

Hedonik Regresyon

Foundations • The Valuation Engineer

Bert Craytor • 7 Ağustos 2026

Sürüm 1.0.0

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

Türkçe sürüm — yapay zekâ desteğiyle hazırlanmış ve editoryal denetimden geçirilmiştir; İngilizce sürüm esas alınır.

Biçimsel tanım. Hedonik regresyon, hedonik fiyat fonksiyonunun heterojen mallar piyasasında gözlenen işlemlerden kestirilmesidir; modele giren niteliklerin örtük fiyatlarının görgül kestirimlerini verir.

Sezgisel çerçeve. Önceki maddeler hedonik fiyat fonksiyonunu ve örtük fiyatlarını kuramsal nesneler olarak betimledi. Fonksiyon ilke olarak vardır, ama uygulamada yalnızca sonlu bir işlem kümesi gözlemleriz: el değiştirmiş demetler ve el değiştirdikleri fiyatlar. Hedonik regresyon, ikisi arasındaki görgül köprüdür.

Kurulum açıktır. Her işlem bir $(\mathbf{z}_i, p_i)$ çifti sağlar: işlem gören demetin nitelik vektörü ve satış fiyatı. Bu çiftlerin toplamı, (gözlenmemiş) hedonik fiyat fonksiyonundan alınmış bir örneklemdir. Regresyon bu örneklem üzerine bir işlevsel biçim uydurur ve fonksiyonu indeksleyen parametreleri kestirir. Parametreler kestirildikten sonra, uydurulan fonksiyonun kısmi türevleri örtük fiyat kestirimlerini verir.

Regresyon çalıştırılmadan önce üç modelleme kararı verilmelidir:

1. *Modele hangi nitelikler girer?* Bu seçim, kestirilen fonksiyonun boyutlarını belirler. İlgili bir niteliğin dışlanması, onun etkisinin dahil edilen ilişkili niteliklerin katsayılarına sızmasına yol açar (dışlanmış değişken yanlılığı).
2. *Hangi işlevsel biçim?* Doğrusal, log-doğrusal, log-log, polinom, temel-genişletilmiş, yarı parametrik, makine öğrenmesi. Bu seçim, kestirilen fonksiyonun alabileceği biçimi ve geri kazanabileceği örtük fiyat örüntülerini kısıtlar.
3. *Hangi kestirici?* EKK (en küçük kareler), genelleştirilmiş en küçük kareler, dayanaklı M-kestiricileri, nicelik regresyonu, mekânsal ekonometrik kestiriciler, cezalı regresyon. Bu seçim, kestiricinin neyi eniyilediğini ve çıkarımını hangi varsayımların desteklediğini belirler.

Her seçim metodolojik olarak esastır ve her biri savunulabilirlik sonuçları taşır.

Değerleme uzmanları kavramla nerede karşılaşır. Hedonik regresyon, toplu değerlemenin temeli ve çoğu otomatik değerlendirme modelinin (ODM) motorudur. Bir ilçe değerlendirme dairesi yılda 100,000 parseli yeniden değerlediğinde, işi yapan şey yakın tarihli satışlar üzerine kurulmuş bir hedonik regresyondur. Zillow ya da Redfin otomatik bir kestirim ürettiğinde, bunu hedonik tarzda bir model (çoğu zaman EKK yerine bir ağaç topluluğu, ama kavramsal olarak yine bir hedonik kestirici) üretmektedir. Bir kreditorün ikincil piyasa ODM'si bir taşınmazı yeniden finansman için geçirdiğinde ya da ellediğinde, aynı aygıt devrededir.

Nihai değer işareti satış karşılaştırmasından gelse bile, tek taşınmaz değerlendirme de giderek daha çok hedonik regresyondan yararlanmaktadır. Bir mahalle örneklemini üzerine kurulmuş bir regresyondan kestirilen örtük fiyatlar, düzeltme büyüklüklerini değerlendirme uzmanının sezgisinden ya da kestirme yüzde kurallarından çok daha savunulabilir biçimde destekleyebilir. Regresyon katsayıları, tabloda uygulanan düzeltmeler için kanaat değil kanıt hâline gelir.

Savunulabilirlik açısından neden önemlidir. Regresyona dayalı kestirimler, eşleştirilmiş satış çözümlemesinden farklı bir savunulabilirlik profili taşır. Üstünlükleri: regresyon her seferinde iki satış yerine mevcut tüm satışları kullanır; bir niteliğin etkisini, aynı anda başka birçok niteliği denetleyerek yalıtabilir; ve her örtük fiyattaki belirsizliği niceleyen standart hatalar üretir. Ödünleştirmeleri: regresyon, gerçek fiyat fonksiyonuna uymayabilecek bir işlevsel biçim dayatır; eşleştirilmiş satış çözümlemesinden daha çok veri gerektirir; ve sonuçları örneklem ile belirleme seçimine duyarlı olabilir.

Regresyona dayalı değerlemede birkaç savunulabilirlik sorusu yinelenir:

Örneklem uygun mu? Regresyon, hedonik fonksiyonu örneklemin piyasa kesitinde ve zaman diliminde işlediği biçimiyle kestirir. İç kesim satışları üzerine kurulup sahil kesimindeki bir konu taşınmaza uygulanan bir regresyon, istatistikler ne kadar temiz görünürse görünsün bir savunulabilirlik başarısızlığıdır.

Belirleme uygun mu? Doğrusal bir belirleme güçlü bir varsayımdır. Gerçek fonksiyon belirli bir eğiğin üzerinde GLA'ya azalan getiriler sergiliyorsa, doğrusal uyum o aralıktaki örtük fiyatları yanlış kestirecektir. Tanılamalar, artık grafikleri ve alternatif belirlemeler savunulabilir iş akışının parçasıdır.

Standart hata ne anlama gelir? Bir standart hata, varsayılan belirleme altındaki örnekleme değişkenliğini yakalar. Model yanlış belirlemesini, dışlanmış değişken yanlışlığını ya da örneklem seçimini yakalamaz. Yanlış bir belirleme üzerindeki küçük bir standart hata, doğru bir belirleme üzerindeki büyük bir standart hatadan daha tehlikelidir.

Sonuç diğer yaklaşımlarla üçgenleniyor mu? Savunulabilir, regresyona dayalı bir örtük fiyat; eşleştirilmiş satış çözümlemesinin önerdiğiyle, maliyet yaklaşımı bileşen maliyetlerinin ima ettiğiyle ve piyasayı bilen değerlendirme uzmanlarının inandığıyla tutarlı olmalıdır. Bu üçüyle de çelişen bir regresyon katsayısı, yayımlamaya değil araştırmaya davettir.

Çözümlü değerlendirme örneği. Artık tüm hedonik regresyon aygıtını sekiz Pacifica emsali üzerinde çalıştırabiliriz. Veri kümesi `shared/data/pacifica_comps.csv` adresinde bulunmaktadır.

R kodu:

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

Çıktı:

```
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
```

Yorum. Uydurulan model, hedonik fiyat fonksiyonunu şöyle kestirir:

$$\hat{p}(\mathbf{z}) = 80,777 + 575 z_{\text{GLA}} + 22.70 z_{\text{arsa}} + 86,179 z_{\text{manzara}} + 49,824 z_{\text{durum}}.$$

Her katsayı, karşılık gelen niteliğin kestirilen örtük fiyatıdır. GLA fit kare başına \$575 katkı verir (en dar standart hata ve en güçlü anlamlılıkla); manzara bir Pacifica taşınmazına yaklaşık \$86,000 katar; bir birimlik ek durum (örneğin “iyi”den “mükemmel”e geçmek) kabaca \$50,000 değerindedir.

Birkaç çekince hemen göze çarpar:

Örneklem çok küçüktür. Sekiz gözlem ve dört öngörücü, geriye üç artık serbestlik derecesi bırakır. Standart hatalar buna bağlı olarak geniştir ve arsa, manzara ile durumun p -değerleri, iktisadi makullüklerine karşın alışımlı 0,05 eşiğine ulaşamaz. Bu, bu niteliklerin önemsiz olduğunun kanıtı değil, küçük örneklem yapaylığıdır. 8 yerine 80 emsalle p -değerleri çarpıcı biçimde küçülürdü.

Genel uyum güçlüdür. Düzeltilmiş $R^2 = 0,93$ ve 0,012’lik model p -değeri, dahil edilen niteliklerin örneklemdeki fiyat değişiminin çoğunu birlikte açıkladığını gösterir. Güçlü ortak uyumun zayıf bireysel anlamlılıkla birleşmesi, ders kitaplarındaki çoklu doğrusal bağlantı işaretidir; burada yapısaldır: GLA, arsa ve durum örneklemde birlikte değişir.

Bir artık öne çıkar. En büyük artık +\$46,455’tir ve H emsaline karşılık gelir. Diğer yedi artığın tamamı $\pm \$18,000$ bandına düşer. H emsaline ilişkin bir şey, uydurulan fonksiyon tarafından yakalanmamaktadır — model, onun fiyatını sistematik olarak, sıradaki en büyük artığın neredeyse üç katı kadar düşük öngörmektedir.

Bu aykırı artık gürültü değildir. H emsalinin gözlenmemiş bir niteliğinin, dahil edilen değişkenlerin geri kazanamayacağı bir biçimde onun fiyatına katkı verdiğinin işaretidir. Gizil değişken maddesi tam da bu soruyu ele alır.

Eşleştirilmiş satış akıl yürütmesiyle doğrulama. 004 numaralı maddedeki eşleştirilmiş satış çözümlemesi, A ile B emsalleri arasında yalnızca manzara ve arsadan gelen \$127,039’luk bir fark öngörmüştü. Regresyonun örtük fiyatları özdeştir (çünkü o hesabın kaynağıdır). Regresyon, eşleştirilmiş satış mantığını sekiz emsalin tamamına aynı anda genelleştirir, ama bu örneklem yüklüğünde temelde yeni bir şey üretmez; eşleştirilmiş çözümlemeyi doğrularken, eşleştirilmiş satış akıl yürütmesinin veremeyeceği standart hataları raporlar.

Kaynakça.

- 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.
- Triplett, J. E. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*. OECD Publishing.

Çapraz başvurular: heterojen mal; nitelikler uzayı; hedonik fiyat fonksiyonu; örtük fiyat; gizil değişken.

Ek: Prolog biçimselleştirmesi

Aşağıdaki makine tarafından okunabilir yardımcı dosya (005-hedonic-regression.pl) bu maddeyi derginin Prolog bilgi tabanında biçimselleştirir. Ayrıca bu makalenin yanında göz atılabilir bir dizgi nüshası (galley) olarak yayımlanır.

```

%% 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, B1, _, _, _),
    fitted_coefficient(view, Bv, _, _, _),
    fitted_coefficient(cond, Bc, _, _, _),
    Predicted is B0 + Bg * GLA + B1 * 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).
```