

Department of Social Systems and Management

Discussion Paper Series

No. 1182

マクロ - ミクロ統合に基づく不動産賃料推定モデルの開発

by

金 春愛, 黄 嘉平, 住田 潮, 盧 韶南

Chun'ai JIN, Jiaping HUANG, Ushio SUMITA, Shaonan LU

August 2007

UNIVERSITY OF TSUKUBA

Tsukuba, Ibaraki 305-8573

JAPAN

# マクロ - ミクロ統合に基づく不動産賃料推定モデルの開発

## Development of Apartment Rent Estimation Model Based on an Integrated Macro-Micro Approach

金 春愛\*, 黄 嘉平\*, 住田 潮\*\*, 盧 韶南\*

Chun'ai Jin\*, Jiaping Huang\*, Ushio Sumita\*\*, Shaonan Lu\*

\* : 筑波大学大学院 システム情報工学研究科 社会システム工学専攻

\*\* : 筑波大学大学院 システム情報工学研究科

**Abstract** Since apartment rents involve sophisticated interactions among the general features of the apartment and the environmental factors, it is extremely difficult to estimate them accurately. The purpose of this paper is to develop a novel algorithm for estimating apartment rents based on an integrated macro-micro approach. A macro model based on the linear regression is combined with a micro model evaluating the land value of the apartment and the value of the nearest station. Numerical experiments reveal that the integrated model provides a noticeable improvement over both the macro model and the micro model when each is implemented alone.

**Keywords** Apartment rents, Estimation, Integrated macro-micro approach, Land value, Station value

## 1 はじめに

推定モデルの有用性は、一般的に、推定値の正確性そのものよりも、モデルの開示する定性的性質に負うところが大きい。しかし、需要予測に基づく発注・在庫管理や通信システムのトラフィック予測等、推定値がそのままマネジメントの意思決定に影響を与え、結果を左右する場合も少なからず存在する。推定アルゴリズム開発の観点から見ると、推定値の正確性が直接的に要求される為、後者においては遥かに困難な問題に直面することになる。賃料推定問題はこの後者に属する問題であり、個別的な物件情報に基づいて賃料を正確に推定することは極めて難しい（例えば、田原（2005）、宮ヶ原（2006）、堤、吉田、瀬谷、川口（2007）等を参照）。

賃料を決定する要因としては、専有面積、部屋数、建物構造に関する情報、築年月、間取りタイプ、設備等、物件の一般的な特性に基づく要因と、所在地や最寄駅までの距離等に依存して決まる要因とに大別される。前者については、線形推定に基づくマクロ的な方法論が有効である一方、後者については、局所的な情報を局所的に分析するミクロ的な方法論が重要となる。本論文の目的は、マクロ的な方法論とミクロ的な方法論を統合的に結び付け、正確性を目指す賃料推定アルゴリズムを開発することにある。

データとしては、『第1回 JAREFE 賃料データ分析コンペティション』においてアットホーム株式会社より提供された東京都23区540物件に関する情報を用いる。このデータを後述する方法によって5組に分割し、4組を推定アルゴリズムのパラメータを決定する為の学習データとして活用し、残る1組を推定アルゴリズムの正確性を検証するテストデータとする。これを5つの可能な組合せ全てについて繰り返し、それらの結果に基づいて最終的に推定アルゴリズムを確立することになる。

賃料推定マクロモデルの開発に当たっては、学習データに基づき、マクロ的な賃料決定要因を説明変数、賃料を被説明変数とする線形推定モデルを構築し、そのパラメータ推定を行う。説明変数は専有面積や建物構造等、物件の一般的特性を表現するものに加え、後述するようにミクロモデルから算出される『駅価値』を採用する。線形推定モデルの精度は、当然、学習データの被説明変数の値域の幅や分散の大小に依存し、それらが小さければ小さいほど正確性が向上すると期待できる。この観点から、データの全体集合を5（賃料価格帯）×3（地域）=15のカテゴリーに分割し、それぞれのカテゴリーについて個別にパラメータ推定を行うこととする。

賃料推定ミクロモデルの中核は、局所的な環境情報に基づき、物件の位置する場所の土地価値を算出するアルゴリズムにある。物件の土地価値は局所的に賃料推定値

を決定することを可能にし、このミクロ的賃料推定値は、当該物件が前述したマクロモデルのどのカテゴリーに所属するかを特定する。一方、ある駅を最寄駅とする物件集合について、土地価値の平均と分散を算出し、『駅価値』を定義することができる。この情報を、上述したように、付加的な説明変数として対応するカテゴリーのマクロモデルで活用し、学習データより線形推定モデルのパラメータが決定される。

テストデータから与えられる物件情報に対し、まず、ミクロモデルを実行し、ミクロ的賃料推定値を決定する。この値を用いて、当該物件がマクロモデルのどのカテゴリーに属するかを特定し、対応する線形推定モデルのパラメータ値によってマクロ的賃料推定値が算出される。このミクロ的賃料推定値とマクロ的賃料推定値を、学習データに基づいて決定される最適加重を用いて線形結合し、最終的な賃料推定値とする。

本論文の構成は以下の通りである。第2章で用いられるデータを記述し、第3章でマクロ分析が必要となるカテゴリーを導入する。第4章ではミクロ分析の詳細を述べ、第5章でマクロ分析を論じ、以上に基づき、第6章でマクロ・ミクロモデルの統合化についてまとめる。更に、4組の学習データと1組のテストデータの可能な組合せ全てについて賃料推定アルゴリズムを実施した結果に基づき、各カテゴリーごとの線形推定モデルのパラメータ値を最終的に決定する方法を論じる。第7章で、与えられた賃料データに基づく数値結果を報告し、最後に、第8章で全体のまとめと今後の研究の方向性について述べる。

## 2 データ概要

本論文の分析に用いられるデータは、2006年1月から5月の間、アットホーム株式会社に登録されている東京都23区内の居住用賃貸物件に関する賃料データである。総データ件数は540件で、各区に均等に配分されている。各物件に関する情報は：1) 直接的に金銭に係るもの：賃料（円/月）、礼金・敷金・保証金の月数、管理費（円/月）等；2) 物件の所在場所に係るもの：所在地、沿線名称、最寄駅、最寄駅までの距離に関する情報等；3) 物件の構造に係るもの：専有面積（平米）、部屋数、建物構造に関する情報、築年月等、の3つに大別される。与えられた情報の内、実際の分析に用いられる情報とその記法を表1から表3にまとめておく。

第1章で述べたように、本論文の目的は、マクロモデルとミクロモデルを統合的に接合し、正確性を期する賃

表1 所与データ1

直接的に金銭に係る情報	
$RP_i$	賃料
$PC_i$	初期費用の月数、ここでは礼金、敷金、保証金の月数の和
$OC_i$	月当たり家賃外費用、ここでは月当たり管理費と駐車場料金の和

表2 所与データ2

物件の所在場所に係る情報	
$S$	駅の集合
$A, B$	個々の駅, $A, B \in S$
$NS_i$	物件 $i$ の最寄駅
$DS_i(A)$	物件 $i$ から駅 $A$ までの距離
$WTTS_i$	物件 $i$ の最寄駅までの徒歩時間（分）
$WTBS_i$	物件 $i$ の最寄バス停までの徒歩時間（分）
$WDTS_i$	物件 $i$ の最寄駅までの徒歩距離（m）
$WDBS_i$	物件 $i$ の最寄バス停までの徒歩距離（m）
$BTTS_i$	物件 $i$ の最寄駅までのバス乗車時間（分）

表3 所与データ3

物件の構造に係る情報	
$AR_i$	専有面積
$NR_i$	部屋数, $NR_i \in NR \stackrel{\text{def}}{=} \{1, 2, 3\}$
$T_i$	建物構造名称
$BY_i$	築年

料推定アルゴリズムを開発することにある。特に、物件の土地価値を導出しミクロ的賃料推定値を算出する点に特色があり、これを実現する為には、所与の物件データだけでは明らかに不十分である。例えば、物件の建物構造によって決まる耐久年数や工事平米単価、消費者物価指数等、外部に情報源を持つデータが必要であり、更に、物件の平米当たりの土地価値や駅価値等、本論文で独自に導入される指標・概念も重要な役割を果たす。分析に用いられる外部データと独自の指標・概念をそれぞれ表4と表5にまとめておく。

## 3 データ・カテゴリーの導入と学習・テストデータの構造化

分析対象となる賃料データは、地域的には一様に分布しており、全データ540物件をまとめて見ると、平均

表 4 外部データ

$EY_i$	物件 $i$ の耐久年数 <sup>1</sup> , $EY_i = EY(T_i)$
$SM2R_i$	物件 $i$ の構造平米単価 <sup>2</sup> , $SM2R_i = SM2R(T_i, BY_i)$
$BC_i$	物件 $i$ の原価償却後の年当たり建築費用
$CPI_t$	$t$ 年における消費者物価指数 <sup>3</sup>
$WAL$	人間の 1 分間歩く平均距離, 80 (m/分)
$BUS$	バスの 1 分間進む平均距離 <sup>4</sup> , 250 (m/分)

表 5 独自に導入される指標・概念

$EC_i$	物件 $i$ の土地価値
$\widehat{EC}_i$	調整後の物件 $i$ の平米当たり土地価値
$Z_i$	物件 $i$ の所属する地域カテゴリー, $Z_i \in \{A, B, C\}$
$I(i)$	物件 $i$ の所属する推定賃料価格帯カテゴリー
$r_i$	物件 $i$ に対する対費用収益率
$\widehat{RP}_i$	物件 $i$ の予測賃料
$\mu_{SV}(A)$	駅 $A$ を最寄駅とする物件の平米当たり平均土地価値
$\sigma_{SV}(A)$	駅 $A$ を最寄駅とする物件の平米当たり土地価値の標準偏差
$\mu_{DS}(A)$	駅 $A$ を最寄駅とする物件から駅 $A$ までの平均距離

147,154 円, 標準偏差 124,106 となっている. 指数分布は平均と標準偏差が等しいという特徴を持っているが, 図 1 に示す賃料データの 2 万円刻みのヒストグラムを見ると, 指数分布にかなり近いことが分かる. しかし, 本論文では推定値の正確性を目指しており, 極めて大きい値域の幅 (最小値 38,000 円, 最大値 1,180,000 円, その差 1,142,000 円) を持つデータ集合に対して, ただ 1 つの線形推定モデルを当て嵌めることには明らかに無理がある. ここでは, 賃料と地域の二軸方向についてカテゴリー化することを試み, 各カテゴリー内において値域の幅と分散を減少させることを目指す.

賃料データを, 賃料と地域の二軸方向についてカテゴリー化することを考える. 賃料に関する分類は, 表 6 に示す価格帯を採用する. 各物件の属する価格帯とその所在地 (経度, 緯度) を, 図 2 に表示する. これに東京都 23 区の地図を重ね合わせ, 以下の 3 つの地域に区分する: 1)  $A$  地域: 高価格帯の物件が密集している区; 2)  $B$

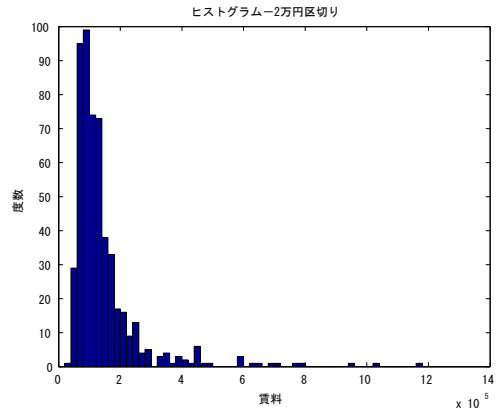


図 1 賃料データヒストグラム

表 6 賃料価格帯カテゴリー

I	100,000 円未満
II	100,000 円以上 200,000 円未満
III	200,000 円以上 300,000 円未満
IV	300,000 円以上 500,000 円未満
V	500,000 円以上

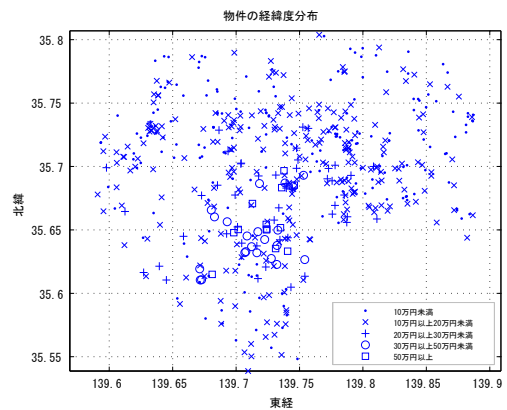


図 2 物件のプロット

地域: 中間価格帯の物件の多い区;  $C$  地域: 低価格帯の物件しかない区. 具体的には, 表 7 に与える区分が成立し, その結果を地図上に表したものが図 3 である. これら  $5 \times 3 = 15$  カテゴリーに属する物件数を表 8 に示す. 30 万円以上の価格帯に属する物件を持つのは  $A$  地域のみであり,  $C$  地域には 20 万円未満の価格帯に属する物件しか存在しないことに注意して置く. また, 各カテゴリーの値域, 平均, 標準偏差を表 9 に示すが, カテゴリー  $[V, A]$  を除くと, 賃料データ全体に比較して飛躍的に縮小していることが分かる.

開発された賃料推定アルゴリズムの精度を検証すると同時に, データサイズが十分でない点を克服する為に,

表7 地域カテゴリー

A 地域	品川, 渋谷, 新宿, 千代田, 港, 目黒
B 地域	江東, 杉並, 墨田, 世田谷, 台東, 中央, 豊島, 文京
C 地域	足立, 荒川, 板橋, 江戸川, 大田, 葛飾, 北, 中野, 練馬

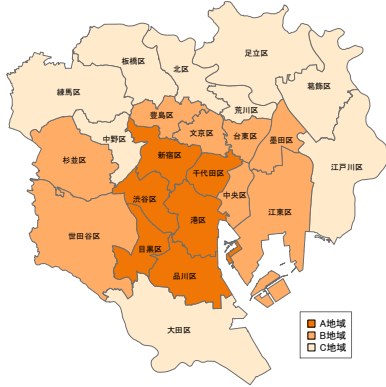


図3 地域カテゴリー分け後の東京都23区

表8 各カテゴリーに属する物件数

	A 地域	B 地域	C 地域	合計
I	25	65	115	205
II	54	103	94	251
III	29	20	0	49
IV	22	0	0	22
V	13	0	0	13
合計	143	188	209	540

全データ集合を5組に分割する。ここで、4組を推定アルゴリズムのパラメータを決定する為の学習データとして活用し、残る1組をテストデータとして検証する。更に、この手続きを可能な組合せ全てについて5回繰返し、それらの結果を統合することで最終的にアルゴリズムを確立することにする。しかし、このデータ分割を行うと、その駅を最寄駅とする物件が一件以下の駅が半数近くになり、後述する『駅価値』を算出する為に必要な情報が著しく不足することになる。この為、マイクロモデルについては全データを学習データとし、分割方式はマクロモデルのみに適用することとする。5組への分割は、各区で均等であり、更に、表8に示される各カテゴリーについても均等であることが望ましい。そこで、まず、各区の物件を乱数を用いて5組に均等分割し、各組のカテゴリ

表9 各カテゴリーの賃料値域・平均・標準偏差(単位:万)

		A 地域	B 地域	C 地域
I	R	[3.8, 9.5]	[5.0, 9.9]	[4.5, 9.9]
	M	7.76	7.66	7.51
	SD	1.24	1.15	1.43
II	R	[10.0, 19.4]	[10.0, 19.5]	[10.0, 18.7]
	M	13.65	14.25	12.53
	SD	2.81	2.68	2.08
III	R	[20.0, 29.8]	[20.0, 28.9]	-
	M	23.96	23.47	-
	SD	2.67	2.82	-
IV	R	[30.0, 47.5]	-	-
	M	39.56	-	-
	SD	5.12	-	-
V	R	[50.0, 118.0]	-	-
	M	74.85	-	-
	SD	19.70	-	-
全体	R	[3.8, 118.0], M=14.72, SD=12.41		

注) R: 値域, M: 平均, SD: 標準偏差

リー上の物件分布が著しく異なった場合には微調整を加え、最終的に表10から表14に示す分割を採用することとした。

表10 組別物件数: 1組目

	A 地域	B 地域	C 地域	合計
I	5	10	23	38
II	14	23	19	56
III	6	4	0	10
IV	2	0	0	2
V	2	0	0	2
合計	29	37	42	108

#### 4 賃料推定マイクロモデル

本章では、学習データに基づき、表1から表3に記された物件情報と表4に示される外部データを用いて、駅価値とミクロ的賃料推定値を算出するアルゴリズムを開発する。まず、賃料を構成する主要要素を吟味し、物件*i*に対する投資家の対費用期待利益率 $r_i$ を既知として賃料算出式を導く。この式で、物件*i*の賃料を所与とすると、単位面積当たりの土地価値が決定される。この結果

表 11 組別物件数：2 組目

	A 地域	B 地域	C 地域	合計
I	7	14	19	40
II	7	17	24	48
III	9	5	0	14
IV	4	0	0	4
V	2	0	0	2
合計	29	36	43	108

表 12 組別物件数：3 組目

	A 地域	B 地域	C 地域	合計
I	2	13	28	43
II	12	23	14	49
III	6	2	0	8
IV	4	0	0	4
V	4	0	0	4
合計	28	38	42	108

表 13 組別物件数：4 組目

	A 地域	B 地域	C 地域	合計
I	5	13	26	44
II	9	22	15	46
III	6	3	0	9
IV	6	0	0	6
V	3	0	0	3
合計	29	38	41	108

表 14 組別物件数：5 組目

	A 地域	B 地域	C 地域	合計
I	6	15	19	40
II	12	18	22	52
III	2	6	0	8
IV	6	0	0	6
V	2	0	0	2
合計	28	39	41	108

を用いて、駅価値に関する平均と標準偏差の推定式を提案し、更に、賃料が未知の場合でも、物件  $i$  の土地価値を推定することを可能にする方法論を確立する。

上記の土地価値推定値を賃料算出式に代入することにより、逆に、物件  $i$  の賃料を推定することができる。当然、得られた賃料推定値と学習データとして与えられた賃料は一致しない。ここで、対費用期待利益率を共有する物件集合を特定し、その集合上で相対誤差の二乗和を最小化する形で対費用期待利益率を決定し、その値と学習データとして与えられた賃料を賃料算出式に代入することにより、物件  $i$  の土地価値が最終的に推定される。次いで、推定された土地価値を用いて、駅価値に関する平均と標準偏差を改めて推定し、説明変数の 1 つとしてマクロモデルに導入することになる。

一方、テストデータから与えられる賃料未知の物件に対しては、上述した駅価値に関する平均と標準偏差の最終推定値を用いて土地価値推定モデルを起動させ、得られた土地価値の推定値と、相対誤差の二乗和を最小化する対費用期待利益率を賃料算出式に代入することにより、ミクロ的賃料推定値が確定する。この推定値により、マクロモデルにおける当該物件のカテゴリーが特定されることになる。以下、順を追って、賃料推定ミクロモデルの詳細を論じる。

#### 4.1 賃料・土地価値算出式

不動産投資家は、当然、収益を求める目的で不動産に資金を投入する。本論文では、定期的に賃料を受け取ることで収益を期待する投資者のみを考慮し、投資後すぐに売却するような投資者を想定しない。この仮定に基づいて、当該年度における物件  $i$  の賃料  $RP_i$ 、単位面積当たりの土地価値  $EC_i$ 、投資家の対費用期待利益率  $r_i$  が以下の関係式を満たすものとする。

$$RP_i \times 12 = \left\{ BC_i \cdot \frac{CPI_{2006}}{CPI_{BY_t}} + EC_i \cdot AR_i \right\} \cdot (1 + r_i) \quad (4.1.1)$$

上式の左辺は、当該年度における物件  $i$  からの賃料収入を示す。右辺は、当該年度における物件  $i$  に対する投資額に対して、利益率が  $r_i$  となることを主張している。ここで、 $BC_i$  は物件  $i$  の原価償却後の年当たり建築費用を意味し、次式で与えられる。

$$BC_i = \frac{SM2R_i \times AR_i}{BY_i} \times 0.9 \quad (4.1.2)$$

$SM2R_i$  は物件  $i$  の構造平米単価であり、注 2 に示す政府データに基づき、物件  $i$  の建築構造タイプ  $T_i$  (木造、ALC 等) と築年  $BY_i$  から計算することができる。計算

結果を表としてまとめたものを、付録1に与えて置く。定額減価償却を想定し、償却年度は物件*i*の耐久年数 $EY_i$ とする。 $EY_i$ は注1に示す政府データより、物件*i*の建築構造タイプ $T_i$ によって表15のように定まる。償却年度における残存価値は築年時の建設費の10%とする。

(4.1.1) 式の説明に戻ると、建築費用 $BC_i$ は築年時の金銭価値に基づいており、消費者物価指数を用いて当該年度の購買力に変換している。右辺の第2項は、当該年度における物件*i*の土地価値を表しており、建築費用と足し合わせることにより、投資額の当該年度における価値となる。

表15 建物構造とその耐久年数<sup>1</sup>

構造	法定耐久年数	実耐久年数
木造	22年	33年
ALC <sup>5</sup>	27年	41年
鉄骨造	27年	41年
RC <sup>6</sup>	47年	71年
SRC <sup>7</sup>	47年	71年
PC <sup>8</sup>	47年	71年

物件*i*の賃料 $RP_i$ は学習データによって与えられており、もし投資家の対費用期待利益率 $r_i$ が既知であるとするならば、(4.1.1) 式を単位面積当たりの土地価値 $EC_i$ について解くことにより、

$$EC_i = \frac{\frac{RP_i \times 12}{1+r_i} - BC_i \times \frac{CPI_{2006}}{CPI_{BY_i}}}{AR_i} \quad (4.1.3)$$

を得る。

#### 4.2 駅価値推定モデルと土地価値・賃料推定式

駅の集合を $S$ とし、駅 $A \in S$ を最寄駅とする物件の集合を

$$N(A) = \{i \mid NS_i = A\}$$

と定義する。更に、 $N(A)$ に属する物件に関し、(4.1.3) 式で求められる土地価値によって、駅 $A$ の価値に関する平均と標準偏差を

$$\mu_{SV}(A) = \frac{1}{|N(A)|} \sum_{i \in N(A)} EC_i \quad (4.2.1)$$

$$\sigma_{SV}(A) = \sqrt{\frac{\sum_{i \in N(A)} (EC_i - \mu_{SV}(A))^2}{|N(A)| - 1}} \quad (4.2.2)$$

と定める。但し、任意の有限集合 $X$ に対して、 $|X|$ は要素の数を表わす。

$|N(A)| = 1$ の場合は、 $A$ を最寄駅とする物件は1件しかなく、標準偏差は0となる。しかし、テストデータに $A$ を最寄駅とする物件が現れる可能性があり、推定の根拠として標準偏差0を放置することは好ましくない。このような事態を避ける為、駅価値に関する標準偏差が0ではない駅の集合から、可能であれば $A$ を含む最も小さい凸包を構成する駅を3つ選び、それが不可能である場合は、その集合に属する駅で $A$ に最も近い2つの駅を選ぶ。こうして選ばれた駅の集合を $J(A)$ とし、 $P_B$ と $\mu_2(B)$ を

$$P_B = \frac{DS_B(A)}{\sum_{C \in J(A)} DS_C(A)}, \quad B \in J(A)$$

$$\mu_2(B) = \frac{1}{|N(B)|} \sum_{i \in N(B)} EC_i^2$$

と定義する。ここで、 $DS_B(A)$ は駅 $B$ から駅 $A$ までの距離を表わし、Google Earthを用いて求めた。集合 $N(A)$ がただ1つの物件を含む場合、 $P_B$ と $\mu_2(B)$ を用いて $\mu_{SV}(A)$ と $\sigma_{SV}(A)$ を以下の通りに定める。

$$\mu_{SV}(A) = \sum_{B \in J(A)} P_B \cdot \mu_{SV}(B) \quad (4.2.3)$$

$$\sigma_{SV}(A) = \sqrt{\sum_{B \in J(A)} P_B \cdot \mu_2(B) - \mu_{SV}^2(A)} \quad (4.2.4)$$

$\mu_{SV}(A)$ と $\sigma_{SV}(A)$ が与えられたとき、賃料を知ることなく得られる物件*i*の土地価値推定値 $\widehat{EC}_i$ を、次式によって定める。

$$\widehat{EC}_i = \mu_{SV}(A) + \left\{ 1 - \frac{DS_i(A)}{\mu_{DS}(A)} \right\} \cdot \sigma_{SV}(A) \quad (4.2.5)$$

ここで、 $DS_i(A)$ は物件*i*とその最寄駅 $A$ との距離(m)であり、所与データに在る徒歩あるいはバスによる所要距離もしくは所要時間より、以下の式で計算される。

$$DS_i(A) = (WTTS_i + WTBS_i) + WAL \times (WDTS_i + WDBS_i) + BUS \times BTTS_i \quad (4.2.6)$$

また、 $\mu_{DS}(A)$ は $N(A)$ に属する物件の駅 $A$ までの距離の平均である。

(4.2.5) 式の意味を考えると、物件*i*の土地価値の基準値は最寄駅の駅価値に関する平均 $\mu_{SV}(A)$ によって与えられ、 $N(A)$ に属する物件の $A$ 駅までの平均距離に対して、物件*i*がより近い場合は土地価値は $\sigma_{SV}(A)$ に比例して高まり、遠ければ逆に $\sigma_{SV}(A)$ に比例して低下することになる。ここで、投資家の対費用期待利益率 $r_i$

が与えられたとすると、(4.2.5) 式から求まる  $\widehat{EC}_i$  を (4.1.1) 式に代入することにより、以下の賃料推定式を得る。

$$\widehat{RP}_i = \left\{ BC_i \cdot \frac{CPI_{2006}}{CPI_{BY_t}} + \widehat{EC}_i \cdot AR_i \right\} \frac{1+r_i}{12}$$

#### 4.3 対費用期待利益率推定モデル

これまでの議論によって、学習データと投資家の対費用期待利益率  $r_i$  が与えられた時、賃料推定値  $\widehat{RP}_i$  を導出することができる。この値は、当然、学習データで与えられる賃料  $RP_i$  とは異なる。そこで、対費用期待利益率を共有する物件集合を考え、その集合上で  $\widehat{RP}_i$  と  $RP_i$  の相対誤差の二乗和を最小にするべく、対費用期待利益率を決定することを考える。

対費用期待利益率は、物件の部屋数 (1,2,3) と地域カテゴリー ( $\mathcal{A}, \mathcal{B}, \mathcal{C}$ ) によって定まると仮定する。すなわち、部屋数が同じで同じ地域に属する物件は、全て同じ対費用期待利益率を共有するものとする。これを数学的に記述すると、 $k \in NR = \{1, 2, 3\}$  と  $\mathcal{Z} \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$  に対し、

$$ER(k, \mathcal{Z}) = \{i \mid NR_i = k, \mathcal{Z}_i = \mathcal{Z}\}$$

で定まる物件集合  $ER(k, \mathcal{Z})$  を考え、

$$i, j \in ER(k, \mathcal{Z}) \Rightarrow r_i = r_j = r(k, \mathcal{Z})$$

と仮定することになる。この仮定の下では、各  $(k, \mathcal{Z})$  の組合せについて、相対誤差の二乗和を最小にする対費用期待利益率  $r^*(k, \mathcal{Z})$  を求めれば十分である。すなわち、 $\widehat{RP}_i$  が対費用期待利益率  $r$  の関数であることを強調する為に  $\widehat{RP}_i(r)$  と書くこととすると、

$$r^*(k, \mathcal{Z}) = \underset{r \geq 0}{\operatorname{argmin}} \left\{ \sqrt{\frac{\sum_{i \in ER(k, \mathcal{Z})} \left( \frac{\widehat{RP}_i(r) - RP_i}{RP_i} \right)^2}{|ER(k, \mathcal{Z})|}} \right\} \quad (4.3.1)$$

を求める。以上の議論を、簡潔に表 16 にまとめて置く。

表 16 対費用期待利益率

$NR_i$	$\mathcal{A}$ 地域	$\mathcal{B}$ 地域	$\mathcal{C}$ 地域
1	$r^*(1, \mathcal{A})$	$r^*(1, \mathcal{B})$	$r^*(1, \mathcal{C})$
2	$r^*(2, \mathcal{A})$	$r^*(2, \mathcal{B})$	$r^*(2, \mathcal{C})$
3	$r^*(3, \mathcal{A})$	$r^*(3, \mathcal{B})$	$r^*(3, \mathcal{C})$

#### 4.4 土地価値と駅価値の最終推定値とマクロモデルへの接合

学習データが与えられた時、前節までの議論に基づいて、 $k \in NR = \{1, 2, 3\}$  と  $\mathcal{Z} \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$  の各組合せについて対費用期待利益率  $r^*(k, \mathcal{Z})$  が求まる。これらの値と学習データから得られる賃料  $RP_i$  を用いて、物件  $i$  の土地価値は最終的に以下のように推定される。

$$EC_i^* = \frac{\frac{RP_i \times 12}{1+r^*(NR_i, \mathcal{Z}_i)} - BC_i \times \frac{CPI_{2006}}{CPI_{BY_t}}}{AR_i}$$

次いで、駅価値に関する平均と標準偏差も、 $EC_i^*$  を用いて

$$\mu_{SV}^*(A) = \frac{1}{|N(A)|} \sum_{i \in N(A)} EC_i^* \quad (4.4.1)$$

$$\sigma_{SV}^*(A) = \sqrt{\frac{\sum_{i \in N(A)} (EC_i^* - \mu_{SV}^*(A))^2}{|N(A)| - 1}} \quad (4.4.2)$$

と推定されることになる。 $|N(A)| = 1$  の場合は、(4.2.3) と (4.2.4) の手続きに準じて処置するものとする。求められた  $\mu_{SV}^*(A)$  は、マクロモデルにおいて 1 つの説明変数としての役割を果たす。

テストデータから与えられる賃料未知の物件  $j$  に対しては、 $j$  の最寄駅である  $NS_j$  に対して (4.4.1) 式と (4.4.2) 式から  $\mu_{SV}^*(NS_j)$  と  $\sigma_{SV}^*(NS_j)$  を求め、土地価値の推定値  $EC_j^*$  を

$$EC_j^* = \mu_{SV}^*(NS_j) + \left\{ 1 - \frac{DS_j(NS_j)}{\mu_{DS}(NS_j)} \right\} \cdot \sigma_{SV}^*(NS_j)$$

によって定める。次いで、これらの値を用いてミクロ的賃料推定値  $RP_j^*$  を

$$RP_j^* = \left\{ BC_j \cdot \frac{CPI_{2006}}{CPI_{BY_t}} + EC_j^* \cdot AR_j \right\} \times \frac{1+r^*(NR_j, \mathcal{Z}_j)}{12}$$

と決定する。これにより、 $RP_j^*$  の属する価格帯カテゴリー  $I(j)$  が  $RP_j^* \in I(j) \in \{I, II, III, IV, V\}$  として定まり、次いでマクロモデルにおいて物件  $j$  の所属するカテゴリー  $[I(j), \mathcal{Z}_j]$  が特定されることになる。

## 5 賃料推定マクロモデル

本章では、賃料価格帯  $I \in \{I, II, III, IV, V\}$  と地域  $\mathcal{Z} \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$  に対して第 3 章で定義されたカテゴリー  $[I, \mathcal{Z}]$  を考え、各カテゴリーの内部でマクロ的に賃料を推定する線形重回帰モデルを論じる。被説明変数は賃料  $PR_i$



表 17 重回帰分析における説明変数

$AR_i$	物件 $i$ の専有面積
$T_i$	物件 $i$ の建物構造
$PC_i$	物件 $i$ の初期費用
$OC_i$	物件 $i$ の家賃外費用
$BY_i$	物件 $i$ の築年数
$NR_i$	物件 $i$ の部屋数
$DS_i(NS_i)$	物件 $i$ からその最寄駅までの距離
$\mu_{SV}^*(NS_i)$	物件 $i$ の最寄駅価値

であり、表 17 に示す説明変数を採用する。ここで、建物構造は表 15 に示されている実耐久年数に従って数値化を行う。具体的に、木造 (33 年) を 1, ALC と鉄骨造 (41 年) を 2 とし、残る RC, PC と SRC (71 年) を 3 とする。初期費用は礼金、敷金及び保証金の合計月数となっており、従って離散的な値を取る。また、家賃外費用は毎月の管理費と駐車場料金の合計金額であり、連続な値を取る変数である。物件から最寄駅までの距離は (4.2.6) 式で与えられる。また、最寄駅価値は、前章で論じた賃料推定マイクロモデルより得られた結果を用いる。これを式で表すと、

$$RP_i = \beta_0 + \beta_{AR}AR_i + \beta_T T_i + \beta_{PC}PC_i + \beta_{OC}OC_i + \beta_{BY}BY_i + \beta_{LA}LA_i + \beta_{DS}DS_i(NS_i) + \beta_{SV} \cdot \mu_{SV}^*(NS_i) + \varepsilon.$$

となる。モデル選択においては、AIC<sup>9</sup>が最小となるように選択し、また、モデルの決定係数を重視し、各変数の信頼性はある程度無視している。

賃料推定マクロモデルより得られるパラメータ値を

$$(\beta_0^{**}, \beta_{AR}^{**}, \beta_T^{**}, \beta_{PC}^{**}, \beta_{OC}^{**}, \beta_{BY}^{**}, \beta_{NR}^{**}, \beta_{DS}^{**}, \beta_{SV}^{**})$$

と定義する。このとき、賃料が未知である物件  $j$  のマクロ的賃料推定値  $RP_j^{**}$  は、次式で与えられる。

$$RP_j^{**} = \beta_0^{**} + \beta_{AR}^{**}AR_j + \beta_T^{**}T_j + \beta_{PC}^{**}PC_j + \beta_{OC}^{**}OC_j + \beta_{BY}^{**}BY_j + \beta_{LA}^{**}LA_j + \beta_{DS}^{**}DS_j(NS_j) + \beta_{SV}^{**} \cdot \mu_{SV}^*(NS_j).$$

## 6 マクロ・マイクロモデルの統合

本章では、マクロ・マイクロ統合に基づく不動産賃料推定アルゴリズムの全容をまとめる。第 4・5 章で論じたマイクロモデルとマクロモデルの概要と、その統合化を図 4 に示す。第 3 章で述べたように、マイクロモデルでは全データを学習データとし、学習データとテストデータの

分割による検証はマクロモデルのみを対象に実施されることを注意しておく。

マイクロモデルとマクロモデルは、それぞれの推定結果の線形結合を取ることによって最終的に統合されるが、その為には線形結合の係数を決定する必要がある。即ち、物件  $i$  のマイクロ的賃料推定値  $RP_i^*$  とマクロ的賃料推定値  $RP_i^{**}$  の線形結合によって、統合的賃料推定値  $RP_i$  を次式で求めるとき、その係数  $\alpha$  を特定しなければならない。

$$RP_i = \alpha \times RP_i^* + (1 - \alpha) \times RP_i^{**}$$

第 4 章で述べた対費用期待利益率  $r^*$  の推定方法に倣って、 $\alpha$  を特定することを考える。学習データにおける全物件の集合を  $L$  とし、 $\alpha$  に依存することを強調する為に統合的賃料推定値を  $RP_i(\alpha)$  と書くとする、 $RP_i(\alpha)$  と賃料データ  $RP_i$  の相対誤差の二乗を最小にするべく、

$$\alpha^* = \operatorname{argmin}_{-1 \leq \alpha \leq 2} \left\{ \sqrt{\frac{1}{|L|} \sum_{i \in L} \left( \frac{RP_i(\alpha) - RP_i}{RP_i} \right)^2} \right\}$$

によって  $\alpha^*$  を決定する。

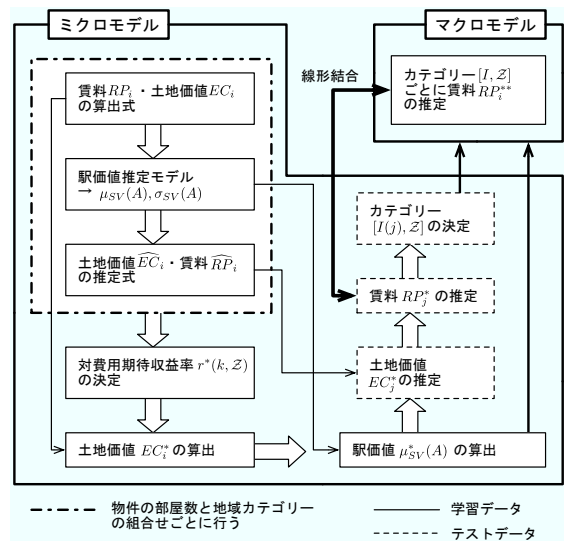


図 4 マクロ・マイクロモデルの統合化

マクロモデルに対しては、全データ集合を 5 組に分割した上で、4 組を線形モデルのパラメータ決定の為に学習データ、残る 1 組をテストデータとし、この手続きを可能な組合せ全てについて 5 回繰返し、その精度を検証している。最終的なパラメータの決定に当たっては、5 回の検証の内、相対誤差の二乗を最小にするものを、15 カテゴリーの各々について別個に決定するものとする。

## 7 数値結果

本章では、数値結果について報告する。まず、全データを学習データとして用い、マイクロモデルによって算出された対費用期待利益率の推定値  $r^*(k, Z)$  を表 18 にまとめる。高価格帯の物件が密集している  $A$  地域で最も値が低く、中間価格帯の物件の多い  $B$  地域がそれに続き、低価格帯の物件しかない  $C$  地域で最も値が高くなっている。また、各地域について、部屋数が多くなると値が低くなる。これは多額の資金を投資する不動産投資家が、利益率よりも利益の絶対額で満足する傾向を持つことを示唆する。表 18 に与えられる対費用期待利益率によって達成される (4.3.1) の評価式の最小値を各カテゴリーについてまとめたものを表 19 に示す。地域については、価格帯の高い地域ほど値は高いが、価格帯そのものについては単調性は認められない。全データに対する評価値は 0.175 と比較的高く、必ずしも満足のいくものではない。しかし、後述するように、マイクロモデルから『駅価値』を算出し、それをマクロモデルにおける線形推定モデルの説明変数に加えることによって、マクロモデルの予測精度は『駅価値』を用いない場合に比べてかなり改善されており、マイクロモデルを考慮する意義は十分に認められる。

表 18 対費用期待利益率の推定値

$NR_i$	$A$ 地域	$B$ 地域	$C$ 地域
1	0.040	0.200	0.240
2	0.001	0.060	0.050
3	0.001	0.001	0.001

表 19 賃料推定マイクロモデルの最小評価値

	$A$ 地域	$B$ 地域	$C$ 地域
I	0.212	0.153	0.191
II	0.187	0.159	0.139
III	0.177	0.168	-
IV	0.267	-	-
V	0.185	-	-
全データ	0.175		

マクロモデルのパラメータ推定結果を各カテゴリーについてまとめたものを付録 2 に与える。その内、相対誤差評価に基づく予測精度のみを抽出したものを表 20 に示す。マイクロモデルと比較して、飛躍的に良い結果を得

ていることが分かる。しかし、表 21 に与えられる『駅価値』を用いない線形推定モデルの推定値と比べてみると、『駅価値』を用いた方がかなり良い値を与えており、マイクロモデルに基づく駅価値推定の有効性を示唆している。最後に、マイクロモデルとマクロモデルの線形結合によって得られる  $\alpha^*$  と予測精度を表 22 にまとめる。マイクロモデルとマクロモデルの個別的な結果から推測されるように、マクロモデルに対する荷重が大きくなっているが、統合モデルが若干の改善を示していることは、今後の研究に希望を持たせるものである。

表 20 賃料推定マクロモデルの予測精度

	$A$ 地域	$B$ 地域	$C$ 地域
I	0.105	0.086	0.099
II	0.090	0.107	0.086
III	0.080	0.099	-
IV	0.102	-	-
V	0.084	-	-
全データ	0.09525		

表 21 賃料推定マクロモデル (駅価値未使用) の最小評価値

	$A$ 地域	$B$ 地域	$C$ 地域
I	0.179	0.103	0.118
II	0.147	0.120	0.099
III	0.076	0.102	-
IV	0.098	-	-
V	0.144	-	-
全データ	0.11804		

## 8 まとめ

本論文では、個別的な物件情報に基づいて賃料を正確に推定するという難題に対し、専有面積、部屋数等、物件の一般的な特性を説明変数とする線形推定に基づくマクロモデルと、所在地や最寄駅までの距離等、局所的な情報を局所的に分析するマイクロモデルを統合し、新たな不動産賃料推定モデルを開発した。マイクロモデルによって得られた結果は、それ自体として必ずしも満足のいくものではなかったが、マクロモデルの結果と線形結合を取ることで、それぞれの個別的結果を改善するかなり良い予測精度が得られた。また、マイクロモデルから算出される『駅価値』を説明変数として加えた線形モデルの予測精度は、それを用いない線形モデルの予測精度

表 22 統合モデル予測結果

		A 地域	B 地域	C 地域
I	$\alpha^*$	-0.124	0.191	0.027
	予測精度	0.093	0.083	0.096
II	$\alpha^*$	0.037	0.047	0.114
	予測精度	0.092	0.101	0.089
III	$\alpha^*$	0.047	0.079	-
	予測精度	0.070	0.085	-
IV	$\alpha^*$	0.064	-	-
	予測精度	0.094	-	-
V	$\alpha^*$	-0.019	-	-
	予測精度	0.099	-	-
学習データ		0.09225		
テストデータ		0.10269		
全データ		0.09415		

に比較してかなりの改善を示している。これらを総合すると、マイクロモデルは未だ改善の余地が残されている一方、マクロモデルとマイクロモデルの統合化は大きな潜在可能性を秘めていると言える。今後の研究では、(4.2.1)式から(4.2.4)式で与えられた『駅価値』推定モデルの精緻化を図り、この潜在可能性を顕在化させることを目指す。

**謝辞** 『第1回 JAREFE 賃料データ分析コンペティション』でお世話になった先生方、特に早稲田大学川口有一朗先生、明海大学前川俊一先生、筑波大学堤盛人先生、千葉商科大学吉田靖先生に心からの謝意を表明する。本研究は文部科学省科学研究費補助金（基礎研究(C)17510114）の助成を受けている。

## 注

<sup>1</sup>財務省「減価償却資産の耐用年数等に関する省令（昭和四十年大蔵省令第十五号）」

<sup>2</sup>国土交通省より毎年度公表する「建築統計年報」の「構造別：建築物の数、床面積の合計、工事費予定額」により、構造別総建築金額と総建築面積より算出した

<sup>3</sup>総務省統計局「平成17年基準消費者物価接続指数中分類指数表10大費目指数（東京都区部）-年平均指数昭和30年～平成18年（総合）」より、ただし、東京都持家の帰属家賃を除く

<sup>4</sup>公営交通事業協会「大都市におけるバスの評定速度の平均値」で与えられたものを使用する

<sup>5</sup>Autoclaved Light-weight Concrete, 軽量気泡コンクリート

<sup>6</sup>Reinforced Concrete, 鉄筋コンクリート

<sup>7</sup>Steel Reinforced Concrete, 鉄骨鉄筋コンクリート

<sup>8</sup>Prestressed Concrete, プレストレストコンクリート

<sup>9</sup>赤池情報量基準

## 参考文献

- [1] 田原拓治 (2005), 『賃料 [地代・家賃] 評価の実際』, 株式会社プログレス.
- [2] 堤盛人, 吉田靖, 瀬谷創, 川口有一郎 (2007), 「空間的相関とデータ欠損問題を考慮した不動産賃料推定モデル」, 『日本不動産金融工学学会 (JAREFE) 2007年大会予稿集』.
- [3] 宮ヶ原光正 (2006), 『新・不動産鑑定評価要説 [七訂版]』, 税務経理協会.

## 付録 1. 建物構造別平米単価

建築年	木造	ALC	鉄骨造	RC	SRC	PC
1960年	9,100	13,400	13,400	21,400	30,900	30,900
1961年	10,300	14,900	14,900	23,900	39,500	39,500
1962年	12,200	15,900	15,900	27,200	40,900	40,900
1963年	13,500	14,600	14,600	27,100	41,300	41,300
1964年	15,100	16,600	16,600	29,500	49,100	49,100
1965年	16,800	17,900	17,900	30,300	45,000	45,000
1966年	18,200	17,800	17,800	30,600	42,400	42,400
1967年	19,900	19,600	19,600	33,700	43,600	43,600
1968年	22,200	21,700	21,700	36,200	48,600	48,600
1969年	24,900	23,600	23,600	39,000	50,900	50,900
1970年	28,000	26,100	26,100	42,900	54,300	54,300
1971年	31,200	30,300	30,300	47,200	61,200	61,200
1972年	34,200	32,400	32,400	50,200	61,600	61,600
1973年	45,300	42,200	42,200	64,300	77,600	77,600
1974年	61,800	55,700	55,700	90,100	113,000	113,000
1975年	67,700	60,500	60,500	97,400	126,400	126,400
1976年	70,300	62,100	62,100	98,200	114,600	114,600
1977年	74,100	65,300	65,300	102,000	121,800	121,800
1978年	77,900	70,100	70,100	105,900	122,400	122,400
1979年	82,500	75,400	75,400	114,300	128,900	128,900
1980年	92,500	84,100	84,100	129,700	149,400	149,400
1981年	98,300	91,700	91,700	138,700	161,800	161,800
1982年	101,300	93,900	93,900	143,000	170,900	170,900
1983年	102,200	94,300	94,300	143,800	168,000	168,000
1984年	102,800	95,300	95,300	141,700	161,200	161,200
1985年	104,200	96,900	96,900	144,500	172,200	172,200
1986年	1,069,200	102,600	102,600	149,500	181,900	181,900
1987年	110,000	108,400	108,400	156,600	191,800	191,800
1988年	116,500	117,300	117,300	175,000	203,600	203,600
1989年	123,100	128,400	128,400	193,300	237,300	237,300
1990年	131,700	147,400	147,400	222,900	286,700	286,700
1991年	137,600	158,700	158,700	246,800	329,800	329,800
1992年	143,500	162,400	162,400	245,600	333,700	333,700
1993年	150,900	159,200	159,200	227,500	300,300	300,300
1994年	156,900	148,400	148,400	212,800	262,900	262,900
1995年	156,600	143,200	143,200	199,000	228,800	228,800
1996年	158,300	143,690	143,690	198,000	229,700	229,700
1997年	161,000	141,000	141,000	201,000	233,000	233,000
1998年	158,600	138,700	138,700	203,800	225,600	225,600
1999年	159,300	139,400	139,400	197,900	200,900	200,900
2000年	159,000	132,300	132,300	182,600	204,300	204,300
2001年	157,200	136,400	136,400	177,800	186,100	186,100
2002年	153,600	135,000	135,000	180,500	195,200	195,200
2003年	152,700	131,400	131,400	179,500	187,300	187,300
2004年	152,100	130,600	130,600	176,100	190,100	190,100
2005年	151,900	132,800	132,800	171,500	185,700	185,700
2006年	153,100	133,000	133,000	179,400	173,200	173,200

単位：円/平米

付録 2. マクロモデル推定・予測結果

カテゴリー	変数名	Const	AR <sup>***</sup>	T	PC <sup>***</sup>	OC	BY	NR	DS <sup>**</sup>	( $\mu_{st}$ ) <sup>***</sup>
(I,A)	Estimate	-35111.47	2295.5174	-	-	-	-397.1578	-	-17.8979	2.3525
	Std. Error	18556.2801	206.5979	-	-	-	222.492	-	6.382	0.3498
	t value	-1.892	11.111	-	-	-	-1.785	-	-2.804	6.726
	予測精度	0.104621553								
(I,B)	Estimate	15254.607	1330.11	-	2034.544	-	-355.919	-	-6.954	1.085
	Std. Error	10909.112	138.486	-	1103.102	-	114.535	-	2.608	0.255
	t value	1.398	9.605	-	1.844	-	-3.108	-	-2.666	4.256
	予測精度	0.0861945								
(I,C)	Estimate	7713.2414	1076.612	-	3382.9943	0.385	-157.744	-	-4.1703	1.1972
	Std. Error	6683.8657	69.9173	-	911.6851	0.1391	85.8435	-	1.3706	0.2343
	t value	1.154	15.398	-	3.711	2.768	-1.838	-	-3.043	5.109
	予測精度	0.098568293								
(II,A)	Estimate	-44359.6684	2381.7001	-	-	-	-548.6109	-	-22.3614	2.642
	Std. Error	16851.4548	206.7201	-	-	-	191.064	-	6.1618	0.2977
	t value	-2.632	11.521	-	-	-	-2.871	-	-3.629	8.873
	予測精度	0.090259327								
(II,B)	Estimate	-8973.6307	1750.9681	6940.7829	-	-	-1071.3775	-4799.1935	-	2.2241
	Std. Error	19687.2222	166.8067	4645.5225	-	-	202.3984	2830.3314	-	0.3765
	t value	-0.456	10.497	1.494	-	-	-5.293	-1.696	-	5.907
	予測精度	0.106997782								
(II,C)	Estimate	-11535.0897	1336.309	9443.8397	4127.843	0.1928	-859.5306	-	-9.6996	1.5733
	Std. Error	20732.618	122.8515	5106.6326	1885.9826	0.1196	194.5175	-	3.8677	0.371
	t value	-0.556	10.877	1.849	2.189	1.612	-4.419	-	-2.508	4.24
	予測精度	0.086280501								
(III,A)	Estimate	170960.222	1043.65	-	-	0.5878	-670.6829	-	-	-
	Std. Error	22706.9068	364.7107	-	-	0.2188	345.8021	-	-	-
	t value	7.529	2.862	-	-	2.686	-1.939	-	-	-
	予測精度	0.079907678								
(III,B)	Estimate	-54635.1304	1448.6294	-	-14152.2092	0.6998	-1886.3097	31854.5536	-	4.286
	Std. Error	114880.1804	579.4753	-	12353.3007	0.4097	1470.9725	28611.051	-	2.2168
	t value	-0.476	2.5	-	-1.146	1.708	-1.282	1.113	-	1.933
	予測精度	0.099407781								
(IV,A)	Estimate	236803.7	1897.1	-	-	-	-2519.4	-	-	-
	Std. Error	44266.4	512.8	-	-	-	1490.1	-	-	-
	t value	5.35	3.699	-	-	-	-1.691	-	-	-
	予測精度	0.101752522								
(V,A)	Estimate	-428612.94	3335.755	-	-	-	-	-	-	13.092
	Std. Error	159457.198	891.086	-	-	-	-	-	-	2.049
	t value	-2.688	3.743	-	-	-	-	-	-	6.389
	予測精度	0.084369214								
全データ	予測精度	0.095245								

有意水準：\*\*\*: 0.00, \*\*: 0.01, \*: 0.05, .: 0.1