

賃貸/分譲住宅価格ヘドニック回帰分析での変数選択と重共線性問題

Variable Selection and Multicollinearity in Hedonic Regression for Housing Price Analysis

刈屋武昭*, 小林裕樹**

Takeaki Kariya (thekariya70@gmail.com), Yuki Kobayashi

* 城西国際大学 国際アドミニストレーション研究科

**三井ホームエステート株式会社

賃貸/分譲住宅を市場価格に基づいてヘドニック回帰分析によるプライシングモデルを構築する過程では、多くの場合基本的な統計的な判断方式として、 t 値による変数・モデル選択が行われる。しかし、そのモデル選択過程では、各推定値の t 値は説明変数であるヘドニック属性間のマルチコ関係に大きな影響を受ける。その結果、見かけ上の有効なプライシングモデルを選択する可能性がある。この論文では、 t 値とマルチコの構造的関係を明らかにし、築年数、階建てなど、ヘドニック属性変数のマルチコ関係が起こりやすい構造を議論して、安定的なヘドニック・プライシング・モデルを定式化する枠組を構築することを狙う。

Key Words: t 値の分解 (Decomposition of t -value), 重共線性 (Multicollinearity), ヘドニック回帰 (Hedonic regression)

1 はじめに

紙幅の都合上、本稿の詳細な考え方は、刈屋・小林・清水『賃貸・分譲住宅の価格分析法の考え方と実際』プログレス社(以下 KKS(2017))を参照されたい。KKS(2017)では、賃貸/分譲住宅の市場価格をその住宅の多様な属性の価値に帰属させるヘドニック分析に基づいて、その考え方と安定的で有効な価格モデルの構築法を解説している。本稿では、その視点から、 t 値の分解による属性変数の選択問題とマルチコ(重共線性)問題を議論する。

個別物件の価格の違いを属性の違いで説明しようとするミクロ的な視点からのヘドニック回帰モデル分析では、 N 物件の(適当な変換後の)価格 y_n と定数項

を含めた属性説明変数のデータ

$(x_{0n}, x_{1n}, \dots, x_{K-1n}), (n=1, \dots, N)$

$(x_{0n} \equiv 1)$ に基づいて

$$(1.1) y_n = \beta_0 + \beta_1 x_{1n} + \dots + \beta_{K-1} x_{K-1n} + u_n$$

と定式化する。ここで誤差項の平均は 0、分散一定 σ^2 、系列相関 0 を仮定する。

(1.1) の定数項は属性ダミー変数と重共線的関係をもちやすくなる。

ヘドニック分析では、その価格を不動産の属性に帰属させて、属性の価値としての係数 β_k を最小 2 乗法で推定(推定値 $\hat{\beta}_k$)し、賃料を K 個の属性の価値の合計として表現する。例えば、 x_{1n} が駅から徒歩分であれば、推定値 $\hat{\beta}_1$ (マイナスの符号を持つ)は 1 分あたりの徒歩分属性の(マイナスの)価値(=価格)となる。このように、推定される係数の符号は事前