

特徵空間

繁體中文版

基礎概念 (Foundations) • The Valuation Engineer

Bert Craytor • 2026 年 8 月 4 日

版本 1.0.0

DOI: [10.67330/3yymh010](https://doi.org/10.67330/3yymh010)

繁體中文譯本，由人工智慧輔助翻譯並經編輯部審校；英文版為權威版本。

形式定義。一類異質商品的特徵空間 (characteristics space)，是以買方賦予價值的各項屬性為座標的向量空間；在這一空間中，每件商品由其屬性水準構成的向量來表示。

直觀闡釋。上一條目確立了異質商品是特徵的獨特組合。自然的下一個問題是：我們應當如何形式化地描述並比較這些組合？答案出自 Lancaster (1966)：把每件商品視為向量空間中的一個點，該空間的座標軸正是買方在意的各項特徵。

在這種表示下，一棟帕西菲卡 (Pacifica) 住宅不再是「帕克帕西菲卡的一棟房子」，而是一個向量，其座標包括建築面積、地塊面積、景觀、狀況、學區、距海遠近等等。若兩筆房地產在所有這些屬性上的取值都相同，它們將佔據特徵空間中的同一個點；在任何屬性上存在差異的房地產則佔據不同的點。空間的幾何結構編碼了兩筆房地產之間的接近或差異程度，並涵蓋了買方作出區分的全部維度。

空間的維數取決於所研究市場中哪些屬性是相關的。對典型的住宅估價而言，相關空間可能有十來個維度：實體維度（建築面積、地塊、臥室、衛浴、屋齡、品質、狀況）、區位維度（學區、景觀、距便利設施或不利設施的遠近）以及法律／經濟維度（使用分區、地役權）。對商用不動產估價而言，相關維度轉向現金流屬性（租約結構、租戶品質、出租率），但底層的幾何思想是相同的。

特徵空間之所以有用，在於它把實務工作者的定性直覺（「這些可比案例與估價對象相似」）轉化為定量命題（「這些可比案例在特徵空間中距估價對象很近」）。一旦定義了距離，每一個銷售比較決策都可以被當作關於空間中位置的陳述來審視。

估價師在何處遇到它。估價師無時無刻不在特徵空間中工作，只是通常沒有叫出它的名字。銷售比較網格的每一列是一條座標軸；每個可比案例是一個由座標值構成的列向量；每一項調整都是在保持其他軸不變的情況下沿一條軸移動的嘗試。

我們將在本期通篇使用的八個帕西菲卡可比案例直接演示了這一構造。每個可比案例都是由建築面積、地塊面積、景觀與狀況張成的四維特徵空間中的一個點：

可比案例	建築面積（平方英尺）	地塊（平方英尺）	景觀	狀況	成交價格
A	1,650	9,000	有	良好	\$1,425,000
B	1,650	7,200	無	良好	\$1,280,000
C	1,850	7,500	無	良好	\$1,407,500
D	1,450	7,000	無	良好	\$1,155,000
E	1,700	8,000	有	優秀	\$1,470,000
F	1,550	7,100	無	一般	\$1,177,500
G	1,700	10,500	無	良好	\$1,392,500
H	1,600	7,300	無	良好	\$1,312,500

上一條目中的可比案例 A 與 B 是前兩列；其餘六個案例對範例作了擴展，我們將在本期餘下部分充分利用這些擴展。

每個可比案例都是 $\mathbb{R}^4$ 中的一個點：

- 可比案例 A 位於 (1650, 9000, 1, 2)。
- 可比案例 B 位於 (1650, 7200, 0, 2)。
- 可比案例 G 位於 (1700, 10500, 0, 2)。

景觀與狀況座標以數值編碼（景觀取 $\{0, 1\}$ ，狀況取 $\{1, 2, 3\}$ ）；價格不是特徵空間的座標，而是空間中位置的函數。

為何關乎可辯護性。特徵空間框架使「可比案例選擇究竟是什麼」變得更加清晰。可比案例不只是「一筆相似的房地產」，而是特徵空間相關區域中靠近估價對象的一個點。可比案例在空間中距估價對象越近，所需的調整越小，得出的指示價值也越可靠。

這一重新表述暴露了非正式網格所掩蓋的若干實務問題：

- 哪些維度重要？座標軸的選擇是一個建模決策。只有當市場確實對景觀單獨定價時，把「景觀」設為一條軸才有意義；只有當某個市場區隔對「風水朝向」定價時，把它設為一條軸才有意義。可辯護的估價要求為座標軸本身辯護，而不僅僅是為軸上的取值辯護。
- 適用什麼距離度量？兩個可比案例的價格可能相差 \$200,000，因為它們在特徵空間中相距很遠；但距離的計算需要對每條軸進行尺度化。未經變換的歐氏距離把一平方英尺建築面積等同於一平方英尺地塊，而市場並不這樣看。
- 可比案例是否位於價格函數表現良好的區域？如果估價對象落在特徵空間中取樣稀疏的角落，就沒有任何可比案例靠近它，價值指示必然是外推性的。

估價實例演算。八個帕西菲卡可比案例佔據了四維特徵空間中一個相對緊緻的區域。考慮這樣一個問題：對於位於 (1,700, 7,500, 1, 2) 的假想估價對象——建築面積 1,700 平方英尺、地塊 7,500 平方英尺、有景觀、狀況良好——哪個可比案例「最近」？

如果我們不加考慮地使用座標軸未經尺度化的歐氏距離，地塊維度會佔據主導地位，因為其數值 (7,000–10,500) 遠大於景觀 (0–1) 或狀況 (1–3)。各案例到估價對象的距離如下：

可比案例	距離（未尺度化歐氏距離）	備註
A	1,501	景觀相符；地塊佔主導
B	308	景觀不符；地塊接近
C	230	景觀不符；地塊接近
D	559	景觀不符；建築面積偏差
E	502	景觀相符；狀況不同
F	422	景觀與狀況均不同
G	3,000	景觀不符；地塊相距甚遠
H	224	景觀不符；距離最近

按原始距離，可比案例 H 是最近鄰。但可比案例 H 沒有景觀，而估價對象有。樸素的度量誤導了我們：地塊維度的數值尺度擠掉了類別型的景觀維度，而後者對價格實際上更重要。

一個可辯護的距離度量應當以各條軸的隱含價格 (implicit price) 來尺度化——粗略地說，就是該軸每變動一個單位值多少美元。一旦座標軸按隱含價格尺度化，景觀（約值 \$100,000）便不再被地塊（約每平方英尺 \$25）所淹沒。於是，在景觀上與估價對象相符的可比案例 A 便浮現為恰當的首選可比案例。

這正是後續條目中發展的隱含價格工具所提供的東西：對特徵空間的一種有原則的尺度化，使「接近」從幾何上的猜測變為源自市場的計算。

參考文獻。

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2), 132–157.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34–55.

交叉引用：異質商品 (*heterogeneous good*)；特徵價格函數 (*hedonic price function*)；隱含價格 (*implicit price*)。

附錄：Prolog 形式化

以下機器可讀的配套檔案 (002-characteristics-space.pl) 在本刊的 Prolog 知識庫中對本條目進行了形式化。該檔案同時作為可瀏覽的 Prolog 版本隨本文一併發布。

```
%% 002-characteristics-space.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 002
%% Title: Characteristics Space
%% Author: Bert Craytor
%% Concept: characteristics_space
%% 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(characteristics_space, "Characteristics Space", economic_theory,
% "The characteristics space of a class of heterogeneous goods is
% the vector space whose coordinates index the attributes that
% buyers value, and in which each good is represented by the
% vector of its attribute levels.", published, 'dictionary-2026.1').
%
% term(heterogeneous_good, ...).
% term(hedonic_price_function, ...).
% term(implicit_price, ...).
%
% generalized_by(characteristics_space, lancaster_1966).
% depends_on(characteristics_space, heterogeneous_good).
% -----

:- use_terms([ characteristics_space, heterogeneous_good,
                hedonic_price_function, implicit_price,
                comp_selection, sales_comparison_approach,
                real_estate ]).
:- use_relations([related/2, depends_on/2, generalized_by/2]).

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

article_concept(characteristics_space).
article_issue(volume(1), number(1), year(2026)).
```

```

article_id('foundations.2026.002').

% =====
% LOCAL VOCABULARY
% =====
%
% Narrow article-local field names + the dimensions of the worked
% example's characteristics space.

local_term(dimension).
local_term(coordinate).
local_term(euclidean_distance).
local_term(distance_metric).
local_term(nearest_neighbor).

% =====
% FORMAL ATTRIBUTION
% =====

% The construction is due to Lancaster (1966), as noted in citations.
attributed_construction(characteristics_space, lancaster_1966).

% =====
% THE ARTICLE'S CLAIM ABOUT REPRESENTATION
% =====
%
% A heterogeneous good becomes a point in a vector space. Each axis is
% a characteristic buyers value; each good is the vector of its
% attribute levels.

representation_of(heterogeneous_good, point_in_vector_space).

% Examples of dimensions relevant to residential appraisal.
residential_dimension(gla_sqft).
residential_dimension(lot_sqft).
residential_dimension(view).
residential_dimension(cond).
residential_dimension(bedrooms).
residential_dimension(baths).
residential_dimension(age).
residential_dimension(quality).
residential_dimension(school_district).
residential_dimension(distance_to_amenity).
residential_dimension(zoning).

% Examples of dimensions relevant to commercial appraisal — included
% to make the article's point that the dimensions shift with market.
commercial_dimension(lease_structure).
commercial_dimension(tenant_quality).
commercial_dimension(occupancy).
commercial_dimension(cash_flow_profile).

% =====
% APPRAISER ENCOUNTERS
% =====
%
% Mapping appraisal practice onto the characteristics-space framework.

practice_to_space_correspondence(

```

```

    sales_comparison_grid_line, coordinate_axis,
    "A grid line is an axis.").
practice_to_space_correspondence(
    comp_row, row_vector,
    "A comp is a vector of coordinate values.").
practice_to_space_correspondence(
    adjustment, movement_along_axis,
    "An adjustment moves along one axis holding others fixed.").

% =====
% WORKED EXAMPLE
% =====
%
% The 8 Pacifica comps are points in a 4D characteristics space spanned
% by GLA, lot_sqft, view, cond_numeric. Bring in the shared dataset.

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

% Axes of the chosen space for this article's analysis.
axis(gla_sqft).
axis(lot_sqft).
axis(view).
axis(cond_numeric).

% Numeric encoding for non-numeric axes.
encode(view, no, 0).
encode(view, yes, 1).
encode(cond_numeric, 1, 1).
encode(cond_numeric, 2, 2).
encode(cond_numeric, 3, 3).

%! coord(?Property, ?Axis, ?Numeric) is nondet.
%
% Numeric coordinate of Property along Axis. Handles symbolic attribute
% values by encoding them.

coord(P, A, N) :-
    attribute_value(P, A, V),
    ( number(V) -> N = V ; encode(A, V, N) ).

%! point(?Property, -Vector) is det.
%
% Vector representation of Property in the 4D characteristics space.

point(P, [GLA, Lot, View, Cond]) :-
    coord(P, gla_sqft, GLA),
    coord(P, lot_sqft, Lot),
    coord(P, view, View),
    coord(P, cond_numeric, Cond).

% =====
% THE HYPOTHETICAL SUBJECT
% =====
%
% The article tests distance metrics against this hypothetical property.

hypothetical_subject(subject).
attribute_value(subject, gla_sqft, 1700).
attribute_value(subject, lot_sqft, 7500).
```

```

attribute_value(subject, view, yes).
attribute_value(subject, cond_numeric, 2).

% =====
% NAIVE EUCLIDEAN DISTANCE (the article's straw distance metric)
% =====
%
% Computes the unscaled Euclidean distance between Property and the
% hypothetical subject. Used to demonstrate that unscaled distances
% mislead: the lot dimension dominates because its numeric range is
% much larger than view (0..1) or cond (1..3).

naive_distance_to_subject(P, D) :-
    point(P, V_P),
    point(subject, V_S),
    sum_of_squares(V_P, V_S, S),
    D is sqrt(S).

sum_of_squares([], [], 0).
sum_of_squares([X|Xs], [Y|Ys], S) :-
    sum_of_squares(Xs, Ys, S0),
    D is X - Y,
    S is S0 + D * D.

% =====
% THE ARTICLE'S CENTRAL OBSERVATION
% =====
%
% Under naive Euclidean distance, Comp H is the nearest neighbor. But
% Comp H has no view and the subject does; view matters more to price
% than the naive metric admits. The fix (developed in entries 003-005)
% is to scale each axis by its implicit price.

claim(naive_metric_misleads,
    "Unscaled Euclidean distance crowds out small-range axes (like view) under large-
    ↪ range axes (like lot_sqft). A market-defensible distance scales each axis by its
    ↪ implicit price.").

% =====
% DEFENSIBILITY QUESTIONS RAISED
% =====
%
% Three questions the characteristics-space framing makes explicit.

defensibility_question(which_dimensions,
    "The choice of axes is a modeling decision. Defending an appraisal requires
    ↪ defending the axes, not just the values on them.").
defensibility_question(distance_metric,
    "Untransformed Euclidean distance treats one sqft of GLA as equivalent to one sqft
    ↪ of lot, which the market does not. Distance requires a scaling per axis.").
defensibility_question(extrapolation_region,
    "If the subject sits in a sparsely-sampled corner of the space, no comp is truly
    ↪ close and the indication is extrapolative.").

% =====
% INFERENCE RULES
% =====

%! same_point(?A, ?B) is nondet.

```

```
%
% Two properties occupy the same point in the characteristics space.
% (Doesn't occur in the running example, but the article notes that it
% would if all coordinates matched.)

same_point(A, B) :-
    point(A, V),
    point(B, V),
    A \== B.

%! nearest_neighbor_by_naive_distance(-P) is semidet.
%
% The comp closest to the subject under the naive Euclidean metric.
% Used to illustrate the article's point that the naive metric chooses
% Comp H, not Comp A.

nearest_neighbor_by_naive_distance(P) :-
    findall(D-Q,
        ( property(Q), Q \== subject, naive_distance_to_subject(Q, D) ),
        Pairs),
    keysort(Pairs, [_-P|_]).

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

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