明力の大きさを比較する方法としては、重回帰分析が最も一般的な手法の一つとして挙げられる。これまでも様々なモデルによる重回帰分析を用いて、目的変数に対して説明力の大きい説明変数を明らかにする研究は数多く見られ、例えば線形重回帰分析を用いた空き家の発生要因の推定（金森ほか、2015）^{文 19)}や、非線形重回帰分析を用いた平均寿命の決定要因の推定（田辺・鈴木、2014）^{文 20)}などが挙げられる。ただし説明変数が複数になる回帰分析（重回帰分析）では、多重共線性を考慮する必要がある上に、本研究のように説明変数の数が非常に多い場合、有意な結果が得られなくなる場合がある^{注 3)}。

そこで本研究ではこれらの課題を解決するために、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

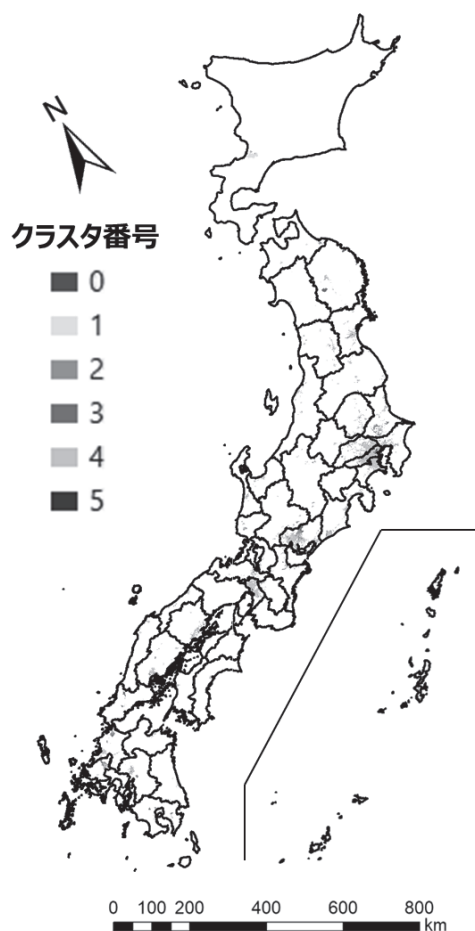


図 3-2 日本全国の郵便番号単位クラスタリング結果（6 分類）^{注 4)}

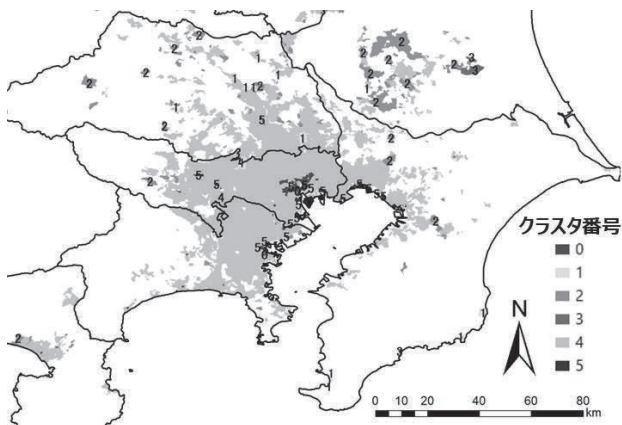


図 3-3 関東地方の郵便番号単位クラスタリング結果（6 分類）^{注 4)}

「スパースモデリング」^{文 21)} の Lasso 回帰を採用^{注 5)} した。スパースモデリング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ふ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忠実に再現することができる^{文 22)}。これは本研究の場合、表 3-1 に示すように説明変数が 82 種類にもおよぶ高次元で複雑なデータでも、より単純な予測式をあえて選ぶことで、説明力が殆ど無い説明変数を自動的に排除し、人間が認知しうるレベルに結果を落とし込むプロセスと言える。すなわち同

表 3-2 各クラスタの特徴（6 分類の場合）

クラスタ番号	該当件数	平均最低保証家賃 [万円]	各クラスタの変数の平均値上位/下位					主な分布地域
			1 位	2 位	3 位	4 位	5 位	
0	1241	5.27	徒歩圏の商業集積地域の店舗数合計の平均値	徒歩圏の各種飲食店の件数	建物階数	最近隣の商業集積地域の平均店舗数	用途地域：商業地域の割合	主要都市の中心部
1	2940	3.56	最低保証家賃	都道府県庁所在地までの距離	中京大都市圏までの距離	最近隣の小学校までの距離	京阪神大都市圏までの距離	地方都市の中心部～近郊
2	1255	3.17	最近隣の商業集積地域までの距離	最近隣の郵便局までの距離	最近隣の診療所までの距離	用途地域：工業地域の割合	最近隣の各種飲食店までの距離	ニュータウン
3	110	3.15	最近隣の公園までの平均距離	最近隣のスーパーまでの平均距離	最近隣のドラッグストアまでの平均距離	最近隣の商業集積地域までの平均距離	最近隣の診療所までの平均距離	中山間地域 漁村 リゾート地など
4	23388	3.82	木造率	建築年代～1981年の建物の割合	平均最大出火率	平均最大倒壊率	用途地域：第一種住居地域の割合	地方都市の郊外
5	109	3.08	主要道路までの平均距離	最近隣の観光資源が歴史・文化の割合	最近隣の観光資源が自然（行事・イベント）の割合	最近隣の中学校までの距離	用途地域：工業地域の割合	大都市の臨海部 大規模小売店舗周辺

表 3-3 各クラスタの特徴（33 分類の場合）

クラスタ番号	該当件数	平均最低保証家賃 [万円]	主な該当地域
0	1500	3.15	郊外（主に関東・東北地方）
1	107	3.43	ニュータウン（主に東日本）
2	3345	3.75	全国の都市近郊（クラス7より都市中心部に近くクラス21より郊外の地域）
3	166	3.28	ニュータウン（主に西日本）
4	848	5.19	中心市街地（業務地区）
5	531	3.19	郊外（山口県の太平洋側）
6	48	3.08	郊外（長野県）
7	2564	3.52	全国の都市近郊（クラス2より郊外の地域）
8	400	3.8	全国（津波被害が特に大きくなる地域）
9	337	3.91	全国の都市近郊（観光資源：都市型が充実した地域）
10	161	3.68	全国の都市近郊（準住居地域の割合が高い地域）
11	680	3.19	ニュータウン（山間部や沿岸部の比較的新興な地域）
12	852	3.67	全国の都市近郊（特に人口・建物密度が高い地域）
13	398	3.11	札幌市の郊外
14	33	5.48	臨海部・工業地帯
15	450	4.46	札幌市（郊外を除く）
16	63	3.16	中山間地域
17	747	5.05	中心市街地（若年者が特に多い地域）
18	6	3	都市公園周辺（建物数が非常に少ない地域）
19	26	3.62	全国の都市近郊（建物数が非常に少ない地域）
20	1134	5.52	全国の都市近郊（東京都・大阪府の近郊）
21	8028	3.81	全国の都市近郊（クラス2よりさらに都市中心部に近い地域）
22	1	3	奄美市（名瀬周辺）
23	322	3.21	郊外（クラス7に類似するが若者がより多い地域）
24	1514	3.75	全国の都市近郊（クラス21に類似するが若者と高齢者がより多い地域）
25	34	3	中山間地域・離島（特に特徴が無い地域）
26	1385	4	全国の都市近郊（大規模小売店舗へのアクセスが特に良い地域）
27	257	5	中心市街地（繁華街）
28	121	4.14	全国の都市近郊（総合病院へのアクセスが良い地域）
29	1215	4.09	全国の都市近郊（総合病院へのアクセスが良く若者が多い地域）
30	47	5.63	乗降客数が特に多い大規模な駅周辺の住商混合地域
31	170	3.36	全国の都市近郊（平野部のみ分布）
32	1552	3.28	全国の都市近郊（クラス21に類似し洪水のリスクが小さい地域）
33	1	3	大規模工場

手法を採用することで、クラスタごとに説明力の強い最適な説明変数を抽出してモデル構築が可能になる。そしてクラスタごとに説明力の強い説明変数を明らかにできる。その結果、地域ごとにどの家賃形成要素が最低保証家賃の決定により大きな影響を与えているのかを明らかにすることができるようになった。

以上の方法を6種類および34種類に分類したクラスタに適用し、それぞれのクラスタの特徴、すなわち最低保証家賃の決定要因となる家賃形成要素を明らかにした。

5. 結果

5.1 6分類の結果

図 5-1 に 6 分類したそれぞれのクラスタにおける家賃形成要素ごとの最低保証家賃に対する説明力を示す。

①住宅の性能：都市郊外（クラスタ 1（以下「C1」）、C4）で建物面積が広く、高層建物であるほど最低保証家賃に正に寄与する一方で、建物容積が大きいと最低保証家賃に負に寄与する傾向がみられた^{注6)}。また同じく都市郊外では木造率が高いと最低保証家賃に負に寄与した。それ以外の地域では①住宅の性能はあまり最低保証家賃に影響を及ぼしていない可能性が示された。

②狭域立地：生活利便施設の種類によって最低保証家賃に正に寄与するものとそうでないものが見られた。交通については都市部で駅へのアクセスの良さが最低保証家賃に正の寄与を示した。また高齢化率が低く子供や若年者の割合が高い地域ほど最低保証家賃が高くなった。安心・安全や環境の影響はあまり見られなかった。

③広域立地：立地特性、地域の知名度ともに最低保証家賃への影響力は殆ど見られなかった。一方、三大都市圏や都道府県庁所在地への距離はクラスタによって大きく異なり、東京大都市圏・京阪神大都市圏までの距離は、主要都市の郊外（C1）や中山間地域（C3）では正の寄与、主要都市の中心部（C0）や地方都市の郊外（C4）では負の寄与が見られた。一方、中京大都市圏の場合は、その影響がほぼ逆転した。このような現象はクラスタを 6 種類にしか分類していないため、全国の類似したクラスタがお互いに三大都市圏までの距離で影響を及ぼしあった結果であると考えられる。そのため地域性をより考慮した 34 類型での結果が必要になることが示唆された。

④災害リスク情報：何れのクラスタにおいても建物倒壊率が高いほど正に寄与する結果が得られた。このことから現在の最低保証家賃は地震による建物倒壊リスクが殆ど考慮されずに決定されている可能性があることが分かった。一方、洪水による浸水リスクは特に郊外で最低保証家賃に負の寄与をする。また津波の想定浸水深や土砂災害警戒区域の指定は最低保証家賃に殆ど影響を与えていないことが分かった。ただしこれらのファクターは日本全国の都道府県で公開されているわけではないため、今後これらに関する全国の情報を追加することで、結果が大きく変わる可能性がある。

ただし何れのクラスタも全国的に分布しており、前述したように 6 分類のクラスタでは全国の平均的な傾向は把握出来るものの、地域性は殆ど考慮できない。そこで 34 分類のそれぞれのクラスタにおいてもスパースモデリングを実施することとした。

5.2 34 分類の結果

34 分類した各クラスタをスパースモデリングしたそ

れぞれの結果から、最低保証家賃に対して正の寄与をする家賃形成要素上位 5 件と、負に寄与する家賃形成要素下位 5 件をそれぞれ抽出し、その度数分布を作成した^{注7)}（図 5-2^{注8)}。「正に寄与する回数」が大きいほど、より多くのクラスタで最低保証家賃に対して正に寄与、すなわち最低保証家賃を高くする要因となっており、逆に「負に寄与する回数」が大きいほど、最低保証家賃に対して負に寄与、すなわちすなわち最低保証家賃を低くする要因となっているといえる。

①住宅の性能：築年代の新しく建物面積が広いほど正の寄与が高くなった。また概ね築年代が古く木造率が高い場合に負の寄与が高くなるものの、正に寄与する場合も見られた。また建物階数の大きさは正よりも負に寄与することが多いことも分かった。この結果から住まい手にとっての住宅の価値を決める要素の 1 つである住宅の性能の高さは最低保証家賃の形成に寄与する地域と、そうではない地域が見られることが分かった。

②狭域立地：生活利便施設については概ね何れの施設についても施設へアクセスの良さが最低保証家賃に対して正に寄与することが分かった。ただしドラッグストアや各種飲食店、郵便局については負に寄与する場合もあることが分かった。交通については駅へのアクセスの良さは正に寄与をするものの、主要道路までの距離やバス停へのアクセスの良さは他の家賃形成要素と比べて相対的に最低保証家賃の形成にはあまり寄与しないことが分かった。特に都市の中心市街地やその近郊で駅の重要性が高かった。人口は周辺人口が多いほど負に寄与する場合が多かった。人口密度が高い人口密集地の住宅は家賃に負の影響を及ぼすものも見られる。安心・安全については警察署までのアクセスが最低保証家賃に対して正・負両方の寄与をしていた。警察署が近いことは安心・安全に寄与する一方で、昼夜間わない緊急車両の出入りによる騒音が負の影響を与えている可能性がある。環境は近隣に都市公園や観光資源があることが最低保証家賃に正に寄与することが分かった。また景観地区に指定されることが負に寄与する地域があることも分かった。景観地区は良好な景観が保証される一方で、建築物の規模・形態に規制がかかるため建物機能の改善・更新が難しくなることが原因であるものと考えられる。

③広域立地：立地特性については商業地域では最低保証家賃に正に寄与している一方で、それ以外の用途地域は正にも負にも寄与していた。特に住宅の建て方や建設可能な非住居系建物に規制があり、中心市街地から一定の距離がある第一種低層住居専用地域、第一種中高層住居専用地域、第一種住居地域や、準工業地域では負に寄

●●×●● クラスタ0（主要都市の中心部） —■— クラスタ1（地方都市の中心部～近郊） ——— クラスタ2（ニュータウン）
 ●●●●● クラスタ3（中山間地域等） ——— クラスタ4（地方都市の郊外） —◆— クラスタ5（大都市臨海部等）

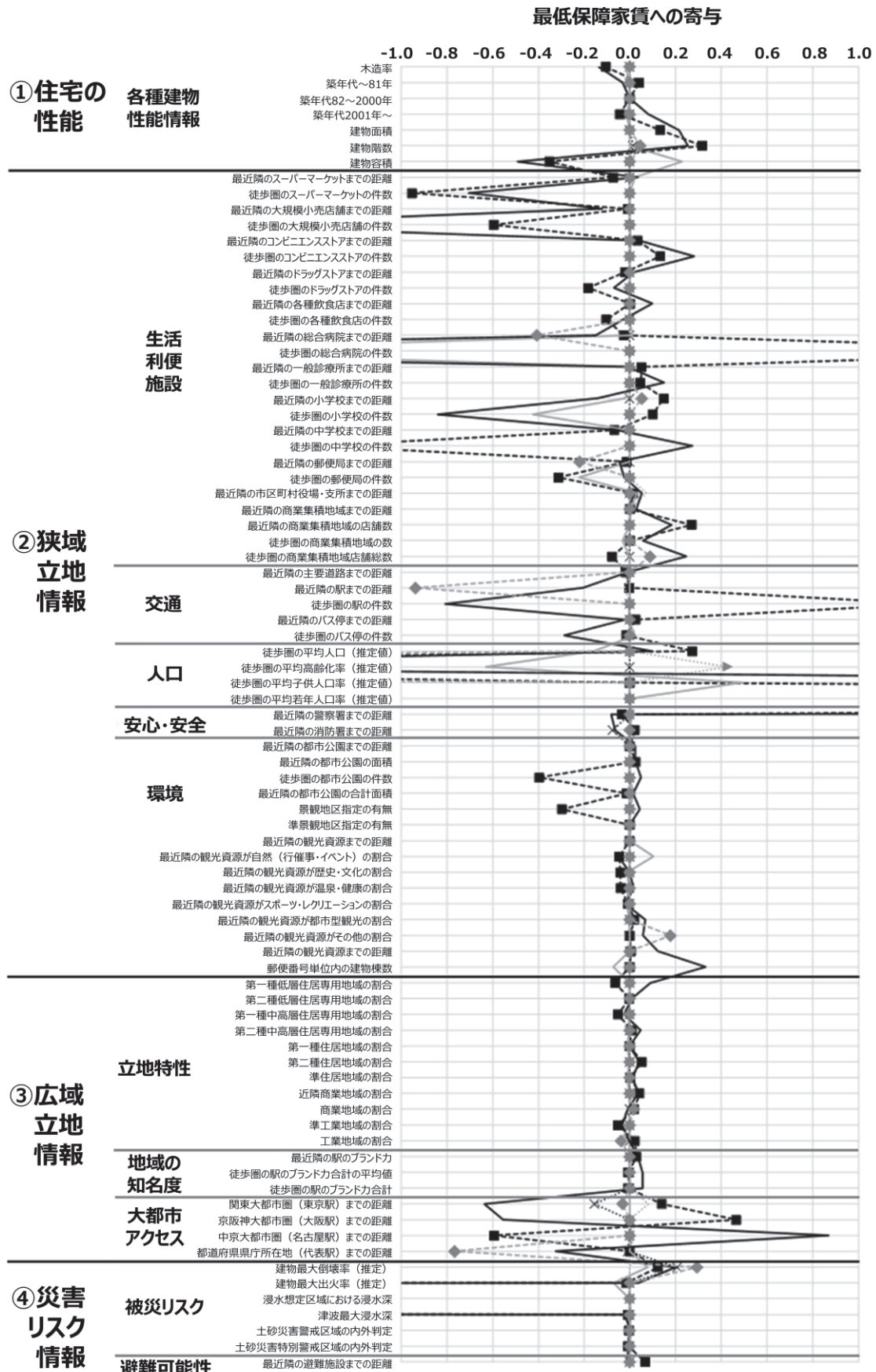


図 5-1 スパースモデリングの結果に基づくそれぞれのクラスタで最低保証家賃に対する説明力（6 分類の場合）

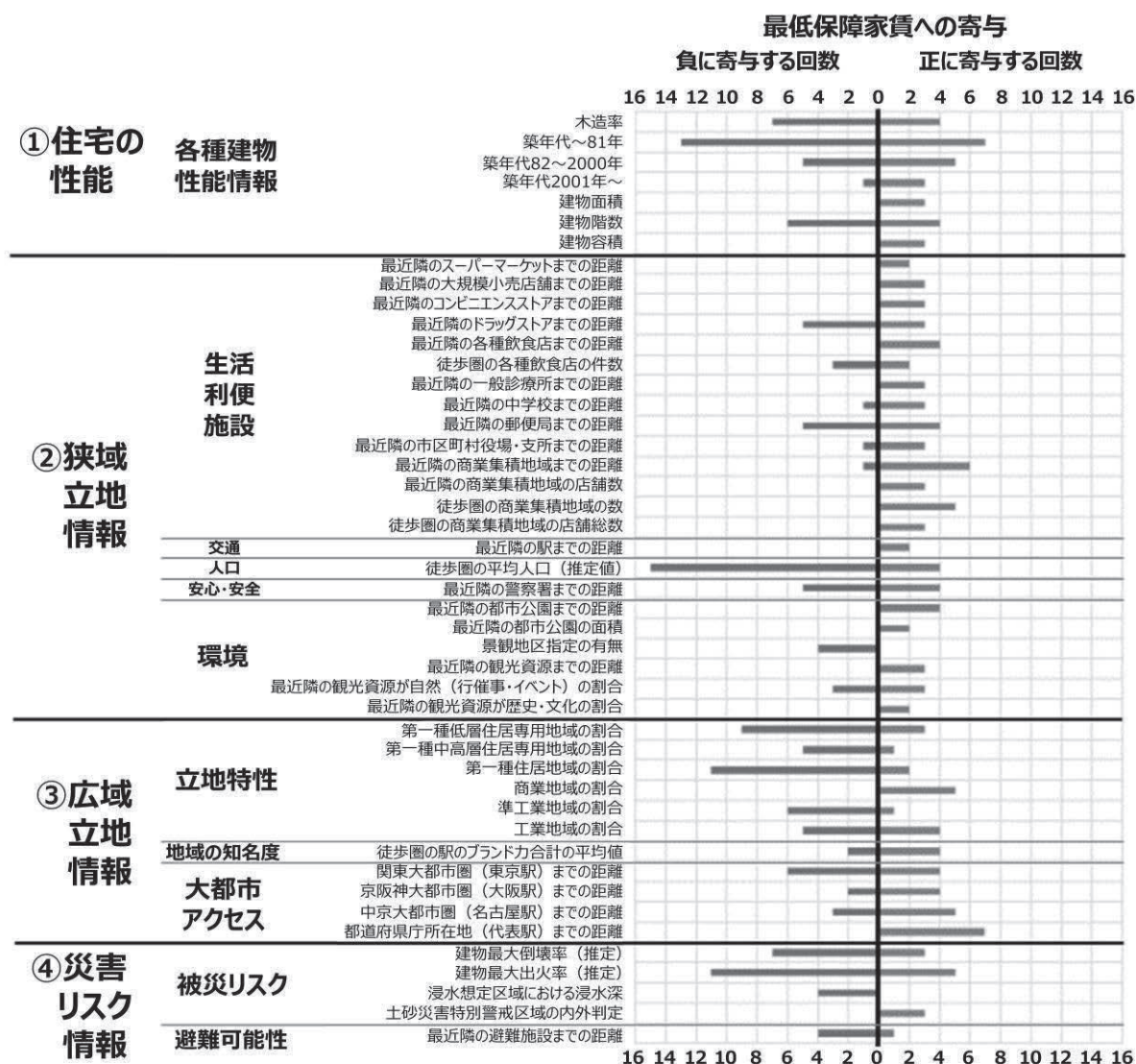


図 5-2 最低保証家賃に対して正の寄与をする家賃形成要素上位 5 件と
負に寄与する家賃形成要素下位 5 件をそれぞれ抽出した結果の度数分布（34 分類の場合）

与する傾向が高かった。地域の地名度はほぼ関東大都市圏や京阪神大都市圏の中心市街地やその郊外においてのみ正に寄与していた。大都市アクセスについては、関東大都市圏、京阪神大都市圏、中京大都市圏までの距離はそれぞれ、それらの都市圏の周辺地域では正に寄与し、それらの地域から離れるに連れて影響がないか、あるいは負に寄与した。一方、都道府県庁所在地への距離は負に寄与する地域は一つも見られなかった。これらの結果から東京や大阪では地名のブランド力や東京、大阪の各都心へのアクセス性が最低保証家賃の形成に大きく影響を与えており、また地方においても各地の道府県所在地へのアクセス性という広域立地に関する要素が家賃形成に正の影響を与えていることが分かった。

④災害リスク情報：最大倒壊率、最大出火率ともに最低保証家賃に負に寄与する場合と正に寄与する場合があった。すなわち地震に伴うリスクが評価されている地域と、

そうではない地域が混在する現状が明らかになった。例えば東京や大阪の近郊では、最大倒壊率や最大出火率が高いにも関わらず、最低保証家賃に正に寄与する地域が広く広がっていた。これは地震による被災リスクが存在するにも関わらず、最低保証家賃が高く査定されていることを意味している。一方、洪水による浸水想定区域については負のみに寄与しており、洪水のリスクは家賃に反映されていると言える。土砂災害特別警戒区域は正にのみ寄与している。これは土砂災害のリスクが存在する地域は同時に斜面地に開発された住宅地であり、自然と眺望に恵まれているため、土砂災害のリスクよりも住宅地としての価値が反映されたものと考えられる。避難所についてもアクセスが良い方が負に寄与するケースが多くなっている。以上のように災害リスクについては、洪水以外はあまり考慮されておらず、特に大都市の近郊ではそのリスクが無視されがちなが分かった。

5.3 分析結果のまとめ

以上の分析結果をまとめると、以下の通りである。

- ・住まい手にとっての価値（①住宅の性能、②狭域立地情報）について：「①住宅の性能」の高さは最低保証家賃の形成に正に寄与する地域と、そうではない地域が見られ、必ずしも高性能な住宅の家賃が高くなるとは限らないことが分かった。「②狭域立地情報」は概ね生活利便施設へのアクセスの良さが最低保証家賃に正に寄与した。また人口密度が低く都市公園や観光地への近さも正に寄与した。さらに景観地区に指定されると負に寄与した。
- ・③広域立地情報について：立地特性（用途地域）は指定される用途地域によって異なる。また地域の知名度は大都市の中心市街地やその郊外においてのみ正に寄与することが分かった。さらに三大都市圏と都道府県庁所在地までのアクセス性も正に寄与した。
- ・④災害リスクについて：地震による倒壊・火災、土砂災害のリスク、また避難所へのアクセス性はあまり考慮されていない。ただし洪水のリスクは最低保証家賃に負に寄与した。津波のリスクは正にも負にも現れなかった。

本報告書の紙面の都合上、1つ1つのクラスタの詳細な結果については割愛せざるを得ないが、本研究で開発した手法と全国データを用いることで、地域ごとに是正すべき家賃形成要素を推定・把握する技術が実現した。

6. おわりに

本研究では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、さらにそれらに家賃形成要素と最低保証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郵便番号単位に集計した後にクラスタリングし、さらに各クラスタでスパースモデリングを実施することで、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ことが可能になった。

その結果、地域により違いはあるものの、全国的には家賃形成において住宅の性能の高さや、地震・津波のリスクが必ずしも考慮されず、大規模な都市では地域の知名度が影響を与えていた。これは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状況であり、今後の我が国の住宅価値形成においては是正が求められる要素といえるだろう。

なお本研究をさらに有用なものにしていくためには、課題も数多い。まず本研究では戸建てと共同住宅の区別や、居住者のタイプ（ファミリー、単身者など）、住宅の詳細な性能（間取りや各種設備など）に関する情報は考慮されていない。今後はこれらの情報を収集する方法を考える必要がある。また家賃形成要素の更なる充実も必

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、また景観や日当たりといった情報も家賃形成要素となりうる。さらにある地域で特有に効くローカル性の高い家賃形成要素^{注9)}も存在するかもしれない。

そしてマイクロジオデータの強みを最大限に発揮するためにも、将来的には非集計、すなわち建物単位の家賃データを用いた分析の実施を考えたい。これは不動産情報企業との協力で実現出来る可能性がある。また本研究の海外展開も視野に入れている。例えば大韓民国^{注10)}では、我が国よりも空間的にも時間的にも高精細なマイクロジオデータや、統計データ、不動産関連データが整備されており、大韓民国において本研究の手法を適用可能であることが分かった。今後海外調査をさらに進めることで海外展開の可能性も広く調査したいと考えている。

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、学術的に挑戦的かつ社会的要請が高い課題挑んだ。また本研究の手法とデータの発展性・拡張性の高さは、我が国の今後の住宅政策の立案を支援できるものと考えている。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、一般社団法人 移住・住みかえ支援機構（JTI）から「全国郵便番号別最低保証家賃査定データ」の提供を受けました。また JTI 代表理事の大垣尚司先生からは本研究を進める上で、数多くのアドバイスを頂きました。さらに株式会社ゼンリンジオインテリジェンスから「郵便番号地図データ」の提供を、本研究を実施する上で必要な様々な空間情報データ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. 122 の一環として提供を受けました。加えて大韓民国の空間情報に関する聞き取り調査でご協力頂いた、韓国国土研究院国土情報研究本部の Youngjoo Lee センター長と同研究院のメンバーの皆様、また金沢市での現地調査および、札幌市への電話聞き取り調査に応じて下さった現地の不動産会社の皆様に、ここに記して謝意を表します。

<注>

- 1) ～1981年（新耐震基準前）、1982年～2000年（新耐震基準後かつ木造住宅における建築基準法改正前）2001年～（木造住宅における建築基準法改正後）の3区分に分類した。
- 2) 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ている。これは Gap 統計量の計算や、スパースモデリングを行う際も同様である。
- 3) クラスタを構成する郵便番号単位の数（サンプルサイズ）よりも説明変数の数のほうが多いクラスタ（サンプルサイズ<説明変数の数）では、全ての説明変数を用いた回帰分析を行うこ

とは不可能となる（例：クラスタ 14, 18 など）。

4) 以下の URL において図 3-2 および図 3-3 のカラー版と、Web GIS による 6 分類及び 34 分類のクラスターデータの閲覧が可能。
<http://akiyama-lab.jp/yuki/jusoken/jusoken2018.html>

5) スパースモデリングの Lasso 回帰を用いる場合、多数の説明変数の中から目的変数に対して潜在的に説明力が高い少数の説明変数を取り出すことが出来る。

6) 建物容積が大きいと最低保証家賃に負に寄与する理由は、都市郊外（C1, C4）に多数分布する大規模な公団住宅や、昭和時代に大規模開発されたマンション群によるものと考えられる。

7) 34 分類のスパースモデリングの結果を図 5-1 のように示すと図が煩雑になり視認性が失われるためにこのような結果抽出を行った。また 34 分類のクラスターに対するスパースモデリングの場合、概ね上位下位 5 位以内の家賃形成要素を用いることで、それぞれのクラスターの最低保証家賃をほぼ説明できる。

8) 図 5-2 では複数箇所のクラスターにおいて上位下位 1~5 位として出現した家賃形成要素のみ集計した。

9) 例えば現地調査を行った金沢市では最近隣の道路の消雪パイプの有無、また不動産店への電話聞き取り調査を行った札幌市では駐車場のロードヒーティングの有無が家賃形成要素の 1 つとなることが分かった。

10) 韓国国土研究院国土情報研究本部を訪問し（2017 年 9 月 4~5 日）、大韓民国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。

<参考文献>

- 1) 都市開発協会：中高層住宅の価格とサラリーマン世帯の年収との乖離，労務事情，Vol. 37, No. 969, pp. 72~78, 都市開発協会，2000. 4
- 2) Home's Press：資産価値とは立地である，http://www.homes.co.jp/cont/press/opinion/opinion_00029/。
- 3) 浅見泰司，近藤英心：建物名称に含まれる地名の分布による地区ブランド力の分析，地理情報システム学会講演論文集，Vol. 10, pp. 39~43, 地理情報システム学会，2001. 10
- 4) 林 良嗣，鈴木康弘〔編著〕，秋山祐樹ほか〔分担執筆者〕：レジリエンスと地域創生，明石書店，2015. 03
- 5) 奥西一夫：過去の地震と津波の被害，国土問題，Vol. 74, pp. 23~25, 国土問題研究会，2013. 1
- 6) 秋山祐樹：GIS を活用した時空間ビッグデータ分析—マイクロデータで見る我が国の現在と将来—，不動産研究，Vol. 59, No. 1, pp. 18~28, 一般財団法人日本不動産研究所，2017. 1
- 7) 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，Vol. 42, pp. 86~97, 都市住宅学会，2003. 8
- 8) 浅田義久：経済学で都市住宅問題をどこまで実証できるか，都市住宅学，Vol. 62, pp. 100~104, 都市住宅学会，2008. 8
- 9) 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, Vol. 19, pp. 297-308, Taylor & Francis Online, 2015. 10

10) 小川芳樹：マイクロジオデータを利用した大規模地震・津波被害リスクの推定とリスク分布から見た災害想定シナリオの重要度評価に関する研究，博士論文，2016 年東京大学博士論文，2016, 9

11) Akiyama, Y., Takada, T. and Shibasaki, R.: Development of Micropopulation Census through Disaggregation of National Population Census. CUPUM2013 conference papers, #110, CUPUM2013, 2013. 7

12) 古賀陽子，本間里見，位寄和久，友松裕希：熊本市における商店街の類型化と衰退要因に関する考察：商店街の活性化に関する研究，日本建築学会研究報告 九州支部，3, 計画系，Vol. 52, pp. 413-416, 日本建築学会，2013. 3

13) 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

14) 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

15) Pelleg, D. and Moore, A.: X-means: extending k-means with efficient estimation of the number of clusters, Proceedings of the 17th International Conference on Machine Learning, pp. 293-302, ICML, 2000. 6

16) 志津綾香，松田眞一：クラスター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

17) 秋山祐樹，柴崎亮介：商業集積地域の定量的分類に関する研究，第 24 回地理情報システム学会講演論文集，C-7-1, 地理情報システム学会，2015. 10

18) 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

19) 金森有子，有賀敏典，松橋啓介：空き家率の要因分析と将来推計，都市計画論文集，Vol. 50, No. 3, pp. 1017-1024, 日本都市計画学会，2015. 10

20) 田辺和俊，鈴木孝弘：サポートベクターマシンを用いた世界各国の平均寿命の決定要因の実証分析，厚生指標 Vol. 61, No. 13, pp. 23-30, 厚生労働統計協会，2014. 11

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

22) 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