

ヘドニック回帰

(日本語版)

基礎概念 (Foundations) • The Valuation Engineer

Bert Craytor • 2026年 8月 3日

版 1.0.0

DOI: [10.67330/ams1t139](https://doi.org/10.67330/ams1t139)

日本語訳。人工知能の支援を得て翻訳し、編集部が確認した。英語版を正本とする。

形式的定義。 ヘドニック回帰 (hedonic regression) とは、異質財市場における観測された取引からヘドニック価格関数を推定することであり、モデルに含まれる特性の潜在価格について経験的な推定値を与えるものである。

直観的理解。 これまでの項目では、ヘドニック価格関数とその潜在価格を理論的な対象として記述してきた。関数は原理的には存在するが、実際に観測できるのは有限個の取引、すなわち取引された財の組合せと、その取引された価格だけである。ヘドニック回帰は、この両者を結ぶ経験的な橋渡しである。

設定は単純である。各取引は $(\mathbf{z}_i, p_i)$ という組を与える。すなわち、取引された財の特性ベクトルと、その売買価格である。これらの組の集まりは、(観測されない) ヘドニック価格関数から得られた一つの標本である。回帰は、その標本に関数形を当てはめ、関数を特徴づけるパラメータを推定する。パラメータが推定されれば、当てはめられた関数の偏導関数が潜在価格の推定値を与える。

回帰を実行する前に、三つのモデル化上の選択を行わなければならない。

1. どの特性をモデルに含めるか。この選択が、推定される関数の次元を決める。関連する特性を欠落させると、その効果は相関する被包含特性の係数へと漏れ出す (欠落変数バイアス)。
2. どのような関数形か。線形、対数線形、両対数、多項式、基底展開、セミパラメトリック、機械学習などがある。この選択は、推定された関数を取りうる形状と、復元しうる潜在価格のパターンとを制約する。
3. どのような推定量か。OLS、一般化最小二乗法、ロバスト M 推定量、分位点回帰、空間計量経済学的推定量、罰則付き回帰などがある。この選択は、推定量が何を最適化しているのか、そしてその推測をどのような仮定が支えているのかを決める。

いずれの選択も方法論的に実質的であり、それぞれが防御可能性に帰結を残す。

鑑定評価者が出会う場面。 ヘドニック回帰は大量評価 (mass appraisal) の基礎であり、大半の自動評価モデル (AVM) のエンジンである。郡の課税評価官が毎年 10 万筆の土地を再評価するとき、その仕事を担っているのは直近の取引に対するヘドニック回帰である。Zillow や Redfin が自動推定値を示すとき、それを生み出しているのはヘドニック型のモデルである (OLS ではなく決定木アンサンブルであることが多いが、概念的には依然としてヘドニック推定量である)。貸し手の流通市場向け AVM がある不動産について借換えの可否を判定するときにも、同じ仕組みが関与している。

個別不動産の鑑定評価もまた、ますますヘドニック回帰に依拠するようになっていく。最終的な価格の判定が取引事例比較法から得られる場合であってもそうである。近隣の標本に対する回帰から推定された潜在価格は、鑑定評価者の直観や経験則的な百分率よりもはるかに防御可能な形で、補正の大きさを裏づけることができる。回帰係数は、グリッドに適用される補正について、意見ではなく証拠となる。

なぜ防御可能性にとって重要か。 回帰に基づく推定は、ペア売買分析とは異なる防御可能性の特性を持つ。利点は次のとおりである。回帰は一度に二件ずつではなく、利用可能なすべての取引を用いる。他の多くの特性を同時に制御しながら、一つの特性の効果を分離できる。そして、それぞれの潜在価格の不確実性を定量化する標準誤差を生み出す。その代償は次のとおりである。回帰は、真の価格関数と一致しないかもしれない関数形を課す。ペア売買分析よりも多くのデータを要する。そして、その結果は標本と定式化の選択に敏感でありうる。

回帰に基づく鑑定評価では、いくつかの防御可能性上の問いが繰り返し現れる。

標本は適切か。 回帰は、標本が属する市場セグメントと期間においてヘドニック関数がどのように働くかを推定する。内陸の取引で推定した回帰を海岸沿いの対象不動産に適用することは、統計量がどれほど整って見えようとも、防御可能性の失敗である。

定式化は適切か。 線形の定式化は強い仮定である。真の関数がある閾値を超えたところで GLA に対して収穫逦減を示すのであれば、線形の当てはめはその領域で潜在価格を誤って推定することになる。診断、残差プロット、代替的な定式化は、防御可能な作業手順の一部である。

標準誤差は何を意味するか。 標準誤差は、想定された定式化のもとでの標本変動を捉える。それは定式化の誤り、欠落変数バイアス、標本選択の問題を捉えない。誤った定式化における小さな標準誤差は、正しい定式化における大きな標準誤差よりも危険である。

結果は他の手法と三角測量的に整合するか。 防御可能な回帰に基づく潜在価格は、ペア売買分析が示唆するもの、原価法の構成要素原価が含意するもの、そしてその市場に精通した鑑定評価者が信じるものと整合していなければならない。これら三つのすべてと食い違う回帰係数は、公表せよという合図ではなく、調査せよという招待である。

鑑定評価実務の例題。 ここで、八つのパシフィカ比較事例に対してヘドニック回帰の装置全体を動かすことができる。データセットは `shared/data/pacifica_comps.csv` にある。

R コード：

```
comps <- read.csv("shared/data/pacifica_comps.csv")
fit <- lm(price ~ gla + lot + view + cond, data = comps)
summary(fit)
```

出力：

```
Call:
lm(formula = price ~ gla + lot + view + cond, data = comps)

Residuals:
    Min       1Q   Median       3Q      Max
-17995   -8243   -5433   -1409   46455

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  80777.07  164509.35   0.491   0.6571
gla           574.93    109.76   5.238   0.0135 *
lot           22.70     10.38   2.187   0.1165
view        86179.03  31342.97   2.750   0.0708 .
cond        49824.33  27761.27   1.795   0.1706
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30481 on 3 degrees of freedom
Multiple R-squared:  0.9707, Adjusted R-squared:  0.9317
F-statistic: 24.86 on 4 and 3 DF, p-value: 0.01231
```

解釈。 当てはめられたモデルは、ヘドニック価格関数を次のように推定する。

$$\hat{p}(\mathbf{z}) = 80,777 + 575 z_{\text{GLA}} + 22.70 z_{\text{lot}} + 86,179 z_{\text{view}} + 49,824 z_{\text{cond}}$$

各係数は、対応する特性について推定された潜在価格である。GLA は 1 平方フィートあたり \$575 を寄与し（標準誤差が最も狭く、有意性が最も強い）、眺望はパシフィカの不動産に約 \$86,000 を寄与する。状態が一単位上がる（たとえば「良好」から「優良」へ移る）ことは、およそ \$50,000 の価値がある。

いくつかの留保がただちに見て取れる。

標本はきわめて小さい。観測値八つに対して説明変数が四つであるため、残差自由度は三つしかない。標準誤差はそれに応じて広く、敷地・眺望・状態の p 値は、経済的にもっともらしいにもかかわらず、慣行的な 0.05 の閾値に達しない。これは小標本による人為的結果であって、これらの特性が重要でないことの証拠ではない。比較事例が 8 件ではなく 80 件であれば、 p 値は劇的に小さくなるであろう。

全体としての当てはまりは強い。自由度調整済み $R^2 = 0.93$ とモデルの p 値 0.012 は、含まれた特性が標本内の価格変動の大部分を共同で説明していることを示す。強い結合的当てはまりと弱い個別的有意性との組合せは多重共線性の教科書的な兆候であり、ここではそれは構造的である。GLA、敷地、状態が標本の中で共変しているからである。

一つの残差が際立っている。最大残差は +\$46,455 であり、比較事例 H に対応する。他の七つの残差はすべて $\pm \$18,000$ の帯の中に収まる。比較事例 H の何かが、当てはめられた関数によって捉えられていない。モデルはその価格を、次に大きい残差のほぼ三倍にあたる大きさで、系統的に過小予測しているのである。

この外れ値としての残差は雑音ではない。それは、比較事例 H の観測されない特性が、含まれた変数では復元しえない仕方での価格に寄与しているという信号である。潜在変数の項目が取り上げるのは、まさにこの問いである。

ペア売買推論との照合。項目 004 のペア売買分析は、眺望と敷地だけから比較事例 A と B の間に \$127,039 の差を予測した。回帰の潜在価格はそれと同一である（その計算の出所がまさにこの潜在価格だからである）。回帰はペア売買の論理を八つの比較事例すべてに同時に一般化するが、この標本規模では本質的に新しいものを生まない。ペア分析を追認しつつ、ペア売買推論では得られない標準誤差を報告するにとどまる。

参考文献。

- Court, A. T. (1939). Hedonic Price Indexes with Automotive Examples. In *The Dynamics of Automobile Demand*. General Motors Corporation.
- Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In *The Price Statistics of the Federal Government*. National Bureau of Economic Research.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34–55.
- Triplet, J. E. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*. OECD Publishing.

相互参照：異質財 (*heterogeneous good*); 特性空間 (*characteristics space*); ヘドニック価格関数 (*hedonic price function*); 潜在価格 (*implicit price*); 潜在変数 (*latent variable*)。

付録：Prolog による形式化

以下の機械可読な補助ファイル (005-hedonic-regression.pl) は、本誌の Prolog 知識ベースにおけるこの項目の形式化である。このファイルは本文とともに、閲覧可能な Prolog ガレーとしても発行される。

```
%% 005-hedonic-regression.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 005
```

```

%% Title: Hedonic Regression
%% Author: Bert Craytor
%% Concept: hedonic_regression
%% Version: 0.1.0-draft

% =====
% DICTIONARY IMPORT
% =====

:- use_module('../tools/dictionary/dictionary').
:- dictionary_release('dictionary-2026.1').

% --- From the dictionary (loaded via :- use_module above) -----
% term(hedonic_regression, "Hedonic Regression", econometric_method,
%      "Hedonic regression is the estimation of the hedonic price
%      function from observed transactions in a heterogeneous-goods
%      market, yielding empirical estimates of the implicit prices of
%      the characteristics that enter the model.",
%      published, 'dictionary-2026.1').
%
% term(mass_appraisal, ...).
% term(automated_valuation_model, ...). % synonym: avm
% term(omitted_variable_bias, ...).
% term(multicollinearity, ...).
% term(residual, ...).
% term(ordinary_least_squares, ...).
%
% introduced_by(hedonic_regression, court_1939).
% introduced_by(hedonic_regression, griliches_1961).
% formalized_by(hedonic_regression, rosen_1974).
% depends_on(hedonic_regression, hedonic_price_function).
% depends_on(hedonic_regression, implicit_price).
% -----

:- use_terms([ hedonic_regression, hedonic_price_function, implicit_price,
               characteristics_space, heterogeneous_good, latent_variable,
               mass_appraisal, automated_valuation_model,
               omitted_variable_bias, multicollinearity, residual,
               ordinary_least_squares,
               sales_comparison_approach, paired_sales_analysis, adjustment ]).
:- use_relations([related/2, depends_on/2,
                  introduced_by/2, formalized_by/2]).

% =====
% ARTICLE SUBJECT
% =====

article_concept(hedonic_regression).
article_issue(volume(1), number(1), year(2026)).
article_id('foundations.2026.005').

% =====
% LOCAL VOCABULARY
% =====

local_term(specification).
local_term(coefficient).
local_term(standard_error).
local_term(p_value).

```

```

local_term(r_squared).
local_term(adjusted_r_squared).
local_term(residual_standard_error).
local_term(degrees_of_freedom).
local_term(f_statistic).
local_term(significance_level).
local_term(diagnostic).

% =====
% THE EMPIRICAL BRIDGE
% =====
%
% Hedonic regression is the bridge from theoretical  $p(z)$  to a sample of
% observed  $(z_i, p_i)$  pairs.

setup_of_regression(
    each_transaction_supplies_pair("(z_i, p_i): characteristics vector and sale price
    ↪"),
    sample_from("the unobserved hedonic price function"),
    fits("a functional form to that sample, estimating parameters")).

% =====
% THREE MODELING CHOICES (per the article)
% =====

modeling_choice(which_characteristics,
    "Determines the dimensions of the estimated function. Omission causes
    ↪ omitted_variable_bias.").
modeling_choice(functional_form,
    "Constrains what shape the estimated function can take. Linear, log-linear,
    ↪ polynomial, basis-expanded, semiparametric, ML.").
modeling_choice(estimator,
    "OLS, GLS, robust, quantile, spatial, penalized. Determines optimization criterion
    ↪ and inference assumptions.").

% =====
% PRACTICE CONTEXTS WHERE REGRESSION IS THE ENGINE
% =====

regression_engine_context(mass_appraisal,
    "Assessor revaluation of 100,000 parcels annually.").
regression_engine_context(automated_valuation_model,
    "Zillow/Redfin/lender AVMs; tree ensembles are conceptually hedonic estimators.").
regression_engine_context(single_property_appraisal,
    "Regression coefficients support adjustments more defensibly than appraiser
    ↪ intuition.").

% =====
% DEFENSIBILITY PROFILE: REGRESSION vs PAIRED-SALES
% =====

advantage_over_paired_sales(uses_all_data,
    "Regression uses all available sales rather than two at a time.").
advantage_over_paired_sales(simultaneous_control,
    "Isolates one characteristic while controlling for many others.").
advantage_over_paired_sales(quantified_uncertainty,
    "Produces standard errors quantifying uncertainty in each implicit price.").

tradeoff(functional_form_constraint,

```

```

"Regression imposes a form that may not match the true price function.").
tradeoff(data_requirement,
  "Requires more observations than paired-sales analysis.").
tradeoff(sample_specification_sensitivity,
  "Results can be sensitive to the choice of sample and specification.").

% =====
% RECURRING DEFENSIBILITY QUESTIONS
% =====

defensibility_question(sample_appropriateness,
  "The regression estimates the function for the sample's market segment and time
  ↳ period; mismatch is a defensibility failure regardless of statistical fit.").
defensibility_question(specification_appropriateness,
  "Linear is a strong assumption; diagnostics, residual plots, and alternative
  ↳ specifications belong in the defensible workflow.").
defensibility_question(standard_error_interpretation,
  "Standard errors capture sampling variability under the assumed specification, NOT
  ↳ misspecification, omitted-variable bias, or sample selection.").
defensibility_question(triangulation,
  "Regression coefficients should be consistent with paired-sales, cost components,
  ↳ and local appraiser opinion; isolated coefficients are an invitation to
  ↳ investigate.").

% =====
% WORKED EXAMPLE: lm(price ~ gla + lot + view + cond)
% =====

:- consult('..../kb/pacifica_comps').

% The fitted model.
fit_call("lm(formula = price ~ gla + lot + view + cond, data = comps)").
fit_estimator(ordinary_least_squares).
fit_software(r_lm).

% Fitted coefficients, with standard errors and significance from
% the R output reproduced in the article.
fitted_coefficient(intercept, 80777.07, 164509.35, 0.491, 0.6571, ' ').
fitted_coefficient(gla, 574.93, 109.76, 5.238, 0.0135, '*').
fitted_coefficient(lot, 22.70, 10.38, 2.187, 0.1165, ' ').
fitted_coefficient(view, 86179.03, 31342.97, 2.750, 0.0708, '.').
fitted_coefficient(cond, 49824.33, 27761.27, 1.795, 0.1706, ' ').

% Diagnostics from the R output.
residual_summary(min, -17995).
residual_summary(q1, -8243).
residual_summary(median, -5433).
residual_summary(q3, -1409).
residual_summary(max, 46455).

model_summary(residual_standard_error, 30481).
model_summary(degrees_of_freedom, 3).
model_summary(r_squared, 0.9707).
model_summary(adjusted_r_squared, 0.9317).
model_summary(f_statistic, 24.86).
model_summary(model_p_value, 0.01231).

% =====
% INTERPRETIVE CAVEATS FROM THE ARTICLE

```

```
% =====

caveat(tiny_sample,
  "n=8, four predictors, three residual df. Standard errors are correspondingly wide
  ↳; lot/view/cond fail to reach the 0.05 threshold despite economic plausibility.
  ↳Small-sample artifact.").

caveat(structural_multicollinearity,
  "Adj R^2 = 0.93 with model p = 0.012 indicates strong joint fit; weak individual
  ↳significance is a textbook multicollinearity signature. Here GLA, lot, and
  ↳condition covary in the sample.").

caveat(comp_h_residual,
  "Comp H's residual is +$46,455; the other seven fall within +/- $18,000. The model
  ↳systematically under-predicts Comp H by ~3x the next-largest residual. NOT noise.
  ↳Signal of an unobserved characteristic.").

% =====
% RELATING TO PAIRED-SALES (entry 004)
% =====
%
% The regression's implicit prices ARE the values used in entry 004's
% paired-sales decomposition of A vs B. The regression generalizes that
% logic to all 8 comps simultaneously while reporting standard errors
% that paired-sales reasoning cannot produce.

confirms_paired_sales_from(entry_004,
  "A vs B differential of $127,039 from view and lot alone — coefficients are the
  ↳same; regression confirms the paired analysis while adding standard errors.").

% =====
% RESIDUAL OF COMP H POINTS FORWARD TO ENTRY 006
% =====

points_forward_to(latent_variable,
  "Comp H's anomalous +$46,455 residual is the signal that an unobserved
  ↳characteristic is contributing to price. Entry 006 takes up the question.").

% =====
% PREDICTED PRICES FROM THE FITTED MODEL
% =====
%
% Computes predicted price from the fitted coefficients. The residual
% for each comp is sale_price - predicted_price.

predicted_price(P, Predicted) :-
  coord(P, gla_sqft, GLA),
  coord(P, lot_sqft, Lot),
  coord(P, view, View),
  coord(P, cond_numeric, Cond),
  fitted_coefficient(intercept, B0, _, _, _, _),
  fitted_coefficient(gla, Bg, _, _, _, _),
  fitted_coefficient(lot, Bl, _, _, _, _),
  fitted_coefficient(view, Bv, _, _, _, _),
  fitted_coefficient(cond, Bc, _, _, _, _),
  Predicted is B0 + Bg * GLA + Bl * Lot + Bv * View + Bc * Cond.

residual_of(P, R) :-
  sale_price(P, Observed),
```

```
predicted_price(P, Predicted),
R is Observed - Predicted.

coord(P, A, N) :-
  attribute_value(P, A, V),
  ( number(V) -> N = V
  ; A == view, V == yes -> N = 1
  ; A == view, V == no -> N = 0
  ; V = N
  ).

% =====
% CROSS-REFERENCES
% =====

cross_reference(heterogeneous_good).
cross_reference(characteristics_space).
cross_reference(hedonic_price_function).
cross_reference(implicit_price).
cross_reference(latent_variable).
```