

Valuation Engineer Journal

價值的工程學——可量測、有節制、可重現。

第 1 卷 · 第 1 期 · 2026 年夏季



本期內容

不動產估價中學理與實務的落差

Residual Constraint Approach (RCA) —— 架構與流程

earthUI、glmnetUI 與 mgcvUI——開放的建模介面

基礎概念 (F0–F6) 異質商品 · 特徵空間 ·

特徵價格函數 · 隱含價格 · 特徵價格迴歸 · 潛在變數

近期判決 美國 · 歐盟 · 拉丁美洲

Valuation Engineer Journal

第 1 卷 · 第 1 期

2026 年夏季

ISSN 3143-5033

主編 William Bert Craytor, SRA

目次

宗旨	2
社論	3
創刊號弁言	4
致編者的信	6
邀您來信	7
論文	8
不動產估價中學理與實務的落差	9
Residual Constraint Approach (RCA) —— 架構與流程	38
earthUI——earth (MARS) 套件的互動式介面	85
glmnetUI——glmnet (Lasso 與彈性網) 套件的互動式介面	125
mgcvUI——mgcv (廣義加法模型) 套件的互動式介面	158
基礎概念	193
總覽 (F0)	194
異質商品 (F1)	200
特徵空間 (F2)	207
特徵價格函數 (F3)	214
隱含價格 (F4)	220
特徵價格迴歸 (F5)	226
潛在變數 (F6)	233
近期判決	239
近期判決——美國 (R1)	240
近期判決——歐盟 (R2)	243
近期判決——拉丁美洲 (R3)	248

宗旨

《*Valuation Engineer Journal*》刊載關於價值工程的嚴謹且面向實務的研究與案例：將技術判斷與經濟結果相連的方法、模型與量測。本刊尤重可重現性、假設的公開與論述的清晰。

社論

欄目：編者的話

創刊號致辭

（繁體中文版）

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。

William Bert Craytor, SRA — 主編

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: 10.67330/0tymfa32

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 創用 CC 姓名標示 4.0 國際授權條款（CC BY 4.0）的條款發布。

摘要。第 1 卷第 1 期的編者的話：本刊為何存在，五篇文章與三個欄目各自貢獻了什麼，以及一份共襄其事的邀約。

關鍵詞：編者的話，估價工程，創刊號

歡迎閱讀《*The Valuation Engineer Journal*》（估價工程師學刊）創刊號。

本刊之所以存在，源於一道鴻溝。估價理論長期以來把估價表述為對市場行為的嚴謹測量，然而日常實務在很大程度上仍然依靠人工挑選可比案例、以及僅憑判斷估算的調整值。與此同時，那些能夠測量市場究竟為房地產特徵支付了多少的統計工具——特徵價格迴歸（hedonic regression）、現代樣條模型（spline model）、經交叉驗證的機器學習——已經成熟了數十年，卻極少被實際估價工作所採納。「估價工程師」（valuation engineer）是我們對彌合這道鴻溝的從業者的稱呼：把持證估價師的專業能力，與資料工程、統計學以及可重現的作業規程結合在一起。

本期的五篇文章奠定了工作基礎。開篇文章考察理論與實務之間的鴻溝本身，並論證本行業自身的準則早已要求比其日常方法所能提供的更多的測量。第二篇——也是本刊全部意圖的核心——完整呈現了殘差約束法（Residual Constraint Approach）：這一規程讓迴歸通道去測量可以測量的部分，把無法測量的部分規訓為一份明確的殘差分解，並對總額加以約束，使判斷只能對價值加以細化，而絕不能憑空創造價值。隨後是三篇軟件文章，記錄 earthUI 及其用於懲罰迴歸與廣義可加模型的配套介面——本刊所依託的開放原始碼工具——它們在付印前不久合併為一個應用程式，各篇文章中的出版者按對此均有說明。

三個欄目與這些文章相伴。*Foundations*（基礎）欄目以七則短文（F0–F6）開篇，梳理出一條概念鏈：從異質商品出發，經由特徵空間、特徵價格函數、隱含價格與特徵價格迴歸，直至潛藏於殘差之中的潛變數。每則短文都註明其學術淵源，擁有各自的 DOI，並以一段機器可讀的 Prolog 形式化表述收尾——這是一個知識庫的開端，我們打算逐期將其培育為本領域精確而共通的術語體系。*Recent Court Decisions*（近期法院判例）欄目綜述美國、歐盟與拉丁美洲中與估價實務相關的裁判，因為法律環境本身就是這

一工程問題的組成部分。*Letters to the Editor*（讀者來信）欄目開啟通信專欄：為使用而設計的方法，總會以其作者從未預料的方式受到檢驗，而這些發現——以及由此引發的分歧——正屬於此處。

本期的每一項內容都以 CC BY 4.0 授權開放取用出版，同時提供 PDF 與 HTML 兩種格式，配有 Crossref DOI，並在 Zenodo 上獨立保存。我們誠邀投稿——文章、Foundations 條目、來往信函以及不同意見——歡迎估價師、經濟學家與工程師共同參與。

William Bert Craytor, SRA

主編

致編者的信

欄目：讀者來信

誠邀來信

（繁體中文版）

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。

William Bert Craytor, SRA — 主編

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: 10.67330/r55mw896

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 創用 CC 姓名標示 4.0 國際授權條款（CC BY 4.0）的條款發布。

摘要。主編開啟「讀者來信」專欄：本刊希望從讀者那裡收到什麼，以及如何寄送。

關鍵詞：讀者來信，通信往來，同儕評論

一份只顧自己說話的刊物，只是半份刊物。本專欄正是為另一半而設。

The Valuation Engineer Journal（估價工程師學刊）刊發的是意在被使用的方法——而被使用的方法，總會以其作者從未預料的方式受到檢驗。如果您應用過殘差約束法（Residual Constraint Approach），並發現其中某一步在您所在的市場行不通，我們想要這樣的來信。如果某則 Foundations（基礎）條目對某個術語的界定比您的實務所需更為狹窄，我們想要這樣的來信。如果我們所綜述的某項法院判例，在該轄區的從業者讀來另有意味，我們同樣想要這樣的來信。不同意見、複現、修正與純粹的糾錯，都在本專欄的範圍之內；來自實務的簡短報告也是——一項奏效的技術、一個已被解決的資料難題、一次值得提醒他人的失敗。

規矩不多。來信應當署名、簡明（一頁為宜）、有的放矢——請註明您所針對的文章、章節或條目，並像在工作底稿中那樣給出證據。我們可能會為篇幅與清晰起見作出編輯，但始終會徵得作者同意。來信與本刊其他所有內容一樣，依據 CC BY 4.0 授權開放取用出版；其作品被討論的作者，將獲得在同一期作出回應的機會。

請透過本刊網站（journal.valuation-engineer.com）或直接向主編寄送來信。本刊收到的第一批真正的來信，將標誌著它從一份出版物變為一場對話的時刻——我們對此滿懷期待。

William Bert Craytor, SRA

主編

論文

房地產估價中的理論與實務落差：功能、邊界與張力的結構化闡述（繁體中文版）

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。

William Bert Craytor

第 1 卷 第 1 期 · 2026 年 5 月 21 日

發表於 2026 年 7 月 1 日

DOI: [10.67330/x7pbf13](https://doi.org/10.67330/x7pbf13)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 [創用 CC 姓名標示 4.0 國際授權條款 \(CC BY 4.0\)](#) 的條款發布。

摘要。房地產估價事關重大公共利益，但其官方理論——以過去與當前條件下的市場價值意見為核心——與估價產品在當代金融流程中的實際用途之間的張力正日益加劇。本文分五個階段建構這一論證，每個階段均配有明確、可查詢的 Prolog 知識庫：前兩個階段確立估價所承擔的社會與經濟功能；第三個階段指出其主要最終產品——市場價值意見；第四個階段考察估價相對於量化預測與金融預測的時間取向和方法論取向，以及將兩者分隔開來的監管邊界。這些內容匯聚為第五個階段的核心主張：在抵押貸款中，按貸款價值比折算的市場價值實際上充當著未來法拍價值的隱性替代指標——這是一種面向未來的用途，而現值理論並不予以承認。許多估價師並未意識到這一擔保品估值角色，因而未能以應有的審慎態度對待市場價值的未來走勢——本文旨在揭示這一問題，並歡迎後續研究。

關鍵詞：房地產估價，市場價值，估價理論，知識表示，估價準則，法拍價值

目錄

1 社會重要性	1
2 經濟重要性	6
3 市場價值估計作為估價最終產品的重要性	10
4 估價活動的時間取向	13
5 市場價值 × LTV 作為法拍價值的替代指標	19
6 結論	25

1 社會重要性

房地產及其他資產估價透過協助買方、賣方、貸款機構、保險公司、投資人、法院、稅務機關以及社會大眾，就財產及其他形式財富的移轉、融資、課稅與管理作出知情且公平的決定，發揮著不可或缺的社會與經濟功能。高品質的估價以證據、分析與專業判斷為基礎，提供獨立且論據充分的價值意見，從而促進市場透明度、公眾信心、資本的高效配置以及金融機構的穩定。(LaCour-Little and Malpezzi 2003; K'Akumu and Larsen 2024; Hendershott and Kane 1995; Lee, Peng, and Liao 2020)

相反地，低品質的估價可能在多個時間層面造成嚴重後果。短期內，不準確的估價可能使個別買方或賣方因買價過高、賣價過低、稅負過重或不當的放款決定而蒙受直接經濟損失。中期內，系統性的估價缺陷可能扭曲市場訊號、損害授信審核、助長投機行為、導致不公平的課稅或保險承保，並加劇訴訟與交易層面的不信任。長期而言，估價品質的持續缺失可能動搖公

眾對市場與制度的信心，導致金融不穩定、資源錯置、社區衰敗或過度投資，並造成波及家庭、企業與社區的更廣泛的社會不公。(LaCour-Little and Malpezzi 2003; Mayer and Nothaft 2022)

由於房地產及其他資產往往佔個人、企業與公共財富的相當大比重，估價品質不僅僅是私人層面的技術問題，而是關乎經濟韌性、公平與長期公共福祉的重大公共利益事項。(Mayer and Nothaft 2022)

```
%% appraisal_social.pl
%%
%% A Prolog knowledge base encoding the social and public-interest role
%% of real estate and other asset appraisals. Companion to
%% appraisal_economics.pl, which covers the strictly economic axis.
%%
%% All domain-specific predicates are prefixed with `social_` so this
%% module can be loaded alongside appraisal_economics.pl (and any other
%% companion KBs) without name collisions, and so cross-KB rules can
%% mention both vocabularies unambiguously.
%%
%% Conceptual model:
%%   - Appraisals support social DECISION TYPES.
%%   - Each decision involves identifiable PARTIES.
%%   - Appraisal QUALITY (high / low) drives EFFECTS across time HORIZONS.
%%   - Emphasis: equity, public confidence, public interest.
%%
%% Author: generated for Bert Craytor, ValEngr / RCA context.

:- module(appraisal_social,
    [ social_horizon/1
    , social_party/1
    , social_decision_type/1
    , social_involves/2           % DecisionType, Party
    , social_effect/4            % Quality, Horizon, Domain, Effect
    , social_contribution/1
    , social_hallmark/1
    , social_equity_principle/1
    , social_public_interest_claim/1
    , social_positive_effect/3
    , social_negative_effect/3
    , social_horizon_summary/2
    , social_consequences_of_quality/2
    , social_parties_affected_by/2
    , social_public_harm/2
    , social_private_harm/2
    , social_appraisal_is_public_interest/0
    ]).

:- discontiguous social_effect/4.
:- discontiguous social_involves/2.

%% -----
%% 1. Time horizons
%% -----

social_horizon(short_term).
social_horizon(medium_term).
social_horizon(long_term).
```

```

%% -----
%% 2. Parties whose decisions depend on appraisals
%%
%% The opening sentence enumerates these explicitly.
%% -----

social_party(buyers).
social_party(sellers).
social_party(lenders).
social_party(insurers).
social_party(investors).
social_party(courts).
social_party(taxing_authorities).
social_party(general_public).

%% -----
%% 3. Types of decisions appraisals inform
%% -----

social_decision_type(property_transfer).
social_decision_type(financing).
social_decision_type(taxation).
social_decision_type(stewardship).
social_decision_type(insurance_coverage).
social_decision_type(adjudication).          % courts: condemnation, divorce, estate
social_decision_type(public_administration).

social_involves(property_transfer,    buyers).
social_involves(property_transfer,    sellers).
social_involves(financing,            lenders).
social_involves(financing,            buyers).
social_involves(taxation,              taxing_authorities).
social_involves(taxation,              general_public).
social_involves(stewardship,          investors).
social_involves(stewardship,          general_public).
social_involves(insurance_coverage,    insurers).
social_involves(insurance_coverage,    buyers).
social_involves(adjudication,          courts).
social_involves(public_administration, taxing_authorities).
social_involves(public_administration, general_public).

%% -----
%% 4. What high-quality appraisals contribute (second sentence)
%% -----

social_contribution(market_transparency).
social_contribution(public_confidence).
social_contribution(efficient_capital_allocation).
social_contribution(stability_of_financial_institutions).

%% Characteristics of a high-quality appraisal opinion of value.
social_hallmark(independence).
social_hallmark(well_supported).
social_hallmark(grounded_in_evidence).
social_hallmark(grounded_in_analysis).
social_hallmark(grounded_in_professional_judgment).

%% -----

```

```

%% 5. Effects of appraisal quality by horizon
%% -----

%% ---- SHORT TERM: high quality ----
social_effect(high, short_term, property_transfer, fair_pricing_for_parties).
social_effect(high, short_term, financing, appropriate_lending_decisions).
social_effect(high, short_term, taxation, accurate_individual_assessment).

%% ---- SHORT TERM: low quality ----
social_effect(low, short_term, property_transfer, buyer_overpayment).
social_effect(low, short_term, property_transfer, seller_underselling).
social_effect(low, short_term, taxation, excessive_individual_taxation).
social_effect(low, short_term, financing,
    ↳ inappropriate_lending_decisions).
social_effect(low, short_term, property_transfer,
    ↳ immediate_financial_harm_to_individuals).

%% ---- MEDIUM TERM: high quality ----
social_effect(high, medium_term, markets, clear_market_signals).
social_effect(high, medium_term, financing, sound_credit_underwriting).
social_effect(high, medium_term, taxation, equitable_tax_assessment).
social_effect(high, medium_term, insurance_coverage,
    ↳ equitable_insurance_coverage).
social_effect(high, medium_term, adjudication, reduced_litigation_pressure).

%% ---- MEDIUM TERM: low quality ----
social_effect(low, medium_term, markets, distorted_market_signals).
social_effect(low, medium_term, financing,
    ↳ impaired_credit_underwriting).
social_effect(low, medium_term, markets,
    ↳ encouragement_of_speculation).
social_effect(low, medium_term, taxation, inequitable_taxation).
social_effect(low, medium_term, insurance_coverage,
    ↳ inequitable_insurance_coverage).
social_effect(low, medium_term, adjudication, increased_litigation).
social_effect(low, medium_term, markets, transactional_distrust).

%% ---- LONG TERM: high quality ----
social_effect(high, long_term, institutions, public_confidence_in_institutions).
social_effect(high, long_term, markets, public_confidence_in_markets).
social_effect(high, long_term, institutions, financial_stability).
social_effect(high, long_term, communities, equitable_community_outcomes).
social_effect(high, long_term, public_interest, economic_resilience).
social_effect(high, long_term, public_interest, social_fairness).
social_effect(high, long_term, public_interest, long_term_public_good).

%% ---- LONG TERM: low quality ----
social_effect(low, long_term, markets, eroded_market_confidence).
social_effect(low, long_term, institutions, eroded_institutional_confidence).
social_effect(low, long_term, institutions, financial_instability).
social_effect(low, long_term, markets, resource_misallocation).
social_effect(low, long_term, communities, neighborhood_decline).
social_effect(low, long_term, communities, neighborhood_overinvestment).
social_effect(low, long_term, communities, social_inequities_for_households).
social_effect(low, long_term, communities, social_inequities_for_businesses).
social_effect(low, long_term, communities, social_inequities_for_communities).

%% -----

```

```

%% 6. Equity principles and public-interest claims
%% -----

social_equity_principle(informed_decisions).
social_equity_principle(equitable_decisions).
social_equity_principle(fair_treatment_of_parties).
social_equity_principle(independence_of_valuer).

social_public_interest_claim(
    'Real estate and other assets are a substantial portion of personal, corporate,
    ↪ and public wealth.').
social_public_interest_claim(
    'Appraisal quality is not merely a private technical concern.').
social_public_interest_claim(
    'Appraisal quality is tied to economic resilience.').
social_public_interest_claim(
    'Appraisal quality is tied to fairness.').
social_public_interest_claim(
    'Appraisal quality is tied to the long-term public good.').

%% Top-level conclusion of the passage.
social_appraisal_is_public_interest :-
    social_public_interest_claim(_).    % at least one such claim exists

%% -----
%% 7. Derived / query predicates
%% -----

social_positive_effect(Horizon, Domain, Effect) :-
    social_effect(high, Horizon, Domain, Effect).

social_negative_effect(Horizon, Domain, Effect) :-
    social_effect(low, Horizon, Domain, Effect).

social_horizon_summary(Horizon, Effects) :-
    findall(Quality-Effect,
        social_effect(Quality, Horizon, _Domain, Effect),
        Effects).

social_consequences_of_quality(Quality, List) :-
    findall(consequence(Horizon, Domain, Effect),
        social_effect(Quality, Horizon, Domain, Effect),
        List).

social_parties_affected_by(DecisionType, Parties) :-
    setof(P, social_involves(DecisionType, P), Parties).

%% Harms that fall on the public rather than on transaction principals.
social_public_harm(Horizon, Effect) :-
    social_effect(low, Horizon, Domain, Effect),
    member(Domain, [markets, institutions, communities, public_interest,
    ↪ taxation]).

%% Private (transaction-level) harms.
social_private_harm(Horizon, Effect) :-
    social_effect(low, Horizon, Domain, Effect),
    member(Domain, [property_transfer, financing, insurance_coverage]).

```

2 經濟重要性

房地產及其他資產估價透過支持信用市場、投資決策、稅制、財務報告、保險承保、遺產管理、公司交易以及公共政策規劃的高效運作，發揮著基礎性的經濟作用。可靠的估價有助於建立可信的擔保品價值與市場價值衡量標準，使資本得以在借款人、貸款機構、投資人、政府與企業之間更有效率地流動。(LaCour-Little and Malpezzi 2003; Ebrahim and Hussain 2010)

短期內，準確的估價可以降低交易不確定性、改善放款決策、促進公平協商，並有助於防止因高估或低估而產生的直接經濟損失。中期內，估價品質透過影響抵押貸款風險、投資配置、租稅公平、保險充足性以及市場參與者的定價行為，左右著更廣泛市場的穩定與效率。在此期間，薄弱或存在系統性偏誤的估價可能助長資產泡沫膨脹、損害放款組合、扭曲開發誘因，並加大違約或訴訟風險。(LaCour-Little and Malpezzi 2003; Hendershott and Kane 1992; Griffin and Maturana 2016)

長期來看，估價品質的累積經濟後果更為顯著，影響著資本形成模式、基礎建設投資、跨世代財富移轉、公共收入、住宅可負擔性以及整個金融體系的韌性。估價實務的長期缺陷可能導致資本的慢性錯置、銀行體系不穩定、投資人信心下降以及土地與生產性資產的低效率利用；而嚴格的估價準則則有助於支撐穩定的市場、穩健的經濟規劃與長期經濟成長。(Ghent, Torous, and Valkanov 2019; Hendershott and Kane 1992; Ebrahim and Hussain 2010)

```
%% appraisal_economics.pl
%%
%% A Prolog knowledge base encoding the economic role of real estate
%% and other asset appraisals across short-, medium-, and long-term
%% horizons, including both the consequences of high-quality appraisal
%% practice and the consequences of weak or biased practice.
%%
%% All domain-specific predicates are prefixed with `economic_` so this
%% module can be loaded alongside appraisal_social.pl (and any other
%% companion KBs) without name collisions, and so cross-KB rules can
%% mention both vocabularies unambiguously.
%%
%% Conceptual model:
%%   - Appraisals support a set of economic FUNCTIONS.
%%   - Each function operates over one or more time HORIZONS.
%%   - Appraisal QUALITY (high / low) drives EFFECTS on those functions.
%%   - Effects are POSITIVE (benefits) when quality is high and
%%     NEGATIVE (harms) when quality is low/systematically biased.
%%
%% Author: generated for Bert Craytor, ValEngr / RCA context.

:- module(appraisal_economics,
    [ economic_horizon/1
    , economic_function/1
    , economic_supports/2
    , economic_effect/4
    ↪ Effect
    , economic_stakeholder/1
    , economic_benefits/2
    , economic_positive_effect/3
    , economic_negative_effect/3
    , economic_horizon_summary/2
    , economic_consequences_of_quality/2
    , economic_functions_at_horizon/2
    , economic_stakeholders_affected_by/2
    , economic_reliable_appraisals_improve/1
    , economic_appraisal_quality_matters_at/1
    ]).
```

```

:- discontinuous economic_effect/4.
:- discontinuous economic_supports/2.

%% -----
%% 1. Time horizons
%% -----

economic_horizon(short_term).
economic_horizon(medium_term).
economic_horizon(long_term).

%% -----
%% 2. Economic functions supported by appraisals
%% -----

economic_function(credit_markets).
economic_function(investment_decisions).
economic_function(taxation_systems).
economic_function(financial_reporting).
economic_function(insurance_underwriting).
economic_function(estate_administration).
economic_function(corporate_transactions).
economic_function(public_policy_planning).

%% -----
%% 3. Stakeholders whose decisions depend on appraisals
%% -----

economic_stakeholder(borrowers).
economic_stakeholder(lenders).
economic_stakeholder(investors).
economic_stakeholder(governments).
economic_stakeholder(businesses).
economic_stakeholder(insurers).
economic_stakeholder(taxpayers).
economic_stakeholder(homeowners).

%% Capital flows more efficiently when appraisals are reliable.
economic_benefits(borrowers, credit_markets).
economic_benefits(lenders, credit_markets).
economic_benefits(investors, investment_decisions).
economic_benefits(governments, taxation_systems).
economic_benefits(governments, public_policy_planning).
economic_benefits(businesses, corporate_transactions).
economic_benefits(businesses, financial_reporting).
economic_benefits(insurers, insurance_underwriting).
economic_benefits(taxpayers, taxation_systems).
economic_benefits(homeowners, credit_markets).

%% -----
%% 4. Which functions are most salient at each horizon
%% -----

economic_supports(credit_markets, short_term).
economic_supports(investment_decisions, short_term).
economic_supports(corporate_transactions, short_term).

```

```

economic_supports(credit_markets,      medium_term).
economic_supports(investment_decisions, medium_term).
economic_supports(taxation_systems,     medium_term).
economic_supports(insurance_underwriting, medium_term).

economic_supports(credit_markets,      long_term).
economic_supports(taxation_systems,     long_term).
economic_supports(public_policy_planning, long_term).
economic_supports(financial_reporting,  long_term).
economic_supports(estate_administration, long_term).

%% -----
%% 5. Effects of appraisal quality, by horizon and function
%% -----

%% ---- SHORT TERM: high quality ----
economic_effect(high, short_term, credit_markets,
    ↪ reduced_transactional_uncertainty).
economic_effect(high, short_term, credit_markets,
    ↪ improved_lending_decisions).
economic_effect(high, short_term, corporate_transactions, fair_negotiations).
economic_effect(high, short_term, investment_decisions,
    ↪ prevention_of_immediate_losses).

%% ---- SHORT TERM: low quality ----
economic_effect(low, short_term, credit_markets,
    ↪ transactional_uncertainty).
economic_effect(low, short_term, credit_markets,      poor_lending_decisions).
economic_effect(low, short_term, investment_decisions,
    ↪ losses_from_overvaluation).
economic_effect(low, short_term, investment_decisions,
    ↪ losses_from_undervaluation).

%% ---- MEDIUM TERM: high quality ----
economic_effect(high, medium_term, credit_markets,      stable_mortgage_risk).
economic_effect(high, medium_term, investment_decisions,
    ↪ efficient_investment_allocation).
economic_effect(high, medium_term, taxation_systems,     tax_equity).
economic_effect(high, medium_term, insurance_underwriting,
    ↪ adequate_insurance_coverage).
economic_effect(high, medium_term, investment_decisions,
    ↪ rational_market_pricing).

%% ---- MEDIUM TERM: low quality ----
economic_effect(low, medium_term, investment_decisions, inflated_asset_bubbles).
economic_effect(low, medium_term, credit_markets,
    ↪ impaired_lending_portfolios).
economic_effect(low, medium_term, public_policy_planning,
    ↪ distorted_development_incentives).
economic_effect(low, medium_term, credit_markets,      increased_default_risk).
economic_effect(low, medium_term, corporate_transactions,
    ↪ increased_litigation_risk).

%% ---- LONG TERM: high quality ----
economic_effect(high, long_term, investment_decisions, sound_capital_formation).
economic_effect(high, long_term, public_policy_planning,
    ↪ sound_infrastructure_investment).

```

```

economic_effect(high, long_term, estate_administration,
↳ orderly_intergenerational_wealth_transfer).
economic_effect(high, long_term, taxation_systems,      stable_public_revenues).
economic_effect(high, long_term, public_policy_planning,
↳ supported_housing_affordability).
economic_effect(high, long_term, credit_markets,
↳ financial_system_resilience).
economic_effect(high, long_term, investment_decisions,
↳ long_term_economic_growth).

%% ---- LONG TERM: low quality ----
economic_effect(low, long_term, investment_decisions,
↳ chronic_capital_misallocation).
economic_effect(low, long_term, credit_markets,      banking_instability).
economic_effect(low, long_term, investment_decisions,
↳ reduced_investor_confidence).
economic_effect(low, long_term, public_policy_planning, inefficient_land_use).
economic_effect(low, long_term, public_policy_planning,
↳ inefficient_productive_asset_use).

%% -----
%% 6. Derived / query predicates
%% -----

economic_positive_effect(Horizon, Function, Effect) :-
    economic_effect(high, Horizon, Function, Effect).

economic_negative_effect(Horizon, Function, Effect) :-
    economic_effect(low, Horizon, Function, Effect).

economic_horizon_summary(Horizon, Effects) :-
    findall(Quality-Effect,
        economic_effect(Quality, Horizon, _Function, Effect),
        Effects).

economic_consequences_of_quality(Quality, List) :-
    findall(consequence(Horizon, Function, Effect),
        economic_effect(Quality, Horizon, Function, Effect),
        List).

economic_functions_at_horizon(Horizon, Functions) :-
    setof(F, economic_supports(F, Horizon), Functions).

economic_stakeholders_affected_by(Function, Stakeholders) :-
    setof(S, economic_benefits(S, Function), Stakeholders).

%% Overall claim: reliable appraisals improve allocative efficiency.
economic_reliable_appraisals_improve(capital_flow_efficiency).
economic_reliable_appraisals_improve(market_stability).
economic_reliable_appraisals_improve(long_term_economic_growth).

%% Top-level qualitative claim: appraisal quality matters at every horizon.
economic_appraisal_quality_matters_at(Horizon) :-
    economic_horizon(Horizon),
    \+ \+ economic_effect(high, Horizon, _, _),
    \+ \+ economic_effect(low, Horizon, _, _).

```

3 市場價值估計作為估價最終產品的重要性

市場價值估價在擔保品風險分析中發揮著重要作用，因為它就財產或資產在正常市場條件下可能實現的價格提供了專業形成的估計，從而為評估信用曝險與潛在法拍結果的貸款機構、投資人和擔保品分析師提供了基礎性基準。(Agarwal, Ben-David, and Yao 2015) 儘管由於不利的銷售條件、縮短的出售期間、法律費用、實體耗損或負面市場情緒，法拍價值 (foreclosure value) 或清算價值可能與市場價值相差甚遠，擔保品分析師仍經常以市場價值估價作為在壓力情境下估算回收潛力的起點。(Campbell, Giglio, and Pathak 2011; Pennington-Cross 2006; Clauretie and Daneshvary 2009)

透過分析市場價值、貸款餘額、貸款價值比 (loan-to-value ratio, LTV)、市場趨勢、房產狀況、流動性以及當地供需狀況之間的關係，擔保品分析師評估借款人違約情況下損失發生的可能性及其嚴重程度。(Qi and Yang 2009)

短期內，準確的市場價值估價有助於貸款機構作出審慎的授信決定並建立適當的擔保品安全邊際。中期內，它們透過協助機構評估經濟循環中擔保品充足性的變化，支持投資組合風險管理、貸款證券化、準備金配置與法令遵循。長期而言，持續可靠的估價實務透過降低過度槓桿、擔保不足的放款以及抵押資產系統性錯誤定價的風險，為金融體系穩定作出貢獻——這些風險會在市場承壓時期放大法拍損失和更廣泛的經濟動盪。(Eriksen et al. 2019)

```
%% appraisal_collateral.pl
%%
%% A Prolog knowledge base encoding the role of market value (MV)
%% appraisals in collateral risk analysis. Companion to
%% appraisal_economics.pl and appraisal_social.pl.
%%
%% Naming convention:
%%   - Predicates that have direct counterparts in the companion KBs
%%     (horizon, effect, summary, consequences, etc.) are prefixed
%%     `collateral_` so this module can be loaded alongside the others
%%     without name collisions, and so bridge rules can mention each
%%     vocabulary unambiguously.
%%   - Predicates that are unique to the collateral-analysis domain
%%     (value_basis, divergence_factor, risk_input, etc.) are NOT
%%     prefixed. They have no analogues in the other KBs and no
%%     collision risk, so the prefix would only add noise.

:- module(appraisal_collateral,
    [ collateral_horizon/1
    , collateral_analyst/1
    , collateral_effect/4                                % Quality, Horizon, Domain,
    ↪ Effect
    , collateral_positive_effect/3
    , collateral_negative_effect/3
    , collateral_horizon_summary/2
    , collateral_consequences_of_quality/2
    , value_basis/1
    , uses/2                                              % Analyst, ValueBasis
    , divergence_factor/1
    , stress_scenario/1
    , risk_input/1
    , risk_output/1
    , inputs_for/2
    , mv_is_benchmark_for/1
    ]).

:- discontiguous collateral_effect/4.
:- discontiguous uses/2.
```

```

%% -----
%% 1. Time horizons (namespaced -- parallel to economic/social_)
%% -----

collateral_horizon(short_term).
collateral_horizon(medium_term).
collateral_horizon(long_term).

%% -----
%% 2. Who uses MV appraisals in collateral risk analysis
%% -----

collateral_analyst(lenders).
collateral_analyst(investors).
collateral_analyst(collateral_analysts).
collateral_analyst(portfolio_risk_managers).
collateral_analyst(regulators).
collateral_analyst(securitization_arrangers).

%% -----
%% 3. Value bases distinguished in the passage
%% -----

value_basis(market_value).           % "normal market conditions"
value_basis(foreclosure_value).       % distressed
value_basis(liquidation_value).       % distressed, accelerated

%% Which value bases each analyst type works with.
uses(lenders, market_value).
uses(investors, market_value).
uses(collateral_analysts, market_value).
uses(collateral_analysts, foreclosure_value).
uses(collateral_analysts, liquidation_value).
uses(portfolio_risk_managers, market_value).
uses(securitization_arrangers, market_value).
uses(regulators, market_value).

%% Top-level claim of the opening sentence.
mv_is_benchmark_for(credit_exposure).
mv_is_benchmark_for(foreclosure_outcomes).
mv_is_benchmark_for(recovery_potential).

%% -----
%% 4. Why foreclosure / liquidation value diverges from MV
%% -----

divergence_factor(distressed_marketing_conditions).
divergence_factor(accelerated_sale_period).
divergence_factor(legal_costs).
divergence_factor(physical_deterioration).
divergence_factor(adverse_market_sentiment).

%% Stressed scenarios in which divergence factors apply.
stress_scenario(borrower_default).
stress_scenario(foreclosure).
stress_scenario(forced_liquidation).
stress_scenario(market_downturn).

```

```

%% -----
%% 5. Inputs the collateral analyst combines with MV, and the
%%      outputs they produce
%% -----

risk_input(market_value).
risk_input(loan_balance).
risk_input(loan_to_value_ratio).
risk_input(market_trends).
risk_input(property_condition).
risk_input(liquidity).
risk_input(local_supply_demand).

risk_output(likelihood_of_loss).
risk_output(severity_of_loss).
risk_output(recovery_estimate).
risk_output(loss_given_default).

%% inputs_for(+Output, -Inputs)
%% All risk inputs are used to produce each risk output.
inputs_for(Output, Inputs) :-
    risk_output(Output),
    setof(I, risk_input(I), Inputs).

%% -----
%% 6. Effects of appraisal quality by horizon
%% -----

%% ---- SHORT TERM: high quality ----
collateral_effect(high, short_term, underwriting, prudent_underwriting_decisions).
collateral_effect(high, short_term, underwriting, appropriate_collateral_margins).

%% ---- SHORT TERM: low quality ----
collateral_effect(low, short_term, underwriting,
    ↪ imprudent_underwriting_decisions).
collateral_effect(low, short_term, underwriting, inadequate_collateral_margins).
collateral_effect(low, short_term, underwriting, mispriced_credit_exposure).

%% ---- MEDIUM TERM: high quality ----
collateral_effect(high, medium_term, portfolio,
    ↪ effective_portfolio_risk_management).
collateral_effect(high, medium_term, portfolio, sound_loan_securitization).
collateral_effect(high, medium_term, portfolio, appropriate_reserve_allocation).
collateral_effect(high, medium_term, regulatory, regulatory_compliance).
collateral_effect(high, medium_term, portfolio,
    ↪ cycle_aware_collateral_evaluation).

%% ---- MEDIUM TERM: low quality ----
collateral_effect(low, medium_term, portfolio,
    ↪ weakened_portfolio_risk_management).
collateral_effect(low, medium_term, portfolio, mispriced_securitizations).
collateral_effect(low, medium_term, portfolio, inadequate_loss_reserves).
collateral_effect(low, medium_term, regulatory, regulatory_noncompliance_risk).
collateral_effect(low, medium_term, portfolio, misjudged_collateral_adequacy).

%% ---- LONG TERM: high quality ----
collateral_effect(high, long_term, system, financial_system_stability).

```

```

collateral_effect(high, long_term, system, restraint_on_excessive_leverage).
collateral_effect(high, long_term, system, well_secured_lending).
collateral_effect(high, long_term, system, accurate_systemic_collateral_pricing).
collateral_effect(high, long_term, system, dampened_foreclosure_amplification).

%% ---- LONG TERM: low quality ----
collateral_effect(low, long_term, system, financial_system_instability).
collateral_effect(low, long_term, system, excessive_leverage).
collateral_effect(low, long_term, system, poorly_secured_lending).
collateral_effect(low, long_term, system, systemic_collateral_mispricing).
collateral_effect(low, long_term, system, amplified_foreclosure_losses).
collateral_effect(low, long_term, system,
↳ amplified_economic_disruption_under_stress).

%% -----
%% 7. Derived / query predicates (namespaced)
%% -----

collateral_positive_effect(Horizon, Domain, Effect) :-
    collateral_effect(high, Horizon, Domain, Effect).

collateral_negative_effect(Horizon, Domain, Effect) :-
    collateral_effect(low, Horizon, Domain, Effect).

collateral_horizon_summary(Horizon, Effects) :-
    findall(Quality-Effect,
            collateral_effect(Quality, Horizon, _Domain, Effect),
            Effects).

collateral_consequences_of_quality(Quality, List) :-
    findall(consequence(Horizon, Domain, Effect),
            collateral_effect(Quality, Horizon, Domain, Effect),
            List).

```

4 估價活動的時間取向

估價師通常就某一特定生效日期形成時點性的市場價值意見——回顧性的（過去）、當前的，或在明確限制下前瞻性的（未來）——而不是對較長時間區間作出機率性預測。這種時間與方法論取向是受監管的估價實務與量化分析師、金融分析師、投資分析師及經濟學家的工作之間的核心區別之一。（Quan and Quigley 1991）

在大多數估價業務中，特別是依據《專業估價實務統一準則》（Uniform Standards of Professional Appraisal Practice, USPAP）或《國際評價準則》（International Valuation Standards, IVS）等準則進行的市場價值估價，估價師給出的是一個經調和的點估計（或在允許的情形下給出一個窄幅區間），錨定於某一生效日期的可觀察市場證據、當時的市場條件和公認的方法。USPAP 明確允許價值結論以具體金額、數值區間或與某一基準的關係的形式給出，但經調和的意見仍然是與該日期掛鉤的一個可信的點位指示。（The Appraisal Foundation 2024; International Valuation Standards Council 2025）

儘管估價師必然會納入市場參與者對趨勢、供需、折舊、資本化率或折現率以及去化速度的合理預期，但這樣做是為了支持生效日期上有市場依據的價值，而非預測投機性波動。前瞻性價值意見（例如開發案完工後的穩定價值）需要明示的特殊假設或假定條件，並且仍被限定於一個明確的未來生效日期，且須有源自市場的支持。

相比之下，量化分析師、金融分析師與投資分析師經常透過涵蓋各時間區間內一系列可能結果的估計來處理不確定性，運用機率分布、蒙地卡羅模擬、情境分析、敏感度檢定或風險值（Value at Risk, VaR）等指標。他們的關注點往往是預期的未來表現、經風險調整的報酬，或

特定主體在不同期限內的投資價值，而不是某一日期上的單一市場價值點位。(Hull 2018)

這類工作還預設了估價師執照既不要求也不認證的技術能力。例如，擬合多元自適應迴歸樣條 (MARS) 模型需要控制模型複雜度的判斷力，使得強勁的樣本內擬合 (高 R^2) 不是以過擬合為代價換來的，而是能夠延續到保留資料上 (高交叉驗證 R^2)，而這又以對底層演算法的切實理解為前提。(Friedman 1991; Steurer, Hill, and Pfeifer 2021) 除這類量化能力之外，這項工作還要求具備處理高度結構化概念、辨析細微語意差別以及吸收陌生方法的概念能力。這些是技能而非資格：持照估價師可能具備它們，無照分析師也可能欠缺它們——這正是它們本身並不能決定某項任務落在監管界線哪一側的原因。

估價師也可以從事有限的面向未來的工作，例如在折現現金流分析中預測收入流、空置率、費用和復歸價值，或在開發或訴訟情境中估計前瞻性價值。只要立足於當前市場證據，並遵循強調可支持性、避免無依據投機的專業準則，這些內容就仍處於估價的邊界之內。(Diaz and Wolverton 1998; Gallimore 1996; Crosby, Lizieri, and McAllister 2010) 現代估價實務越來越多地融合統計工具、情境考量與風險分析——尤其是在大量估價、自動估價模型或投資組合業務中——但核心交付成果通常仍是時點性的市場價值意見。(International Association of Assessing Officers 2017; Office of the Comptroller of the Currency and Board of Governors of the Federal Reserve System and Federal Deposit Insurance Corporation and National Credit Union Administration and Consumer Financial Protection Bureau and Federal Housing Finance Agency 2024; Alexandrov, Goodman, and Neal 2023)

可以這樣概括：估價實務主要是證據導向的，面向從過去到現在（或有限前瞻）的點估計；而金融與量化分析則更具機率性、分布性，並且明確面向各個時間跨度上的未來。(Geltner 1991) 估價師問的是：“在現有證據和既定假設下，截至指定生效日期，最有依據的市場價值指示是什麼？”金融與量化分析師則更常問：“未來一段時期內可能出現何種結果區間、機率幾何，以及應如何配置資源？”

這條邊界並非絕對，並且隨著市場複雜性的提高而持續演變。儘管如此，在目的（市場價值與投資價值之別）、證據要求和監管框架方面，根本性差異依然存在。

然而，真正起作用的界線不是由技能而是由法律劃定的：決定性的問題只是某項任務在法律上是否要求持有估價師執照。量化分析師、金融分析師、經濟學家、投資分析師、投資組合經理人及相關專業人士可以在不持有房地產估價師執照的情況下，合法地估計、預測、建模或分析未來資產價值——包括分布式或區間式的預測——只要該工作不構成受監管的估價。(Murphy 2012)

關鍵區別在於：

1. 價值類型（例如市場價值點估計，還是投資價值或機率性預測）；
2. 時間焦點（特定生效日期，還是某一時間區間或期限）；
3. 輸出形式（點估計或窄幅區間，還是完整的機率分布）；
4. 預期用途以及該工作是否以估價的名義出具；以及
5. 是否涉及要求遵守 USPAP 等準則的受監管放款或法律情境。

這一框架在很大程度上源於美國 1989 年《金融機構改革、復甦與執行法》(FIRREA) 等銀行監管法規，其目的在於確保聯邦相關交易能夠獲得可信、獨立的市場價值意見。(United States Congress 1989; K'Akumu and Larsen 2024) 許多高度複雜的估值活動——股票研究目標價、可行性研究、投資組合風險建模、選擇權定價或預測分析——都不在執照管轄範圍之內，因為它們服務於投資、預測或諮詢目的，而非受監管的估價職能。

```
%% appraisal_scope_regulation.pl
%%
%% Knowledge base for the boundary of regulated appraisal practice.
%%
%% TWO TIERS, deliberately kept separate:
%%
%% TIER 1 -- THE DEFINING TEST (regulation cluster). A task is a
%% regulated appraisal IFF it is one for which an appraiser's
%% license is legally required. requires_license/1 is that test;
%% it -- and only it -- decides regulated_appraisal/1 vs.
```

```

%%      unregulated_valuation/1.
%%
%%      TIER 2 -- DESCRIPTIVE TENDENCIES (scope cluster).  How the two
%%      populations -- `appraisal' and `quant_financial' -- typically
%%      differ in approach, temporal scope, output form, and skill set.
%%      These are CORRELATES, not the boundary: the populations overlap
%%      (a licensed appraiser may do quant work; a quant may be
%%      unlicensed), and nothing here determines licensure.  No claim is
%%      made about intelligence -- only about training and methodology.
%%
%%      `temporal scope' and `approach' are DISTINCT axes: an approach has a
%%      typical (defeasible) temporal mode, but the two are not the same
%%      viewpoint -- which is why an appraiser using an AVM (appraisal
%%      approach, distributional methods) is possible.

:- module(appraisal_scope_regulation,
  [ % --- TIER 1: the defining test (regulation) ---
    regulation_framework/1
  , regulation_framework_origin/2
  , regulation_distinguishing_factor/1
  , regulation_triggers_licensure/1
  , regulation_definition_of_appraisal/1
  , regulation_unlicensed_activity/2
  , regulation_licensed_activity/2
  , regulation_legal_exposure_condition/1
  , requires_license/1
  , exempt_from_appraisal_license/1
  , regulated_appraisal/1
  , unregulated_valuation/1
  , regulatory_asymmetry/2

    % --- TIER 2: descriptive tendencies (scope) ---
  , valuation_approach/1
  , scope_profession/1
  , scope_orientation/2
  , temporal_scope/1
  , temporal_mode/1
  , scope_temporal_focus/2
  , scope_output_form/2
  , scope_skill/1
  , scope_skill_demanded_by/2
  , scope_inputs_considered/2
  , scope_future_oriented_context/2
  , scope_professional_constraint/1
  , scope_contrast/4
  , scope_question/2
  , scope_convergence_area/1
  ]).

:- disjoint regulation_unlicensed_activity/2.
:- disjoint regulation_licensed_activity/2.
:- disjoint scope_contrast/4.

%% =====
%% TIER 1 -- THE DEFINING TEST (the boundary)
%% =====

regulation_framework(firrea_1989).

```

```

regulation_framework(uspap).
regulation_framework(ivs).
regulation_framework(state_licensing).
regulation_framework(state_certification).

regulation_framework_origin(firrea_1989,      savings_and_loan_crisis).
regulation_framework_origin(uspap,            firrea_1989).
regulation_framework_origin(state_licensing, firrea_1989).

regulation_distinguishing_factor(kind_of_value).
regulation_distinguishing_factor(purpose_of_valuation).
regulation_distinguishing_factor(temporal_focus).
regulation_distinguishing_factor(output_form).
regulation_distinguishing_factor(represented_as_appraisal).
regulation_distinguishing_factor(federally_related_transaction_involvement).

regulation_triggers_licensure(federally_related_mortgage_transaction).
regulation_triggers_licensure(work_represented_as_appraisal).
regulation_triggers_licensure(work_meeting_state_appraisal_definition).
regulation_triggers_licensure(regulated_lending_transaction).

regulation_definition_of_appraisal(formal_opinion_of_value).
regulation_definition_of_appraisal(concerning_identified_real_property).
regulation_definition_of_appraisal(developed_per_recognized_methods).
regulation_definition_of_appraisal(for_compensation).

regulation_unlicensed_activity(equity_analyst,      stock_price_targets).
regulation_unlicensed_activity(investment_bank,
    ↪ enterprise_value_estimates).
regulation_unlicensed_activity(hedge_fund,
    ↪ commercial_real_estate_appreciation_forecasts).
regulation_unlicensed_activity(economist,
    ↪ housing_market_trend_projections).
regulation_unlicensed_activity(insurance_analyst,
    ↪ replacement_cost_estimates).
regulation_unlicensed_activity(developer,          feasibility_studies).
regulation_unlicensed_activity(internal_corporate_analyst,
    ↪ asset_impairment_estimates).
regulation_unlicensed_activity(internal_corporate_analyst,
    ↪ portfolio_value_estimates).
regulation_unlicensed_activity(quantitative_analyst,
    ↪ predictive_valuation_models).
regulation_unlicensed_activity(avm_developer,
    ↪ algorithmic_residential_price_estimates).

regulation_licensed_activity(licensed_appraiser,
    ↪ market_value_for_federally_related_transaction).
regulation_licensed_activity(licensed_appraiser,
    ↪ formal_opinion_of_value_for_compensation).
regulation_licensed_activity(licensed_appraiser,
    ↪ appraisal_for_regulated_mortgage_lending).
regulation_licensed_activity(licensed_appraiser,
    ↪ appraisal_represented_as_such_under_state_law).

regulation_legal_exposure_condition(holding_self_out_as_licensed_appraiser).
regulation_legal_exposure_condition(preparing_appraisal_as_defined_by_state_law).
regulation_legal_exposure_condition(involved_in_transaction_requiring_licensed_app
    ↪ raisal).

```

```

%% --- The test itself: does the TASK require an appraiser's license? ---
requires_license(Task) :-
    regulation_triggers_licensure(Task), !.
requires_license(Task) :-
    regulation_licensed_activity(_, Task), !.

exempt_from_appraisal_license(Task) :-
    regulation_unlicensed_activity(_, Task),
    \+ requires_license(Task).

:- dynamic activity_satisfies_definition/1.

activity_satisfies_definition(Task) :-
    regulation_licensed_activity(_, Task).

regulated_appraisal(Task) :-
    activity_satisfies_definition(Task),
    requires_license(Task).

unregulated_valuation(Task) :-
    regulation_unlicensed_activity(_, Task),
    \+ regulated_appraisal(Task).

%% Meta-tension: the regulatory asymmetry the passage closes on.
regulatory_asymmetry(current_market_value_for_lending,
    future_oriented_valuation_for_investment).
regulatory_asymmetry(heavily_regulated_appraisal_practice,
    lightly_regulated_quantitative_practice).
regulatory_asymmetry(consumer_and_lending_protection_focus,
    investment_advisory_focus).

regulatory_asymmetry_origin(banking_regulation).
regulatory_asymmetry_origin(consumer_protection).
regulatory_asymmetry_origin(legal_evidentiary_concerns).
regulatory_asymmetry_origin_disclaimer(
    'Not based on any exclusive claim that appraisers alone possess valuation
    ↪ expertise.').

%% =====
%% TIER 2 -- DESCRIPTIVE TENDENCIES (NOT the boundary)
%%
%% Overlapping populations; correlates of who typically lands on each
%% side. None of these determine licensure -- Tier 1 does.
%% =====

%% --- Axis: valuation approach (descriptive populations) ---
valuation_approach(appraisal).
valuation_approach(quant_financial).

scope_profession(appraiser).
scope_profession(quantitative_analyst).
scope_profession(financial_analyst).
scope_profession(investment_analyst).
scope_profession(portfolio_manager).
scope_profession(economist).
scope_profession(investment_strategist).
scope_profession(financial_forecaster).

```

```

scope_orientation(appraisal,      evidentiary).
scope_orientation(quant_financial, probabilistic).

%% --- Axis: temporal scope (its OWN viewpoint, distinct from approach) ---
temporal_scope(retrospective).
temporal_scope(current).
temporal_scope(prospective_point).
temporal_scope(future_range_horizon).

temporal_mode(point_in_time).
temporal_mode(distribution_over_horizon).

%% Typical (defeasible) pairing of approach to temporal mode -- a LINK
%% between two distinct axes, not an identity.
scope_temporal_focus(appraisal,      point_in_time).
scope_temporal_focus(quant_financial, distribution_over_horizon).

%% --- Axis: output form ---
scope_output_form(appraisal,      point_estimate).
scope_output_form(appraisal,      narrow_range).
scope_output_form(quant_financial, probability_distribution).
scope_output_form(quant_financial, range_of_outcomes).

%% --- Axis: skills / methodological toolkit (training, NOT intelligence) ---
scope_skill(mars_without_overfitting).
scope_skill(fit_vs_generalization_balance).    % balancing  $R^2$  against  $CV-R^2$ 
scope_skill(algorithmic_understanding).
scope_skill(handling_structured_concepts).
scope_skill(discerning_fine_distinctions_in_meaning).
scope_skill(assimilating_new_methods).

%% Skills typically demanded by advanced quantitative valuation; an
%% appraisal license neither requires nor certifies them.
scope_skill_demanded_by(quant_financial, mars_without_overfitting).
scope_skill_demanded_by(quant_financial, fit_vs_generalization_balance).
scope_skill_demanded_by(quant_financial, algorithmic_understanding).
scope_skill_demanded_by(quant_financial, handling_structured_concepts).
scope_skill_demanded_by(quant_financial, discerning_fine_distinctions_in_meaning).
scope_skill_demanded_by(quant_financial, assimilating_new_methods).

%% --- Inputs the appraiser weighs to support a value at the effective date ---
scope_inputs_considered(appraisal, current_market_evidence).
scope_inputs_considered(appraisal, observable_transactions).
scope_inputs_considered(appraisal, prevailing_market_conditions).
scope_inputs_considered(appraisal, recognized_methodologies).
scope_inputs_considered(appraisal, market_trends).
scope_inputs_considered(appraisal, anticipated_supply_demand).
scope_inputs_considered(appraisal, depreciation).
scope_inputs_considered(appraisal, capitalization_rates).
scope_inputs_considered(appraisal, absorption_patterns).
scope_inputs_considered(appraisal, expectations_already_priced_in).

%% --- Bounded future-oriented contexts that remain within appraisal ---
scope_future_oriented_context(income_property, future_income_streams).
scope_future_oriented_context(income_property, vacancy_rates).
scope_future_oriented_context(income_property, expenses).
scope_future_oriented_context(income_property, reversion_values).

```

```

scope_future_oriented_context(income_property, dcf_analysis).
scope_future_oriented_context(development, prospective_stabilized_value).
scope_future_oriented_context(conservation, future_damages).
scope_future_oriented_context(environmental, future_damages).
scope_future_oriented_context(insurance, future_damages).
scope_future_oriented_context(litigation, future_damages).
scope_future_oriented_context(eminent_domain, future_damages).
scope_future_oriented_context(general, remaining_economic_life).
scope_future_oriented_context(general, anticipated_market_reactions).
scope_future_oriented_context(market_analysis, price_trend_direction).

%% --- Professional constraints on any future-oriented component ---
scope_professional_constraint(supportability).
scope_professional_constraint(market_evidence).
scope_professional_constraint(extraordinary_assumptions_or_hypothetical_conditions ↪
↪ _disclosed).
scope_professional_constraint(avoidance_of_unsupported_speculation).

%% --- Contrasts across the descriptive axes:
↪ scope_contrast(A,B,Dimension,Distinction) ---
scope_contrast(appraisal, quant_financial, orientation,
evidentiary_vs_probabilistic).
scope_contrast(appraisal, quant_financial, temporal_mode,
point_in_time_vs_distribution_over_horizon).
scope_contrast(appraisal, quant_financial, output_form,
point_or_narrow_range_vs_probability_distribution).
scope_contrast(appraisal, quant_financial, purpose,
market_value_vs_investment_worth).
scope_contrast(appraisal, quant_financial, skill_set,
appraisal_licensing_toolkit_vs_advanced_statistical_toolkit).
scope_contrast(appraisal, quant_financial, typical_methods,
sales_comparison_cost_income_vs_simulation_option_pricing_var).

%% --- The orienting question each population asks (verbatim from the passage) ---
scope_question(appraisal,
'What is the most supportable market value indication as of the specified
↪ effective date, given current evidence and defined assumptions?').
scope_question(quant_financial,
'What range of outcomes is likely over a future period, with what
↪ probabilities, and how should resources be allocated?').

%% --- Where the two converge (appraisal borrows distributional methods) ---
scope_convergence_area(complex_commercial_valuation).
scope_convergence_area(mass_appraisal).
scope_convergence_area(automated_valuation_models).
scope_convergence_area(institutional_portfolio_analysis).

```

5 市場價值 × LTV 作為法拍價值的替代指標

前面各節闡明了估價的用途、估價的產出，以及估價的理論邊界與監管邊界之所在；本節則觸及它們一路鋪陳而來的那個張力。

抵押貸款與擔保品分析背後一個更深層的實際現實是：就實際經濟功能而言，用於抵押貸款授信審核的市場價值估價往往充當著對未來某一持有期內預期擔保品可回收性的隱性替代指標，儘管在形式上該估價被表述為截至特定生效日期的當前市場價值意見。(Foster and Van Order 1984; Kau et al. 1995; Deng, Quigley, and Van Order 2000)

80% 的貸款價值比並非隨意設定。從歷史上看，它在一定程度上是作為保護貸款機構的風險緩衝而演變形成的，用以抵禦：

- 溫和的市場下行、
- 法拍費用、
- 清算折價、
- 持有成本、
- 法律費用、
- 應計利息、
- 以及違約後處分擔保品所需時間內的不確定性。(Foote, Gerardi, and Willen 2008; Qi and Yang 2009)

這一緩衝也解釋了為何貸款金額本身——市場價值與允許的貸款價值比之乘積——可以被解讀為貸款機構對可回收法拍價值的工作估計：將貸款上限設定為市場價值的某一比例，實際上是在押注強迫出售所得不會跌破貸款餘額。然而，這一折減應當有多大，遠比 80% 這個整數所暗示的更難確定。對法拍折價的估計因市場、方法和時期而差異巨大；一項被頻繁引用的強迫出售研究將其定在比可比正常交易價格低約 27% 的水準 (Campbell, Giglio, and Pathak 2011)，而其他分析——在校正房產狀況、區位和樣本選擇之後——報告的幅度則大不相同。(Pennington-Cross 2006; Clauretie and Daneshvary 2009)

相比之下，常規的 80% 門檻本質上是一個全國統一的單一參數，被套用於千差萬別的地方市場和信用循環的先後階段。它無法追蹤任何特定市場的折價，因此只是一個非常粗略的、一體適用的近似，而非從市場價值到法拍回收額的精確映射。真要說的話，它還是一個偏薄的緩衝：20% 的折減低於上述研究報告的 27% 折價，因此即便尚未扣除前文列舉的處分成本，常規貸款的最高額度就可能超過強迫出售的毛收入。

在借款人權益過薄、不足以構成這一緩衝的情形下，便以抵押貸款保險取而代之。對此最好的理解不是把它當作一個政策旋鈕，而是一種風險分配的權衡：有一部分借款人資產薄弱但收入穩定，向他們承作 LTV 超過 80% 的貸款所帶來的增量違約與損失風險由保險人預測並承擔——民間管道是高 LTV 常規貸款上的私人抵押貸款保險，公共管道則是聯邦住宅管理局 (Federal Housing Administration)。然而，使這種權衡成為常態的市場本身是聯邦住宅政策的產物：主導常規貸款市場的政府贊助企業依其特許章程被禁止在缺乏此類信用增強的情況下購買 LTV 超過 80% 的貸款 (United States Code 1970)。保險終止的條件同樣是制度選擇而非風險計算的結果，並且在不同方案之間差異懸殊。¹

因此，儘管估價師在官方口徑上估計的是“當前市場價值”，貸款機構在經濟上卻是透過一個前瞻性風險視角來解讀這一價值的。實際上，貸款機構間的問題更接近於：

“如果這位借款人在可預見的未來某個時候違約，擔保品是否仍可能足以涵蓋未償債務和與法拍相關的損失？”

這在本質上是一個面向未來的問題。

這有賴於以下兩者之間刻意的概念區分：

- 估價師的“當前市場價值”意見，以及
- 貸款機構在實務中將該意見用作抵禦未來擔保品減值的預測緩衝。

從歷史與法律上看，監管機關曾力圖防止估價師變成投機性預測者，因為：

- 投機性預測曾是以往放款危機的成因之一，
- 法院希望估價與可觀察的證據掛鉤，
- 而銀行監管機關則尋求標準化、可辯護的估價實務。

於是，估價理論朝著以下方向演進：

¹根據 1998 年《屋主保護法》(Homeowners Protection Act)，常規貸款中由借款人支付的私人抵押貸款保險在 LTV 降至 78% 時自動終止，在 80% 時可申請終止。(United States Congress 1998) 聯邦住宅管理局的保費規則更為嚴格：對於 2013 年 6 月 3 日及以後承作、原始貸款價值比高於 90% (例如頭期款 3.5%) 的貸款，年度保費貫穿貸款整個存續期，而 LTV 不高於 90% 的貸款則在十一年後取消。(U.S. Department of Housing and Urban Development 2013) 因此，許多 FHA 借款人只能透過轉貸改為接近 80% LTV 的常規貸款來擺脫抵押貸款保險。

- 現值、
- 當前市場證據、
- 明確的生效日期、
- 以及受約束的假設。

與此同時，貸款機構、投資人和擔保品分析師卻在機率性的未來風險框架中例行使用這些同樣的估價。

現實中，抵押貸款授信審核是一個綜合系統，涉及：

1. 當前市場估值、
2. 未來風險估計、
3. 違約機率、
4. 法拍時點、
5. 清算損失嚴重度、
6. 總體經濟預期、
7. 以及投資組合壓力分析。

估價師正式貢獻的主要是第 (1) 項，而貸款機構在內部將這一結果與其餘各項相結合。

這一區分在實務中變得模糊。如果貸款機構只有在相信擔保品未來大機率能保有足夠價值時才肯放款，那麼估價實際上就間接地充當了未來價值保護模型的一部分——即使估價準則在正式表述上迴避這樣描述它。

隨著以下因素的出現，這一張力愈發突出：

- 自動估價模型 (AVM)、
- 機構投資組合分析、
- 證券化模型、
- 機器學習估值系統、
- 巴塞爾資本建模、
- 以及大規模抵押貸款壓力測試。

現代擔保品分析日益將財產價值視為隨時間變化的機率量，而非單點靜態估計。(Basel Committee on Banking Supervision 2017; Steurer, Hill, and Pfeifer 2021; Board of Governors of the Federal Reserve System 2021)

可以認為，傳統的估價監管反映了一種歷史性妥協：

- 估價師提供有紀律的、基於證據的現值之錨，
- 而金融機構承擔未來風險解讀的責任。

這種分工在現代金融體系中是否仍然完全自治，是一個正當且日益受到討論的問題。

由此可以得出對估價實務本身的一個現實啟示。大多數估價師並未意識到這一擔保品估值角色——貸款機構將市場價值用作前瞻性緩衝的做法在估價業務內部很難看到——因而沒有以該角色所要求的、對未來價值走勢的審慎態度來對待自己的市場價值結論。認識到現值意見實際上是如此被使用的，是彌合這一落差的第一步，也是本文希望為後續研究開啟的領域。

```
%% appraisal_proxy.pl
%%
%% A Prolog knowledge base encoding the thesis that market value
%% appraisals, while formally framed as present-value opinions, function
%% in practice as implicit proxies for future collateral recoverability
%% within mortgage underwriting and lender risk management.

:- module(appraisal_proxy,
    [ % --- central thesis ---
      proxy_thesis/1

      % --- LTV-as-buffer mechanism ---
      , proxy_ltv_buffer_against/1
```

```

, proxy_historical_ltv_threshold/1

% --- formal vs practical interpretation ---
, proxy_formal_framing/1
, proxy_practical_interpretation/1
, proxy_underlying_question/1

% --- the artificial separation ---
, proxy_separation_party/2      % Party, Role
, proxy_doctrine_practice_gap/2 % FormalSide, PracticalSide

% --- historical compromise ---
, proxy_historical_reason/1
, proxy_doctrine_emphasis/1
, proxy_routine_practical_use/1

% --- modern intensifiers ---
, proxy_modern_driver/1
, proxy_modern_treatment/1

% --- responsibility allocation ---
, proxy_appraiser_responsibility/1
, proxy_lender_responsibility/1

% --- meta-claim about coherence ---
, proxy_open_question/1

% --- practical implication for appraisal practice ---
, proxy_practical_implication/1
, proxy_appraiser_blindspot/1
, proxy_recommended_practice/1
, proxy_further_work/1

% --- underwriting system view ---
, underwriting_component/2      % Number, Component
, underwriting_responsibility/2 % Component, Party
]).

:- discontiguous proxy_thesis/1.
:- discontiguous underwriting_component/2.

%% =====
%% PART 1: THE CENTRAL THESIS
%% =====

proxy_thesis(
  'A market value appraisal used in mortgage underwriting often serves as an
  ↪ implicit proxy for expected collateral recoverability over a future holding
  ↪ period.').
proxy_thesis(
  'The appraisal is formally framed as a present-market-value opinion as of a
  ↪ specific effective date.').
proxy_thesis(
  'The lender economically interprets the appraisal through a forward-looking
  ↪ risk lens.').
proxy_thesis(

```

```

    'There is a deliberate conceptual separation between the appraiser's
    ↪ current-market-value opinion and the lender's use of it as a forecast
    ↪ buffer.').
proxy_thesis(
    'The appraisal indirectly functions as part of a future-value protection model,
    ↪ even though appraisal standards formally avoid describing it that way.').
proxy_thesis(
    'Whether the separation between appraisal doctrine and lender practice remains
    ↪ fully coherent in modern financial systems is an open and increasingly
    ↪ debated question.').

%% =====
%% PART 2: THE LTV-AS-BUFFER MECHANISM
%% =====

proxy_historical_ltv_threshold(eighty_percent).

proxy_ltv_buffer_against(moderate_market_declines).
proxy_ltv_buffer_against(foreclosure_expenses).
proxy_ltv_buffer_against(liquidation_discounts).
proxy_ltv_buffer_against(carrying_costs).
proxy_ltv_buffer_against(legal_expenses).
proxy_ltv_buffer_against(accrued_interest).
proxy_ltv_buffer_against(uncertainty_during_disposition).

%% =====
%% PART 3: FORMAL FRAMING VS. PRACTICAL INTERPRETATION
%% =====

proxy_formal_framing(present_market_value_opinion).
proxy_formal_framing(opinion_as_of_specific_effective_date).
proxy_formal_framing(grounded_in_current_market_evidence).

proxy_practical_interpretation(forward_looking_risk_lens).
proxy_practical_interpretation(forecast_buffer_against_future_impairment).
proxy_practical_interpretation(input_to_future_collateral_recoverability_estimate).

proxy_underlying_question(
    'If this borrower defaults sometime in the foreseeable future, is the
    ↪ collateral likely to remain sufficient to cover the outstanding debt and
    ↪ foreclosure-related losses?').

%% =====
%% PART 4: THE ARTIFICIAL SEPARATION
%% =====

proxy_separation_party(appraiser, opinion_of_current_market_value).
proxy_separation_party(lender, forecast_buffer_against_future_impairment).
proxy_separation_party(appraiser, evidence_based_present_value_anchor).
proxy_separation_party(lender, future_risk_interpretation).

proxy_doctrine_practice_gap(present_market_value_opinion,
                             implicit_future_value_proxy).
proxy_doctrine_practice_gap(opinion_as_of_effective_date,
                             forecast_over_foreseeable_holding_period).
proxy_doctrine_practice_gap(evidence_based_present_anchor,
                             future_risk_buffer).
proxy_doctrine_practice_gap(static_single_point_estimate,

```

```

        probabilistic_value_over_time).

%% =====
%% PART 5: HISTORICAL COMPROMISE
%% =====

proxy_historical_reason(speculative_forecasting_contributed_to_past_crises).
proxy_historical_reason(courts_required_observable_evidence).
proxy_historical_reason(regulators_sought_standardized_defensible_practices).

proxy_doctrine_emphasis(present_value).
proxy_doctrine_emphasis(current_market_evidence).
proxy_doctrine_emphasis(defined_effective_dates).
proxy_doctrine_emphasis(constrained_assumptions).

proxy_routine_practical_use(lenders_use_in_probabilistic_future_risk_frameworks).
proxy_routine_practical_use(investors_use_in_probabilistic_future_risk_frameworks).
proxy_routine_practical_use(collateral_analysts_use_in_probabilistic_future_risk_f
↪ rameworks).

%% =====
%% PART 6: MODERN INTENSIFIERS
%% =====

proxy_modern_driver(avms).
proxy_modern_driver(institutional_portfolio_analytics).
proxy_modern_driver(securitization_models).
proxy_modern_driver(machine_learning_valuation_systems).
proxy_modern_driver(basel_capital_modeling).
proxy_modern_driver(large_scale_mortgage_stress_testing).

proxy_modern_treatment(probabilistic_over_time).
proxy_modern_treatment(distributional_rather_than_point_estimate).
proxy_modern_treatment(dynamic_rather_than_static).

%% =====
%% PART 7: ALLOCATION OF RESPONSIBILITY UNDER THE COMPROMISE
%% =====

proxy_appraiser_responsibility(disciplined_present_value_anchor).
proxy_appraiser_responsibility(evidence_based_valuation).
proxy_appraiser_responsibility(current_market_value_opinion).

proxy_lender_responsibility(future_risk_interpretation).
proxy_lender_responsibility(default_probability_estimation).
proxy_lender_responsibility(foreclosure_timing_estimation).
proxy_lender_responsibility(liquidation_severity_estimation).
proxy_lender_responsibility(portfolio_stress_analysis).
proxy_lender_responsibility(macroeconomic_expectation_integration).

%% =====
%% PART 8: OPEN QUESTION (META-CLAIM)
%% =====

proxy_open_question(
    'Whether the separation between disciplined present-value appraisal and
    ↪ lender-side future-risk interpretation remains fully coherent in modern
    ↪ financial systems.').

```

```

proxy_open_question(
    'Whether modern probabilistic-over-time collateral analytics make the
    ↪ formal/practical separation untenable in some sectors.').

%% =====
%% PART 9: THE INTEGRATED MORTGAGE UNDERWRITING SYSTEM
%% =====

underwriting_component(1, current_market_valuation).
underwriting_component(2, future_risk_estimation).
underwriting_component(3, default_probability).
underwriting_component(4, foreclosure_timing).
underwriting_component(5, liquidation_severity).
underwriting_component(6, macroeconomic_expectations).
underwriting_component(7, portfolio_stress_analysis).

underwriting_responsibility(current_market_valuation, appraiser).
underwriting_responsibility(future_risk_estimation, lender).
underwriting_responsibility(default_probability, lender).
underwriting_responsibility(foreclosure_timing, lender).
underwriting_responsibility(liquidation_severity, lender).
underwriting_responsibility(macroeconomic_expectations, lender).
underwriting_responsibility(portfolio_stress_analysis, lender).

%% =====
%% PART 10: PRACTICAL IMPLICATION FOR APPRAISAL PRACTICE
%% =====

proxy_practical_implication(
    'Most appraisers are not aware that market value functions, via the
    ↪ loan-to-value ratio, as a proxy for future foreclosure value.').
proxy_practical_implication(
    'The lender''s forward-looking use of market value is rarely visible from
    ↪ within the appraisal assignment.').
proxy_practical_implication(
    'Appraisers therefore do not treat market value conclusions with the caution
    ↪ toward future value trends that the collateral-valuation role warrants.').

proxy_appraiser_blindspot(collateral_valuation_role).
proxy_appraiser_blindspot(forward_looking_use_of_market_value).
proxy_appraiser_blindspot(market_value_as_ltv_proxy_for_foreclosure_value).

proxy_recommended_practice(recognize_how_present_value_opinion_is_consumed).
proxy_recommended_practice(treat_market_value_with_caution_re_future_trends).

proxy_further_work(closing_the_doctrine_practice_gap).
proxy_further_work(equipping_appraisers_for_the_collateral_valuation_perspective).

```

6 結論

前述知識庫所編碼的五段論述，從五個互補的視角描述了同一個職業。孤立地看，每段論述都是一份自成一體的闡述：它們所承載的社會與公共利益關切、估價所支持的經濟功能、其在擔保品風險分析中的操作角色、將其與相鄰量化工作分隔開來的監管邊界，以及其正式表述與實際用途之間的張力。合在一起，並藉助 Prolog 將結構顯性化之後，它們描述的是某種更有意味的東西——一個其官方理論與經濟功能被刻意分置的職業，其內在自洽性正日益受到如今施加於其產出之上的分析工具的考驗。

前三個知識庫共享一個共同模式：effect(Quality, Horizon, Domain, Effect)。每一個都將估價品質框定為一種因果輸入，其後果在短期、中期與長期跨度上傳播，區別僅在於解讀這些後果所經由的領域。經濟視角透過信用市場、投資配置、租稅和資本形成來解讀；社會視角透過公平、公眾信心和社區結果來解讀；擔保品視角透過授信審核、投資組合管理和金融體系穩定來解讀。共享模式並非巧合。它反映了貫穿三段論述的一個實質性主張：估值工作的品質是有後果的，這些後果隨時間放大，而同樣的底層失效模式會因所關注利益主體的不同而以不同方式顯現。謂詞參數側的共享原子——尤其是品質與時間跨度詞彙——使三套詞彙表得以互通，從而像系統性高估這樣的單一失效能夠同時經由經濟上的錯誤定價、社會上的不公與擔保品上的錯誤計量得到追蹤。

第四個知識庫轉換了論域。它描述的不是估價品質的後果，而是估價在方法論和法律意義上是什麼。其謂詞是分類性的而非因果性的，其推理機制回答的是管轄問題而非影響問題。這一轉換是必要的，因為前三段論述的後果式框架預設了對被評價活動的定義，而這一定義本身是有爭議的。scope 謂詞記錄估價師做什麼——基於證據的現值意見，輔以邊界狹窄的前瞻性成分——而 regulation 謂詞記錄該活動落入持照執業範圍的條件。前向鏈推理規則(requires_license/1、exempt_from_appraisal_license/1、regulated_appraisal/1)提供了一種機制，用以詢問某項估值工作落在受監管邊界之內還是之外——這個問題是前三個知識庫預設已經得到回答的。

第五個知識庫提出了一個論題。其他知識庫或描述、或分類、或追蹤後果，而 proxy 知識庫則論證：現值估價與未來價值預測之間審慎的理論分隔——即 regulation 知識庫所記錄的那種分隔——與估價在抵押貸款中的實際經濟功能存在張力。80% 的貸款價值比慣例、它所緩衝的七類具體風險、逐字記錄的貸款機構的底層問題（“如果這位借款人在可預見的未來某個時候違約，擔保品是否仍可能足以...”），以及由七個組成部分構成的綜合授信審核系統，都支持同一個主張：估價在形式上是現值意見，在經濟上是未來價值緩衝。促成這一安排的歷史性妥協——估價師充當有紀律的現值之錨，貸款機構充當未來風險解讀的中樞——在估價執照制度形成的歷史條件下是一種可以辯護的因應，但文中列舉的現代強化因素（AVM、機器學習估值系統、巴塞爾資本建模、證券化分析、大規模壓力測試）將財產價值作隨時間變化的機率化處理。這些工具把在理論上被分隔的層次重新拉回同一條運算流程，層與層之間的概念分隔也因此愈發難以維繫。

這種調和是否可以實現、以何種條件實現，正是 proxy 段落收束時提出的問題。這些知識庫並未回答它。但它們確實為回答該問題提供了一份明確、可查詢的相關考量清單：估值品質在各時間跨度與領域中的後果、決定哪些後果落入受監管實務的管轄邊界，以及該邊界如今承壓的結構性原因。這份闡述是進一步建構更具體論證——關於方法論、關於準則改革、關於估價實務與更廣泛的量化估值體系之間關係——的基礎。

關於知識表示的說明

有三個結構性模式值得明確記錄，因為它們產生於編碼過程本身，而非源自任何單獨一段論述。

第一，這一知識庫家族並不是扁平的。五個知識庫中有三個共享同一模式；一個提供決定該模式何時適用的管轄性限定；一個就正式表述與經濟功能之間的落差提出論題。將五者結合的橋接層並不把它們當作同一矩陣的五個直欄。它把前三者視為對後果的描述，把第四個視為對這些描述在何處有效的約束，把第五個視為對該約束本身承壓之處的識別。這種非對稱結構是特性而非缺陷：它映照了專業實務的實際組織方式——在其中，描述性主張、管轄性主張與批判性主張佔據著確實不同的認識論角色。

第二，區分各模組的前綴紀律 (economic_、social_、collateral_、scope_/regulation_、proxy_/underwriting_) 只有在存在衝突風險之處才真正發揮作用。在各模組間有直接對應物的謂詞才配得上前綴；某一概念領域獨有的謂詞則不需要。參數側的共享取值詞彙——品質、時間跨度和領域原子——恰恰是各模組得以互通的關鍵，給這些取值加前綴反而會破壞這一點。由此浮現的模式具有普遍性：謂詞加前綴，參數共享原子，並且在謂詞為單一視角獨有時採用選擇性前綴而非一律加前綴。

第三，proxy 知識庫所識別的核心張力與 scope/regulation 知識庫所識別的監管不對稱密切相關，卻又彼此不同。兩者透過平行的二元謂詞——regulatory_asymmetry/2 與

proxy_doctrine_practice_gap/2——加以記錄，並刻意採用不同的名稱，因為它們描述的是同一底層安排的不同軸線。監管不對稱關乎誰受監管、監管程度多深；理論與實務落差關乎對估價師產品的正式表述與其實際經濟用途之間的落差。兩種張力相互強化。監管把估價師推向現值理論；經濟功能把他們的產出拉向未來價值用途。合起來的圖景是：這一職業的官方自我描述，與其產品所匯入的分析流程之間，正變得越來越難以調和。

參考文獻

- Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao (2015). “Collateral Valuation and Borrower Financial Constraints: Evidence from the Residential Real Estate Market”. In: *Management Science* 61.9, pp. 2220–2240. DOI: [10.1287/mnsc.2014.2002](https://doi.org/10.1287/mnsc.2014.2002).
- Alexandrov, Alexei, Laurie Goodman, and Michael Neal (Jan. 2023). *Reengineering the Appraisal Process: Better Leveraging Both Automated Valuation Models and Manual Appraisals*. Tech. rep. Urban Institute. URL: <https://www.urban.org/sites/default/files/2023-01/Reengineering%20the%20Appraisal%20Process.pdf>.
- Basel Committee on Banking Supervision (2017). *Basel III: Finalising Post-Crisis Reforms*. Standards d424. Bank for International Settlements. URL: <https://www.bis.org/bcbs/publ/d424.pdf>.
- Board of Governors of the Federal Reserve System (Mar. 2021). *Dodd-Frank Act Stress Test 2021: Supervisory Stress Test Methodology*. Tech. rep. Board of Governors of the Federal Reserve System. URL: <https://www.federalreserve.gov/publications/files/2021-april-supervisory-stress-test-methodology.pdf>.
- Campbell, John Y., Stefano Giglio, and Parag Pathak (2011). “Forced Sales and House Prices”. In: *American Economic Review* 101.5, pp. 2108–2131. DOI: [10.1257/aer.101.5.2108](https://doi.org/10.1257/aer.101.5.2108).
- Clauretie, Terrence M. and Nasser Daneshvary (2009). “Estimating the House Foreclosure Discount Corrected for Spatial Price Interdependence and Endogeneity of Marketing Time”. In: *Real Estate Economics* 37.1, pp. 43–67. DOI: [10.1111/j.1540-6229.2009.00235.x](https://doi.org/10.1111/j.1540-6229.2009.00235.x).
- Crosby, Neil, Colin M. Lizieri, and Patrick M. McAllister (2010). “Means, Motive and Opportunity? Disentangling Client Influence on Performance Measurement Appraisals”. In: *Journal of Property Research* 27.2, pp. 181–201. DOI: [10.1080/09599916.2010.499014](https://doi.org/10.1080/09599916.2010.499014).
- Deng, Yongheng, John M. Quigley, and Robert Van Order (2000). “Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options”. In: *Econometrica* 68.2, pp. 275–307. DOI: [10.1111/1468-0262.00110](https://doi.org/10.1111/1468-0262.00110).
- Diaz Julian, III and Marvin L. Wolverton (1998). “A Longitudinal Examination of the Appraisal Smoothing Hypothesis”. In: *Real Estate Economics* 26.2, pp. 349–358. DOI: [10.1111/1540-6229.00749](https://doi.org/10.1111/1540-6229.00749).
- Ebrahim, M. Shahid and Sikandar Hussain (2010). “Financial Development and Asset Valuation: The Special Case of Real Estate”. In: *Journal of Banking & Finance* 34.1, pp. 150–162. DOI: [10.1016/j.jbankfin.2009.07.011](https://doi.org/10.1016/j.jbankfin.2009.07.011).
- Eriksen, Michael D. et al. (2019). “The Influence of Contract Prices and Relationships on Appraisal Bias”. In: *Journal of Urban Economics* 111, pp. 132–143. DOI: [10.1016/j.jue.2019.03.001](https://doi.org/10.1016/j.jue.2019.03.001).
- Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen (2008). “Negative Equity and Foreclosure: Theory and Evidence”. In: *Journal of Urban Economics* 64.2, pp. 234–245. DOI: [10.1016/j.jue.2008.07.006](https://doi.org/10.1016/j.jue.2008.07.006).
- Foster, Chester and Robert Van Order (1984). “An Option-Based Model of Mortgage Default”. In: *Housing Finance Review* 3.4, pp. 351–372.
- Friedman, Jerome H. (1991). “Multivariate Adaptive Regression Splines”. In: *The Annals of Statistics* 19.1, pp. 1–67. DOI: [10.1214/aos/1176347963](https://doi.org/10.1214/aos/1176347963).
- Gallimore, Paul (1996). “Confirmation Bias in the Valuation Process: A Test for Corroborating Evidence”. In: *Journal of Property Research* 13.4, pp. 261–273. DOI: [10.1080/095999196368781](https://doi.org/10.1080/095999196368781).

- Geltner, David M. (1991). "Smoothing in Appraisal-Based Returns". In: *The Journal of Real Estate Finance and Economics* 4.3, pp. 327–345. DOI: [10.1007/BF00161933](https://doi.org/10.1007/BF00161933).
- Ghent, Andra C., Walter N. Torous, and Rossen I. Valkanov (2019). "Commercial Real Estate as an Asset Class". In: *Annual Review of Financial Economics* 11, pp. 153–171. DOI: [10.1146/annurev-financial-110118-123121](https://doi.org/10.1146/annurev-financial-110118-123121).
- Griffin, John M. and Gonzalo Maturana (2016). "Who Facilitated Misreporting in Securitized Loans?" In: *The Review of Financial Studies* 29.2. Evidence of appraisal inflation and its consequences, pp. 384–419. DOI: [10.1093/rfs/hhv130](https://doi.org/10.1093/rfs/hhv130).
- Hendershott, Patric H. and Edward J. Kane (July 1992). *Office Market Values During the Past Decade: How Distorted Have Appraisals Been?* NBER Working Paper 4128. Available at SSRN: <https://ssrn.com/abstract=226786>. National Bureau of Economic Research. DOI: [10.3386/w4128](https://doi.org/10.3386/w4128). URL: <https://www.nber.org/papers/w4128>.
- (1995). "U.S. Office Market Values During the Past Decade: How Distorted Have Appraisals Been?" In: *Real Estate Economics* 23.2, pp. 101–116. DOI: [10.1111/1540-6229.00660](https://doi.org/10.1111/1540-6229.00660).
- Hull, John C. (2018). *Options, Futures, and Other Derivatives*. 10th ed. New York: Pearson. ISBN: 978-0-13-447208-9.
- International Association of Assessing Officers (2017). *Standard on Mass Appraisal of Real Property*. International Association of Assessing Officers, Kansas City, MO. URL: https://www.iaao.org/wp-content/uploads/Standard_on_Mass_Appraisal.pdf.
- International Valuation Standards Council (2025). *International Valuation Standards (IVS)*. International Valuation Standards Council, London. Published 31 January 2024; effective 31 January 2025. URL: <https://www.ivsc.org/>.
- K'Akumu, Owiti A. and James E. Larsen (2024). "A Primer on Regulations and the Practice of Residential Property Appraisal". In: *Journal of Real Estate Practice and Education*. Focus on regulatory and public interest aspects. DOI: [10.1080/15214842.2024.2377869](https://doi.org/10.1080/15214842.2024.2377869).
- Kau, James B. et al. (1995). "The Valuation at Origination of Fixed-Rate Mortgages with Default and Prepayment". In: *The Journal of Real Estate Finance and Economics* 11.1, pp. 5–36. DOI: [10.1007/BF01097934](https://doi.org/10.1007/BF01097934).
- LaCour-Little, Michael and Stephen Malpezzi (2003). "Appraisal Quality and Residential Mortgage Default: Evidence from Alaska". In: *The Journal of Real Estate Finance and Economics* 27.2, pp. 211–233. DOI: [10.1023/A:1024728420837](https://doi.org/10.1023/A:1024728420837).
- Lee, Ming-Hsiang, Chien-Wen Peng, and Hui-Fen Liao (2020). "An Analysis of Objectivity in the Real Estate Appraisal Process". In: *International Real Estate Review* 23.4, pp. 483–504.
- Mayer, Yanling G. and Frank E. Nothaft (2022). "Appraisal Overvaluation: Evidence of Price Adjustment Bias in Sales Comparisons". In: *Real Estate Economics* 50.3, pp. 862–881. DOI: [10.1111/1540-6229.12351](https://doi.org/10.1111/1540-6229.12351).
- Murphy, Edward V. (2012). *Regulation of Real Estate Appraisers*. CRS Report RS22953. Congressional Research Service. URL: <https://sgp.fas.org/crs/misc/RS22953.pdf>.
- Office of the Comptroller of the Currency and Board of Governors of the Federal Reserve System and Federal Deposit Insurance Corporation and National Credit Union Administration and Consumer Financial Protection Bureau and Federal Housing Finance Agency (2024). *Quality Control Standards for Automated Valuation Models*. Final Rule, 89 Fed. Reg. 64538 (Aug. 7, 2024). URL: <https://www.federalregister.gov/documents/2024/08/07/2024-16197/quality-control-standards-for-automated-valuation-models>.
- Pennington-Cross, Anthony (2006). "The Value of Foreclosed Property". In: *Journal of Real Estate Research* 28.2, pp. 193–214. DOI: [10.1080/10835547.2006.12091177](https://doi.org/10.1080/10835547.2006.12091177).
- Qi, Min and Xiaolong Yang (2009). "Loss Given Default of High Loan-to-Value Residential Mortgages". In: *Journal of Banking & Finance* 33.5, pp. 788–799. DOI: [10.1016/j.jbankfin.2008.09.010](https://doi.org/10.1016/j.jbankfin.2008.09.010).

- Quan, Daniel C. and John M. Quigley (1991). “Price Formation and the Appraisal Function in Real Estate Markets”. In: *The Journal of Real Estate Finance and Economics* 4.2, pp. 127–146. DOI: [10.1007/BF00173120](https://doi.org/10.1007/BF00173120).
- Steurer, Miriam, Robert J. Hill, and Norbert Pfeifer (2021). “Metrics for Evaluating the Performance of Machine Learning Based Automated Valuation Models”. In: *Journal of Property Research* 38.2, pp. 99–129. DOI: [10.1080/09599916.2020.1858937](https://doi.org/10.1080/09599916.2020.1858937).
- The Appraisal Foundation (2024). *Uniform Standards of Professional Appraisal Practice (USPAP), 2024 Edition*. The Appraisal Foundation, Appraisal Standards Board, Washington, DC. Effective January 1, 2024. URL: <https://www.appraisalfoundation.org/>.
- U.S. Department of Housing and Urban Development (2013). *Revision of Federal Housing Administration Policies Concerning Cancellation of the Annual Mortgage Insurance Premium*. Mortgagee Letter 2013-04, Federal Housing Administration. Effective June 3, 2013; the annual MIP continues for the life of loans with original LTV above 90%, and for 11 years at or below 90%. URL: <https://www.hud.gov/sites/documents/13-04ml.pdf>.
- United States Code (1970). *Federal National Mortgage Association and Federal Home Loan Mortgage Corporation Charter Acts*. 12 U.S.C. § 1717(b)(2) (Fannie Mae) and 12 U.S.C. § 1454(a)(2) (Freddie Mac). Prohibit GSE purchase of conventional mortgages exceeding 80% LTV absent credit enhancement (private mortgage insurance, seller recourse/repurchase, or a retained seller participation).
- United States Congress (1989). *Financial Institutions Reform, Recovery, and Enforcement Act of 1989*. Pub. L. No. 101-73, 103 Stat. 183; Title XI codified at 12 U.S.C. §§ 3331–3356. URL: <https://www.govinfo.gov/content/pkg/STATUTE-103/pdf/STATUTE-103-Pg183.pdf>.
- (1998). *Homeowners Protection Act of 1998*. Pub. L. No. 105-216, 112 Stat. 897; codified at 12 U.S.C. §§ 4901–4910. Automatic termination of borrower-paid private mortgage insurance at 78% LTV; cancellation may be requested at 80%.

殘差約束法 (RCA)：框架與規程

本文為 2025 年 3 月 5 日最初發表於 Zenodo 的論文之更新版的繁體中文譯本。譯文藉助人工智慧輔助完成並經編輯部審校；英文版為權威版本。

William Bert Craytor, SRA

第 1 卷 第 1 期 · 2026 年夏季

發表於 2026 年 7 月 1 日

DOI: [10.67330/kgebf46](https://doi.org/10.67330/kgebf46)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據創用 CC 姓名標示 4.0 國際授權條款 (CC BY 4.0) 的條款發布。

摘要。傳統估價方法的一個重大侷限在於其依賴過時的手工技術，例如配對比較分析 (matched pair analysis)。這類方法對不動產特徵的價值貢獻視而不見：傳統的銷售網格 (sales grid) 沒有價值貢獻一欄，只有調整一欄。而事實上，價值貢獻正是建構數學約束的必要基礎——這些約束能夠防止傳統估價師的高估與低估。

殘差約束法 (Residual Constraint Approach, RCA) 透過整合一個多階段估價流程來推進傳統的成交案例比較法 (Sales Comparison Approach, SCA)：該流程運用多元自適應迴歸樣條 (Multivariate Adaptive Regression Splines, MARS)，並對特徵價值貢獻施加嚴格的數學約束，尤其針對未被測量的特徵——諸如美學、設計、狀況與品質等潛變數——其價值在傳統估價中通常受制於估價師的主觀判斷與偏誤。

RCA 中的數學約束依賴於對 MARS 迴歸的專業運用，以估計可測量不動產特徵對成交價格的價值貢獻。然而，這一估計通常只能解釋實際成交價格的約 80%，不包括設計、功能效用、狀況和品質等主觀評估的特徵。在舊金山灣區 (San Francisco Bay Area)，代表這些主觀成分的殘差平均約占不動產總價值的 20%。80/20 這一數字反映的是作者在舊金山灣區典型資料條件下的經驗，不應被解讀為跨市場不變的常數。在可實現 R^2 較低的市場中，本框架的應對之道是識別並測量未被測量的價值驅動因素——例如沿海社區中與海洋的距離、山坡社區中的海拔高度，以及當 MLS 資料未提供時從外部來源收集的類似變數——而不是接受這一侷限並報告更寬的誤差區間。雖然在其他情境下 MARS 殘差通常被視為估計誤差，但在不動產估價中，殘差是潛變數價值的間接度量。儘管這初看似似乎不切實際，但已有證明表明：只要各描述性成分的貢獻之和等於殘差，殘差就可以被有意義地分解為這些成分——且不影響最終的估價結果。

這一方法論源自作者二十年的實證應用經驗，透過在 R 與 Python 環境中對 MARS 的迭代實作逐步發展而成。

關鍵詞：不動產估價，住宅估價，殘差約束法，MARS，多元自適應迴歸樣條，成交案例比較法，殘差分析，潛變數，自動估價

目錄

1 引言	3
2 背景	5
3 估價工程師	7
4 假設	8

5 RCA 工作流程	10
5.1 MARS 分析的類型	10
5.2 處理工作流程	10
5.3 資料工作流程	11
5.3.1 MARS 的資料要求	11
5.3.2 規程：已成交交易掛牌資訊	12
5.4 專案資料夾結構	14
5.4.1 規程：多次 MARS (earth) 執行	15
5.5 核心 Excel 活頁簿	16
5.6 備份	16
6 活頁簿記號	17
6.1 描述維度	17
6.1.1 專案維度 (“a”)	18
6.1.2 執行版本維度 (“r”)	18
6.1.3 工作流階段 (“w”)	19
6.1.4 Excel 活頁簿工作表 (“s”)	19
6.1.5 不動產掛牌資訊 (“p”) ——MLSData 列	20
6.1.6 MLS 變數 (“v”) ——MLSData 欄	20
6.1.7 最終記號示例	20
6.1.8 理由	21
6.1.9 矩陣?	21
6.1.10 第一次簡化	21
6.1.11 第二次簡化	22
7 RCA 銷售網格	23
8 變數	26
8.1 變數來源本體	26
8.2 變數應用本體	27
8.3 二階與三階項	28
8.3.1 調整變數	28
8.3.2 匯總變數	29
8.3.3 特殊變數 & 記號	29
8.3.4 分解殘差	30
9 RCA 核心證明	31

10 MARS 模型	32
10.1 模型變數與交互作用的「 g 」函數	32
10.2 g 函數的索引	32
10.3 繪圖維度	33
10.4 最終模型方程	33
10.5 繪製一個模型所需的圖形數量	34
11 資料特徵	36
11.1 殘差的用途	36
11.2 作為近似迴歸變數的狀況與品質評級	36
11.3 資料錯誤	37
11.4 什麼是準確性？	38
11.5 作為異常值的估價對象	39
11.6 CVR2 在應對意外資料方面的重要性	39
12 殘差	41
12.1 排序與評分	41
12.2 CQA-殘差曲線	43
12.3 CQA 曲線的特徵	43
12.3.1 CQA 曲線各組成部分的分析	43
12.3.2 市場動態與社會含義	44
12.4 估價對象殘差	44
12.5 RCA 價值結論	45
12.5.1 一般考慮	45
12.5.2 複核、審計與投訴	45
13 結論	46

1 引言

本文闡述了一種現代方法的邏輯與數學基礎，該方法旨在準確估計住宅及其他不動產的市場價值（market value）。這一方法源自 20 年來將多元自適應迴歸樣條（Multivariate Adaptive Regression Splines, MARS）軟件應用於北加州住宅——尤其是舊金山灣區各社區——的實務經驗，在估價精度上實現了顯著進步。MARS 能夠嫺熟地捕捉不動產特徵與價值之間的非線性關係，同時高效地識別並優先處理影響最大的變數，從而提升估價品質。這一能力使估價師（appraiser）能夠以遠快於傳統方法的速度發現新市場區域中的價值動態。MARS 構成了本文提出的殘差約束法（Residual Constraint Approach, RCA）的骨幹。RCA 方法之所以得名，源於其嚴謹的框架：

- 它對已測量和未測量不動產特徵的價值貢獻（value contribution）均施加嚴格約束：

1. 所有由 MARS 導出的價值貢獻，包括殘差（residual）在內，其總和必須等於每個可比交易案例（comparable）的淨售價。
 2. 殘差可以分解為具有指定價值的子成分，前提是這些子成分的總和等於該可比交易案例的殘差。
- RCA 透過將主觀特徵——如美學吸引力、狀況和品質——的總價值與殘差掛鉤，最大限度地減少了估計此類特徵時的偏誤；當由勝任的估價工程師（valuation engineer）構建 MARS 模型時，殘差能夠被準確量化。

儘管 RCA 主要為住宅不動產設計——在此類不動產中主觀特徵對市場價值影響顯著——但它在更廣泛的資產估價應用中亦具潛力。鑑於美國住宅不動產市場的估計價值約達 \$45 兆，準確的估價至關重要。Fannie Mae 指出，典型估價與實際售價的偏差約為 5%，但這一數字掩蓋了偏誤問題，因為估價師往往面臨與售價保持一致的壓力。此外，由於 60%–70% 的估價用於再融資——缺乏可比售價——RCA 所能提供的精度之必要性便顯而易見。美國聯準會 2012 年的分析估計，2008 年危機前有 15%–20% 的估價與售價的偏差超過 $\pm 10\%$ ，這一範圍如今很可能仍適用於再融資估價，並因不動產類型和市場而異。準確性為何重要？對於低貸款價值比（LTV）的交易——例如信用良好的買方在 \$1 million 的住宅上支付 \$500,000 頭期款——貸款機構可能免除估價，認為法拍風險極小。然而，跳過估價的買方有多付的風險，往往在購買之後才發現缺陷。這凸顯了 RCA 這類穩健估價方法的價值。量化估價的準確性仍然頗具挑戰。研究者缺乏針對「市場價值」偏差的直接度量指標，只能依賴間接代理。傳統估價方法根植於主觀判斷，且受估價師勝任程度參差不齊的影響，使分析更加複雜。相關策略包括：

1. Tzioumis (2016) 研究了 11 個州 4,782 名估價師的 259,436 份估價報告，發現三分之一的估價與合約價的偏差在 1% 以內，一半高出合約價 1%–3%，少數平均高出 5%——表明存在向上偏誤。該研究識別出兩類估價師：「按市場價值」型（接近合約價）和「高於市場價值」型（持續偏高）。
2. Krivorotov and LaCour-Little (2020) 發現 90% 的購屋估價達到或超過合約價，凸顯了「錨定偏誤」——向某一期望價值靠攏的傾向——在再融資情境中尤為明顯。

這些研究揭示了估價實務中的偏誤，而傳統工具的侷限性又放大了這種偏誤。如果估價師普遍勝任且無偏誤，他們的估價結果應當趨同；某些群體持續的高估（overvaluation）表明存在系統性問題。

上述偏誤研究及其指向的系統性模式，反映了傳統成交案例比較法（Sales Comparison Approach, SCA）實務的一個更深層特徵，而這正是 RCA 旨在解決的問題。傳統方法在兩個認識論地位截然不同的管道上運作。在可測量特徵管道上——GLA、房間數、土地面積、房齡、位置——估價師被期望透過配對分析或對可比交易案例的局部線性迴歸，客觀地推導調整（adjustment）值。勝任的估價師會做這項工作，由此得到的調整至少大體上能在資料面前站得住腳。但即便在這一管道上，傳統的網格也只記錄調整；它甚至沒有一欄用於記錄價值貢獻。傳統的調整也並非作為價值貢獻之差來計算——它們是透過配對或類似程序直接估計出來的，底層的價值貢獻從未被計算過。相比之下，在 RCA 中，每一項調整都是估價對象（subject property）與可比交易案例在給定特徵上的價值貢獻之差——證明 II 正是精確刻畫了這一結構的後果。在主觀特徵管道上——狀況、施工品質、設計、功能效用、美學——可用的支持程序從單薄到近乎空洞不等。主觀特徵調整可以透過對直接可比集合之外不動產的配對分析、訴諸公認的同行慣例，或引用估價師在該市場中有記錄的經驗來支持。這三者形式上都符合專業標準。但沒有一個能達到可測量管道常規實現的統計嚴謹性。

其結果是，標準的要求與方法所能提供的支持之間存在程序上的不對稱。標準要求對每一項調整——無論可測量還是主觀——都提供支持。而方法在一個管道上提供穩健的支持，在另一個管道上則只有一系列較弱的程序，從樣本量為一的軼事式配對，到引用其他估價師的習慣做法，不一而足。誠實而勝任的估價師對景觀調整進行細緻的配對分析，其提供的支持顯著優於毫無依據地援引同行慣例者；但在標準表格上比較這兩份估價報告的審查者卻難以區分二者，

因為表格記錄的是調整值，而非其推導的嚴謹程度。方法之所以容忍這種不對稱，是因為對於主觀特徵，一直沒有更強的程序可供執業估價師普遍使用。

RCA 對這兩個管道都有改進，但方式在範疇上截然不同。在可測量管道上，MARS 迴歸能夠捕捉配對分析和普通線性迴歸無法表達的非線性關係、變數交互作用和基底函數（basis function）結構。MARS 模型是針對市場區域內全部合格銷售資料（通常為數百宗）擬合的，而非傳統網格中的三到九個可比交易案例，因此由 MARS 導出的價值貢獻既客觀，又比其所取代的方法準確得多。在主觀管道上，RCA 所做的事情與任何傳統支持程序在性質上都不相同。它根本不試圖為單個主觀特徵估計調整值。所有主觀特徵——潛變數（latent variable）——的貢獻共同表現為殘差：即每個可比交易案例的實際淨售價與其 MARS 預測價之差，可直接由資料計算，無需估價師輸入。將該殘差分解為具名成分——設計、功能效用、景觀、狀況、品質等——受到其總和必須等於殘差總額的約束，並且根據證明 II（第 9 節），任何滿足該約束的分解都會得出相同的價值結論。工作流程中剩下的判斷，是將估價對象放置在可比交易案例殘差排名分布中的位置，這一判斷是對照可檢查的排名做出的，並在報告中明確記錄。一份薄弱的 RCA 估價報告仍然可能出現——缺乏經驗的 VE 可能構建出過擬合的 MARS 模型、對估價對象做出不合理的定位，或忽視資料中的異常——但其薄弱之處是審查者透過檢查模型和排名即可識別的，而不是標準表格審查所無法揭示的那種薄弱。

除再融資之外，估價師對屋主負有提供準確估價的義務，尤其是在高價值購置中，錯誤可能造成數萬乃至數十萬的損失。然而，許多人不願在估價上花費 \$600–\$1,000，懷疑其附加價值。RCA 這類先進方法雖然初期資源投入較大，卻有望帶來更高的精度。隨著方法的改進和經驗的積累，其效率——以及採用率——應會不斷提高。

2 背景

作者的教育背景為數學、資訊科學與統計學。其職業為軟件工程師，兼具估價與會計方面的輔助經驗。他自 2002 年起斷斷續續從事估價工作，並自約 2004 年起將 MARS 應用於住宅不動產的調整值確定和價格趨勢分析。

使用 MARS 進行估價是估價領域的「深度經驗」。MARS 生成的模型能比任何其他統計方法都更細緻地反映哪些變數對市場價值有貢獻，以及這些變數取值的變化如何影響售價。因此，在使用 MARS 估價的過程中，估價師不斷考察由 MARS 模型揭示的變數取值與售價之間的因果關係。

出於透明性的考慮，RCA 方法最適用於具備以下特徵的住宅區域：

1. 不動產彼此差異很大的複雜社區

- 較老的住宅
- 更新程度參差不齊
- 多種不同的建築風格與設計
- GLA、土地面積、海拔和景觀等不動產屬性差異較大。

2. 複雜的市場狀況

- 快速變化的市場狀況
- 類型多樣的買方

3. 良好的資料來源：RCA 規程（protocol）基於資料探勘，需要充足的高品質資料。雖然它也可用於資料稀疏的區域，但可能不值得為此投入所需的時間和成本。

4. 需要準確的市場價值估計。RCA 在多數情況下應能提供 $\pm 1-2\%$ 的精度。估價工程師報告應就最終價值結論的精度條件加以說明。如果資料或資訊不足，或未能透過資料探勘發現

市場中的一致模式，精度就會受損，且應將這一情況告知客戶。具體而言，較高的 R2 和 CVR2（交叉驗證 R2，透過將 MLS 資料集反覆隨機劃分為訓練集和測試集計算得到）、有意義的 MARS 模型，以及排名殘差中不動產 CQA（狀況-品質-吸引力）特徵所呈現的強模式，是評估價值結論準確性與可靠性的主要手段。

過去 20 餘年間這些 RCA 規程的發展，依託於 MLSListings Inc. (Sunnyvale, CA) 及其他目前可透過 MLSListings 存取的 MLS 資料，這些資料佔作者資料分析經驗的 95% 以上。作者的專業能力源自在北加州 15 個主要郡進行的大量估價工作，並在 10 個次要郡積累了額外經驗。

RCA 方法在滿足以下兩個關鍵條件的市場中最為有效：

1. 高品質估價資料的豐富可得性
2. 複雜、異質的住宅區域，其特徵為：
 - 多樣的建築風格
 - 程度不一的不動產更新
 - 跨越數十年的開發模式

加利福尼亞州各郡經驗

本文所作論述基於在以下北加州各郡進行不動產估價的經驗：

San Mateo	Santa Clara	Marin	Napa	Monterey	Santa Rosa
Alameda	San Francisco	Contra Costa	Sacramento	El Dorado	San Joaquin
Stanislaus	Placer	Solano	Mendocino	Lake	Colusa
Merced	Nevada	San Benito	Santa Cruz	Sutter	Yuba

上述區域基本涵蓋矽谷及其周邊各郡，並向外延伸至更多鄉村地區，包括中央谷地和內華達山麓。應當指出，與美國其他欠發達地區的許多估價師所能獲得的資料相比，作者始終擁有充足且相對準確的資料。

所用的 MLS 資料經過 RESO 認證，符合 RCF (RESO Common Format)。儘管有 RCF 認證，一些不動產的、通常屬次要重要性的資料有時仍會缺失或有誤，其原因或是最初稅務評估員的錄入錯誤，或是後續銷售經紀人的 MLS 輸入錯誤。這些錯誤往往遵循某種模式，例如特徵不存在時乾脆不做輸入。舉例來說，在沒有車棚時，「carport spaces」一欄不輸入 0，而是什麼都不填。在車棚罕見的地區，分析師可以用 0 替代缺失資料，以便得到有用的車棚調整值。雖然 R/earth 及其他 MARS 實作可以處理缺失資料，但如何處理由估價工程師 (Valuation Engineer, VE) 決定。

MLSListings 的資料也可從其他一些州（從佛羅里達到夏威夷）的 MLS 獲得。這對於測試規程和程式碼在不同州的執行效果很有用。

3 估價工程師

定義 1：估價工程師

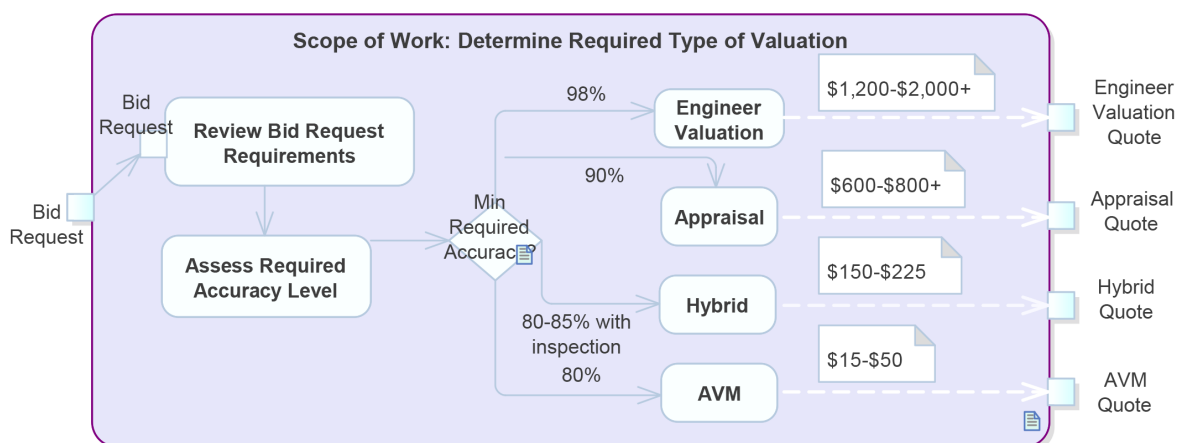
為避免歧義，「估價工程師」（Valuation Engineer，簡稱「VE」）一詞指這樣的人員：他既具備持照估價師的專業能力，又在運用資料探勘、統計學、程式設計、詳細規程和人工智慧來確定複雜不動產市場價值方面擁有高階的知識、技能與經驗。還應認識到，鑑於既已證實的詐欺可能性，估價是一項需要某種形式執業許可的工作。

貸款機構通常提供以下幾種估價選項：

1. 自動估價模型（AVMs）——成本最低的選項
2. 混合估價——由無執照查勘員實地查勘，持照估價師負責監督
3. 標準估價——由持照估價師完成

通常不在估價工作考慮範圍之內的，是工程師級估價。具備這類技能組合的估價師實在不多，但市場應當存在對他們的需求。藉助 RCA 方法（MARS 加約束），他們能夠提供顯著更高的準確性，並幾乎保證相當良好的客觀性。

對於較新住宅社區中的合規不動產，標準估價通常已經足夠。然而，對於翻新歷史長達 30 年以上區域中的老宅，則建議採用工程師估價。在這兩種方法之間的選擇應遵循基於不動產特徵和市場狀況的結構化決策流程。



4 假設

假設 1：地理區域 & 不動產類型

本文的範圍主要聚焦於加利福尼亞州大都市及城市地區的獨戶住宅與共管公寓。雖然成交案例比較法也適用於其他不動產類型和地理區域，但在對農業用地、多戶不動產和商業不動產進行估價時，收益法和成本法通常發揮更主導的作用。在此類情況下，RCA 方法可能需要修改，或適用性有限。

假設 2：資料儲存

CSV 和 Excel 在 MLS 系統中被大量使用，並被假定為向 RCA 程式饋送資料的主要方式；當然，也可以使用經紀商資料源將資料持續送入 SQL 資料庫。備份和中間儲存常使用 SQL 資料庫，其中扮演重要角色的兩個是免費開源的 SQLite 和 PostgreSQL。文字或 RTF 檔案也用於報告片段，PNG、JPG、TIFF 檔案用於圖形。需要注意的是，許多人也使用 Microsoft 的 SQL Server，但其免費版「SQL Server Express」限制很大（例如只能使用 1 個 CPU 核心），而其他版本用於商業用途則需支付高昂的許可費。加利福尼亞州的不動產資料庫可能非常龐大，連接（join）操作可能非常耗時，因此對大型連接而言，一台 8 核以上、配備 128G 記憶體電腦會很有用。當需要查找任意邊界區域內的所有不動產時，SQL 資料庫同樣重要。因此，舉例來說，如果使用群集分析來劃定社區，就會傾向於使用某種 SQL 資料庫。然而，不同資料庫處理地理資訊的內建函數各不相同，因而互不相容。在這一系列關於 RCA 的文章中，複雜的 SQL 任務假定使用 PostgreSQL，簡單任務使用 SQLite。

假設 3：MARS 處理使用「R/earth」。

2017 年之前，Salford Systems 的 MARS 軟件是一個價格親民的解決方案，每年 \$270。然而，在 Minitab 收購 Salford Systems 之後，MARS 被打包進其 SPM 軟件包，年費高達 \$16,000——這一價位令大多數小企業望而卻步。幸運的是，由 Stephen Milborrow 維護的 R 語言「earth」包 (Milborrow 2024a) 已發展到與 Minitab 的 SPM MARS 相當乃至更強的水準。earth 包還具有額外優勢：它免費可得，與 R 全面的統計生態系統無縫整合，並可透過 Python 藉助 Pandas 存取，為不同工作流程提供了出色的靈活性。與 MARS 相關的討論均假定具體使用 R/earth。

假設 4：符號約定

來自 CSV 檔案和 Excel 試算表等外部來源的原始資料用標準拉丁字元表示。這包括多重掛牌服務 (MLS) 不動產資料、使用 R/earth 包的 MARS (多元自適應迴歸樣條) 組態設定，以及專案特定資料。這一記法同樣適用於它們在 R 或 Python 中對應的資料框表示。

在討論不動產特徵及其各種類型和子集時，記法改用希臘字母，以增強精確性和分析上的清晰度。雖然這種數學記法看起來更為嚴格，但它有一個關鍵目的：將抽象的不動產特徵與其原始資料表示區分開來。按照慣例，希臘字元始終表示不動產特徵，而拉丁字元表示試算表資料結構。例如，符號 M 表示 MLS 資料矩陣，其列為各不動產，欄為相應的特徵。

假設 5：程式設計語言

在本領域，R 和搭配 Pandas 的 Python 是主導語言，在對速度要求較高時則使用 C 或 C++。R/earth 本身即以 R 和 C 編寫。

假設 6：索引

索引方式在不同程式設計語言與數學之間可能各不相同，這可能令人困惑：

- 數學和 R 從 1 開始索引。
- C、C#、C++ 和 Python 從 0 開始。

在本文中，為保持一致和簡單，所有索引均從 0 開始（如 Python 或 C）。具體規則如下：

- 對於列表、矩陣或資料框：
 - 索引 0 保留給「估價對象」（即被研究的主要不動產）。
 - 示例：在以列表表示不動產的表格中，第 0 列是估價對象，第 0 欄存放所有不動產的唯一 ID。
 - 實用提示：排列不動產時將估價對象放在最前（第 0 列），其餘不動產按銷售日期從新到舊排列，並按順序編號（0、1、2 等）。
- 當沒有估價對象時：
 - 如果資料結構不包含不動產（例如僅有變數或函數），則跳過索引 0，索引從 1 開始。
 - 給 R 使用者的特別說明：由於 R 使用 1 基索引，在用 R 實作本文的數學內容時，需將所有索引加 1，以從 0 基調整為 1 基。

5 RCA 工作流程

理解 RCA 工作流程是進行有效分析的關鍵。整個過程遵循一系列結構化的分析任務與決策節點

5.1 MARS 分析的類型

MARS 分析的工作流程取決於僅需進行市場分析，還是需要對特定不動產進行估值或估價。請注意，對特定不動產的估價通常需要對其所在市場區域進行分析。

- 市場分析：側重於郡、市或界定區域等較大範圍。這種方式不涉及具體的估價對象，因此無需得出價值結論。工作流程以模型開發、驗證以及生成配套分析結果為核心。一旦確定並記錄了最優模型，流程即告結束。
- 單一不動產分析涵蓋市場分析的全部內容，並進一步確定特定不動產的價值。這與傳統估價方法有著根本區別。在傳統方法中，價值通常透過對 3-9 個可比交易案例進行調整並計算加權平均值得出。相比之下，RCA 主要透過兩個部分來確定價值：

1. MARS 模型的預測值
2. 估計的不動產特定殘差（連同相應的 CQA 評分）

在 RCA 中，銷售網格的主要作用是透過展示各變數的貢獻，來解釋估價對象與可比交易案例之間的價值差異。RCA 的一個關鍵優勢在於其穩定性：儘管客戶可能會對可比交易案例的選擇提出異議，但無論選擇哪些可比交易案例用於展示，價值結論都保持一致，因為所有成交案例（無論是 15 個還是 600 個）都是使用同一模型進行調整的。雖然較早的成交案例會被自動調整至當前市場狀況，但一般而言仍以近期成交案例為佳。隨著時間跨度的拉長，未被測量的市場特徵變得愈發重要，模型捕捉價格因素的效果也隨之減弱。

5.2 處理工作流程

整體處理工作流程可以用以下步驟來描述：

1. 工作範圍。
2. 建立初始專案資料夾與檔案。
3. 下載 MLS 資料。關於所需不動產記錄數量以及日期範圍限制的設定，請參閱規程 1.1-1.3。
4. 完成各設定工作表：迴歸變數、交互項、聚合項、計算項。
5. 執行 MARS (R/earth) 迴歸以建立價格模型。
6. 由 MARS 模型生成各不動產的價值貢獻與殘差（純殘差以及殘差/SF）。
7. 按殘差/SF 對不動產進行排序。
8. 為每個可比交易案例建立 CQA 評分。
9. 假設已經生成了一份真實反映各不動產整體吸引力的品質排序，則透過研究排序中的總體模式趨勢，為估價對象在排序中找到最合適的位置，並為其賦予 CQA 評分。
10. 將相應的殘差/SF 賦予估價對象，即殘差/SF · (估價對象) GLA。
11. 將 MARS 價格估計值與所賦殘差相加，得出估價對象的價值結論。

12. 將估價對象的殘差分解為描述性組成部分及其價值貢獻。
13. 計算可比交易案例由模型產生的價值貢獻與調整額，以及所有小計和調整後成交價格。
14. 調整後成交價格將等於估價對象的最終價值結論。
15. 選取 6–12 個最相似的不動產納入銷售網絡。
16. 將可比交易案例的殘差總額分解為殘差組成部分的價值貢獻，分解完成後將殘差總值設為 0。
17. 重新計算所有小計以及銷售網絡中可比交易案例的調整後成交價格。核對數值。
18. 將銷售網絡上傳至報告。

請注意，採用 RCA 時，最終價值結論並非直接來源於銷售網絡中的可比價值。在 RCA 方法下，銷售網絡僅僅成為一種解釋工具，用於說明為何那些看似與估價對象相似的不動產，其成交價格會高於或低於最終價值結論或估價對象的評估價值。必須換一種方式來看待這一點：估價對象的價值結論取決於 MARS 模型的價格估計值，以及估價工程師所估計的殘差。但這兩個量在很大程度上依賴於參與 R/earth 迴歸的全部不動產，該迴歸生成了不動產殘差的排序，而估價工程師透過將估價對象置於該排序中的某一位置，確定了其 CQA 評分，進而確定了其殘差。銷售網絡中的調整實際上是由 RCA 所施加的數學約束決定的；儘管殘差分解在一定程度上取決於估價工程師的判斷，但分解出的價值貢獻合計必須始終等於相應的殘差。我們從證明 2 可知，雖然殘差各組成部分的實際數值可能因估價工程師的判斷而異，但對於給定不動產，其合計值是完全受約束的，因此無論估價工程師如何改動殘差分解，都不會改變該不動產的調整後成交價格。

5.3 資料工作流程

請注意，本文僅對 RCA 方法給出高層級的工作流程。其他文章將深入探討 RCA 分析的各個階段與子階段。

5.3.1 MARS 的資料要求

我們需要指定銷售交易的起始日期，並根據估價或估值的基準日確定截止日期。這通常需要考慮每一交互度（大於 1 時）所需迴歸變數與交互項數量的估計值。

5.3.2 規程：已成交交易掛牌資訊

規程 1.1：確定所需 MLS 已成交交易掛牌資訊的數量

本規程說明如何計算輸入 MARS 迴歸所建議的已成交交易掛牌資訊的最低數量。你需要先做一個粗略估計，隨後執行一階、二階乃至三階的 MARS 以生成初始模型，然後按照下述規則審查資料列數，判斷是否需要更多記錄。本規程假設各自變數之間是 100% 相互獨立的，而這在不動產領域幾乎不可能成立；例如，GLA 與房間數量通常高度相關，GLA 也很可能與土地面積存在相關性。這就涉及到「自由度 (DOF)」的計算。每一個真正獨立的變數佔一個自由度。但是，當所謂的自變數之間存在依賴關係時，準確計算 DOF 需要進行一定的分析，這已超出本文的討論範圍。只需記住，由於變數之間很可能存在共線性，下述估計值可能偏高。

1. 對於一階項（無交互項），每個變數約 ~ 15 列的經驗規則效果良好
2. 對於每一個允許的雙向交互項，應額外增加約 15-30 處不動產
3. 對於更高階的交互項（3 個及以上變數），每項應有 30 處以上不動產。

不過，這並非硬性規則。實際所需數量可能取決於：

- 資料中的訊噪比
- 所要建模的關係的複雜程度
- 反應變數的變異程度
- 資料的品質

示例：

因此，如果你有：

- 10 個主變數
- 5 個允許的雙向交互項
- 2 個允許的三向交互項

則可按如下方式計算所需成交案例的最低數量：

$$\begin{aligned}
 N &= (10 \text{ 個變數} * 15 \text{ 列/變數}) + \\
 &\quad (5 \text{ 個雙向交互項} * 22 \text{ 列/雙向交互項}) + \\
 &\quad (2 \text{ 個三向交互項} * 30 \text{ 列/三向交互項}) \\
 &= 320 \text{ 列}
 \end{aligned}$$

或者，也可以採用以下做法：在對一組資料執行 R/earth 之後，可以用模型中基底函數的數量加一作為自由度的估計值，再將該數值乘以 20-30，即可得到所需輸入記錄的數量。

規程 1.2：銷售期間應盡量避免購屋者偏好發生變化

在收集歷史不動產資料時，如果市場偏好的變化能夠透過適當的變數加以捕捉，則對時間跨度沒有嚴格限制。在處理市場中建築風格演變時，請考慮以下要點：主流建築風格的變化（例如從加州牧場式向地中海式的轉變）可以透過以下方式建模：

- 社區（subdivision）識別碼
- MLS 區域程式碼
- 建造日期區間

然而，如果這些風格變化無法透過特定變數得到充分捕捉，你將面臨一個重要約束：銷售日期通常用於調整市場狀況，若同時用它來控制建築風格偏好，會使這兩種截然不同的效應相互混淆。在這種情況下，較為可取的做法是二者擇一：

1. 將資料收集範圍限制在較新建築風格出現之後的時期
2. 將可比交易案例的選擇限制在具有相同建築風格的社區之內

這樣既能確保分析在不動產特徵上保持一致性，又仍然允許對市場狀況隨時間的變化進行恰當調整。或者，至少在這種情形下，假設你能夠從資料中提取出建築風格，可以考慮這兩個變數之間的交互作用。然而，這種提取並非總是可行的。

規程 1.3：若處於成交稀少的區域，則盡力而為

在使用 MARS 對銷售資料建模時，進行有意義的分析大約需要 15 筆交易。如果成交案例更少（例如只有 5 處不動產），可以將每條記錄複製兩次，以達到 15 處不動產的門檻。採用這種做法時的重要注意事項：

1. 僅使用一階迴歸，以降低過擬合
2. 由於資料集過於有限，忽略或跳過交叉驗證
3. 在這種情形下，MARS 的表現很可能優於配對比較法或標準線性迴歸方法，但要注意，複製資料會影響不確定性度量與統計顯著性的有效性

最好的長期解決方案是透過補充真實成交案例來擴充資料集。不過，如果必須在資料有限的情況下繼續進行工作，資料複製在承認其統計侷限性的前提下，不失為一種可行的臨時方案。

5.4 專案資料夾結構

有必要討論一下專案資料夾結構。建議為市場分析和單一不動產估價分別設立單獨的資料夾。這兩個資料夾具有相似的子資料夾結構：

```

RCA_Projects/
├── MarketAnalysis/
│   ├── Burlingame_20240501/
│   ├── Pacifica_20240310/
│   ├── RedwoodCity_20240615/
│   └── ...
├── SingleFamily/
│   ├── Pacific_Rosewood_123_20240510/
│   │   ├── Code/
│   │   │   ├── R/
│   │   │   │   └── ...
│   │   │   └── Python/
│   │   │       └── ...
│   │   └── Data/
│   │       ├── MLS_Original.xlsx
│   │       ├── MLS/
│   │       │   ├── VER202402510153248/
│   │       │   │   ├── MLS.xlsx
│   │       │   │   ├── MLS_Stage2.xlsx
│   │       │   │   ├── MLS_Stage3.xlsx
│   │       │   │   ├── MLS_Stage4.xlsx
│   │       │   │   ├── MLS_Stage5.xlsx
│   │       │   └── Documents/
│   │       │       ├── LinearModels.pdf
│   │       │       ├── Cum_Dist.pdf
│   │       │       ├── Model.txt
│   │       │       ├── PartContrib.pdf
│   │       │       ├── Pairs.pdf
│   │       │       ├── Residual1.png
│   │       │       ├── Residual2.png
│   │       │       ├── ResidualSF.png
│   │       │       ├── Rpart.pdf
│   │       │       ├── VariablesImp.pdf
│   │       │       ├── VariableVsSP.pdf
│   │       │       └── ...
│   │       └── ...
│   └── ...
└── ...

```

第 1 階段處理將建立版本資料夾及其子資料夾。它會複製 `MLS_Original.xlsx`——儘管該試算表理應永不更改，而是精確反映在某一日期和時間從 `MLS` 下載的原始內容；但在週期較長的專案中，仍有可能需要重新從 `MLS` 下載資料。接著，此階段會將經過擴展的 `MLS.xlsx` 活頁簿連同其設定參數複製到版本資料夾中。版本資料夾可以有很多個，每個都對應一次完整的 `MARS` 執行。

隨後的各個階段將對活頁簿進行欄的新增、排序及其他更改，並在估價完成時將更新後的版本以 `MLS_Stage(N)` 的形式寫入專案的 `Data` 資料夾。再次強調，根目錄下的 `MLS.xlsx` 處於持續變動之中，因此版本副本是對原始 `MLS` 活頁簿的一種持續演化，包含變數的刪除與新增以及其他更改。處理用的 `R` 或 `Python` 程式碼很可能以 `Data` 資料夾中的 `MLS.xlsx` 活頁簿作為

輸入執行。這樣，我們始終保有每一版 RCA 執行背後全部資料的副本。

版本可以根據需要不斷建立，直至無法再做出改進為止，此時選出最佳版本作為最終版本。可以在資料夾名稱末尾加上「_Final」，以便瀏覽目錄時一目瞭然。一些包含有趣模型或競爭性模型的舊版本通常應予保留。每個版本資料夾都包含若干文字、圖示和圖形檔案，可嵌入到報告中。最終的 `MLS_Stage5.xlsx` 活頁簿包含報告所需的全部工作表。特別是，它可能包含一張用於轉入 Fannie Mae 銷售網格的試算表。它還必然包含聚合項背後的計算或分解過程。這可能是 RCA 方法中最耗時的部分——在一個困難的市場區域中苦苦尋找完美模型的過程。

重建銷售網格所需的一切都必須複製到版本資料夾中，包括所使用的 R 或 Python 程式碼——如果程式碼經常變動的話。

5.4.1 規程：多次 MARS (earth) 執行

規程 2.1：利用過擬合執行精細剖析市場區域的定價因素

在最初的 earth 執行中，應當有策略地利用過擬合來發現潛藏的價格因素。例如，一個由 800 筆成交案例構成的市場區域中，可能只有 3 處不動產帶有工作間 (workshop)，而在這些孤立案例中，工作間可能對最終成交價格產生重大影響。僅進行前向傳遞的執行會產生一個過擬合模型，從中可以看出這些工作間的近似價值貢獻。然而，由於 800 筆成交中只有 3 筆這樣的案例，工作間相關項將在後向傳遞中被剔除，在交叉驗證下更是必然如此。可以將僅前向傳遞賦予工作間的價值記錄下來，轉入潛變數分析，供銷售網格中出現的任何帶工作間的不動產使用。此類過擬合模型還能夠為許多其他資料稀薄的特徵給出詳細估計，而這些特徵本會被後向傳遞剪枝或交叉驗證執行所剔除。

將每一次這樣的探索性執行都儲存為獨立的版本資料夾：所記錄的價值貢獻是估價工程師日後分解殘差時會想要引用的證據。當然要明白，估價工程師最終必須依靠剪枝和交叉驗證，才能使最終模型對迴歸從未見過的不動產——例如後來出現的估價對象——保持穩健。過擬合執行是用於研究的工具；它們絕不是最終模型。

規程 2.2：在不同 earth() 參數下進行多次交叉驗證執行

市場區域在結構和規模上各不相同，估價工程師通常需要透過一系列執行來進行試驗：以瞭解哪些變數對價格影響最大；判斷現有的已測變數集合是否足以估計價值，抑或必須補充新資料（如距水域的距離、海拔、距購物中心的距離等）；並發現新型的交互關係，同時將其他交互關係作為低效或具有誤導性的項予以排除。最終，估價工程師希望在訓練資料上獲得良好的 R^2 ，並在測試資料上取得盡可能好的 CVR^2 。要達到這一目標，可能需要收集新資料、提高現有資料的準確性，或根據市場區域的規模與特點（例如針對可用的自由度——參見規程 1.1）調整 `earth()` 參數。

這正是上述專案結構支持多個執行版本資料夾的原因：每個候選模型都連同其參數及其 R^2/CVR^2 一並保留，從而可以比較各次執行之間的演變軌跡。若某一版本的 R^2 上升而其 CVR^2 停滯甚至下降，則說明該版本擬合的是噪聲而非市場信號，而正是版本歷史使這一點得以顯現。

規程 2.3：利用 Prolog 展開備註及描述性特徵欄

儘管這一功能仍處於一定的試驗階段，earthUI 現已內建 Prolog 引擎（vProlog），並使用定子句文法（DCG）來解析 MLS 的自由文字欄——首先是 Public Remarks（公開備註），此外還有經紀人備註以及以逗號分隔的特徵列表，例如房間特徵或既有建築物。這種展開會將「owned solar」「3 car garage」「40x50 workshop」之類的短語轉化為可供迴歸使用的普通特徵欄，而以此方式探勘出的資料稀薄特徵可直接輸入潛變數分析（參見後文關於將狀況與品質評級作為近似迴歸變數的小節；該功能本身在本期的 earthUI 配套文章中有詳細說明）。

文字展開會成倍擴大候選預測變數集合，因而必須進行多次 earth/RCA 執行加以測試：展開出的欄必須在模型中反覆取捨試驗，在存在共線性時合併為派生的計數或布林變數，並對照 CVR2 檢驗其承載的是市場信號而非噪聲。應將每一次此類試驗視為獨立的執行版本資料夾，與規程 2.1 和 2.2 的做法完全一致。此外，由於解析文法本身仍在開發之中，估價工程師在將提取出的欄用於模型之前，應先將其與源文字進行核對。

5.5 核心 Excel 活頁簿

由於大多數 MLS 服務將不動產資料下載為 Excel CVS 工作表，Excel 可能是儲存資料及各類設定參數最便捷的方式。每個專案都有一個名為 MLS.xlsx 的核心 Excel 活頁簿，其中包含一張名為「MLSDData」的工作表，用於儲存從 MLS 下載的已成交不動產銷售掛牌資訊，此外還有若干設定與參數工作表：

```
Workbook: MLS.xlsx/
├─ Sheet: Project Data
├─ Sheet: MLSDData
├─ Sheet: Regression Variables
├─ Sheet: Calculations
├─ Sheet: Allowed Interactions
├─ Sheet: Aggregations
└─ Sheet: (MARS run parameters)
```

原始資料儲存在一個名稱類似 MLS_Original.xlsx 的活頁簿中。應當以不同名稱——比如「MLS.xlsx」——製作一個副本，由 R 或 Python 的 RCA 程式載入。這兩個活頁簿都存放在 Data 資料夾下。此後，Data/MLS.xlsx 活頁簿很可能會經歷多輪迭代執行，嘗試不同的變數組合、MARS 參數以及其他設定的更改。每次迭代都會建立一個以日期和時間命名的新輸入/輸出資料夾，格式為 VER_YYYYMMDDHHMMSS。每個資料夾都會儲存輸入的 MLS.xlsx 活頁簿的一個副本，並包含四個處理階段（第 2–5 階段）各自生成的 Excel 活頁簿，以及輸出的文字文件、圖示和圖形。請注意，最後一個階段即第 5 階段，可能會根據客戶類型和工作範圍（SOW）包含若干供報告使用的其他工作表。

5.6 備份

在將新建或修改後的 Excel 試算表寫入專案版本資料夾時，還可以在專案資料夾中向 SQLite 或 PostgreSQL 資料庫進行備份。資料庫檔案在後續時期的審閱或參考使用過程中，較不容易被意外損壞。

6 活頁簿記號

6.1 描述維度

活頁簿 “M” 大約有六個描述維度：

1. 估價專案 “a”
2. 執行版本 “r”
3. 工作流階段 “w”
4. 工作表 “s”
5. 不動產掛牌資訊 “p”
6. 變數集 “v”

因此，特定活頁簿 M 中的某個儲存格可以描述為：

$$\begin{matrix} r_* \\ w_* \end{matrix} \begin{matrix} a_* \\ W \\ s_* \end{matrix} \begin{matrix} p_* \\ v_* \end{matrix} \quad (1)$$

(注：“*” 表示某個有效索引)

或：

$$\begin{matrix} r_k \\ w_m \end{matrix} \begin{matrix} a_g \\ W \\ s_h \end{matrix} \begin{matrix} p_i \\ v_j \end{matrix} \quad \begin{matrix} i = 1, \dots, n_p \\ j = 1, \dots, n_v \\ k = 1, \dots, n_r \\ m = 1, \dots, n_m \\ g = 1, \dots, n_g \\ h = 1, \dots, n_h \end{matrix} \quad (2)$$

其中：

- n_p = 列數 (取決於工作表)
- n_v = 欄數 (取決於工作表)
- n_r = 執行版本數
- n_m = 工作流階段數
- n_g = 分析專案數
- n_h = 活頁簿工作表數

請理解，“W” 指的是整個 Excel 活頁簿，它包含六個核心資料工作表，每個工作表都有自己的表格；在執行相應的 R 或 Python 程式時，這些表格會被載入到各自獨立的內部資料框中。R 和 Python 都有內部函數，可以根據名稱查找它屬於資料框的哪一行或哪一欄。例如，如果在下標 v_j 中， $v = (\text{“GLA”}, \text{“LotSize”}, \text{“BathRms”}, \text{“BedRms”}, \dots)$ ，且 $j=0$ ，則 $v_j = \text{“GLA”}$ 。R 或 Python 都能夠在內部將 “GLA” 解析為其所儲存的欄，例如，這可能是相應資料框的第 15 欄。因此，在本文中我們對名稱陣列進行索引，把名稱傳遞給執行中的程式，程式再將該名稱重新索引為內部資料框中實際的列或欄索引。這有點複雜。但為了精確起見，這樣做是必要的，也是最方便的。這是分析師角色與開發者角色之間的「職責分離」；儘管在現實中，估價工程師 (VE) 可能同時承擔這兩個角色。請注意，R 中的 “data frame” 在 Python 中稱為 “dataframe”。

另請注意，MARS 執行時參數工作表並不是 W 的一個維度，因為 W 關係到輸入 MARS 的資料版本，而與決定 MARS/earth 如何處理這些資料的參數無關。

為清晰起見，本文給出了一套複雜的多維活頁簿記號，它與資料活頁簿各工作表所載入到的內部 R 或 Python 資料框更直接地對應。這是一套「沉重」的記號。但當我們進入證明等環節

時，透過對所處理對象作出假設，就可以拋掉大部分包袱。儘管如此，這套記號有助於我們從一開始就更好地總覽 workflow 和資料。

我們先從將要處理的活頁簿集合的記號開始。

6.1.1 專案維度 (“a”)

一旦工作範圍 (SOW) 獲得批准，就需要建立一個專案資料夾來儲存輸入和輸出資料，包括設定參數、輸出圖形、圖表、試算表和報告片段。VE 應當用簡單的順序 ID (例如整數) 標示所有專案，這樣一旦某份報告丟失，就會一目瞭然。如果專案被取消，應當有一致的方式在相應專案資料夾內註明該專案已被取消，但絕不應刪除該資料夾。我們假設專案 ID 從 0 開始並逐次加 1，儘管也存在其他可能。我們進一步假設所有有效的專案 ID 都儲存在由處理程式建立的有序集合 “a” 中。

估價專案資料位於活頁簿的專案工作表中，但也可以儲存在某個中心 SQL 資料庫中，

以下是建議的專案欄位：

1. Local ID：為公司順序生成的短整數 ID。它非常適合作為儲存用的主鍵——並且可以由資料庫在每個後續專案中自動生成。
2. Global GUID：如果需要，可以使用以下網址建立一個 GUID (全域唯一 ID)

<https://www.uuidgenerator.net/>

3. 不動產位置 (市場分析) 或地址
4. 估價生效日期
5. 委託日期
6. 查勘日期
7. 原始報告日期
8. 第一次更新報告日期
9. 第二次更新報告日期
10. 第三次更新報告日期
11. 客戶名稱
12. 客戶地址
13. 貸款機構名稱
14. 貸款機構地址
15. 屋主名稱
16. 屋主地址
17.

專案 ID 列表示例：

```
a = ("1",
      "2",
      "...",
      "141",
      "142")
```

6.1.2 執行版本維度 (“r”)

工作流的第二步是對輸入執行 MARS 迴歸。這很可能需要數次乃至多次執行，因為我們會調整各種參數以改進模型。每次執行都會為輸出建立一個新的版本資料夾，其名稱基於執行的日期和時間。例如，

```
r = ("ver20241210_083124",
      "ver20241209_151130",
      "ver20241208_160944")
```

```
latest_version: r1 = "ver20241210_083124"
```

具有競爭力的模型和輸出的有希望的版本應當保留，直到確定最終版本為止；如果它們仍有某些用處，或許還可以保留更久。你可能希望修改版本資料夾名，在名稱末尾加上“Final”，以表明它是最終版本。

6.1.3 workflow階段 (“w”)

RCA workflow 通常分為五個獨立的階段，如下所列。每個階段必須成功完成後才能開始下一個階段。VE 應能夠在某個階段結束時停止處理，先審查結果再繼續。這是必要的，因為市場區域相當複雜，而且往往頗為獨特。可能會遇到需要採用新方法的異常問題。VE 必須能夠進入 R 或 Python 程式碼進行必要的修改，甚至嘗試新方法。而且在這種情況下，他可能希望進行試驗並嘗試不同的方法，這就需要對他正在處理的同一階段進行多次執行。當然，VE 也應當可以用一條命令執行所有階段——但有一些前提（例如，三階聚合必須已透過之前的 Stage 2 執行確定）。這意味著每個階段完成時，都應把其全部內部資料存入檔案或資料庫，以便下一個階段可以透過把上一階段的資料重新載入到記憶體中來啟動，就好像上一階段仍駐留在記憶體中一樣。

以下是各個階段：

```
w = ("stage1", -設定
      "stage2", -MARS 處理
      "stage3", -最終銷售網格
      "stage4", -報告片段
      "stage5") -審查/清理
```

```
stage2: w2 = "stage2"
```

6.1.4 Excel 活頁簿工作表 (“s”)

在 RCA 核心 Excel 活頁簿中，我們有以下工作表，每個工作表都映射到一個內部資料框：

```
s = ("MLSDData",
      "Regression Variables",
      "Calculations",
      "Interactions",
      "Aggregations 1st & 3rd Degree",
      "Aggregations 2nd Degree")
```

之所以把二階聚合放在單獨的工作表上，是因為它們最容易透過一個對稱表來定義：該表展示所有可能的交互對，允許為每一對指定一個聚合變數，將其相應的價值貢獻和調整聚合（相加）到該變數上。三階聚合無法這樣處理，因為那需要一個三維對稱表，而這在 Excel 中是不可能的。此外，一個 20x20x20 的表將有 8,000 個儲存格需要處理。更簡單的做法是把任何三階項的名稱直接列在包含一階聚合的表中，例如：

Variable	Aggregate To
AboveGradeFinArea	GLA
BelowGradeFinArea	BGLA
UnfinArea	UBA
LotSize	LotSize
...	...
GLA___LotSize	GLA
GLA___DateOfSale	GLA
BathRms___DateOfSale	GLA
Lat___Long	Location
Lat___Long___GLA	Location
...	...

6.1.5 不動產掛牌資訊 (“p”) ——MLSDData 列

不動產掛牌資訊構成 MLSDData 工作表的列。它們不會出現在其他工作表中，直到 Stage 4——在該階段，將使用最終選定的可比交易案例生成最終的報告片段。以下是不動產 ID 的示例列表：

$$p = (\text{“MLS823211”, “MLS823456”, “MLS823542”, “MLS823721”, “...”})$$

6.1.6 MLS 變數 (“v”) ——MLSDData 欄

關於變數的處理，後文還會詳述。此處只需指出：MLSDData 工作表通常有 30-40 個類型各異的變數，但通常只有 10-20 個會被選作 MARS 迴歸的自變數。這些變數也出現在其他工作表中，不過通常是作為可供提交給 MARS 的所有可能變數的列表。變數列表的一個示例是：

$$v = (\text{“GLA”, “SalePrice”, “BedRms”, “BathRms”, “GarageSF”, “...”})$$

6.1.7 最終記號示例

讓我們把上述記號匯總成一個最終示例。使用上面的列表，令

$$\begin{matrix} r_1 \\ w_1 \end{matrix} \overset{a_{141}}{W}_{s_0} \begin{matrix} p_2 \\ v_0 \end{matrix} \quad (3)$$

請注意，這裡我們使用 0 起始索引，即假定列表的第一個元素的索引為 0。Python 也使用 0 起始索引，而 R 使用 1 起始索引。

因此，上式表示我們所討論的是以下儲存格的值：

- 變數 v_0 ，即欄 “GLA”
- 不動產 p_2 ，即 ID “MLS823542”
- 工作表 s_0 ，即 “MLSDData”
- workflow 階段 w_1 ，即 “stage2”

- 版本 r_1 ，即
“ver20241209_151130”
- 專案 a_{141} ，即 “142”

上述記號也可以寫成：

$$\begin{array}{ccc} \text{“ver20241209_151130”} & \overset{\text{“142”}}{W} & \text{“MLS823542”} \\ \text{“stage2”} & \text{“MLSDData”} & \text{“GLA”} \end{array} \quad (4)$$

6.1.8 理由

「為什麼不在記號中直接使用索引或數字呢？」我們對記號有兩個要求：

1. 子集。例如，分析人員經常選取列表的子集，比如變數的子集。舉例來說，MLSDData 工作表中可能有 35-40 個變數，但某次執行只使用其中的 10-20 個。從一次執行到下一次執行，VE 會頻繁更改提交給 MARS/earth 的變數。
2. 迭代。例如， $i = 1, 2, \dots, n_1$
我們的演算法、規程和證明通常需要能夠從頭到尾遍歷有序列表中的所有項——或挑出某個特定的項。

因此，我們需要一個字母來表示項的子集或列表，並且它必須帶有另一個下標，供我們遍歷列表中的各項，以給出用於索引資料框列或欄的字串。

6.1.9 矩陣？

那麼線性代數中的矩陣呢？矩陣只能處理數字，即整數或實數。R 和 Python 都支持矩陣這種資料結構。然而，我們在 MLS 資料中處理的許多變數是所謂的「字串」或字元資料。即使有些變數只是數字，它們往往也不是實數，而只是名稱。例如，某城市的 MLS 區域程式碼可能是（“660”、“661”、“662”、“663”、...、“672”）。Earth 迴歸允許把變數指定為「因子」變數，此時這些數字僅被當作名稱處理，不含任何順序。在這種情況下，生成的模型只會為每個數字提供一個價值貢獻，而不是一個在序列上取值的函數。其他例子包括表示狀況的 “C1”、...、“C6” 和表示品質的 “Q1”、...、“Q6”。

6.1.10 第一次簡化

上述針對核心 Excel 專案活頁簿的記號很適合把各部分聯繫起來，但用於解釋演算法或數學證明就顯得太沉重了。在做繁瑣的具體工作時，我們可以假設一切都發生在給定的專案、版本、階段和工作表之內。如果所處理的對象尚不明確，我們直接說明即可。一旦指明了所處理的對象，我們真正需要的只是：

$$M_{s_h}^{p_i v_j} \quad \begin{array}{l} i = 1, 2, \dots, n_p \\ j = 1, 2, \dots, n_v \\ h = 1, 2, \dots, n_s \end{array} \quad (5)$$

其中：

n_p	=	不動產數量
n_v	=	輸入變數數量
n_s	=	輸入工作表數量

6.1.11 第二次簡化

更方便的做法是按工作表所承載的資料框的具體類型來重新命名它們，而不是使用索引。對每個工作表而言，列和欄具有不同的功能。因此，W 這一資料框集合被拆分為 M、V、A、I 和 C 資料框，如下所示：

◇ **MLS 資料**

$$M_{v_j}^{p_i} = W_0^{p_i} \quad \begin{matrix} i = 0, 1, \dots, n_p \\ j = 0, 1, \dots, n_v \end{matrix} \quad (6)$$

其中：
 n_p = 不動產數量
 n_v = 輸入變數數量

注意：估價對象預設始終位於不動產列表首位

在“MLSDData” Excel 試算表及相應的資料框中，第一列始終保留給估價對象。它絕不會與其下方的其他 MLS 不動產一起參與排序。它也從不作為自變數資料輸入 MARS 迴歸，而是作為目標資料，即依變數輸入。不動產列表同樣如此。在所有不動產列表中，除非另有說明，索引為 0 的第一個元素始終被假定為估價對象的 ID。

即使任務是為某個沒有估價對象的市場區域建立模型，第一列也仍會載入一些虛擬的不動產資料。這使得撰寫程式、給出演算法或進行證明都方便得多。因此，如果我們用 c 表示不動產列表，則 c_0 是估價對象，而 c_i ($i = 1, \dots, n_p$) 是將輸入 MARS/earth 迴歸的可比掛牌資訊。

◇ **迴歸變數規格說明**

$$V_{s_j}^{v_i} = W_1^{v_i} \quad \begin{matrix} i = 1, 2, \dots, n_v \\ j = 1, 2, \dots, n_s \end{matrix} \quad (7)$$

其中：
 n_v = 可用變數數量
 n_s = 規格說明數量

注意， $n_s \leq n_v$ ，因為我們通常從可用變數中選取一個子集作為迴歸的輸入。

◇ **計算**

$$C_{c_j}^{v_i} = W_2^{v_i} \quad \begin{matrix} i = 1, 2, \dots, n_v \\ j = 1, 2, \dots, n_c \end{matrix} \quad (8)$$

其中：
 n_v = 變數數量
 n_c = 計算規格數量

◇ **交互**

$$I_{t_j}^{f_i} = W_3^{f_i} \quad \begin{matrix} i = 1, 2, \dots, n_f \\ j = 1, 2, \dots, n_t \end{matrix} \quad (9)$$

其中：
 n_f = “from” 變數數量
 n_t = “to” 變數數量

請注意，資料框 I 是一個對稱資料框，變數按字母順序排列於“from”列和“to”欄。儲存格標記為“1”表示允許相應變數對之間的交互；否則，“0”表示不允許交互。對於二階交互，這相對簡單，只可能存在一種交互。對於三階迴歸，三個變數之間共有三個變數對交互，必須允許其中每一對交互，才能生成三變數交互項（或變數）。

◇ **聚合**

$$A_{v_j}^{v_i} = W_4^{v_i} \quad \begin{matrix} i = 1, 2, \dots, n_v \\ j = 1, 2, \dots, n_v \end{matrix} \quad (10)$$

其中：
 n_v = 變數數量

7 RCA 銷售網格

請記住，RCA 規程需要一個能夠呼叫 R/earth 並執行所有細節步驟的 R 或 Python 程式。這項工作資料密集，絕不可能以手工方式完成。同時，假設程式沒有缺陷，您也需要一個軟件程式來確保所需工作精確無誤地完成，避免人為差錯。

對於從事這項工作的估價工程師來說，能夠根據需要修改此類程式也是一種必要能力，尤其是在面對新的、未曾預見的市場區域、社區、不動產或其他新需求、新問題時。

在 RCA 處理過程中，我們有三種類型的銷售網格：

1. RCA 處理銷售網格，在本文中標記為「MLSDData」，它是一個非常龐大的資料框及其關聯的試算表，可能包含許多中間細節，這些細節最終經過聚合處理，得到下文提到的 RCA 銷售網格。示例太大，無法在此展示。估價工程師很可能會按照各自的偏好來客製化這一試算表。我打算在將來另撰一篇論文，專門討論此類網格的開發。
2. RCA 最終銷售網格，用於非 GSE 報告，其形式類似於 [表 2](#)。它確實在一定程度上模仿了例如 Fannie Mae 的表格，足以讓讀者有一定的熟悉感。不過請注意，它為每個可比交易案例提供的欄數多於 Fannie Mae 網格：每個可比交易案例除「測量值」和「調整」欄之外，還有一個「價值貢獻」欄，估價對象也同樣有一個——這些欄是 Fannie Mae 表格所沒有的。一個更簡單的形式見 [表 1](#)。
3. GSE 上傳銷售網格，用於將銷售網格上傳到諸如 Alamode 之類的軟件中。但請注意，並非所有估價表格軟件包都提供從試算表上傳銷售網格資料的功能。出於必要，此類網格不能包含單獨的價值貢獻欄，而 Fannie Mae 等 GSE 對價值貢獻毫無概念。我們的策略是：向它們提供其所要求的表格，但以 RCA 銷售網格作為支撐，並輔以其他支持性文件，例如各變數聚合的詳細資訊。

RCA 處理銷售網格看起來不像 Fannie Mae 銷售網格，而大致相當於 Fannie Mae 銷售網格的轉置，列與欄互換：每一列包含某一不動產的資料，每一欄包含某一不動產變數的資料。不動產列可以輕易達到數百或數千列，欄數可達 200 欄或更多。第一列保留給變數名，第二列保留給估價對象，即使我們只是在做市場分析也是如此。當試算表被讀入資料框後，工作表的第一列成為 names 屬性，而估價對象所在的第二列則成為資料框的第一列。

RCA 銷售網格中的欄包括識別欄，然後是主要變數、外部變數和合成變數，它們合併後作為自變數輸入 MARS。MARS 從這些變數中輸出它認為對成交價格（也稱依變數或目標變數）有顯著影響的變數。在二階迴歸中，我們可能得到成對的變數，它們作為新的變數發揮作用。在三階迴歸中，我們還可能得到變數三元組，它們作為所謂的自變數發揮作用，儘管它們當然與底層變數高度相關。

表 1: RCA 試算表示例：縱向布局

估價對象				可比交易案例 1			可比交易案例 2		
	測量值	價值貢獻		測量值	價值貢獻	調整	測量值	價值貢獻	調整
成交價格				\$1,150,000			\$1,500,000		
基礎價值		\$500,000			\$500,000	\$0.00		\$500,000	\$0
		已測量特徵			已測量特徵			已測量特徵	
居住面積	1,530	\$463,500		1,350	\$382,500	\$81,000	1,700	\$540,000	-\$76,500
土地面積	6,000	\$200,000		6,500	\$220,000	-\$20,000	7,000	\$240,000	-\$40,000
房齡	53	\$63,600		57	\$68,400	-\$4,800	51	\$61,200	\$2,400
浴室數	2	\$15,000		3	\$30,000	-\$15,000	3	\$30,000	-\$15,000
臥室數	3	\$6,000		4	\$9,000	-\$3,000	4	\$9,000	-\$3,000
測量合計		\$1,248,100.00			\$1,209,900.00	\$38,200		\$1,380,200.00	-\$132,100
		殘差特徵			殘差特徵			殘差特徵	
殘差	CQA 5.8	\$30,000		CQA 2.5	-\$59,900	\$89,900	CQA 8.2	\$119,800	-\$89,800
設計	殘差分解			殘差分解			殘差分解		
狀況		7,000			-\$15,000	\$22,000		\$35,000	-\$28,000
品質		5,000			-\$8,000	\$13,000		\$20,000	-\$15,000
功能效用		8,000			-\$10,000	\$18,000		\$30,000	-\$22,000
配套設施		5,000			-\$12,000	\$17,000		\$15,000	-\$10,000
小計		5,000			-\$14,900	\$19,900		\$19,800	-\$14,800
		30,000			-\$9,900	\$89,900		119,800	-\$89,800
合計				合計			合計		
價值貢獻總計		\$1,278,100.00			\$1,150,000.00			\$1,500,000.00	
調整後成交價格						\$1,278,100.00			\$1,278,100.00

表 2: 死亡日期估價實際使用的 RCA 銷售網絡副本 (已刪除不動產識別資訊)

[illegible]

8 變數

在執行 MARS 的輸入端，我們只有簡單變數，它們是：

1. 從 MLS 試算表載入的不動產特徵。
2. 由 VE 建立的派生變數，用於細化 MLS 變數或新增現有變數的變換。例如，VE 可能傾向於將銷售日期變換為「DaysOffMarket」，其值等於估價生效日期減去銷售日期。
3. VE 可以從第三方來源匯入資料，例如房貸利率。或者，如果 MLS 資料中缺少緯度、經度或海拔資料，他可以從其他來源匯入。這些變數可稱為「外部變數」。

在 MARS 的輸出端，我們處理的價格模型是若干項的組合，這些項：

1. 僅包含一個輸入變數，並給出該變數價值貢獻的方程。這些稱為一階項或（輸出）變數。該項本身被視為一個價值貢獻值並被賦予名稱。
2. 包含兩個輸入變數，是給定兩個變數的交互作用之價值貢獻的方程。這些項可稱為二階項或（輸出）變數。
3. 包含三個輸入變數，是三個變數之間交互作用之價值貢獻的方程。在本文中，這些項可稱為三階項或（輸出）變數。

於是，MARS 的輸出變數即為價值貢獻，包括殘差價值貢獻。它們又在銷售網格計算中被作為變數使用，例如計算調整或繪製圖形。

我們可以有落入兩種基本本體的不同類型變數：（1）變數來源；（2）變數應用，即非殘差與殘差

8.1 變數來源本體

1. 一階變數
 - 1.1 MLS 迴歸變數
 - 1.2 外部迴歸變數
 - 1.3 派生迴歸變數
2. 因子變數：

為了處理因子變數，需要一個能夠識別它們的函數。由於它們是在 Excel 專案設定活頁簿中指定的，因此建立了一個函數：當傳入的變數已被歸類為因子類型時，該函數返回 true。
3. 交互變數：如果階數設定為 2 或更高，這些變數由 MARS 在生成模型時建立。然而，由於解釋上的困難、過擬合的可能性、對自由度的影響以及為確保模型穩健而隨之增加的不動產記錄數量，最多只建議 3 個變數交互。
 - 3.1 二階變數：這些很常見，好的例子有 GLA:BathRms、GLA:LotSize、DateOfSale:BathRms、Latitude _Longitude。
 - 3.2 三階變數：較不常見的交互，提示需要謹慎並可能存在過擬合。
4. 殘差分解變數

8.2 變數應用本體

1. MARS 生成的價值貢獻

1.1 一階變數

1.1.1. MLS 迴歸變數

1.1.2. 外部迴歸變數

1.1.3. 派生迴歸變數

1.2 交互變數

1.2.1. 二階變數

1.2.2. 三階變數

1.3 殘差分解變數

1.4 匯總變數

2. 工程師生成的變數 & 數值

2.1 殘差分解變數

2.1.1. 設計/美學

2.1.2. 景觀

2.1.3. 功能效用

2.1.4. 視野

2.1.5. 狀況

2.1.6. 品質

2.1.7. ...

這些變數組將在下一節中給出其記號並作更詳細的描述

規程 3.1：因子變數

因子變數，或簡稱因子，可出現在所有項中，包括交互項。它們之所以重要，是因為在 MARS 迴歸和繪圖中必須對其進行特殊處理。

雖然取字元值的變數必然是因子，但數值變數必須被標記為因子變數，MARS 才能將其識別為因子。例如，MLS 區域常以數字標示。特別是，舊金山灣區的鄰里區域常以唯一的整數標示。如果這些值作為因子傳給 MARS，那麼 MARS 將嘗試發現某些區域對其邊界內的不動產交易是否具有平均貢獻價值，而不考慮其與其他區域的任何關係。相反，對於普通變數，它會假設隨著變數值的變化，價值貢獻存在某種連續性。這是一個重要的區別。此外，這些因子變數可能與非因子變數一起出現在模型項中，這使繪圖變得困難。具體而言，在繪製含因子的單變數項時，我們將所有此類因子歸並在一起，簡單地繪製一張顯示其正負貢獻的條形圖。

如果因子與其他變數一起出現在某個項中，它們總是以一個可能的取值出現，因為因子變數對給定不動產只能有一個值：如果某區域的值為 662，根據定義，該區域不可能有任何其他值。如果你覺得這令人困惑，那麼更糟的是，在 MARS 生成的模型中，因子並沒有單一的變數名；MARS 會取輸入時給它的因子變數的詞根名，然後透過附加各可能取值來構造新的變數名並建立新欄。於是，Excel 試算表中的單個欄（例如 MlsArea）會為每個可能的取值建立多個資料框欄，例如 MlsArea660、MlsArea661 等。諸如 MlsArea661 這樣的變數，對位於 661 區域的不動產取值為 1，否則為 0。然後，MARS 將把這些欄當作布林變數進行迴歸。MARS 提供的模型中將使用這些修改後的名稱。

你很可能希望將這些模型名稱轉換回其原始名稱，使圖形對喜歡簡單明瞭（至少盡可能簡單）的客戶更易理解。我們看到的將不是 MlsArea661=1，而是 MlsArea=661。然後，我們只需在圖形標題中註明諸如「對於 MLS 區域 = 662」之類的文字，即可繪製各項。這樣，含 M 個因子變數的 N 階項將作為 (N-M) 階項繪製，因子維度則消失於標題之中。

注意，單個項中可能存在兩個或更多因子變數，因此你很可能需要為每一對可能的取值組合繪製一張圖。如果具有三個不同取值的變數 x 與具有兩個不同取值的變數 y 同非因子變數 z 一起出現在同一項中，那麼你將需要 $2 \cdot 3 = 6$ 張圖。一個例子是「MlsArea」+「Architectural Style」，例如舊金山的維多利亞式住宅，其價值可能因其所處的舊金山區域不同而或高或低。幸運的是，在大多數區域，通常只有 1-3 種建築風格作為價值貢獻者重要到足以用於模型（儘管可能存在 6 種以上不同的建築風格）。

8.3 二階與三階項

在二階或三階迴歸中，輸出模型中的某些項可能分別包含兩個或三個變數。例如：

$$\text{BathRms_GLA} = -52 * \max(0, 1200 - \text{GLA}) * \max(0, \text{BathRms} - 1)$$

注意，我們透過按字母順序對變數名排序並將其連接（或許以下劃線相連）來命名這些項。對於模型中產生的變數名三元組，我們對三階項也如此處理。例如「BathRms_GLA_LotSz」。

在第 3 階段，當 RCA 程式為每個不動產計算所有變數（包括交互變數）的貢獻時，這些變數會被加入到主要輸入變數之中。

8.3.1 調整變數

RCA 程式還將在第 3 階段，即價值貢獻計算完成之後，透過用估價對象某一特徵的貢獻減去相應可比不動產該特徵的價值貢獻，來計算所有變數和所有不動產的調整。

8.3.2 匯總變數

匯總變數被定義用來取代變數的某些子集，將它們的所有價值貢獻和調整加總到一個公共變數中，以便在報告中使用。例如，我們可能已將已完工和未完工居住面積都拆分為更細的變數：

1. 地上合法已完工面積
2. 地下合法已完工面積
3. 地上非法已完工面積
4. 地下非法已完工面積
5. 地上未完工面積
6. 地下未完工面積

是的，確實存在這樣做有意義的地區。

MARS 迴歸將對所有這些變數進行迴歸，並生成為它們提供貢獻和調整的模型。但對於表格式報告而言，這些不同的貢獻可能必須合併或匯總為更少的變數。匯總的細節應在附錄中提供。一個更常見的例子是將交互調整匯總到某個主要變數的調整中。例如，我們可以將 GLA_Lotsize 的調整匯總到 GLA 或 LotSize，甚至按 50:50 在兩者之間拆分，具體由匯總工作表決定。

8.3.3 特殊變數 & 記號

π_i	不動產 $i = 1, 2, \dots, n_p$ 的實際淨售價；對於估價對象， π_0 表示價值結論（見證明 I）
α_i	不動產 $i = 1, 2, \dots, n_p$ 的調整後淨售價
ρ_i	不動產 $i = 0, 1, \dots, n_p$ 的 MARS 估計淨售價
β	模型的基準值
$\varphi_{i,j}$	不動產 $i = 0, 1, \dots, n_p$ 的迴歸變數 $j = 1, 2, \dots, n_r$ 的價值貢獻
$\delta_{i,j}$	不動產 $i = 1, 2, \dots, n_p$ 的迴歸變數 $j = 1, 2, \dots, n_r$ 的調整
ϵ_i	不動產 $i = 0, 1, \dots, n_p$ 的總殘差
$\epsilon_{i,c}$	不動產 $i = 0, 1, \dots, n_p$ 的殘差分量價值貢獻 c ， $c = 1, 2, \dots, n_c$
$\xi_{i,c}$	不動產 $i = 0, 1, \dots, n_p$ 的殘差分量 c 的調整， $c = 1, 2, \dots, n_c$

其中

n_p = 可比不動產的數量

n_r = 迴歸變數的數量

n_c = 殘差分解變數的數量

8.3.4 分解殘差

一旦 MARS 迴歸建立了模型，它便會根據 MARS 選定的輸入變數生成估計售價。它用相應的淨售價減去該值，得到每個不動產的殘差值。此時，RCA 規程規定：一旦 VE 確定了最終的可比不動產——通常為 6-12 條不動產掛牌資訊——他需要逐一查看每條不動產掛牌資訊，並在忽略隨機資料誤差可能性的前提下，將計算出的殘差分解為他認為最可能影響價值的特徵（潛變數）。例如，對於某一不動產，他可能有以下分解變數：

1. 設計
2. 功能效用
3. 高於/低於市場的成交
4. 狀況
5. 品質

不同不動產很可能會有一些不同的殘差分解變數，需要將它們收集到一個有序集合中。

估計稀有特徵的貢獻。有些承載價值的特徵對迴歸通道而言過於稀有：在 780 筆交易中僅出現 3 次的工作間無法獲得站得住腳的系數，前向過程通常會完全跳過這樣的變數。這些就是殘差特徵，其貢獻可以在殘差空間中估計：先擬合不含該特徵的模型，然後用普通最小二乘法將模型殘差對被排除的欄進行迴歸。對於二元標誌，系數即為一次性總額貢獻（等價地，即被標記交易的平均殘差偏移）；對於數值欄，則是單位貢獻。當若干相關設施被聯合估計時（例如附屬建築面積與工作間標誌一起估計），迴歸會在它們之間分配貢獻，而不是在兩列分解中重複計算同一筆金額。由不足約十個觀測支持的估計僅具指示性，應透過折舊成本或擴大後的樣本（例如從公開備註中提取的標誌）予以佐證；殘差約束本身會對最終賦予的價值設定上限。earthUI (v0.11+) 以 *Residual Contribution* 索引標籤的形式直接實現了這一方法，包括聯合估計、逐筆交易的殘差證據以及用於工作檔案的記錄說明。

記號：我們使用上表中的記號，得到：

$$\epsilon_i = \epsilon_{i,1} + \epsilon_{i,2} + \dots + \epsilon_{i,n_c} \quad (11)$$

其中 n_c 是殘差分解分量的數量。

VE 必須運用其判斷將殘差分解為給定的分量。有人可能認為這會讓偏見進入最終的價值結論。然而事實並非如此，因為我們有數學證明：只要各分解調整的總和等於殘差，它們就不會影響價值結論。分解的目的不是確立價值，而是解釋殘差相對於其他不動產交易的殘差為何是這個數值。工程師的偏見唯一可能影響最終價值結論的地方，是對估價對象殘差的估計，因為估價對象當然沒有實際售價可用來為其計算殘差。

每個不動產都有其殘差價值貢獻的集合，但這些集合需要合併為覆蓋所有不動產的一個集合 C：

$$\begin{aligned} \text{設：} A_i &= \{\{\epsilon_{i,c}\}_{c=1}^{n_c}\} \\ C &= \{x \mid x \in \bigcup_{i=0}^{n_p} A_i\} \end{aligned}$$

其中 $[\epsilon_{i,c}]$ 是值 $\epsilon_{i,c}$ 的變數名， A_i 是此類變數的集合， n_c 是每個 ϵ_i 中值的數量。

因此，當估價對象與可比不動產的資料分列於不同欄中時，在每個不動產的所有迴歸模型變數價值貢獻之下，我們應看到由 n_c 個相應殘差價值貢獻構成的集合 C 。

9 RCA 核心證明

設：

$$\begin{aligned} n_p &= \# \text{ of properties} \\ n_r &= \# \text{ of regression variables} \\ n_c &= \# \text{ of residual component variables} \end{aligned}$$

證明 I：每一個可比交易案例的調整後售價均等於估價對象的估計售價。

對於任意可比交易案例 i ，調整後售價 α_i 等於淨售價 π_i 加上各項調整：即估價對象與可比交易案例在基礎價值、迴歸變數 $j = 1, 2, \dots, n_r$ 以及殘差上的價值貢獻之差。對於沒有實際成交的估價對象，其淨售價 π_0 定義為價值結論 $\beta + \varphi_{0,1} + \varphi_{0,2} + \dots + \varphi_{0,n_r} + \epsilon_0$ 。於是：

$$\begin{aligned} \alpha_i &= \pi_i + (\beta - \beta) + (\varphi_{0,1} - \varphi_{i,1}) \\ &\quad + (\varphi_{0,2} - \varphi_{i,2}) + \dots + \\ &\quad (\varphi_{0,n_r} - \varphi_{i,n_r}) \\ &\quad + (\epsilon_0 - \epsilon_i) \\ &= \pi_i + (\beta + \varphi_{0,1} + \varphi_{0,2} + \dots \\ &\quad + \varphi_{0,n_r} + \epsilon_0) \\ &\quad - (\beta + \varphi_{i,1} + \varphi_{i,2} + \dots + \\ &\quad \varphi_{i,n_r} + \epsilon_i) \\ &= \pi_i + \pi_0 - \pi_i \\ &= \pi_0 \end{aligned} \tag{12}$$

證明 II：在所有殘差分量價值貢獻之和等於可比交易案例殘差這一約束下，改變殘差的分解方式對 RCA 規程的最終價值結論沒有任何影響。

這幾乎是顯而易見的。然而若不認真思考其中的數學，就很容易忽視這一點，因為傳統估價師關注的是調整而非價值貢獻。在這裡相互抵消的正是價值貢獻——這是一個持續了數十年的疏漏。對於某個可比交易案例 k ($k > 0$)，總殘差調整等於 n_c 個分解殘差分量之和，而後者又等於各價值貢獻之差的和，即等於估價對象總殘差減去可比交易案例總殘差。這與所建立的殘差分量的數量或其所代表的內容無關。只要殘差分量的價值之和對估價對象和所選可比交易案例都等於總殘差， ξ_k 就不會改變。

$$\begin{aligned} \xi_k &= \xi_{k,1} + \xi_{k,2} + \dots + \xi_{k,n_c} \\ &= (\epsilon_{0,1} - \epsilon_{k,1}) + (\epsilon_{0,2} - \epsilon_{k,2}) + \dots + \\ &\quad (\epsilon_{0,n_c} - \epsilon_{k,n_c}) \\ &= (\epsilon_{0,1} + \epsilon_{0,2} + \dots + \epsilon_{0,n_c}) - \\ &\quad (\epsilon_{k,1} + \epsilon_{k,2} + \dots + \epsilon_{k,n_c}) \\ &= \epsilon_0 - \epsilon_k \end{aligned} \tag{13}$$

10 MARS 模型

10.1 模型變數與交互作用的「 g 」函數

RCA 方法的一項核心原則是「解釋性推理」至關重要。遺憾的是，為各個價值貢獻方程提供有意義的圖形——尤其是在使用交互作用和因子變數時——可能頗具挑戰。因此，為繪製這些圖形制定一套清晰而嚴格的規程被視為勢在必行。本文在此給出這一規程，但請理解，該規程面向的是編寫 R 或 Python 程式碼以生成圖形的開發者，對於不參與編碼的讀者而言，可能僅有粗淺的參考意義。

MARS 生成的價格模型是一個基礎常數加上若干附加項之和，每一項涉及 1 到 3 個變數。為了對其進行分析，我們首先將模型拆分為單個的項。然後，將共享同一變數集合、代表單一交互作用的多個項合併為單個交互作用表達式。假設迴歸的最大階數為 3。考慮一個三階迴歸模型，它由一個基礎常數、 n_1 個一階表達式、 n_2 個二階交互作用表達式和 n_3 個三階交互作用表達式組成。這樣我們便得到四組不同的表達式，它們或為單變數表達式，或為多變數交互作用，其繪圖方式取決於兩個因素：（1）它們是否包含因子變數（即分類變數），以及（2）它們所含唯一變數的數量。具體而言：

- 一階表達式（1 個變數）以二維圖形繪製。
- 二階表達式（2 個變數）以三維圖形繪製。
- 三階表達式（3 個變數）以一組三維圖形繪製。
- 若表達式包含因子變數，則每個唯一因子會使圖形的維度降低一維（例如，若其中一個變數是因子，則三維圖形變為二維圖形）。

每個函數 g 的索引提供了確定其繪圖方式所需的資訊，反映了唯一變數的數量以及是否存在因子變數。

10.2 g 函數的索引

每個函數 g 的索引形式為 $f_{j,k}g_k^j$ ，其中：

- $f_{j,k}$ 是第一個索引（置於 g 之前的上標）：
 - j 的取值範圍為 0 到 3，用於標示組別：
 - * $j = 0$ ：基礎常數（是通常索引規則的一個例外）。
 - * $j = 1$ ：一階表達式。
 - * $j = 2$ ：二階表達式。
 - * $j = 3$ ：三階表達式。
 - k 是某一具體函數在其所屬組內的索引（例如， $k = 1$ 表示組 j 中的第一個函數）。
 - $f_{j,k}$ 的值表示該函數中唯一因子變數的數量：
 - * $f_{j,k} = 0$ ：不含因子變數。
 - * $f_{j,k} > 0$ ：因子變數的數量（例如 1、2 或 3）。
- g 的後置索引：
 - g 之後的上標 j （例如 g^j ）重複標註組號，以求清晰。
 - g 之後的下標 k （例如 g_k ）與 $f_{j,k}$ 中的 k 一致，標示該函數在其組內的位置。

雖然 f 與 g 的索引在數值上一致，但它們的作用不同： $f_{j,k}$ 計數因子變數，這對繪圖決策至關重要；而 g 上的 j 和 k 則定義其組別和位置。

10.3 繪圖維度

函數 $f_{j,k} g_k^j$ 的圖形維度 d 按下式計算：

$$d = j - f_{j,k} \quad (14)$$

其中：

- j 是唯一變數的數量（即組號，不考慮常數的情形）。
- $f_{j,k}$ 是因子變數的數量。
- d 是需要繪圖的連續（非因子）變數的數量，最大為 3（因為最大階數為 3）。含有 d 個連續變數的函數需在 $d+1$ 維中繪製： $d=1$ 對應二維圖形， $d=2$ 對應三維圖形（或二維熱力圖）， $d=3$ 則對應一組三維圖形。

例如：

- 若 $j=3$ （3 個變數）且 $f_{j,k}=0$ （無因子變數），則 $d=3-0=3$ ，需要一組三維圖形，每張圖固定其中一個變數。
- 若 $j=3$ 且 $f_{j,k}=1$ （一個因子變數），則 $d=3-1=2$ ，需要針對模型中出現的每個因子變數取值繪製一張三維圖形（或二維熱力圖）。
- 若 $j=0$ （常數），通常無需繪圖，因為它只是一個單一數值。

當 $d=3$ 時（例如不含因子變數的三階表達式），要將完整函數可視化，可能需要多張三維圖形，每張圖將某一變數固定在不同取值上，以展示函數行為的一個子集。

10.4 最終模型方程

MARS 模型被拆分為四組表達式（常數、一階、二階和三階）。索引 $f_{j,k} g_k^j$ 告訴我們：

- j ：組別及變數數量。
- $f_{j,k}$ ：因子變數的數量。
- $d = j - f_{j,k}$ ：圖形維度。

這一結構指導每個 g 函數的可視化方式，確保繪圖方法與表達式的複雜度和變數類型相匹配。

MARS 模型給出的售價估計值 $\hat{P}$ 可以表示為：

$$\begin{aligned} \hat{P} = & f_{0,1} g_1^0 + \sum_{q=1}^{n_1} f_{1,q} g_q^1(x_*) \\ & + \sum_{r=1}^{n_2} f_{2,r} g_r^2(x_*, y_*) \\ & + \sum_{s=1}^{n_3} f_{3,s} g_s^3(x_*, y_*, z_*) \end{aligned} \quad (15)$$

$n_1 = \#$ of first degree terms

$n_2 = \#$ of second degree terms

$n_3 = \#$ of third degree terms

10.5 繪製一個模型所需的圖形數量

要將模型中的所有函數全部可視化，需要多少張圖？答案取決於變數本身、它們之間的交互作用，以及它們是因子變數還是非因子變數。

估價工程師可以在供 R/earth 使用的 Excel 設定檔案中限制變數和交互作用，從而影響哪些交互作用被允許。一種務實的做法是讓 R/earth 評估所有可能的交互作用，並選出對售價估計有顯著貢獻的那些，同時檢查過擬合或共線性等問題。某些交互作用可能被禁用，這會影響更高階的交互作用（例如，若三階交互作用中的某一變數對被排除，則整個組合都被排除）。

對於因子變數（例如具有 12 個取值的 MLS Area 與 GLA 交互），繪圖會變得更為複雜。這可能需要 12 張圖——每個區域一張。若另一個因子（如具有 6 個取值的 Architectural Style）再與 MLS Area 和 GLA 交互，則可能需要多達 $12 \cdot 6 = 60$ 張圖。不過，R/earth 很少會認定所有組合都顯著；通常只有少數關鍵交互作用（例如特定的區域和風格）值得繪圖，而模型會明確指出這些組合。

以下是所需圖形數量的簡明指南：

1. 單變數函數（非因子）：1 張圖。
2. 單變數函數（因子）：1 張將所有因子取值合併展示的條形圖。
3. 基礎函數：無需繪圖。
4. 二階交互作用（無因子）：1 張圖（例如以顏色表示售價的二維熱力圖）。
5. 二階交互作用（1 個因子）：模型中每個因子取值 1 張圖（例如 2 個取值 = 2 張圖）。
6. 二階交互作用（2 個因子）：1 張以因子組合構成矩陣的圖（例如 3 個區域 $\times$ 2 種風格 = 6 個儲存格）。
7. 三階交互作用（無因子）：多張圖（例如按 Bath 數量 = 1、2、3、4 繪製 4 張圖）。
8. 三階交互作用（1 個因子）：模型中每個因子取值 1 張圖。
9. 三階交互作用（2 個因子）：模型中每種因子取值組合 1 張圖。
10. 三階交互作用（3 個因子）：以不同取值最少的那個因子的每個取值各繪 1 張圖，其餘兩個因子用顏色矩陣表示。

更複雜的情形可能需要更高級的繪圖技術。

此外還存在其他可能性，通常需要更為精細的圖形。

由 MARS 生成的完整模型示例見 [圖 1](#)。

11 資料特徵

以下一節討論 MARS。MARS 完全依賴於它所接收的資料，而這些資料大多來自某個 MLS（多重掛牌服務）或稅務評估機構。不過，資料也可能經由其他資料提供商中轉，這些提供商維護並改進資料，然後以一定利潤出售。在某些地區，資料相當可靠；在另一些地區則不那麼可靠。在關鍵情形下，不動產查勘人員必須親自進行觀察和測量，而其準確性也可能受到質疑。

估價工程師需要理解資料代表什麼，以及資料在多大程度上如實地代表了它本應代表的內容。這引出了一系列問題：

- 如何將一個變數歸類為測量值？
- 如果一個變數不是測量值，它還能是什麼？
- 如果迴歸所接收的資料不準確，迴歸又如何能有用？

我們從 MLS 掛牌服務和稅務評估機構獲得的用於迴歸的資料，大多是隱含測量值或計數的數值型資料。其中一部分是邏輯型資料，即真或假。

11.1 殘差的用途

殘差是對未納入迴歸模型的變數所增添價值的一種間接量度。用特徵價格文獻的語言來說，這些未測量的特徵就是潛變數：它們影響成交價格但未被直接測量，多數情況下是因為沒有切實可行的方法對其進行客觀測量。它們的取值只能透過推斷得到，而在實務中，潛變數的取值往往透過主觀判斷來估計——在 RCA 中，這種判斷以殘差排序為錨。透過按殘差或每平方英尺殘差對不動產進行排序，我們得到了一個從吸引力最低到吸引力最高的客觀排名。只要由一位對當地市場有深入理解的熟練分析師構建多元自適應迴歸樣條（MARS）模型，這種排序方法就是有效的。以下是一些附加要求：

- 資料與模型的品質：構建良好的迴歸模型和高品質的資料是獲得可靠殘差的前提。在舊金山灣區，我的目標是建立一個 R^2 約為 75-80%、交叉驗證 R^2 (CV R^2) 為 55-70% 的模型。
- 避免過擬合、限制 R^2 ：在像舊金山灣區這樣房價較高地區的多數成熟市場中，估價工程師應以 R^2 最高約 80% 為目標，因為住宅不動產競爭性價值的很大一部分在於其主觀吸引力，這種吸引力無法量化，因而不能直接納入迴歸模型。 R^2 顯著高於 80% 可能意味著模型過擬合，需要進一步審查。但需要提醒的是：有些市場區域的 R^2 值高於 90% 而並未過擬合；也有一些難以估價的地區，能達到的最好水準可能只有 50-60% 甚至更低。
- 特定地區的高 R^2 ：不過，在某些區域，例如分塊開發的社區或其他同質化的住宅區，由於不動產的同質性，較高的 R^2 值出人意料地常見。

11.2 作為近似迴歸變數的狀況與品質評級

房利美的狀況與品質評級量表（C1-C6 和 Q1-Q6）值得特別說明。這些量表過於粗糙，不能被當作真正的測量特徵；而且由於它們由估價師給出，實務中屬於主觀評級——本身就是潛變數。明確區分兩個層次是有益的：估價師用作迴歸變數的粗糙評級，即粗粒度狀況（condition-1）和粗粒度品質（quality-1）；相對地，從殘差中恢復的細粒度潛變數，即細粒度狀況（condition-2）和細粒度品質（quality-2）。VE 可以將估價師的粗粒度評級作為近似值接受並對其進行迴歸，從而得到近似的價值貢獻，這些價值貢獻可作為細粒度潛變數的初始賦值。在採用這種做法時，規程是從模型估計值中減去這兩個粗粒度變數經迴歸估計的價值貢獻——把它們從銷售網絡迴歸部分下的已測量屬性中移出——並將其加回到殘差之上。此後，再從殘差中移除這部分並將其分配給相應的細粒度潛變數 `res_condition1` 和 `res_quality1`。隨後利用

從 MLS 公開備註 (public remarks) 及相關字串列表中提取的附加資訊，對這些潛變數的取值進行微調。最終結果是，細粒度價值貢獻取代粗粒度價值貢獻：粗糙評級只是近似模型的腳手架，最終報告的狀況與品質取值是從殘差推導出的細粒度取值。

11.3 資料錯誤

根據社交媒體上的評論，加利福尼亞州的多重掛牌服務 (MLS) 資料總體上比美國許多其他地區品質更高。通常，在大多數成熟的大都市地區，MLS 資料的品質應當是可以預期的良好，而在農村地區則較差。

儘管資料中因各種原因偶有錯誤，但只要資料量足夠，這些錯誤一般應相互抵消，對價格模型所用平均值的影響很小。

然而，特定地區可能出現系統性偏誤。例如，考慮這樣一個社區：許多住宅最初建有完整的二層，但部分買方可以選擇拆除二樓房間，為一層的一個或多個房間打造挑高拱頂天花板，從而顯著減少總居住面積 (GLA)。由於種種原因，這些改動後來沒有在評估機構的記錄中更新。結果，MLS 可能顯示 2,400 平方英尺，而實測值只有 2,000 平方英尺。在這類情形下，應在執行多元自適應迴歸樣條 (MARS) 迴歸之前先更正 MLS 資料。需要注意的是，為可比交易案例獲取這類資訊可能需要一些調查工作。發現的差異應予以記錄以備將來參考。

綜上所述，可以合理假設，在 MLS 資料品質較高的地區，溢入殘差的錯誤與設計、狀況、品質等主觀且未測量特徵的影響相比是微不足道的。在估計 RCA 方法的準確性時，這是一個重要的考慮因素。例如，如果某個市場區域的 R^2 值為 80%，那麼價格中約有 20% 歸因於 CQA 殘差。因此，在殘差排序中定位估價對象時 $\pm 10\%$ 的誤差，傳導到最終價值結論上就是 $\pm 10\% \cdot 20\% = \pm 2\%$ 的誤差。這就是提高準確性的途徑！在許多情況下，可以得到誤差遠低於 $\pm 1\%$ 的準確價值估計。

因此，在 MARS 迴歸中把殘差當作未建模變數集體價值的間接指標是合理的，前提是估價工程師具備足夠的當地市場經驗，能夠識別系統性測量偏誤存在於何處以及為何存在。這種方法確保分析準確反映實際不動產價值，同時也承認不動產資料的侷限性與細微之處。

實例演算：在 Sea Ranch 測量未被測量的價值驅動因素

當某個市場區域可實現的 R^2 低於該市場表面同質性所應達到的水準時，本框架的應對方式是識別並測量缺失的變數，而不是接受這一侷限。索諾馬郡的 Sea Ranch 社區直接說明了這一點。

這是「測量而非接受」式應對的一個具體實例。索諾馬郡的 Sea Ranch 社區是一個沿海開發區，到海洋的距離是主要的價值驅動因素：在其他特徵相同的情況下，距離懸崖幾百英尺的住宅與向內陸四分之一英里的住宅，價格差異巨大。該社區的 MLS 資料並未把到海距離記錄為變數。基於現有 MLS 變數擬合的初始 MARS 模型，其解釋力低於該市場表面同質性所應支持的水準，且殘差結構未能按資料中任何可見特徵充分地對不動產排序——這一信號表明模型缺失了一個重要的驅動因素，而不是該市場本身噪聲很大。

補救辦法是直接測量。作者獲取了一張顯示不動產相對海岸線布局的社區地圖，並為數百套住宅測量了到海距離。將測得的距離作為變數加入 MLS 資料集，並重新執行 MARS 迴歸。新變數被模型選為顯著變數，解釋力大幅提升，隨後的殘差排序按與其餘主觀特徵——景觀、狀況、品質——相一致的模式對不動產排序，而不再受此前污染殘差的潛在到海距離信號的影響。在 MLS 變數缺失或不可靠的山坡社區，高程問題也以類似方式得到了處理。

這一方法論原則具有普遍性。當可實現的 R^2 低於該市場表面同質性所應支持的水準時，首要假設是模型缺失了一個重要變數，恰當的應對是弄清它是什麼並直接收集測量資料。資料來源包括社區地圖、GIS 資料、稅務評估記錄、建築管理部門記錄、衛星影像以及實地查勘。這類測量工作的成本是實實在在的，但有限度；一旦為某個市場區域收集了該變數，就可以在該區域的後續估價任務中沿用。RCA 能夠順暢地吸納新變數，因為該框架並不在意哪些變數進入

MARS 迴歸——殘差結構會告訴 VE 新變數是否捕捉到了真實信號，各項價值貢獻也會相應調整。

另一種應對方式——接受較低的 R^2 並報告更寬的誤差區間——在缺失的驅動因素確實無法測量（買方特有的偏好、短暫的市場情緒、交易的特殊情形）時是站得住腳的，但當這些因素只是尚未收集時則不然。面對一個弱模型，VE 的第一步應當是追問缺失的信號是否屬於後者。

11.4 什麼是準確性？

完美準確的估價要求：

1. 符合市場價值：估價對象的假設性銷售交易 100% 符合給定的市場價值定義。
2. 準確而穩健的模型開發：建立一個高品質的 MARS 模型，視地區而定，其 R^2 約為 75-80%，CVR2 約為 55-70%。
3. 良好的資料錯誤分析：已測量變數所有可能的資料錯誤都得到理解和控制，幾乎不溢入殘差。
4. 系統性偏誤核算：資料中任何系統性偏誤都應予以核算和調整。
5. 殘差排名分析：確保按殘差排名的不動產從最低到最高排名呈現出合理的特徵模式。遺產出售、虧本急售和拍賣出售等異常值應予以調查、解釋，並在必要時剔除。
6. 估價對象定位：在透徹理解殘差排序的基礎上，應準確定位估價對象。找出一處吸引力明顯更高的不動產和一處吸引力明顯更低的不動產，將估價對象置於兩者之間並打分。然後盡量收窄較高與較低不動產之間的排名區間。
7. 解釋性推理：絕不為追求準確性而犧牲解釋性推理。估價師必須能夠解釋估價對象與可比不動產之間的價值差異，至少要讓一個具有正常理解能力的人明白。例如，一個 R^2 和 CVR2 達到最優的良好 MARS 模型可用於解釋已測量特徵的價值貢獻，而估價工程師對殘差的分解則可用於解釋其他特徵的價值貢獻。
8. 過擬合：避免模型過擬合，不使用那些單獨或組合起來能夠識別一處或數處不動產的不動產特徵。尤其要避免使用其取值能夠鎖定少數不動產的因子變數。在規模足夠大的資料集中，用於標示社區的因子變數是合適的，但在游泳池極少的地區標示帶游泳池不動產的因子變數則必須慎用。標示唯一不動產的特徵，例如精確的 GIS 坐標、地號或地址，是絕對不允許的。

如果迴歸的 R^2 為 80%，我們預計估價對象價值中約有 20% 歸因於殘差。當然，這只是近似值。從 CQA—殘差曲線可知，殘差分布的大部分集中在 CQA 的低端和高端兩個極值區。基於 CQA—殘差曲線以及低、高 CQA 值處的極端斜率，我們可能預期在這些區域對殘差的估計不那麼準確。然而，這一點被這些區域不動產的相對稀少以及相鄰不動產之間吸引力差異更大這兩個因素所抵消。因此，儘管在中間區段我們談論的影響範圍可能是 $\pm 5\%$ ，在兩端我們更應關注的則是略窄的 $\pm 2 - \pm 3\%$ 範圍。

一般而言，由於 CQA—殘差曲線的非線性，計算 RCA 的準確性有一定難度。但只要有一個合理的殘差排序，就可以逐案計算。CQA 分值的範圍是 0.0 到 10.0。有了良好的 MARS 模型，你的殘差排序應呈現出從吸引力最低到最高的清晰模式，你應該不難把評分控制在總範圍的 $\pm 5\%$ 之內，即 CQA 分值的 ± 0.5 之內。如果迴歸中有 100 處不動產，這相當於把估價對象準確地排入一個 10 處不動產的組內。那麼，參照表 3 中的伯靈格姆銷售資料，讓我們分別針對 VE 估計的 CQA 為 7.5、5.1 和 1.9 的情形計算近似準確性（或誤差），其中他對估價對象的價值結論為 \$3,000,000。注意，由於表 3 只包含所用數百處不動產中的一個樣本，必要時你需要（假設線性）對殘差進行插值。需要明確的是，我們假定 VE 在每種情形下都非常確信，他對估價對象的排名在僅含 100 處不動產的殘差排序中處於 ± 5 處不動產的範圍之內。

CQA= 7.5 ±0.5					
CQA	殘差	絕對差	預期誤差	假設成交價	誤差百分比
8.00	\$257,112				
7.50	\$197,113	\$59,999			
7.00	\$154,615	\$42,498	±\$51,249	\$3,000,000	±1.71%

CQA=5.1 ±0.5					
CQA	殘差	絕對差	預期誤差	假設成交價	誤差百分比
5.60	\$50,911				
5.10	\$19,559	\$31,352			
4.60	-\$28,207	\$47,766	±\$39,559	\$3,000,000	±1.32%

CQA= 1.9 ±0.5					
CQA	殘差	絕對差	預期誤差	假設成交價	誤差百分比
2.40	-\$228,411				
1.90	-\$276,381	\$47,970			
1.40	-\$326,938	\$50,557	±\$49,264	\$3,000,000	±1.64%

最後，驗證準確性的唯一客觀方式，是建立一套嚴格的規程（本文只是粗線條地部分提及）並核查其遵守情況。之所以必要，是因為受市場不完善的影響，實際成交價格可能偏離市場價值。將估價與其後的成交進行比較，可以在一定程度上反映準確性。然而，不動產市場的複雜性和不穩定性，要求由估價師或高度合格的估價工程師依據既定標準和規程給出價值估計。這些標準和規程應當足夠嚴格，以確保勝任的估價工程師能夠各自獨立地得出幾乎相同的價值結論。在大多數能夠獲得良好資料的情況下，這應當是一個可以實現的目標。

11.5 作為異常值的估價對象

既然估計一處不動產的市場價值需要分析近期的銷售交易，那麼就有理由追問：對於一處缺乏現代化改造和維修的較舊的估價對象，近期銷售交易能否很好地代表其各項特徵。在舊金山灣區尤其如此，那裡近期成交的住宅通常都經過翻新，以提高其售價。因此，可比成交案例的平均品質和狀況通常優於那些很久沒有進入市場的老舊住宅。這一價值差異應當對估價對象殘差的估計——換言之，由其 CQA 分值——來捕捉。就估價對象估值中的資料品質與準確性而言，這是一個重要的考慮因素。

11.6 CVR2 在應對意外資料方面的重要性

MARS 迴歸基於對隨機選取的不動產子集多次迭代執行 R/earth，該子集通常佔 MLS 資料集的 80% 到 90%，稱為訓練集。其餘 10% - 20% 的不動產作為測試集。在這一階段，先由訓練資料集建立 MARS 模型，再用它預測測試資料集中不動產的成交價格，然後將預測值與實際成交價格進行比較。這一方法已被資料探勘者公認為構建穩健模型、預測尚未觀察到的不動產成交價格的最有效策略。

然而，CVR2 指標的主要用途並不在於反映迴歸模型對實際市場區域成交的事實準確性，而在於最佳化模型對以指定生效日期為準的假設性不動產交易的預測能力。這些不動產不在訓練資料集中，可能表現出訓練集和測試集不動產資料中都缺失的新的或意外的特徵組合。

在這一背景下，假設一處過去 15 年從未翻新的不動產在生效估價日期前一個月掛牌，我們如何估計其成交價格？一個 CVR2 表現強勁的模型，比一個未經過這種反覆交叉驗證的模型更有可能給出最準確的估計。同樣至關重要的是，資料集中要包含足夠數量、翻新、維修和現代化改造程度各異的可比不動產。

歸根結底，如果我們擁有一個穩健的模型，能夠準確預測可比不動產資料集中各項不動產特徵的價值貢獻，我們就能有效地把新的估價對象放入模型按殘差對這些不動產的排序之中。這一定位使我們能夠為該不動產賦予恰當的 CQA 分值和殘差。

12 殘差

正如「殘差約束法」這一名稱所示，殘差是本方法論的核心。此處的殘差是指不動產的淨成交价格與迴歸分析所預測的成交价格之間的差額。

正殘差表明，基於建築面積（GLA）、土地面積等典型可測量屬性，該不動產的成交價高於其估計價格，暗示其吸引力可能高於平均水準。負殘差則表明該不動產的成交價偏低，可能意味著吸引力較低或存在其他未被捕捉到的負面屬性。

考慮到較大的住宅通常具有較大的殘差，我們可以使用基於 GLA 的每平方英尺殘差進行更規範化的比較。關於這一問題的更深入討論將留待另一篇論文。

12.1 排序與評分

在為已測量變數建立 MARS 模型之後，為每一處不動產計算殘差。這些殘差表示觀測到的淨成交价格與模型預測的成交价格之間的差額。還應額外建立一欄「殘差/平方英尺」，即殘差除以住宅的居住面積——這是基於這樣的認識：一般而言，具體不動產的殘差會隨居住面積的大小而增大。其他可能的統計量在某些社區中或許效果更好，但那是另一個話題。

假設 7：「按殘差排序」

在本文中，凡提及「按殘差排序」，均應理解為按殘差排序、按殘差/平方英尺排序，或按估價工程師認為最適合該市場區域的其他殘差函數排序。作者在舊金山灣區的經驗表明，殘差/平方英尺通常最為可靠，尤其是在處理 GLA 異常小或異常大的估價對象時。

接下來，將各不動產按其殘差從最負到最正進行排序。然後為每一處不動產生成一個「狀況—品質—吸引力」（CQA）評分，取值範圍為 0.0 至 10.0。該評分的計算方法如下：

CQA 評分對應於某一不動產的殘差在所有不動產中的百分位排名。具體而言，評分的計算方法是：確定殘差低於該不動產的不動產所佔的百分比，再將該百分比除以 10。例如，若某一不動產的殘差高於 50% 的不動產的殘差，則其評分為 5.0。

由於四捨五入的原因，可能有一處不動產獲得 10.0 分，即殘差最高的那處不動產。然而，10.0 的滿分在理論上是不可能的，因為從邏輯上講，沒有任何不動產的殘差能夠高於 100% 的不動產（包括其自身）。另一方面，最低分 0.0 是可能出現的，因為至少有一處不動產不存在殘差更低的其他不動產。請注意，當比較的不動產數量很多時，由於四捨五入，可能不止一處不動產得到 0.0 分。

因此，CQA 評分提供了一個直觀的度量，用以衡量某一不動產的屬性與模型預期的吻合程度——評分越高，表明與模型預測的吻合度越好。

必須承認，只有當銷售條件與其市場價值相稱時，CQA 評分才可被視為可比不動產吸引力的間接指標。估價工程師有責任透過對銷售交易進行全面解釋和徹底調查，來查明殘差價值中是否包含高於或低於市場的因素。雖然顯著高於或低於市場的銷售並不常見，但在強制出售、遺囑認證、當事人無行為能力、串通等情形下仍可能發生。

表 3 展示了 MLS 資料的一個半真實節選，連同殘差和 CQA 評分（包括總額和每平方英尺兩種形式）。本研究分析的不動產總數約為 600 處。值得注意的是，若干參與迴歸的變數並未顯示。換言之，所展示的殘差是在 MARS 迴歸已考慮了成交日期等其他變數的影響之後得出的。

表 3: 按殘差排序的可比交易案例

成交價格	估計成交價	殘差	CQA	殘差/平方英尺	CQA SF	GLA	成交日期
\$3,450,000	\$2,318,865	\$1,131,135	10	\$420.50	9.98	2,690	2021-04
\$4,360,757	\$3,399,922	\$960,835	9.9	\$290.28	9.82	3,310	2021-09
\$4,000,000	\$3,134,444	\$865,556	9.8	\$234.57	9.60	3,690	2020-11
\$5,850,000	\$5,150,966	\$699,034	9.7	\$109.91	8.35	6,360	2021-12
\$3,600,000	\$3,023,366	\$576,634	9.5	\$269.46	9.79	2,140	2022-06
\$4,425,000	\$3,987,873	\$437,127	9.2	\$127.44	8.68	3,430	2021-08
\$2,530,000	\$2,164,564	\$365,436	8.9	\$180.02	9.28	2,030	2019-12
\$4,250,000	\$3,884,719	\$365,281	8.8	\$87.18	7.91	4,190	2021-08
\$3,350,000	\$3,034,613	\$315,387	8.5	\$130.87	8.75	2,410	2021-09
\$3,800,000	\$3,522,028	\$277,972	8.2	\$77.43	7.72	3,590	2020-09
\$4,200,000	\$3,963,748	\$236,252	7.8	\$56.12	6.89	4,210	2021-01
\$2,900,000	\$2,702,887	\$197,113	7.5	\$89.19	7.92	2,210	2021-07
\$4,250,000	\$4,103,885	\$146,115	6.9	\$39.38	6.29	3,710	2022-08
\$5,000,000	\$4,864,958	\$135,042	6.8	\$28.25	5.85	4,780	2021-07
\$4,320,000	\$4,199,672	\$120,328	6.5	\$33.90	6.04	3,550	2022-07
\$2,600,000	\$2,479,904	\$120,096	6.5	\$60.96	7.00	1,970	2021-10
\$3,400,000	\$3,301,146	\$98,854	6.2	\$30.32	5.90	3,260	2021-06
\$2,000,000	\$1,901,987	\$98,013	6.2	\$73.14	7.54	1,340	2019-11
\$3,000,000	\$2,919,718	\$80,282	5.9	\$26.50	5.79	3,030	2020-05
\$6,750,000	\$6,699,089	\$50,911	5.6	\$7.75	5.20	6,570	2020-08
\$5,652,000	\$5,601,570	\$50,430	5.5	\$12.83	5.37	3,930	2022-05
\$2,600,000	\$2,579,082	\$20,918	5.2	\$10.51	5.33	1,990	2022-03
\$1,925,000	\$1,905,441	\$19,559	5.1	\$11.51	5.36	1,700	2021-01
\$3,200,000	\$3,209,500	-\$9,500	4.8	-\$3.21	4.88	2,960	2020-11
\$4,150,000	\$4,187,560	-\$37,560	4.5	-\$10.49	4.68	3,580	2021-04
\$2,700,000	\$2,776,510	-\$76,510	3.9	-\$23.61	4.14	3,240	2021-07
\$3,450,000	\$3,567,668	-\$117,668	3.4	-\$31.05	3.90	3,790	2020-09
\$1,975,000	\$2,109,255	-\$134,255	3.3	-\$55.48	3.24	2,420	2020-02
\$1,950,000	\$2,085,431	-\$135,431	3.2	-\$68.40	2.86	1,980	2021-02
\$3,695,000	\$3,859,154	-\$164,154	3.1	-\$47.04	3.56	3,490	2021-05
\$2,288,000	\$2,455,747	-\$167,747	2.9	-\$56.86	3.19	2,950	2020-06
\$2,635,000	\$2,802,781	-\$167,781	2.9	-\$56.87	3.18	2,950	2020-09
\$2,050,000	\$2,217,794	-\$167,794	2.9	-\$82.66	2.46	2,030	2021-07
\$1,550,000	\$1,731,804	-\$181,804	2.7	-\$130.79	1.32	1,390	2021-05
\$2,800,000	\$3,028,411	-\$228,411	2.4	-\$81.58	2.51	2,800	2021-02
\$3,700,000	\$3,976,381	-\$276,381	1.9	-\$76.99	2.62	3,590	2021-08
\$2,517,500	\$2,834,974	-\$317,474	1.5	-\$120.71	1.48	2,630	2019-11
\$2,550,000	\$2,905,330	-\$355,330	1.1	-\$115.74	1.64	3,070	2019-12
\$1,815,000	\$2,263,944	-\$448,944	0.7	-\$273.75	0.15	1,640	2021-11
\$3,711,000	\$4,683,454	-\$972,454	0	-\$226.68	0.37	4,290	2021-12
\$2,225,000	\$3,332,960	-\$1,107,960	0	-\$370.56	0.03	2,990	2021-08
\$3,550,000	\$4,746,489	-\$1,196,489	0	-\$274.42	0.14	4,360	2022-06

12.2 CQA-殘差曲線

將某一社區的 CQA 與殘差/平方英尺作圖，必然得到 圖 2 所示的 CQA-殘差特徵曲線。該圖所示為 2022 年前後加州 Burlingame 的資料。將 CQA 評分與殘差值作圖通常也會觀察到類似的曲線；不過，我更傾向於使用殘差/平方英尺，因為它能夠結合估價對象的建築面積（GLA）來解讀 CQA。將殘差/平方英尺乘以 GLA，即可把某一不動產的 CQA 轉換為相應的殘差價值。

估價工程師必須仔細評估不動產的規模——具體而言是建築面積（GLA）——是否顯著影響殘差的大小。這一評估至關重要，因為它可以決定計算不動產價值的最準確方法。例如，高端廚房電器和豪華衛浴設施可以大幅提升住宅的殘差價值。然而，在廚房和衛生間的尺寸與布局普遍統一的社區中，這些升級可能與 GLA 的相關性並不強。它們對殘差價值的影響，可能更多在於它們所代表的品質，而非它們所佔據的空間。

相反，昂貴的外部特徵和材料能夠顯著提升不動產的門面吸引力，可能與不動產的規模有更直接的關係。較大的住宅可以容納更精緻的外部設施，例如大規模的園林綠化、更優質的屋頂材料或寬敞的露台，這些自然會提升不動產的整體吸引力，進而提升其殘差價值。

因此，估價工程師在分析殘差時必須考慮這些動態因素。他們不僅應評估高品質設施和特徵所帶來的內在價值，還應理解這些改進如何與不動產的物理尺寸相互作用。這種全面的方法可確保估價準確反映不動產的定性與定量屬性。

因此，建議同時繪製 CQA-殘差曲線和 CQA-殘差/平方英尺曲線並研究其差異，這可以揭示某一市場區域中的模式。估價工程師可以讓程式同時計算兩種殘差，並取某種加權平均來計算分配給估價對象的殘差。其影響可能大，也可能不大。

12.3 CQA 曲線的特徵

「狀況—品質—吸引力」（CQA）曲線提供了一個精細的模型，用以理解不動產價值如何隨給定區域內的吸引力、狀況和市場動態而波動。該模型捕捉了影響不動產市場的經濟因素與物理因素之間的相互作用，反映了不同細分市場中不動產估價的微妙現實。

12.3.1 CQA 曲線各組成部分的分析

1. 曲線左側的指數衰減：

當 CQA 評分從約 2 降至 0 時，不動產表現出快速惡化的特徵。曲線的這一部分可以很好地用指數衰減函數來表示，突顯了缺乏適當維護和更新的不動產會以多快的速度喪失其吸引力和價值。木材腐爛、顏色褪去、金屬鏽蝕等物理退化會加速這種衰退，尤其是在未定期維護的住宅中。

2. 中間區段的線性增長：

市場的主體通常屬於這一類別，其吸引力在評分約 2 至 8 之間近乎線性地增長。這一區段反映了那些總體上維護達標、但不一定具備突出品質或區位的不動產。它代表了住房市場中廣泛的中間層級，涵蓋了住房市場的大多數——這些住宅基於其標準特徵和整體狀況進行有競爭力的定價。

3. 曲線右側的指數增長：

在此，曲線在評分 8 至 10 之間呈指數上升，映照出市場中吸引力較高的一端。這一細分市場中的不動產往往屬於較富裕的人群，他們不僅投入於更優質的維護，還投入於能顯著提升不動產吸引力和價值的改進。指數增長反映了此類高品質不動產的稀缺性以及富裕買方的旺盛需求。而且要明白，一套價格較低的住宅在吸引力方面可能評分很高，而一套面積大、價格昂貴的住宅評分卻可能很低。這些動態可透過對 MARS 的專業運用而揭示出來，而大多數估價師通常不會注意到它們。

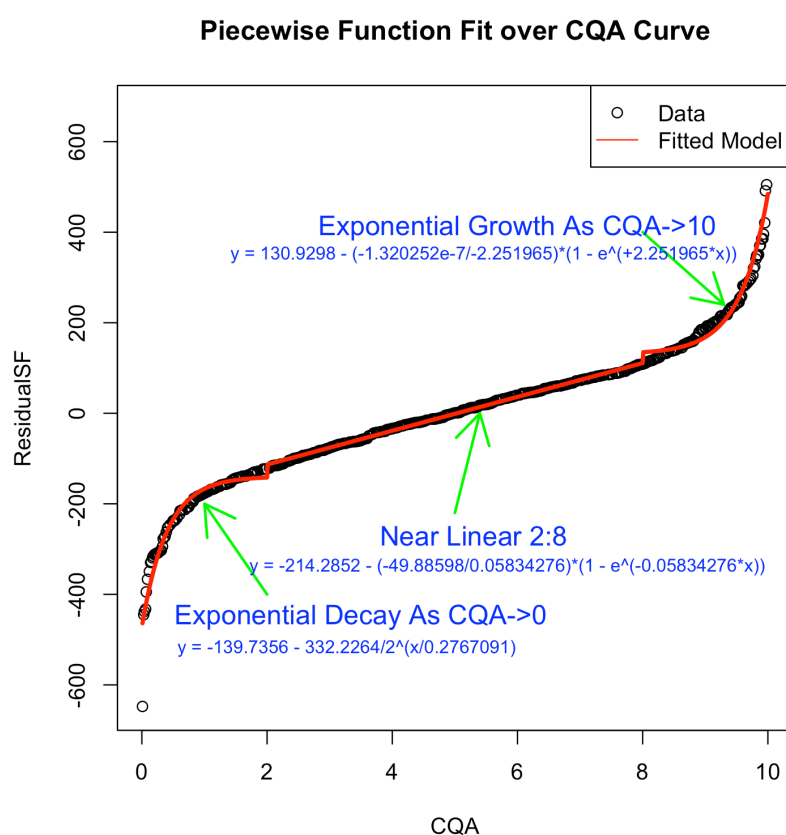
12.3.2 市場動態與社會含義

CQA 曲線隱含地講述了不同區域內相對財富與經濟差距的故事。在聖馬特奧郡這樣的地方，Burlingame 等區域與 Hillsborough、Woodside 或 Atherton 等更富裕社區之間的不動產價值差異，展示了顯著的社會經濟分化。這種差距影響著從不動產維護到市場估值的方方面面，並最終影響買方群體構成和不動產市場趨勢。

此外，還有這樣一個例子：一位相對富有的人士，在自己感到舒適、且出於種種原因更願意居住的中檔社區中投資建造高品質住宅——這凸顯了不動產經濟學的另一個關鍵方面——「過度建造」不動產的概念。此類不動產雖然極具吸引力，但可能在當地找不到能夠支付其在更高價位市場區域本可賣出的價格的買方。

CQA-殘差曲線是每個市場區域所特有的，在進行分析時不應被忽視。

圖 2: CQA-殘差特徵函數



12.4 估價對象殘差

與可比交易案例不同，估價對象並不具有客觀確定的殘差；它必須由估價工程師（VE）估計。這一估計為殘差約束法（RCA）規程引入了一定程度的主觀性，標誌著個人判斷發揮關鍵作用的一個重要環節。

為了估計「狀況—品質—吸引力」（CQA）評分，VE 會細緻地分析已排序的不動產，在排序中尋找可辨識的模式。這一分析通常涉及對每處不動產相關 MLS 照片的詳細審閱。對於大型資料集（如 600 處不動產）而言，逐一詳盡審閱並不現實。相反，VE 運用實務推理來近似判斷估價對象在整體排序中的可能位置。需要大量修繕的不動產（即「待修房」）通常位於排序底部附近，而更新程度高、吸引力強的不動產則往往排在頂部附近。

隨後，VE 採用更聚焦的方法，在排序中「跳躍」，以識別與估價對象吸引力最接近的可比交易案例。隨著範圍的縮小，VE 會更加關注各不動產之間的相似之處，尤其是在品質、狀況、設計美感（包括油漆、木作、石作和窗戶等元素）、功能效用及其他相關特徵等方面

必須理解的關鍵一點是，VE 的判斷從根本上受市場所確立的排序的引導。這確保了該過程中的主觀因素被錨定於市場現實，力求對估價對象作出公平而準確的估價。

12.5 RCA 價值結論

12.5.1 一般考慮

確定 CQA 評分對於確立不動產的估計殘差價值至關重要，因為這兩個要素本質上是相互關聯的。雖然不使用 CQA 也可以對不動產進行評估，但這一指標透過展示該不動產相對於市場區域內其他不動產所處的位置，提供了有價值的背景資訊。例如，CQA 評分為 3.2 表明該不動產的吸引力高於周邊 32% 的可比不動產。如果對所賦評分存在異議，異議方必須查閱排序，並基於照片證據或其他實質性資料來論證其提出的任何調整。如果排序體系是穩健的，且估價工程師（VE）已在此框架內準確定位了該不動產，則不太可能出現重大分歧。

在報告階段，通常會選擇 6 至 12 個近期且相關的可比交易案例來支持估價結論。對這些可比交易案例與估價對象之間的差異進行分析，並對每個變數作相應調整。然後將這些調整應用於可比交易案例的淨成交價格，得出調整後成交價格。正如數學上已證明的那樣，這些調整後價格應收斂於估價對象的估計成交價格。

12.5.2 複核、審計與投訴

一旦確定了估價對象的估計殘差，就可以為所有可比交易案例計算殘差調整。這些調整對 VE 針對其認為相關的特定 CQA 變數所作的進一步手工分解構成約束。必須強調的是，正如證明 II 所確認的，這些細化分解僅具解釋性，並不影響最終估值。

對於不滿意估價價值結論的使用者所提出的投訴，傳統估價師通常不得不浪費時間解釋為什麼某處價格更高或更低的不動產並不具有使用者所期望的價值影響。而使用 RCA，所有可能的可比交易案例（通常為 100-600 個）通常都被評估為與價格結論完全相同的價值，因此展示任意某處不動產的調整幾乎不費力氣。除估價對象殘差的評分外，整個 RCA 過程都是自動化且客觀的。因此，RCA 的另一個優勢在於，針對 VE 主觀判斷不完善之處的批評主要侷限於這一個環節——並且可以得到有效而高效的管理。此外，假設 R^2 約為 80%，估計估價對象殘差時 $\pm 10\%$ 的誤差對最終價值估計的影響僅為 $\pm 2\%$ 的誤差。

在許多情況下，估價對象在不動產殘差排序中的契合度如此之高，以至於可以預期其準確度在 $\pm 1\%$ 以內。

估計殘差與估價對象的 MARS 估計成交價格之和，即得出最終的價值結論，稱為殘差約束法（RCA）價值結論。這一過程確保了不動產估價方法的全面性與透明度，增強了人們對評估結果準確性與公正性的信心。

13 結論

殘差約束法是一個具有三項實質性貢獻的框架：其一，源於在整個市場而非小規模可比交易案例集合上擬合模型的結構性抗錨定性質（證明 I）；其二，一項不變性結果，使殘差分解得以作為解釋層而不擾動價值結論（證明 II 及其後的注記）；其三，一項經驗規律——CQA-殘差特徵曲線——它似乎以特徵性的「指數—線性—指數」形態在各市場區域中反覆出現。該框架的侷限性同樣具有實質性：殘差分解具有不變性但不具識別性；CQA 定位步驟依賴尚未經正式評分者間一致性研究驗證的 VE 判斷；其準確性主張尚待與替代方法進行基準比較。

本文所呈現的是一個框架和一套規程。本文尚未呈現、但目前正在進行的更大研究計畫所提供的，是一條可執行、可審計追溯的管線，它以具有可辯護預設設定的軟件將該規程付諸操作。這條管線——用於 MARS 基底函數階段的 CRAN 上的 earthUI 包、用於在這些基底函數上構建正則化結構模型的 glmnetUI 包，以及用於空間平滑的 mgcvUI 包——是本期姊妹文章的主題，也是從「作為方法的 RCA」通往「作為可交付成果的 RCA」的橋梁。估價監管中的理論與實務鴻溝（Craytor, 2026, 本期）說明了這座橋梁為何重要：估價產出在經濟上被當作前瞻性風險輸入來使用，而一個使現值意見可審計、可重現的框架，是對這種前瞻性用途進行任何誠實探討的前提條件。RCA 是這一更大計畫中的一步棋。它並非整個計畫，本文也並未如此宣稱。

參考文獻

- Ambrose, Brent W. et al. (2025). “Do Appraiser and Borrower Race Affect Mortgage Collateral Valuation?” In: *Review of Finance*. DOI: [10.1093/rof/rfaf046](https://doi.org/10.1093/rof/rfaf046).
- Craytor, William Bert (2026). “The Doctrine–Practice Gap in Real Estate Appraisal: A Structured Account of Functions, Boundaries, and Tensions”. In: *Valuation Engineer Journal* 1.1.
- Friedman, Jerome H. (1991). “Multivariate Adaptive Regression Splines”. In: *The Annals of Statistics* 19.1, pp. 1–141.
- Hastie, Trevor, Robert Tibshirani, and Jerome H. Friedman (2017). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second. Springer Series in Statistics. New York, NY: Springer. ISBN: 978-0-387-84857-0.
- James, Gareth et al. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer Texts in Statistics 103. New York: Springer. ISBN: 978-1-4614-7137-0.
- Krivorotov, George and Michael LaCour-Little (2020). “AVM versus Appraisal-Based Underwriting in Refinance Mortgages: The Trade-off Between Noise and Bias”. In: *SSRN Electronic Journal*. DOI: [10.2139/ssrn.3603327](https://doi.org/10.2139/ssrn.3603327).
- Kuhn, Max and Kjell Johnson (2016). *Applied Predictive Modeling*. Corrected at 5th printing. New York: Springer. ISBN: 978-1-4614-6848-6.
- Milborrow, Stephen (2024a). *Earth: Multivariate Adaptive Regression Splines*. R Package. URL: <https://CRAN.R-project.org/package=earth>.
- (2024b). *Notes on the Earth Package*. Tech. rep. URL: <http://www.milbo.org/doc/earth-notes.pdf>.
- (2024c). *Variance Models in Earth*. Tech. rep. URL: <http://www.milbo.org/doc/earth-varmod.pdf>.
- Perry, Andre M., Jonathan Rothwell, and David Harshbarger (2018). *The Devaluation of Assets in Black Neighborhoods*. Tech. rep. Brookings Institution.
- Perry, Andre M., Jonathan Rothwell, J. Williamson, et al. (2024). *How Racial Bias in Appraisals Affects the Devaluation of Homes in Majority-Black Neighborhoods*. Tech. rep. Brookings Institution.

- Property Appraisal and Valuation Equity Task Force (2022). *Action Plan to Advance Property Appraisal and Valuation Equity*. Tech. rep. U.S. Department of Housing and Urban Development.
- Tzioumis, Konstantinos (2016). “Appraisers and Valuation Bias: An Empirical Analysis”. In: *Real Estate Economics* 45.3, pp. 679–712. DOI: [10.1111/1540-6229.12133](https://doi.org/10.1111/1540-6229.12133).
- Williamson, Jake and Mark Palim (2022). *Appraising the Appraisal: A Closer Look at Divergent Appraisal Values for Black and White Borrowers Refinancing Their Home*. Tech. rep. Fannie Mae.
- Zhang, Wengang (2020). *MARS Applications in Geotechnical Engineering Systems: Multi-Dimension with Big Data*. Singapore: Springer. ISBN: 978-981-13-7421-0.

earthUI : earth (MARS) 套件的互動式介面

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。

William Bert Craytor

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: [10.67330/7fj41b65](https://doi.org/10.67330/7fj41b65)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 [創用 CC 姓名標示 4.0 國際授權條款 \(CC BY 4.0\)](#) 的條款發布。

摘要。 earthUI 是 R 語言 earth 套件的圖形使用者介面，該套件實作了多元自適應迴歸樣條 (MARS, multivariate adaptive regression splines)。它提供三種用途模式——通用預測建模、房地產估價和市場區域分析——並引導使用者完成資料匯入、模型配置、模型擬合，以及對 earth 診斷資訊和圖形輸出的解讀。本文記錄了 earthUI 的資料格式要求、建模工作流程、輸出顯示介面以及完整的功能參考。

關鍵詞： MARS，多元自適應迴歸樣條，earth，R，圖形使用者介面，房地產估價

出版者按 (2026 年 7 月)：本文描述的是 earthUI 0.10.0。付印前不久，earthUI 0.11.0 吸收了配套應用 glmnetUI 和 mgcvUI (見隨後兩篇文章)，將 MARS、彈性網絡和 GAM 工作流程整合到單一應用程式中。本文所述的介面與工作流程不受此次整合的影響。

目錄

1	引言	3
1.1	什麼是 earthUI ?	3
1.2	一個套件，三個常式	4
1.3	三種用途模式	4
1.4	房地產專用功能	4
1.5	安裝 earthUI 與 vProlog	5
1.6	快速入門	5
2	MLS 輸入資料要求	7
2.1	檔案格式與結構	7
2.2	估價模式所需的欄	7
2.3	特殊欄命名約定	8
2.4	資料品質與完整性	8
2.5	估價對象列的位置	8
3	通用模式	9
3.1	概述	9
3.2	側邊欄工作流	9
3.3	主面板頁籤	9
3.4	設定持久化	10
3.5	深色模式	10
3.6	語言環境與區域設定	10
3.6.1	支援的國家	11
3.6.2	紙張覆蓋設定	11
3.6.3	儲存預設設定	11

4	估價模式	12
4.1	估價對象列的處理	12
4.2	生效日期與銷售日齡	12
4.3	特殊欄指定	12
4.4	殘差貢獻（稀有特徵估計）	13
4.5	RCA 調整概述	13
5	市場區域分析模式	15
5.1	與估價模式的區別	15
5.2	何時使用市場模式	15
6	變數選擇	16
6.1	目標變數	16
6.2	預測變數表	16
6.3	資料類型檢測與覆蓋	16
6.4	Factor 與 Linear 標誌	16
6.5	特殊欄類型參考	17
6.6	多個目標變數	17
7	參數選擇	18
7.1	基本參數	18
7.2	前向過程	18
7.3	允許的交互	19
7.4	剪枝	19
7.5	交叉驗證	19
7.6	子集過濾	19
7.7	推薦參數值	20
7.7.1	前向過程參數	20
7.7.2	剪枝參數	21
7.7.3	交叉驗證參數	21
7.7.4	變異數模型	22
7.7.5	Earth 輸出檔案	22
7.8	設定預設值	22
8	擬合 Earth 模型	23
8.1	Fit 按鈕	23
8.2	擬合模態框與追蹤輸出	23
8.3	擬合日誌	23
9	下載資料	24
9.1	輸出欄	24
9.2	欄排序與 Excel 格式	24
9.3	CQA 評分	25
9.4	排序與估價對象列	25
9.5	mgcvUI 自動匯出	25
9.6	glmnetUI 匯入	25
9.6.1	主要優勢	25
9.6.2	重要提示：權重欄必須一致	26
9.6.3	工作流程	26

10 RCA 計算與下載	27
10.1 開啟 RCA 對話方塊	27
10.2 CQA 評分插值	27
10.3 輸出欄	27
10.4 RCA Adjustments 頁籤	28
11 銷售比較網格	29
11.1 概述與用途	29
11.2 可比案例選擇模態框	29
11.3 網格佈局	29
11.4 售價、讓價優惠與 Net SP	30
11.5 分組列（位置、地塊、屋齡）	30
11.6 模型變數列	30
11.7 CQA Residual 與殘差特徵列	30
11.8 調整後售價公式	30
11.9 工作表保護與可編輯儲存格	31
11.10 在 Excel 中使用網格	31
12 下載報告	32
12.1 生成 Quarto 報告	32
12.2 轉換 Quarto 報告	32
12.3 報告內容	32
13 後處理模組：彈性網絡與 GAM	34
13.1 為什麼需要輔助擬合？	34
13.2 啟動與共享狀態	34
13.3 彈性網絡常式	34
13.4 GAM 常式	34
13.5 一個實用的工作流	34
14 示範資料集：Appraisal_1.csv	36
14.1 描述	36
14.2 資料欄	36
14.3 建議的快速上手步驟	37
15 實驗性功能：Prolog 文字提取與推導規則	38
15.1 它的作用	38
15.2 環境要求：vProlog (Trealla)，而非 SWI-Prolog	38
15.3 啟用與執行	38
15.4 推導規則	39
15.5 帶版本號的活動資料檔案	39
15.6 編程介面與隨附檔案	39
15.7 相關實驗：氣候區先驗	40
15.7.1 超越加利福尼亞：轄區包與市場區域模組	40

1 引言

1.1 什麼是 earthUI？

earthUI 是 R 語言 earth 套件的圖形使用者介面，該套件實作了多元自適應迴歸樣條 (MARS, multivariate adaptive regression splines) (Friedman 1991)。建模引擎是 Stephen Milborrow 的

earth 套件 (Milborrow 2024), Milborrow (n.d.[a]) 與 Milborrow (n.d.[b]) 對其有詳細的說明文件。它以本機 Shiny 應用程式的形式執行——無需登入、無需伺服器、也沒有帳號。您從 R 中啟動它，匯入資料集 (CSV 或 Excel)，配置模型，然後互動式地進行擬合。

該應用程式提供了完整的工作流程：資料匯入、變數配置、模型擬合、診斷圖、變數重要性、模型方程，以及可下載的 HTML、Word 或 PDF 格式報告。

1.2 一個套件，三個常式

從當前版本起，earthUI 套件還包含了它的兩個後處理模組，這兩個模組在早期版本中是獨立的套件 glmnetUI 和 mgcvUI。安裝 earthUI 即安裝了全部三個常式；每個常式都作為獨立的本機 Shiny 應用程式執行：

- `earthUI::launch()` ——本文所描述的 earth (MARS) 應用程式，執行在連接埠 7878 上。`earth()` 是核心估值引擎。
- `earthUI::launch_glmnet()` ——彈性網後處理模組 (連接埠 7879)，它匯入 earth 模型的基函數，並在 lasso / 彈性網正則化下重新估計它們 (參見下文「glmnetUI Import」)。
- `earthUI::launch_mgcv()` ——GAM 後處理模組 (連接埠 7880)，它以 earth 模型的節點位置作為 mgcv 平滑項的起點 (參見下文「mgcvUI Auto-Export」)。

這三個常式共享同一個專案樹、同一個設定資料庫和同一套支撐基礎設施，因此在一個應用程式中開啟的專案可以立即供其他應用程式使用。glmnet 和 mgcv 常式是對 earth 模型價值結論的佐證，而不是相互競爭的工作流程：預期的使用模式是先擬合並完善 earth 模型，然後在需要佐證時針對同一專案執行這些模組。

1.3 三種用途模式

earthUI 以專案的形式組織工作——專案是有名稱的實體，將您的輸入檔案、輸出和設定集中儲存在一個固定位置。每個專案都帶有一個 **Purpose** (用途)，在建立專案時於 New Project wizard (新增專案精靈) 中選擇。共有三種模式：

- **General** (通用) ——適用於任何類型的總體或資料集的 Earth 迴歸。這是預設模式。它提供完整的 MARS 建模工作流程，不含任何特定領域的附加功能。
- **For Appraisal** (估價用途) ——為房地產估價量身打造的 Earth 迴歸。增加了針對單一房產估值的功能，包括估價對象 (subject) 處理、特殊欄指定、殘差約束法 (RCA)，以及銷售比較網格 (Sales Comparison Grid) 的生成。
- **Market Area Analysis** (市場區域分析) ——為市場區域研究量身打造的 Earth 迴歸。增加了在既定市場中分析成組房產的功能，而不是對單一估價對象進行估值。

在這三種模式下，核心建模引擎完全相同——您始終是在擬合一個 Earth (MARS) 模型。用途設定控制的是哪些附加工具和介面元素可用。

每個專案的用途是固定的。專案的用途——連同其國家、州、縣和城市——在專案建立時即被鎖定。切換到另一個專案會將 earthUI 重置為乾淨的預設狀態 (已匯入的資料、模型結果、頁籤、變數配置和 earth 參數全部清空)，然後恢復該專案針對其用途儲存的配置。

1.4 房地產專用功能

當選擇 **For Appraisal** 或 **Market Area Analysis** 時，earthUI 會啟用若干為房地產分析設計的功能：

- **特殊欄指定**——每個預測變數都可以被標記為某種特殊角色，例如 `contract_date`、`dom`、`concessions`、`latitude`、`longitude`、`living_area`、`lot_size`、`actual_age`、`effective_age`、`area`、`site_dimensions` 或 `display_only`。這些指定控制該欄在擬合、輸出和銷售網格生成過程中的處理方式。完整參考請參見第 6 章。
- **經緯度取整**——被指定為 `latitude` 或 `longitude` 的欄會自動取整到小數點後 3 位，以防止過擬合。

- **Sale Age 欄**——當某欄被指定為 `contract_date` 且提供了 Effective Date (生效日期) 時，earthUI 會計算一個 `sale_age` 欄 (銷售日期與生效日期之間的天數)，並將其作為預測變數代入。
- **RCA 計算 (僅限估價模式)**——在估價模式下，模型擬合完成後，earthUI 可以計算殘差約束法 (RCA) 輸出，生成逐個可比銷售案例的調整、淨 / 總調整匯總，以及估價對象的調整後銷售價格。
- **銷售比較網格 (僅限估價模式)**——生成一個多工作表的 Excel 活頁簿，其中包含可比案例調整、殘差特徵分解和調整後銷售價格的公式。參見第 11 章。

提示：*General* 用途模式適用於任何資料集——金融、科學、工程或房地產。估價模式和市場模式只是為房地產專業人士增加了便利功能。

1.5 安裝 earthUI 與 vProlog

earthUI 已發布在 CRAN 上，因此標準安裝只需在 R 中執行一行命令：

```
install.packages("earthUI")
```

本文描述的是當前版本。如果您閱讀本文時 CRAN 上仍是較舊的版本，可以從 GitHub 安裝開發版 (注意 `subdir` 參數——該套件位於儲存庫的 `pkg/` 資料夾中)：

```
remotes::install_github("wcraytor/earthUI", subdir = "pkg")
```

vProlog (可選)。實驗性的備註處理功能 (參見本文末尾附近的「Prolog Text Extraction」) 使用 vProlog 套件——一個自包含的嵌入式 Prolog 引擎。它被有意設計為非強制依賴：earthUI 在沒有它的情況下也能安裝並執行，Prolog 相關功能只是在其安裝之前保持停用狀態。vProlog 通過作者的 *r-universe* 儲存庫分發，該儲存庫為 Windows 和 macOS 提供預編譯的二進位檔案 (無需編譯器工具鏈)：

```
install.packages("vProlog",
  repos = c("https://wcraytor.r-universe.dev",
    getOption("repos")))
```

不需要系統級的 Prolog 安裝 (無論是 SWI-Prolog 還是其他)——Trealla 引擎已編譯進套件本身。本文中的其他所有內容在沒有 vProlog 的情況下均可正常使用。

1.6 快速入門

使用 earthUI 的步驟：

1. 安裝套件——按上文所述從 CRAN 安裝。
2. 啟動應用程式：在 R 中執行 `earthUI::launch()`，或從命令行執行 `Rscript -e 'earthUI::launch()'`。應用程式將在您的網頁瀏覽器中開啟，連接埠為 7878。您也可以直接在瀏覽器中造訪 `http://localhost:7878` 來開啟應用程式。(後處理模組以同樣的方式啟動：`launch_glmnet()` 使用連接埠 7879，`launch_mgcv()` 使用連接埠 7880。)
3. 建立或開啟專案——在側邊欄的第 1 節中進行。New Project wizard 會記錄專案名稱、Purpose (General、For Appraisal 或 Market Area Analysis)、位置 (國家、州、縣、城市)，以及可選的初始資料檔案，該檔案會被複製到專案的 `in/` 資料夾中。
4. 匯入資料——在第 2 節中進行，該節列出活動專案 `in/` 資料夾中的 CSV 和 Excel 檔案。您通過 Finder 或檔案總管手動新增的檔案，在點選 Refresh (重新整理) 後會顯示出來。
5. 配置變數——選擇目標變數和預測變數，設定資料類型，並指定任何特殊欄角色。
6. 設定 Earth 參數——`degree`、`nprune`、子集過濾器以及其他 `earth()` 參數。
7. 擬合模型——點選「Fit Earth Model」(擬合 Earth 模型)，並在主面板中查看結果。

8. 匯出——將預測結果下載為 Excel，生成 Quarto 報告包，或（在估價模式下）計算 RCA 調整和銷售比較網格。

配置通過第 3 節和第 4 節中的「Save current as default」（將當前設定儲存為預設）按鈕儲存在伺服器端的專案資料庫中，以專案和用途為鍵（參見第 3 章「Settings Persistence」）。

2 MLS 輸入資料要求

對於房地產估價和市場分析工作流程，您的輸入資料通常來自多重掛牌服務（MLS，Multiple Listing Service）的匯出檔案。本章描述預期的檔案結構以及 earthUI 可以使用的欄。

2.1 檔案格式與結構

earthUI 接受 CSV 和 Excel (.xlsx、.xls) 檔案。匯入時，欄名會被自動轉換為 snake_case——例如，「Living SqFt」變為 living_sqft，「Contract Date」變為 contract_date，「Sale Price」變為 sale_price。這種規範化確保了整個工作流程中欄引用的一致性。匯入時使用的 CSV 分隔符號和小數點符號由區域設定決定（參見第 3 章「Locale & Regional Settings」）。

您的資料檔案應為平面表格，每個房產佔一行，每個屬性佔一欄。檔案的第一列必須包含欄標題。

2.2 估價模式所需的欄

雖然 earthUI 可以處理任何欄集合，但完整的估價工作流程（RCA 調整 + 銷售比較網格）在您的 MLS 匯出檔案中包含以下欄時效果最佳：

欄	特殊類型	用途
Sale Price	（目標變數）	模型的反應變數
Contract Date	contract_date	用於計算 sale_age（距生效日期的天數）
Listing Date	listing_date	在沒有 DOM 欄時，與合約日期一起用於計算 DOM
Days on Market	dom	在市天數；顯示在銷售網格中
Concessions	concessions	銷售讓步；銷售網格中 Net SP = Sale Price – Concessions
Living Area (SF)	living_area	啟用每平方英尺殘差 (residual_sf、cqa_sf)
Lot Size	lot_size	在銷售網格中歸入「Site Size Dimensions」列
Site Dimensions	site_dimensions	在銷售網格中與地塊面積歸為一組
Latitude	latitude	取整到小數點後 3 位；用於鄰近度計算和 Location 組
Longitude	longitude	取整到小數點後 3 位；用於鄰近度計算和 Location 組
Area ID	area	在銷售網格中歸入「Loc: Long Lat Area」列
Actual Age	actual_age	在銷售網格中歸入「Actual Age Effective Age」列
Effective Age	effective_age	在銷售網格中與實際屋齡歸為一組
Parcel Number	（無）	顯示在銷售網格的 APN 列中
Listing ID	（無）	顯示在銷售網格的 APN 列中
Address	display_only	顯示在銷售網格的 Address 列中；不參與模型

試算表的欄名可以使用外語——「特殊」名稱使用英語，以便 R 程式能對它們進行特殊處理。除此之外，給定的欄名會出現在迴歸模型、圖表，以及（在進行估價時）中間銷售網格 (Intermediate Sales Grid) 中。

並非所有欄都是必需的。earthUI 會自行適應——如果某欄缺失，相應的功能會被直接省略。例如，如果沒有指定 concessions 欄，銷售網格中的 Net SP 列將顯示未扣除讓步的銷售價格。不過，對於房地產定價模型，強烈建議提供某些欄以獲得可接受的擬合效果：

1. **Sale Age**——合約銷售日期與估價或分析的生效日期之間的天數。如果使用多年的銷售歷史資料，尤其是超過 5 年的時段，sale_age 往往在銷售價格估計中扮演核心角色。事實上，它常常重要到沒有它 earthUI 根本無法給出任何模型。
2. **Living Area**——也稱作「Living Sqft」、「GLA」（gross living area，總居住面積）等。這是銷售價格的另一個主要決定因素。

3. **Total Bath Count**——全衛、四分之一衛、半衛和四分之三衛浴的總數。例如，兩個全衛加一個半衛的值為 2.5。
4. **Garage Bays** 或 **Garage Area**——車庫車位數量或車庫面積（平方英尺）。
5. **Lot Size**——房產的土地面積，通常以平方英尺或英畝為單位。
6. **Longitude**、**Latitude**，如果有的話再加上 **Area ID**。這些變數的調整將在銷售網格中合併為單一的 Location 調整項。

2.3 特殊欄命名約定

earthUI 通過欄的特殊類型指定而非欄名來識別欄。您可以在 MLS 匯出檔案中為欄取任何名稱——關鍵在於您在變數配置表（第 6 章）中指定了正確的特殊類型。

例如，您的 MLS 可能將居住面積匯出為「GLA」、「Living SqFt」、「liv_area」或「gross_living_area」。匯入之後（此時它已變為 snake_case），您只需在 Special dropdown（特殊類型下拉式選單）中將其指定為 living_area。此後，無論其原始名稱如何，earthUI 都會將其用於每平方英尺殘差計算和銷售網格分組。

2.4 資料品質與完整性

- **遺漏值 (NA)**：任何預測變數或目標變數欄中含有 NA 值的列會在擬合前自動刪除。na.action 在內部始終設定為 na.fail，NA 的刪除在呼叫 earth() 之前完成。
- **日期欄**：日期應採用 R 可解析的格式（例如 2025-06-15、06/15/2025、June 15, 2025）。當至少 80% 的值能夠成功解析時，earthUI 會自動檢測日期欄。兩位數年份（例如 06/15/25）會被讀取為最接近的匹配年份，因此近期的 MLS 匯出檔案可以正確解析；四位數年份始終按字面值處理。
- **數值欄**：銷售價格、居住面積、地塊面積、讓步及類似欄位必須是數值型。如果您的 MLS 匯出的價格帶有貨幣符號或千分位分隔符號（例如「\$350,000」或「350.000,00」），您可能需要在匯入前清理這些資料。
- **因子欄**：區域 ID、風格或狀況等分類變數應包含數量適中的唯一值。earthUI 會自動檢測因子，但您可以覆蓋該檢測結果。

提示：匯入後請查看預測變數表中的 T/F/NA 欄——它顯示每欄的 true/false/NA 計數。遺漏值超過 30% 的欄會以紅色標出。請考慮排除高遺漏率的欄，或在匯入前清理資料。

2.5 估價對象列的位置

在 **Appraisal**（估價）模式下，第 1 列必須是估價對象 (subject)。其餘所有列均為可比銷售案例。估價對象的銷售價格可以留空 (NA) 或設為任意值——無論如何，earthUI 在擬合時都將其視為 NA。

在 **Market Area Analysis** 模式下，將估價對象放在第 1 列是可選的，且不會自動排除任何列。若要將估價對象所在列（或任何其他列）排除在擬合之外，可在指定的權重欄中將其值設為零，或使用子集過濾器。

在 **General** 模式下，沒有特殊的列處理——所有列一視同仁。

3 通用模式

3.1 概述

通用模式 (General Purpose mode) 是啟動 earthUI 時的預設模式。它為任何資料集——而不僅僅是房地產資料——提供完整的多元自適應迴歸樣條 (MARS, multivariate adaptive regression splines) 建模工作流。您可以將 earthUI 用於科學資料、金融分析、工程研究，或任何希望探索變數間非線性關係與交互作用的迴歸問題。

在通用模式下，介面會省略房地產專用功能（特殊欄、銷售日齡、座標捨入、RCA）。側邊欄經過精簡，聚焦於變數選擇、參數配置、模型擬合與匯出。

3.2 側邊欄工作流

側邊欄按編號組織為可折疊的分區，引導您從專案選擇一路走到匯出：

1. Project (專案) ——選擇一個現有專案，或通過 New Project (新增專案) 精靈建立新專案；該精靈會記錄用途、國家和行政區位置（建立後全部鎖定）。當前活動專案決定了輸入檔案的讀取位置以及所有輸出的寫入位置——不存在單獨的輸出資料夾欄位。

2. Import Data (匯入資料) ——從活動專案的 in/ 資料夾中選擇 CSV 或 Excel 檔案；Refresh (重新整理) 連結會重新掃描該資料夾，以發現手動新增的檔案。對於包含多個工作表的 Excel 檔案，會出現工作表選擇器。欄名會被自動轉換為 snake_case。

3. Variable Configuration (變數配置) ——目標變數選擇器（支援多個目標變數）、Effective Date (生效日期)，以及帶有 Include、Factor、Linear 和 Latent 核取方塊的預測變數表。詳見第 6 章。

4. Earth Call Parameters (Earth 呼叫參數) ——earth() 函數的全部參數：degree、penalty、nk、剪枝方法、交叉驗證等。完整的參數參考見第 7 章。

Execute Prolog Processing (執行 Prolog 處理) ——位於第 4 與第 5 分區之間的一個可選、無編號分區，除非在 Settings (設定) 中啟用了實驗性的 Prolog 功能，否則處於停用狀態。見第 15 章。

5. Fit Earth Model (擬合 Earth 模型) ——一個隨機種子欄位和非同步執行模型的按鈕。擬合的詳細說明見第 8 章。

6. Download Estimated Target Variable(s) & Residuals (下載估計的目標變數與殘差) ——（在估價模式和市場模式下標題為「Download Estimated Sale Prices & Residuals」，即下載估計售價與殘差）將預測值、殘差、CQA 評分以及各 g 函數的貢獻匯出為一個 Excel 檔案。見第 9 章。

7–8. RCA Adjustments and Sales Grid (RCA 調整與銷售網格) ——僅限估價模式。見第 10 章和第 11 章。

9. Generate Quarto Report (生成 Quarto 報告) ——將一個自包含的 Quarto 報告包 (.qmd 原始檔及資源檔案) 寫入專案的輸出資料夾。在非估價模式下編號為 7。見第 12 章。

10. Convert Quarto Report (轉換 Quarto 報告) ——將任意 .qmd 檔案渲染為 HTML、Word 或 PDF。見第 12 章。

3.3 主面板頁籤

擬合完成後，主面板提供十個頁籤：

頁籤	內容
Data (資料)	匯入資料的預覽。在估價模式下，拆分為「Subject Property」(估價對象)和「Comparable Sales」(可比銷售案例)兩個表格。
Equation (方程)	以 LaTeX/MathJax 顯示的擬合模型方程，展示每個基函數及其係數。
Summary (摘要)	關鍵指標 (R^2 、GRSq、GCV、RSS、 $CV R^2$) 和係數表。
Variable Importance (變數重要性)	預測變數重要性得分的長條圖和排名表。
Contribution (貢獻)	g 函數貢獻圖。從下拉式選單中選擇一個 g 函數；單變數預測變數顯示折線圖，雙變數交互項 ($\text{degree} \geq 2$) 顯示 3D 透視圖和等高線圖。
Correlation (相關性)	預測變數相關性的熱圖。
Diagnostics (診斷)	三幅圖：殘差 vs 擬合值、常態 Q-Q 圖，以及實際值 vs 預測值。
RCA Adjustments (RCA 調整)	(僅估價模式) 各可比銷售案例的殘差調整百分比 (Residual Adj %)、淨調整百分比 (Net Adj %) 和總調整百分比 (Gross Adj %) 的直方圖，在第 7 步 (Calculate RCA Adjustments) 完成後填充。
Residual Contribution (殘差貢獻)	通過將模型殘差對模型從未使用過的欄進行聯合迴歸，估計稀有特徵或被排除特徵的市場貢獻 (見 subsection 4.4)。
ANOVA	ANOVA 分解表，顯示每個基函數對 RSS 的貢獻。
Earth Output (Earth 輸出)	來自 'earth::summary()' 的原始文字輸出，包括剪枝過程的詳細資訊。

3.4 設定持久化

earthUI 將您的配置儲存在伺服器端的專案資料庫中 (位於 regProj 根目錄下的 `projects.sqlite`)，以專案和用途為鍵。儲存是按鈕驅動的：第 3 分區中的「Save current as default」(將當前設定儲存為預設) 按鈕儲存目標變數、生效日期和預測變數配置，而第 4 分區中的按鈕儲存 earth 呼叫參數和 Allowed Interactions (允許的交互項) 矩陣。每個按鈕只更新自己負責的那一半，因此它們絕不會相互覆蓋，並且擬合不會自動儲存——您可以自由試驗而不必擔心覆蓋已儲存的預設設定。一個專案的多個輸入檔案 (一個小型測試抽樣、完整資料集) 共享同一份配置；而且由於設定存放在專案目錄樹中，當您同步、備份或共享 regProj 資料夾時，設定會隨之遷移。

3.5 深色模式

點選右上角的月亮/太陽圖示可在淺色與深色主題之間切換。主題偏好儲存在本機儲存中，並在工作階段之間持續生效。

3.6 語言環境與區域設定

earthUI 通過基於國家的語言環境系統支援國際化的數字、日期和 CSV 格式約定。頂部選單欄 **Settings** (設定，齒輪圖示) 選單中的 **Country** (國家) 下拉式選單可為 31 個受支援的國家選擇預設。每個預設配置以下內容：

- **CSV 分隔符號**——美國/英國/日本使用逗號 (,)，歐洲大部分國家使用分號 (;)，因為那裡逗號被用作小數點。
- **小數點符號**——句點 (.) 或逗號 (,)。
- **千分位分隔符號**——逗號 (美國/英國/日本)、句點 (德國/義大利/西班牙)、空格 (芬蘭/法國/波蘭/波羅的海國家/烏克蘭/俄羅斯) 或撇號 (瑞士)。

- 日期格式——MM/DD/YYYY（美國）、DD/MM/YYYY（歐洲大部分國家）或 YYYY-MM-DD（瑞典/立陶宛/日本/加拿大）。
- 紙張尺寸——Letter（美國/加拿大/墨西哥）或 A4（其他所有國家）。

3.6.1 支援的國家

國家	CSV 分隔符號	小數點	千分位分隔符號	日期	紙張
美國、加拿大 (csv)	,	.	,	MDY/YMD	Letter
英國、愛爾蘭、澳洲、紐西蘭	,	.	,	DMY	A4
德國、奧地利、比利時、荷蘭、義大利、西班牙、葡萄牙、丹麥、巴西	;	,	.	DMY	A4
法國、波蘭、捷克、挪威、拉脫維亞、愛沙尼亞	;	,	(空格)	DMY	A4
芬蘭、瑞典	;	,	(空格)	DMY/YMD	A4
瑞士	;	,	,	DMY	A4
烏克蘭、俄羅斯	;	,	(空格)	DMY	A4
土耳其	;	,	.	DMY	A4
日本、韓國	,	.	,	YMD	A4
墨西哥	,	.	,	DMY	Letter
立陶宛	;	,	(空格)	YMD	A4

3.6.2 紙張覆蓋設定

在國家選擇器下方，**Paper**（紙張）下拉式選單（Letter 或 A4）允許您在不切換國家的情況下覆蓋報告頁面尺寸。CSV 分隔符號、小數點符號和日期格式順序仍遵循所選國家的預設；當您更改國家時，紙張覆蓋設定會重置為該國家的預設值。

3.6.3 儲存預設設定

點選 Settings 選單中的 **Save**（儲存）可全域儲存您的語言環境偏好。這些預設設定適用於今後所有工作階段，無論您開啟哪個專案。專案級設定（目標變數、預測變數、參數）單獨儲存在專案資料庫中；語言環境預設設定則通過使用者級 SQLite 資料庫在所有專案間持續生效。這種兩級設計面向審計事務所等需要處理來自多個國家資料的機構——將您最常用的國家設為預設，需要時再進行調整。

提示：圖表座標軸和斜率標籤上的數字格式會自動適配您的語言環境。德國語言環境使用句點作為千分位分隔符號（200.000），芬蘭使用空格（200 000），瑞士使用撇號（200'000）。不顯示任何貨幣符號——earthUI 與貨幣無關。

4 估價模式

當您選擇 **For Appraisal**（用於估價）作為用途（Purpose）時，earthUI 會將自身配置為面向單一物業估值。第 3 章描述的所有功能仍然可用；本章僅涵蓋估價模式特有的附加功能。

4.1 估價對象列的處理

在估價模式下，資料集的第 1 列是估價對象（**subject**）物業，其餘所有列均為可比銷售案例。您的輸入檔案必須按此方式組織（見第 2 章）。估價對象的售價可以留空或設為任意值——earthUI 在擬合時會自動將其視為 NA。

匯入後，Data 頁籤會拆分為兩個部分：**Subject Property**（估價對象物業，第 1 列）和 **Comparable Sales**（可比銷售案例，第 2 列及之後）。第 1 列始終被排除在模型擬合之外——通知訊息「Skipping row 1 (subject). Fitting on N rows.」（跳過第 1 列（估價對象）。在 N 列上擬合。）會對此加以確認。擬合後，模型仍會為估價對象列生成預測值，在輸出中顯示為 `est_<target>`。

4.2 生效日期與銷售日齡

在每種用途模式下，Variable Configuration 分區中都會出現一個 **Effective Date**（生效日期）欄位（預設為當天日期）——通用分析和市場分析同樣有生效日期，而不僅僅是估價——但它只在估價模式和市場模式下才會被使用。在這兩種模式下，如果您在 Special（特殊欄）下拉式選單中將某欄指定為 `contract_date`，earthUI 會計算一個 `sale_age` 欄——即每筆銷售的合約日期與生效日期之間的整數天數。該欄會取代原始日期欄作為預測變數。

在指定合約日期後第一次點選 Fit 時，earthUI 會建立 `sale_age` 欄，並通知您再次點選 Fit 以將其納入模型。

4.3 特殊欄指定

在估價模式和市場模式下，Variable Configuration 表中的每個預測變數旁都會出現一個 **Special**（特殊）下拉式選單。特殊類型及其作用的完整列表如下：

特殊類型	作用
no	預設——無特殊處理
contract_date	觸發基於生效日期的 <code>sale_age</code> 自動計算
listing_date	在不存在 <code>dom</code> 欄時，與合約日期一起用於計算 DOM
dom	標示「在售天數」(Days on Market) 欄；顯示在銷售網格 (Sales Grid) 中
concessions	標示銷售讓步 (concessions)；用於銷售網格中的淨售價公式 (售價 - 讓步)
latitude	數值捨入到小數點後 3 位；用於 Haversine 鄰近度計算及銷售網格中的 Location 分組
longitude	數值捨入到小數點後 3 位；用於 Haversine 鄰近度計算及銷售網格中的 Location 分組
area	在銷售網格中歸入「Loc: Long Lat Area」列
living_area	在下載輸出中啟用按平方英尺計算的殘差 (<code>residual_sf</code> 、 <code>cqa_sf</code>)
lot_size	在銷售網格中歸入「Site Size Dimensions」列
site_dimensions	在銷售網格中與地塊面積歸為一組
actual_age	在銷售網格中歸入「Actual Age Effective Age」列
effective_age	在銷售網格中與實際屋齡歸為一組
display_only	該欄包含在匯出中，但排除在模型擬合之外（允許多欄）
remarks	（實驗性）自由文字備註欄，由可選的 Prolog 步驟解析為特徵欄（允許多欄）；見第 15 章
list	（實驗性）逗號分隔的欄，由可選的 Prolog 步驟拆分為獨熱 (one-hot) TRUE/FALSE 欄（允許多欄）；見第 15 章

每種特殊類型只允許指定一欄 (`display_only`、`remarks` 和 `list` 除外)。將某個特殊類型指定給第二欄時，會自動從第一欄上清除該指定。變數名旁邊會出現一個藍色小徽章，顯示其被指定的特殊類型。

4.4 殘差貢獻（稀有特徵估計）

對於因過於稀有而無法進入迴歸通道的特徵——例如僅在少數幾筆銷售中出現的工作間 (workshop)——它們屬於 RCA 殘差層。**Residual Contribution**（殘差貢獻）頁籤通過普通最小二乘法，將擬合模型的殘差對模型從未使用過的欄進行迴歸，以估計這些特徵的市場貢獻。選擇多欄時會對它們進行聯合估計，因此相關的配套設施（例如 `outbuilding_sqft` 加上 `pr_has_workshop`）會分攤貢獻，而不會重複計算。二元標誌欄產生一次性總額估計，並列出每筆帶該標誌的銷售的殘差作為證據；數值欄產生每單位貢獻；報告的標準誤為一個 SE（通常以 $t \geq 2$ 作為顯著性判斷標準）。支持觀測數少於十個的項會在生成的工作底稿註記中自動獲得 CAUTION（警示）標記；該註記以一個段落記錄方法、估計值和注意事項，供報告使用。只有模型實際使用過的預測變數才會被排除在候選列表之外——一個曾提供給擬合但從未被前向過程選中的稀有標誌欄仍然符合條件，因為模型沒有吸收它的任何價值。

4.5 RCA 調整概述

Calculate RCA Adjustments & Download（計算 RCA 調整並下載）按鈕（側邊欄第 7 分區，僅在估價模式下擬合完成後可見）計算每個可比銷售案例相對於估價對象的市場推導調整值。完整的 RCA 工作流在第 10 章中描述。

計算完 RCA 調整後，**Generate Sales Grid & Download**（生成銷售網格並下載）按鈕（側邊欄第 8 分區）變為可用。銷售比較網格 (Sales Comparison Grid) 工作流在第 11 章中描述。

提示：您為估價對象指定的 *CQA* 評分決定了殘差分布中有多大比例歸屬於估價對象。評分 5.00 會將估價對象置於可比銷售案例的中位數位置。

5 市場區域分析模式

當您選擇 **Market Area Analysis**（市場區域分析）作為用途時，earthUI 提供與估價模式相同的房地產專用功能（特殊欄、銷售日齡、座標捨入），但其定位是分析一組物業，而不是對單一估價對象進行估值。

5.1 與估價模式的區別

- 不自動排除估價對象——與估價模式不同，市場模式不會自動將第 1 列排除在擬合之外。所有列都會被納入；若要將估價對象列排除在擬合之外，可在指定的權重欄中為其賦零值（如示範資料集所示），或使用子集過濾器。
- 沒有 **RCA** 或銷售網格分區——「Calculate RCA Adjustments & Download」和「Generate Sales Grid & Download」步驟不可用。市場模式專注於模型擬合與輸出，而非逐個可比案例的調整。
- 側邊欄編號——由於沒有 RCA 和銷售網格步驟，Generate Quarto Report 分區的編號為 7 而非 9（Convert Quarto Report 分區仍為 10）。

5.2 何時使用市場模式

以下情形適合使用市場區域分析模式：

- 為某個社區或市場區域構建迴歸模型，以理解價值驅動因素
- 分析建築面積、屋齡、地塊面積和區位等變數如何影響一組物業的售價
- 為市場狀況分析或社區界定準備支持材料
- 使用的資料集第 1 列包含一個估價對象物業，但您希望可以自由選擇將其納入或排除在擬合之外

提示：市場模式也適用於一般的房地產迴歸場景——您需要特殊欄功能（銷售日齡、座標捨入），但不需要 **RCA** 調整工作流。

6 變數選擇

側邊欄第 3 節——**Variable Configuration**（變數配置）——用於選擇哪些欄參與模型，以及它們如何被處理。

6.1 目標變數

第 3 節頂部的 **Target (response) variable(s)**（目標（反應）變數）下拉式選單列出資料集中的每一欄。選擇一欄可擬合標準的單反應模型；選擇多欄則擬合多反應模型。選擇多個目標變數時，模型使用 `earth(cbind(y1, y2, ...) ~ .)` 同時擬合所有反應。

被選為目標變數的欄會自動從預測變數列表中排除。

6.2 預測變數表

在目標變數選擇器下方，一個表格列出其餘每一欄，包含以下欄位：

欄	說明
Variable （變數）	欄名（若被截斷，完整名稱顯示在工具提示中），前面帶有一個 $\times/+$ 切換按鈕（見下文）。在估價/市場模式下，藍色徽章顯示所分配的特殊類型。
Type （類型）	資料類型下拉式選單： <code>numeric</code> 、 <code>integer</code> 、 <code>character</code> 、 <code>logical</code> 、 <code>factor</code> 、 <code>Date</code> 、 <code>POSIXct</code>
Include （納入）	核取方塊——將此欄作為預測變數納入模型
Factor （因子）	核取方塊——將此欄視為分類變數
Linear （線性）	核取方塊——強制僅以線性方式進入模型（不生成鉸鏈函數）
Latent （潛變數）	核取方塊——將此欄完全排除在模型之外（既不進入 <code>earth</code> ，也不包含在匯出到 <code>glmnetUI</code> / <code>mgcvUI</code> 的資料中），保留用於潛變數分析
Special （特殊類型）	下拉式選單（僅估價/市場模式）——見下文「特殊欄類型參考」
T/F/NA	每欄三個計數：對邏輯欄為 <code>TRUE/FALSE/NA</code> ；對數值欄為 $\neq 0/=0/NA$ 。當遺漏值超過 30% 時以紅色顯示

表格上方的提示列解釋了這些縮寫：「Type = 欄資料類型，Inc = 作為預測變數納入，Factor = 視為分類變數，Linear = 僅線性（無鉸鏈），Special = 欄角色（例如 `contract_date`）。」

軟性停用而非刪除。每個變數名旁的 $\times$ 切換按鈕會將該列變灰，並使該欄不參與模型，但不會將其從資料中刪除；按鈕隨之變為綠色的 $+$ ，可重新啟用該欄。欄永遠不會被銷毀。

6.3 資料類型檢測與覆蓋

`earthUI` 在匯入時自動檢測資料類型。數值、整數、邏輯、因子和日期欄都會被識別。看起來像日期的字元欄（至少 80% 的值能按常見日期格式解析）會被歸類為 `Date`。

你可以通過更改 **Type** 下拉式選單來覆蓋任何檢測結果。當你將某欄改為 `character` 或 `factor` 時，Factor 核取方塊會自動勾選。更改類型會影響該欄傳遞給 `earth()` 函數的方式。

6.4 Factor 與 Linear 標誌

- **Factor**——勾選後，該欄被視為分類變數。這適用於 `style`、`area_id` 或 `grade` 等表示離散組別而非連續測量值的欄。因子欄以指示（虛擬）變數的形式進入模型。
- **Linear**——勾選後，該欄被強制以線性方式進入模型——不會為其建立鉸鏈（分段線性）函數。當你確定某個變數與目標變數之間存在嚴格的線性關係時，此選項很有用。

6.5 特殊欄類型參考

在估價和市場模式下，**Special** 下拉式選單提供以下選項。每種類型最多只能分配給一欄（`display_only`、`remarks` 和 `list` 除外，它們允許多欄）：

日期與時間類型：

- `contract_date`——觸發自動計算 `sale_age`。原始日期欄被替換為一個整數欄，度量銷售日期與 Effective Date（評估基準日）之間的天數。
- `listing_date`——當未指定明確的 `dom` 欄時，用作計算在市天數（DOM = 合約日期 - 掛牌日期）的後備來源。
- `dom`——標示在市天數（Days on Market）欄。顯示在 Sales Grid（銷售比較網格）的 APN 列和 Date of Sale 列中。

貨幣類型：

- `concessions`——標示銷售讓步（賣方信貸、買方激勵等）。在 Sales Grid 中用於計算淨售價： $\text{Net SP} = \text{售價} - \text{讓步}$ 。

面積與位置類型：

- `latitude`——值自動四捨五入到小數點後 3 位以防止過擬合。用於 Haversine 鄰近度計算（估價對象（subject）到每個可比銷售案例的距離），並歸入 Sales Grid 的 Location 列。
- `longitude`——與 `latitude` 相同的捨入處理。用於鄰近度計算和 Location 組。
- `area`——通常為社區或區域識別碼。與 `latitude`、`longitude` 一起歸入 Sales Grid 的 “Loc: Long | Lat | Area” 列。
- `living_area`——啟用下載輸出中的每平方英尺殘差計算（`residual_sf` 和 `cqa_sf`）。
- `lot_size`——歸入 Sales Grid 的 “Site Size | Dimensions” 列。
- `site_dimensions`——與 `lot_size` 一起歸入 Sales Grid（例如 “75x120”）。

屋齡類型：

- `actual_age`——歸入 Sales Grid 的 “Actual Age | Effective Age” 列。
- `effective_age`——與 `actual_age` 一起歸入 Sales Grid。

顯示類型：

- `display_only`——該欄包含在 Excel 匯出中，但完全不參與模型擬合。適用於地址欄位、MLS 編號、宗地編號或其他不應作為預測變數的參考資料。允許多欄使用此指定。

文字提取類型（實驗性）：

- `remarks`——一個自由文字欄（例如 `public_remarks`），由可選的 Prolog 步驟解析為特徵欄。在 Settings（設定）中啟用該功能之前不起任何作用。允許多欄。參見第 15 章。
- `list`——一個逗號分隔的欄（例如 `existing_structures`），由同一步驟拆分為獨熱編碼的 `<col>_<value>` TRUE/FALSE 欄。允許多欄。參見第 15 章。

6.6 多個目標變數

選擇多個目標變數會擬合一個多反應 Earth 模型。選擇多個目標變數時：

- 模型同時預測所有反應，各目標變數共享基函數
- **wp (response weights)**（反應權重）按鈕被啟用，允許為每個目標變數分配數值權重
- 變異數模型（`varmod.method`）被停用（多反應模型不支援）
- 結果頁籤會顯示每個反應的指標、方程和診斷圖

提示：指定為 `display_only` 的欄保留在資料集中並出現在 Excel 匯出中，但完全不出現在預測變數表中。適用於 ID 欄、地址或其他參考欄位。

7 參數選擇

側邊欄第 4 節——**Earth Call Parameters** (Earth 呼叫參數)——提供對 `earth()` 函數所接受的全部參數的存取。每個參數都帶有一個藍色說明圖示 (?) 及工具提示說明。參數按可折疊的小節組織。

7.1 基本參數

參數	預設值	說明
subset	(空)	列過濾表達式。見下文「子集過濾」。
weights	NULL	用於個案(列)權重的欄選擇器。僅列出數值欄。
wp	NULL	多目標模型的反應權重。按鈕會開啟一個對話方塊，為每個目標變數提供一個數值輸入(預設各為 1.0)。
keepxy	關閉	在模型對象中保留 <code>x</code> 、 <code>y</code> 、 <code>subset</code> 和 <code>weights</code> 。
trace glm	0 無	擬合輸出的追蹤級別 (0 到 5)。可選的 GLM 族： <code>gaussian</code> 、 <code>binomial</code> 或 <code>poisson</code> 。
degree	1	最大交互階數。將 <code>degree</code> 設為 ≥ 2 會自動啟用交叉驗證，並顯示 Allowed Interactions (允許的交互) 矩陣。
penalty	2	每個節點的 GCV 懲罰。值越高，模型越簡單。

7.2 前向過程

參數	預設值	說明
nk	自動	剪枝前的最大模型項數。自動填充為所載入資料的推薦值(見下文「推薦參數值」)。
thresh	0.001	前向步驟閾值。值越小，允許的模型項越多。
minspan	自動	節點之間的最小間隔。自動填充為推薦值。負值設定每個預測變數的最大節點數。
endspan	自動	端部間隔——節點到資料邊緣的最小距離。自動填充為推薦值。
newvar.penalty	0	引入新變數的懲罰(鼓勵重用已有的預測變數)。
fast.k	20	快速 MARS 演算法中考慮的父項數量。
fast.beta	1	控制快速 MARS 的老化因子。

7.3 允許的交互

當 **degree** 設為 2 或更高時，基本參數下方會出現一個交互矩陣。它是一個上三角形的核取方塊網格，每個預測變數對對應一個核取方塊。勾選表示允許這兩個預測變數交互；取消勾選則禁止該特定交互。

點選預測變數名稱（列或欄標籤）可切換該變數的所有交互。頂部的 **Allow All**（全部允許）和 **Clear All**（全部清除）核取方塊提供批量控制。矩陣使用固定表頭，滾動時列、欄標籤保持可見。

一條資訊提示會提醒你：「交互項會增加過擬合的風險。交叉驗證已啟用（10 折）。」

7.4 剪枝

參數	預設值	說明
pmethod	backward	剪枝方法：backward、none、exhaustive、forward、seqrep 或 cv。
nprune	NULL	剪枝後的最大模型項數。留空則自動選擇。

7.5 交叉驗證

參數	預設值	說明
nfold	10	交叉驗證折數。設為 0 可停用交叉驗證。
ncross	20	交叉驗證重複次數。
stratify	開啟	對交叉驗證樣本進行分層，使每一折具有相似的反應分布。

當 $\text{degree} \geq 2$ 時，交叉驗證會自動啟用（**nfold** 設為 10），以幫助防範交互項帶來的過擬合。

7.6 子集過濾

subset 文字輸入框接受一個 R 表達式，用於篩選哪些列參與模型擬合。你可以直接輸入表達式（例如 `sale_age < 365 & area_id == 460`），也可以使用 **Build filter...**（構建過濾器...）按鈕以視覺化方式構建。

Build Filter Dialog（構建過濾器對話方塊）——點選“Build filter...”開啟一個引導式對話方塊。每個條件列包含一個欄下拉式選單、運算子（<、>、<=、>=、==、!=），以及一個根據欄類型自適應的值輸入框：數值欄為數字輸入，日期欄為日期選擇器，字元/因子欄為唯一值下拉式選單。各條件通過 AND（&）或 OR（|）連接運算子相連。底部的預覽會顯示表達式以及匹配的列數。點選 **Apply**（應用）將表達式插入文字輸入框。

日期欄在手動表達式中必須使用 `as.Date("...")` 或 `as.POSIXct("...")` 包裝。Build Filter 對話方塊會自動處理這一點。

子集過濾是非破壞性的——被排除的列仍保留在資料集中，並在 Excel 匯出中獲得預測值。

提示：將 OR 條件與 AND 組合時，請使用括號進行分組。若不加括號， $A \& B / C$ 會被解釋為 $(A \& B) / C$ ，這可能不符合你的本意。

7.7 推薦參數值

earthUI 會在第 4 節的關鍵參數下方顯示推薦值。這些推薦值會根據擬合列數 (n) 和所選預測變數數 (p) 動態更新。這些公式來源於 Friedman 的 MARS 論文、earth 的內部演算法以及經驗測試。**nk**、**minspan** 和 **endspan** 欄位會在資料載入時預填充其推薦值；手動編輯的值會被保留。

7.7.1 前向過程參數

nk (剪枝前的最大模型項數)：

$$\text{nk} = \min(100, \max(21, 2p + 1, \lfloor n/10 \rfloor))$$

n	$p=5$	$p=10$
30	21	21
50	21	21
100	21	21
200	21	21
500	50	50
1000	100	100
1500	100	100

Earth 的預設值 $\min(200, \max(20, 2p)) + 1$ 未考慮資料集規模，對大資料集而言可能過小，會在前向過程充分探索所有預測變數之前就對其加以限制。 $\lfloor n/10 \rfloor$ 一項允許大約每 10 個觀測值對應一個模型項——這是避免過度參數化的標準經驗法則。100 的上限可避免收益遞減。

minspan (節點之間的最小觀測數)：

$$\text{minspan} = \min(16, \lfloor 5 + n/50 \rfloor)$$

endspan (距資料邊界的最小觀測數)：

$$\text{endspan} = \min(16, \lfloor 5 + n/28 \rfloor)$$

n	minspan	endspan
30	5	6
50	6	6
100	7	8
200	9	12
300	11	15
500	15	16
1500	16	16

Earth 自動計算的 minspan 使用 Friedman 的公式 43： $\lfloor (-\ln(-\ln 0.95) + \ln(p \cdot n)) / (2.5 \ln 2) \rfloor$ ，它按 $\ln(np)$ 縮放，對大資料集增長過慢。在 $n=1500$ 時它僅得出 ≈ 7 ，使每個連續預測變數約有 245 個候選節點位置——太多了，會讓前向過程去擬合噪聲。推薦公式的目標是在大 n 時每個預測變數約有 100 個候選位置，而在小 n 時（此時 earth 的自動計算經過了充分檢驗）沿用其自動計算結果。endspan 設定得略大於 minspan，以提供額外的邊界保護，這在資料尾部稀薄時（房地產資料中很常見）很有幫助。

penalty (每個節點的 GCV 懲罰)：

$$\text{penalty} = \begin{cases} 3 & \text{if degree} > 1 \\ 2 & \text{if degree} = 1 \end{cases}$$

這正是 earth 自身的預設值——加性模型為 2，交互模型為 3（更高的懲罰補償了更大的搜索空間）。

newvar.penalty（引入新預測變數的懲罰）：

推薦值：當預測變數之間存在相關性時（例如居住面積與臥室數、浴室數、地塊面積相關）取 0.1。這會使前向過程傾向於重用模型中已有的預測變數，而不是引入相關的替代變數。其機制是：每個新預測變數的 RSS 改進量會乘以 $1/(1 + \text{penalty})$ 。當 **newvar.penalty** = 0.1 時，新變數的 RSS 改進必須比在已有變數上再加一個節點至少高出 10%。這會產生預測變數更少、更簡單的模型，且不影響最終的係數估計（懲罰在選擇完成後即被移除）。若預測變數之間不相關，則保持為 0。注意，earth 不支援在使用個案權重的同時設定非零的 **newvar.penalty**——一旦指定了權重欄，earthUI 就會停用該輸入（強制為 0）。

7.7.2 剪枝參數

pmethod：推薦值：**backward**。預設的基於 GCV 的後向剪枝是確定性的——相同的資料總是產生相同的模型，不依賴任何隨機種子。這對可重現性很重要，在估價工作中尤其如此。cv 方法使用交叉驗證進行剪枝，會引入對種子的依賴。

nprune：推薦值：留空（NULL）。讓 GCV 選擇最優的模型項數。設定 **nprune** 會施加一個硬性上限，凌駕於 GCV 的判斷之上。

7.7.3 交叉驗證參數

nfold（交叉驗證折數）：按樣本量區間推薦：

n （擬合列數）	nfold
0–299	5
300–399	6
400–499	7
500–599	8
600–699	9
700+	10

折數越多，每一折的訓練資料越多（估計偏差更小），但留出的測試折會隨之縮小。該表隨著樣本增大逐級升至標準的 10 折，同時保持每個測試折足夠大（約 50 列以上）以獲得穩定的單折指標；對小樣本則回退到 5 折——此時 CVR^2 估計仍有噪聲，因此應提高 **ncross**（例如 15–20）以平均掉變異數。上限 10 反映的是：每增加一折都是一次完整的 earth 擬合，而診斷收益微乎其微。當 **pmethod** = "backward" 時，交叉驗證不影響模型——它只計算診斷性的 CVR^2 ，並為變異數模型提供殘差。

ncross（交叉驗證重複次數）：按樣本量區間推薦：

n （擬合列數）	ncross
0–199	20
200–399	16
400–699	12
700–1199	10
1200–2999	7
3000+	5

重複可以平均掉折劃分中的運氣成分，從而穩定 CVR^2 ；樣本越大所需重複越少，因為單次交叉驗證本身已經更穩定。重複交叉驗證的殘差還會輸入變異數模型。

7.7.4 變異數模型

varmod.method：推薦值：**lm**——Milborrow 的預設值：用於預測區間的穩健線性變異數模型（離散度是從帶噪聲的殘差中估計的，因此簡單模型的泛化能力更好）。如果 **lm** 在你的資料上無法收斂，可嘗試 **rlm** 或 **const**（同變異數，總能收斂），或退回到 **earth**（也能處理非線性的離散度）。它要求 **nfold** > 1 且 **ncross** > 1。

7.7.5 Earth 輸出檔案

每次成功擬合後，**earthUI** 會將 **<filename>_earth_output_<timestamp>.txt** 寫入輸出資料夾。該檔案包含模型項、匯總統計量（ R^2 、 $GRSq$ 、 $CVRSq$ ）、變異數模型詳情和追蹤日誌。每次擬合建立一個檔案，為比較不同參數配置提供累積記錄。

7.8 設定預設值

第 4 節頂部的選項按鈕控制載入哪一組預設值：

- **Use last settings for input file**（使用該輸入檔案的上次設定）——恢復你上次針對這個特定檔案使用的設定
- **Use default settings**（使用預設設定）——應用你儲存的自訂預設值
- **Earth defaults**（Earth 預設值）——將所有參數重置為原廠值

點選第 4 節底部的 **Save current as default**（將當前設定儲存為預設）以儲存當前的 **earth** 參數和 **Allowed Interactions** 矩陣；第 3 節中對應的按鈕儲存目標變數、評估基準日和預測變數配置。兩者都按專案和用途儲存在專案資料庫中。

8 擬合 Earth 模型

8.1 Fit 按鈕

側邊欄第 5 節包含 **Fit Earth Model** (擬合 Earth 模型) 按鈕，以及一個 **Random seed** (隨機種子) 欄位。該種子使交叉驗證折的分配具有可重現性——下拉列表可調出最近使用過的 10 個種子，擬合完成後所用的種子會記錄在 Earth Output (Earth 輸出) 頁籤中。在擬合之前，earthUI 會驗證您的配置——必須選定一個目標變數，並且至少包含一個預測變數。在估價/市場模式下，緯度和經度欄會被四捨五入，銷售時距則根據估價基準日和合約日期 (若已指定) 計算得出。

8.2 擬合模態框與追蹤輸出

當您點選 Fit 時，會出現一個深色模態覆蓋層，其中包含：

- 標題欄——顯示「Fitting Earth Model」，附帶以秒計數的已用時間計時器，以及一個可取消背景擬合的 **Abort** (中止) 按鈕
- 追蹤日誌——一個可滾動的等寬文字區域，即時顯示 `earth()` 函數的輸出：資料集資訊、前向遍歷進度、交叉驗證折以及完成時間
- 顏色編碼——預設行為綠色，交叉驗證行為黃色，錯誤為紅色

earthUI 使用 `callr::r_bg()` 非同步擬合模型，將計算放在單獨的 R 列程中執行。這使應用程式在長時間擬合期間保持回應。一個 300 毫秒的輪詢觀察器從背景列程讀取輸出並將其流式寫入追蹤日誌。

擬合完成後，會出現「Done in X.Xs」訊息，模態框上會新增一個關閉按鈕 (X)，並移除背景遮罩，以便結果頁籤可以按需渲染。待可見輸出渲染完畢後，模態框會自動關閉——您也可以提前關閉它。

擬合成功時，Fit 按鈕上會附加一個綠色勾號。如果發生錯誤，錯誤訊息會顯示在追蹤日誌中，並彈出一條通知。

8.3 擬合日誌

每次擬合 (無論成功或失敗) 都會在您的輸出資料夾中寫入一個日誌檔：

- 檔案名：`<datafile>_earth_log_<YYYYMMDD_HHMMSS>.txt`
- 內容：時間戳記、擬合過程中的全部追蹤輸出，以及所有錯誤訊息
- 位置：當前活動專案的輸出資料夾

提示：如果未安裝 `callr`，earthUI 會回退到帶簡單進度條的同步擬合。安裝 `callr` 可獲得完整的追蹤輸出體驗。

9 下載資料

擬合完成後，側邊欄第 6 節——在通用模式下標題為「Download Estimated Target Variable(s) & Residuals」(下載估計的目標變數與殘差)，在估價和市場模式下標題為「Download Estimated Sale Prices & Residuals」(下載估計的銷售價格與殘差)——可下載一個包含預測值和診斷資訊的 Excel 檔案。該輸出用於第 7 步 (RCA)，為估價對象 (subject) 指定 CQA (狀況—品質—吸引力) 評分。輸出按 `residual_sf` 和 `cqa_sf` 排序，以幫助您評估估價對象在排名中的位置。如果模型品質良好，那麼各房產應按照未納入迴歸的殘差特徵，從吸引力最低到吸引力最高進行排序。中間值應約為 0，下半部分為負值，上半部分為正值。您應當在排名底部附近發現品質最差的住宅，即「待修房」(fixers)，而最好的住宅位於頂部。通常會有一些異常情形的例外，例如法拍、虧本急售 (short sale)、遺囑認證 (與繼承相關的) 銷售，以及因工作變動或其他原因需要的快速出售。對異常案例的調查通常會揭示出價格異常的相關原因。

按鈕標籤會根據用途模式而變化：

- 通用/市場：**Download Output (Excel)** (下載輸出 (Excel))
- 估價：**Download Intermediate Output (Excel)** (下載中間輸出 (Excel))

檔案名格式為 `<datafile>_modified_<YYYYMMDD_HHMMSS>.xlsx`。

9.1 輸出欄

對於每個目標變數，將追加以下各欄：

欄名	說明
<code>est_<target></code>	模型預測值，例如 <code>est_sale_price</code> (保留 1 位小數)
<code>residual</code>	實際值減去預測值 (1 位小數)
<code>cqa</code>	狀況—品質—吸引力評分 (2 位小數，範圍 0–10)
<code>residual_sf</code>	殘差除以居住面積 (若已指定，1 位小數)
<code>cqa_sf</code>	通過 <code>residual_sf</code> 排名計算得出的 CQA (2 位小數)
<code><variable>_contribution</code>	每個 g 函數的貢獻值 (1 位小數，每個 g 函數一欄)
<code>basis</code>	截距貢獻值，對所有房產相同 (1 位小數)
<code>calc_residual</code>	實際值減去 (basis + 所有貢獻值) ——驗證欄 (1 位小數)

對於多目標模型，會新增 `_<i>` 後綴以區分各目標對應的欄。

9.2 欄排序與 Excel 格式

排名欄被放置在輸出檔案的最左側位置，順序為：`residual_sf`、`cqa_sf`、`residual`、`cqa`。這樣便於瀏覽排序後的列表，並評估估價對象在 CQA 分布中的位置。

這些欄應用了以下 Excel 數字格式：

欄名	格式	示例
<code>residual_sf</code>	數值，2 位小數	12.34
<code>cqa_sf</code>	數值，2 位小數	7.25
<code>residual</code>	數值，0 位小數	15,200
<code>cqa</code>	數值，2 位小數	6.80

9.3 CQA 評分

CQA（狀況—品質—吸引力）評分將每一列的殘差與所有其他殘差進行排名比較。對於給定的列，CQA 是帶符號殘差小於該列的列所佔的百分比乘以 10。由此得到一個 0–10 的評分尺度，其中：

- 高 CQA ($\approx 9-10$) ——該房產的成交價遠高於模型預測（較大的正殘差）
- 低 CQA ($\approx 0-1$) ——該房產的成交價遠低於預測（較大的負殘差）
- CQA ≈ 5 ——殘差接近所有殘差的中位數

當指定了 `living_area` 欄時，`cqa_sf` 提供基於每平方英尺殘差的同類排名。

9.4 排序與估價對象列

在估價和市場模式下，可比銷售案例行按 `residual_sf` 降序排序（若未指定居住面積則按 `residual` 排序）。估價對象所在列（第 1 列）保持在第 1 位，不參與排序。在通用模式下，各列按其原始順序匯出。

這種排序方式與最左側的排名欄相結合，使估價師能夠快速瀏覽從預測偏高最多到預測偏低最多的可比銷售案例，並評估估價對象被賦予的 CQA 評分在該分布中的位置。

在呼叫 `predict()` 之前，預測資料中的因子水準會與訓練資料對齊。含有未見過的因子水準的列將產生 NA 預測值。

提示：`calc_residual` 欄應始終等於 `residual`（在捨入誤差範圍內）。如果兩者差異顯著，則表明存在值得調查的計算問題。

9.5 mgcvUI 自動匯出

每次成功擬合且 $\text{degree} \leq 2$ 時，`earthUI` 會自動將完整的結果對象儲存為 `.rds` 檔案，存入專案的輸出資料夾。檔案名遵循如下模式：`<datafile>_earthUI_result_<YYYYMMDD_HHMMSS>.rds`。

該檔案可由 `mgcv` 後處理模組載入——該模組包含在 `earthUI` 套件中，通過 `earthUI::launch_mgcv()` 啟動——使用 `readRDS()` 讀取。GAM 常式使用 `earth` 模型的節點位置和基函數作為 GAM 平滑項的起點，從而實現從 MARS 到 GAM 建模的無縫過渡。

$\text{degree} > 2$ 的模型會被跳過，因為 `mgcvUI` 僅支援兩兩交互。當 `earthUI` 以 Trilogy 模式（從 Trilogy UI）啟動時，額外的 **Lock Model Output**（鎖定模型輸出）區塊可讓您將某次擬合的 `.rds`——連同其預測變數基和共享的 CQA 設定——鎖定為供 `glmnetUI` 和 `mgcvUI` 匯入的規範 `earth` 模型。

9.6 glmnetUI 匯入

為 GAM 常式儲存的同一 `.rds` 檔案也可匯入彈性網絡後處理模組——同樣包含在 `earthUI` 套件中，通過 `earthUI::launch_glmnet()` 啟動。在 `glmnet` 應用中，導航至 **Section 2: Import from earthUI**（第 2 節：從 `earthUI` 匯入），使用 Browse（瀏覽）按鈕選擇該 `.rds` 檔案。

`glmnetUI` 採用標準的 `earth::model.matrix()` 方法：`earth` 模型的基函數（鉸鏈、交互項和線性項）成為 `glmnet` 設計矩陣的欄。`glmnet` 隨後對這些基欄執行正則化迴歸，通過 lasso/彈性網絡對其進行選擇和收縮。這將 `earth` 的自適應基構造與 `glmnet` 的正則化相結合。

9.6.1 主要優勢

- 可解釋的交互項：`earth` 基於鉸鏈的交互項（包括三階項）具有清晰的幾何含義，而 `glmnet` 原生的叉積交互項在法庭或審計場合難以解釋。
- 自動非線性：`earth` 的鉸鏈函數可以捕捉使用原始預測變數的標準 `glmnet` 模型會遺漏的非線性關係。
- 變數選擇：`glmnet` 可以進一步修剪 `earth` 基，剔除沒有貢獻的鉸鏈項。

9.6.2 重要提示：權重欄必須一致

警告

如果 earthUI 使用了權重欄（例如，用 `weight=0` 排除估價對象所在列），則必須在 glmnetUI 的變數配置中指定相同的權重欄。如果權重設定不同，列數將無法對齊，earth 基也就無法附加到 glmnet 的模型矩陣中。

9.6.3 工作流程

1. 在 earthUI 中擬合一個 earth 模型（任意階數）。
2. `.rds` 結果檔案會自動儲存到專案的輸出資料夾。
3. 在 glmnetUI 中，開啟 **Section 2: Import from earthUI**，瀏覽至該 `.rds` 檔案。
4. 確認兩個應用中使用的是同一資料檔案和同一權重欄。
5. 在第 5 節中配置 glmnet 參數（當 earthUI 匯入處於活動狀態時，交互矩陣會自動停用）。
6. 點選 **Fit Model**（擬合模型）。earth 基欄將出現在 Equation（方程）、Coefficients（係數）和 Contributions（貢獻）頁籤中。

10 RCA 計算與下載

Calculate RCA Adjustments & Download (計算 RCA 調整並下載) 按鈕位於側邊欄第 7 節，僅在估價模式下且模型擬合完成後可見。RCA (Residual Constraint Approach, 殘差約束法) 生成相對於估價對象 (subject) 的、源自市場的逐個可比案例調整。

10.1 開啟 RCA 對話方塊

點選該按鈕會開啟一個小型模態對話方塊，其中包含：

- **Score type** (評分類型) —— 選項按鈕，用於選擇 **CQA** 或 **CQA per SF** (僅當已指定 **living_area** 欄時，SF 選項才可用)。如果選擇「CQA per SF」，則根據你為估價對象指定的 CQA 評分，其殘差評分將是其居住面積乘以與給定 CQA_SF 評分匹配的 **residual_sf**。
- **CQA value** (CQA 值) —— 用於輸入估價對象的 CQA 評分 (0.00–10.00)。該評分表示你認為估價對象在可比案例品質分布中所處的位置。
- **Generate** (生成) 按鈕——計算 RCA 輸出並以 Excel 格式下載。在輸入有效評分之前，該按鈕保持停用狀態。

10.2 CQA 評分插值

當你點選 **Generate** 時，earthUI 會根據可比案例的 CQA/殘差對來插值估價對象的殘差：

1. 將可比案例的 CQA 評分和殘差按 CQA 升序排序
2. 通過線性插值 (`stats::approx()`) 將你輸入的 CQA 值映射為一個殘差
3. 如果使用 CQA per SF，則將每平方英尺殘差乘以估價對象的居住面積，換算回總殘差
4. 估價對象的估計價值計算為：模型預測值 + 插值殘差

10.3 輸出欄

RCA Excel 檔案 (<datafile>_adjusted_<YYYYMMDD_HHMMSS>.xlsx) 包含所有中間輸出欄，另加：

欄	說明
subject_value	模型預測值 + 插值殘差 (僅第 1 列)
subject_cqa	你輸入的 CQA 評分 (僅第 1 列)
<variable>_adjustment	估價對象貢獻減去可比案例貢獻 (按 g 函數逐項計算)
residual_adjustment	估價對象殘差減去可比案例殘差
net_adjustments	所有調整之和 (貢獻 + 殘差)
gross_adjustments	所有調整絕對值之和
residual_pct	殘差調整佔可比案例售價的比例
net_adj_pct	淨調整佔可比案例售價的比例
gross_adj_pct	總調整佔可比案例售價的比例
adjusted_sale_price	可比案例售價 + 淨調整

各調整欄告訴你，對於每個可比案例，模型將差異歸因於各變數的金額。**adjusted_sale_price** 是應用所有模型推導調整後的可比案例售價——這組調整後售價應聚集在估價對象的估計價值附近。

10.4 RCA Adjustments 頁籤

RCA 輸出計算完成後，主面板中的 **RCA Adjustments** (RCA 調整) 頁籤會顯示各可比案例的 Residual Adj %、Net Adj % 和 Gross Adj % 分布的直方圖（多目標模型則每個目標一組），從而可以快速直觀地檢查調整的聚集程度。

提示：CQA 評分 5.00 表示估價對象處於可比案例的中位數位置。評分高於 5 表示估價對象品質高於平均水準；低於 5 則表示低於平均水準。

11 銷售比較網格

11.1 概述與用途

銷售比較網格 (Sales Comparison Grid) 是在估價模式下 (側邊欄第 8 節)、計算 RCA 調整之後生成的 Excel 活頁簿。它以結構化的網格形式將估價對象與選定的可比案例並排呈現，並附帶自動計算調整和每個可比案例調整後售價的 Excel 公式。

該網格專為估價師的工作底稿而設計——它將 Earth 模型推匯出的迴歸調整與可編輯儲存格相結合，估價師可在其中將 CQA 殘差分配到具體的房產特徵上。輸出檔案名為 SalesGrid_<YYYYMMDD_HHMMSS>.xlsx。

11.2 可比案例選擇模態框

點選「Generate Sales Grid & Download」(生成銷售網格並下載) 會開啟一個模態對話方塊，供你選擇要包含的可比案例：

- 推薦可比案例已預先勾選。這些可比案例的總調整百分比低於售價的 25%，按銷售時間由近及遠排序 (最近的銷售在前)，上限為 30 個。
- 其他可比案例欄在下方，預設不勾選。它們的總調整達到或超過 25%，但仍可供選擇 (最多顯示 50 個，按總調整百分比升序排序)。
- 最多可選擇 30 個可比案例，最多生成 10 張工作表 (每張工作表 3 個可比案例)。

確認選擇後，所選可比案例按總調整百分比升序排序，因此最相似的可比案例會出現在 Sheet 1 上。

11.3 網格佈局

每張工作表採用 20 欄佈局，容納估價對象加 3 個可比案例。各實體的欄為：

實體	欄 (每個實體 5 欄)
估價對象	標籤、事實值 1、事實值 2、事實值 3、價值貢獻 (VC)
每個可比案例	事實值 1、事實值 2、事實值 3、價值貢獻 (VC)、調整

各列從上到下依次為：

1. 標題——工作表名稱 (例如「Intermediate Sales Comparable Grid — Sheet 1 of 3」)
2. 表頭——「Subject」與各可比案例的欄標題及地址
3. 地址——完整的房產地址
4. APN | MLS# | DOM | Subj.Prox (宗地號 | 掛牌編號 | 在售天數 | 與估價對象的距離)——宗地號、掛牌編號、在售天數，以及與估價對象的 Haversine 距離 (英里)
5. Sales Price | Concess. | Net SP (售價 | 讓價優惠 | 淨售價)——售價、讓價優惠，以及 Net Sale Price (淨售價) 公式
6. Regression Features (迴歸特徵) 表頭列
7. BASE VALUE (基準值)——模型截距
8. Date of Sale | OffMkt | OnMkt (成交日期 | 下市 | 上市)——合約日期、銷售距今時間和 DOM
9. 分組列 (視情況而定)——Location (位置)、Site Size (地塊面積) 和/或 Age (屋齡) 組
10. 模型變數列——每個預測變數一列 (不含已分組的變數)
11. 空白分隔列
12. Residual Features (殘差特徵) 表頭列
13. CQA|Residual (CQA | 殘差)——CQA 評分 + 剩餘殘差公式
14. 殘差特徵列——命名列加空白列，供估價師填寫

15. **Total VC / Net Adjustment** (總 VC / 淨調整)
16. **Net Adjustment %** (淨調整百分比)
17. **Gross Adjustment %** (總調整百分比)
18. **Adjusted Sale Price** (調整後售價) ——公式列
19. 版權頁腳

11.4 售價、讓價優惠與 Net SP

第 5 行為每個可比案例顯示三個值：

- **Sale Price**——來自資料的可比案例售價
- **Concessions**——指定為 **concessions** 的欄中的值 (未指定時為 0)
- **Net SP**——一個 Excel 公式： $\text{Sale Price} - \text{Concessions}$

估價對象欄的售價顯示為「N/A」(因為未知)，若有讓價優惠值則一並顯示。

11.5 分組列 (位置、地塊、屋齡)

當指定了某些特殊類型且這些變數出現在模型中時，earthUI 會建立將相關變數合併的分組列：

Loc: Long | Lat | Area (位置：經度 | 緯度 | 區域) ——當 **longitude**、**latitude** 或 **area** 中的任一變數出現在模型中時出現。顯示每個組成變數的事實值、合併的價值貢獻 (各 VC 之和)，以及針對可比案例的調整 (估價對象合併 VC – 可比案例合併 VC)。VC 儲存格以淺藍色背景標示。

Site Size | Dimensions (地塊面積 | 尺寸) ——當 **lot_size** 或 **site_dimensions** 出現在模型中時出現。結構與 **Location** 組相同。

Actual Age | Effective Age (實際屋齡 | 有效屋齡) ——當 **actual_age** 或 **effective_age** 出現在模型中時出現。結構相同。

被分組列佔用的變數將從下方的單個模型變數列中排除，以防止調整合計中的重複計算。

11.6 模型變數列

在分組列下方 (若無分組列，則直接在 **BASE VALUE** 下方)，每個剩餘的模型預測變數佔一列，顯示：

- **事實值**——估價對象和每個可比案例的實際資料值
- **價值貢獻 (VC)** ——該列 **g** 函數對預測值的貢獻
- **調整 (僅可比案例)** ——估價對象 VC 減去可比案例 VC

11.7 CQA|Residual 與殘差特徵列

CQA|Residual 列顯示每個房產的 CQA 評分，並包含剩餘殘差公式——即尚未分配到具體特徵的殘差部分。公式為：

$$\text{Remaining Residual} = \text{Total Residual} - \sum (\text{Residual Feature VCs below})$$

其下是殘差特徵列——5 個命名列 (**Location**、**View/Appeal**、**Condition**、**Quality**、**Other**，即位置、景觀/吸引力、狀況、品質、其他) 加 6 個空白列。這些是輸入儲存格，供估價師為模型未捕捉到的特徵填寫價值貢獻。隨著估價師填入數值，**Remaining Residual** 公式的值會自動減少。

對於每個可比案例，調整欄包含一個公式：估價對象特徵 VC – 可比案例特徵 VC。

11.8 調整後售價公式

Adjusted Sale Price 列包含 Excel 公式：

- 估價對象：從 BASE VALUE 到最後一個殘差特徵列的所有價值貢獻儲存格之和。這代表模型對估價對象的總預測值，加上估價師分配的殘差特徵。
- 可比案例：Net SP + 從第一個分組列或變數列到最後一個殘差特徵列的所有調整儲存格之和。這給出應用所有迴歸推導調整和估價師填寫調整後的可比案例售價。

11.9 工作表保護與可編輯儲存格

每張工作表均已受保護，以防止公式和資料被意外修改。保護鎖定了：

- 所有公式儲存格（Net SP、剩餘殘差、調整、調整後售價等）
- 所有資料儲存格（地址、售價、事實值、來自模型的價值貢獻）
- 標籤和表頭

唯一未鎖定（可編輯）的儲存格是殘差特徵價值貢獻輸入格——即 CQA|Residual 列下方供估價師填寫分解值的儲存格。這些儲存格以淺黃色背景標示，表示可編輯。

11.10 在 Excel 中使用網格

1. 開啟檔案——使用 Excel 或相容的試算表應用程式。
2. 審核迴歸調整——模型推導的調整已預先填入並鎖定。
3. 分配殘差——在殘差特徵列（黃色儲存格）中，填寫你對剩餘殘差中歸因於各特徵（位置、景觀、狀況、品質等）的金額的判斷。隨著分配的進行，Remaining Residual 公式的值會相應減少。
4. 檢查 **Adjusted Sale Price**——該公式會隨你填入殘差特徵值而自動更新。
5. 多張工作表——如果選擇了 3 個以上的可比案例，請在工作表之間切換。每張工作表的左欄均為同一估價對象，並配以 3 個不同的可比案例。

提示：目標是將 CQA 殘差分配到各特徵列，直至 *Remaining Residual* 接近於零。這表明你能夠解釋每個可比案例的迴歸估計值與其實際售價之間的差異。

12 下載報告

模型擬合完成後，報告生成是一個兩步式的 Quarto 工作流：**Generate Quarto Report**（生成 Quarto 報告，估價模式下為側邊欄第 9 節，其他模式下為第 7 節）寫出一個自包含的報告包，而 **Convert Quarto Report**（轉換 Quarto 報告，第 10 節）將其——或任何其他 Quarto 檔案——渲染為最終格式。

12.1 生成 Quarto 報告

點選 **Generate Quarto Report** 會在專案輸出資料夾下的 `<base>_qmd/` 中寫出一個自包含的 Quarto 包，其中包含已填充內容的 `.qmd` 原始檔、所有預生成的圖表資源（PNG 和 PDF）、報告資料，以及一個 Word 參考模板。此步驟不進行任何渲染——該報告包可以手工編輯、通過 Quarto includes 與其他 Quarto 原始檔合併，或使用 `quarto preview` 單獨預覽。

12.2 轉換 Quarto 報告

第 10 節提供一個檔案選擇器（預設指向當前專案最近生成的報告包——但你可以瀏覽到任意 `.qmd` 檔案）以及格式核取方塊：

- **HTML**——一個 `.html` 檔案，帶有 KaTeX 數學公式渲染、內嵌圖片和目錄。使用 Flatly Bootstrap 主題和 Roboto Condensed 字體。
- **Word**——一個適合編輯和分發的 `.docx` 檔案。包含目錄、頁碼，並使用自訂參考模板以保持樣式一致。
- **PDF**——使用 LuaLaTeX 排版，輸出達到專業品質。紙張尺寸遵循區域設定（Letter 或 A4）。包含用於大型交互矩陣的橫向頁面，並使用 Roboto Condensed 搭配 Latin Modern Math 排版公式。僅當檢測到 LaTeX 發行版時才會提供 PDF 選項。

點選 **Convert**（轉換）以渲染所選格式。渲染在背景執行——一個模態對話方塊顯示已用時間和 Quarto 進度，應用保持可回應狀態。完成後會有通知確認檔案位置。

所有操作（模型擬合、資料下載、RCA 計算、銷售網格生成和報告渲染）都會連同起訖時間戳記和耗時記錄到輸出資料夾中的 `<datafile>_earthui_log.txt`，以便排查問題和監控效能。

12.3 報告內容

每份生成的報告都包含以下各節：

1. 資料集描述——觀測數量、目標變數、所用預測變數、分類預測變數
2. 模型規格——階數（degree）、交叉驗證狀態、項數、模型中的預測變數
3. 允許的交互項（僅當 $\text{degree} \geq 2$ 時）——帶勾選標記的交互矩陣，按輸出類型格式化
4. 結果摘要—— R^2 、GRSq、GCV、RSS，以及 $CV R^2$ （若啟用了交叉驗證）
5. 模型方程——以 LaTeX 記號表示的完整 Earth 模型方程
6. 係數與基函數——所有項、其係數以及定義它們的基函數的表格
7. 變數重要性——預測變數重要性得分的長條圖和排名表
8. g 函數貢獻——每個 g 函數的圖：單變數項為折線圖，雙變數交互項為 3D 透視圖和等高線圖
9. 相關矩陣——預測變數相關性的熱圖
10. 診斷——殘差對擬合值（Residuals vs Fitted）、常態 Q-Q 圖，以及實際值對預測值（Actual vs Predicted）圖
11. ANOVA 分解——顯示每個基函數對 RSS 貢獻的表格
12. Earth 輸出——來自 `earth::summary()` 的原始文字輸出，包括剪枝過程的細節

對於多反應模型，第 4–8 節和第 10 節會針對每個目標變數重複一次。

提示：生成報告需要 *quarto R* 套件和 *Quarto* 安裝。若要輸出 *PDF*，還必須安裝 *LaTeX* 發行版（例如通過 `tinytex::install_tinytex()`）——否則 *PDF* 核取方塊會被隱藏。

13 後處理模組：彈性網絡與 GAM

earthUI 套件附帶兩個後處理模組——一個基於 `glmnet` 的彈性網絡 (elastic net) 常式和一個基於 `mgcv` 的廣義可加模型 (GAM) 常式。本期各有一篇文章對它們分別做了深入論述；本節說明它們在 earthUI 工作流中的用途，以及三個常式如何協作。

13.1 為什麼需要輔助擬合？

earth 模型給出價值結論。後處理模組的存在是為了印證它：由方法學上不同的引擎、針對同一估價對象 (subject)、基於相同專案資料產生的兩個額外估計。當對 earth 基函數進行彈性網絡重估計與獨立的 GAM 平滑擬合得出實質上相同的價值時，估價師的結論便不再僅僅依賴於單一模型的假設——這在複核和訴訟場景中是一項實實在在的優勢。當它們不一致時，這種分歧本身就是診斷資訊，通常指向資料稀薄、某個過度靈活的項，或某個值得重新審視的離群可比銷售案例。兩個模組都不凌駕於 earth 模型之上；它們是佐證證據，並有意不做加權呈現。

13.2 啟動與共享狀態

每個常式都是一個獨立的本機 Shiny 應用：

- `earthUI::launch_glmnet()`——彈性網絡應用，連接埠 7879。
- `earthUI::launch_mgcv()`——GAM 應用，連接埠 7880。

三個應用讀取同一專案樹和同一設定資料庫，因此你在 earthUI 中開啟的專案在模組中立即可用——連同 earthUI 為其儲存的沿用欄位：Response (反應/目標變數)、Effective Date (估價基準日)、RCA 估價對象的 CQA 類型和取值，以及居住面積 (living-area) 指定。這些無需重新輸入；模組會讀取它們，使同樣的取值貫穿全部三次擬合。

13.3 彈性網絡常式

`glmnet` 常式將 earth 模型的基函數——其鉸鏈函數、交互項和線性項——作為設計矩陣的欄匯入，並在套索迴歸 (lasso) / 彈性網絡正則化下對其重新估計 (其機制及權重欄的注意事項參見上文「`glmnetUI Import`」)。其結果保留了 earth 可解釋、具有幾何意義的項，同時讓正則化器收縮或剔除那些不能證明自身價值的項。完整論述參見本期的 `glmnetUI` 文章。

13.4 GAM 常式

`mgcv` 常式走的是另一個方向：它不複用 earth 的基函數，而是重新擬合帶懲罰的平滑項，以 earth 模型的節點位置作為起點 (參見上文「`mgcvUI Auto-Export`」)。由於 GAM 在不同的平滑範式下重新推導函數形式，它與 earth 模型的一致性真正獨立的印證，而非回聲。支援 $\text{degree} \leq 2$ 的 earth 模型。完整論述參見本期的 `mgcvUI` 文章。

13.5 一個實用的工作流

1. 先在 earthUI 中擬合并打磨 earth 模型，直到你對其滿意為止。每次成功的擬合 ($\text{degree} \leq 2$) 都會自動匯出供模組使用的 `.rds` 檔案。
2. 針對同一專案開啟 `glmnet` 和 `mgcv` 常式，匯入 earth 結果並進行擬合。一次只在一個常式中工作——擬合耗時較長，且各常式通過共享專案進行協調，因此合理的節奏是順序進行，而非並行。
3. 自由迭代。在你於各應用之間切換時，每個應用都會保留其結果，也沒有任何環節強制你在過程中生成報告。
4. 只有當三個結果都令你滿意時，再回到各個常式生成其 Quarto 報告。每種方法都會將其擬合和價值結論登記到專案的三部曲 (trilogy) 記錄中，使三份報告可追溯到同一次比較執行。

提示：這些模組在任何層面上都是可選的：*earthUI* 自身的工作流——擬合、*RCA*、銷售網格、報告——沒有它們也是完整的。當某個結論需要印證時再使用它們：棘手的估價對象、資料稀薄的市場細分，或很可能面臨複核的委託任務。

14 示範資料集：Appraisal_1.csv

earthUI 包含一個示範 MLS 資料集，用於探索估價工作流。可通過編程方式載入：

```
demo_file <- system.file("extdata", "Appraisal_1.csv", package = "earthUI")
df <- import_data(demo_file)
```

或者，在 Shiny 應用中，將其複製到專案的 in/ 資料夾（或在建立專案時將其作為初始資料檔案附加），然後在第 2 節（Import Data，匯入資料）中選擇它。

14.1 描述

該檔案包含來自一份模擬 MLS 匯出的 1,501 條住宅銷售記錄，外加第 1 列的 1 個估價對象 (subject)（共 1,502 列）。資料代表一個多區域市場中的獨棟住宅銷售，涵蓋多種房屋面積、屋齡和位置。

這不是真實資料，但基於北加州一個真實感很強的社區。所有識別資訊均已被更改或刪除。

14.2 資料欄

欄名	類型	特殊類型	描述
weight	numeric	—	觀測權重 (0 = 不參與擬合)
id	numeric	display_only	內部記錄 ID
property_id	numeric	display_only	MLS 房產識別碼
listing_id	character	display_only	MLS 掛牌編號
parcel_number	character	display_only	縣稅務評估員宗地號 (APN)
street_address	character	display_only	房產地址
city_name	character	display_only	城市
postal_code	character	display_only	郵遞區號 (ZIP)
county_name	character	display_only	縣
contract_date	Date	contract_date	銷售合約日期 (用於計算 sale_age)
sale_age	numeric	—	從合約日期到估價基準日的天數 (預先計算)
coe_date	Date	display_only	過戶完成 (close of escrow) 日期
listing_status	character	display_only	掛牌狀態 (例如 "Sold")
sale_price	numeric	(target)	銷售價格——反應變數
rent	numeric	—	月租金 (用於多目標模型)
list_price	numeric	display_only	掛牌價格
original_list_price	numeric	display_only	原始掛牌價格
living_sqft	numeric	living_area	總居住面積 (平方英尺)
beds_total	integer	—	臥室數量
baths_total	numeric	—	浴室總數 (例如 2.5 = 2 個全衛 + 1 個半衛)
lot_size	numeric	lot_size	地塊面積 (平方英尺)
area_id	integer	area	MLS 區域識別碼
area_text	character	display_only	區域名稱
age	numeric	actual_age	屋齡 (年)
year_built	integer	display_only	建成年份
latitude	numeric	latitude	緯度 (建模時四捨五入到小數點後 3 位)
longitude	numeric	longitude	經度 (建模時四捨五入到小數點後 3 位)
latitude6	numeric	display_only	全精度緯度
longitude6	numeric	display_only	全精度經度
garage_spaces	integer	—	車庫車位數
fp_count	integer	—	壁爐數量
no_of_stories	numeric	—	樓層數
style	character	—	建築風格
view	character	—	景觀類型 (例如 "Neighborhood"、"Hills")
days_on_market	integer	dom	在市天數
listing_date	Date	listing_date	掛牌日期
sale_concessions	numeric	concessions	賣方讓步 (concessions) 金額

14.3 建議的快速上手步驟

1. 啟動 earthUI：`earthUI::launch()`
2. 在 New Project（新增專案）精靈中建立一個 Purpose（用途）為 **For Appraisal**（用於估價）的專案，並附加 `Appraisal_1.csv` 作為初始資料檔案
3. 在第 2 節（Import Data）中選擇 `Appraisal_1.csv`
4. 選擇 `sale_price` 作為目標變數
5. 按上表所示指定特殊類型
6. 納入預測變數：`sale_age`、`living_sqft`、`baths_total`、`lot_size`、`area_id`（作為因子）、`age`、`latitude`、`longitude`、`garage_spaces`
7. 將 `degree` 設為 1，點選 **Fit Earth Model**（擬合 Earth 模型）
8. 下載中間輸出（第 6 步），審閱 CQA 排名
9. 計算 RCA 調整（第 7 步），CQA 得分約為 ~5.00
10. 生成銷售比較網格（第 8 步）

提示：在匯入你自己的 *MLS* 資料之前，示範資料集是探索 *earthUI* 全部功能的好方法。上面建議的預測變數通常能得到 R^2 高於 0.85 的模型。

15 實驗性功能：Prolog 文字提取與推導規則

0.10.0 版本引入了一項可選且確屬實驗性的功能，用於從 MLS 匯出資料的描述性文字欄中提取可供模型使用的特徵——首要的是「Public Remarks」（公開備註）欄，此外還包括經紀人備註以及以逗號分隔的特徵列表，例如房間特徵或既有建築物。提取引擎是 Prolog，具體而言是確定子句文法（**Definite Clause Grammars, DCG**）：在備註（remarks）文字中識別諸如「move in ready」「3 car garage」或「owned solar」等短語的規則，均以 Prolog 文法規則的形式編寫。

該功能仍在開發之中。目前它可以部分用於提取——隨附的文法已能識別一組實用的特徵——但要將其適配到您所在 MLS 的用語習慣，您需要能夠閱讀並編寫 Prolog，尤其是 DCG。如果這不適合您，可以完全跳過本章：該功能預設關閉，關閉時 earthUI 的其餘部分工作方式完全不變。

Warning

這是一項處於積極開發中的實驗性功能。其提取文法既不完整，也未經過面向實際估價工作的驗證——在將提取出的欄用於模型之前，必須將其與原始文字核對。要有效使用該功能，需要具備 Prolog 和 DCG 的實際工作知識。它所做的一切都是附加且可逆的；如果出現任何異常，只需關閉 Settings 中的開關並照常繼續即可。

15.1 它的作用

MLS 房源在自由文字和分隔列表欄位中承載著大量與價值相關的資訊。這一可選步驟通過兩種特殊的欄角色（見第 6 章），將這些文字轉化為迴歸可以使用的普通欄：

- **remarks** 欄由 Prolog DCG 文法解析，每個提取出的屬性生成一個特徵欄。欄前綴由源欄名稱各單詞的首字母構成：public_remarks 生成 pr_* 欄，agent_remark 生成 ar_* 欄，依此類推。
- **list** 欄（例如 existing_structures = "Shed, Workshop"）被拆分為獨熱編碼的 <col>_<value> TRUE/FALSE 欄。這一部分是純 R 實作，完全不需要 Prolog。

隨附的文法（inst/prolog/mls_remarks.pl）目前可提取的內容包括：帶有 1–5 級 condition_score 的分級狀況等級；景觀標誌與景觀類型；廚房與衛浴翻新標誌；硬木及其他地板材質；升級項目列表及數量；車庫車位數與車庫面積；泳池、水療池及其他室外設施；壁爐數量；附屬建築（棚屋、穀倉、工作間、casita、客房、犬舍等）及從「40x50」這類尺寸中解析出的面積；房車停放與儲物空間；太陽能自有與租賃之分、電池儲能；地塊面積（含英畝到平方英尺的換算）；與主體相關聯的交易背景價格金額（降價、HOA 會費、太陽能租賃或買斷、Mello-Roos）；區位線索；以及取值範圍為 [-1, 1] 的亮點 / 顧慮情感評分。

15.2 環境要求：vProlog (Trealla)，而非 SWI-Prolog

備註解析執行於 **vProlog** R 套件之上——一個自包含的 Trealla Prolog 引擎。它不在 CRAN 上；請按照本文開頭「Installing earthUI and vProlog」（安裝 earthUI 與 vProlog）一節所示，從作者的 r-universe 儲存庫安裝（提供 Windows 和 macOS 的預編譯二進位套件，原始碼位於 github.com/wcraytor/vProlog）。無需也不會使用任何系統級 Prolog 安裝（無論是 SWI-Prolog 還是其他實作）。如果缺少 vProlog，備註解析會被跳過並給出提示，而 list 欄拆分仍照常執行。輔助函數 vProlog_available() 可報告引擎與隨附文法是否均已就緒。

15.3 啟用與執行

1. 開啟 Settings（設定，齒輪圖示），勾選 **Enable Prolog / list column processing**（啟用 Prolog / 列表欄處理）。該開關預設關閉，附有「進階功能——面向具備 Prolog 專

業知識的使用者」的說明，按專案記憶設定，並顯示在 Project Information（專案資訊）對話方塊中。

2. 在預測變數表中，將相關欄的 **Special**（特殊）角色設為 **remarks** 或 **list**（每種角色都可以指定給多個欄）。
3. 使用位於側邊欄第 4 節與第 5 節之間的 **Execute Prolog Processing**（執行 Prolog 處理）區域。它提供兩個步驟：

1. **Expand lists & remarks**（展開列表與備註）——對每個 **remarks** 欄執行 DCG 文法，並對每個 **list** 欄進行獨熱拆分。展開是增量式的：只處理自上次執行以來新指定的欄，一旦所有已指定的欄都已展開，按鈕即被停用。解析過程由進度條追蹤。

2. **Execute derivation rules**（執行推導規則）——在展開後的欄上應用您自己的 Prolog **derive/2** 規則（見下文）。規則在執行前會經過語法檢查和試執行驗證。

15.4 推導規則

對於每一資料列，earthUI 將該列的各欄斷言為 **cell(Name, Value)** 事實，並收集您的 **derive(Col, Val)** 規則的所有解——每個不同的 **Col** 都成為一個新欄，且每欄以找到的第一個值為準，因此可以在後面附加一條後備子句。輔助謂詞（**count_true/2**、**count_false/2**、**count_true_prefix/2**、**cell_or/3**、**is_na/1**）支援聚合與 NA 處理，而 **drop/1** 事實會在規則執行後丟棄相應欄。例如：

```
% Merge a collinear cluster of solar flags into one boolean.
derive(solar_any, B) :-
    ( cell(pr_has_solar, true)
    ; cell(pr_solar_owned, true)
    ; cell(pr_has_battery, true) ) -> B = true ; B = false.

% Count amenities instead of OR-ing them (numeric column).
derive(amenity_count, N) :- count_true_prefix(pr_has_, N).

% Threshold -> boolean, with a fallback clause.
derive(big_outbuilding, true) :- cell(pr_outbuilding_sqft, S), S >= 1000.
derive(big_outbuilding, false).

% Discard a raw granular column once consolidated.
drop(pr_price_solar_lease).
```

規則存放在每個專案獨立的 **<project>_rules.pl** 檔案中，該檔案由一份帶完整註釋的模板初始化，並通過 **Edit derivation rules...**（編輯推導規則……）在一個可浮動、可拖動、也可停靠到側邊欄的程式碼編輯器中編輯。**Check**（檢查）按鈕可在不執行規則的情況下對其進行驗證（語法掃描、載入檢查，以及針對資料第 1 列的試執行）。

15.5 帶版本號的活動資料檔案

任一步驟完成後，增補後的資料框都會以 **<original_name>_<YYMMDD_HHMMSS>.xlsx** 的形式寫入專案的輸入資料夾，並成為該專案的活動資料檔案——重新開啟專案時載入的即是增補後的資料，因此展開結果永遠不會被重複計算，而 **glmnetUI** 和 **mgcvUI** 使用的也正是同一個檔案。原始的備註欄會被自動排除在模型之外；生成的欄則不會被自動納入預測變數。文字解析可能產生大量欄，因此請將您（暫時）不希望進入迴歸的欄標記為 **Latent**（潛在），或用 **×** 將其停用（見第 6 章）。

15.6 編程介面與隨附檔案

整個流程均已匯出，可供批次處理使用：**execute_prolog_processing(df, remarks_cols, list_cols, rules_text)** 負責編排各個處理階段，其底層基於 **remarks_to_columns()**、

`split_list_column()`、`apply_user_rules()` 和 `validate_rules()`。位於 `inst/prolog/` 下隨附的 Prolog 原始檔包括：`mls_remarks.pl`（DCG 文法與特徵聚合）、`mls_analyze.pl`（用於批量解析的檔案 IO 驅動）、`mls_rules.pl`（可供推導規則使用的輔助謂詞），以及 `climate_regions_ca.pl`（見下文）。

15.7 相關實驗：氣候區先驗

一項配套的實驗性功能針對資料稀薄的特徵——即那些因僅出現在極少數可比銷售案例中而無法由迴歸估計的特徵（例如蒸發式冷風機）。三個 R 輔助函數——`climate_region_for()`、`market_area_profile()` 和 `climate_feature_priors()`——將專案所在位置（縣+城市）歸入七個購房者認知層面的加利福尼亞氣候區之一（這是對加州能源委員會 16 個建築氣候區的實用性歸並），並按區域返回 HVAC 與製冷特徵（供暖、空調、蒸發式冷風機、全屋換氣扇、空氣過濾）的序數化貢獻價值先驗。Prolog 模組 `climate_regions_ca.pl` 鏡像了同一套分類，可在推導規則內部使用；兩側均由同一個規範的 R 源生成，因此不會產生偏差。這些先驗旨在為特徵估值提供參考，並作為對估計係數的合理性核查——區域歸屬（尤其是橫跨多個區域的縣）應由估價師複核。

15.7.1 超越加利福尼亞：轄區包與市場區域模組

加利福尼亞是第一個轄區，而非邊界。`earthUI` 在國際範圍內被使用，氣候先驗知識正被重構為轄區包（**jurisdiction packs**）：自包含、資料化的事實表（分區定義、行政單元歸屬、特徵先驗），以專案系統的國家代碼和行政層級代碼為鍵，因此一個包可以描述一個美國的州、一個德國的聯邦州（Land），乃至直接描述一個小國。與之配套的市場區域模組（**market-area modules**）層則捕捉某一具名市場區域的在地用語與更精細的先驗。

二者的設計目標都是由 Valuation Engineer 和估價師在一線現場、用他們自己的語言編寫——並且，從現實出發，可借助 LLM 完成：世界上並沒有多少 Prolog 程式設計師，因此現場產出的成果被有意設計為表格而非程式。包格式是一個有說明文件的模式（schema），並附有完整的成品示例（即加利福尼亞本身），其明確目的就是讓 Claude 這樣的助手能夠根據對某一市場的自然語言描述起草出一個有效的包；`earthUI` 會從包生成底層的 Prolog，並且一個具有阻斷性的驗證器（模式、覆蓋度、引用完整性和試執行檢查）必須通過後，包才能被啟用。每個現場編寫的包都必須攜帶出處資訊——署名作者、日期以及所引用的權威依據——因此為這些知識背書的是負有責任的專業人員而非起草工具，報告也能準確引用所使用的包及其版本。

提示：一個實用的模式：先展開備註，再用 `derive/2` 規則將細粒度的 `pr_*` 標誌整合為少數幾個派生的計數或布林欄，用 `drop/1` 丟棄已整合的原始標誌，只保留派生欄作為候選預測變數。這樣可以保持預測變數列表精簡，並使共線性可控。

參考文獻

- Friedman, Jerome H. (1991). "Multivariate Adaptive Regression Splines". In: *The Annals of Statistics* 19.1, pp. 1–141. DOI: [10.1214/aos/1176347963](https://doi.org/10.1214/aos/1176347963).
- Milborrow, Stephen (Oct. 1, 2024). *earth: Multivariate Adaptive Regression Splines*. R package. URL: <https://CRAN.R-project.org/package=earth>.
- (n.d.[a]). *Notes on the earth Package*. URL: <http://www.milbo.org/doc/earth-notes.pdf>.
- (n.d.[b]). *Variance Models in earth*. URL: <http://www.milbo.org/doc/earth-varmod.pdf>.

glmnetUI : glmnet (Lasso 與彈性網絡) 套件的互動式介面

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。

William Bert Craytor

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: [10.67330/ja8s2k39](https://doi.org/10.67330/ja8s2k39)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 [創用 CC 姓名標示 4.0 國際授權條款 \(CC BY 4.0\)](#) 的條款發布。

摘要。 glmnetUI 是 R 語言 glmnet 套件的圖形使用者介面 (graphical user interface)，該套件透過 Lasso、嶺迴歸與彈性網絡 (elastic net) 罰則來擬合正則化的廣義線性模型 (generalized linear model)。它提供三種用途模式——通用預測建模、不動產估價與市場區域分析——並引導使用者完成資料匯入、模型設定、交叉驗證擬合、係數路徑檢視以及報告下載。本文記錄了 glmnetUI 的資料格式要求、建模工作流程、輸出顯示介面以及完整的功能參考。

關鍵詞： Lasso，嶺迴歸，彈性網絡，正則化迴歸，glmnet，R，圖形使用者介面，不動產估價

出版者按 (2026 年 7 月)：付印前不久，glmnetUI 已併入 earthUI (0.11.0 及以後版本)，本文所述的彈性網絡工作流程現與 MARS 和 GAM 工作流程一同由 earthUI 提供。glmnetUI 仍以其最終發布形態 (0.5.0) 保留在 <https://github.com/wcraytor/glmnetUI>；後續開發將在 earthUI 中繼續。本文適用於已發布的 glmnetUI。

目錄

1	引言	3
1.1	歷史與背景	3
1.1.1	Lasso	3
1.1.2	嶺迴歸與彈性網絡	4
1.1.3	glmnet 套件	4
1.1.4	與 MARS (earth) 和 GAM (mgcv) 的比較	4
1.1.5	建議的工作流程	5
1.2	什麼是 glmnetUI ?	5
1.3	三種用途模式	5
1.4	不動產專用功能	6
1.5	快速上手	6
2	MLS 輸入資料要求	7
2.1	檔案格式與結構	7
2.2	估價模式所需的欄位	7
2.3	特殊欄位命名慣例	8
2.4	資料品質與完整性	8
2.5	標的列的位置	8
3	通用模式	9
3.1	概述	9
3.1.1	跳過第一列	9
3.1.2	按用途分別儲存設定	9

3.2	側邊欄工作流程	9
3.3	專案	10
3.4	從 earthUI 匯入 (選用)	10
3.5	主面板分頁	10
3.6	設定的保存	11
3.7	深色模式	11
3.8	地區與區域設定	11
3.8.1	儲存預設值	12
4	估價模式	13
4.1	標的列的處理	13
4.2	基準日與成交時距	13
4.3	特殊欄位指定	13
4.4	RCA 調整概覽	13
5	市場區域分析模式	14
5.1	與估價模式的差異	14
5.2	何時使用市場模式	14
6	變數選擇	15
6.1	目標變數	15
6.2	預測變數表	15
6.3	資料型別偵測與手動覆寫	15
6.4	預期符號與約束的施加	15
6.5	特殊欄位型別參考	16
7	參數選擇	17
7.1	Alpha (混合參數)	17
7.2	Lambda 的選擇	17
7.3	Family (分布族)	17
7.4	Standardize (標準化)	17
7.5	符號約束	18
7.6	鬆弛 Lasso	18
7.7	隨機種子	18
7.8	交互項矩陣	18
7.9	禁止進入主效應 (僅作交互項)	19
7.10	進階參數	19
7.11	設定預設值	19
8	擬合模型	20
8.1	擬合按鈕	20
8.2	擬合過程中發生了什麼	20
9	結果分頁	21
9.1	Data Preview (資料預覽)	21
9.2	Equation (方程式)	21
9.3	Correlation (相關性)	21
9.4	Summary (摘要)	21
9.5	Coefficients (係數)	21
9.6	Variable Importance (變數重要性)	22
9.7	Contributions (貢獻)	22
9.8	ANOVA (變異數分析)	22

9.9	Diagnostics (診斷)	22
9.10	Glmnet Output (Glmnet 輸出)	22
9.11	Report (報告)	23
10	下載資料	24
10.1	輸出欄位	24
10.2	CQA 分數	24
11	RCA 計算與下載	25
11.1	開啟 RCA 對話方塊	25
11.2	CQA 分數內插	25
11.3	輸出欄位	25
12	銷售比較表格	26
12.1	可比案例選擇對話方塊	26
12.2	活頁簿版面	26
12.3	工作表保護與剩餘特徵儲存格	26
13	下載報告	27
13.1	兩階段 Quarto 工作流程	27
13.2	報告格式	27
13.3	報告內容	27
14	與 earthUI 的比較	29
14.1	主要差異	29
14.2	共有的功能	29
14.3	何時該用哪一個	29
15	示範資料集：Appraisal_1.csv	30
15.1	說明	30
15.2	各欄位	30
15.3	建議的快速上手流程	30
16	系統需求與疑難排解	32
16.1	支援的平台	32
16.2	選用的相依套件	32
16.3	各平台說明	32
16.4	功能降級	32
16.5	疑難排解	32

1 引言

1.1 歷史與背景

1.1.1 Lasso

Lasso (Least Absolute Shrinkage and Selection Operator, 最小絕對收縮與選擇運算子) 由 Robert Tibshirani 於 1996 年提出。這個名稱是對牛仔套索 (lasso) 的有意雙關——套索是一種既能約束又能選擇的繩索。Tibshirani 之所以選用這個名字，是因為該方法真的會「套住」係數，約束它們的絕對值，並把其中一部分拉到恰好為零，從而選出哪些變數留在模型之中。

Tibshirani 的洞見在於給最小平方目標函數加上一個 L_1 罰則 (penalty, 係數絕對值之和)。嶺迴歸 (ridge regression, Hoerl 與 Kennard, 1970) 所用的 L_2 罰則只會把係數向零收縮而永

遠遠不到零；與之不同， L_1 罰則產生的是稀疏解，其中許多係數恰好等於零。這使得 Lasso 既是一種正則化 (regularization) 方法，也是一種變數選擇方法。

1.1.2 嶺迴歸與彈性網絡

嶺迴歸 (ridge regression) 由 Arthur Hoerl 與 Robert Kennard 於 1970 年提出。它添加的是 L_2 罰則 (係數平方之和)，會按比例收縮所有係數，但把所有變數都保留在模型中。當預測變數之間存在相關性 (多重共線性) 時，嶺迴歸很有效，但它不會簡化模型。

彈性網絡 (elastic net) 由 Hui Zou 與 Trevor Hastie 於 2005 年提出，它把兩種罰則結合起來。混合參數 α 控制兩者的配比： $\alpha = 1$ 為純 Lasso， $\alpha = 0$ 為純嶺迴歸，介於其間的取值則給出折衷方案。彈性網絡克服了 Lasso 的一個侷限：當預測變數高度相關時，Lasso 往往只選中其中一個而忽略其餘，而彈性網絡則會把相關的預測變數成組保留。

1.1.3 glmnet 套件

R 套件 `glmnet` 由史丹佛大學的 Jerome Friedman、Trevor Hastie 與 Robert Tibshirani 開發。它使用極其高效的座標下降演算法為廣義線性模型 (generalized linear model, GLM) 實作彈性網絡正則化 (Friedman, Hastie, and Tibshirani 2010)，其後又擴展到 Cox 比例風險模型 (Simon et al. 2011) 以及全部廣義線性模型族 (Tay, Narasimhan, and Hastie 2023)。該套件可處理 gaussian (線性)、binomial (羅吉斯)、poisson 及其他 GLM 族。

`glmnet` 的主要特性包括：

- 正則化路徑：一次運算即可在整個 λ 網絡上擬合模型，比在每個 λ 處分別擬合快得多。
- 交叉驗證 (cross-validation)：`cv.glmnet()` 透過 k 摺交叉驗證自動選出最佳的正則化強度。
- 係數界限：支援對單個係數設定上限和下限 (符號約束)。
- 鬆弛 lasso (relaxed lasso)：在選中的變數上以減弱的罰則或完全不加罰則重新擬合模型，以減小偏誤。

1.1.4 與 MARS (earth) 和 GAM (mgcv) 的比較

`earth` 應用程式及其兩個後處理模組 (`glmnet` 與 `mgcv`) 使用不同的建模引擎。下表總結了它們的主要差異：

特性	<code>glmnet</code> (彈性網絡)	<code>earth</code> (MARS)	<code>mgcv</code> (GAM)
模型類型	線性 (帶正則化)	分段線性 (自適應樣條)	平滑非線性 (罰則樣條)
非線性	僅能透過交互項或基底函數展開實現	透過帶資料驅動節點的鉸鏈函數自動實現	透過平滑函數自動實現 (薄板、三次樣條等)
變數選擇	內建，經由 L_1 罰則 (Lasso)	內建，經由前向 / 後向逐步法配合 GCV 修剪	選用，經由雙重罰則收縮 (<code>select=TRUE</code>)
交互項	兩兩交叉乘積 (不易解釋)	基於鉸鏈函數的乘積 (可解釋，最高 3 階)	張量積平滑項 (<code>te()</code> 、 <code>ti()</code>)
可解釋性	係數即線性權重，逐個易於說明	g 函數按變數展示分段線性偏效應	平滑的偏效應曲線，非常直觀
最適用於	高維資料、變數選擇、小樣本	自動識別非線性與交互作用	平滑的非線性關係、靈活建模
RCA 調整	逐變數的線性調整	經由 g 函數的分段線性調整	經由偏效應的平滑調整
過擬合控制	交叉驗證選定的 λ	基於 GCV 的修剪	REML/GCV 平滑參數估計

MARS (Multivariate Adaptive Regression Splines, 多元自適應迴歸樣條) 由 Jerome Friedman 於 1991 年提出。它透過從資料中自適應地選擇鉸鏈函數 $\max(0, x - k)$ 及其節點位置來建構分段線性模型。其 R 實作是 Stephen Milborrow 編寫的 **earth** 套件。MARS 的長處在於無需使用者指定即可自動發現非線性關係與交互作用。它的 **g** 函數 (按變數歸組的基底項) 提供了高度可解釋的偏效應曲線。

GAM (Generalized Additive Models, 廣義可加模型) 由 Hastie 與 Tibshirani 於 1990 年正式確立。其 R 實作 **mgcv** 由 Simon Wood 編寫, 提供帶自動平滑度選擇 (經由 REML 或 GCV) 的罰則迴歸樣條。GAM 產生的平滑偏效應曲線便於視覺化和解釋。當 **earthUI** 把節點位置匯出到 **mgcvUI** 時, **earth** 得出的節點會作為 GAM 平滑項的起始點, 從而把 MARS 的自適應節點定位與 GAM 的平滑估計結合起來。

1.1.5 建議的工作流程

對於不動產估價 (real estate appraisal) 以及其他需要可解釋、可辯護模型的類似應用：

1. 從 **earthUI** 入手, 發現資料中的重要變數、非線性關係與交互作用。**earth** 的自動基底函數選擇與 **g** 函數能提供極佳的初步洞察。
2. 用 **mgcvUI** 加以細化, 如果您希望得到更平滑的偏效應。把 **earthUI** 的節點匯入 **mgcvUI**, 即可從分段線性模型平順過渡到平滑模型。
3. 使用 **glmnetUI**, 當您需要力度更大的變數選擇 (Lasso)、預測變數數量接近或超過觀測數量, 或者希望得到一個帶正則化、便於辯護的純線性模型時。

1.2 什麼是 glmnetUI ?

glmnetUI 是 R 語言 **glmnet** 套件的圖形使用者介面 (graphical user interface)。它以本機 Shiny 應用程式的形式執行——無需登入、無需伺服器、也沒有帳號。您從 R 中啟動它, 匯入資料集 (CSV 或 Excel), 設定模型, 然後互動式地進行擬合。

從當前版本起, **glmnetUI** 不再是一個單獨安裝的套件: 它作為彈性網絡後處理模組隨 **earthUI** 套件一起發布, 用 `earthUI::launch_glmnet()` 啟動。應用程式本身——以及本文所述的一切——並無改變; 不同的只是安裝方式和啟動指令。三個常式 (**earth**、**glmnet**、**mgcv**) 共享同一個專案樹、同一個設定資料庫以及同一套支撐基礎設施。

該應用程式提供了完整的工作流程: 資料匯入、變數設定、模型擬合、診斷圖、變數重要性、模型方程式, 以及可下載的 Word、PDF 或 HTML 格式報告。

彈性網絡迴歸結合了兩種正則化技術:

- 嶺迴歸 ($\alpha = 0$): 按比例把所有係數向零收縮。當您認為所有變數都相關時較為合適。
- Lasso 迴歸 ($\alpha = 1$): 可以把係數收縮到恰好為零, 從而實現自動變數選擇。產生的模型更簡潔、更易解釋。
- 彈性網絡 ($0 < \alpha < 1$): 兩種做法的混合。當預測變數之間相關時很有用。

1.3 三種用途模式

每個 **glmnetUI** 專案都有一個 **Purpose** (用途), 在建立專案時選定 (第 1 節, *Project* (專案))。當前專案的用途決定哪些工具和介面元素會出現; 若要在另一種模式下工作, 請建立或開啟一個具有該用途的專案。三種模式為:

- **General** (通用) ——適用於任何類型的母體或資料集的彈性網絡迴歸。這是預設模式。它提供完整的正則化迴歸工作流程, 不含任何領域專用的附加功能。
- **For Appraisal** (用於估價) ——為不動產估價量身打造的彈性網絡迴歸。增加了針對單一不動產估值的功能, 包括估價標的處理、特殊欄位指定以及殘差約束法 (Residual Constraint Approach, RCA)。
- **Market Area Analysis** (市場區域分析) ——為市場區域研究量身打造的彈性網絡迴歸。增加了分析既定市場中成組不動產的功能, 並可選擇排除標的列。

在這三種模式下，核心建模引擎完全相同——您始終是在擬合一個彈性網絡（glmnet）模型。用途設定只控制哪些附加工具和介面元素可用。

1.4 不動產專用功能

當選擇 **For Appraisal** 或 **Market Area Analysis** 時，glmnetUI 會啟用若干為不動產分析設計的功能：

- **特殊欄位指定**——每個預測變數都可以標記一個特殊角色，例如 `contract_date`、`dom`、`concessions`、`latitude`、`longitude`、`living_area`、`lot_size`、`actual_age`、`effective_age`、`area`、`site_dimensions` 或 `display_only`。這些指定控制該欄在擬合與輸出過程中的處理方式。
- **經緯度取整**——被指定為緯度或經度的欄會自動四捨五入到小數點後 3 位，以防止過擬合。
- **Sale Age 欄（成交時距）**——當某欄被指定為 `contract_date` 且提供了 Effective Date（估價基準日）時，glmnetUI 會計算一個 `sale_age` 欄（成交日期與基準日之間的天數），並用它替代原欄作為預測變數。
- **RCA 計算（僅限估價模式）**——在估價模式下，擬合模型之後 glmnetUI 可以計算殘差約束法（RCA）輸出，得出逐個可比案例的調整值、淨 / 毛調整彙總，以及估價標的的調整後成交價格。

提示：*General* 用途模式適用於任何資料集——金融、科學、工程或不動產。估價模式與市場模式只是為不動產專業人士增加了一些便利功能。

1.5 快速上手

使用 glmnetUI 的步驟：

1. 在 R 中安裝 **earthUI**：`install.packages("earthUI")` ——glmnet 常式已包含在內。（關於開發版和選用元件，參見本期 earthUI 一文中的「Installing earthUI and vProlog」。）
2. 啟動應用程式：在 R 中執行 `earthUI::launch_glmnet()`。應用程式會在您的網頁瀏覽器中以連接埠 7879 開啟。您也可以直接在瀏覽器中前往 <http://localhost:7879>。應用程式會記住您上次使用的用途模式並自動還原。
3. 在側邊欄第 1 節建立或開啟專案。專案設定用途和工作地點（國家、州 / 省、縣、市）及其輸入 / 輸出資料夾。專案存放在共享的 `regProj` 根目錄下，並與姊妹應用程式 earthUI 和 mgcvUI 共用。開啟一個具有不同用途的專案，就是切換模式的方式。
4. 在第 2 節匯入資料，從當前專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。把檔案放入該資料夾後，點選 **Refresh**（重新整理）即可列出它們。
5. 從 **earthUI** 匯入（選用）——第 3 節允許您匯入一個 earthUI 結果 `.rds` 檔案，以便把 earth 的鉸鏈基底函數與 glmnet 的正則化結合使用。若沒有 earthUI 結果，可跳過此步。
6. 設定變數——選擇目標變數和預測變數，設定資料型別、預期符號，並指定任何特殊欄位角色。
7. 設定 **glmnet** 參數——`alpha`、`lambda`、`family`、符號約束、交互項及其他選項。
8. 擬合模型——點選「Fit Glmnet Model」（擬合 Glmnet 模型），並在主面板中檢視結果。
9. 匯出——將預測結果下載為 Excel，產生並轉換 Quarto 報告，或（在估價模式下）計算 RCA 調整值與銷售比較表格（Sales Comparison Grid）。

設定會自動儲存在瀏覽器的本機儲存空間中，並在您重新載入同一個輸入檔案時還原。

2 MLS 輸入資料要求

對於不動產估價與市場分析工作流程，您的輸入資料通常來自多重上市服務（Multiple Listing Service，MLS）匯出檔案。本章說明預期的檔案結構以及 glmnetUI 可以使用的欄位。

2.1 檔案格式與結構

glmnetUI 接受 CSV 與 Excel (.xlsx、.xls) 檔案。匯入時，欄名會自動轉換為 snake_case——例如，「Living SqFt」變為 living_sqft，「Contract Date」變為 contract_date，「Sale Price」變為 sale_price。這種規範化確保了整個工作流程中欄位參照的一致性。匯入時使用的 CSV 分隔符號和小數點記號由地區設定決定（參見第 3 章「Locale & Regional Settings」）。

您的資料檔案應當是一張平面表格，每列一筆不動產，每欄一個屬性。檔案的第一列必須包含欄標題。

2.2 估價模式所需的欄位

雖然 glmnetUI 可以處理任意一組欄位，但完整的估價工作流程若具備以下各欄會更為得力：

欄位	特殊型別	用途
Sale Price（成交價格）	（目標變數）	模型的反應變數
Contract Date（合約日期）	contract_date	用於計算 sale_age（距基準日的天數）
Listing Date（掛牌日期）	listing_date	在沒有 DOM 欄時，與合約日期一起計算 DOM
Days on Market（在市天數）	dom	在市天數；顯示於匯出結果中
Concessions（成交讓利）	concessions	成交讓利；淨成交價 = 成交價格 - 讓利
Living Area (SF)（居住面積）	living_area	啟用每平方英尺殘差 (residual_sf、cqa_sf)
Lot Size（基地面積）	lot_size	基地面積欄
Site Dimensions（基地尺寸）	site_dimensions	與基地面積歸為一組
Latitude（緯度）	latitude	四捨五入到小數點後 3 位
Longitude（經度）	longitude	四捨五入到小數點後 3 位
Area ID（區域編號）	area	市場區域 / 街廓識別碼
Actual Age（實際屋齡）	actual_age	不動產的屋齡
Effective Age（有效屋齡）	effective_age	不動產的有效屋齡
Address（地址）	display_only	顯示於匯出結果中；不進入模型

試算表的欄名可以使用外語——「特殊」名稱之所以用英文，是為了讓 R 程式能對它們作特殊處理。除此之外，您所給定的欄名會出現在迴歸模型、圖形中，以及（若在做估價）輸出報告中。

並非所有欄位都是必需的。glmnetUI 會自行調適——若缺少某一欄，相應的功能就直接省略。不過，對於不動產定價模型，為達到可接受的擬合效果，強烈建議具備以下幾欄：

1. **Sale Age** (成交時距) ——合約成交日期與估價或分析基準日之間的天數。若使用跨年度的成交歷史資料，尤其是超過 5 年的期間，`sale_age` 在估計成交價格時往往起核心作用。
2. **Living Area** (居住面積) ——也常被稱為「Living Sqft」、「GLA」(gross living area，總居住面積) 等。這是成交價格的又一主要決定因素。
3. **Total Bath Count** (衛浴總數) ——全套、四分之一套、半套和四分之三套衛浴的總數。例如，兩套全衛浴加一套半衛浴記為 2.5。
4. **Garage Bays** (車位數) 或 **Garage Area** (車庫面積) ——車庫停車位數量或車庫面積。
5. **Lot Size** (基地面積) ——不動產的土地面積，通常以平方英尺或英畝計。
6. **Longitude** (經度)、**Latitude** (緯度)，以及 (若有) **Area ID** (區域編號)。這些位置變數有助於模型解釋地理上的價格差異。

2.3 特殊欄位命名慣例

glmnetUI 透過特殊型別指定而非欄名來識別欄位。在 MLS 匯出檔案中，您可以隨意命名各欄——關鍵在於您在變數設定表 (第 6 章) 中指定了正確的特殊型別。

例如，您的 MLS 可能把居住面積匯出為「GLA」、「Living SqFt」、「liv_area」或「gross_living_area」。匯入之後 (欄名轉為 snake_case)，您只需在 Special (特殊) 下拉式選單中把它指定為 `living_area`。此後無論其原始名稱如何，glmnetUI 都會用它來計算每平方英尺的殘差。

2.4 資料品質與完整性

- 遺漏值 (NA)：任何預測變數欄或目標欄中含 NA 的資料列，會在擬合前自動剔除。
- 日期欄：符合常見日期格式的字元欄在匯入時會自動解析。當偵測不到 4 位年份時，優先按 2 位年份格式解析。
- 數值欄：成交價格、居住面積、基地面積、讓利等欄位必須為數值。若您的 MLS 匯出的價格帶有貨幣符號或千位分隔符號 (例如「\$350,000」或「350.000,00」)，可能需要在匯入前先行清理。
- 因子欄：文字欄、TRUE/FALSE 欄和 R 因子欄會自動按類別變數處理——迴歸只能把它們當作因子。而數值欄 (例如數字形式的區域碼或風格碼) 只有在您勾選其 **Factor** (因子) 核取方塊時才會變成因子；glmnetUI 絕不會根據數值的取值範圍推斷其為類別變數，因此像 `bath_count` 這樣的離散數值預測變數仍保持連續。Special 角色與是否作為因子無關。

提示：匯入後請檢視預測變數表中的 NA 一欄。遺漏值很多的欄位可能導致大量資料列被剔除。可以考慮排除高 NA 欄位，或在匯入前清理資料。

2.5 標的列的位置

在 **Appraisal** (估價) 模式下，第 1 列必須是估價標的。其餘各列均為可比成交案例。標的列不參與模型擬合。擬合完成後，模型仍會為標的列產生預測值。

在 **Market Area Analysis** (市場區域分析) 模式下，把標的放在第 1 列是選用的。

在 **General** (通用) 模式下沒有特殊的資料列處理——所有資料列一視同仁。

3 通用模式

3.1 概述

啟動 glmnetUI 時的預設模式即為通用 (General Purpose) 模式。它為任何資料集——而不只是不動產資料——提供完整的彈性網絡迴歸工作流程。您可以把 glmnetUI 用於科學資料、金融分析、工程研究，或任何迴歸問題。

在通用模式下，介面會略去不動產專用功能（特殊欄位、成交時距、座標取整、RCA）。側邊欄經過精簡，專注於變數選擇、參數設定、模型擬合與匯出。

3.1.1 跳過第一列

開啟通用或市場專案後，側邊欄頂部會出現一個 **Skip first row**（跳過第一列）核取方塊。勾選後，第 1 列將不參與模型擬合。當第 1 列包含一個不應影響模型的標的或參照觀測值時，這一選項很有用。在估價模式下，第 1 列（估價標的）始終自動排除。

3.1.2 按用途分別儲存設定

設定（預測變數選擇、模型參數、交互項）會按輸入檔案與用途模式的每種組合分別儲存。在通用、估價和市場模式之間切換時，各模式的設定互不影響、各自保留。

3.2 側邊欄工作流程

側邊欄按編號、可收合的小節組織，引導您從專案選擇一路走到報告匯出。多數小節只在開啟使用中的專案後才出現，其中幾個還與專案的用途有關。輸出資料夾由使用中的專案推導得出——沒有單獨的輸出資料夾欄位。

1. **Project**（專案）——建立或開啟專案。專案設定用途和工作地點（國家、州 / 省、縣、市）及其輸入（<os>_in/）和輸出資料夾。專案存放在共享的 **regProj** 根目錄下（預設 `~/regProj`，可用 `REGPROJ_ROOT` 環境變數覆寫），並與姊妹應用程式 **earthUI** 和 **mgcvUI** 共用，因此三者的模型可以並存。點選 **Close Project**（關閉專案）即可切換。

2. **Import Data**（匯入資料）——從使用中專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。透過 Finder / 檔案總管把檔案放入該資料夾，然後點選 **Refresh**（重新整理）列出它們。對於含多個工作表的 Excel 檔案，會出現工作表選擇器。欄名會自動轉換為 `snake_case`。

3. **Import from earthUI**（從 earthUI 匯入，選用）——匯入一個 earthUI 結果 `.rds` 檔案，以便把 earth 的鉸鏈基底函數與 glmnet 的彈性網絡正則化結合使用。瀏覽按鈕的樣式與第 2 節的檔案選擇器一致。此步驟為選用——若沒有可匯入的 earthUI 結果，可以跳過。

4. **Variable Configuration**（變數設定）——目標變數選擇器，帶 Include（納入）、Factor（因子）和 Force（強制保留）核取方塊的預測變數表，以及 Sign（符號）下拉式選單。Special（特殊）欄——用於指定合約日期、座標等——只在估價與市場模式下出現。詳見第 6 章。

5. **glmnet Call Parameters**（glmnet 呼叫參數）——全部模型設定：alpha、lambda、family、standardize、符號約束、鬆弛 lasso、隨機種子、交互項矩陣，以及進階參數（lambda.min.ratio、nlambda、CV 損失度量、收斂閾值、最大疊代次數、截距）。完整的參數參考見第 7 章。

6. **Fit Glmnet Model**（擬合 Glmnet 模型）——執行模型的按鈕。結果會顯示在右側的各個分頁中。

7. **Download Output**（下載輸出）——把預測值、殘差、CQA 分數和逐變數貢獻匯出為 Excel 檔案，存入專案的輸出資料夾。模型擬合後，在所有用途模式下均可用。

8. **Calculate RCA Adjustments & Download**（計算 RCA 調整值並下載）——為每個可比案例計算逐變數的 RCA 調整值，並透過 CQA 內插得出標的殘差。僅在估價與市場模式下、擬合之後出現。詳見第 11 章。

9. Generate Sales Grid & Download (產生銷售比較表格並下載) ——產生中間的銷售比較表格，按毛調整百分比對可比案例排序。僅在估價模式下、計算完 RCA 調整值之後出現。詳見第 12 章。

10. Generate Quarto Report (產生 Quarto 報告) ——把一個自足的 Quarto (.qmd) 報告包——含全部圖形與表格——寫入專案的輸出資料夾。模型擬合後即可使用。

11. Convert Quarto Report (轉換 Quarto 報告) ——把產生的 .qmd 報告包轉譯為 HTML、Word 或 PDF。PDF 需要安裝 LaTeX；若未偵測到，則隱藏 PDF 選項。詳見第 13 章。

在市場模式與通用模式下，僅限估價的步驟會被跳過，兩個報告小節的編號也相應調整。

3.3 專案

glmnetUI 把工作組織為共享 regProj 根資料夾下的專案 (earthUI 與 mgcvUI 使用的是同一棵目錄樹)。每個專案擁有一個輸入資料夾 (<os>_in) 和按引擎劃分的輸出資料夾 (例如 <os>_out_glmnet)。

要建立專案，請點選第 1 節中的 + New... (新建)，並提供：

- **Purpose** (用途) ——General、Appraisal 或 Market Area。
- **Location** (地點) ——專案的政治 / 行政層級。需要填寫哪些層級取決於您選擇的國家，因此各國的命名慣例都能自動支援——美國專案使用州 / 縣 / 市，法國專案使用 region/département/commune，日本專案使用都道府縣 / 市或區 / 町丁，依此類推。
- **Project Name** (專案名稱) ——一個簡短標籤，限 8 個字元 (字母、數字、_ 和 -)。

應用程式不會原樣使用您輸入的名稱：它會在前面加上地點代碼和建立日期時間，構造出一個唯一的帶時間戳記的資料夾名，形式為 <country>_<admin level codes>_<YYYYMMDD-HHMMSS>_<your name> ——例如，一個名為 burl 的美國專案會變成 us_ca_081_burlin_20260604-131500_burl。8 個字元的限制只適用於您鍵入的那一部分。只有新建的專案才會帶上時間戳記前綴；磁碟上已有的專案保留原名，並且照常完全可用。

3.4 從 earthUI 匯入 (選用)

側邊欄的第 3 節——**Import from earthUI** (從 earthUI 匯入) ——允許您匯入一個 earthUI 結果 .rds 檔案。這使得 glmnetUI 能夠把 earth 的鉸鏈基底函數 (分段線性項) 與 glmnet 的彈性網絡正則化結合起來，把 earth 的自動非線性識別與 glmnet 的變數選擇和係數約束合二為一。

匯入 earthUI 結果後，驅動 glmnet 擬合的是 earth 模型的預測變數，而不是 Include 核取方塊，因為 glmnet 是在 earth 的基底函數展開上擬合的。變數設定表會明確體現這一點：**Include** 與 **Type** 被鎖定，earth 模型未使用的預測變數所在資料列被停用，並有一條提示說明預測變數集來自 earth 模型。**Force** 與 **Sign** 對 earth 的預測變數仍可編輯，因為它們在擬合時仍會作用於 earth 基底函數。

瀏覽按鈕的樣式與第 2 節的檔案選擇器一致。此步驟完全選用——若沒有 earthUI 結果檔案，直接跳過第 3 節進入第 4 節即可。

3.5 主面板分頁

擬合完成後，主面板提供以下分頁：

分頁	內容
Data Preview (資料預覽)	匯入資料的預覽。在估價模式下，分為估價標的表與可比成交案例表。
Equation (方程式)	以 LaTeX/MathJax 顯示的擬合模型方程式，展示每個非零係數與交互項。
Correlation (相關性)	數值型預測變數相關係數的熱圖（擬合前即可檢視）。
Summary (摘要)	關鍵指標 (R^2 、調整 R^2 、 GR^2 、 $CV R^2$ 、RMSE、MAE) 與估價指標 (COD、PRD、中位數比率)。
Coefficients (係數)	非零係數表，附符號違例警示。
Variable Importance (變數重要性)	預測變數重要性 ($ \beta \times sd(x)$) 的長條圖與排序表。可用時支援 plotly 互動式滑鼠停留。
Contributions (貢獻)	帶斜率標註的逐變數偏效應圖。數值型預測變數為散布圖 + 擬合線，因子為盒鬚圖，交互項為三維曲面 / 散布圖 / 熱圖。
ANOVA (變異數分析)	變異數分解表：逐變數的平方和、估模型平方和的百分比、係數。
Diagnostics (診斷)	四幅圖：係數路徑、CV 誤差、實際值對預測值、殘差對擬合值。
Glmnet Output (Glmnet 輸出)	原始模型輸出：所用隨機種子、模型打印輸出、選定的 λ 、 $\lambda_{min}/1se$ 、 γ （若使用鬆弛 lasso），以及完整的係數向量。
Report (報告)	報告設定與匯出欄位。
RCA Adjustments (RCA 調整)	RCA 分析結果（估價模式，計算完成後）。

3.6 設定的保存

glmnetUI 會自動把您的設定儲存到瀏覽器的本機儲存空間中，以輸入檔名作為索引鍵。當您重新載入同一個檔案時，所有設定都會還原：目標變數選擇、預測變數核取方塊、資料型別、預期符號、glmnet 參數以及交互項矩陣。上次使用的用途模式也會全域儲存，並在重新啟動應用程式時還原。有三個選項可供選擇：

- **Use last settings for input file**（使用該輸入檔案的上次設定）——還原為此檔案儲存的設定
- **Use default settings**（使用預設設定）——套用您儲存的全域預設值
- **glmnet defaults** (glmnet 預設值) ——重設為出廠預設值 ($\alpha = 1$ 、 $CV \lambda_{1se}$ 、10 摺、gaussian、開啟標準化)

Save current as default（將目前設定存為預設）按鈕可把目前的全部設定儲存為全域預設值。

3.7 深色模式

點選右上角的月亮 / 太陽圖示，即可在淺色與深色主題之間切換。主題偏好儲存在本機儲存空間中，跨工作階段保持。所有介面元素（表格、圖形、卡片、按鈕）都會透過 CSS 變數適配所選主題。

3.8 地區與區域設定

glmnetUI 透過基於國家的地區設定系統支援各國的數字與 CSV 格式慣例。標題列中的 **Settings**（設定）下拉式選單提供面向 31 個支援國家的 **Country**（國家）與 **Paper**（紙張）

選擇器。每個預設會設定：

- **CSV 分隔符號**——美國 / 英國 / 日本用逗號 (,)，歐洲多數國家用分號 (;)，因為那裡逗號被用作小數點。
- **小數點記號**——句點 (.) 或逗號 (,)。
- **千位分隔符號**——逗號 (美國 / 英國 / 日本)、句點 (德國 / 義大利 / 西班牙)、空格 (芬蘭 / 法國 / 波蘭 / 波羅的海國家 / 烏克蘭 / 俄羅斯)，或撇號 (瑞士)。
- **紙張尺寸**——Letter (美國 / 加拿大 / 墨西哥) 或 A4 (其他所有地區)。

3.8.1 儲存預設值

點選 **Save as my default** (存為我的預設值)，即可全域儲存您的地區設定偏好。這些預設值將適用於今後所有工作階段，無論您載入哪個資料檔案。

提示：圖形座標軸和斜率標註上的數字格式會自動適配您的地區設定。德語地區用句點作千位分隔符號 (200.000)，芬蘭用空格 (200 000)，瑞士用撇號 (200'000)。不顯示任何貨幣符號——*glmnetUI* 與幣別無關。

4 估價模式

當您把 Purpose 選為 **For Appraisal**（用於估價）時，glmnetUI 會針對單一不動產估值進行自我設定。第 3 章所述的全部功能依然可用；本章只講估價模式特有的附加內容。

4.1 標的列的處理

在估價模式下，資料集的第 1 列即估價標的，其餘各列均為可比成交案例。您的輸入檔案必須據此組織（參見第 2 章）。標的的成交價格可以留空或填任意值——glmnetUI 在擬合時會自動把它當作 NA 處理。

匯入之後，Data Preview 分頁會分為兩部分：**Subject Property**（估價標的，第 1 列）與 **Comparable Sales**（可比成交案例，第 2 列及以後）。第 1 列始終不參與模型擬合。擬合完成後，模型仍會為標的列產生預測值。

4.2 基準日與成交時距

在估價與市場模式下，變數設定小節中會出現一個 **Effective Date**（估價基準日）欄位（預設為當天日期）。如果您在 Special 欄下拉式選單中把某欄指定為 `contract_date`，glmnetUI 就會計算一個 `sale_age` 欄——即每筆成交的合約日期與基準日之間的整數天數。該欄會作為預測變數加入模型。

當基準日發生變化時，`sale_age` 會自動重新計算。

4.3 特殊欄位指定

在估價與市場模式下，變數設定表中每個預測變數都會出現一個 **Special**（特殊）下拉式選單。特殊型別及其作用的完整參考見第 6 章。

4.4 RCA 調整概覽

Calculate RCA Adjustments & Download（計算 RCA 調整值並下載）按鈕（側邊欄第 8 節，在估價與市場模式下擬合之後可見）會為每個可比案例計算相對於標的的市場推導調整值。完整的 RCA 工作流程見第 11 章。

提示：您賦予標的的 *CQA* 分數決定了殘差分布中有多大比例歸屬於標的。分數為 5.00 時，標的位於各可比案例的中位數處。

5 市場區域分析模式

當您把 Purpose 選為 **Market Area Analysis**（市場區域分析）時，glmnetUI 提供與估價模式相同的不動產專用功能（特殊欄位、成交時距、座標取整），但其取向是分析一組不動產，而非為單一標的估值。

5.1 與估價模式的差異

- 標的列處理更為靈活——在市場模式下，第 1 列不會被自動當作估價標的。
- 沒有銷售比較表格——「Calculate RCA Adjustments & Download」（第 8 節）步驟在市場模式下可用，但銷售比較表格步驟（第 9 節）僅限估價模式。

5.2 何時使用市場模式

在以下情形下適合採用市場區域分析模式：

- 為某個街廓或市場區域建立迴歸模型，以了解價值驅動因素
- 分析面積、屋齡、基地面積和位置等變數如何影響一組不動產的成交價格
- 為市場狀況分析或街廓界定準備佐證資料
- 處理某個資料集時，您需要特殊欄位功能（成交時距、座標取整），但不需要 RCA 調整工作流程

提示：當您需要特殊欄位功能但不需要 RCA 調整工作流程時，市場模式同樣適用於一般性的不動產迴歸分析。

6 變數選擇

側邊欄的第 4 節——**Variable Configuration**（變數設定）——是您選擇哪些欄位參與模型以及如何處理它們的地方。

6.1 目標變數

第 4 節頂部的 **Target (response) variable**（目標 / 反應變數）下拉式選單會列出資料集中的每一欄。請選擇一欄作為反應變數（例如成交價格）。目標欄會自動從預測變數清單中排除。

6.2 預測變數表

在目標變數選擇器下方，有一張表列出其餘每一欄，包含以下欄位：

欄位	說明
Variable （變數）	欄名
Type （型別）	資料型別下拉式選單：numeric、integer、character、Date、POSIXct
Include （納入）	核取方塊——把該欄作為預測變數納入模型
Factor （因子）	核取方塊——把該變數當作類別因子處理（在模型矩陣中產生虛擬變數欄）
Force （強制保留）	核取方塊——強制進入模型（罰則因子 = 0，絕不會被 lasso 剔除）
Special （特殊）	下拉式選單（僅估價 / 市場模式）——參見下文的特殊欄位型別參考
Sign （符號）	預期的係數符號：正、負或不限
NAs	遺漏值計數

6.3 資料型別偵測與手動覆寫

glmnetUI 在匯入時會自動偵測資料型別，能夠識別數值型、整數型和日期型的欄位。看起來像日期的字元欄（符合常見日期格式模式）會被歸類為 **Date**。

您可以透過修改 **Type** 下拉式選單來覆寫任何偵測結果。更改型別會影響該欄在模型矩陣中的編碼方式。

文字欄、TRUE/FALSE 欄和 R 因子欄會自動按類別變數處理。數值欄只有在您勾選其 **Factor** 核取方塊時才會變成因子——glmnetUI 絕不會根據數值的取值範圍推斷其為類別變數，因此像衛浴數量這樣的離散數值預測變數除非您另行指定，否則保持連續。Factor 設定透過 localStorage 持續保存。

6.4 預期符號與約束的施加

對每個預測變數，**Sign** 下拉式選單用於指定您預期其係數為正、為負還是無限方向。擬合之後：

- 符號與預期不符的係數會在係數表中被標記出來
- 符號警示會顯示在 Summary 分頁中

當 glmnet 參數（第 7 章）中啟用了 **Enforce Sign Constraints**（施加符號約束）核取方塊時，模型會被強制產生與預期符號一致的係數。這是透過 glmnet 中的 upper.limits 與 lower.limits 實現的。

6.5 特殊欄位型別參考

在估價與市場模式下，**Special** 下拉式選單提供以下選項：

Weighting (加權)：

- **weight** ——觀測權重欄（所有模式下均可用；只允許指定一欄；權重為 0 的資料列不參與擬合）

Date & Time Types (日期與時間型別)：

- **contract_date** ——觸發依據基準日自動計算 **sale_age**
- **listing_date** ——在計算在市天數時作為備用來源
- **dom** ——標識在市天數 (Days on Market) 欄
- **sale_age** ——指定一個已有的成交時距欄，這樣任何一欄都可以充當成交時距，而不必讓 glmnetUI 從 **contract_date** 計算

Monetary Types (貨幣型別)：

- **concessions** ——標識成交讓利（賣方補貼、買方誘因等）

Transaction Types (交易型別)：

- **sale_type** ——標識成交 / 交易型別欄

Size & Location Types (面積與位置型別)：

- **latitude** ——數值自動四捨五入到小數點後 3 位
- **longitude** ——與緯度作相同的取整處理
- **area** ——市場區域或街廓識別碼（例如 **area_id**、**neighborhood**、**market_area**）
- **living_area** ——啟用按平方英尺計的殘差計算 (**residual_sf** 與 **cqa_sf**)
- **lot_size** ——基地面積欄
- **site_dimensions** ——與基地面積歸為一組（例如「75x120」）

Age Types (屋齡型別)：

- **actual_age** ——不動產的屋齡欄
- **effective_age** ——不動產的有效屋齡

Display Types (僅顯示型別)：

- **display_only** ——該欄會包含在 Excel 匯出結果中，但完全不參與模型擬合。可用於地址欄位、MLS 編號、地號或其他參考資料。可以有多欄採用此指定。

提示：被指定為 *display_only* 的欄位仍保留在資料集中並出現在 Excel 匯出結果裡，但完全不進入預測變數表。可用於 ID 欄、地址或其他參考欄位。

7 參數選擇

側邊欄的第 5 節——**glmnet Call Parameters** (glmnet 呼叫參數) ——提供彈性網絡模型的全部設定選項。每個參數旁都有一個藍色說明圖示 (?), 滑鼠停留可看到說明。

7.1 Alpha (混合參數)

alpha 參數控制正則化的類型：

Alpha 取值	類型	行為
0	嶺迴歸	按比例收縮所有係數；絕不會把任何係數置為零
1	Lasso	可以把係數置為恰好為零 (變數選擇)
0 至 1	彈性網絡	嶺迴歸與 lasso 的混合

有兩種 alpha 選擇方法：

- **Fixed** (固定) ——透過滑桿設定單一 alpha 值 (預設：1.0)
- **Grid Search** (網格搜尋) ——測試多個 alpha 值，選出 CV 誤差最小的那個。可設定取值範圍與取值個數。

7.2 Lambda 的選擇

lambda 控制正則化的強度：

方法	說明
交叉驗證	推薦。使用 k 摺 CV 測試多個 lambda 值。
手動	若您有充分理由，可輸入指定的 lambda 值。

使用交叉驗證時，有兩種 lambda 取法：

- **lambda.1se** (預設) ——在最小值一個標準誤範圍內的最大 lambda。產生更簡潔、正則化更強的模型。估價工作推薦使用。
- **lambda.min** ——使交叉驗證誤差最小的 lambda。產生擬合最好但可能更複雜的模型。

Number of CV Folds (CV 摺數) 參數控制交叉驗證時資料的劃分方式 (預設：10 摺)。

7.3 Family (分布族)

為您的反應變數選擇分布族：

分布族	適用情形
gaussian	連續型反應 (例如成交價格)。最常用。
binomial	二元結果 (例如已售 / 未售)。
poisson	計數資料 (例如成交筆數)。

7.4 Standardize (標準化)

勾選時 (預設)，所有預測變數在擬合前都會被縮放為平均數 0、標準差 1。這確保罰則對所有預測變數一視同仁，而不受其原始尺度影響。係數會按原始尺度傳回。通常應保持開啟。

7.5 符號約束

啟用 **Enforce Sign Constraints** 後，係數會被約束為與變數表（第 6 章）中設定的預期符號一致：

- 「positive」（正）符號強制係數 ≥ 0
- 「negative」（負）符號強制係數 ≤ 0
- 設為「either」（不限）的變數不受約束

這是透過 `glmnet()` 中的 `upper.limits` 與 `lower.limits` 參數實現的。

7.6 鬆弛 Lasso

常規 lasso 對變數選擇和係數估計使用同一個罰則，這可能把重要係數過度收縮到零。鬆弛 lasso (relaxed lasso) 把這兩步分開：

1. 首先，透過 lasso 選擇變數（由罰則決定哪些係數得以保留）
2. 然後，對保留下來的變數以更小的罰則或不加罰則重新擬合，得到偏誤更小的估計

Gamma 滑桿控制鬆弛的程度：

Gamma	效果
0	完全鬆弛（OLS 重新擬合，完全不作收縮）
1	不作鬆弛（與常規 glmnet 相同）
介於其間	在鬆弛擬合與帶罰則擬合之間取折衷

在使用交叉驗證時，gamma 會被自動選取以獲得最佳表現。

7.7 隨機種子

Random seed（隨機種子）文字方塊預填了一個隨機整數，用於使交叉驗證的摺分配可重現——擬合前會呼叫 `set.seed()`，所用的種子會顯示在 Glmnet Output 分頁以及擬合狀態訊息中。每次擬合之後會自動產生一個新種子（使用系統時間，而非 R 的亂數產生器）。目前檔案最近使用的 5 個種子會以可點選連結的形式列出，便於快速調回；載入新檔案時種子歷史會重設。

7.8 交互項矩陣

一個可互動的上三角矩陣讓您控制哪些變數對之間允許產生交互作用。每個儲存格內有一個核取方塊——勾選表示交互項 $x1:x2$ 會被納入模型矩陣。

- **Allow All**（全部允許）——勾選所有交互變數對
- **Clear All**（全部清除）——取消勾選所有變數對（新檔案的預設狀態是全部不勾選）
- **點選變數名**——切換該變數的所有交互項

交互項的選擇會按輸入檔案分別儲存到瀏覽器 localStorage，並在您重新載入同一檔案時還原。

只要啟用了任何交互項，匯出的報告中就會包含一張 **Enabled Interactions**（已啟用交互項）矩陣（兩個座標軸都是預測變數，納入 $x1:x2$ 乘積項之處打勾），以橫向頁面呈現，與 earthUI 的 Allowed Interactions 矩陣相對應。請注意，lasso 仍可能把某個已啟用項的係數收縮為零。

提示：當預測變數很多時，lasso 會自動把不需要的交互項歸零。先用「Allow All」起步，再讓模型自行選擇，是一種合理的做法。

警告

啟用 glmnet 交互項會產生在法庭或稽核場合難以解釋和說明的調整值。請謹慎使用交互項，並核實所得調整值是否合理。

7.9 禁止進入主效應（僅作交互項）

在交互項矩陣中對某個變數名按右鍵，即可禁止它以主效應的身份進入模型。變數名後會出現一個粗體的「1」標記。被禁止時：

- 該變數會從主公式項中移除
- 該變數仍可出現在交互項 (x1:x2) 中
- 再次按右鍵即可解除禁止
- 該狀態按輸入檔案透過 localStorage 持續保存

一個典型用例：在為應隨不動產面積按比例變化的時間調整建模時，禁止 sale_age 進入主效應，只允許它與 living_area 產生交互。這樣得到的是 sale_age:living_area 交互項，而不是一個平坦的 sale_age 係數——即時間增值率隨房屋面積而異。

7.10 進階參數

glmnet 呼叫參數底部有一個可收合的「Advanced」（進階）小節，其中是一些附加設定。它們採用合理的預設值，但仍然可見，以便所有模型設定都能為法庭或稽核留下書面記錄。

- **Lambda Min Ratio** —— 正則化路徑中最小 lambda 與最大 lambda 之比。預設值：0.00001。對於小資料集（觀測數 <50），取較大的值（例如 0.01）可能表現更好。
- **Number of Lambda Values** —— 路徑中 lambda 值的個數。預設值：100。
- **CV Loss Metric** —— 交叉驗證所用的損失函數：MSE（gaussian 的預設值）、MAE（對離群值穩健）或 Deviance（偏差）。
- **Convergence Threshold** —— 座標下降的收斂閾值。預設值：1e-07。
- **Max Iterations** —— 座標下降的最大疊代次數。預設值：100,000。若出現收斂警告，可調大此值。
- **Fit Intercept** —— 是否包含截距（常數）項。預設：開啟。在估價 RCA 中，截距即「基準值」。

7.11 設定預設值

有三個選項控制載入檔案時參數如何初始化：

- **Use last settings for input file** —— 還原上次處理此檔案時所用的確切設定
- **Use default settings** —— 套用您儲存的全域預設值
- **glmnet defaults** —— 重設為出廠設定（alpha = 1、CV lambda.1se、10 摺、gaussian、開啟標準化）

點選 **Save current as default**，即可把目前全部設定存為今後檔案所用的全域預設值。

8 擬合模型

8.1 擬合按鈕

側邊欄的第 6 節包含 **Fit Glmnet Model** (擬合 Glmnet 模型) 按鈕。點選它即以您目前的設定執行模型。

8.2 擬合過程中發生了什麼

當您點選 Fit 時，glmnetUI 會：

1. 準備資料——剔除含 NA 的資料列，施加權重 (若已指定)，排除標的列 (估價模式下)，並把因子變數編碼為虛擬變數欄。
2. 建構模型矩陣——產生設計矩陣，含交互項矩陣中所有被允許的交互項。
3. 施加約束——若啟用了符號約束，則為係數設定 `upper.limits` 與 `lower.limits`。
4. 執行交叉驗證——若選擇了 CV，則呼叫 `cv.glmnet()` 尋找最佳 `lambda`。若啟用了鬆弛 lasso，則呼叫 `cv.glmnet(..., relax = TRUE)`。
5. 擬合最終模型——儲存擬合好的模型、係數、預測值與殘差。
6. 更新所有結果分頁——填入 Summary、Coefficients、Equation、Variable Importance、Contributions、ANOVA 與 Diagnostics。

Fit 按鈕下方的狀態訊息會顯示觀測數量、被排除的資料列數以及選定的 `lambda`。

9 結果分頁

9.1 Data Preview (資料預覽)

以互動式 DataTable 顯示匯入的資料。在估價 / 市場模式下，預覽分為兩張表：**Subject Property** (估價標的，第 1 列) 與 **Comparable Sales** (可比成交案例，第 2 列及以後)。

每個儲存格都保持單行顯示，並按欄寬截斷，因此較寬的自由文字欄位 (如不動產備註) 不會把資料列撐得很高。按兩下任一儲存格可彈出視窗，顯示其完整、未截斷的內容。

9.2 Equation (方程式)

顯示透過 MathJax 以 LaTeX 轉譯的擬合模型方程式。展示所有非零係數，並採用規範的數學排版：

- 正係數以 + 號顯示
- 負係數以 - 號顯示
- 交互項顯示為 $x_1 \times x_2$
- 數字帶千位逗號，並保留適當的小數位數

9.3 Correlation (相關性)

一張熱圖矩陣，展示所有數值型預測變數與反應變數之間的 Pearson 相關係數。資料匯入後立即可用——無需擬合模型。

- 色彩漸層：藍色 (-1) → 白色 (0) → 紅色 (+1)
- 每個儲存格中印出相關係數值
- 文字與座標軸尺寸隨變數數量自適應
- 有助於在擬合前識別多重共線性

9.4 Summary (摘要)

以卡片形式顯示的模型擬合統計量：

- R^2 ——被解釋變異數的比例
- $\text{Adj } R^2$ ——按預測變數個數調整後的 R^2
- GR^2 ——廣義 R^2 ，由 glmnet 在選定 lambda 處的偏差比得出
- $\text{CV } R^2$ ——交叉驗證 R^2 (使用 CV 時)
- **RMSE** ——均方根誤差
- **MAE** ——平均絕對誤差

在估價 / 市場模式下，還會顯示以下卡片：

- **Median Ratio** —— (預測值 / 實際值) 的中位數
- **COD** ——離散係數 (Coefficient of Dispersion，越低越好)
- **PRD** ——價格相關差異 (Price-Related Differential)

當訓練 R^2 超出 $\text{CV } R^2$ 達 0.1 以上時，會出現過擬合警示。

卡片下方有一張帶符號警示的係數表，為方便查閱而重複了 Coefficients 分頁的內容。

9.5 Coefficients (係數)

一張列出所有非零係數的表，顯示：

- 變數名
- 係數值
- 預期符號 (來自變數設定)
- 若係數方向與預期不符，則給出符號警示標記

9.6 Variable Importance (變數重要性)

標準化的係數量級：對每個預測變數計算 $|\beta| \times \text{sd}(x)$ ，因子變數則跨其虛擬變數欄彙總。有兩種顯示模式：

- 互動式 **plotly** (安裝了 **plotly** 套件時)：水平長條圖，滑鼠停留提示顯示精確的重要性、係數與相對百分比。
- 靜態 **ggplot2** (備用方案)：水平長條圖，座標軸標籤帶千位逗號。

圖表下方有一張 **DataTable**，顯示 **Variable**、**Importance**、**Coefficient** 與 **Relative %**。

9.7 Contributions (貢獻)

逐變數的偏效應圖，展示每個預測變數對模型預測值的貢獻。請從下拉式選單中選擇一個變數。

對數值型預測變數：一張該變數的 x 值對其貢獻 ($\beta \times x$) 的散布圖，附一條紅色線段和一個斜率標註。斜率標註顯示每單位的邊際效應 (例如 $+1,234.56/\text{unit}$)，對於緯度 / 經度這類取值範圍很小的變數，單位會自適應調整 (例如 $+0.12/0.001$)。副標題顯示截距 (基準) 值。

對因子變數：一張盒鬚圖，展示各因子水準上貢獻值的分布。

對雙變數交互項：三幅上下疊放的視覺化圖——聯合貢獻的三維曲面圖 (可用時為互動式 **plotly**，否則為靜態 **persp()**)、按貢獻著色的散布圖，以及跨兩個變數取值範圍的平均貢獻熱圖。

9.8 ANOVA (變異數分析)

變異數分解表，為每個預測變數顯示：

- **SS** ——平方和： $\sum (\text{contribution}_j - \overline{\text{contribution}_j})^2$
- **% of Model SS** ——佔模型總平方和的百分比
- **Coefficient** ——該預測變數的主導係數

表中含一列截距和一列 **TOTAL MODEL** (模型合計)。交互項 (名稱中含「:」者) 作為單獨的組顯示。

9.9 Diagnostics (診斷)

四幅診斷圖，採用大字級 (基準字級 16pt)、15 個座標軸刻度，標籤帶千位逗號 (不使用科學記號)：

1. 係數路徑 (Coefficient Path) ——展示隨著 $\log(\lambda)$ 增大，所有係數如何變化。曲線按變數著色。有助於直觀理解正則化路徑與變數選擇過程。
2. **CV 誤差** (CV Error) ——交叉驗證誤差隨 $\log(\lambda)$ 的變化。散點為 CV 誤差平均數；誤差槓為 ± 1 個標準誤。虛線垂直線標出 $\lambda_{\min}$ (藍色) 與 λ_{1se} (綠色)。
3. 實際值對預測值 (Actual vs Predicted) ——觀測值與預測值的散布圖。散點應聚集在 45 度虛線參照線附近。散布越寬，說明預測誤差越大。
4. 殘差對擬合值 (Residuals vs Fitted) ——殘差與擬合值的散布圖。應呈現圍繞零線 (虛線) 的隨機散布。若出現某種模式，則提示模型設定有誤 (例如非線性、異質變異數)。

每幅圖都有一個 **Download PNG** 按鈕，可按 150 DPI 儲存。

9.10 Glmnet Output (Glmnet 輸出)

原始模型輸出，顯示：所用隨機種子、模型打印輸出、選定的 λ 、 $\lambda_{\min}/\lambda_{1se}$ 、 γ (若啟用鬆弛 **lasso**)，以及完整的係數向量。

9.11 Report (報告)

設定報告欄位（估價師姓名、不動產地址、報告日期、檔案編號），並可直接從該分頁匯出為 Word (.docx) 或 PDF。完整的基於 Quarto 的報告工作流程（HTML、Word 或 PDF）見第 13 章。

10 下載資料

擬合之後，**Download Output (Excel)**（下載輸出，Excel）按鈕（側邊欄第 7 節，在所有用途模式下均可用）會把含預測值與診斷資訊的 Excel 檔案匯出到專案的輸出資料夾。該小節標題在估價與市場模式下為「Download Estimated Sale Prices & Residuals」（下載估計成交價格與殘差），在通用模式下為「Download Estimated Target Variable(s) & Residuals」（下載估計目標變數與殘差）。

10.1 輸出欄位

欄位	說明
est_<target>	模型預測值（例如 est_sale_price）
residual	實際值 - 預測值
cqa	狀況- 品質- 吸引力分數（Condition-Quality-Appeal，0-10 分制）
residual_sf	殘差 / 居住面積（若已指定 living_area）
cqa_sf	依據 residual_sf 排名計算出的 CQA
<var>_contribution	各變數對預測值的貢獻
basis	截距值（所有資料列相同）

成功完成後，下載按鈕上會出現一個白色勾號。

10.2 CQA 分數

CQA 把每一資料列的殘差與其餘各列相比較，按 0-10 分制排名：

- **CQA 高（~9-10）**：成交價遠高於預測值——很可能狀況 / 品質 / 吸引力更優
- **CQA 低（~0-1）**：成交價遠低於預測值——很可能狀況較差或屬不良資產處分
- **CQA ~5**：接近中位數

在估價 / 市場模式下，若指定了 living_area 欄，各資料列會按 residual_sf 遞減排序。

11 RCA 計算與下載

RCA (Residual Constraint Approach, 殘差約束法) 工作流程在 **Appraisal** (估價) 與 **Market Area Analysis** (市場區域分析) 模式下、模型擬合之後可用。後續的銷售比較表格 (第 12 章) 僅限估價模式。

11.1 開啟 RCA 對話方塊

點選側邊欄第 8 節的 **Calculate RCA Adjustments & Download** 按鈕。然後選擇：

- 分數類型：**CQA** 或 **CQA per SF** (每平方英尺 CQA)。若選擇「CQA per SF」，標的的殘差會按居住面積縮放。
- 輸入標的的 CQA 分數 (0.00–9.99，預設 5.00)

11.2 CQA 分數內插

1. 對各可比案例的 CQA 分數與殘差進行排序
2. 用線性內插把您給定的 CQA 值對應到一個殘差
3. 標的價值 = 模型預測值 + 內插所得殘差

11.3 輸出欄位

欄位	說明
subject_value	模型預測值 + 內插所得殘差
<var>_adjustment	標的貢獻 – 可比案例貢獻
residual_adjustment	標的殘差 – 可比案例殘差
net_adjustments	所有調整值之和
gross_adjustments	各調整值絕對值之和
adjusted_sale_price	可比案例成交價格 + 淨調整額

計算成功後，按鈕上會出現一個白色勾號。

12 銷售比較表格

銷售比較表格 (Sales Comparison Grid) 僅在 **Appraisal** (估價) 模式下、計算完 RCA 調整值 (第 8 步) 之後可用。它會產生一份格式化的 Excel 活頁簿，可直接編入估價報告。

12.1 可比案例選擇對話方塊

點選側邊欄第 9 節的 **Generate Sales Grid & Download** 按鈕，會彈出一個強制回應對話方塊，列出 RCA 輸出中的所有可比成交案例。可比案例分為兩組：

- **Recommended comps** (推薦的可比案例，預先勾選)：毛調整額 < 成交價格的 25%，按成交時距排序 (最新在前)。
- **Other comps** (其他可比案例，預設不勾選)：其餘的可比案例，按成交時距排序。

最多可選 30 個可比案例，然後點選 **Generate Sales Grid** 產生活頁簿。

12.2 活頁簿版面

產生的活頁簿最多含 10 個工作表，每個工作表並排放置 3 個可比案例：

- 表頭列：標的與可比案例的地址、成交價格、成交日期以及在市天數
- 分組列：位置 (緯度、經度、區域)、基地面積 (基地面積、基地尺寸) 與屋齡 (實際屋齡、有效屋齡)，帶組標題與子項目
- 模型變數列：每個非零模型係數一列，顯示標的貢獻、各可比案例貢獻以及調整百分比
- 剩餘特徵列：留空且未鎖定的儲存格，供估價師填寫模型未涵蓋的特徵 (例如狀況、品質、景觀)
- 彙總列：淨調整百分比、毛調整百分比與調整後成交價格，均帶有即時的 Excel 公式

12.3 工作表保護與剩餘特徵儲存格

每個工作表都設有密碼保護以防誤改，但剩餘特徵的取值儲存格是明確解鎖的。這樣估價師既能填入模型之外的特徵取值，又能保留由公式驅動的調整計算。

13 下載報告

13.1 兩階段 Quarto 工作流程

從側邊欄匯出報告採用兩階段的 Quarto 工作流程：

1. **Generate Quarto Report**（產生 Quarto 報告，側邊欄第 10 節）——把一個自足的 Quarto 報告包（.qmd 原始檔加上圖形、report_data.rds 與 reference.docx）寫入專案的 glmnet 輸出資料夾。
2. **Convert Quarto Report**（轉換 Quarto 報告，側邊欄第 11 節）——把 .qmd 檔案轉譯為您選擇的輸出格式。路徑預設指向最近產生的報告包，但您也可以瀏覽選擇任何一個 .qmd。

13.2 報告格式

轉換步驟中有三種格式可選：

- **HTML** ——自足的 HTML 文件。在所有平台上都能使用，除 pandoc 外無需額外軟件。推薦的預設選項。
- **Word (.docx)** ——格式化文件，便於編輯與分送。在所有平台上均可使用。
- **PDF** ——透過 Quarto 配合 LaTeX 產生。紙張尺寸遵循地區設定（Letter 或 A4）。需要安裝 LaTeX（參見系統需求附錄）。若未偵測到 LaTeX，格式核取方塊中會自動隱藏 PDF 選項。

有 Quarto 可用時使用 Quarto 產生報告，否則改用 rmarkdown。Report 分頁另外還提供直接匯出為 Word 和匯出為 PDF 的按鈕。

13.3 報告內容

報告涵蓋所有分頁的內容：

- 資料集說明（檔名、觀測數、反應變數、預測變數、用途）
- 模型設定（alpha、lambda、family、standardize、是否鬆弛、非零係數）
- 結果摘要（ R^2 、調整 R^2 、廣義 R^2 、CV R^2 、RMSE、MAE）
- 模型方程式（以 LaTeX 轉譯）
- 係數表
- 變數重要性（長條圖與表格）
- 與 Contributions 分頁一致的逐變數貢獻圖：
 - 數值型預測變數：帶線性擬合線的散布圖
 - 因子變數：按水準分組的盒鬚圖
 - 交互項：按貢獻著色的散布圖、平均貢獻熱圖，以及靜態三維曲面快照（persp）
- Enabled Interactions（已啟用交互項）矩陣（啟用了任何交互項時）——兩個座標軸都是預測變數，納入 $x_1:x_2$ 乘積項之處打勾，以橫向頁面呈現
- 相關係數矩陣熱圖
- ANOVA 分解表
- 診斷圖：
 - 係數路徑
 - 交叉驗證誤差
 - 實際值對預測值
 - 殘差對擬合值
 - 常態 Q-Q 圖
- 模型輸出

所有座標軸標籤與色彩圖例中的數字均帶千位逗號（不使用科學記號）。
報告產生成功後，報告按鈕上會出現一個白色勾號；轉譯過程中會顯示已用時間計時器。

14 與 earthUI 的比較

earth 應用程式及其 glmnet 後處理模組同屬一套迴歸建模工具。它們共用相同的資料格式、特殊欄位型別、RCA 工作流程和示範資料集，但使用不同的建模引擎。

14.1 主要差異

特性	glmnetUI	earthUI
方法	彈性網絡 (glmnet)	MARS/Earth (earth)
變數選擇	經由 lasso 罰則自動進行	經由前向 / 後向修剪自動進行
係數關係	僅限線性效應	分段線性 (鉸鏈函數)
交互項	經由交互項矩陣兩兩組合	最高 3 階
係數符號	可透過約束強制施加	無法直接強制施加
鬆弛擬合	可使用鬆弛 lasso	不適用
Alpha 調校	可用網格搜尋	不適用 (無混合參數)
銷售比較表格	可用 (第 9 步)	可用 (第 8 步)
g 函數圖	不適用	歸組項視覺化
報告格式	HTML、Word、PDF	HTML、Word、PDF

14.2 共有的功能

兩個工具都提供：

- CSV 與 Excel 匯入，並把欄名轉換為 snake_case
- 支援 31 個國家的地區設定，含 CSV 分隔符號、小數點記號與紙張尺寸
- 特殊欄位指定 (contract_date、living_area、area、latitude、longitude 等)
- 依據基準日自動計算 sale_age
- 深色 / 淺色模式切換
- 透過 localStorage 持續保存設定
- 相關性熱圖、變數重要性、貢獻圖、ANOVA、診斷圖
- CQA 評分與 RCA 調整工作流程
- 帶可比案例選擇和格式化 Excel 輸出的銷售比較表格
- 相容同一批示範資料集 (Appraisal_1.csv)

14.3 何時該用哪一個

- 使用 glmnetUI，當您想要一個簡潔、可解釋、並能自動選擇變數的線性模型時。當預測變數很多而您希望由模型挑出最重要的那些時，彈性網絡尤其合適。
- 使用 earthUI，當您預期存在非線性關係（例如居住面積對價格的影響在不同面積區間有所不同），或需要用鉸鏈函數刻畫閾值效應時。earth 模型還能刻畫更高階的交互作用（2 階和 3 階）。此外，earthUI 還提供一項基於 Prolog/DCG 的實驗性備註擷取功能，這在 glmnetUI 中沒有對應項。
- 兩者並用，以在同一份資料上比較線性模型與非線性模型。若兩者給出的 R^2 相近，出於可解釋性考量，可能更宜採用較簡潔的 glmnet 模型。

15 示範資料集：Appraisal_1.csv

15.1 說明

示範資料集隨 glmnetUI 一同發布（並與 earthUI 共用）。可用以下方式定位：

```
demo_file <- system.file("extdata", "Appraisal_1.csv", package = "glmnetUI")
```

該檔案包含 1,502 筆住宅成交紀錄（外加第 1 列的 1 筆估價標的），取自一份模擬的 MLS 匯出檔案。資料代表一個多區域市場中的獨棟住宅成交，涵蓋不同的不動產面積、屋齡與位置。

這並非真實資料，但以北加州一個真實街廓為原型。所有身分識別資訊均已更改或移除。

15.2 各欄位

欄位	型別	特殊型別	說明
weight	numeric	weight	觀測權重（0 = 不參與擬合）
id	numeric	display_only	內部紀錄 ID
property_id	numeric	display_only	MLS 不動產識別碼
listing_id	character	display_only	MLS 掛牌編號
parcel_number	character	display_only	縣估稅官地號（APN）
street_address	character	display_only	不動產地址
city_name	character	display_only	城市
postal_code	character	display_only	郵遞區號
county_name	character	display_only	縣
contract_date	Date	contract_date	成交合約日期（用於計算 sale_age）
sale_age	numeric	—	從合約日期到基準日的天數
sale_price	numeric	（目標變數）	成交價格——反應變數
living_sqft	numeric	living_area	總居住面積（平方英尺）
beds_total	integer	—	臥室數量
baths_total	numeric	—	衛浴總數（例如 2.5 = 2 套全衛浴 + 1 套半衛浴）
lot_size	numeric	lot_size	基地面積（平方英尺）
area_id	integer	area	MLS 區域識別碼
age	numeric	actual_age	不動產屋齡（年）
latitude	numeric	latitude	緯度（入模前四捨五入到小數點後 3 位）
longitude	numeric	longitude	經度（入模前四捨五入到小數點後 3 位）
garage_spaces	integer	—	車庫車位數
days_on_market	integer	dom	在市天數
listing_date	Date	listing_date	掛牌日期
sale_concessions	numeric	concessions	賣方讓利

15.3 建議的快速上手流程

1. 啟動 glmnet 常式：earthUI::launch_glmnet()
2. 新建一個用途為 **For Appraisal** 的專案（第 1 節）
3. 把 Appraisal_1.csv 放入專案的輸入資料夾，並在第 2 節中選中它（必要時點選 **Refresh**）
4. 選擇 sale_price 作為目標變數
5. 按上表所示指定各欄位的特殊型別

6. 納入以下預測變數：sale_age、living_sqft、baths_total、lot_size、area_id（作為因子）、age、latitude、longitude、garage_spaces
7. 保持 $\alpha = 1$ (Lasso)，使用 CV 配 lambda.1se
8. 點選 **Fit Glmnet Model**
9. 檢視 Summary、Coefficients 與 Diagnostics 分頁
10. 下載輸出（第 7 步），檢視 CQA 排名
11. 以約 ~5.00 的 CQA 分數計算 RCA 調整值（第 8 步）
12. 用推薦的可比案例產生銷售比較表格（第 9 步）

16 系統需求與疑難排解

16.1 支援的平台

glmnetUI 可在 **macOS**、**Windows** 與 **Linux** 上執行。強烈建議使用 RStudio Desktop (2023.06 及以上版本)——它內建產生報告所需的 **pandoc**。

16.2 選用的相依套件

- **LaTeX** (TinyTeX、MiKTeX 或 MacTeX)：僅用於產生 PDF 報告。若未偵測到，PDF 選項會自動隱藏。安裝方式：`tinytex::install_tinytex()`
- **sysfonts + showtext**：為圖形提供 Roboto Condensed 字型。若不可用或處於離線狀態，會自動改用系統無襯線字型。
- **earth** ($\geq 5.3.0$)：僅在使用 earthUI 匯入管道時需要。

16.3 各平台說明

macOS：問題最少。安裝 TinyTeX 以產生 PDF：`tinytex::install_tinytex()`

Windows：與 RStudio 配合良好。公司配發或權限受限的機器可能對暫存目錄有限制——應用程式會在主控台給出警告，但仍會繼續執行，只是設定不能持續保存。

Linux：可能需要安裝系統程式庫：`sudo apt install libcurl4-openssl-dev libssl-dev libxml2-dev libsqlite3-dev libfontconfig1-dev`

16.4 功能降級

缺少的元件	表現
沒有 LaTeX	隱藏 PDF 選項。HTML 與 Word 仍可使用。
沒有網路 / 字型載入失敗	使用系統無襯線字型。主控台記錄相應訊息。
沒有 RSQLite / 檔案系統唯讀	設定不能持續保存。應用程式照常執行。
暫存目錄不可寫入	報告產生失敗並給出明確錯誤。請把 TMPDIR 設為可寫入的位置。

16.5 疑難排解

「PDF option not available」(PDF 選項不可用)：在 R 中執行 `tinytex::install_tinytex()`，然後重新啟動應用程式。

「Settings will not persist」(設定無法保存)：檢查使用者資料目錄的權限 (macOS：`~/Library/Application Support/R/glmnetUI/`，Linux：`~/.local/share/R/glmnetUI/`，Windows：`%APPDATA%/R/data/glmnetUI/`)。

連接埠 7879 已被佔用：執行 `lsof -ti:7879 | xargs kill` (macOS/Linux) 或 `Stop-Process -Id (Get-NetTCPConnection -LocalPort 7879).OwningProcess` (Windows PowerShell)。

參考文獻

- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent". In: *Journal of Statistical Software* 33.1, pp. 1–22. DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Simon, Noah et al. (2011). "Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent". In: *Journal of Statistical Software* 39.5, pp. 1–13. DOI: [10.18637/jss.v039.i05](https://doi.org/10.18637/jss.v039.i05).

Tay, J. Kenneth, Balasubramanian Narasimhan, and Trevor Hastie (2023). “Elastic Net Regularization Paths for All Generalized Linear Models”. In: *Journal of Statistical Software* 106.1, pp. 1–31. DOI: [10.18637/jss.v106.i01](https://doi.org/10.18637/jss.v106.i01).

mgcvUI : mgcv (廣義加法模型) 套件的互動式介面

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。

William Bert Craytor

第 1 卷，第 1 期 · 2026 年夏季

發表日期：2026 年 7 月 1 日

DOI: [10.67330/xtv1pa25](https://doi.org/10.67330/xtv1pa25)

© 2026 William Bert Craytor. Published by Pacific Vista Net.

本開放取用文章依據 [創用 CC 姓名標示 4.0 國際授權條款 \(CC BY 4.0\)](#) 的條款發布。

摘要。 mgcvUI 是 R 語言 mgcv 套件的圖形使用者介面 (graphical user interface)，該套件使用帶自動平滑度選擇 (smoothness selection) 的懲罰迴歸樣條 (penalized regression spline) 來擬合廣義加法模型 (generalized additive model, GAM)。它提供三種用途模式——通用預測建模、不動產估價與市場區域分析——並引導使用者完成資料匯入、平滑項設定、模型擬合、診斷圖與效應圖檢視以及報告下載。本文記錄了 mgcvUI 的資料格式要求、建模工作流程、輸出顯示介面以及完整的功能參考。

關鍵詞： 廣義加法模型，GAM，懲罰樣條，平滑化，mgcv，R，圖形使用者介面，不動產估價

出版者按 (2026 年 7 月)：付印前不久，mgcvUI 已併入 earthUI (0.11.0 及以後版本)，本文所述的 GAM 工作流程現與 MARS 和彈性網絡工作流程一同由 earthUI 提供。mgcvUI 仍以其最終發布形態 (0.3.0) 保留在 <https://github.com/wcraytor/mgcvUI>；後續開發將在 earthUI 中繼續。本文適用於已發布的 mgcvUI。

目錄

1	引言	4
1.1	歷史與背景	4
1.1.1	廣義加法模型	4
1.1.2	mgcv 套件	4
1.1.3	與 MARS (earth) 和彈性網絡 (glmnet) 的比較	4
1.1.4	earthUI-mgcvUI 流程	5
1.1.5	建議的工作流程	5
1.2	什麼是 mgcvUI ?	5
1.3	三種用途模式	6
1.4	不動產專用功能	6
1.5	入門	6
2	MLS 輸入資料要求	8
2.1	檔案格式與結構	8
2.2	估價模式的必需欄位	8
2.3	特殊欄位命名約定	9
2.4	資料品質與完整性	9
2.5	標的列的位置	9
3	通用模式	10
3.1	概述	10
3.2	側邊欄工作流程	10
3.3	從 earthUI 匯入 (選用)	11
3.4	主面板頁籤	11

3.5	設定的持久化	12
3.6	深色模式	12
3.7	區域與地區設定	12
4	估價模式	14
4.1	標的列的處理	14
4.2	估價基準日與銷售時距	14
4.3	特殊欄位指定	14
4.4	RCA 調整概覽	14
5	市場區域分析模式	15
5.1	與估價模式的差異	15
5.2	何時使用市場模式	15
6	變數選擇	16
6.1	目標變數	16
6.2	反應變數變換	16
6.3	預測變數表	16
6.4	資料型別偵測與覆寫	17
6.5	特殊欄位類型參考	17
7	參數選擇	18
7.1	參數預設組	18
7.2	Family (分配族)	18
7.3	Method (方法)	18
7.4	Gamma (平滑度乘子)	18
7.5	交叉驗證	19
7.6	Select (收縮)	19
7.7	預設基底	19
7.8	預設 k (基底維度)	19
7.9	張量類型	19
7.10	允許的交互	20
7.11	因子-平滑交互	20
7.12	進階參數	20
8	擬合模型	21
8.1	Fit 按鈕	21
8.2	擬合過程中發生了什麼	21
8.3	背景擬合與進度顯示	21
9	結果頁籤	22
9.1	Data (資料)	22
9.2	Equation (方程)	22
9.3	Summary (摘要)	22
9.4	Variable Importance (變數重要性)	22
9.5	貢獻圖	22
9.5.1	單變數平滑項 — $s(x)$	23
9.5.2	因子-平滑項 — $s(x, \text{by} = \text{factor})$	23
9.5.3	交互項 — $ti(x_1, x_2)$ 與 $te(x_1, x_2)$	23
9.5.4	參數項 — 線性項與因子項	23
9.6	Correlation (相關性)	23
9.7	Diagnostics (診斷)	23

9.8	RCA Adjustments (RCA 調整)	23
9.9	Anova (變異數分析)	24
9.10	Mgcv Output (mgcv 輸出)	24
9.11	Sign Check (符號檢查)	24
9.12	Concurvity (共曲性)	24
9.13	Basis Check (基底檢驗)	24
10	下載資料	25
10.1	輸出欄位	25
10.2	CQA 得分	25
11	RCA 計算與下載	26
11.1	開啟 RCA 對話方塊	26
11.2	CQA 得分內插	26
11.3	輸出欄位	26
12	銷售比較網格	27
12.1	可比案例選擇對話方塊	27
12.2	活頁簿版式	27
12.3	工作表保護與殘差儲存格	27
13	下載報告	28
13.1	第 1 步：產生 Quarto 報告	28
13.2	第 2 步：轉換 Quarto 報告	28
13.3	報告內容	28
14	匯出模型函數	29
14.1	R 函數	29
14.2	Python 函數	29
14.3	C++ 標頭檔	29
14.4	SQLite 資料庫	29
15	與 earthUI 和 glmnetUI 的比較	30
15.1	主要差異	30
15.2	共有功能	30
15.3	建議的工作流程	30
16	示範資料集：Appraisal_1.csv	31
16.1	說明	31
16.2	各欄位	31
16.3	建議的快速上手步驟	32
17	系統需求與疑難排解	33
17.1	受支援的平台	33
17.2	R 版本	33
17.3	必需的 R 套件	33
17.4	選用相依套件	33
17.5	與平台相關的說明	33
17.5.1	macOS	33
17.5.2	Windows	34
17.5.3	Linux	34
17.6	優雅降級	34
17.7	疑難排解	34

17.7.1 「PDF option is not available」(PDF 選項不可用)	34
17.7.2 「Settings will not persist」(設定將不會儲存)	35
17.7.3 「Diagnostics plot unavailable」(診斷圖不可用)	35
17.7.4 「Report error: number of items to replace...」(報告錯誤：待替換的項數)	35
17.7.5 7880 連接埠已被佔用	35
17.7.6 在 Linux 上安裝套件失敗	35

1 引言

1.1 歷史與背景

1.1.1 廣義加法模型

廣義加法模型 (generalized additive model, GAM) 由 Trevor Hastie 與 Robert Tibshirani 在其 1990 年的專著中正式提出。GAM 對線性模型作了推廣：把每一個線性項 $\beta_j x_j$ 替換為一個平滑函數 (smooth function) $f_j(x_j)$ ，於是模型成為：

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_p(x_p) + \varepsilon$$

每個 f_j 都用懲罰迴歸樣條 (penalized regression spline) 從資料中估計得到。這些平滑函數使模型無需使用者事先指定函數形式即可刻畫非線性關係。懲罰項控制每個平滑項的擺動程度 (wiggliness)，在保留足夠彈性以追蹤真實模式的同時防止過擬合。

1.1.2 mgcv 套件

R 套件 `mgcv` (Mixed GAM Computation Vehicle) 由巴斯大學的 Simon Wood 開發 (Wood 2017)。它於 2001 年首次發布，如今已是 GAM 在 R 中的標準實作，並隨每一份 R 安裝一同發行。其底層方法見於一系列論文，涵蓋穩定的多平滑參數估計 (Wood 2004)、薄板迴歸樣條 (Wood 2003)、快速穩定的 REML 估計 (Wood 2011) 以及一般平滑模型的選擇 (Wood, Pya, and Säcken 2016)。主要特性包括：

- 自動平滑度選擇 (smoothness selection)：每一項的平滑參數都作為模型擬合的一部分加以估計，可採用 REML (受限最大概似)、GCV (廣義交叉驗證) 或 ML (最大概似)。
- 多種基底類型：薄板迴歸樣條 (tp)、三次迴歸樣條 (cr)、P 樣條 (ps) 與 B 樣條 (bs)，各具不同性質。
- 張量積平滑：`te()` 與 `ti()` 可實現變數間的平滑交互，而不受維度詛咒之困。
- 變數選擇：`select = TRUE` 選項會附加一個額外懲罰，可將平滑項收縮至零，從而實際完成變數選擇。
- 共曲性診斷 (concurvity diagnostics)：度量平滑項中與多重共線性相對應的現象。

1.1.3 與 MARS (earth) 和彈性網絡 (glmnet) 的比較

`earth` 應用程式及其兩個後處理模組 (`glmnet` 與 `mgcv`) 使用不同的建模引擎。下表概括了主要差異：

特性	mgcv (GAM)	earth (MARS)	glmnet (彈性網絡)
模型類型	平滑非線性 (懲罰樣條)	分段線性 (自適應樣條)	線性 (帶正則化)
非線性	透過平滑函數自動實現 (薄板、三次等)	透過帶資料驅動節點的鉸鏈函數自動實現	僅能透過交互項或基底展開實現
變數選擇	選用, 透過雙重懲罰收縮實現 (select=TRUE)	內建, 透過前向/後向逐步法配合 GCV 修剪實現	內建, 透過 L_1 懲罰 (Lasso) 實現
交互項	張量積平滑 (te()、ti())	基於鉸鏈函數的乘積 (可解釋, 最高三階)	兩兩交叉乘積 (難以解釋)
可解釋性	平滑的偏效應曲線; 非常直觀	g 函數按變數給出分段線性偏效應	係數即線性權重; 易於逐一說明
最適用於	平滑的非線性關係、彈性建模	自動偵測非線性與交互	高維資料、變數選擇、小樣本
RCA 調整	透過偏效應給出平滑調整	透過 g 函數給出分段線性調整	按變數給出線性調整
過擬合控制	REML/GCV 平滑參數估計	基於 GCV 的修剪	交叉驗證選取的 lambda

MARS (Multivariate Adaptive Regression Splines, 多元自適應迴歸樣條) 由 Jerome Friedman 於 1991 年提出。它透過從資料中自適應地選擇鉸鏈函數及其節點 (knot) 位置來建構分段線性模型。其 R 實作是 Stephen Milborrow 編寫的 **earth** 套件。

彈性網絡 (elastic net) 迴歸由 Friedman、Hastie 與 Tibshirani (2010) 實作於 **glmnet** 套件中, 它把 Lasso (L_1) 懲罰與嶺迴歸 (L_2) 懲罰結合起來, 同時完成變數選擇與係數收縮。

1.1.4 earthUI-mgcvUI 流程

一套強有力的工作流程把 earthUI 的自動節點發現與 mgcvUI 的平滑估計結合起來。當 earthUI 匯出一個 .rds 結果檔案而 mgcvUI 將其匯入時:

1. **earth** 發現節點 —— MARS 為每個變數找出由資料驅動的鉸鏈位置 (變點)。
2. **mgcvUI** 以其為錨點 —— earth 的節點成為 GAM 中三次迴歸樣條 (cr 基底) 的基礎, 使平滑項一開始就擁有反映資料真實結構的節點位置。
3. **mgcv** 精化擬合 —— 懲罰樣條估計把 earth 的分段線性擬合加以平滑化, 在保留 earth 所發現的整體形狀的同時, 給出平滑的偏效應曲線。

這條流程把探索性建模 (earth) 與驗證性建模 (GAM) 銜接起來, 兼取兩種方法之長。

1.1.5 建議的工作流程

對於不動產估價及其他要求模型可解釋、可辯護的類似應用:

1. 先用 **earthUI** 發現資料中的重要變數、非線性關係與交互作用。earth 的自動基底選擇與 g 函數能提供極好的初步洞見。
2. 再用 **mgcvUI** 精化, 如果你希望得到更平滑的偏效應。把 earthUI 的節點匯入 mgcvUI, 即可從分段線性模型順暢過渡到平滑模型。
3. 使用 **glmnetUI**, 當你需要激進的變數選擇 (Lasso)、預測變數數目接近或超過觀測數目, 或者你希望得到一個純線性且帶正則化、便於辯護的模型時。

1.2 什麼是 mgcvUI?

mgcvUI 是 R 語言 mgcv 套件的圖形使用者介面 (graphical user interface)。它以本機 Shiny 應用程式的形式執行——無需登入、無需伺服器、也沒有帳號。你從 R 中啟動它, 匯入資料集 (CSV 或 Excel), 設定模型, 然後互動式地擬合 GAM。

自目前版本起，mgcvUI 不再是單獨安裝的套件：它隨 earthUI 套件一同發行，作為其中的 GAM 後處理模組（Post Processing Module），用 `earthUI::launch_mgcv()` 啟動。應用程式本身——以及本文所述的一切——並無改變；改變的只是安裝與啟動指令。三個例程（`earth`、`glmnet`、`mgcv`）共用一棵專案樹、一個設定資料庫以及同一份支撐基礎設施。

該應用程式提供完整的工作流程：資料匯入、含平滑項設定的變數設定、帶背景處理的模型擬合、診斷圖、平滑偏效應曲線，以及可下載的 Word、PDF 或 HTML 格式報告。

廣義加法模型把每個預測變數的線性關係 βx 替換為平滑函數 $f(x)$ 。這意味著：

- 無需指定函數形式——每種關係的形狀由資料本身決定。
- 平滑的偏效應——每個預測變數的貢獻是一條平滑曲線而非直線，便於看清某變數在不同取值處如何影響反應變數。
- 自動平滑度選擇——mgcv 透過 REML 或 GCV 確定每條曲線應有多擺動，在擬合優度與過擬合之間取得平衡。

1.3 三種用途模式

每個專案都有一個用途（purpose），在側邊欄第 1 節（**Project**，專案）中建立專案時選定，此後即被鎖定。目前作用中專案的用途決定哪些工具與介面元素出現；若要在別的模式下工作，請建立或開啟一個具有該用途的專案。三種模式為：

- **General**（通用）——適用於任何類型母體或資料集的 GAM 迴歸。這是預設模式，提供完整的 GAM 工作流程，不含任何領域專用的附加功能。
- **For Appraisal**（用於估價）——為不動產估價量身打造的 GAM 迴歸。增加了單一不動產估價所特有的功能，包括估價標的處理、特殊欄位指定以及殘差約束法（Residual Constraint Approach, RCA）。
- **Market Area Analysis**（市場區域分析）——為市場區域研究量身打造的 GAM 迴歸。增加了在既定市場中分析成組不動產的功能，並可選擇處理標的列。

在三種模式下，核心建模引擎完全相同——你始終是在透過 `mgcv::gam()` 擬合 GAM。用途設定只控制哪些附加工具與介面元素可用。

1.4 不動產專用功能

當選定 **For Appraisal** 或 **Market Area Analysis** 時，mgcvUI 會啟用若干為不動產分析而設計的功能：

- **特殊欄位指定**——每個預測變數都可以被標記上某種特殊角色，例如 `contract_date`、`dom`、`concessions`、`latitude`、`longitude`、`living_area`、`lot_size`、`actual_age`、`effective_age`、`area`、`site_dimensions` 或 `display_only`。這些指定決定該欄位在擬合與輸出過程中的處理方式。
- **Sale Age 欄位**（銷售時距）——當某欄位被指定為 `contract_date` 且提供了 `Effective Date`（估價基準日）時，mgcvUI 會計算一個 `sale_age` 欄位（成交日期與基準日之間的天數），並以它作為預測變數代入模型。
- **RCA 計算**（僅估價模式）——在估價模式下，模型擬合完成後 mgcvUI 可計算殘差約束法（RCA）輸出，給出逐個可比案例的調整額、淨調整與總調整彙總，以及估價標的的調整後銷售價格。

提示：*General* 用途模式適用於任何資料集——金融、科學、工程或不動產。估價模式與市場模式只是為不動產專業人士增添了一些便利功能。

1.5 入門

使用 mgcvUI 的步驟：

1. 在 R 中安裝 **earthUI**：`install.packages("earthUI")` ——**mgcv** 例程已包含在內。(開發版與選用元件請參見本期 **earthUI** 一文中的「Installing earthUI and vProlog」一節。)
2. 啟動應用程式：在 R 中執行 `earthUI::launch_mgcv()`。應用程式會在你的網頁瀏覽器中透過 7880 連接埠開啟。
3. 在側邊欄第 1 節建立或開啟專案。專案設定用途 (General、For Appraisal 或 Market Area Analysis) 與工作地點 (國家、州、縣、市) 及其輸入/輸出資料夾。專案存放在共用的 **regProj** 根目錄之下，並與同族應用 **earthUI** 和 **glmnetUI** 共用。開啟一個用途不同的專案，正是切換模式的方式。
4. 在第 2 節匯入資料：從目前作用中專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。把檔案放入該資料夾，然後點選 **Refresh** (重新整理) 以列出它們。
5. 從 **earthUI** 匯入 (選用) ——第 3 節允許你匯入一個 **earthUI** 結果 **.rds** 檔案，用 **earth** 發現的節點為 **GAM** 提供初值。若你沒有 **earthUI** 結果，跳過此步即可。
6. 設定變數 ——選擇目標變數與預測變數，設定資料型別，指派特殊欄位角色，並將變數標記為因子型或線性。
7. 設定 **mgcv** 參數 ——**family**、**method**、**gamma**、基底類型、**k**、張量類型、交互項以及其他選項。
8. 擬合模型 ——點選「Fit **Mgcv** **GAM** Model」(擬合 **mgcv** **GAM** 模型)，並在主面板中檢視結果。
9. 匯出 ——把預測結果下載為 Excel，產生並轉換 **Quarto** 報告，或 (在估價模式下) 計算 **RCA** 調整額與銷售比較網絡。

設定會自動儲存在一個 **SQLite** 資料庫中，並在你重新載入同一輸入檔案時復原。

2 MLS 輸入資料要求

對於不動產估價與市場分析工作流程而言，你的輸入資料通常來自多重上市服務（Multiple Listing Service，MLS）的匯出檔案。本章說明預期的檔案結構以及 mgcvUI 可以使用的各欄位。

2.1 檔案格式與結構

mgcvUI 接受 CSV 與 Excel (.xlsx、.xls) 檔案。匯入時，欄位名會自動轉換為 snake_case——例如，「Living SqFt」變為 living_sqft，「Contract Date」變為 contract_date，「Sale Price」變為 sale_price。這一規範化確保整個工作流程中欄位引用的一致性。匯入時使用的 CSV 分隔符與小數點符號由區域設定決定（參見第 3 章「Locale & Regional Settings」，區域與地區設定）。

你的資料檔案應是一張扁平表格，每個不動產佔一行，每個屬性佔一欄。檔案的第一列必須是欄位標題。

2.2 估價模式的必需欄位

雖然 mgcvUI 可以處理任意一組欄位，但完整的估價工作流程若具備以下各欄位將獲益更多：

欄位	特殊類型	用途
Sale Price（銷售價格）	（目標變數）	模型的反應變數
Contract Date（成交日期）	contract_date	用於計算 sale_age（距估價基準日的天數）
Listing Date（掛牌日期）	listing_date	在缺少 DOM 欄位時，與成交日期一起用於計算 DOM
Days on Market（市場停留天數）	dom	市場停留天數；在匯出檔案中顯示
Concessions（成交讓價）	concessions	銷售讓價；淨售價 = 銷售價格 - 讓價
Living Area (SF)（居住面積）	living_area	啟用按平方英尺計的殘差 (residual_sf、cqa_sf)
Lot Size（地塊面積）	lot_size	宗地面積欄位
Site Dimensions（宗地尺寸）	site_dimensions	與地塊面積歸為一組
Latitude（緯度）	latitude	四捨五入至 3 位小數
Longitude（經度）	longitude	四捨五入至 3 位小數
Area ID（區域識別碼）	area	市場區域 / 街區識別碼
Actual Age（實際年齡）	actual_age	不動產年齡
Effective Age（有效年齡）	effective_age	不動產的有效年齡
Address（地址）	display_only	在匯出檔案中顯示；不進入模型

試算表的欄位名可以使用外語——而「特殊」名稱使用英文，以便 R 程式對其作特殊處理。除此之外，你所給定的欄位名會出現在迴歸模型、圖形中，若是做估價，還會出現在輸出報告中。

並非所有欄位都是必需的。mgcvUI 會自行適應——若某欄位遺漏，相應功能便自動略去。不過，對於不動產定價模型，以下幾個欄位強烈建議：

1. **Sale Age**（銷售時距）——成交日期與估價基準日之間的天數。若使用跨越數年的銷售歷史，`sale_age` 往往起著核心作用。
2. **Living Area**（居住面積）——也常被稱作「Living Sqft」「GLA」（gross living area，總居住面積）。
3. **Total Bath Count**（衛浴總數）——全衛、四分之一衛、半衛與四分之三衛的總數。
4. **Garage Bays**（車庫車位）或 **Garage Area**（車庫面積）——車庫車位數或建築面積。
5. **Lot Size**（地塊面積）——不動產的土地面積。
6. **Longitude, Latitude**（經度、緯度），以及若有的話 **Area ID**（區域識別碼），用於刻畫地理上的價格變化。

2.3 特殊欄位命名約定

mgcvUI 依據欄位的特殊類型指定而非欄位名來識別各欄位。你可以在 MLS 匯出檔案中隨意命名欄位——關鍵在於你要在變數設定表（第 6 章）中指派正確的特殊類型。

2.4 資料品質與完整性

- **遺漏值 (NA)**：任一預測變數欄位或目標欄位中含 NA 的列會在擬合前自動刪除。mgcvUI 會把資料子集化為模型實際使用的那些欄位，因此未被使用的欄位中的 NA 不會導致丟列。
- **數值欄位**：銷售價格、居住面積、地塊面積等類似欄位必須為數值型。
- **因子欄位**：文字欄位、TRUE/FALSE 欄位以及 R 因子欄位會被自動視為分類變數——它們只能以因子形式進入模型。一個數值欄位（例如數值型的區域碼或風格碼）只有在你勾選其 **Factor**（因子）核取方塊時才會成為因子；mgcvUI 絕不會根據數值的取值範圍推斷其為分類變數，因此像 `bath_count` 這樣的離散數值預測變數仍保持連續（並被平滑化）。Special（特殊）角色與因子的判定彼此獨立。

提示：匯入後請檢視預測變數表中的 NA 欄位。遺漏值很多的欄位可能導致丟列。NA 佔比 $\geq 50\%$ 的欄位會被標紅並自動排除。

2.5 標的列的位置

在 **Appraisal**（估價）模式下，第 1 列必須是估價標的。其餘各列均為可比銷售案例。標的列不參與模型擬合。擬合完成後，模型仍會為標的列產生預測值。

在 **Market Area Analysis**（市場區域分析）模式下，是否把估價標的置於第 1 列可自行決定。

在 **General**（通用）模式下沒有特殊的列處理——所有列一視同仁。

3 通用模式

3.1 概述

啟動 mgcvUI 時，通用 (General Purpose) 模式是預設模式。它為任何資料集——而不僅僅是不動產資料——提供完整的 GAM 工作流程。你可以把 mgcvUI 用於科學資料、金融分析、工程研究，或任何預期存在平滑非線性關係的迴歸問題。

在 General 模式下，介面略去不動產專用功能 (特殊欄位、銷售時距、RCA)。側邊欄被精簡，專注於變數選擇、參數設定、模型擬合與匯出。

3.2 側邊欄工作流程

側邊欄按編號組織為若干可折疊小節，引導你從專案選擇一路走到報告匯出。多數小節只有在開啟一個作用中專案後才出現，其中若干還與專案的用途相關。輸出資料夾由作用中專案推導得出——並沒有單獨的輸出資料夾欄位。

1. Project (專案) ——建立或開啟一個專案。專案設定用途與工作地點 (國家、州、縣、市)，以及它的輸入 (<os>_in/) 與輸出資料夾。專案存放在共用的 regProj 根目錄下 (預設為 ~/regProj，可透過 REGPROJ_ROOT 環境變數或 Settings 下拉選單覆寫)，並與同族應用 earthUI 和 glmnetUI 共用，因此三者的模型可以並列存放。點選 **Close Project** (關閉專案) 即可切換。

2. Import Data (匯入資料) ——從作用中專案的輸入資料夾中選擇一個 CSV 或 Excel 檔案。先透過 Finder/檔案總管把檔案放入該資料夾，然後點選 **Refresh** (重新整理) 以列出它們。對於含多個工作表的 Excel 檔案，會出現一個工作表選擇器。欄位名會自動轉換為 snake_case。

3. Import from earthUI (optional) (從 earthUI 匯入，選用) ——匯入一個 earthUI 結果 .rds 檔案，用 earth 發現的節點位置為 GAM 提供初值 (同時也為變數選擇提供初值)。此步驟是選用的——若你沒有可匯入的 earthUI 結果，跳過即可。

4. Variable Configuration (變數設定) ——目標變數選擇器、反應變數變換 (無、log、log10)，以及帶 Include (納入)、Factor (因子) 和 Linear (線性) 核取方塊的預測變數表。Special (特殊) 欄位只在估價模式與市場模式下出現。詳見第 6 章。

5. MgcV Call Parameters (mgcv 呼叫參數) ——全部模型設定：參數預設組、family、method、gamma、交叉驗證、select、基底類型、k、張量類型、交互矩陣以及進階參數。完整的參數參考見第 7 章。

6. Fit MgcV GAM Model (擬合 mgcv GAM 模型) ——執行模型的按鈕。結果會顯示在右側的各個頁籤中。

7. Download Output (下載輸出) ——把預測值、殘差、CQA 得分以及各變數的貢獻匯出為 Excel 檔案，存入專案的輸出資料夾。模型擬合之後，在所有用途模式下均可使用。

8. Calculate RCA Adjustments & Download (計算 RCA 調整額並下載) ——為每個可比案例計算逐變數的 RCA 調整額，並透過 CQA 對估價標的殘差作內插。僅在估價模式與市場模式下、且已完成擬合後出現。

9. Generate Sales Grid & Download (產生銷售網格並下載) ——產生銷售比較網格 (Sales Comparison Grid)，按總調整百分比建議可比案例。僅在估價模式下出現。

10. Generate Quarto Report (產生 Quarto 報告) ——把一個自包含的 Quarto (.qmd) 報告包——含全部圖形與表格——寫入專案的輸出資料夾。模型擬合之後即可使用。

11. Convert Quarto Report (轉換 Quarto 報告) ——把產生的 .qmd 包渲染為 HTML、Word 或 PDF。PDF 需要安裝 LaTeX；若未偵測到 LaTeX，PDF 選項會被隱藏。參見第 13 章。

在其他模式下，與用途相關的步驟會被略去 (其餘步驟重新編號)：市場模式下沒有銷售網格步驟，通用模式下 RCA 與銷售網格兩個步驟都沒有。

3.3 從 earthUI 匯入 (選用)

側邊欄的第 3 節——**Import from earthUI**——允許你匯入一個 earthUI 結果 .rds 檔案。這就啟用了 earth-mgcv 流程：earth 由資料驅動得出的節點位置成為 GAM 平滑項的錨點。

匯入 earthUI 結果後，mgcvUI 會：

- 顯示一份摘要，列出目標變數、R 平方、項數以及每個變數的節點位置。
- 自動選中「Earth Pipeline」(earth 流程) 參數預設組 (三次迴歸樣條基底， $\gamma = 1.4$)。
- 把目標變數設為與 earth 的目標變數一致。
- 為 earth 的預測變數預先勾選 Include。
- 為 earth 的分類變數預先勾選 Factor，為 earth 的線性預測變數預先勾選 Linear。
- 預先填入 earth 偵測到的交互項 (在交互矩陣中以黃色高亮顯示，並鎖定不允許手工更改)。
- 提供一個 **Export Knots CSV** (匯出節點 CSV) 按鈕，可把 earth 的節點下載為 CSV 檔案。

隨後，變數設定表會顯示一條說明，指出變數選擇 (Include、Linear、Factor) 與樣條節點是由 earth 模型提供初值的。與 glmnetUI 不同——在 glmnetUI 中 glmnet 的擬合被綁定到 earth 的基底上——mgcvUI 是依據你目前的選擇來擬合 GAM 的，earth 只是提供了一個起步優勢，而每一個控制項都仍可編輯。

一個 **Clear** (清除) 按鈕可移除已匯入的 earth 資料，並重置為獨立執行模式。

3.4 主面板頁籤

資料匯入之後，主面板提供以下頁籤 (取決於模型的頁籤在擬合完成後才有內容)：

頁籤	內容
Data (資料)	已匯入資料的互動式 DataTable。在估價模式與市場模式下，分為估價標的表與可比銷售表。
Equation (方程)	用 MathJax 渲染的完整 GAM 公式，以及一張平滑函數定義表。
Summary (摘要)	關鍵指標 (R^2 、 $CV R^2$ 、已解釋離差、AIC、n、平滑項個數) 以及平滑項 / 參數項表格。
Variable Importance (變數重要性)	F 統計量 (平滑項) 與 $ t $ 值 (參數項) 的水平長條圖，以及排序表。
Contribution (貢獻)	互動式 plotly 平滑偏效應曲線，帶信賴帶與 earth 節點標記。
Correlation (相關性)	數值型預測變數相關係數的熱圖 (擬合前即可檢視)。
Diagnostics (診斷)	殘差對擬合值、Q-Q 圖、直方圖、反應值對擬合值 (透過 gratia 產生)。實際值對預測值散點圖。
RCA Adjustments (RCA 調整)	RCA 分析直方圖 (估價 / 市場模式，計算完成後)。
Anova (變異數分析)	參數項與平滑項合併的變異數分析表。
Mgcv Output (mgcv 輸出)	原始文字：計時、公式、模型摘要、family、共曲性。
Sign Check (符號檢查)	earth 與 GAM 的方向一致性檢查 (當已匯入 earth 結果時)。
Concurvity (共曲性)	整體與兩兩共曲性矩陣，帶顏色編碼。
Basis Check (基底檢驗)	<code>gam.check()</code> 的基底維度充分性檢驗。

3.5 設定的持久化

mgcvUI 會自動把你的設定儲存到一個 SQLite 資料庫中，並以輸入檔案名為鍵。當你重新載入同一檔案時，所有設定都會復原：目標變數選擇、預測變數核取方塊、資料型別、mgcv 參數以及交互矩陣。用途是作用中專案的屬性而非全域設定，最近使用的輸入檔案按專案分別記住，並在專案重新開啟時自動載入。

3.6 深色模式

點選右上角的月亮 / 太陽圖示，可在 Nord Light 與 Nord Dark 兩種主題之間切換。主題偏好儲存在 localStorage 中，跨會話保持。所有介面元素都會隨所選主題相應變化。

3.7 區域與地區設定

mgcvUI 透過一套基於國家的區域設定 (locale) 系統支援國際化的數字與 CSV 格式約定。標題欄中的 **Settings** (設定) 下拉選單為 30 多個受支援國家提供 **Country** (國家) 與 **Paper** (紙張) 選擇器。每種預設設定以下各項：

- **CSV 分隔符** —— 美國 / 英國 / 日本用逗號 (,)，歐洲大部分地區用分號 (;)。
- **小數點符號** —— 句點 (.) 或逗號 (,)。
- **千位分隔符** —— 逗號 (美國 / 英國 / 日本)、句點 (德國 / 義大利 / 西班牙)、空格 (芬蘭 / 法國 / 波蘭 / 波羅的海國家)，或撇號 (瑞士)。
- **紙張尺寸** —— Letter (美國 / 加拿大 / 墨西哥) 或 A4 (其他所有地區)。

點選 **Save**（儲存）即可把你的區域偏好全域存儲。Import Data 小節還提供一個按檔案生效的 **Locale**（區域）選擇器，供採用不同約定的一次性匯入使用。同一個 Settings 下拉選單中還有 **regProj root folder**（regProj 根資料夾）欄位，用於重新指定 mgcvUI、earthUI 與 glmnetUI 共用的專案樹位置。

提示：圖形座標軸與斜率標籤上的數字格式會自動適應你的區域設定。德語區域用句點作千位分隔（200.000），芬蘭用空格（200 000），瑞士用撇號（200'000）。介面不顯示任何貨幣符號——mgcvUI 與貨幣無關。

4 估價模式

當作用中專案的用途為 **For Appraisal**（用於估價）時，mgcvUI 會把自身設定為面向單一不動產估價。第 3 章所述的全部功能依然可用；本章只講估價模式特有的附加內容。

4.1 標的列的處理

在估價模式下，資料集的第 1 列即估價標的，其餘各列為可比銷售案例。你的輸入檔案必須據此組織（參見第 2 章）。估價標的的銷售價格可以留空，也可以填任意值——mgcvUI 在擬合時會自動把它當作 NA 處理。

擬合完成後，模型會為標的列產生預測值，這是 RCA 調整工作流程的基礎。

4.2 估價基準日與銷售時距

在估價模式與市場模式下，變數設定小節中會出現一個 **Effective Date**（估價基準日）欄位（預設為當天日期）。若你在 Special 下拉選單中把某欄位指定為 `contract_date`，mgcvUI 就會計算一個 `sale_age` 欄位——即每筆交易的成交日期與基準日之間的整天數。該欄位會作為預測變數加入模型。

當估價基準日改變時，`sale_age` 會自動重新計算。

4.3 特殊欄位指定

在估價模式與市場模式下，變數設定表中的每個預測變數都會出現一個 **Special**（特殊）下拉選單。特殊類型及其作用的完整參考見第 6 章。

4.4 RCA 調整概覽

Calculate RCA Adjustments & Download（計算 RCA 調整額並下載）按鈕（在估價模式與市場模式下、擬合完成後可見）會為每個可比案例相對於估價標的計算由市場推導出的調整額。完整的 RCA 工作流程見第 11 章。

提示：你為估價標的指定的 *CQA* 得分決定了殘差分配中有多大比例被歸屬於估價標的。得分 5.00 會把估價標的置於各可比案例的中位數位置。

5 市場區域分析模式

當作用中專案的用途為 **Market Area Analysis**（市場區域分析）時，mgcvUI 提供與估價模式相同的不動產專用功能（特殊欄位、銷售時距），但其取向是分析一組不動產，而非評估單一估價標的的價值。

5.1 與估價模式的差異

- 標的列的處理更為彈性 —— 在市場模式下，第 1 列不會被自動當作估價標的。
- 沒有銷售網格 —— 「Generate Sales Grid & Download」步驟不可用。

5.2 何時使用市場模式

在以下情形下適合使用市場區域分析模式：

- 為某個街區或市場區域建立 GAM 模型，以理解非線性的價值驅動因素
- 分析建築面積、年齡、地塊面積與區位等變數如何影響一組不動產的銷售價格
- 為市場狀況分析或街區界定準備支撐材料
- 處理這樣一份資料集：你想要特殊欄位功能（銷售時距），但不需要銷售網格工作流程

提示：市場模式對一般性的不動產迴歸也很有用——當你想要特殊欄位功能，但不需要單一不動產估價工作流程時。

6 變數選擇

側邊欄的第 4 節——**Variable Configuration**（變數設定）——是你選定哪些欄位參與模型、以及它們如何被處理的地方。

6.1 目標變數

第 4 節頂部的 **Target (response) variable**（目標 / 反應變數）下拉選單列出資料集中的每一個數值欄位。選擇其中一欄作為反應變數（例如銷售價格）。目標欄位會自動從預測變數清單中排除。

6.2 反應變數變換

在目標變數選擇器下方，**Response Transform**（反應變數變換）下拉選單提供三個選項：

- **None (raw)**（無，原值）——直接使用反應變數。當預測變數與反應變數在原始尺度上近似可加時，此選項合適。
- **Log (natural)**（自然對數）——在擬合前施加自然對數變換。反變換之後，這會使調整額成為比例式（%）而非絕對額（\$）。這對銷售價格尤其有用，因為一筆 \$10,000 的調整在不同價格水準上含義大不相同。
- **Log10**（以 10 為底的對數）——施加以 10 為底的對數變換。

提示：對數變換會把時間調整變為比例式（%）而非絕對額（\$）。對於價格跨度很大的不動產模型，對數變換往往能產生性態更好的殘差。

選擇對數變換後，反應變數中 ≤ 0 的取值會被自動濾除。所有輸出（預測值、貢獻、殘差）都會自動反變換回原始尺度。

6.3 預測變數表

在目標變數選擇器下方，一張可捲動的表格列出其餘所有欄位。核取方塊欄位使用旋轉的直排表頭以求緊湊。欄位的順序為：

欄位	說明
Variable	欄位名（等寬字型，過長時以省略號截斷）。帶邊框儲存格。
Type	資料型別下拉選單：numeric、integer、character、factor、logical、Date、POSIXct。帶邊框下拉選單。
Include	核取方塊——把該欄位作為預測變數納入模型。直排表頭。
Factor	核取方塊——按分類因子處理（為每個水準產生一個單獨的係數）。直排表頭。
Linear	核取方塊——強制使用線性關係（不用平滑項，只用 βx ）。直排表頭。
Special	下拉選單（僅估價 / 市場模式）——參見下文的特殊欄位類型參考。帶邊框下拉選單。
NAs	遺漏值的個數與百分比，帶顏色編碼：紅色（ $\geq 50\%$ ）、橙色（ $\geq 20\%$ ）、灰色（ > 0 ）。帶邊框儲存格。

因子、平滑與線性之別：預設情況下，被納入的數值變數會得到一個平滑項 $f(x)$ 。勾選 **Factor** 會產生一個分類項（每個水準一個係數）。勾選 **Linear** 會產生一個簡單線性項 βx 以取代平滑項。同時被標記為 Factor 與 Linear 的變數按 Factor 處理。

6.4 資料型別偵測與覆寫

mgcvUI 在匯入時會自動偵測資料型別，可識別數值、整數、邏輯、因子與日期欄位。你可以透過更改 **Type** 下拉選單來覆寫任何一項偵測結果。更改類型會影響該欄位在模型中的處理方式。

6.5 特殊欄位類型參考

在估價模式與市場模式下，**Special** 下拉選單提供以下選項：

加權：

- **weight** —— 觀測權重欄位（只允許一欄；權重為 0 的列不參與擬合）

日期與時間類型：

- **contract_date** —— 觸發根據估價基準日自動計算 **sale_age**
- **sale_age** —— 標示一個已預先計算好的銷售時距欄位（距基準日的天數），按原樣使用
- **listing_date** —— 在計算市場停留天數時作為備選來源
- **dom** —— 標示市場停留天數欄位

記錄類型：

- **sale_type** —— 標示用於區分估價標的與可比銷售案例的記錄類型欄位

貨幣類型：

- **concessions** —— 標示銷售讓價（賣方補貼、買方優惠等）

面積與區位類型：

- **latitude** —— 取值自動四捨五入至 3 位小數
- **longitude** —— 與緯度同樣的捨入處理
- **area** —— 市場區域或街區識別碼
- **living_area** —— 啟用按平方英尺計的殘差計算（**residual_sf** 與 **cqa_sf**）
- **lot_size** —— 宗地面積欄位
- **site_dimensions** —— 與地塊面積歸為一組

年齡類型：

- **actual_age** —— 不動產年齡欄位
- **effective_age** —— 不動產的有效年齡

顯示類型：

- **display_only** —— 該欄位會包含在 Excel 匯出中，但完全不參與模型擬合。可用於地址欄位、MLS 編號、宗地編號或其他參考性資料。

提示：被指定為 *display_only* 的欄位仍保留在資料集中並出現在 Excel 匯出裡，但會完全從預測變數表中排除。可用於 ID 欄位、地址或其他參考欄位。

7 參數選擇

側邊欄的第 5 節——**Mgcv Call Parameters** (mgcv 呼叫參數)——提供 GAM 的全部設定選項。每個參數都帶有一個藍色說明圖示 (?)，其提示框給出說明。

7.1 參數預設組

頂部的一個下拉選單提供兩種預設組，為常見工作流程設定合理的預設值：

預設	基底	設定
Standalone (discovery) (獨立執行，探索)	tp (薄板)	k = 自動 (10)，gamma = 1.2，select = TRUE
Earth Pipeline (refinement) (earth 流程，精化)	cr (三次迴歸)	k = 自動 (10)，gamma = 1.4，select = TRUE

匯入 earthUI 節點時會自動選中 **Earth Pipeline** 預設組。earth 節點的整合必須使用三次迴歸樣條 (cr) 基底，因為只有它允許指定確切的節點位置。

7.2 Family (分配族)

為你的反應變數選擇分配族：

Family	適用情形
gaussian	連續型反應變數 (例如銷售價格)。最為常用。
Gamma	變異數與平均數成比例的正值連續型反應變數。
poisson	計數資料 (例如成交筆數)。
binomial	二元結果。
inverse.gaussian	嚴重右偏的正值資料。

7.3 Method (方法)

平滑參數的估計方法：

Method	說明
REML	受限最大概似 (預設，建議)。對過擬合最為穩健。
GCV.Cp	廣義交叉驗證。傾向於選出略為複雜的模型。
ML	最大概似。與 REML 相似，但可能欠平滑。
P-REML	Pearson REML 變體。
P-ML	Pearson ML 變體。
GACV.Cp	廣義近似交叉驗證。

7.4 Gamma (平滑度乘子)

gamma 參數會對平滑準則中的有效自由度作乘法放大，從而促成更平滑 (擺動更少) 的擬合。預設值為 1.2。

取值	效果
1.0	標準平滑度
1.2–1.4	略比預設更平滑（建議）
2.0+	平滑得多——適用於雜訊較大的資料或小樣本

較大的 γ 值透過更重地懲罰複雜度來防範過擬合。Earth Pipeline 預設採用 1.4，以便在精化 earth 的節點時施加額外的平滑。

7.5 交叉驗證

勾選時（預設勾選），mgcvUI 會在擬合後執行 10 折交叉驗證（cross-validation）以計算 CV R 平方。這給出了樣本外預測能力的一個誠實估計。

7.6 Select（收縮）

勾選時（預設勾選），會為每個平滑項附加一個額外懲罰，可將其收縮至零。這就實現了自動變數選擇——不重要的平滑項實際上被移出模型。其實作方式是在 `gam()` 中設定 `select = TRUE`。

7.7 預設基底

平滑項所用的樣條基底函數（basis function）類型：

基底	全稱	說明
tp	薄板迴歸樣條	預設。各向同性（無需安放節點）。是很好的通用選擇。
cr	三次迴歸樣條	允許顯式安放節點。整合 earth 節點時必須使用。
ps	P 樣條	節點等距的 B 樣條基底，配以差分懲罰。計算效率高。
bs	B 樣條	彈性的 B 樣條基底。

7.8 預設 k（基底維度）

基底維度 k 控制每個平滑項擺動程度的上限。取值為 0（預設）表示「自動」——mgcvUI 會視情形取 $k = 10$ 或 earth 節點的個數。

- 較小的 k （3–5）：曲線非常平滑，接近直線。
- 中等的 k （8–15）：彈性適中（通常已足夠）。
- 較大的 k （20 以上）：彈性極高——若無強正則化則有過擬合之虞。

提示：Summary 頁籤中的有效自由度（EDF）告訴你每個平滑項實際用掉了多少自由度。若 EDF 接近 $k - 1$ ，說明基底可能太小——應加大 k 。若 EDF 遠小於 k ，說明該平滑項受到了良好的懲罰。

7.9 張量類型

對於連續變數之間的交互，有兩種張量積類型可用：

類型	函數	說明
ti	ti(x1, x2)	張量交互——只對交互效應建模，主效應由單變數 $s()$ 項分別處理。就可解釋性而言更受建議。
te	te(x1, x2)	張量積——對包含主效應在內的整個聯合效應建模。在 RCA 調整中較難分解。

7.10 允許的交互

一個可折疊的 **Allowed Interactions**（允許的交互）小節會為所有被納入的平滑（非線性、非因子）預測變數顯示一個上三角核取方塊矩陣。每個核取方塊啟用相應變數對之間的張量積平滑。

- **Allow All**（全部允許） / **Clear All**（全部清除）——批次切換按鈕。
- 點選變數名——切換該變數的全部交互。
- 右鍵點選變數名——「Block from main effect」（從主效應中遮蔽）——把該變數標記為僅用於交互（不產生單變數平滑項）。

匯入 earthUI 節點時，earth 偵測到的交互會被預先勾選並鎖定（以黃色背景高亮）。若 earth 模型使用了 $\text{degree} = 1$ （無交互），則會顯示一條提示資訊。

警告

啟用張量交互會大幅提高模型複雜度。請在你擁有領域知識、或有 earth 偵測到的證據支援時才使用交互項。務必核實所得調整額是合理且可以解釋的。

7.11 因子-平滑交互

當因子變數與平滑變數同時被納入時，會出現一個單獨的 **Factor-by-Smooth Interactions**（因子-平滑交互）矩陣。每個核取方塊會產生一個 $s(x, \text{by} = \text{factor})$ 項——即為每個因子水準產生一條單獨的平滑曲線。

7.12 進階參數

一個可折疊的「Advanced」（進階）小節展示若干附加設定：

- **Optimizer**（最佳化器）——平滑參數估計所用的演算法：`outer/newton`（預設）、`outer/bfgs` 或 `efs`（擴展 Fellner-Schall 法）。
- **Scale**（尺度參數）——尺度參數。0（預設）表示由資料估計。對於尺度已知的模型可設為一個確定值。
- **Discrete covariate method**（離散化協變數方法）——針對大型資料集的快速離散化擬合（`discrete = TRUE`）。
- **Threads**（執行緒數）——並行執行緒的數目（1-32）。僅在 `discrete = TRUE` 時有效。

8 擬合模型

8.1 Fit 按鈕

側邊欄第 6 節中有 **Fit MgcV GAM Model** (擬合 mgcv GAM 模型) 按鈕。點選它即以你目前的設定執行模型。

8.2 擬合過程中發生了什麼

當你點選 Fit 時，mgcvUI 會：

1. 校驗設定 —— 檢查是否至少有一個預測變數、以及至少 10 個完整列。若完整列不足 50%，會給出警告（並列出造成大部分 NA 的那些欄位）。
2. 準備資料 —— 只保留與模型相關的欄位，刪除含 NA 的列，應用權重（若已指定），排除標的列（在估價模式下），並施加反應變數變換。
3. 建構 GAM 公式 —— 為數值預測變數構造平滑項 $s(x, bs="cr", k=10)$ ，為分類變數構造因子項，為被標記為 Linear 的變數構造線性項，並為所選交互構造張量積。earth 的節點會作為顯式的節點位置嵌入到 cr 基底的項中。
4. 處理邊界情形 —— 剔除常數變數，把二值預測變數轉換為因子，把 k 限制在不同取值個數減 1 以內，並在需要更多節點時用資料分位數補充 earth 的節點。
5. 擬合 GAM —— 以構造好的公式、family、method、gamma、select 及其他參數呼叫 `mgcv::gam()`。
6. 執行交叉驗證（若已啟用） —— 10 折交叉驗證，計算樣本外 R 平方。
7. 更新所有結果頁籤 —— Summary、Equation、Contribution 圖、Diagnostics 以及其他所有頁籤都會被填充。

8.3 背景擬合與進度顯示

當 callr 套件可用時，模型擬合會在一個背景行程中執行，以保持應用程式的回應性。一個暗色主題的終端式進度浮層會顯示：

- 計時器，顯示已耗費的秒數 / 分鐘數。
- 按顏色編碼的追蹤日誌：藍色表示狀態，黃色表示交叉驗證折的進度，紅色表示錯誤，綠色表示結果。
- 一個 **Abort**（中止）按鈕，可在需要時終止擬合行程。
- 一個 **Close**（關閉）按鈕，在擬合完成後出現。

若 callr 不可用，擬合將同步執行（應用程式會凍結，直到擬合完成）。

擬合成功後，Fit 按鈕上會出現一個白色勾號，同時狀態列顯示：「R-sq = X, CV R-sq = X, Dev = X%, AIC = X, n = X.」

9 結果頁籤

9.1 Data (資料)

以互動式 DataTable 顯示已匯入的資料，支援橫向捲動，每頁 15 列。在估價模式與市場模式下，該頁籤分為一張 **Subject Property**（估價標的，第 1 列）表和一張 **Comparable Sales**（可比銷售案例，其餘各列）表；在 General 模式下只顯示一張表。

每個儲存格都只顯示一行文字，並按欄寬截斷，這樣很寬的自由文字欄位（例如不動產備註）就不會把整列拉得很長。點選可選取其中一列；雙擊任一儲存格可彈出視窗，顯示其完整、未經截斷的內容。

9.2 Equation (方程)

分兩部分顯示已擬合的 GAM 模型：

模型方程（由 MathJax 渲染）：

- 左端顯示反應變數及其相應的變換（log、log10）與連結函數。
- 右端顯示截距值、以 f_i (變數) 形式給出的平滑項、按變數分組並標明水準數的因子項，以及帶係數值的線性項。
- family 與 method 顯示在方程上方。

Smooth Function Definitions（平滑函數定義）表：

- 項的公式字串（例如 `s(living_sqft, bs="cr", k=8)`）
- 變數名
- 類型（s、te、ti 或 linear）
- 基底函數類型
- k 值
- earth 節點位置（若已匯入）

9.3 Summary (摘要)

頂部橫向排列六張指標卡片：

- **R-squared** —— 已解釋離差的比例
- **CV R-squared** —— 交叉驗證 R 平方（若已啟用交叉驗證）
- **Dev. Explained %** —— 已解釋離差的百分比
- **AIC** —— 赤池資訊準則
- **n** —— 所用觀測的個數
- **Smooths** —— 平滑項的個數

卡片下方：

- **Smooth Terms** 表（平滑項）—— 項、EDF（有效自由度）、Ref.df、F 統計量、p 值
- **Parametric Terms** 表（參數項）—— 項、估計值、標準誤、t 值、p 值

9.4 Variable Importance (變數重要性)

一張水平長條圖，顯示各模型項的相對重要性：

- 平滑項：按 F 統計量排序（藍色條）
- 參數項：按 $|t|$ 值排序（綠色條）

圖下方的 DataTable 列出項、類型、EDF、統計量與 p 值，按統計量降序排列。

9.5 貢獻圖

每一個對預測值有貢獻的模型項都有一張互動式 plotly 貢獻圖，顯示在一張帶邊框的卡片中。該頁籤會為所有類型的項給出圖形：

9.5.1 單變數平滑項 — $s(x)$

- 藍色曲線，帶 95% 信賴帶
- 灰色散點，顯示偏殘差（用於未作對數變換的模型）
- 紅色垂直虛線，位於 earth 節點處（若已匯入 earth 結果）
- 懸停提示：變數取值、貢獻（\$）、95% 信賴區間、變化率（每單位的斜率）

對於經對數變換的模型，縱軸會用公式 $\bar{y} \times (e^{f(x)} - 1)$ 反變換為以美元計的貢獻，使曲線可以直接按美元數額來解讀。

對於緯度／經度變數，斜率按每 0.001 度而非每 1 度給出，因為前者是更有意義的地理增量。

9.5.2 因子—平滑項 — $s(x, \text{by} = \text{factor})$

每個因子水準一條彩色曲線，疊繪在同一張圖上。這有助於看清某個預測變數的效應如何隨組別（例如不同街區）而變化。

9.5.3 交互項 — $ti(x_1, x_2)$ 與 $te(x_1, x_2)$

用熱圖顯示二維貢獻曲面。色標為：紅色（負）→ 白色（零）→ 藍色（正）。懸停時同時顯示兩個變數的取值與貢獻數額。

9.5.4 參數項 — 線性項與因子項

- 數值型線性項：資料點的散點圖，疊加擬合直線
- 因子項：長條圖，顯示每個因子水準的平均貢獻

9.6 Correlation（相關性）

一張熱圖矩陣，顯示所有數值型預測變數之間的 Pearson 相關係數。資料匯入後即可檢視——無需擬合模型。

- 顏色漸變：藍色（-1）→ 白色（0）→ 紅色（+1）
- 每個儲存格中標出相關係數值
- 有助於在擬合之前識別共曲性（concurvity，多重共線性在平滑項中的對應現象）

9.7 Diagnostics（診斷）

兩組診斷視覺化：

診斷面板（透過 `gratia::appraise` 產生）：四張圖——殘差對擬合值、Q-Q 圖、殘差直方圖，以及反應值對擬合值。

實際值對預測值散點圖：觀測值對預測值，附一條 45 度虛線參考線。對於經對數變換的模型，取值會被反變換回原始尺度。

9.8 RCA Adjustments（RCA 調整）

RCA 計算完成後顯示三張直方圖（僅限估價 / 市場模式）：

- **Residual Adjustment %**（殘差調整百分比）——殘差調整額佔銷售價格百分比的分配
- **Net Adjustment %**（淨調整百分比）——淨調整額的分配
- **Gross Adjustment %**（總調整百分比）——總調整額的分配

每張直方圖的副標題中給出平均數、中位數與標準差，並以虛線標出平均數與中位數的參考位置。

9.9 Anova (變異數分析)

一張合併的變異數分析表，把 GAM 摘要中的參數項與平滑項匯合在一起。顯示項、類型（參數項或平滑項）以及所有相關統計量。

9.10 Mgcvc Output (mgcv 輸出)

原始文字輸出，內容包括：

- 擬合計時
- GAM 公式
- 模型的 print 輸出
- 完整的 `summary(model)` 輸出
- family 與連結函數
- 整體共曲性度量

9.11 Sign Check (符號檢查)

當已匯入 earthUI 節點時，該頁籤會把每個變數在 earth 模型中的方向與其在 GAM 中的方向作比較：

- **earth_direction**：遞增、遞減或混合（依據鉸鏈函數的符號判定）
- **gam_direction**：遞增、遞減或混合（由平滑項在 100 點網格上的導數估計得出）
- **consistent**：TRUE/FALSE
- **warning**：對任何不一致之處的說明

狀態列會顯示「All directions consistent」（所有方向一致）或存在不一致的變數個數。

9.12 Concurvity (共曲性)

共曲性是多重共線性在平滑項中的對應現象。該頁籤顯示：

- 整體共曲性表：各列分別對應 worst / observed / estimate 三種度量，各欄對應每個平滑項。帶顏色編碼：白色 (< 0.5)、黃色 ($0.5-0.8$)、紅色 (> 0.8)。
- 兩兩共曲性表：最壞情形的兩兩矩陣，採用相同的顏色編碼。

「worst」一列中高於約 0.8 的取值表明，該平滑項可以被其他平滑項很好地逼近，提示存在冗餘。

9.13 Basis Check (基底檢驗)

`mgcv::gam.check()` 的輸出：針對每個平滑項的基底維度充分性檢驗。若某個平滑項的 k 過小，該檢驗會以一個較小的 p 值將其標示出來，提示你應當加大 k 。

10 下載資料

擬合完成後，**Download Output (Excel)**（下載輸出，Excel）按鈕（側邊欄第 7 節）會把含預測值與診斷資訊的 Excel 檔案匯出到作用中專案的 mgcv 輸出資料夾。

10.1 輸出欄位

欄位	說明
basis	反變換後的截距值（所有列相同）
<var>_contribution	各平滑項對預測值的貢獻。對於對數模型，會按比例分配為美元數額。
est_<target>	模型預測值（例如 est_sale_price），若適用則已從對數尺度反變換
residual	實際值 – 預測值
calc_residual	校驗用：實際值 – 各項貢獻之和（反變換後）
cqa	狀況–品質–吸引力（Condition-Quality-Appeal，CQA）得分（0–10 分制）
residual_sf	殘差 / 居住面積（若已指定 living_area）
cqa_sf	由 residual_sf 排序計算得到的 CQA

在估價模式下，第 1 列（估價標的）的 residual/cqa/cqa_sf 均置為 NA。當指定了 living_area 欄位時，各列按 residual_sf 降序排列。

成功完成後，下載按鈕上會出現一個白色勾號。

10.2 CQA 得分

CQA 把每一列的殘差與其餘所有列相比較，按 0–10 分制排序評分：

- 高 CQA（約 9–10）：成交價遠高於預測值——很可能狀況 / 品質 / 吸引力較優
- 低 CQA（約 0–1）：成交價遠低於預測值——很可能狀況較差或屬於困境出售
- CQA 約為 5：接近中位數

11 RCA 計算與下載

RCA (Residual Constraint Approach, 殘差約束法) 工作流程在 **Appraisal** (估價) 與 **Market Area Analysis** (市場區域分析) 模式下、模型擬合之後可用。估價標的必須位於第 1 列。

11.1 開啟 RCA 對話方塊

點選側邊欄第 8 節中的 **Calculate RCA Adjustments & Download** 按鈕，會彈出一個模式對話方塊，其中包含：

- **Score type** (得分類型)：CQA 或 CQA per SF (若已指定 living_area)。按平方英尺計的選項會用居住面積對估價標的的殘差作縮放。
- **Subject CQA Score** (估價標的 CQA 得分)：接受 0.00 至 10.00 之間的取值。該欄位初始為空；若已匯入過 earthUI 模型，你在 earthUI 中使用的得分類型與取值會從共用的專案設定中沿用過來。

點選 **Generate** (產生) 即可計算並下載。Excel 檔案會寫入作用中專案的 mgcv 輸出資料夾 (<os>_out_mgcv)。

11.2 CQA 得分內插

1. 先對各可比案例的 CQA 得分與殘差排序
2. 用線性內插 (approx()) 把你給定的 CQA 值映射為一個殘差
3. 估價標的價值 = 模型預測值 + 內插所得殘差
4. 權重為 0 的列按與估價標的具有相同殘差來處理

11.3 輸出欄位

欄位	說明
subject_value	模型預測值 + 內插所得殘差
subject_cqa	使用者輸入的 CQA 得分
<var>_adjustment	估價標的貢獻 - 可比案例貢獻 (以美元計)
residual_adjustment	估價標的殘差 - 可比案例殘差
net_adjustments	全部調整額之和
gross_adjustments	全部調整額絕對值之和
residual_adj_pct	殘差調整額佔銷售價格的百分比
net_adj_pct	淨調整額佔銷售價格的百分比
gross_adj_pct	總調整額佔銷售價格的百分比
adjusted_sale_price	實際銷售價格 + 淨調整額

計算成功完成後，按鈕上會出現一個白色勾號。

12 銷售比較網格

銷售比較網格 (Sales Comparison Grid) 僅在 **Appraisal** (估價) 模式下、且已計算 RCA 調整額之後可用。它會產生一份格式化的 Excel 活頁簿，適合納入估價報告之中。

12.1 可比案例選擇對話方塊

點選 **Generate Sales Grid & Download** 按鈕會開啟一個模態對話方塊，列出所有可比銷售案例。可比案例分為兩組：

- **Recommended comps** (建議的可比案例，已預先勾選)：總調整額 < 銷售價格的 25%，按銷售時距排序 (最近的在前)。最多 30 個。
- **Other comps** (其他可比案例，預設不勾選)：其餘的可比案例，按總調整百分比排序。

選定你的可比案例後，點選 **Generate Sales Grid** 即可建立活頁簿。

必需的特殊類型：必須已指定 `contract_date` 與 `living_area`。建議的特殊類型：`latitude`、`longitude` 與 `lot_size`——沒有它們網格也能產生，但某些欄位會留空。

12.2 活頁簿版式

產生的活頁簿含多個工作表，每張表並排容納 3 個可比案例：

- 表頭列：估價標的與可比案例的地址、銷售價格、成交日期以及市場停留天數
- 分組列：區位 (緯度、經度、區域)、宗地面積 (地塊面積、宗地尺寸) 與年齡 (實際年齡、有效年齡)，帶組標題與子條目
- 模型變數列：每個平滑項一列，顯示估價標的貢獻、各可比案例貢獻以及調整額
- 殘差特徵列：空白且未鎖定的儲存格，供估價師填寫模型未能刻畫的特徵 (例如狀況、品質、景觀)
- 彙總列：淨調整百分比、總調整百分比與調整後銷售價格，均帶有可即時計算的 Excel 公式

12.3 工作表保護與殘差儲存格

每張工作表都受到保護以防誤改，但殘差特徵的取值儲存格被顯式解鎖。這既讓估價師能夠為模型之外的特徵輸入取值，又保留了由公式驅動的調整額計算。

13 下載報告

報告產生採用兩步式 Quarto 工作流程，取代了早期版本中的一步式報告。

13.1 第 1 步：產生 Quarto 報告

Generate Quarto Report（產生 Quarto 報告）按鈕（模型擬合後可用）會把一個自包含的 .qmd 包寫入作用中專案的 mgcv 輸出資料夾。該包含有 Quarto 原始檔、預先渲染好的圖形檔案（PNG 與 PDF）、一個裝載了全部表格與模型元資料的 report_data.rds，以及用於 Word 輸出的 reference.docx 樣式模板。由於該包是自包含的，你可以在渲染之前編輯 .qmd 原始檔，或把它與其他 Quarto 文件組合起來。

圖形在產生該包時即已預先渲染；若某張圖渲染失敗，程式會寫入一張佔位圖，包中其餘部分照常產生。

13.2 第 2 步：轉換 Quarto 報告

Convert Quarto Report（轉換 Quarto 報告）小節透過 quarto::quarto_render() 把一個 .qmd 檔案渲染為一種或多種輸出格式。它預設指向最近產生的那個包，但 **Browse**（瀏覽）按鈕允許你選擇任意 .qmd 檔案。可用格式有三種：

- **HTML** —— 自包含的 HTML 文件。在所有平台上均可使用，無需額外軟件。建議作為預設選項。
- **Word (.docx)** —— 使用隨包提供的 reference.docx 樣式模板渲染。適合編輯與分發。
- **PDF** —— 需要安裝 LaTeX（參見下文的系統需求）。若未偵測到 LaTeX，PDF 選項會自動從格式核取方塊中隱去。

早期的一步式渲染器 render_gam_report() 與 export_gam_docx() 仍保留，可供程式化呼叫。

13.3 報告內容

產生的報告包括：

- 資料集與模型：資料檔案、樣本數、family、method 以及擬合資訊
- 結果摘要： R^2 、 $CV R^2$ 、已解釋離差、AIC 及相關的擬合指標
- 模型方程：模型方程，其中平滑項用 f_i 記號表示、線性項給出顯式係數，並附平滑函數定義。較長的方程在 PDF 中會跨多行折排。
- 摘要表：平滑項（EDF、F、p 值）與參數項（估計值、標準誤、t、p 值）
- 已啟用的交互：模型中所含張量交互項與因子—平滑交互項的矩陣（僅在存在交互項時顯示）
- 變數重要性：按 F 統計量（平滑項）或 $|t|$ （參數項）排序的長條圖
- 平滑效應（所有類型的項）：
 - 單變數平滑圖（互動式圖形的 ggplot2 靜態版本）
 - 因子—平滑的多線條圖
 - 交互熱圖（ t_i 與 t_e 項）
 - 參數項圖（數值型用散點圖，因子型用長條圖）
- 預測變數相關性熱圖
- 診斷：殘差診斷以及實際值對預測值的散點圖
- mgcv 摘要：已擬合模型的完整文字輸出
- 記號附錄：一個不編號的結尾小節，說明報告中所用的價值貢獻函數記號

成功產生後，Generate Quarto Report 按鈕上會出現一個白色勾號。

14 匯出模型函數

mgcvUI 可以把已擬合 GAM 的平滑函數匯出為多種語言的獨立程式碼。每個平滑項都在一個 200 點網格上求值，並以帶線性內插的查表形式匯出，從而使該模型可在 R 之外使用，且不依賴 mgcv。

在目前版本中，函數匯出是透過程式化方式進行的。匯出用的輔助函數接受一個已擬合的結果物件（由 `fit_gam()` 產生），並寫出下述檔案：

- `generate_r_code()`、`generate_python_code()` 與 `generate_cpp_code()` ——各語言的原始檔
- `export_functions_sqlite()` ——SQLite 查表資料庫
- `export_functions_zip()` ——把這四種格式的任意組合打包成一個 zip 壓縮包

14.1 R 函數

檔案：`gam_functions.R`

- 頭部註解，含反應變數、 R^2 、CV R^2 與模型公式。
- 一個 INTERCEPT 常數。
- 每個平滑變數對應一個 `approxfun()`，帶預先算好的 x/y 向量。
- 線性項係數以具名常數形式給出。
- 用法：把截距、各個 `g_var(value)` 以及各個 `coef_var * value` 相加。

14.2 Python 函數

檔案：`gam_functions.py`

- numpy 匯入語句。
- INTERCEPT 常數。
- 每個變數對應：`_x_var` 與 `_fx_var` 陣列，以及使用 `np.interp()` 的 `g_var(x)` 函數。
- 線性項係數以 `COEF_*` 常數形式給出。

14.3 C++ 標頭檔

檔案：`gam_functions.hpp`

- 僅含標頭檔，使用 `#pragma once`。
- `gam_functions` 命名空間。
- 使用 `std::lower_bound` 的 `interp()` 輔助函數。
- 每個變數對應：一個帶 `static const` 陣列的內聯函數。
- 線性項係數以 `constexpr double` 形式給出。

14.4 SQLite 資料庫

檔案：`gam_functions.sqlite`

- 表 `smooth_grids`：欄位為 `(variable, x, fx)` ——即 200 點查表。
- 表 `model_info`：`response`、`r_squared`、`cv_rsqr`、`dev_explained`、`aic`、`n_obs`、`family`、`method`、`formula`。
- 表 `linear_terms`：`variable`、`coefficient`。
- 表 `intercept`：截距值。

15 與 earthUI 和 glmnetUI 的比較

earth 應用程式及其後處理模組構成一套用於迴歸建模的工具集。它們共用相同的資料格式、特殊欄位類型、RCA 工作流程與示範資料集，但使用不同的建模引擎。

15.1 主要差異

特性	mgcvUI	earthUI	glmnetUI
方法	GAM (mgcv)	MARS/Earth (earth)	彈性網絡 (glmnet)
關係形態	平滑非線性	分段線性	線性
變數選擇	select=TRUE 收縮	逐步修剪	Lasso 懲罰
交互項	張量積 (te/ti)	鉸鏈函數乘積 (2-3 階)	兩兩乘積
偏效應	帶信賴區間的平滑曲線	g 函數 (分段)	線性斜率
earth 流程	從 earth 匯入節點	節點的來源	匯入鉸鏈函數基底
程式碼匯出	R、Python、C++、SQLite	R、Python、C++、SQLite	—
報告格式	Word、PDF、HTML	Word、PDF、HTML	Word、PDF、HTML

有一項 earthUI 獨有的差異值得一提：earthUI 含有一個實驗性的 Prolog/DCG 功能，可從自由文字的 MLS 備註中提取結構化的不動產屬性。

15.2 共有功能

三個工具都提供：

- 共用的 regProj 專案樹（預設為 ~/regProj）——專案、輸入檔案以及按方法分列的輸出在同族應用之間共用
- CSV 與 Excel 匯入，並作 snake_case 欄位名轉換
- 支援 30 多個國家的區域設定，涵蓋 CSV 分隔符、小數點符號與紙張尺寸
- 特殊欄位指定 (contract_date、living_area、area、latitude、longitude 等)
- 依據估價基準日自動計算 sale_age
- 深色 / 淺色模式切換 (Nord 主題)
- 設定的持久化 (SQLite 與 / 或 localStorage)
- 相關性熱圖、變數重要性、診斷
- CQA 評分與 RCA 調整工作流程
- 帶可比案例選擇與格式化 Excel 輸出的銷售比較網格
- 相容同一批示範資料集 (Appraisal_1.csv)

15.3 建議的工作流程

- 先用 earthUI 發現重要變數、非線性關係與交互作用。earth 的自動建模提供了最佳的起點。
- 再用 mgcvUI 精化，如果你希望得到更平滑的偏效應。匯入 earthUI 的節點即可無縫過渡。
- 使用 glmnetUI，當你需要一個帶激進變數選擇或係數符號約束的純線性模型時。
- 在同一份資料上比較三者，以評估非線性模型所增添的複雜度是否值得。

16 示範資料集：Appraisal_1.csv

16.1 說明

該示範資料集與 earthUI 和 glmnetUI 共用。它位於 mgcvUI 原始碼的 demo_mls/ 資料夾中；若已安裝 earthUI，則可用：

```
demo_file <- system.file("extdata", "Appraisal_1.csv", package = "earthUI")
```

該檔案包含約 1,500 條住宅銷售記錄，取自一份模擬的 MLS 匯出，標的不動產位於第 1 列，可比銷售案例位於其餘各列。資料代表一個多區域市場中的獨棟住宅銷售，不動產面積、年齡與區位各不相同。

這並非真實資料，但以北加州一個實際存在的街區為藍本。所有識別資訊均已改寫或刪除。

16.2 各欄位

欄位	類型	特殊類型	說明
sale_type	character	sale_type	用於區分估價標的與可比銷售案例的記錄類型碼
weight	numeric	weight	觀測權重 (0 = 不參與擬合)
id	numeric	display_only	內部記錄 ID
property_id	numeric	display_only	MLS 不動產識別碼
listing_id	character	display_only	MLS 掛牌編號
parcel_number	character	display_only	縣估價官宗地編號 (APN)
street_address	character	display_only	不動產地址
city_name	character	display_only	市
postal_code	character	display_only	郵遞區號
county_name	character	display_only	縣
contract_date	Date	contract_date	銷售合約日期 (用於計算 sale_age)
sale_age	numeric	—	從合約日期至估價基準日的天數 (已預先計算)
coe_date	Date	display_only	交割日期
listing_status	character	display_only	掛牌狀態 (例如「Sold」, 已售)
sale_price	numeric	(目標變數)	銷售價格——反應變數
rent	numeric	—	月租金 (用於多目標模型)
list_price	numeric	display_only	掛牌價格
original_list_price	numeric	display_only	初始掛牌價格
living_sqft	numeric	living_area	以平方英尺計的總居住面積
ppsf	numeric	display_only	每平方英尺單價
beds_total	integer	—	臥室數
baths_total	numeric	—	衛浴總數 (例如 2.5 = 2 個全衛 + 1 個半衛)
lot_size	numeric	lot_size	以平方英尺計的地塊面積
area_id	integer	area	MLS 區域識別碼
area_text	character	display_only	區域名稱
age	numeric	actual_age	以年計的不動產年齡
year_built	integer	display_only	建成年份
latitude	numeric	latitude	緯度 (入模前四捨五入至 3 位小數)
longitude	numeric	longitude	經度 (入模前四捨五入至 3 位小數)
latitude6	numeric	display_only	全精度緯度
longitude6	numeric	display_only	全精度經度
garage_spaces	integer	—	車庫車位數
fp_count	integer	—	壁爐數
no_of_stories	numeric	—	樓層數
style	character	—	建築風格
view	character	—	景觀類型 (例如「Neighborhood」街區景、「Hills」山景)
days_on_market	integer	dom	市場停留天數
listing_date	Date	listing_date	掛牌日期

欄位	類型	特殊類型	說明
sale_concessions	numeric	concessions	賣方讓價

16.3 建議的快速上手步驟

1. 啟動 mgcv 例程：`earthUI::launch_mgcv()`
2. 在第 1 節中新建一個用途為 **For Appraisal** 的專案
3. 把 `Appraisal_1.csv` 放入該專案的輸入資料夾，並在第 2 節中選中它（必要時點選 **Refresh**）
4. 選擇 `sale_price` 作為目標變數
5. 如需比例式調整，可將 Response Transform 設為 **Log (natural)**
6. 按上表所示指派特殊類型
7. 納入以下預測變數：`sale_age`、`living_sqft`、`baths_total`、`lot_size`、`area_id`（作為因子）、`age`、`latitude`、`longitude`、`garage_spaces`
8. 保留預設設定：`REML`、`gamma = 1.2`、`tp` 基底、`k = 自動`、`select = TRUE`
9. 點選 **Fit Mgcvm GAM Model**
10. 檢視 Summary、Contribution（平滑圖）與 Diagnostics 頁籤
11. 下載輸出（第 7 步），檢視 CQA 排名
12. 以約 ~5.00 的 CQA 得分計算 RCA 調整額（第 8 步）
13. 用建議的可比案例產生銷售比較網格（第 9 步）

17 系統需求與疑難排解

17.1 受支援的平台

mgcvUI 可在 **macOS**、**Windows** 與 **Linux** 上執行。本應用程式在 macOS 配合 RStudio 的環境中開發，也主要在該環境中測試。與平台相關的說明見下文。

17.2 R 版本

要求 $R \geq 4.1.0$ 。強烈建議使用 RStudio Desktop (2023.06 或更高版本) ——它自帶 **Quarto** 與 **pandoc** (渲染報告所需)，並為啟動本應用程式提供了便利的環境。

17.3 必需的 R 套件

以下套件會作為相依套件自動安裝：

套件	用途
shiny	Web 應用程式框架
mgcv	GAM 模型擬合
valengrCore	共用的估價核心 (特殊角色、earthUI 設定沿用)
ggplot2, plotly	靜態與互動式貢獻圖
gratia	平滑項提取與診斷
DT	互動式資料表
quarto, rmarkdown	Quarto 報告的產生與渲染
readxl	Excel 檔案匯入
shinyFiles	檔案與資料夾瀏覽器
RSQLite, DBI	設定與專案資料庫 (SQLite)
jsonlite	變數狀態的序列化

17.4 選用相依套件

元件	何時需要
LaTeX (TinyTeX、MiKTeX 或 MacTeX)	僅用於產生 PDF 報告。若未偵測到 LaTeX 安裝，PDF 選項會自動從格式下拉選單中隱去。安裝指令： <code>tinytex::install_tinytex()</code>
sysfonts + showtext	供 ggplot2 圖形使用的 Roboto Condensed 字型。若不可用或處於離線狀態，應用程式會自動退回到系統無襯線字型。
earth ($\geq 5.3.0$)	僅在使用 earthUI 匯入管道時需要。獨立使用時不需要。
callr	背景模型擬合 (保持應用程式的回應性)。沒有它則擬合約步執行。
writexl, openxlsx	Excel 下載 (模型輸出、RCA、銷售網格)。遺漏時應用程式會提示。
bslib	Bootstrap 5 主題 (Nord Light/Dark)。
officer	早期的程式化 Word 匯出 (<code>export_gam_docx()</code>)。

17.5 與平台相關的說明

17.5.1 macOS

問題最少。系統庫建議使用 Homebrew 安裝。若需要 PDF 報告：

```
# In R:
install.packages("tinytex")
tinytex::install_tinytex()
```

17.5.2 Windows

配合 RStudio Desktop 執行良好。要點：

- 企業／受限機器：若暫存目錄或使用者資料目錄受到限制，設定的持久化與報告產生可能失敗。應用程式會在 R 控制台中顯示警告，但仍繼續執行，只是不儲存設定。
- **LaTeX**：從 R 中安裝 TinyTeX（指令與上文 macOS 相同），或從 miktex.org 安裝 MiKTeX。
- 過長的檔案路徑：Windows 有 260 個字元的路徑長度上限。請保持工作目錄路徑簡短。

17.5.3 Linux

各發行版之間差異最大。Ubuntu/Debian 使用者在編譯 R 套件之前可能需要先裝好系統庫：

```
sudo apt install libcurl4-openssl-dev libssl-dev \
  libxml2-dev libsqlite3-dev libfontconfig1-dev
```

若需 PDF 報告：使用 `tinytex::install_tinytex()`，或從系統套件管理器安裝 `texlive-full`。

對於無顯示設備的無頭伺服器，`plotly` 互動式圖形在報告中會以靜態圖形作為退回方案渲染。`Shiny` 應用程式本身仍需要一個網頁瀏覽器連接。

17.6 優雅降級

`mgcvUI` 的設計使其即便在缺少選用元件時也能工作：

遺漏的元件	行為
無 LaTeX	PDF 選項從報告格式下拉選單中隱去。HTML 與 Word 仍然可用。
無網絡 / 字型載入失敗	Roboto Condensed 字型由系統無襯線字型替代。控制台記錄一條提示資訊。
無 RSQLite / 唯讀檔案系統	設定不會跨會話儲存。控制台顯示警告。除此之外應用程式正常執行。
<code>gratia::appraise</code> 失敗	診斷頁籤與報告中顯示佔位資訊。可在控制台中改用 <code>mgcv::gam.check()</code> 。
暫存目錄不可寫	報告產生失敗，並給出明確的錯誤訊息。可透過把 <code>TMPDIR</code> 環境變數設為可寫位置來解決。

17.7 疑難排解

17.7.1 「PDF option is not available」(PDF 選項不可用)

未偵測到 LaTeX 安裝。請在 R 中執行：

```
install.packages("tinytex")
tinytex::install_tinytex()
```

然後重啟應用程式，PDF 選項便會出現。

17.7.2 「Settings will not persist」(設定將不會儲存)

SQLite 資料庫目錄無法建立或無法寫入。請檢查以下位置的權限：

- macOS/Linux : ~/Library/Application Support/R/mgcvUI/ 或 ~/.local/share/R/mgcvUI/
- Windows : %APPDATA%/R/data/mgcvUI/

17.7.3 「Diagnostics plot unavailable」(診斷圖不可用)

`gratia::appraise()` 函數在複雜模型(含許多張量交互項、非標準分配族)上偶爾會失敗。這是一個已知的上游問題。變通辦法：在 R 控制台中執行 `mgcv::gam.check(model)` 以直接檢視診斷圖。

17.7.4 「Report error: number of items to replace...」(報告錯誤：待替換的項數……)

這是早期版本中一個與 `gratia` 套件相關的已知問題。請確保你安裝的是最新版本的 `mgcvUI`。目前版本會捕獲該錯誤，並在產生報告時為診斷部分放置一個佔位內容。

17.7.5 7880 連接埠已被佔用

若應用程式因 7880 連接埠被佔用(來自上一次會話)而無法啟動，請執行：

```
# macOS/Linux:
lsof -ti:7880 | xargs kill
```

```
# Windows (PowerShell):
Stop-Process -Id (Get-NetTCPConnection -LocalPort 7880).OwningProcess
```

17.7.6 在 Linux 上安裝套件失敗

缺少系統庫。請安裝上文 Linux 一節中列出的開發標頭檔，然後重新執行 `install.packages()`。

參考文獻

- Wood, S. N. (2003). "Thin-plate regression splines". In: *Journal of the Royal Statistical Society (B)* 65.1, pp. 95–114. DOI: [10.1111/1467-9868.00374](https://doi.org/10.1111/1467-9868.00374).
- (2004). "Stable and efficient multiple smoothing parameter estimation for generalized additive models". In: *Journal of the American Statistical Association* 99.467, pp. 673–686. DOI: [10.1198/016214504000000980](https://doi.org/10.1198/016214504000000980).
- (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". In: *Journal of the Royal Statistical Society (B)* 73.1, pp. 3–36. DOI: [10.1111/j.1467-9868.2010.00749.x](https://doi.org/10.1111/j.1467-9868.2010.00749.x).
- (2017). *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman and Hall/CRC.
- Wood, S. N., N. Pya, and B. Säfken (2016). "Smoothing parameter and model selection for general smooth models (with discussion)". In: *Journal of the American Statistical Association* 111, pp. 1548–1575. DOI: [10.1080/01621459.2016.1180986](https://doi.org/10.1080/01621459.2016.1180986).

基礎概念

概述

繁體中文版

基礎概念 (*Foundations*) • *The Valuation Engineer*

Bert Craytor (主編) • 2026 年 8 月 4 日

版本 1.0.0

DOI: [10.67330/82vssv81](https://doi.org/10.67330/82vssv81)

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。

《基礎概念》(*Foundations*) 專欄發表對現代估價方法論基礎概念的精確、可引用的闡述。各條目選取執業估價師會遇到的術語——有時以其他名稱出現，有時甚至未被察覺——並將其置於形式化的脈絡之中：這些概念從何而來、其技術含義是什麼、在實務中出現於何處，以及它們為何關乎可辯護性。

創刊號確立了特徵價格 (hedonic) 框架的六個核心概念，其編排使每個條目都建立在前面條目所確立的詞彙之上。這一系列從《異質商品》(*Heterogeneous Good*，即房地產之所以需要估價方法論的根本屬性) 出發，經由《特徵空間》(*Characteristics Space*) 與《特徵價格函數》(*Hedonic Price Function*，即組織房地產所代表的屬性組合的形式化工具)，再到《隱含價格》(*Implicit Price*) 與《特徵價格迴歸》(*Hedonic Regression*，即如何從市場證據中還原每個屬性的邊際貢獻)，最後到達《潛變數》(*Latent Variable*，即任何特徵價格模型中未被還原的殘差訊號究竟代表什麼)。

一個貫穿全部六個條目的通篇範例——來自加利福尼亞州帕西菲卡 (Pacifica, California) 某社區的八個可比案例——構成本期的教學主線。

為何是這六個概念，為何按此順序

六個條目的選擇並非任意。每個概念都是現代特徵價格方法論的承重構件，每個概念都確實出現在日常估價實務之中，而且每個概念被誤解或未被命名的頻率之高，使得將其寫成可引用的形式具有實際價值。

條目的排序由概念間的依賴關係決定。異質商品是根基：它命名了房地產（以及中古車、勞動市場和其他若干商品）的一種屬性，正是這一屬性使整個特徵價格框架成為必要。特徵空間是將異質商品表示為向量的形式化工具；沒有它，任何進一步的論述都無法精確展開。特徵價格函數是將特徵空間中的點映射到價格的均衡對象。隱含價格是它的偏導數——銷售比較網格上的每一項調整都可以被解讀為對這些邊際貢獻的斷言，無論估價師是否如此稱呼它們。特徵價格迴歸是從觀測到的交易中還原隱含價格的經驗估計量；正是在這一條目中，理論轉變為資料分析。潛變數則從迴歸殘差開始的地方接續：那些未被觀測的屬性，其效應以有結構的訊號而非可丟棄的雜訊的形式存留於殘差之中。

按此順序閱讀的讀者，應當能獲得一條連續的思想線索：從「估價方法論究竟為何存在？」一直到「我們應當如何看待擬合模型中的殘差？」——以作者的經驗而言，大多數估價師需要多年實務的零散積累才能走完這段旅程。

通篇範例

一個由加利福尼亞州帕西菲卡八個住宅可比案例構成的資料集貫穿全部六個條目：

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

這個範例被有意設計得足夠小，可以徒手推理；同時又足夠豐富，可以演練每一個概念。條目 001 考察可比案例 A 與 B 之間 \$145,000 的差價，將其視為兩項特徵（景觀與地塊）糾纏於單一配對銷售觀測中的聯合貢獻。條目 002 將全部八個可比案例置於一個四維特徵空間中的點，並說明樸素的距離度量會對它們作出錯誤排序。條目 003 為該社區的特徵價格函數提出一個函數形式。條目 004 以解析方式還原隱含價格，並將最初的 A 與 B 差價分解為景觀與地塊貢獻的 \$127,039，另有 \$18,000 未獲解釋。條目 005 在 R 中正式擬合該模型，並報告完整的 `lm()` 輸出，包括標準誤與殘差。條目 006 考察模型無法解釋的那一個殘差——可比案例 H 上 +\$46,455 的偏差——並將其與潛變數的解釋框架聯繫起來，這一框架正是作者在別處發展的殘差約束法（Residual Constraint Approach）方法論的基礎。

該資料集可在本期的儲存庫中取得，路徑為 `shared/data/pacifica_comps.csv`，供希望重現任何計算的讀者使用。

如何閱讀本期

初次接觸特徵價格方法論的讀者應按順序閱讀六個條目。條目間的依賴是實實在在的：每個條目都假定讀者已掌握前面條目所確立的詞彙，跳躍式閱讀會招致條目本身無法預料的困惑。

先前接觸過特徵價格迴歸的讀者——包括大多數修過高等統計學課程、使用過自動估價模型 (AVM) 或讀過大量估價文獻的估價師——可將本期用作溫習與參考。各條目在寫作上力求可被獨立引用；本刊其他文章若提及隱含價格，可透過 DOI 連結到條目 004，而無需重新推導該概念。

對殘差約束法方法論特別感興趣的讀者，應先讀條目 006 以理解潛變數的解釋框架，然後回到條目 003–005 了解殘差約束法 (RCA) 所擴展的特徵價格工具，再到條目 001–002 了解支撐一切的基礎詞彙。

關心什麼使估價調整具有可辯護性這一問題的讀者會發現：依次閱讀各條目的「為何關乎可辯護性」部分，便相當於一篇關於基於調整的估價方法之方法論地位的短文。這篇短文並未在本期任何地方被明確地組裝起來，但它是每個條目的潛在主題。

編輯立場

《基礎概念》不是教科書式專欄。每個條目短到可以一次讀完，面向已經熟稔估價實務的讀者寫作，旨在給出概念的精確表述而非面面俱到的闡釋。其目標是提供一個穩定的參照物，供本刊其他文章引用。

條目經編輯部審閱，並可能附有特邀領域專家的評論。條目以創用 CC 姓名標示授權條款 (Creative Commons Attribution) 發布，並透過 Crossref（作為規範引用）與 Zenodo（作為版本化典藏）兩個管道獲得 DOI。每個條目都可能隨時間推移而被打磨——小幅修訂更新 Zenodo 版本，重大修訂則同時更新兩個 DOI 並附編輯說明描述改動。Crossref DOI 始終解析到最新版本；特定歷史版本可透過 Zenodo 定址。

未來各期將把《基礎概念》的詞彙擴展到相鄰領域：空間相依、條件期望值、內生性、遺漏變數偏誤、可識別性，以及區分代理變數與潛變數的形式化工具。歡迎讀者就值得闡述的概念提出建議，來函請寄編輯部。

本期六個條目：

1. 異質商品 (Heterogeneous Good) —— 估價方法論為何必須存在。
2. 特徵空間 (Characteristics Space) —— 屬性組合的向量表示。
3. 特徵價格函數 (Hedonic Price Function) —— 從特徵到價格的均衡映射。
4. 隱含價格 (Implicit Price) —— 每項特徵的邊際市場價格。
5. 特徵價格迴歸 (Hedonic Regression) —— 從觀測交易中估計該函數，附完整的 R 實例演算。
6. 潛變數 (Latent Variable) —— 從殘差中還原有結構的訊號。

參考文獻。

Court, A. T. (1939). Hedonic Price Indexes with Automotive Examples. In *The Dynamics of Automobile Demand*. General Motors Corporation.

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.

Craytor, B. (2025). Residual Constraint Approach: Framework & Protocol. Zenodo. [doi:10.5281/zenodo.14787917](https://doi.org/10.5281/zenodo.14787917)

附錄：Prolog 形式化

以下機器可讀的配套檔案 (000-overview.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。

```
%% 000-overview.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Overview
%% Title: Overview
%% Author: Bert Craytor, Editor
%% Piece ID: foundations_overview
%% Kind: editorial_piece (not a concept; sits outside the dep graph)
%%
%% The Overview is an editorial piece, not a defined concept. Its .pl
%% companion records the issue's structure (which entries appear in
%% what reading order), the running example, and the reader pathways
%% the editor proposes for readers with different backgrounds.

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

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

% --- From the dictionary (loaded via :- use_module above) -----
% term(foundations_overview, "Overview", foundations_editorial,
```

```

%      "Editorial overview of a Foundations issue. Frames the concepts
%      covered, situates them within appraisal practice, and is not
%      itself a defined concept in the dependency graph.",
%      published, 'dictionary-2026.1').
%
% Plus the six concept terms covered by this issue (heterogeneous_good,
% characteristics_space, hedonic_price_function, implicit_price,
% hedonic_regression, latent_variable).
% -----

:- use_terms([ foundations_overview,
                heterogeneous_good, characteristics_space,
                hedonic_price_function, implicit_price,
                hedonic_regression, latent_variable,
                residual_constraint_approach ]).
:- use_relations([depends_on/2]).

% =====
% PIECE METADATA
% =====

article_piece(foundations_overview).
article_kind(overview).
article_issue(volume(1), number(1), year(2026)).
article_id('foundations.2026.overview').

% =====
% ISSUE STRUCTURE: the six concept entries in dependency-ordered sequence
% =====
%
% The published reading order. Each entry assumes vocabulary established
% in the previous ones. This is the same as queries:reading_order/1
% derived from the depends_on graph; recording it here lets readers and
% downstream tools query the issue's structure without invoking the
% topological-sort machinery.

issue_entry(1, heterogeneous_good).
issue_entry(2, characteristics_space).
issue_entry(3, hedonic_price_function).
issue_entry(4, implicit_price).
issue_entry(5, hedonic_regression).
issue_entry(6, latent_variable).

% Each entry's one-line theme as stated in the Overview's closing list.
entry_theme(heterogeneous_good,
    "Why valuation methodology must exist at all.").
entry_theme(characteristics_space,
    "The vector representation of bundle attributes.").
entry_theme(hedonic_price_function,
    "The equilibrium mapping from characteristics to prices.").
entry_theme(implicit_price,
    "The marginal market price of each characteristic.").
entry_theme(hedonic_regression,
    "Estimating the function from observed transactions, with a full R worked example
    ↪.").
entry_theme(latent_variable,
    "Recovering structured signal from residuals.").

% =====

```

```
% THE RUNNING EXAMPLE
% =====
%
% A single dataset of 8 Pacifica residential comparables runs through
% all six entries. The data lives in kb/pacifica_comps.pl and is
% consulted by each entry that uses it. The Overview's role here is
% only to point at it.

running_example_dataset(pacifica_comps).
running_example_n_comps(8).
running_example_location(pacifica_california).

% Which entries exercise the dataset.
:- consult('..../kb/pacifica_comps').

% =====
% READER PATHWAYS
% =====
%
% The editor proposes four pathways through the issue, calibrated to
% the reader's background. These are advisory orderings; the canonical
% reading order is issue_entry/2 above.

reader_pathway(new_to_hedonic_methodology,
    "Read entries 1-6 in order. Dependencies are real and skipping invites confusion
    ↪ .",
    [heterogeneous_good, characteristics_space, hedonic_price_function,
     implicit_price, hedonic_regression, latent_variable]).

reader_pathway(prior_exposure_to_hedonic_regression,
    "Use as a refresher and atomic reference. Read in any order.",
    [hedonic_regression, hedonic_price_function, implicit_price,
     characteristics_space, heterogeneous_good, latent_variable]).

reader_pathway(interested_in_rca_methodology,
    "Read 006 first for the framing, then 003-005, then 001-002.",
    [latent_variable, hedonic_price_function, implicit_price,
     hedonic_regression, heterogeneous_good, characteristics_space]).

reader_pathway(interested_in_adjustment_defensibility,
    "Read the 'Why it matters for defensibility' sections in entry order. Forms an
    ↪ implicit short essay.",
    [heterogeneous_good, characteristics_space, hedonic_price_function,
     implicit_price, hedonic_regression, latent_variable]).

% =====
% EDITORIAL POSTURE
% =====

editorial_posture(
    "Foundations is not a textbook department. Each entry is short, written for an
    ↪ audience already fluent in appraisal practice, and designed to give a precise
    ↪ formulation rather than a comprehensive exposition.").

publication_model(
    [ license(cc_by),
      crossref_doi(canonical_citation),
      zenodo_doi(versioned_archive),
      revision_policy("Minor revisions bump Zenodo version; major revisions bump both
```

```
↪DOIs with editorial note.") ]).

% =====
% FUTURE DIRECTIONS
% =====
%
% Concepts the editor flags as deserving treatment in future Foundations
% issues. Recorded here so the journal's planning is machine-queryable.

future_topic(spatial_dependence).
future_topic(conditional_expectation).
future_topic(endogeneity).
future_topic(omitted_variable_bias_extended).
future_topic(identification).
future_topic(proxy_vs_latent_variable).
```

異質商品

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/s1scsv65](https://doi.org/10.67330/s1scsv65)

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。 異質商品 (heterogeneous good) 是指不可相互替換的商品：每一件商品都是買方分別賦予價值的特徵的獨特組合，而這些特徵本身並不單獨交易。

直觀闡釋。 同質商品 (homogeneous good) 是任何一件都可以的商品：一桶 WTI 原油、一蒲式耳二號黃玉米、電網上的一千瓦小時電力、某支股票的一股。這些商品背後的分級標準或規格的存在，正是為了使單件商品可以相互替換，從而使價格適用於類別而非特定的那一件。異質商品則是具體那一件很重要的商品——是這棟房子，而不是那棟房子；是這輛中古車，而不是旁邊那一輛。

在技術意義上，兩個特點使商品成為異質商品。第一，買方在意多個可區分的屬性，而這些屬性在市場上的各件商品之間存在差異。第二，這些屬性因商品的生產技術而被捆綁在一起，無法單獨購買。你不可能單獨購買 200 平方英尺的建築面積；你不可能脫離一筆土地而取得一個學區。特徵成捆而來，市場為整個組合定價，而不是為特徵本身定價。

房地產是異質商品的典型代表，並且異常鮮明地體現了這一屬性。區位特徵是完全不可複製的：任何兩筆土地都不可能佔據同一位置，因此每一筆房地產至少擁有一項屬性（其座標）將它與市場上的所有其他房地產區分開來。實體、法律與環境屬性又在此之上疊加了更多的變異維度，使得整體的特徵向量對每一筆房地產而言都是獨一無二的。

估價師在何處遇到它。市場比較法 (sales comparison approach) 中的每一項調整，都是對估價對象與可比案例乃異質商品這一事實的隱含承認。如果房地產是同質的，成交價格本身就可以直接比較，根本不需要比較網格。網格是一種透過對組合構成差異作出調整，來調和不同組合價格的嘗試。值得注意的是網格不做什麼：它從不將成交價格分解為各特徵的價值貢獻。傳統網格沒有價值貢獻一欄——只有逐項直接估計的調整。

成本法透過從組成部分重建組合來因應異質性——基礎、構架、裝修、折舊——並對每一部分分別定價。收益法則透過將組合抽象為單一的現金流輪廓來因應異質性，由資本化率與收益率把組合層面的異質性吸收進一個源自市場的乘數。三種估價方法都是對「房地產是異質商品」這一根本事實的方法論回應。它們的區別在於如何因應這一事實，而不在於是否面對它。

估價師在非正式場合同樣會遇到異質性：每當他們思考可比案例的選擇、配對銷售分析 (paired-sales analysis) 或「同類相比」的比較時。尋找在儘可能多的屬性上相似的可比案例，這一直覺正是實務工作者對異質性的回應：將估價對象與可比案例之間的組合距離最小化，組合中未被觀測的成分對成交價格差異的貢獻便隨之減小。

為何關乎可辯護性。 房地產的異質性正是估價方法論必須作為一門學科存在的原因。如果房地產可以相互替換，市場價值 (market value) 便可從標價直接觀測，估價將淪為查表。正因為並非如此，價值必須從相關但不相同的房地產的價格中推斷得出，而每一步推斷都是一次需要說明理由的方法論選擇。

正式承認異質性，可以釐清 USPAP 及各準則條款所關注的可辯護性所依賴的若干實務問題。可比案例的選擇，是在特徵空間中挑選與估價對象足夠接近、以致從中還原的隱含價格可以推廣的樣本。一項調整是否合理，取決於其背後的隱含價格能否從市場證據中得到支持。各估價方法

之間的最終調和，反映的是估價師對哪種方法最可信地因應了這筆房地產的異質性的判斷。一旦「異質商品」被點明為深層條件，上述每一個可辯護性問題都會變得更加清晰。

估價實例演算。 加利福尼亞州帕西菲卡 (Pacifica, California) 同一社區最近的兩筆成交：

屬性	可比案例 A	可比案例 B
建築面積 (平方英尺)	1,650	1,650
臥室 / 衛浴	3 / 2	3 / 2
地塊 (平方英尺)	9,000	7,200
景觀	有	無
成交價格	\$1,425,000	\$1,280,000

在同質商品的視角下，這是同一筆房地產：建築面積相同、房間數相同、社區相同。在異質商品的視角下，它們是兩個不同的組合，在買方分別賦予價值的兩項特徵（地塊與景觀）上存在差異。\$145,000 的價格差是這兩項屬性的聯合貢獻——兩個隱含價格糾纏在同一個配對之中。

要將隱含價格逐一還原，需要更多的配對，最好是每次只改變一項屬性。再有三四筆涵蓋有景觀 / 無景觀、大地塊 / 小地塊各種組合的成交，兩個隱含價格便開始可以分別識別。當足夠多的可比案例提供了足夠的變異時，特徵價格迴歸 (hedonic regression) 便可正式完成這一解纏；而當變異有限時，估價師只能侷限於區間化的配對銷售推理，其可靠性明確取決於各組合在特徵空間中彼此有多接近。

這個例子說明了為什麼異質性是根本條件：即使是兩棟「完全相同」的房子也並不相同，而估價的方法論工具之所以存在，正是為了處理它們之間的距離。

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：特徵空間 (*characteristics space*)；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 (001-heterogeneous-good.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。

```
%% 001-heterogeneous-good.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 001
%% Title: Heterogeneous Good
%% Author: Bert Craytor
%% Concept: heterogeneous_good
%% Version: 0.1.0-draft
%%
%% This file is the Prolog galley for the article: it encodes the
%% article's main claims, worked example, and inference rules in a form
%% that other Prolog code in the journal can consult, query, and verify
%% against. Read alongside the entry.tex / entry.pdf prose.
```

```

% =====
% DICTIONARY IMPORT
% =====
%
% The journal-wide dictionary is the source of truth for term definitions
% and relation schemas. The use_module below pulls those facts in at
% runtime; the comments that follow reproduce the relevant entries so a
% reader skimming this file in isolation can see what's been imported.

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

% --- From the dictionary (loaded via :- use_module above) -----
% term(heterogeneous_good, "Heterogeneous Good", economic_theory,
%     "Heterogeneous goods are goods that are not interchangeable
%     because each is a distinct bundle of characteristics that buyers
%     value separately, and whose characteristics are not separately
%     traded.", published, 'dictionary-2026.1').
%
% 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(implicit_price, "Implicit Price", economic_theory,
%     "The implicit price of a characteristic is the partial derivative
%     of the hedonic price function with respect to that characteristic,
%     evaluated at a point in the characteristics space; it represents
%     the marginal market price of an incremental unit of the
%     characteristic, holding all other characteristics fixed.",
%     published, 'dictionary-2026.1').
%
% term(hedonic_price_function, ...).
%
% relation(depends_on, 2, "depends on", [foundations_concept, foundations_concept],
%     foundations_methodology, "...", accepted, 'dictionary-2026.1').
% relation(related, 2, ...).
%
% These facts are NOT declared in this file; the use_module above is
% what makes them available at runtime.
% -----

:- use_terms([ % Foundations concepts (terms/foundations.pl)
    heterogeneous_good, characteristics_space, implicit_price,
    hedonic_price_function,
    % Practice terms (terms/valuation-practice.pl)
    real_estate, homogeneous_good,
    sales_comparison_approach, cost_approach, income_approach,
    comp_selection, paired_sales_analysis, adjustment,
    reconciliation, uspap, standards_rule ]).
:- use_relations([related/2, depends_on/2]).

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

article_concept(heterogeneous_good).
article_issue(volume(1), number(1), year(2026)).

```

```

article_id('foundations.2026.001').

% =====
% ARTICLE-LOCAL VOCABULARY
% =====
%
% Terms the article needs that are NOT yet in the dictionary, and that
% the editor decided are too narrow to promote. These are field names
% for the worked-example property records — not concepts in the
% journal's ontology.

local_term(gla_sqft).
local_term(bedrooms).
local_term(baths).
local_term(lot_sqft).
local_term(view).

% =====
% FORMAL DEFINITION (restated as Prolog rule)
% =====
%
% From the formal definition: a good is heterogeneous IF buyers care
% about multiple distinguishable attributes that vary across items in
% the market, AND those attributes are bundled by production rather
% than separately traded.

%! heterogeneous(?Good) is nondet.
%
% True when Good satisfies both bundling criteria. The two predicates
% buyers_value_multiple_attributes/1 and attributes_bundled_by_production/1
% must be asserted for specific goods elsewhere; here they characterize
% the formal definition itself.

heterogeneous(Good) :-
    buyers_value_multiple_attributes(Good),
    attributes_bundled_by_production(Good).

% =====
% INTUITIVE FRAMING — CONTRASTS WITH HOMOGENEOUS GOODS
% =====

% Canonical examples of homogeneous goods cited in the article.
canonical_homogeneous_example(wti_crude).
canonical_homogeneous_example(no_2_yellow_corn).
canonical_homogeneous_example(kilowatt_hour).
canonical_homogeneous_example(common_stock_share).

% Their distinguishing feature: grading or specification standards make
% individual items interchangeable.
interchangeable_via_grading(wti_crude).
interchangeable_via_grading(no_2_yellow_corn).
interchangeable_via_specification(kilowatt_hour).
interchangeable_via_specification(common_stock_share).

% Canonical heterogeneous goods.
canonical_heterogeneous_example(real_estate).
canonical_heterogeneous_example(used_car).
canonical_heterogeneous_example(unique_machinery).

```

```

% Why real estate exhibits heterogeneity unusually strongly:
% locational characteristics are perfectly non-replicable.
buyers_value_multiple_attributes(real_estate).
attributes_bundled_by_production(real_estate).
non_replicable_attribute(real_estate, location).

% =====
% APPRAISER ENCOUNTERS
% =====
%
% The three classical valuation approaches are all methodological
% responses to heterogeneity. They differ in HOW they navigate the
% bundle, not in WHETHER they confront it.

valuation_approach(sales_comparison_approach).
valuation_approach(cost_approach).
valuation_approach(income_approach).

responds_to_heterogeneity(sales_comparison_approach,
    "Reconciles the prices of distinct bundles by adjusting for composition
    ↪ differences; the grid records only adjustments, estimated directly, and never
    ↪ decomposes a price into feature value contributions.").
responds_to_heterogeneity(cost_approach,
    "Reconstructs the bundle from priced components: foundation, framing, finishes,
    ↪ depreciation.").
responds_to_heterogeneity(income_approach,
    "Abstracts the bundle into a single cash-flow profile; cap/yield rates absorb
    ↪ heterogeneity.").

% Practitioner-level encounters with heterogeneity.
practitioner_response_to_heterogeneity(comp_selection,
    "Minimize bundle distance between subject and comp in characteristics space.").
practitioner_response_to_heterogeneity(paired_sales_analysis,
    "Isolate an attribute's contribution by holding other attributes constant.").
practitioner_response_to_heterogeneity(adjustment,
    "Quantify the price impact of one differing attribute at a time.").

% =====
% DEFENSIBILITY
% =====
%
% Why naming "heterogeneous good" sharpens USPAP defensibility questions.

defensibility_consequence(comp_selection,
    "Comps must be close enough in characteristics space that recovered implicit
    ↪ prices generalize.").
defensibility_consequence(adjustment,
    "Each adjustment must be supportable from market evidence of the implicit price."
    ↪ .
defensibility_consequence(reconciliation,
    "Choice among approaches reflects which best navigates THIS property's
    ↪ heterogeneity.").

% Root claim: valuation methodology exists as a discipline BECAUSE
% real estate is heterogeneous. If it were homogeneous, value would be
% directly observable from posted prices.
root_claim(
    "Valuation methodology exists as a discipline because real estate is heterogeneous
    ↪; in the homogeneous case, value would be directly observable.").

```

```

% =====
% WORKED EXAMPLE — PACIFICA COMPS
% =====
%
% Two recent sales in the same Pacifica, California neighborhood. The
% example demonstrates that even "identical" houses are not identical:
% the $145,000 price spread is the joint contribution of differences in
% lot size and view, two implicit prices entangled in a single pair.

% The 8-comp Pacifica dataset is shared across entries 001-006 and
% lives in kb/pacifica_comps.pl. Consulting it here brings property/1,
% in_market/2, attribute_value/3, and sale_price/2 into this article's
% namespace.
:- consult('../..kb/pacifica_comps').

% --- Derived facts from the example -----

% The attributes that differ between the two comps.
differing_attribute(comp_a, comp_b, lot_sqft).
differing_attribute(comp_a, comp_b, view).

% The price differential and what it's attributable to.
price_differential(comp_a, comp_b, 145000).
attributed_to(price_differential(comp_a, comp_b, 145000),
    [implicit_price_of(lot_sqft), implicit_price_of(view)]).

% --- Article's argument about identification -----

% A single pair cannot separately identify the two implicit prices —
% they are entangled in the single $145K observation.
joint_identification_only(price_differential(comp_a, comp_b, 145000),
    [lot_sqft, view]).

% Separate identification requires more pairs varying one attribute at
% a time. With sufficient variation, hedonic regression performs the
% disentanglement formally; with limited variation, the appraiser uses
% bracketed paired-sales reasoning, whose reliability depends on how
% close the bundles are in characteristics space.
identification_strategy(few_pairs,
    "Bracketed paired-sales reasoning. Reliability depends on bundle proximity in
    ↪ characteristics space.").
identification_strategy(many_pairs_with_variation,
    "Hedonic regression (see hedonic_regression)").

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

%! comparable(?A, ?B) is nondet.
%
% A and B are comparable when they are distinct properties in the same
% market. This is the operational definition of "comp" used implicitly
% in the article.

comparable(A, B) :-
    property(A),
    property(B),
    A \== B,

```

```

    in_market(A, M),
    in_market(B, M).

%! attribute_difference(?A, ?B, ?Attr, ?DA, ?DB) is nondet.
%
% A and B differ on Attr, with DA the value at A and DB the value at B.
% Use this to enumerate the dimensions on which two comps actually vary.

attribute_difference(A, B, Attr, DA, DB) :-
    comparable(A, B),
    attribute_value(A, Attr, DA),
    attribute_value(B, Attr, DB),
    DA \== DB.

%! identifiable_implicit_price(?Attr) is nondet.
%
% An attribute's implicit price is identifiable from the observed comps
% if there exists at least one pair of comparables differing ONLY in
% that attribute. With our two-comp example this fails for lot_sqft AND
% view (both differ in the same pair), which is exactly the article's
% point.

identifiable_implicit_price(Attr) :-
    comparable(A, B),
    findall(X, attribute_difference(A, B, X, _, _), Diffs),
    Diffs == [Attr].

% =====
% CROSS-REFERENCES
% =====
%
% Concepts the article cross-references (corresponding to the
% \foundationsref{} macros in entry.tex).

cross_reference(characteristics_space).
cross_reference(hedonic_price_function).
cross_reference(implicit_price).

```

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: 10.67330/3yyymh010

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).
```

特徵價格函數

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: [10.67330/9x1a0489](https://doi.org/10.67330/9x1a0489)

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。特徵價格函數 (hedonic price function) 是從異質商品特徵空間中的點到成交價格的均衡映射，由競爭性市場中買方出價函數與賣方要價函數的包絡構成。

直觀闡釋。我們現在有了兩個要素：異質商品（特徵的組合）與特徵空間（每個組合是其中一點的向量空間）。特徵價格函數將二者結合起來。它是為特徵空間中每個點 $\mathbf{z}$ 賦予一個價格的函數 $p(\mathbf{z})$ ，告訴我們市場為具有該特定屬性組合的商品支付多少。

Rosen (1974) 的關鍵洞見在於： $p(\mathbf{z})$ 不是由任何單個買方或賣方選定的，而是從市場均衡中湧現出來的。每個買方都有一個出價函數 (bid function)，表達在其偏好與所得給定的情況下，願為任一組合支付的最高價格。每個賣方都有一個要價函數 (offer function)，表達在其成本與替代方案給定的情況下，願意接受的最低價格。特徵價格函數是這些出價與要價的包絡——買賣雙方實際成交所沿的邊界。

對不動產而言， $p(\mathbf{z})$ 回答的是這樣一個問題：如果你給定一個完整的特徵組合，市場會為它給出什麼價格？答案並非任意。它反映了市場中全體買方的總體偏好與全體賣方的總體成本與替代方案，並經由買方與組合的均衡匹配來調節。

特徵價格函數是特徵價格分析的核心理論對象。隱含價格是它的偏導數。特徵價格迴歸是對它的經驗估計。任何銷售比較調整的可辯護性最終都可歸結為關於這一函數的一個命題。

估價師在何處遇到它。估價師無時無刻不在對 $p(\mathbf{z})$ 進行推理——同樣，通常沒有叫出它的名字。若干實務活動都是對特徵價格函數的隱性操作：

市場比較法中的調整。當估價師因估價對象有景觀而可比案例 B 沒有，而將可比案例 B 上調 \$100,000 時，估價師是在斷言一個關於 $p(\mathbf{z})$ 的命題：即在其他特徵保持不變的情況下，當景觀特徵從 0 變為 1 時，函數值增加 \$100,000。這是關於 p 在景觀維度上局部斜率的陳述。

各估價方法之間的調和。成本法自下而上構建價值指示，將複製組合的各組成部分成本相加。收益法從現金流推導價值，而在均衡狀態下，現金流反映產生類似現金流的組合的價格。市場比較法則直接從相似組合的觀測成交中讀取價格。在運作良好的市場中，三種方法應當收斂於 $p(\mathbf{z})$ 上的同一個點。當它們出現分歧時，估價師必須判斷哪種方法對估價對象在特徵空間中所處位置的函數測量得最可靠。

市場分析與時間調整。隨著供給、需求、融資與預期的變化，特徵價格函數會隨時間發生移動。每年 6% 的增值率是關於 $p(\mathbf{z})$ 整體如何移動的命題；而「有景觀溢價的不動產比無景觀不動產增值更快」的發現，則是關於函數在特徵空間中非均勻移動的命題。

為何關乎可辯護性。明確點出特徵價格函數，可以釐清若干根本性的可辯護性問題：

函數是否存在？Rosen 框架假定存在一個競爭性市場，且交易量足以使均衡函數有意義。在交易稀薄的市場、快速變化的市場，或買方群體特質化的次市場中，該函數可能定義不清或不穩定。在此類市場中執業的估價師應當承認，其推斷工具是在一個不穩固的理論基礎上運行的。

函數是否處處相同？Rosen 的框架為每個市場區隔給出一個函數，而不是全域一個函數。帕西菲卡 (Pacifica) 與舊金山 (San Francisco) 是不同的市場，各有不同的特徵價格函數；即便在同一座城市內，獨棟住宅與公寓大廈也是不同的次市場。將次市場混為一談的估價師，實際上是在隱含地假設同一個函數橫跨兩者，而這通常是錯誤的。

p 取什麼函數形式？該函數不必是線性、對數線性或任何其他特定形式。Rosen 只證明了 p 是出價與要價的包絡；他沒有給定其形狀。經驗上， p 表現出非線性（新增一間臥室的邊際價值往往隨建築面積而變化）、交互作用（景觀溢價取決於地塊朝向）以及閾值效應（低於某一品質門檻的廚房可能招致與其狀況評分不成比例的折價）。可辯護的估計要求要麼採用能夠從市場中得到支持的函數形式，要麼使用能夠從資料中還原函數形狀的彈性估計量。

估價實例演算。對於所研究的帕西菲卡社區，我們將使用一個簡單的可加函數形式，下一條目將對其作更充分的演練：

$$p(\mathbf{z}) = \beta_0 + \beta_{\text{GLA}} z_{\text{GLA}} + \beta_{\text{lot}} z_{\text{lot}} + \beta_{\text{view}} z_{\text{view}} + \beta_{\text{cond}} z_{\text{cond}}.$$

這一形式內嵌了兩個強假設：每項特徵的貢獻對其水準是線性的，且各特徵之間不存在交互作用。兩個假設都是可檢驗的，也經常被違背；我們在此將其作為出發點採用。

對於本期運行示例中的八個可比案例，後續條目的分析將估計出一組 β 值，得到大致如下的函數：

參數	估計值
β_0 (截距)	\$80,777
β_{GLA} (每平方英尺)	\$575
β_{lot} (每平方英尺)	\$22.70
β_{view} (有/無)	\$86,179
β_{cond} (每級)	\$49,824

將這些值代入可加形式，該函數便可為四維特徵空間中的任何一點預測價格。代入可比案例 A 的座標，得到的預測價格接近其觀測到的 \$1,425,000；代入位於 (1700, 7500, 1, 2) 的假想估價對象，則得到該不動產的價值指示。

本條目的要點不在於這些具體數字，而在於概念上的一步：在我們估計它之前，該函數在原則上就已存在。函數是市場的屬性。我們的估計是還原它的一種嘗試。這種還原可能是不完全的、有偏的或含雜訊的，接下來的三個條目所討論的正是如何負責任地完成這一還原。

參考文獻。

- Court, A. T. (1939). Hedonic Price Indexes with Automotive Examples. In *The Dynamics of Automobile Demand*. General Motors Corporation.
- 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.

交叉引用：異質商品 (*heterogeneous good*)；特徵空間 (*characteristics space*)；隱含價格 (*implicit price*)；特徵價格迴歸 (*hedonic regression*)。

附錄：Prolog 形式化

以下機器可讀的隨附檔案 (003-hedonic-price-function.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。

```

%% 003-hedonic-price-function.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 003
%% Title: Hedonic Price Function
%% Author: Bert Craytor
%% Concept: hedonic_price_function
%% 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_price_function, "Hedonic Price Function", economic_theory,
% "The hedonic price function is the equilibrium mapping from
% points in the characteristics space of a heterogeneous good to
% transaction prices, formed as the envelope of buyer bid
% functions and seller offer functions in a competitive market.",
% published, 'dictionary-2026.1').
%
% term(bid_function, "Bid Function", economic_theory, ...).
% term(offer_function, "Offer Function", economic_theory, ...).
% term(sales_comparison_approach, ...).
% term(cost_approach, ...).
% term(income_approach, ...).
%
% formalized_by(hedonic_price_function, rosen_1974).
% depends_on(hedonic_price_function, heterogeneous_good).
% depends_on(hedonic_price_function, characteristics_space).
% -----

:- use_terms([ hedonic_price_function, heterogeneous_good,
               characteristics_space, implicit_price, hedonic_regression,
               bid_function, offer_function,
               sales_comparison_approach, cost_approach, income_approach,
               adjustment, reconciliation ]).
:- use_relations([related/2, depends_on/2, formalized_by/2]).

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

article_concept(hedonic_price_function).
article_issue(volume(1), number(1), year(2026)).
article_id('foundations.2026.003').

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

local_term(equilibrium).
local_term(envelope).
local_term(functional_form).
local_term(market_segment).
local_term(linearity).
local_term(separability).

```

```

local_term(time_adjustment).
local_term(coefficient).
local_term(intercept).

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

attributed_formalization(hedonic_price_function, rosen_1974).

% =====
% CONSTRUCTION: ENVELOPE OF BID AND OFFER FUNCTIONS
% =====
%
% The Rosen (1974) construction:  $p(z)$  is the envelope of buyer bid
% functions and seller offer functions.

primitive_of_construction(bid_function, buyer).
primitive_of_construction(offer_function, seller).

constructed_as_envelope_of(hedonic_price_function,
                           [bid_function, offer_function]).

emerges_from(hedonic_price_function, market_equilibrium).

% Property that flows from the construction: the function is not
% chosen by any single market participant.
property(hedonic_price_function, not_set_by_individual_participant).
property(hedonic_price_function, equilibrium_object).
property(hedonic_price_function, defined_per_market_segment).

% =====
% APPRAISER ENCOUNTERS — IMPLICIT MANIPULATIONS OF  $p(z)$ 
% =====

implicit_manipulation_of_hedonic(
    adjustment,
    sales_comparison_approach,
    "Asserts a local slope of  $p$  in the relevant characteristic dimension.").
implicit_manipulation_of_hedonic(
    reconciliation,
    multi_approach_valuation,
    "Three approaches should converge on the same point of  $p$ ; divergence requires
    ↪ editorial judgment.").
implicit_manipulation_of_hedonic(
    time_adjustment,
    market_analysis,
    " $p$  shifts over time; appreciation rates are statements about that shift.").

% =====
% DEFENSIBILITY QUESTIONS
% =====
%
% The article names three questions the hedonic-price-function framing
% sharpens.

defensibility_question(function_existence,
    "In thin or rapidly-changing markets, the equilibrium function may be poorly
    ↪ defined; the inferential machinery operates on a shaky foundation.").

```



```
; A == view, V == no -> N = 0
; V = N
).

% =====
% THE CENTRAL CONCEPTUAL MOVE
% =====
%
% The function exists, in principle, before estimation. The function is
% a property of the market; estimation is recovery.

claim(function_precedes_estimation,
    "The hedonic price function is a property of the market and exists conceptually
    ↪ prior to any specific fitted model. Estimation is recovery, and recovery may be
    ↪ partial, biased, or noisy.").

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

cross_reference(heterogeneous_good).
cross_reference(characteristics_space).
cross_reference(implicit_price).
cross_reference(hedonic_regression).
```

隱含價格

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/md6wke17](https://doi.org/10.67330/md6wke17)

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。一項特徵的隱含價格 (implicit price) 是特徵價格函數對該特徵的偏導數，在特徵空間中某一點取值；它代表在其他特徵保持不變的情況下，該特徵增加一個單位的邊際市場價格。

直觀闡釋。特徵價格函數 $p(\mathbf{z})$ 為每個組合賦予一個價格。然而，估價實務中的大多數問題並不關乎完整組合的價格，而是關乎組合中某個具體組成部分的價格。景觀值多少錢？多出一百平方英尺能增加多少？「良好」與「優秀」狀況之間的差異貢獻多少？

隱含價格回答的正是這些問題。這一術語源自 Rosen (1974)，他證明了差異化產品的觀測價格連同其特徵「定義了一組隱含的或『特徵』價格」；Lancaster (1966) 則提供了將商品視為特徵組合的底層觀點。對每項特徵 z_i 而言，隱含價格是在其他所有特徵保持不變時， $p(\mathbf{z})$ 隨 z_i 變化的變化率：

$$\pi_i(\mathbf{z}) = \frac{\partial p(\mathbf{z})}{\partial z_i}.$$

隱含價格的單位是美元每單位特徵——建築面積每平方英尺多少美元、地塊每平方英尺多少美元、狀況每提升一級多少美元。對於景觀這樣的二元特徵，隱含價格就是從無景觀 (0) 變為有景觀 (1) 所支付的全部溢價。

「隱含」一詞點明了核心事實：這些價格沒有被標價。沒有人以 \$575 出售「一平方英尺建築面積」。這一價格隱含在確實被標價的均衡組合價格之中，而特徵價格分析工具的存在正是為了把它提取出來。

隱含價格是特徵價格框架中與估價直接相關的產出。銷售比較網絡中的調整，原則上就是隱含價格的估計值乘以特徵差異。每一項調整都是對特徵空間中相關點處 π_i 取值的一次押注。

估價師在何處遇到它。銷售比較調整網絡就是一張隱含價格應用表，即便估價師並不這樣稱呼它。網絡中「因建築面積比估價對象少 200 平方英尺而 +\$10,000」的調整，是在斷言：在特徵空間的相關位置上，建築面積的隱含價格為每平方英尺 \$50。「因景觀較差而調整 \$100,000」的網絡條目，是在斷言二元景觀特徵的隱含價格為 \$100,000。市場比較法的全部可辯護性，都取決於這些隱含價格能否從市場中得到支持。然而在實踐中，估價師是直接估計調整額的——依據配對案例、經驗或同業做法——從未計算過底層的價值貢獻；傳統網絡只記錄調整本身。隱含價格的解讀是對這些調整所斷言內容的一種詮釋，而不是對其推導方式的描述。

這一命名也闡明了為什麼配對銷售分析是推導調整的基礎技術。恰好只在一項特徵上不同的兩個可比案例，可以直接估計該特徵的隱含價格：價格差除以特徵差。配對銷售分析的可靠性，取決於這兩個案例在特徵空間中是否真的足夠接近，以致從中還原的隱含價格能夠推廣到估價對象。

隱含價格有三個對實務重要的性質：

它們是局部的。在 1,650 平方英尺處的建築面積隱含價格，未必與在 2,500 平方英尺處相同。特徵價格函數一般是非線性的，其偏導數隨特徵空間中的位置而變化。由 1,500 平方英尺的配對銷售分析推導出的調整，未必適用於 3,000 平方英尺的估價對象。

它們是有條件的。海濱次市場中景觀的隱含價格，與內陸次市場中的並不相同。偏導數是在其他特徵保持不變的情況下求取的，但該導數的取值取決於其他特徵被固定在什麼位置。

它們依賴於市場。隱含價格是均衡價格，因而反映交易發生時買方群體的偏好。即使特徵本身沒有變化，市場的變動也會改變隱含價格。時間調整實際上是在承認：昨天的隱含價格未必是今天的。

為何關乎可辯護性。隱含價格的視角揭示了非正式網格可能掩蓋的調整可辯護性問題：

你處在特徵空間中的什麼位置？只有在估價對象於特徵空間中所處位置附近估計出的隱含價格才是可辯護的。由內陸可比案例推導出的景觀溢價被套用於海濱估價對象，等於在函數行為未知的特徵空間區域上作外推。

局部幾何結構是什麼？如果兩項特徵存在交互作用——例如景觀在更大的地塊上更值錢——那麼將它們分別作可加處理就是一種模型設定錯誤。景觀的隱含價格依賴於地塊面積，忽略交互作用的調整低估了函數的實際結構。

隱含價格是如何還原的？配對銷售分析、迴歸係數、業界經驗法則與「專業判斷」指向的是同一個量——隱含價格——但其認識論地位截然不同。估價師的可辯護性取決於說明每項調整所使用的隱含價格出自哪個來源。

估價實例演算。對於採用上一條目可加設定的帕西菲卡 (Pacifica) 運行示例，隱含價格就是線性函數的係數：

$$\pi_{\text{GLA}} = \beta_{\text{GLA}}, \quad \pi_{\text{lot}} = \beta_{\text{lot}}, \quad \pi_{\text{view}} = \beta_{\text{view}}, \quad \pi_{\text{cond}} = \beta_{\text{cond}}.$$

使用條目 003 的估計係數：

特徵	隱含價格	單位
建築面積	\$575	每平方英尺
地塊	\$22.70	每平方英尺
景觀	\$86,179	每 (有/無)
狀況	\$49,824	每級

這些隱含價格蘊含了一次配對銷售分析。回到條目 001 中可比案例 A 與可比案例 B 的差異：

- 可比案例 A 有景觀 ($\Delta z_{\text{view}} = +1$)；貢獻 +\$86,179。
- 可比案例 A 的地塊大 1,800 平方英尺 ($\Delta z_{\text{lot}} = +1,800$)；貢獻 $1,800 \times \$22.70 = +\$40,860$ 。
- 其他所有特徵相同。
- 預測價格差： $\$86,179 + \$40,860 = \$127,039$ 。

觀測到的價格差為 \$145,000。隱含價格分解解釋了其中約 \$127,000；其餘 \$18,000 反映了小資料集中的抽樣變異（更一般地說，反映了潛變數條目將要討論的那類未觀測因素）。

有兩個操作層面的要點值得注意。第一，線性假設承擔著實實在在的工作：在非線性的 $p(\mathbf{z})$ 下，同樣的地塊面積差在小地塊上可能比在大地塊上更值錢，上述可加分解便會有偏。第二，地塊每平方英尺 \$22.70 的隱含價格是增加一平方英尺的邊際價值，而不是地塊的總價值。一塊 9,000 平方英尺的地塊對組合價格的總貢獻為 $9,000 \times \$22.70 = \$204,300$ ，但再增加一平方英尺只使組合價格提高 \$22.70。

這一區分——邊際與總量——是日常估價語言中大量混淆的根源，也是明確命名隱含價格所帶來的最重要的概念收穫之一。

參考文獻。

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

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

- Freeman, A. M., III (1979). Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues. *Scandinavian Journal of Economics*, 81(2), 154–173.
- Taylor, L. O. (2003). The Hedonic Method. In *A Primer on Nonmarket Valuation* (pp. 331–393). Kluwer Academic Publishers.
- Palmquist, R. B. (2005). Property Value Models. In *Handbook of Environmental Economics* (Vol. 2, pp. 763–819). North-Holland.
- Triplet, J. E. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*. OECD Publishing.

交叉引用：異質商品 (*heterogeneous good*)；特徵空間 (*characteristics space*)；特徵價格函數 (*hedonic price function*)；特徵價格迴歸 (*hedonic regression*)。

附錄：Prolog 形式化

以下機器可讀的隨附檔案 (004-implicit-price.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。

```
%% 004-implicit-price.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 004
%% Title:   Implicit Price
%% Author:  Bert Craytor
%% Concept: implicit_price
%% 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(implicit_price, "Implicit Price", economic_theory,
%      "The implicit price of a characteristic is the partial
%      derivative of the hedonic price function with respect to that
%      characteristic, evaluated at a point in the characteristics
%      space; it represents the marginal market price of an
%      incremental unit of the characteristic, holding all other
%      characteristics fixed.", published, 'dictionary-2026.1').
%
% term(hedonic_price_function, ...).
% term(paired_sales_analysis, ...).
% term(adjustment, ...).
%
% formalized_by(implicit_price, rosen_1974).
% depends_on(implicit_price, hedonic_price_function).
% -----

:- use_terms([ implicit_price, hedonic_price_function, characteristics_space,
               heterogeneous_good, hedonic_regression,
               sales_comparison_approach, paired_sales_analysis, adjustment,
               comp_selection ]).
:- use_relations([related/2, depends_on/2, formalized_by/2]).
```

```

% Provenance: the term originates with Rosen (1974).
formalized_by(implicit_price, rosen_1974).

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

article_concept(implicit_price).
article_issue(volume(1), number(1), year(2026)).
article_id('foundations.2026.004').

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

local_term(partial_derivative).
local_term(marginal).
local_term(total).
local_term(locality).
local_term(conditionality).
local_term(market_dependence).

% =====
% IMPLICIT PRICE AS PARTIAL DERIVATIVE
% =====
%
%  $\pi_i(z) = \partial p(z) / \partial z_i$ 

implicit_price_definition(
    "Implicit price of characteristic i at point z = partial derivative of p with
    ↪ respect to z_i, evaluated at z.").

unit_of_implicit_price(continuous, "dollars per unit of the characteristic").
unit_of_implicit_price(binary, "the full premium between the two states").

% =====
% THREE PROPERTIES OF IMPLICIT PRICES (per the article)
% =====

property_of_implicit_prices(local,
    "The implicit price of GLA at 1,650 sf is not necessarily the same as at 2,500 sf.
    ↪ p is generally nonlinear; partial derivatives vary with location.").
property_of_implicit_prices(conditional,
    "Implicit price depends on where the OTHER characteristics are held fixed. View
    ↪ premium on a beachfront submarket differs from inland.").
property_of_implicit_prices(market_dependent,
    "Implicit prices are equilibrium prices reflecting the current buyer pool. Market
    ↪ shifts change implicit prices even when characteristics do not.").

% =====
% IMPLICIT-PRICE / ADJUSTMENT EQUIVALENCE
% =====
%
% Every adjustment in a sales-comparison grid can be READ as asserting an
% implicit price applied to a characteristic difference. This is an
% interpretation of what the adjustment asserts, not a description of how
% traditional adjustments are derived in practice.

adjustment_as_implicit_price_application(

```

"Interpretation: adjustment magnitude = (implicit price of attribute at relevant
 ↳ point) * (characteristic difference between comp and subject) --- what the
 ↳ adjustment asserts, not how it is derived.").

adjustment_derivation_in_practice(

"Traditional adjustments are estimated directly --- from matched pairs, experience
 ↳ , or peer practice --- without computing the underlying value contributions; the
 ↳ traditional grid records only the adjustments themselves.").

```
% =====
% LINEAR-FORM IMPLICIT PRICES (from entry 003's adopted specification)
% =====
%
% Under the additive linear specification, implicit prices ARE the
% regression coefficients. The article reproduces them:

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

% Implicit price per characteristic, as inherited from the additive
% form of p adopted in entry 003.
implicit_price_of(gla_sqft, 575.00, "per square foot").
implicit_price_of(lot_sqft, 22.70, "per square foot").
implicit_price_of(view, 86179.03, "per yes/no").
implicit_price_of(cond, 49824.33, "per step on the ordinal scale").

% =====
% PAIRED-SALES DECOMPOSITION: COMP A vs COMP B
% =====
%
% Returns to the article-001 paired comparison and decomposes it using
% implicit prices.
%
% Comp A and Comp B share GLA, bedrooms, baths, condition.
% They differ in: view (+1) and lot_sqft (+1800).
% Predicted differential = 86179 + 40860 = 127039.
% Observed differential = 145000.
% Unexplained = 17961 (~$18,000).

paired_sales_decomposition(comp_a, comp_b,
  [ contribution(view, 86179),
    contribution(lot_sqft, 40860) ],
  127039,
  145000,
  17961).

% Derived from the data + coefficients (for the linter to type-check
% against the dataset rather than the hand-computed values).
contribution_for_pair(A, B, Attr, Contribution) :-
  coord(A, Attr, VA),
  coord(B, Attr, VB),
  Diff is VA - VB,
  Diff =\= 0,
  implicit_price_of(Attr, Pi, _),
  Contribution is Diff * Pi.

% Tiny coord helper for symbolic view/cond values.
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
).

% =====
% MARGINAL vs TOTAL CONTRIBUTION
% =====
%
% The article emphasizes the distinction. The implicit price of lot is
% $22.70 per square foot (marginal); the TOTAL contribution of Comp A's
% 9,000-sf lot is 9,000 * 22.70 = $204,300.

distinction(marginal_vs_total,
    "Implicit price is MARGINAL ($22.70 per additional sqft of lot). Total
    ↪ contribution of a 9,000-sf lot is 9,000 * $22.70 = $204,300. Confusing them is a
    ↪ common source of error in everyday appraisal language.").

%! total_contribution(?Property, ?Attribute, -Total) is det.
%
% Total dollar contribution of a characteristic to a property's bundle
% price under the linear specification.

total_contribution(P, A, Total) :-
    coord(P, A, V),
    implicit_price_of(A, Pi, _),
    Total is V * Pi.

% =====
% DEFENSIBILITY QUESTIONS RAISED
% =====

defensibility_question(location_in_space,
    "An implicit price is defensible only when estimated near the subject's location
    ↪ in characteristics space; a view premium from inland comps applied to a beachfront
    ↪ subject extrapolates into unknown territory.").
defensibility_question(local_geometry,
    "If view interacts with lot size, treating them as additively separable
    ↪ misspecifies the function. The implicit price of view depends on lot.").
defensibility_question(epistemic_provenance,
    "Paired-sales analysis, regression coefficient, industry rule, and professional
    ↪ judgment all point to the same quantity but have different epistemic standing.
    ↪ Defensibility requires declaring the source.").

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

cross_reference(heterogeneous_good).
cross_reference(characteristics_space).
cross_reference(hedonic_price_function).
cross_reference(hedonic_regression).

```

特徵價格迴歸

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

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。特徵價格迴歸 (hedonic regression) 是根據異質商品市場中觀測到的交易對特徵價格函數進行的估計，得到進入模型的各項特徵的隱含價格的經驗估計值。

直觀闡釋。前面的條目將特徵價格函數及其隱含價格描述為理論對象。該函數在原則上存在，但在實務中我們只能觀測到有限的一組交易：成交的組合，以及它們成交的價格。特徵價格迴歸就是連接二者的經驗橋梁。

設定很直接。每筆交易提供一個 $(\mathbf{z}_i, p_i)$ 對：成交組合的特徵向量及其成交價格。這些資料對的集合是（未被觀測的）特徵價格函數的一個樣本。迴歸 (regression) 為該樣本擬合一個函數形式，估計出刻畫該函數的參數。參數估出之後，擬合函數的偏導數便給出隱含價格的估計值。

在迴歸執行之前，必須作出三項建模選擇：

1. 哪些特徵進入模型？這一選擇決定了被估函數的維度。遺漏相關特徵會使其效應滲入與之相關的已納入特徵的係數之中（遺漏變數偏誤）。
2. 什麼函數形式？線性、對數線性、雙對數、多項式、基底函數展開、半參數、機器學習。這一選擇限定了被估函數可以呈現的形狀，以及它能還原哪些隱含價格模式。
3. 什麼估計量？普通最小平方法 (OLS)、廣義最小平方法、穩健 M 估計、分位數迴歸、空間計量估計量、懲罰迴歸。這一選擇決定了估計量在最佳化什麼，以及哪些假設支撐其推論。

每項選擇在方法論上都有實質意義，也都帶有可辯護性上的後果。

估價師在何處遇到它。特徵價格迴歸是大量估價 (mass appraisal) 的基礎，也是大多數自動估價模型 (AVM) 的引擎。當郡稅務估價員每年對 10 萬筆地塊重新估值時，承擔工作的是對近期成交所作的特徵價格迴歸。當 Zillow 或 Redfin 給出自動估計時，產出它的是一個特徵價格式模型（往往是樹集成而非 OLS，但在概念上仍是特徵價格估計量）。當貸款機構的次級市場 AVM 判定某宗房地產的再融資通過與否時，涉及的也是同一套機制。

單宗房地產估價也越來越多地借助特徵價格迴歸，即便最終的價值指示來自銷售比較。由社區樣本迴歸估出的隱含價格，能夠遠比估價師直覺或經驗法則百分比更可辯護地支持調整幅度。迴歸係數成為網格中所用調整的證據，而非意見。

為何關乎可辯護性。基於迴歸的估計具有與配對銷售分析不同的可辯護性輪廓。優勢在於：迴歸使用全部可得成交，而非一次只用兩筆；它可以在同時控制許多其他特徵的情況下分離出單項特徵的效應；它給出量化每個隱含價格不確定性的標準誤。代價在於：迴歸強加了一個可能與真實價格函數不符的函數形式；它比配對銷售分析需要更多資料；其結果可能對樣本與模型設定的選擇敏感。

基於迴歸的估價中反覆出現若干可辯護性問題：

樣本是否恰當？迴歸估計的是特徵價格函數在樣本所處市場區隔與時間段內的運作。用內陸成交作迴歸再套用於海濱估價對象，無論統計結果多麼漂亮，都是可辯護性上的失敗。

模型設定是否恰當？線性設定是一個強假設。如果真實函數在某一門檻以上對建築面積呈現報酬遞減，線性擬合就會在該區間內錯估隱含價格。診斷、殘差圖與備選設定是可辯護工作流程的組成部分。

標準誤意味著什麼？標準誤刻畫的是在假定的模型設定下的抽樣變異。它不涵蓋模型設定錯誤、遺漏變數偏誤或樣本選擇問題。錯誤設定下的小標準誤，比正確設定下的大標準誤更危險。

結果能否與其他方法互相印證？一個可辯護的、基於迴歸的隱含價格，應當與配對銷售分析的提示一致，與成本法的構件成本所暗示的一致，也與熟悉該市場的估價師的看法一致。與三者都相左的迴歸係數是一份邀請函——邀請你去調查，而不是去發表。

估價實例演算。現在我們可以在八個帕西菲卡 (Pacifica) 可比案例上演練完整的特徵價格迴歸機制。資料集位於 `shared/data/pacifica_comps.csv`。

R 程式碼：

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

輸出：

```
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
```

解讀。擬合模型將特徵價格函數估計為：

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

每個係數都是對應特徵的隱含價格估計值。建築面積每平方英尺貢獻 \$575（標準誤最窄、顯著性最強）；景觀為一宗帕西菲卡房地產貢獻約 \$86,000；狀況每提升一級（例如從「良好」到「優秀」）約值 \$50,000。

若干注意事項一望即知：

樣本極小。八個觀測、四個預測變數，只剩三個殘差自由度。標準誤相應地很寬，地塊、景觀與狀況的 p 值儘管在經濟上合乎情理，卻未達到常規的 0.05 門檻。這是小樣本的假象，而不是這些特徵無關緊要的證據。若有 80 個而非 8 個可比案例， p 值將大幅縮小。

整體擬合很強。調整後 $R^2 = 0.93$ ，模型 p 值為 0.012，表明所納入的特徵聯合解釋了樣本中大部分價格變異。聯合擬合強而單項顯著性弱的組合，是多重共線性的教科書式訊號；在這裡它是結構性的：建築面積、地塊與狀況在樣本中共同變動。

一個殘差格外突出。最大殘差為 +\$46,455，對應可比案例 H。其餘七個殘差全部落在 $\pm \$18,000$ 的區間內。可比案例 H 的某種特質沒有被擬合函數捕捉——模型對其價格的系統性低估，幾乎是第二大殘差的三倍。

這個離群殘差不是雜訊。它是一個訊號：可比案例 H 的某項未觀測特徵正以納入變數無法還原的方式對其價格作出貢獻。潛變數條目討論的正是這一問題。

與配對銷售推理的相互驗證。條目 004 的配對銷售分析僅憑景觀與地塊就預測出可比案例 A 與 B 之間 \$127,039 的差異。迴歸給出的隱含價格與之完全相同（因為它們正是那次計算的來源）。迴歸將配對銷售邏輯同時推廣到全部八個可比案例，但在這一樣本量下並未產生任何本質上新的東西；它確認了配對分析，同時報告了配對銷售推理無法給出的標準誤。

參考文獻。

- 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.

交叉引用：異質商品 (*heterogeneous good*)；特徵空間 (*characteristics space*)；特徵價格函數 (*hedonic price function*)；隱含價格 (*implicit price*)；潛變數 (*latent variable*)。

附錄：Prolog 形式化

以下機器可讀的配套檔案 (005-hedonic-regression.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。

```
%% 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).

```

潛變數

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: [10.67330/11n3g168](https://doi.org/10.67330/11n3g168)

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。潛變數 (latent variable) 是一項未被觀測的屬性，它影響觀測到的結果，其取值必須由它對可觀測變數的作用推斷得出，而無法直接測量。

直觀闡釋。前五個條目在處理特徵向量 $\mathbf{z}$ 時，彷彿它可以被完整測量：建築面積、地塊、景觀、狀況等等，全都以資料欄位的形式可得。而在實務中，沒有任何特徵向量是完整的。有些屬性不在資料之中，是因為它們難以測量（室內裝修品質、平面配置的流暢度、微區位的朝向細節），是因為標準資料來源中根本沒有它們（MLS 紀錄不會記載採光品質、隔音效果，也不會記載近期未取得許可的改建），或者是因為它們尚未被認為相關（某些社區層面的特徵，只有在分析揭示之後才被發現能夠預測價格）。

這些未被觀測卻相關的屬性，就是潛變數。它們存在，它們影響價格，但它們不在我們迴歸的資料欄位中。它們的作用不會因為未被測量而消失；這些作用會累積到模型的殘差裡，從已納入變數的視角看它們像是「雜訊」，但從底層資料生成過程的視角看，它們其實是有結構的訊號。

潛變數是理想化的特徵價格框架與雜亂的估價資料現實之間的橋梁。明確承認它們的存在，是誠實地思考模型不確定性、殘差結構以及可觀測特徵解釋力邊界的前提。

估價師在何處遇到它。估價師在三類反覆出現的情境中遇到潛變數。

品質與狀況調整。作為估價類別的「品質」與「狀況」，本身就是把一個多維潛在構念壓縮到單一序數量表上的嘗試。UAD 3.6 中的 C3 狀況評級，意在概括一宗房地產的遞延維修、飾面磨損、機電系統年限與呈現狀態——四到十項底層屬性被濃縮為一個數字。這個單一數字的概括是潛在構念的代理變數；估價師的功力之一，就在於判斷該代理在每個具體案例中對底層真相的捕捉程度。

網格中未獲解釋的殘差。當三個經調整的可比案例給出價值指示，彼此的差距卻超出預期時，這一分歧本身是有資訊的。它很少全部是抽樣雜訊；更常見的情況是，它反映了估價師尚未辨識、尚未調整的某項屬性。追查這種殘餘分歧——「為什麼這個可比案例的指示值比其他高出 \$50,000？」——正是估價師對潛變數推斷的非正式版本。

誰也看不見的近期改建。一宗房地產可能剛做過廚房翻新、屋頂更換，或建成了經許可的附屬住宅單元 (ADU)，而這些尚未反映在 MLS 或稅務估價資料中。從估價師的視角看，這項改建實地可見，卻不在資料之中。從迴歸的視角看，這項改建是潛在的。估價師必須把實地觀察與資料驅動的指示相互調和，二者都應收斂於該潛在屬性的真實貢獻。

為何關乎可辯護性。潛變數帶來了若干可辯護性方面的考量，而條目 005 中的可加線性模型並未直接處理這些問題。

遺漏變數偏誤。如果某個潛變數與一個已觀測變數相關，它的效應就會滲入該已觀測變數的係數。如果近期成交的帕西菲卡 (Pacifica) 住宅往往擁有翻新過的廚房（一項潛在屬性），而房屋成交的合約日期又在資料之中，那麼成交時點（即 `sale_age`，「距成交天數」）的係數就會吸收一部分廚房翻新效應。`sale_age` 的隱含價格貢獻估計值因此被扭曲——它反映的是 `sale_age` 加上未被測量的廚房翻新狀態的一部分。可辯護的估價必須承認，已觀測變數的係數是以模型中缺失的內容為條件的。

殘差的解讀。標準迴歸理論把殘差視為均值為零的高斯雜訊。潛變數視角則把殘差視為雜訊與訊號的混合：一部分是抽樣變異，另一部分是未觀測屬性的累積效應。分辨二者在方法論上具有實質意義。若某殘差與位置、與成交日期、與估價師身分，或與任何其他可觀測量相關，就提示相應的潛變數可以被部分還原。

模型選擇。納入更多已觀測變數是遺漏變數偏誤的顯見對策，但最要緊的那些潛變數，往往恰恰是最難加入模型的（狀況的細顆粒、近期改建的品質、社區微效應）。可辯護的因應方式有時是承認模型永遠無法捕捉它們，從而把未獲解釋的變異數歸入殘差，而不是強行塞進額外的變數。

估價實例演算。條目 005 的迴歸對可比案例 H 產生了一個異常殘差：模型對可比案例 H 價格的低估達 \$46,455，而其餘七個殘差都落在 $\pm \$18,000$ 的區間內。可比案例 H 的已觀測特徵——建築面積 1,600、地塊 7,300、無景觀、狀況「良好」——把它牢牢置於樣本的中間位置；其可觀測輪廓中沒有任何東西預示異常高的價格。然而它的成交價卻比模型所說的高出 \$46,455。

自然的假設是：可比案例 H 具有某項潛在屬性，它對價格有貢獻，卻沒有出現在資料中。在本期資料集的建構中，情況恰恰如此：可比案例 H 經歷了一次近期廚房翻新，為其組合價值增加約 \$60,000。這項翻新不在迴歸的變數清單之內，因此其貢獻累積在殘差中，而不是體現在某個係數上。估計出的殘差（\$46,455）略小於真實的潛在貢獻（\$60,000），因為迴歸會把潛在效應的一小部分分攤到其他係數上，在相關維度上使它們向上偏高。

由此可得三點操作性結論：

把離群殘差當作證據，而不是誤差。對可比案例 H 的殘差，第一個恰當的反應不是把它作為離群值剔除，而是去調查這宗房地產的何種特質解釋了這一低估。在本例中，實地查勘或建築許可紀錄檢索就會揭示那次翻新。模型的失敗具有診斷價值。

認識到其餘係數存在輕微偏誤。景觀、狀況與地塊的係數，透過與可比案例 H 已觀測特徵之間的相關性，吸收了翻新效應的一小部分。在 $n = 8$ 時這一偏誤很小，隨著可比案例與共變數的增加還會進一步縮小，但它並不為零。

殘差整體承載著可還原的結構。這正是作者所提出的殘差約束法（Residual Constraint Approach, RCA）的實質主張（Craytor, 2025）：在住宅估價資料中，殘差在系統上並非獨立同分布的高斯雜訊。它們反映了未觀測屬性的結構化貢獻——狀況的細顆粒、品質的細顆粒、功能效用以及其他潛在構念——而這些貢獻可以透過把殘差本身當作有待建模、而非棄置的潛變數訊號來部分還原。

RCA 對狀況與品質的處理，使這一兩層結構得以落實。估價師的序數評級——即上文討論的粗顆粒代理變數，可標記為狀況-1（狀況粗顆粒）與品質-1（品質粗顆粒）——粗糙到幾乎可以把它們本身也視為潛變數，但它們仍可作為迴歸元納入，以得到近似的隱含價格貢獻。隨後，該協定把這些估計出的貢獻從銷售網格迴歸部分的已測屬性中移出、送回殘差，再由殘差重新分配給細顆粒潛變數狀況-2（狀況細顆粒）與品質-2（品質細顆粒），並對照 MLS 公開備註等補充證據加以精細化。粗顆粒評級提供了初始的價值分配；細顆粒潛變數則承載可還原的結構。最終，細顆粒的價值貢獻取代粗顆粒的價值貢獻。

這裡的單一可比案例演示，是 RCA 直覺最簡單的情形。有了更多資料與更豐富的建模，殘差中的結構可以被劃分為若干可辨識的潛在因子，每個因子都以標準特徵價格設定無法企及的方式對組合的市場價格作出貢獻。潛變數不是需要被控制掉的麻煩；它們是可辯護估價建模的下一個前沿。

參考文獻。

- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *American Journal of Psychology*, 15(2), 201–292.
- Bollen, K. A. (2002). Latent Variables in Psychology and the Social Sciences. *Annual Review of Psychology*, 53, 605–634.
- Craytor, B. (2025). Residual Constraint Approach: Framework & Protocol. Zenodo. doi:10.5281/zenodo.14787917

交叉引用：特徵空間 (*characteristics space*)；特徵價格函數 (*hedonic price function*)；隱含價格 (*implicit price*)；特徵價格迴歸 (*hedonic regression*)。

附錄：Prolog 形式化

以下機器可讀的配套檔案 (006-latent-variable.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。

```
%% 006-latent-variable.pl
%%
%% Prolog companion to Foundations Vol. 1, Issue 1, Article 006
%% Title: Latent Variable
%% Author: Bert Craytor
%% Concept: latent_variable
%% 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(latent_variable, "Latent Variable", measurement_theory,
%      "A latent variable is an unobserved attribute that influences
%      observed outcomes, whose value must be inferred from its
%      effects on observable variables rather than measured directly.",
%      published, 'dictionary-2026.1').
%
% term(residual, ...).
% term(omitted_variable_bias, ...).
% term(proxy_variable, ...).
% term(residual_constraint_approach, ...). % synonym: rca
%
% introduced_by(latent_variable, spearman_1904).
% formalized_by(latent_variable, bollen_2002).
% introduced_by(residual_constraint_approach, craytor_2025).
% depends_on(latent_variable, characteristics_space).
% -----

:- use_terms([ latent_variable, residual, omitted_variable_bias,
               proxy_variable, residual_constraint_approach,
               characteristics_space, hedonic_price_function,
               implicit_price, hedonic_regression ]).
:- use_relations([related/2, depends_on/2,
                  introduced_by/2, formalized_by/2]).

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

article_concept(latent_variable).
article_issue(volume(1), number(1), year(2026)).
article_id('foundations.2026.006').

% =====
```

```

% LOCAL VOCABULARY
% =====

local_term(unobserved).
local_term(noise).
local_term(signal).
local_term(structured_signal).
local_term(uad_3_6).
local_term(c3_condition_rating).
local_term(mls).
local_term(permit_records).
local_term(neighborhood_micro_effect).

% =====
% THE CENTRAL CLAIM
% =====

claim(latent_variables_accumulate_in_residuals,
    "Unobserved-but-relevant attributes do not vanish when omitted from the model;
    ↪ their effects accumulate in the residual, appearing as noise from the model's
    ↪ perspective but as structured signal from the data-generating process's
    ↪ perspective.").

claim(residuals_are_mixture,
    "Standard regression theory treats residuals as iid Gaussian noise; the latent-
    ↪ variable view treats them as a mixture of sampling noise AND signal from
    ↪ unobserved attributes.").

% =====
% THREE CONTEXTS WHERE APPRAISERS ENCOUNTER LATENT VARIABLES
% =====

appraiser_encounter(quality_and_condition_proxies,
    "UAD 3.6 'C3 condition' is a single ordinal scale collapsing a multi-dimensional
    ↪ construct (deferred maintenance, finish wear, mechanical-systems age, presentation
    ↪). The rating is a proxy_variable for a latent construct.").

appraiser_encounter(unexplained_grid_disagreement,
    "When three adjusted comps disagree by more than expected, the disagreement is
    ↪ rarely all sampling noise; usually reflects an attribute not yet identified or
    ↪ adjusted for.").

appraiser_encounter(unobserved_recent_improvement,
    "A recent remodel, roof replacement, or permitted ADU buildout may be observable
    ↪ in person but not in MLS/assessor data — observable to the appraiser but latent
    ↪ to the regression.").

% =====
% DEFENSIBILITY ISSUES
% =====

defensibility_issue(omitted_variable_bias,
    "If a latent variable is correlated with an observed variable, its effect leaks
    ↪ into the observed variable's coefficient. Observed-variable coefficients are
    ↪ conditional on what is missing from the model.").

defensibility_issue(residual_interpretation,
    "Treat residuals as a mixture, not pure noise. A residual that correlates with
    ↪ location, sale date, appraiser identity, or any other observable suggests the

```

```

    ↪ latent variable can be partially recovered." ).

defensibility_issue(model_selection_limits,
    "The latent variables that matter most are often hardest to add (condition fine-
    ↪ grain, quality, micro-locational effects). Sometimes the defensible response is to
    ↪ allocate the unexplained variance to the residual rather than force-fit
    ↪ additional variables." ).

% =====
% WORKED EXAMPLE: COMP H'S ANOMALOUS RESIDUAL
% =====

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

% Comp H's regression residual from entry 005.
comp_h_residual_dollars(46455).

% The latent attribute revealed by investigation.
comp_h_latent_attribute(kitchen_remodel).
comp_h_latent_attribute_value_estimate(60000).

% Why the estimated residual is smaller than the latent attribute's
% true contribution: regression spreads small portions of the latent
% effect across correlated coefficients, biasing them upward.
residual_smaller_than_truth_explanation(
    "Regression spreads small portions of the latent effect across the other
    ↪ coefficients via correlations, biasing them slightly upward in correlated
    ↪ dimensions." ).

% =====
% THREE OPERATIONAL CONSEQUENCES
% =====

operational_consequence(outlier_residual_is_evidence,
    "The first response to Comp H's residual is to investigate, not to drop. An in-
    ↪ person inspection or permit search reveals the remodel; the model's failure has
    ↪ diagnostic value." ).

operational_consequence(other_coefficients_slightly_biased,
    "View, condition, and lot coefficients absorbed small fractions of the remodel
    ↪ effect via correlations with Comp H's observed characteristics. Bias is small at n
    ↪ =8 but not zero." ).

operational_consequence(residuals_carry_recoverable_structure,
    "Residuals in residential appraisal data are systematically not iid Gaussian. They
    ↪ reflect structured contributions of unobserved attributes that can be partially
    ↪ recovered: this is the Residual Constraint Approach intuition." ).

% =====
% CONNECTION TO RESIDUAL CONSTRAINT APPROACH (RCA)
% =====
%
% The single-comp illustration here is the simplest case of the RCA
% intuition. With richer modeling, residual structure is partitioned
% into recognizable latent factors.

framework_reference(residual_constraint_approach, craytor_2025).

rca_intuition(

```

```

    "Treat the residual itself as a latent-variable signal to be modeled rather than
    ↪ discarded. Latent variables are not a nuisance to be controlled away; they are the
    ↪ next frontier of defensible valuation modeling.")).

rca_factor_examples(
    [condition_fine_grain, quality_fine_grain, functional_utility])).

% Two-level condition/quality structure (RCA protocol; mirrors the prose).
rating_levels(condition, condition_coarse_grain, condition_fine_grain).
rating_levels(quality, quality_coarse_grain, quality_fine_grain).

coarse_to_fine_grain_protocol(
    "Regress on the appraiser's coarse-grain C1-C6/Q1-Q6 ratings to obtain approximate
    ↪ value contributions; move those contributions out of the measured attributes back
    ↪ into the residual; reassign them to the fine-grain latent variables and refine
    ↪ against MLS public remarks. In the end, the fine-grain value contribution replaces
    ↪ the coarse-grain value contribution.")).

% =====
% INFERENCE RULE: anomalous-residual detection
% =====
%
% A property's residual is "anomalous" if its magnitude exceeds the
% typical residual band reported in entry 005 (+/- $18,000 for the other
% seven comps). Useful for flagging properties whose hedonic prediction
% suggests a latent attribute warrants investigation.

typical_residual_band(18000).

anomalous_residual(P, R) :-
    comp_h_residual_dollars(R0), % stand-in: this entry references the
                                % residual computed in entry 005; future
                                % integration would pull from there.

    P == comp_h,
    R = R0.

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

cross_reference(characteristics_space).
cross_reference(hedonic_price_function).
cross_reference(implicit_price).
cross_reference(hedonic_regression).

```

近期判決

近期法院判決——美國

近期法院判決 • *The Valuation Engineer*

William Bert Craytor • 第 1 卷，第 1 期 • 2026 年夏季

DOI: 10.67330/fz34f711

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。

按轄區分列的案例摘要

California (加利福尼亞州)

The Retail Property Trust v. Orange County Assessment Appeals Board No. 1. California Court of Appeal (加利福尼亞州上訴法院)，2026 年 4 月判決。

上訴法院處理的問題是：對於 COVID-19 疫情期間政府命令所施加的通行限制，能否依據 California Revenue and Taxation Code section 170 (《加利福尼亞州稅收與稅務法典》第 170 條) 取得財產稅估價減免。作為原告的購物中心主張，疫情期間通行受限構成該第 170 條意義上的對財產的「損害」(damage)，因而有權取得估價減免。

對估價師的意義。該判決意見對第 170 條中「損害」一詞的解釋，將影響今後基於外因造成的財產效用減損而提出的減免主張，無論其起因是公共衛生命令、基礎建設工程，還是其他影響通行的政府行為。凡受託就通行限制對商用不動產價值的影響發表意見的估價師，都應研讀法院關於「效用受限」——區別於實體損傷——是否構成可取得估價減免之損害的說理。

Colorado (科羅拉多州)

Cheroutes v. Rocket Mortgage & Solidifi. U.S. District Court, D. Colorado (美國科羅拉多地區聯邦地區法院)；審理中，2026 年有進展更新。

該案起訴狀主張，2021 年對 Denver (丹佛) 一處雙併住宅所作的估價存在違反 Fair Housing Act (《公平住房法》) 的估價歧視：該不動產被估值約 \$640,000，比上一年度的估價值低出逾 \$220,000。此前一名聯邦法官已駁回被告請求駁回起訴的聲請 (motion to dismiss)，裁定 Rocket Mortgage 與估價管理公司 Solidifi 必須就相關主張進行答辯，而不能以貸款機構和估價管理公司 (AMC) 無須為獨立計費估價師的行為承擔責任為由推卸責任。該案在 2026 年處於證據開示 (discovery) 階段。

對估價師的意義。該案是目前檢驗估價管理公司與貸款機構是否須依據《公平住房法》就計費估價師行為承擔責任的主要活躍載體。若法院最終就實體問題判定該估價管理公司與貸款機構承擔責任，將大幅擴大因估價而面臨公平住房主張的主體範圍，並很可能傳導至估價管理公司與估價師之間的委任契約之中。

Florida (佛羅里達州)

針對向住宅借款人收取的 AMC 費用的佛羅里達州集體訴訟 (class action)。2026 年 1 月 2 日提起。

該起訴狀對估價管理公司向住宅抵押貸款借款人收取的費用提出質疑，主張支付給計費估價師的費用與在交割時向借款人收取的費用之間的加價，超出了聯邦法律確立的限額，且揭露不充分。

該案處於 Dodd-Frank (《陶德—法蘭克法》)「慣常且合理」費用要求 (12 U.S.C. § 1639e) 與 Truth in Lending Act (《貸款真實性法》)、Real Estate Settlement Procedures Act (《不動產交割程序法》) 項下消費者揭露制度的交叉地帶。

對估價師的意義。該訴訟檢驗的是估價管理公司報酬結構；若獲准作為集體訴訟並就實體問題作出裁判，可能重塑全行業的揭露實務。其結果可能透過將「估價管理公司支付額」與「向借款人收取額」之間的價差置於司法審視之下，間接影響經由估價管理公司承接業務的計費估價師的費用形成環境。

Georgia (喬治亞州)

Green Rock 保育地役權 (conservation easement) 系列案。 United States Tax Court (美國聯邦稅務法院)，判決意見見 *Tax Notes* 2026 年 3 月 2 日的評述。

稅務法院在評估喬治亞州 Piedmont (皮埃蒙特) 地區相鄰地塊 (Green Rock Property 與 Green Rock South) 上設定的保育地役權時，逐一檢視了最高最佳使用 (highest and best use) 檢驗的各項要件。申請人一方的專家斷言，在設定地役權之前，該等不動產的最高最佳使用是經營花崗岩採石場，並採用現金流量折現分析估算出設定地役權前的公平市場價值約為 \$13.75 million，而設定地役權後的價值為 \$180,000。

法院對該檢驗的每一項要件都表示保留。在法律許可性方面，法院不認為將土地改劃為允許採石作業的分區在合理機率上可以實現，並特別指出：Green Rock Property 的分區類別變更是在委託新的保育地役權估價後不久取得的——這一先後順序令法院難以採信。在實體可能性方面，法院認為申請人的地質證據不足以證明在 Harman Road 地塊上可以開發出可行的採石場，儘管 Piedmont 地區總體上已知存在可開採的花崗岩。法院進一步認為，即便假定法律許可性與實體可能性均已成立，在未滿足該分析其餘要件的情況下，這也不具有決定意義。

對估價師的意義。該判決意見是有序、逐項適用四要件最高最佳使用檢驗的一個範例。它強化了 *Ranch Springs, LLC v. Commissioner*, 164 T.C. No. 6 (2025) 與 *Beaverdam Creek Holdings, LLC v. Commissioner*, T.C. Memo. 2025-53 所確立的立場：即便擬議用途在周邊區域的總體可行性有據可查，稅務法院也不會接受帶有投機性質的最高最佳使用主張。編製地役權估價的估價師應針對每一項要件建立針對具體地塊的證據，而不應依賴區域性的或類推支持的可行性論證；對於時間上的先後模式——例如與委託估價同期取得的分區變更——尤應格外審慎，因為這類情形容易招致司法上的懷疑。

Maryland (馬里蘭州)

Baltimore (巴爾的摩市) 徵收 Pimlico Race Course (皮姆利科賽馬場)。 土地徵收程序，2025–2026 年進行中。

巴爾的摩市依據徵收權取得 Pimlico Race Course，由此引出關於專用型設施的繼續經營價值與不動產價值之間關係的公正補償問題。由於 Pimlico 的營運業務與其實體設施在功能上不可分割，常規的不動產可比銷售分析用處有限，而收益法估價則高度依賴於賽馬營運將持續進行這一假設。

對估價師的意義。凡受託評估營運業務與不動產在功能上不可分割的單一用途設施——例如賽馬場、影視製片廠、主題樂園、碼頭，以及某些醫療或旅館類資產——的估價師，都值得關注該程序的進展。訴訟程序中形成的方法論，很可能影響今後徵收案件中公正補償計算如何處理繼續經營增值部分。

針對估價師的聯邦《公平住房法》主張

Braxton v. Bank of America. 2026 年 4 月提起。

一名屋主起訴 Bank of America（美國銀行）及一名 Virginia（維吉尼亞州）執業的估價師，主張其行為構成違反 Fair Housing Act（《公平住房法》）的種族歧視。原告聲稱，估價現場有她本人及其承租人在場，且該估價師的舉止表現出種族敵意。該案延續了此前估價歧視訴訟所形成的模式，包括 Maryland 聯邦地區法院的 *Connolly/Mott v. loanDepot* 案，以及加利福尼亞州 Marin County（馬林郡）的 *Austin v. Miller* 案，即由個人原告針對估價師、估價管理公司與貸款機構提起《公平住房法》主張。

對估價師的意義。針對個體估價師的估價歧視案件持續以一定速度提起，加之聯邦層面在差別性影響執法上的收縮，顯示責任暴露的格局已經改變：私人原告與州一級的公平住房執法機構，如今已取代聯邦機構成為估價歧視責任的主要來源。估價師應檢視其文件記錄做法，確保可比案例選取、調整值推導與最終價值調和在事後審視下均能獨立成立；同時應意識到，現場勘查本身的言辭或舉止，往往正是原告主張所依據的事實基礎。

近期法院判決——歐盟

近期法院判決 • *The Valuation Engineer*

William Bert Craytor • 第 1 卷，第 1 期 • 2026 年夏季

DOI: 10.67330/rlynly02

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。

範圍與資料來源

本綜述取材於可取得的公開一手資料——最主要包括歐盟法院（Court of Justice of the European Union）的官方判例資料庫（curia.europa.eu）、歐洲人權法院（European Court of Human Rights）的 HUDOC 資料庫、德國聯邦財政法院（Bundesfinanzhof）公布的裁判（bundesfinanzhof.de）、德國聯邦憲法法院（bverfg.de）、用於法國判例的 Légifrance，以及英國的 Upper Tribunal (Lands Chamber)（上級審裁處土地庭）——並參考了專業與學術出版物中的二手評述。以各國語言作出的裁判均已閱讀原文並以英文加以描述；案件名稱保留其原有拼寫與變音符號。

僅限訂閱的法律檢索資料庫（juris.de、beck-online、Westlaw UK、Lamy、Aranzadi）未予使用。最高法院層級以下的國內法院裁判僅作選擇性收錄，優先選取具有明確先例價值或方法論意義者。本文所載案件引證，在用於實務依據或進一步出版之前，均應對照法院的官方裁判彙編予以核實。

本基線一輯將 Germany（德國）、France（法國）與 United Kingdom（英國）作為主要國家轄區。Italy（義大利）、Spain（西班牙）、Netherlands（荷蘭）與北歐各國在本輯中未單獨綜述，但將視材料情況在其後的季度各輯中補入。Switzerland（瑞士）雖不屬於歐盟，但因瑞士估價準則對歐洲實務總體上具有很強的方法論影響，將在今後各輯中予以綜述。

歐洲監管背景

歐盟層面的兩項監管進展對估價師的執業環境有實質影響，儘管兩者本身都不是法院判決：

EU AI Act 與自動估價模型。 Regulation (EU) 2024/1689 (EU AI Act，歐盟《人工智慧法》) 已於 2024 年生效，適用於高風險人工智慧系統的條款分階段生效，至 2026 年 8 月全部到位。2026 年 5 月 19 日，European Commission（歐盟執行委員會）就高風險人工智慧系統的分類發布了指引草案，供利害關係人諮詢。該法 Annex III（附件三）將與自然人有關的信用狀況評估列入八類推定高風險應用場景之中，執委會的指引草案對這一類別作了詳細展開。由於用於抵押貸款核貸、抵押貸款組合監控與保險核保的自動估價模型（AVM）在功能上與相關貸款的信用狀況評估不可分割，歐盟境內由信用機構或保險機構營運的 AVM，在大多數配置下都將被歸類為高風險人工智慧系統。由此產生的合規後果包括：強制性的風險管理制度、資料治理控制、日誌留存不少於六個月的自動記錄、使營運者能夠理解並推翻系統決定的人為監督機制，以及上市後監測義務。

對估價師的意義。歐洲估價師應當預期，其貸款機構客戶所營運的、由 AVM 支援的工作流程，將面臨大幅提高的文件記錄與可解釋性要求。凡估價師提供模型輸入、校準資料，或對 AVM 輸出進行人為監督複核之處，這些活動本身即構成貸款機構高風險人工智慧合規態勢的組成部分。

Energy Performance of Buildings Directive (重訂版)。 Directive (EU) 2024/1275，即重訂版《建築能源效率指令》(EPBD)，於 2024 年 5 月 8 日在《歐盟官方公報》上公布，並於 2024 年 5 月 28 日生效。會員國的轉換期限為 2026 年 5 月 29 日。除其他條款外，重訂版 EPBD 要求會員國優先翻修能源效率表現最差的百分之十六的非住宅建築存量與能源效率表現最差的百分之四十三的住宅存量；設定最低能源效率標準 (MEPS)；要求自 2028 年起新建公共建築、自 2030 年起所有新建建築均須為零排放建築；並要求自 2030 年起對新建建築開展全生命週期碳評估。

對估價師的意義。重訂版 EPBD 將透過若干管道影響不動產估價：最低能源效率門檻造就了一類新的、由合規驅動的功能性貶值；能源效率證書 (EPC) 評級成為住宅與商用存量建築價格的實質決定因素；為達到合規所需的翻修成本，必須在成本法以及剩餘經濟壽命的扣減中予以反映。估價師還應預見到，受 Mortgage Credit Directive (《住宅抵押貸款指令》) 約束的貸款機構，將日益要求在擔保品估價中對 EPC 評級與 MEPS 相關的資本支出要求作出明確處理。

後續各輯說明

繼本基線一輯之後的季度各輯將聚焦於變化：新的裁判、此處所列待決事項的處理結果（尤其是德國聯邦憲法法院對 1 BvR 472/26 的處理），以及會員國為將重訂版 EPBD 轉換為國內法而制定的實施立法。瑞士法律框架、義大利與西班牙最高法院裁判的處理，以及法國 *Conseil constitutionnel*（憲法委員會）項下的進展，將視材料情況補入。

按轄區分列的案例摘要

Court of Justice of the European Union (歐盟法院)

GR Real (Case C-351/23)。 歐盟法院 2025 年 6 月 24 日判決。

歐盟法院就一件來自斯洛伐克法院的先決裁判請求作出裁判。該案源於一項庭外抵押貸款強制執程序：在質疑契約條款可能不公平的訴訟尚在進行期間，消費者已因拍賣而失去其家庭住宅。法院裁定，Directive 93/13/EEC（第 93/13/EEC 號指令）項下的消費者保護保障要求，國內程序規則必須允許在設有抵押權的住宅經庭外強制執行出售之前，對契約公平性進行有效的司法審查；凡允許出售在該審查完成之前即不可逆地推進的程序，均與該指令不相容。

對估價師的意義。該裁判並非直接的估價案件，但它重塑了抵押貸款擔保品估價被委託與使用的程序環境。在若干設有庭外強制執程序的會員國（尤其是斯洛伐克、捷克、匈牙利與波羅的海三國），其實際效果是放緩強制執行拍賣流程，並延長貸款機構擔保品估價被依賴的期間。在這些轄區為強制執行情形編製擔保品估價的估價師，應當預期估價報告的使用期限可能長於以往，並可能須面對原估價基準日與最終出售之間的市場變動的檢驗。

European Court of Human Rights (歐洲人權法院)

Baku (巴庫) 建築物拆除案 (Azerbaijan, 亞塞拜然)，2025 年 7 月判決。 歐洲人權法院，《第一議定書》第 1 條。

歐洲人權法院就巴庫一棟建築物的拆除與徵收認定存在侵犯和平享有財產權的行為；該拆除是在當局援引地震風險後下令實施的。史特拉斯堡法院認定，國內當局在確定補償時，未能就其駁回申請人所提交的估價報告給出充分理由，並且不正當地僅採納了法院自行委託的專家意見所載金額。法院共判給 €75,000 的財產損害賠償、€3,000 的非財產上損害賠償，以及 €1,055 的法律費用。

對估價師的意義。該裁判強化了《第一議定書》第 1 條的一條判例脈絡，即要求國內法院在徵收程序中對相互競爭的估價證據給予附理由的斟酌，並且在採信法院委託的專家意見而否定申請人的估價報告時，這一否定本身也必須說明理由。對於從事徵收與公正補償業務的歐洲估價師而言，本案凸顯了一點：當估價報告被提交以支持針對國家的主張時，報告在文件上的可辯護性——支撐可比案例選取、調整值推導與最終價值調和的明示說理——本身即是一項基本權利事項。

France (法國)

Cour de cassation (法國最高法院)，Chambre commerciale (商事庭)，2024 年 6 月 19 日，n° 22-24.169。（專業評述持續討論至 2025 年。）

最高法院在繼承稅 *valeur vénale*（市場價值）爭議的語境下裁定：繼承取得的不動產在應稅事件（*de cujus*，即被繼承人死亡）之後發生的出售，不得被繼承人用於向下修正其原先申報的 *valeur vénale*，即便該事件後的成交價格顯著更低、且出售發生在死亡後數月之內亦然。該裁判適用的原則是：就繼承稅而言，*valeur vénale* 以應稅事件之日為準確定，不因該不動產其後的交易而予以更正。

對估價師的意義。該裁判確認了法國繼承稅實務中嚴格的估價基準日紀律，並限制了事件後成交作為價值追溯證據的相關性。編製法國繼承稅估價的估價師，應基於估價基準日當日或之前的交易建立完備的可比銷售證據，且不應將標的不動產本身在事件後的成交作為佐證證據加以呈現。同樣的說理很可能延伸適用於贈與稅（*droits de donation*）與財富稅（*IFI*）情形。

Cour de cassation，3^e Chambre civile (第三民事庭)，2026 年 1 月 8 日，n° 23-22.803。協商性（非司法）專家報告。

最高法院第三民事庭作出裁判，重新界定了協商性（非司法）專家報告在法國民事訴訟中的法律價值。該裁判承接 2025 年 7 月 18 日 *décret*（法令）所推動的更廣泛程序改革，該法令使調解與協商性爭議解決成為民事程序中的預設規則。法院澄清：由單方當事人依契約委託編製的協商性專家報告，仍可作為證據被採納並被賦予證明力，但須遵循對審原則（*principe du contradictoire*），並須經法院對專家獨立性與方法論的評估。

對估價師的意義。該裁判強化了在法院指定司法專家（*experts judiciaires*）框架之外編製的 *expertises immobilières*（不動產鑑定報告）的證據地位，前提是報告遵循本行業的方法論標準且專家的獨立性有據可查。對於承辦繼承、離婚或商事爭議中爭議性估價的法國估價師而言，該裁判為客戶委託的報告在其後訴訟中被援引與依賴打下了更堅實的基礎。相應地，格外注意記錄獨立性（不存在從屬關係、經濟利益或與客戶的親屬關係）並適用專業標準所規定的方法，也就變得更加重要。

Germany (德國)

Bundesfinanzhof (德國聯邦財政法院)，2025 年 11 月 12 日判決 (II R 25/24、II R 31/24、II R 3/25)。*Grundsteuer Bundesmodell* 三部曲——改革後財產稅估價制度的合憲性。德國聯邦財政法院（*Bundesfinanzhof*，BFH）第二審判庭在三件平行的法律審上訴程序中裁定：自 2025 年 1 月 1 日起施行的改革後不動產稅（*Grundsteuer*）中，德國十一個邦依所謂 *Bundesmodell*（聯邦模式）對住宅不動產估價所適用的收益資本化估價方法（*Ertragswertverfahren*）的相關規定，符合憲法。原告為 North Rhine-Westphalia（北萊茵—西發利亞邦）、Saxony（薩克森邦）與 Berlin（柏林）的公寓所有權人，他們以該套規整合化的估價規則違反一般平等原則（《基本法》第 3 條第 1 款）並產生系統性失準的價值為由，對其稅額認定提出質疑。

聯邦財政法院認為，立法者所建立的估價制度在概念上以市場價值為取向（*konzeptionell einer Verkehrswertorientierung folgt*），其設計意圖是在所有被估價對象上平均而言逼近「客觀現實的不

動產價值」，並處於公平市場價值 (*gemeiner Wert*) 的一個區間之內。法院認定，立法者在權衡該估價方案的目標與此類方案必然伴隨的不平等之間，並未超出其被允許的裁量餘地；並且，為避免透過自動化定期更新而再次形成「估價積壓」而賦予的重大權重，在憲法上是可以允許的。聯邦財政法院還強調，納稅人仍保有由該院此前 2024 年 5 月 27 日的 *Beschlüsse* (裁定) (II B 78/23 與 II B 79/23) 所確立的選擇權，即依據 § 198 BewG (《估價法》第 198 條) 透過合格估價證明較低的公平市場價值 (*gemeiner Wert*)，但受此限制：該較低價值須由公開委任的不動產估價師運用公認估價方法出具的估價報告予以證實。

由 *Bund der Steuerzahler Deutschland* (德國納稅人協會) 與 *Haus & Grund Deutschland* (德國不動產業主協會) 支持的一項憲法訴願 (*Verfassungsbeschwerde*)，現以案號 1 BvR 472/26 待聯邦憲法法院 (*Bundesverfassungsgericht*) 審理。

對估價師的意義。這一三部曲是本綜述期間最重要的德國不動產估價判決。對估價師而言，主要的實務後果是對 § 198 BewG「較低價值」估價 (*Verkehrswertgutachten zum Nachweis des niedrigeren gemeinen Werts*) 的需求上升——即在依 *Bundesmodell* 確定的 *Grundsteuerwert* (不動產稅價值) 看似超出標的不動產實際市場價值的情形下。聯邦財政法院對這一救濟途徑的明確認可，加之其對 *Bundesmodell* 的總體尊重，意味著估價報告如今已成為納稅人個體從系統性高估中取得救濟的實際運作機制。估價師應預期此類報告的委託將持續增加，並應確保符合 ImmoWertV 2021 (《不動產估價條例》2021) 的方法論與 § 198 BewG 的要求。待決於聯邦憲法法院的憲法訴願 (1 BvR 472/26) 對該框架是否會有進一步變動具有實質意義，值得持續關注，預計將成為下一季度一輯的主要議題之一。

Bundesfinanzhof，2025 年 10 月 7 日判決，IX R 26/24。土地與建築物之間的購價分攤 (*Kaufpreisaufteilung*)。在一件涉及住宅出租不動產折舊 (*Absetzung für Abnutzung*, AfA) 計提基數的所得稅案件中，聯邦財政法院第九審判庭處理了如何在不可折舊的土地部分 (*Grund und Boden*) 與可折舊的建築物部分 (*Gebäude*) 之間分攤契約購價的方法問題。原告曾以 *Bodenrichtwerte* (基準地價) 所依據的資料基礎就標的不動產的具體區位並無可靠記載為由，對下級法院對其的援用提出質疑。

對估價師的意義。該裁判延續了聯邦財政法院的一條判例脈絡 (尤其是上文所引 2025 年 11 月 12 日的 BFH II R 25/24，以及此前 2022 年 9 月 20 日的 BFH IX R 12/21)，確認 ImmoWertV 是購價分攤所公認的估價原則來源。本案還表明了 *Bodenrichtwerte* 所依據的資料基礎的重要性：凡 *Gutachterausschüsse* (專家估價委員會) 就具體區位並無足夠可比銷售資料者，*Bodenrichtwert* 即可能因資料品質問題而受到質疑。對於編製購價分攤或折舊基數意見的估價師而言，這意味著對 *Bodenrichtwert* 資料基礎的記錄——包括其底層的 *Kaufpreissammlung* (成交價格彙編)——成為可辯護分析的組成部分。

United Kingdom (英國)

Mehrunnisa Alikhan Ahmed Khan v Wolverhampton City Council。Upper Tribunal (Lands Chamber) (上級審裁處土地庭)，2025 年。

上級審裁處土地庭確認，評估強制收購 (compulsory purchase) 補償的正確估價基準日，是該不動產依據一般歸屬聲明 (general vesting declaration, GVD) 歸屬於收購主管機關之日。審裁處還處理了同期估價與追溯估價之間的關係：在本案中，地方主管機關在 GVD 之後隨即作出的初次估價為 £150,000，而 2025 年在審裁程序中編製的第二份估價——以業主發出通知後三個月的假定歸屬日為基準——則為 £120,000。儘管審裁處確認了法律上的估價基準日，它仍偏向採用在時間上更接近 GVD 作出之日所計算的數額，從而使請求人取得較高的支付額。

對估價師的意義。該裁判對從事強制收購業務的英國估價師具有方法論上的啟示意義。雖然法律

上的估價基準日是歸屬之日，但審裁處在同一法律估價基準日之下，偏向同期估價而非數年之後編製的追溯估價，這顯示同期市場證據即便在歸屬之日形式上不可採，實際上仍具有分量。估價師應確保在 GVD 前後編製的估價報告有充分的支撐證據作為記錄，以充當主要記錄，並預見到其後的審裁程序可能優先採信同期報告而非事後的追溯重構。

近期法院判決——拉丁美洲

近期法院判決 • *The Valuation Engineer*

William Bert Craytor • 第 1 卷，第 1 期 • 2026 年夏季

DOI: 10.67330/ywghzg47

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。

範圍與資料來源

本綜述取材於可取得的公開一手資料——最主要包括美洲人權法院（Inter-American Court of Human Rights）判例庫（corteidh.or.cr）、巴西聯邦最高法院（STF）與高等法院（STJ）入口網站（portal.stf.jus.br、stj.jus.br）、墨西哥 *Semanario Judicial de la Federación*（聯邦司法週刊，sjfsemanal.scjn.gob.mx）、阿根廷 *Corte Suprema de Justicia de la Nación*（國家最高法院，csjn.gov.ar）、智利 *Poder Judicial*（司法機關）入口網站（poderjudicial.cl）與 *Diario Constitucional*，以及哥倫比亞 *Corte Constitucional*（憲法法院，corteconstitucional.gov.co）——並參考了專業與學術出版物中的二手評述。以西班牙語與葡萄牙語作出的裁判均已閱讀原文並以英文加以描述；案件名稱保留其原有拼寫與變音符號。

僅限訂閱的法律檢索資料庫（vLex、La Ley、Thomson Reuters Proview Latin America）未予使用。最高法院層級以下的國內法院裁判僅作選擇性收錄，優先選取具有明確先例價值或方法論意義者。本文所載案件引證，在用於實務依據或進一步出版之前，均應對照法院的官方裁判彙編予以核實。

本基線一輯將 Argentina（阿根廷）、Brazil（巴西）、Chile（智利）、Colombia（哥倫比亞）與 Mexico（墨西哥）作為主要國家轄區。Peru（秘魯）、Venezuela（委內瑞拉）、Ecuador（厄瓜多）、Bolivia（玻利維亞）、Paraguay（巴拉圭）、Uruguay（烏拉圭）以及中美洲與加勒比各轄區在本輯中未單獨綜述，但將視材料情況在其後的季度各輯中補入。委內瑞拉的財產返還業務與古巴制裁相關的估價業務，因其方法論上的獨特性及其與跨境實務的相關性，將在今後一輯中予以綜述。

區域背景

與歐盟綜述不同，本拉丁美洲綜述沒有與 EU AI Act（歐盟《人工智慧法》）和重訂版 Energy Performance of Buildings Directive（《建築能源效率指令》）所代表的超國家監管層相對應的直接類比。拉丁美洲的不動產估價監管絕大部分屬於國家層面；主要的超國家影響來自美洲人權法院依據 American Convention（《美洲人權公約》）第 21 條（財產權）所形成的基本權利判例法，以及本區域多數轄區為財務報導目的採用國際財務報導準則（特別是 IFRS 13《公允價值衡量》）。儘管如此，仍有兩條方法論主線貫穿本區域判例，值得在此提及：

原住民共有財產的估價。拉丁美洲估價實務中最具特色的方法論挑戰，是在返還、徵收與補償情形下對原住民與非洲裔族群共有財產的估價。美洲人權法院的判例法一貫認為，原住民共有財產與個人私有財產在憲法上受到同等保護；但此類財產的估價方法——通常並無可比的公平交易案例，財產由共同體共有且不可轉讓，且該領土的文化與精神意義構成其價值的內在組成部分——仍不成熟。截至 2025 年底，美洲人權法院作出的有利於原住民族的二十四項判決仍未執行或僅部分執行，其中若干涉及尚待解決的補償計算，而這些計算取決於懸而未決的估價方法。

徵收補償中的通貨膨脹、貨幣與指數化。本區域若干轄區，尤其是阿根廷，要求估價方法與賠償計算必須應對高通膨、以外幣計價的 *tasaciones*（估價）以及複雜的指數化制度。阿根廷最高法院認為，原以外幣表示的徵收賠償不得按人為的平價匯率折算為本國貨幣，並且估價必須反映不動產在賠償之時的真實經濟價值。類似關切亦見於委內瑞拉（與財產返還及制裁相關的估價有關），並在巴西通膨高漲時期不時出現。對於跨區域各轄區執業的估價師而言，這些通膨與貨幣方面的考量是估價方法的內在組成部分，其程度是歐洲與美國所沒有的。

後續各輯說明

繼本基線一輯之後的季度各輯將聚焦於變化：美洲人權法院的新裁判、此處所列待決事項的處理結果、哥倫比亞 Jurisdicción Agraria y Rural（農業與農村管轄法院系統）邁向 2027 年投入運作的實施進展，以及各國在 IFRS 13 公允價值衡量實務方面的動向。秘魯、委內瑞拉、厄瓜多與中美洲各轄區將視材料情況補入。將特別關注原住民共有財產估價的方法論發展——在該領域，美洲人權法院的實體裁判已超前於估價師與補償主管機關所能取得的實用估價方法。

按轄區分列的案例摘要

Inter-American Court of Human Rights（美洲人權法院）

Maya Q'eqchi' Agua Caliente vs. Guatemala 與 *Comunidad Garífuna de San Juan y sus Miembros vs. Honduras*。美洲人權法院 2024 年判決，區域判例評述持續至 2025 年。

在兩項同期判決中，美洲人權法院就瓜地馬拉與宏都拉斯在劃界、確權以及防止原住民與非洲裔族群祖傳領土被侵奪方面的義務作出裁判。*Maya Q'eqchi' Agua Caliente* 社區歷經四十五年的行政與司法行動，方才取得其領土的確定性所有權證書；宏都拉斯的 *San Juan* 加里富納（*Garífuna*）社區則面臨祖傳沿海土地持續被侵奪的局面。兩項判決均涉及國家依《美洲人權公約》第 21 條（財產權）承認、劃界並保護原住民共有財產的義務。區域評論者注意到，美洲人權法院在這些裁判中的返還框架，較之其早先判決略趨收緊，尤其體現在其對「善意第三人」（*terceros inocentes*）——即在不知悉原住民主張的情況下取得重疊權利者——的處理上。

對估價師的意義。這些裁判重申：原住民與非洲裔族群的共有財產在整個美洲人權體系內受到憲法保護；在返還或補償情形下對此類財產的估價，若不輔以方法論上的補充，便不能依賴常規的公平交易可比銷售（此類案例通常並不存在）。在原住民返還情形下受託的估價師應當預期，估價必須整合三個不同的組成部分：本區域內物理特徵可比的土地的基礎市場價值；可歸因於原住民使用的機會成本與改良；以及該領土的非經濟性文化與精神意義——美洲人權法院判例法將其視為財產權中可予補償的一個面向。第三部分的方法仍不成熟，是美洲人權法院說理中持續探討的主題。

Mapuche / Consejo de Todas las Tierras vs. Chile。美洲人權法院 2024 年判決（見 2025 年判例評述）。

美洲人權法院就智利國家針對馬普切（*Mapuche*）人 135 名成員的人權侵犯行為作出譴責，該案涉及 *Consejo de Todas las Tierras*（1989–1992 年）的領土索還行動，以及將反恐立法適用於原住民領土主張的做法。該判決擴展了 *Norín Catrínán and others vs. Chile*（2014 年）所確立的原則，並明確承認原住民族在領土主張方面的自決權。

對估價師的意義。儘管本案的主要裁判要旨涉及刑事程序保障而非直接的估價方法，但對原住民在領土方面自決權的承認，形成了一層程序性疊加，影響智利原住民領土返還及相關情形下的估

價工作。在此類情形下受託的估價師應確保其方法納入適當的協商程序，並承認該共同體對其領土價值如何被表述所享有的自決權。

Argentina (阿根廷)

阿根廷關於 **Tribunal de Tasaciones de la Nación** (國家估價法庭) 的原則。經 2025 年判例整合確立。

阿根廷國家最高法院 (CSJN) 與下級聯邦法庭持續發展關於 Tribunal de Tasaciones de la Nación (TTN, 國家估價法庭) 在徵收與 *expropiación inversa* (逆向徵收) 事項中所作估價的原則。既定的 CSJN 原則 (Fallos 313:82, 在 2024–2025 年間反覆適用) 認為，基於該法庭成員的技術能力、其估價所依據的證據基礎以及實務的統一性，TTN 的估價具有推定的分量；欲背離 TTN 的估價，須證明存在明顯錯誤或遺漏。2024–2025 年間適用該原則的聯邦法庭——包括 Juzgado Federal de Oberá (奧韋拉聯邦法院) 在 *Sosa v. Entidad Binacional Yacyretá* (2024 年 6 月，因 Yacyretá 水力發電廠水庫引發的 *expropiación inversa*) 一案中——繼續以 TTN 估價作為賠償的實際運作基準。

對估價師的意義。對阿根廷估價師而言，實際後果是：為對抗 TTN 估價而提交的私人估價報告面臨沉重的舉證責任。要推翻或變更 TTN 的估價，私人報告必須證明 TTN 工作中存在方法上的錯誤或遺漏，而不能僅僅提出另一套估價結論。編製此類報告的估價師，應將方法論批評聚焦於 TTN 的可比案例選取、調整值推導以及對當地市場狀況的處理，而不應僅僅著眼於價值差額的大小。

以外幣計價的 **tasaciones** 與賠償平價。CSJN 關於原以外幣表示的徵收賠償不得按人為的一比一平價折算為阿根廷披索的原則，至 2025 年仍然有效。該原則保護被徵收人免受因通膨或貨幣政策而無法反映不動產在賠償之時真實經濟價值的賠償。這一原則在 2024–2025 年間已就長期進行的徵收程序被多次重申。

對估價師的意義。阿根廷估價師在為跨越顯著通膨或貨幣政策變動期間的徵收事項編製估價時，應明確記錄估價所依據的貨幣基準，並應考慮在估價基準日與預期支付日之間採用適當的指數化方法。未能處理通膨與貨幣風險的報告，不大可能說服適用 CSJN 賠償原則的法庭。

Brazil (巴西)

Supremo Tribunal Federal, RE 922.144 (Tema 865, 一般影響力主題)。巴西聯邦最高法院 2024 年判決，2025 年持續適用。

STF 裁定，*desapropriação* (徵收) 程序中的賠償與《聯邦憲法》第 100 條所確立的聯邦 *precatório* (司法債務令) 制度在憲法上相容。具體而言，公共主管機關為 *imissão provisória na posse* (臨時占有) 目的最初提存的金額，與最終由司法鑑定確定的金額之間的差額，必須透過 *precatório* 制度支付，而非作為即時的補充提存。該裁判在《憲法》第 5 條第 24 項 *justa, prévia e em dinheiro* (公正、事先、以貨幣支付) 的補償命令，與 *precatório* 制度的有序支付要求之間取得平衡。

對估價師的意義。該裁判對受徵收主管機關委託的估價師所編製的初始提存估價具有直接的實務後果。由於初始提存中的任何低估都將在其後透過 *precatório* 制度支付 (這可能涉及大幅延遲)，原告在結構上就有動機主張初始估價被低估。編製初始提存估價的估價師應格外審慎地記錄其方法，以便在其後的司法鑑定面前為該估價辯護，並預見到初始提存與最終司法認定之間的差額，將依 STF 判例法與 Decreto-Lei 3.365/1941 (第 3.365/1941 號法令) 框架計收補償性利息 (*juros compensatórios*)。

Superior Tribunal de Justiça, REsp 2.129.162 (重複性主題)。 巴西高等法院 2025 年 5 月判決。

STJ 就徵收主管機關撤回 *desapropriação* 之訴時的律師費問題，解決了一項重複性爭議。法院確立：撤回起訴案件中的費用應遵循 Decreto-Lei 3.365/1941 第 27 條第 1 款的框架，而非 *Código de Processo Civil* (《民事訴訟法典》) 的一般程序規則。該案源於 Cemig (Companhia Energética de Minas Gerais, 米納斯吉拉斯電力公司) 為一條配電線路尋求行政地役權而提起的訴訟。

對估價師的意義。該裁判屬程序性而非方法論性質，但對行政地役權與 *desapropriação* 事項中估價師受託關係的構造有所影響。在此類程序中被聘為專家證人 (*peritos judiciais*) 的估價師，應確保其費用安排的記錄援引適用的法令框架，而非一般程序規則。

Superior Tribunal de Justiça 關於 *juros compensatórios* 的原則。 2025 年持續適用。

STJ 維持其原則：*desapropriação* 事項中的補償性利息 (*juros compensatórios*) 僅自法院就被徵收財產的權屬作出裁判之後起算。該利息適用於初始提存與最終賠償價值之間的差額，按歷史上確立的年利率百分之十二計算 (STF Súmula 618, 第 618 號判例要旨)，其作用是補償被徵收人在最終賠償支付之前喪失占有的損失。

對估價師的意義。巴西估價師應當預期，他們在司法鑑定中所認定的估價差額——即公共主管機關的初始提存與評估市場價值之間的差額——將成為補償性利息計算的基數。因此，該差額的經濟意義大於其名目差額所顯示的程度，在長期程序中尤其如此。

Chile (智利)

Corte Suprema, Rol N° 31.423-2025。 智利最高法院 2025 年 9 月判決。

智利最高法院駁回了一家不動產開發商 (*inmobiliaria*) 針對國庫 (*Fisco de Chile*) 提起的撤銷之訴 (*casación*) 上訴；該案屬依 Decree-Law N° 2.186 (第 2.186 號法令) 第 12 條進行的特別徵收程序。原告曾主張，專家委員會 (*Comisión de Peritos*) 所定的估價未能反映該不動產的最高最佳使用 (*mayor y mejor uso*) 以及既有建築物與配套工程的價值，並另以既有租賃契約本可產生未來租金收入為由，請求就可得利益損失 (*lucro cesante*) 取得賠償。法院認為，原告的主張實質上是對下級法院證據權衡的不同意見，而非對規範估價的法律規範的真正違反；而下級法院的說理——其依據專家委員會的證據並結合書證與證人證言——已有充分根據。

對估價師的意義。該裁判說明了智利最高法院原則對徵收程序中專家委員會估價所給予的高度尊重，以及僅以方法論分歧為由對估價結論進行撤銷審查的有限空間。在智利徵收事項中編製對立估價的估價師，應著眼於證明委員會工作中可以指明的方法錯誤，而非提出另一套估價說理；同時應意識到，基於既有租賃安排的 *lucro cesante* 主張面臨沉重的舉證要求。

Corte Suprema (Coquimbo 徵收案)。 智利最高法院 2025 年 7 月 17 日判決。

智利最高法院維持了上訴法院提高 Coquimbo (科金博) 徵收案賠償額的裁判：賠償額由國家主張的 UF 2,349.07 (依據 *Conservador de Bienes Raíces*, 不動產登記處的登記成交案例) 提高至 UF 2,969.75 (依據一名私人估價師的報告)。法院強調，上訴法院援用該私人估價師的報告，並結合證人證言以及另一份關於鄰近不動產的估價報告一併考量，屬於 *sana crítica* (合理心證) 的容許行使，而撤銷審查並不延及對此種證據權衡的再行評斷。

對估價師的意義。Coquimbo 一案是 Rol N° 31.423-2025 的對照：既然委員會的估價在受開發商質疑時能夠勝出，那麼私人估價師較高的估價在有充分的方法論說理與佐證證據支撐時，同樣可以勝出。對智利估價師而言，實務上的教訓是：私人估價的方法與文件支撐具有決定意義；僅憑登記成交案例的證據，不足以推翻一份運用了額外可比案例與情境證據、論證充分的私人估價。

Tribunal Constitucional，2025 年 10 月。《稅法典》項下的 *bienes raíces*（不動產）估價。

智利憲法法院駁回了針對 *Código Tributario*（《稅法典》）中規範不動產（*bienes raíces*）為 *impuesto territorial*（土地稅）目的進行估價的條款的合憲性挑戰。持不同意見的大法官認為，被質疑的規範未為受規範者提供充分的機制以對估價提出異議，損害了憲法上關於理性且公正程序的保障。

對估價師的意義。該裁判保留了智利不動產稅現行的行政估價框架，但不同意見的說理指出了現行框架中的一處脆弱之處，今後的憲法或行政法挑戰可能會重新提出。在與不動產稅爭議有關的情形下編製估價的智利估價師，應關注裁判之後的政策環境，因為為改進納稅人異議機制而進行行政改革看來是可能的。

Colombia（哥倫比亞）

Corte Constitucional，對 *Jurisdicción Agraria y Rural* 的核准。哥倫比亞憲法法院 2025 年裁判。

哥倫比亞憲法法院對創設新的 *Jurisdicción Agraria y Rural*（農業與農村管轄法院系統）的立法通過了合憲性審查。該院核准了界定改革對象、其納入司法機關的方式、專門法庭與法院的設立以及預算方面的關鍵條文。該新管轄體系定於 2027 年開始運作，旨在解決土地返還爭議、農業財產爭議，以及因執行 2016 年和平協議而產生的事項。截至核准之日，涉及六萬一千餘公頃的農業爭議仍待現行行政與司法系統處理。

對估價師的意義。該新管轄體系將成為哥倫比亞土地返還程序與農村財產事項所生估價爭議的主要審理場所。在返還情形下執業的哥倫比亞估價師（包括為 *Unidad de Restitución de Tierras*（土地返還局）與 *Instituto Geográfico Agustín Codazzi*（阿古斯丁·科達西地理研究所）工作者）應當預期，新法庭將發展出自身的估價判例，而方法論標準可能自現行以行政為主導的框架演進。從業者應關注 2027 年的運作啟動，以及這些法庭初期案件量的發展。

現行以 **IGAC** 為基礎的土地返還估價框架。2025 年持續適用。

哥倫比亞的土地返還實務繼續以 *Instituto Geográfico Agustín Codazzi*（IGAC）編製的 *avalúo comercial*（商業估價）作為返還與置換情形下補償的主要依據，並在需要時以 *Lonjas de Propiedad Raíz*（不動產同業公會）編製的 *avalúos* 作為補充。該框架體現於憲法法院 2018 年第 SU-034 號裁判（關於履行的程序要求）以及其後的裁判，包括 2024 年第 T-120 號裁判（關於在以實物補償情形下 IGAC 估價與個別受益人偏好之間的關係）。

對估價師的意義。對於從事哥倫比亞土地返還業務的估價師而言，該制度框架把 IGAC 估價置於補償計算的實際運作核心，而 *Lonja de Propiedad Raíz* 的 *avalúos* 則充當補充性或爭議解決機制。即將設立的 *Jurisdicción Agraria y Rural* 可能改變這一制度架構，而 IGAC 估價與 *Lonja* 估價之間的相對分量，正是新法庭可能著手發展的問題之一。

Mexico（墨西哥）

Suprema Corte de Justicia de la Nación，關於 *avalúo* 與成交價差額的 **ISR** 納稅義務。墨西哥最高法院，2026 年初公布的裁判。

墨西哥最高法院（SCJN）全院宣告 *Ley del Impuesto sobre la Renta*（ISR，《所得稅法》）第 125 條、第 130 條第 4 項與第 132 條合憲。這些條款規定：當不動產取得的 *avalúo*（估價）超出契約買賣價格百分之十以上時，其差額被視為買方的應累計所得，因而須繳納 **ISR**。法院認為，該規則並不違反稅收的比例原則與公平原則，理由是以實質低於市場價值的價格取得財產，構成買方財產的客觀增加，因而恰當地構成應稅的經濟利益。

對估價師的意義。該裁判對墨西哥不動產交易以及為其提供支援的估價實務具有重大的實務後果。由 *corredores públicos*（公共經紀人）或其他獲授權專業人士出具的 *avalúos*，在其超出契約價格百分之十以上時，如今對買方產生直接的財稅後果。估價師應當預期，交易類 *avalúos* 的方法將受到更嚴格的審視，且對以特別嚴謹方式記錄可比銷售依據的 *avalúos* 的需求將會上升。該裁判還隱含地提高了在各種估價方法之間作出方法論選擇的分量，因為這一選擇如今可能影響買方的納稅義務。

Suprema Corte de Justicia de la Nación, Amparo en revisión 391/2024。 墨西哥最高法院第二庭 2025 年 6 月 25 日判決。

墨西哥最高法院第二庭審結了一家礦業公司提出的合憲性挑戰，其針對的是一項影響採礦區塊及相關附屬權利（臨時占用、地役權、用水特許優先權與租約續期條款）徵收法律框架的法令。法院認為，該法令並不違反不溯及既往原則，因為受影響的條款所涉及的只是對未來權利的期待，而非既得權利。

對估價師的意義。該裁判對墨西哥礦業權與礦業不動產估價具有重要意義。法院將採礦區塊延期權及相關條款定性為「單純期待」而非既得權利，影響採礦特許權的估價以及對其估價中監管風險的評估。在墨西哥編製礦業不動產或相關權利估價的估價師，應在估價中明確處理監管風險這一面向，並應避免把或有的、依賴於延期的價值當作已完全既得的價值來對待。