

Espaço de características

Tradução para o português do verbete publicado originalmente em inglês. Tradução elaborada com assistência de IA e revisada pela redação; a versão em inglês é a autorizada.

Foundations • The Valuation Engineer

Bert Craytor • 3 de agosto de 2026

Versão 1.0.0

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

Definição formal. O espaço de características de uma classe de bens heterogêneos é o espaço vetorial cujas coordenadas indexam os atributos que os compradores valoram, e no qual cada bem é representado pelo vetor de seus níveis de atributo.

Enquadramento intuitivo. O verbete anterior estabeleceu que um bem heterogêneo é uma cesta distinta de características. A questão seguinte é natural: como descrever e comparar formalmente tais cestas? A resposta, devida a Lancaster (1966), é tratar cada bem como um ponto de um espaço vetorial cujos eixos são as características que interessam aos compradores.

Nessa representação, um imóvel residencial de Pacifica deixa de ser “uma casa em Park Pacifica” e passa a ser um vetor com coordenadas para área construída, área do terreno, vista, estado de conservação, distrito escolar, distância até o oceano e assim por diante. Dois imóveis que compartilhassem todos esses valores de atributo ocupariam o mesmo ponto no espaço de características; imóveis que diferem em qualquer atributo ocupam pontos distintos. A geometria do espaço codifica quão próximos ou quão diferentes são dois imóveis, de um modo que capta todas as dimensões sobre as quais os compradores fazem distinções.

A dimensionalidade do espaço depende de quais atributos são relevantes para o mercado em estudo. Para uma avaliação residencial típica, o espaço relevante pode ter uma dúzia de dimensões: físicas (GLA, terreno, dormitórios, banheiros, idade, padrão construtivo, estado de conservação), locacionais (distrito escolar, vista, proximidade de amenidades ou desamenidades) e jurídico-econômicas (zoneamento, servidões). Para uma avaliação de imóvel comercial, as dimensões relevantes se deslocam para atributos de fluxo de caixa (estrutura de locação, qualidade do locatário, ocupação), mas a ideia geométrica subjacente é a mesma.

O que torna o espaço de características útil é que ele converte uma intuição qualitativa do profissional (“estes elementos comparativos são parecidos com o avaliando”) em uma afirmação quantitativa (“estes elementos comparativos estão próximos do avaliando no espaço de características”). Uma vez definida a distância, toda decisão do método comparativo direto de dados de mercado pode ser examinada como uma afirmação sobre localização no espaço.

Onde os avaliadores o encontram. Os avaliadores trabalham no espaço de características o tempo todo, geralmente sem nomeá-lo. Cada linha de uma planilha de homogeneização é um eixo coordenado; cada elemento comparativo é um vetor-linha de valores de coordenadas; cada ajuste é uma tentativa de mover-se ao longo de um eixo mantendo os demais fixos.

Os oito elementos comparativos de Pacifica que usaremos ao longo desta edição ilustram a construção diretamente. Cada elemento é um ponto em um espaço de características de quatro dimensões, gerado por GLA, área do terreno, vista e estado de conservação:

Elem.	GLA (sf)	Terreno (sf)	Vista	Conserv.	Preço de venda
A	1.650	9.000	sim	bom	\$1.425.000
B	1.650	7.200	não	bom	\$1.280.000
C	1.850	7.500	não	bom	\$1.407.500
D	1.450	7.000	não	bom	\$1.155.000
E	1.700	8.000	sim	excelente	\$1.470.000
F	1.550	7.100	não	médio	\$1.177.500
G	1.700	10.500	não	bom	\$1.392.500
H	1.600	7.300	não	bom	\$1.312.500

Os elementos A e B do verbete anterior são as duas primeiras linhas; os seis restantes ampliam o exemplo de maneiras que exploraremos no restante da edição.

Cada elemento é um ponto em $\mathbb{R}^4$:

- O elemento A está em (1650, 9000, 1, 2).
- O elemento B está em (1650, 7200, 0, 2).
- O elemento G está em (1700, 10500, 0, 2).

As coordenadas de vista e de estado de conservação são codificadas numericamente (vista como $\{0, 1\}$, conservação como $\{1, 2, 3\}$); o preço *não* é uma coordenada do espaço de características, mas uma função da posição dentro dele.

Por que importa para a defensabilidade. O arcabouço do espaço de características aguça o que a seleção de dados de mercado de fato é. Um elemento comparativo não é apenas “um imóvel parecido”, mas um ponto da região relevante do espaço de características próximo do avaliando. Quanto mais próximo do avaliando estiver o elemento no espaço, menores os ajustes exigidos e mais confiável se torna o valor indicado.

Esse reenquadramento expõe várias questões práticas que a planilha informal obscurece:

- *Quais dimensões importam?* A escolha dos eixos é uma decisão de modelagem. Incluir “vista” como eixo só faz sentido se o mercado precifica a vista de forma distinta; incluir “orientação segundo o Feng Shui” só faz sentido se algum segmento de mercado a precifica. A avaliação defensável exige defender os eixos, e não apenas os valores ao longo deles.
- *Qual métrica de distância se aplica?* Dois elementos podem diferir em \$200.000 de preço porque estão distantes um do outro no espaço de características, mas o cálculo da distância exige um escalonamento para cada eixo. A distância euclidiana não transformada trata um pé quadrado de GLA como equivalente a um pé quadrado de terreno, coisa que o mercado não faz.
- *Os elementos estão em uma região em que a função de preços é bem comportada?* Se o avaliando se situa em um canto pouco amostrado do espaço de características, nenhum elemento comparativo está próximo dele e a indicação de valor é necessariamente extrapolativa.

Exemplo de avaliação resolvido. Os oito elementos comparativos de Pacífica ocupam uma região relativamente compacta do espaço de características de quatro dimensões. Considere a questão de qual elemento é o “mais próximo” de um imóvel avaliando hipotético em (1,700, 7,500, 1, 2) — 1.700 sf, terreno de 7.500 sf, com vista, em bom estado de conservação.

Se usarmos ingenuamente a distância euclidiana com eixos não escalonados, a dimensão do terreno domina, porque seus valores (7.000–10.500) são numericamente muito maiores do que os de vista (0–1) ou de conservação (1–3). Distâncias até o avaliando:

Elem.	Distância (euclidiana não escalonada)	Observações
A	1.501	vista coincide; terreno domina
B	308	vista difere; terreno próximo
C	230	vista difere; o mais próximo em distância bruta
D	559	vista difere; GLA fora
E	502	vista coincide; conservação difere
F	422	vista e conservação diferem
G	3.000	vista difere; terreno distante
H	224	vista difere; o mais próximo de todos

Pela distância bruta, o elemento H é o vizinho mais próximo. Mas o elemento H não tem vista, e o avaliando tem. A métrica ingênua nos induziu ao erro: a escala numérica da dimensão do terreno sufocou a dimensão categórica da vista, que na verdade importa mais para o preço.

Uma métrica de distância defensável escalonaria cada eixo por seu preço implícito — grosso modo, por quantos dólares vale uma unidade de variação naquele eixo. Uma vez que os eixos sejam escalonados pelo preço implícito, a vista (que vale ~\$100.000) deixa de ser dominada pelo terreno (~\$25 por pé quadrado). O elemento A, que coincide com o avaliando quanto à vista, emerge então como o elemento comparativo primário apropriado.

É exatamente isso que o maquinário de preços implícitos desenvolvido nos verbetes seguintes fornece: um escalonamento fundamentado do espaço de características, que transforma a “proximidade” de um palpite geométrico em um cálculo derivado do mercado.

Referências.

- 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.

Referências cruzadas: bem heterogêneo; função de preços hedônicos; preço implícito.

Adendo: formalização em Prolog

O material complementar legível por máquina apresentado a seguir (002-characteristics-space.pl) formaliza este verbete na base de conhecimento Prolog da revista. Ele também é publicado como composição final navegável junto com este artigo. O código não é traduzido.

```
%% 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).

```