

Valuation Engineer Journal

Volume 1

1 · 1 · 2026



(RCA):
earthUI, glmnetUI & mgcvUI:

(F0–F6) . . .
. . .
. . .

Valuation Engineer Journal

1 · 1

2026

ISSN 3143-5033

William Bert Craytor, SRA

	2
	3
.....	4
	6
.....	7
	9
.....	10
(RCA):	45
earthUI: earth (MARS)	109
glmnetUI: glmnet ()	159
mgcvUI: mgcv ()	200
	245
(F0)	246
(F1)	252
(F2)	260
(F3)	268
(F4)	275
(F5)	282
(F6)	291
	298
– (R1)	299
– (R2)	303
– (R3)	310

Valuation Engineer Journal , : , , . , .

РУБРИКА: РЕДАКЦИОННАЯ СТАТЬЯ

Добро пожаловать в первый номер

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Уильям Берт Крейтор, SRA — главный редактор

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

DOI: 10.67330/0tymfa32

© 2026 Уильям Берт Крейтор. Published by Pacific Vista Net.

Эта статья открытого доступа распространяется на условиях лицензии Creative Commons Attribution 4.0.

Аннотация. Обращение редакции к читателям 1-го тома, 1-го номера: о том, зачем создан этот журнал, какой вклад вносят пять статей и три рубрики, а также приглашение к участию.

Ключевые слова: редакционная статья, инженерия оценки, первый номер

Добро пожаловать в первый номер журнала «*The Valuation Engineer Journal*».

Этот журнал появился в связи с существующим пробелом. В теории оценки уже давно определяется оценка как систематизированное измерение поведения рынка, однако повседневная практика по-прежнему в значительной степени опирается на ручной подбор сопоставимых аналогов и корректировки, основанные исключительно на субъективном суждении. Между тем, статистические инструменты, способные измерить, сколько рынки фактически платят за характеристики объектов недвижимости — гедоническая регрессия, современные сплайновые модели, машинное обучение с перекрёстной проверкой — развивались на протяжении десятилетий, но практически не нашли применения в оценке. «Инженер по оценке» — так мы называем специалиста, который устраняет этот пробел: он сочетает в себе компетенцию лицензированного оценщика с навыками работы с данными, статистикой и воспроизводимыми протоколами.

Пять статей данного номера закладывают рабочую основу. В вступительной статье рассматривается сам разрыв между теорией и практикой и доказывается, что собственные стандарты профессии уже требуют более широкого спектра измерений, чем позволяют повседневные методы. Во-вторых — и это является центральным элементом всего, что ставит перед собой цель данный журнал, — в полном объеме представлен «Метод

остаточных ограничений» : протокол, который позволяет регрессионному анализу измерить то, что поддается измерению, систематизировать то, что не поддается измерению, в виде явного разбивку по остаткам, а также ограничивает итоговую величину, чтобы суждение могло уточнять, но никогда не придумывать значение. Далее следуют три статьи, посвящённые программному обеспечению, в которых описываются earthUI и сопутствующие интерфейсы для регрессии с штрафными терминами и обобщённых аддитивных моделей — открытые инструменты, на которых основан этот журнал, — которые незадолго до публикации были объединены в одно приложение, как описано в примечаниях издателя к каждой статье.

Статьи сопровождаются тремя разделами. *Раздел «Основы»* открывается семью краткими статьями (F0–F6), в которых прослеживается концептуальная цепочка от неоднородных благ через пространство характеристик, гедоническую функцию цены, неявные цены и гедоническую регрессию до латентных переменных, содержащихся в остатках. Каждая статья указывает на свою научную происхождение, имеет собственный DOI и заканчивается машиночитаемой формализацией на языке Prolog — это начало базы знаний, которую мы намерены пополнять, номер за номером, превращая её в точный общий словарь для данной области. В рубрике *«Недавние судебные решения»* рассматриваются постановления судов Соединённых Штатов, Европейского союза и Латинской Америки, имеющие отношение к практике оценки, поскольку правовая среда является частью инженерной задачи. *«Письма редактору»* открывают столбец переписки: методы, предназначенные для использования, тестируются способами, которые их авторы никогда не предполагали, и именно здесь находят своё место эти выводы — и вызываемые ими разногласия.

Каждый материал в этом выпуске публикуется в открытом доступе по лицензии CC BY 4.0, как в формате PDF, так и в формате HTML, с идентификатором Crossref DOI, а также хранится отдельно на Zenodo. Мы приглашаем к публикации материалы — статьи, записи в разделе «Foundations», переписку и возражения — как от оценщиков, так и от экономистов и инженеров.

Уильям Берт Крейтор, *главный редактор журнала «*
» (SRA)

РУБРИКА: ПИСЬМА РЕДАКТОРУ

Приглашение к переписке

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Уильям Берт Крейтор, SRA — главный редактор

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

DOI: 10.67330/r55mw896

© 2026 Уильям Берт Крейтор. Published by Pacific Vista Net.

Эта статья открытого доступа распространяется на условиях лицензии Creative Commons Attribution 4.0.

Аннотация. Редактор открывает столбец «Письма»: что этот журнал надеется получить от своих читателей и как присылать материалы.

Ключевые слова: письма, переписка, комментарии коллег

Журнал, который только говорит, — это журнал лишь наполовину. Этот столбец создан для остальной половины.

Журнал «Valuation Engineer Journal» публикует методы, которые предназначены для применения — и методы, которые применяются, проходят проверку способами, которые их авторы даже не могли предвидеть. Если вы применили подход «Остаточное ограничение» и обнаружили, что какой-то этап не работает на вашем рынке, мы хотим получить письмо. Если в разделе «Основы» термин определен более узко, чем того требует ваша практика, мы хотим получить письмо. Если судебное решение, которое мы обобщили, интерпретируется иначе практикующим специалистом в данной юрисдикции, мы тоже хотим получить это письмо. Несогласие, воспроизведение, уточнение и простое исправление — всё это входит в сферу нашего внимания; так же как и краткие отчёты из практики — метод, который сработал, решённая проблема с данными, неудача, о которой стоит предупредить других.

Правил не так много. Письма должны быть подписаны, лаконичны (оптимальный объем — одна страница) и конкретны — укажите статью, раздел или статью, к которой вы обращаетесь, и приведите аргументы так же, как вы это делали бы в рабочем файле. Мы можем редактировать тексты с целью сокращения объема и повышения ясности, всегда с согласия автора. Письма публикуются в открытом доступе по лицензии CC BY 4.0, как и всё остальное на этих страницах, а авторам, чьи работы обсуждаются, будет предложена возможность ответить в том же выпуске.

Присылайте письма через сайт журнала (journal.valuation-engineer.com) или напрямую редактору. Первые настоящие письма, которые поступят в этот журнал, станут знаком того, что он превратится из простого издания в площадку для диалога — мы с нетерпением ждем их.

Уильям Берт Крейтор, *главный редактор журнала «*
» (SRA)

Разрыв между теорией и практикой в области оценки недвижимости: структурированный обзор функций, границ и противоречий

Уильям Берт Крейтор

Vol. 1, Issue 1 · Лето 2026

Published 1 июля 2026

DOI: [10.67330/x7pbfd13](https://doi.org/10.67330/x7pbfd13)

© 2026 Уильям Берт Крейтор. Published by Pacific Vista Net.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

Abstract. Оценка недвижимости представляет значительный общественный интерес, однако ее официальная доктрина — основанная на мнениях о рыночной стоимости в прошлых и нынешних условиях — находится во все более растущем противоречии с тем, как результаты оценки фактически используются в современных финансовых механизмах. В данной статье эта проблема раскрывается в пяти этапах, каждый из которых сопровождается явной, доступной для запросов базой знаний на языке Prolog: первые два этапа определяют социальные и экономические функции, которые выполняет оценка; третий определяет её основной конечный продукт — оценку рыночной стоимости; а четвёртый исследует временную и методологическую ориентацию оценки по отношению к количественному и финансовому прогнозированию, а также регуляторную границу, разделяющую их. Все это сходится к центральному утверждению пятого этапа: в ипотечном кредитовании рыночная стоимость, масштабированная по отношению кредита к стоимости, функционирует в качестве неявного прокси-показателя будущей стоимости при обращении взыскания — это перспективное использование, которое доктрина «текущей стоимости» не признает. Многие оценщики, не осознавая этой роли оценки залога, не относятся к рыночной стоимости с той осторожностью в отношении будущих тенденций стоимости, которую она заслуживает — эта проблема и является предметом данного исследования, призванного привлечь внимание к ней и стимулировать дальнейшие исследования.

Keywords: оценка недвижимости, рыночная стоимость, доктрина оценки, представление знаний, стандарты оценки, стоимость при обращении взыскания

Содержание

1	Социальная значимость	2
2	Экономическое значение	6
3	Важность оценок рыночной стоимости как конечного результата оценки	11
4	Временная ориентация деятельности оценщика	15
5	Рыночная стоимость× Коэффициент LTV как приблизительный показатель стоимости при обращении взыскания	23
6	Заключение	30

1 Социальная значимость

Оценка недвижимости и других активов выполняет важную социальную и экономическую функцию, помогая покупателям, продавцам, кредиторам, страховщикам, инвесторам, судам, налоговым органам и широкой общественности принимать обоснованные и справедливые решения в отношении передачи, финансирования, налогообложения и управления недвижимостью и другими формами богатства. Высококачественные оценки способствуют прозрачности рынка, укреплению доверия общественности, эффективному распределению капитала и стабильности финансовых институтов, предоставляя независимые и обоснованные заключения о стоимости, основанные на доказательствах, анализе и профессиональном суждении. (LaCour-Little и Malpezzi 2003; K’Akumu и Larsen 2024; Hendershott и Kane 1995; Lee, Peng и Liao 2020)

Напротив, некачественные оценки могут привести к серьезным последствиям в различных временных горизонтах. В краткосрочной перспективе неточные оценки могут нанести отдельным покупателям или продавцам непосредственный финансовый ущерб в виде переплаты, продажи по заниженной цене, чрезмерного налогообложения или несоответствующих решений о предоставлении кредитов. В среднесрочной перспективе систематические недостатки в оценке могут искажать рыночные сигналы, ухудшать качество кредитования, поощрять спекулятивное поведение, способствовать несправедливому налогообложению или страховому покрытию, а также приводить к росту числа судебных споров и недоверия к сделкам. В долгосрочной перспективе постоянные недостатки в качестве оценки могут подорвать доверие к рынкам и институтам, способствовать финансовой нестабильности, нерациональному распределению ресурсов, упадку районов или чрезмерным инвестициям, а также более широким социальным неравенствам, затрагивающим домохозяйства, предприятия и сообщества. (LaCour-Little и Malpezzi 2003; Mayer и Nothaft 2022)

Поскольку недвижимость и другие активы зачастую составляют значительную часть личного, корпоративного и государственного богатства, качество оценки является не просто частной технической проблемой, а вопросом, представляющим значительный общественный интерес и связанным с экономической устойчивостью, справедливостью и долгосрочным общественным благом. (Mayer и 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
]).

:- disjoint social_effect/4.
:- disjoint 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.
%%% -----

социальная_сторона(покупатели).
социальная_сторона(продавцы).
социальная_сторона(кредиторы).
социальная_сторона(страховщики).
социальная_сторона(инвесторы).
социальная_сторона(суды).
social_party(налоговые органы).
social_party(широкая общественность).

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

тип_социального_решения(передача_собственности).
тип_социального_решения(финансирование).
тип_социального_решения(налогообложение).
тип_социального_решения(управление).
тип_социального_решения(страховое_покрытие).
social_decision_type(adjudication).      % courts: condemnation, divorce, estate
social_decision_type(public_administration).

```

```

social_involves(передача_собственности,    покупатели).
social_involves(передача_собственности,    продавцы).
social_involves(финансирование,            кредиторы).
social_involves(финансирование,            покупатели).
social_involves(taxation,                  taxing_authorities).
social_involves(taxation,                  general_public).
social_involves(stewardship,              investors).
social_involves(stewardship,              general_public).
social_involves(страховое_покрытие,       страховщики).
social_involves(страховое_покрытие,       покупатели).
social_involves(судебное_разрешение_споров, суды).
social_involves(государственное_управление, налоговые_органы).
social_involves(государственное_управление, широкая_общественность).

```

```

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

```

```

social_contribution(прозрачность_рынка).
social_contribution(доверие_общественности).
social_contribution(эффективное_распределение_капитала).
social_contribution(стабильность_финансовых_учреждений).

```

```

%% Characteristics of a high-quality appraisal opinion of value.
social_hallmark(независимость).
social_hallmark(обоснованность).
social_hallmark(основанность на доказательствах).
social_hallmark(основанность на анализе).
social_hallmark(основанность на профессиональном суждении).

```

```

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

```

```

%% ---- SHORT TERM: high quality ----
social_effect(высокий, краткосрочный, передача_права_собственности,
↳ справедливое_ценообразование_для_сторон).
social_effect(высокий, краткосрочный, финансирование,
↳ обоснованные_решения_по_кредитованию).
social_effect(высокий, краткосрочный, налогообложение,
↳ точная_индивидуальная_оценка).

```

```

%% ---- 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, ненадлежащие_решения_о_предоставлении
↳ кредитов).
social_effect(low, short_term, property_transfer, непосредственный_финансовый_ущерб
↳ физическим_лицам).

```

```

%% ---- MEDIUM TERM: high quality ----
social_effect(высокий, среднесрочный, рынки, четкие_рыночные_сигналы).
social_effect(высокий, среднесрочный, финансирование, надежное_кредитное
↳ обеспечение).

```

```

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, поощрение_спекуляции).
social_effect(low, medium_term, taxation, несправедливое_налогообложение).
social_effect(low, medium_term, insurance_coverage, несправедливое_страховое_покрытие).
social_effect(low, medium_term, adjudication, increased_litigation).
social_effect(low, medium_term, markets, transactional_distrust).

%% ---- LONG TERM: high quality ----
social_effect(высокий, долгосрочный, институты, доверие_общественности_к_институтам).
social_effect(высокий, долгосрочный, рынки, доверие_общественности_к_рынкам).
social_effect(высокий, долгосрочный, институты, финансовая_стабильность).
social_effect(high, long_term, communities, справедливые_результаты_для_сообществ).
social_effect(high, long_term, public_interest, экономическая_устойчивость).
social_effect(high, long_term, public_interest, социальная_справедливость).
social_effect(высокий, долгосрочный, общественный_интерес,
↪ долгосрочное_общественное_благо).

%% ---- 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, чрезмерное_инвестирование_в_районы).
social_effect(low, long_term, communities, социальное_неравенство_для_домохозяйств).
social_effect(low, long_term, communities, социальное_неравенство_для_предприятий).
social_effect(low, long_term, communities, социальное_неравенство_для_сообществ).

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

принцип_социальной_справедливости(осознанные_решения).
принцип_социальной_справедливости(справедливые_решения).
принцип_социальной_справедливости(справедливое_отношение_к_сторонам).
принцип_социальной_справедливости(независимость_оценщика).

social_public_interest_claim(
    'Недвижимость и другие активы составляют значительную часть личного, корпоративного
    ↪ и общественного богатства.').
social_public_interest_claim(
    'Качество оценки --- это не просто частная техническая проблема.').
social_public_interest_claim(
    'Качество оценки связано с экономической устойчивостью.').
social_public_interest_claim(
    'Качество оценки связано со справедливостью.').
social_public_interest_claim(
    'Качество оценки связано с долгосрочным общественным благом.').

```



```

%%% 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(Горизонт, Область, Эффект) :-
    social_effect(высокий, Горизонт, Область, Эффект).

social_negative_effect(Горизонт, Область, Эффект) :-
    social_effect(низкий, Горизонт, Область, Эффект).

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(Тип_решения, Стороны) :-
    setof(P, social_involves(Тип_решения, P), Стороны).

%%% 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 Экономическое значение

Оценка недвижимости и других активов играет основополагающую экономическую роль, способствуя эффективному функционированию кредитных рынков, принятию инвестиционных решений, формированию налоговых систем, составлению финансовых отчётов, страхованию, управлению наследством, проведению корпоративных сделок и планированию государственной политики. Надёжные оценки помогают установить достоверные показатели стоимости залога и рыночной стоимости, что позволяет капиталу более эффективно циркулировать между заемщиками, кредиторами, инвесторами, правительствами и предприятиями. (LaCour-Little и Malpezzi 2003; Ebrahim и Hussain 2010)

В краткосрочной перспективе точные оценки снижают неопределенность при совершении сделок, улучшают качество решений по кредитованию, способствуют проведению справедливых переговоров и помогают предотвратить непосредственные финансовые потери, возникающие в результате завышения или занижения стоимости. В среднесрочной перспективе качество оценки влияет на стабильность и эффективность рынков в целом, по-

сколько оказывает влияние на ипотечный риск, распределение инвестиций, справедливость налогообложения, адекватность страхового покрытия и ценообразование участников рынка. Недостоверные или систематически искажённые оценки в этот период могут способствовать образованию пузырей на рынке активов, ухудшению качества кредитных портфелей, искажению стимулов для развития и увеличению риска дефолта или судебных разбирательств. (LaCour-Little и Malpezzi 2003; Hendershott и Kane 1992; Griffin и Maturana 2016)

В долгосрочной перспективе совокупные экономические последствия качества оценки становятся ещё более значимыми, влияя на динамику накопления капитала, инвестиции в инфраструктуру, передачу богатства между поколениями, государственные доходы, доступность жилья и общую устойчивость финансовой системы. Устойчивые недостатки в практике оценки могут способствовать хроническому нерациональному распределению капитала, банковской нестабильности, снижению доверия инвесторов и неэффективному использованию земельных участков и производственных активов, в то время как строгие стандарты оценки способствуют стабильности рынков, рациональному экономическому планированию и долгосрочному экономическому росту. (Ghent, Torous и Valkanov 2019; Hendershott и Kane 1992; Ebrahim и 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.

:- модуль(appraisal_economics,
    [ economic_horizon/1
    , economic_function/1
    , economic_supports/2           % Function, Horizon
    , economic_effect/4           % Quality, Horizon, Function, Effect
    , economic_stakeholder/1
    , economic_benefits/2         % Stakeholder, Function
    , economic_positive_effect/3
    , economic_negative_effect/3
    , economic_horizon_summary/2 % Horizon, Effects
    , economic_consequences_of_quality/2 % Quality, ConsequenceList
    , 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(управление_имуществом).
economic_function(корпоративные_сделки).
economic_function(планирование_государственной_политики).

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

economic_stakeholder(заемщики).
economic_stakeholder(кредиторы).
economic_stakeholder(инвесторы).
economic_stakeholder(правительства).
economic_stakeholder(предприятия).
economic_stakeholder(страховщики).
economic_stakeholder(налогоплательщики).
economic_stakeholder(домовладельцы).

%% Capital flows more efficiently when appraisals are reliable.
economic_benefits(заемщики, кредитные_рынки).
economic_benefits(кредиторы, кредитные_рынки).
economic_benefits(инвесторы, инвестиционные_решения).
economic_benefits(правительства, налоговые_системы).
economic_benefits(правительства, планирование государственной политики).
economic_benefits(предприятия, корпоративные сделки).
economic_benefits(предприятия, финансовая отчётность).
economic_benefits(страховщики, страховое андеррайтинг).
economic_benefits(налогоплательщики, налоговые системы).
economic_benefits(домовладельцы, кредитные рынки).

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

economic_supports(кредитные_рынки, краткосрочные).

```

```

economic_supports(инвестиционные_решения, краткосрочные).
economic_supports(корпоративные_сделки, краткосрочные).

economic_supports(кредитные_рынки, среднесрочный).
economic_supports(инвестиционные_решения, среднесрочный).
economic_supports(налоговые_системы, среднесрочный).
economic_supports(страховое_андеррайтинг, среднесрочный).

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(управление_имуществом, долгосрочный).

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

%% ---- SHORT TERM: high quality ----
economic_effect(высокий, краткосрочный, кредитные_рынки, снижение_неопределённости
↳ при_совершении_сделок).
economic_effect(высокий, краткосрочный, кредитные_рынки, улучшение_качества
↳ решений_о_предоставлении_кредитов).
economic_effect(высокий, краткосрочный, корпоративные_сделки, честные_переговоры).
economic_effect(high, short_term, investment_decisions,
↳ предотвращение_непосредственных_убытков).

%% ---- 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(высокий, среднесрочный, кредитные_рынки, стабильный_ипотечный
↳ риск).
economic_effect(высокий, среднесрочный, инвестиционные_решения, эффективное
↳ распределение_инвестиций).
economic_effect(высокий, среднесрочный, налоговые_системы, налоговая_справедливость).
economic_effect(high, medium_term, страховое_андеррайтинг,
↳ адекватное_страховое_покрытие).
economic_effect(high, medium_term, инвестиционные_решения,
↳ рациональное_рыночное_ценообразование).

%% ---- MEDIUM TERM: low quality ----
economic_effect(низкий, среднесрочный, инвестиционные_решения,
↳ пузыри_на_рынке_активов).
economic_effect(low, medium_term, credit_markets, обесцененные_кредитные_портфели).
economic_effect(low, medium_term, public_policy_planning, искаженные_стимулы_развития).
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(высокий, долгосрочный, инвестиционные_решения,
↳ рациональное_накопление_капитала).

```

```

economic_effect(высокий, долгосрочный, планирование_государственной_политики,
↳ рациональные_инвестиции_в_инфраструктуру).
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, доступность_субсидируемого_жилья).
economic_effect(high, long_term, credit_markets, устойчивость_финансовой_системы).
economic_effect(high, long_term, investment_decisions, долгосрочный_экономический_рост).

%% --- LONG TERM: low quality ---
экономический_эффект(низкий, долгосрочный, инвестиционные_решения,
↳ хроническое_неправильное_распределение_капитала).
экономический_эффект(низкий, долгосрочный, кредитные_рынки,
↳ нестабильность_банковской_системы).
economic_effect(low, long_term, investment_decisions, reduced_investor_confidence).
economic_effect(low, long_term, планирование_государственной_политики,
↳ неэффективное_использование_земельных_ресурсов).
economic_effect(low, long_term, планирование_государственной_политики,
↳ неэффективное_использование_производственных_активов).

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

economic_positive_effect(Горизонт, Функция, Эффект) :-
    economic_effect(высокий, Горизонт, Функция, Эффект).

economic_negative_effect(Горизонт, Функция, Эффект) :-
    economic_effect(low, Горизонт, Функция, Эффект).

economic_horizon_summary(Горизонт, Эффекты) :-
    findall(Качество-Эффект,
        economic_effect(Качество, Горизонт, _Функция, Эффект),
        Эффекты).

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(Функция, Заинтересованные_стороны) :-
    setof(S, economic_benefits(S, Функция), Заинтересованные_стороны).

%% Overall claim: reliable appraisals improve allocative efficiency.
экономически_обоснованные_и_достоверные_оценки_повышают_эффективность_движения
↳ капитала.
экономически_обоснованные_и_достоверные_оценки_способствуют_стабильности_рынка.
экономически_обоснованные_и_достоверные_оценки_способствуют_долгосрочному
↳ экономическому_росту.

%% 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 Важность оценок рыночной стоимости как конечного результата оценки

Оценка рыночной стоимости играет важную роль в анализе рисков, связанных с залоговым обеспечением, поскольку она представляет собой профессионально разработанную оценку вероятной цены, по которой недвижимость или актив могли бы быть реализованы в нормальных рыночных условиях, тем самым служа основополагающим ориентиром для кредиторов, инвесторов и аналитиков по залоговому обеспечению, оценивающих кредитный риск и возможные последствия обращения взыскания. (Agarwal, Ben-David и Yao 2015) Хотя стоимость при обращении взыскания или ликвидационная стоимость могут существенно отличаться от рыночной стоимости из-за неблагоприятных условий реализации, ускоренных сроков продажи, судебных издержек, физического износа или негативных настроений на рынке, аналитики по обеспечению зачастую опираются на оценки рыночной стоимости в качестве отправной точки для оценки потенциала возмещения в условиях стрессовых сценариев. (Campbell, Giglio и Pathak 2011; Pennington-Cross 2006; Clauretie и Daneshvary 2009)

Анализируя взаимосвязь между рыночной стоимостью, остатками по кредитам, отношением кредита к стоимости, рыночными тенденциями, состоянием объектов недвижимости, ликвидностью и местными условиями спроса и предложения, аналитики по залоговому обеспечению оценивают вероятность и размер убытков в случае дефолта заемщика. (Qi и Yang 2009)

В краткосрочной перспективе точные оценки рыночной стоимости помогают кредитора́м принимать обоснованные решения по выдаче кредитов и устанавливать надлежащие маржи по залому. В среднесрочной перспективе они способствуют управлению рисками портфеля, секьюритизации кредитов, формированию резервов и соблюдению нормативных требований, помогая финансовым учреждениям оценивать изменение адекватности залога на протяжении экономических циклов. В долгосрочной перспективе последовательные и надёжные методы оценки способствуют стабильности финансовой системы, снижая риск чрезмерного использования заёмных средств, кредитования с недостаточным обеспечением и системного неправильного ценообразования залоговых активов, что может усилить убытки от обращения взыскания и привести к более широким экономическим потрясениям в периоды рыночного напряжения. (Eriksen и др. 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
%%% -----

аналитик_по_обеспечению(кредиторы).
аналитик_по_обеспечению(инвесторы).
аналитик_по_обеспечению(аналитики_по_обеспечению).
аналитик_обеспечения(менеджеры_по_рискам_портфеля).
аналитик_обеспечения(регулирующие_органы).
аналитик_обеспечения(организаторы_секьюритизации).

%%% -----
%%% 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(кредиторы, рыночная_стоимость).
uses(инвесторы, рыночная_стоимость).
uses(аналитики_обеспечения, рыночная_стоимость).
uses(аналитики_обеспечения, стоимость_при_обращении_к_исполнению).
uses(аналитики_обеспечения, ликвидационная_стоимость).
использует(менеджеров_по_рискам_портфеля, рыночная_стоимость).
использует(организаторов_секьюритизации, рыночная_стоимость).
использует(регулирующие_органы, рыночная_стоимость).

%% 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
%% -----

коэффициент_расхождения(неблагоприятные_рыночные_условия).
коэффициент_расхождения(сокращенный_срок_продажи).
коэффициент_расхождения(судебные_издержки).
коэффициент_расхождения(физический_износ).
коэффициент_расхождения(неблагоприятные_рыночные_настроения).

%% Stressed scenarios in which divergence factors apply.
stress_scenario(дефолт_заемщика).
stress_scenario(обращение_взыскания_на_залог).
stress_scenario(принудительная_ликвидация).
stress_scenario(спад_на_рынке).

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

risk_input(рыночная_стоимость).
risk_input(остаток_по_кредиту).
risk_input(коэффициент_«кредит/стоимость»).
risk_input(рыночные_тенденции).
risk_input(состояние_недвижимости).
risk_input(ликвидность).
risk_input(местный_спрос_и_предложение).

risk_output(вероятность_убытка).
risk_output(размер_убытка).
risk_output(оценка_возмещения).
risk_output(убыток_в_случае_дефолта).

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

```



```

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

%% ---- SHORT TERM: high quality ----
побочный_эффект(высокий, краткосрочный, андеррайтинг,
↳ взвешенные_решения_по_андеррайтингу).
побочный_эффект(высокий, краткосрочный, андеррайтинг,
↳ надлежащие_маржи_обеспечения).

%% ---- SHORT TERM: low quality ----
побочный_эффект(низкий, краткосрочный, андеррайтинг,
↳ неосторожные_решения_по_андеррайтингу).
побочный_эффект(низкий, краткосрочный, андеррайтинг,
↳ недостаточные_маржи_обеспечения).
побочный_эффект(низкий, краткосрочный, андеррайтинг,
↳ неправильно_оцененный_кредитный_риск).

%% ---- MEDIUM TERM: high quality ----
побочный_эффект(высокий, среднесрочный, портфель,
↳ эффективное_управление_рисками_портфеля).
побочный_эффект(высокий, среднесрочный, портфель,
↳ надежная_секьюритизация_кредитов).
побочный_эффект(высокий, среднесрочный, портфель,
↳ надлежащее_распределение_резервов).
collateral_effect(high, medium_term, regulatory, regulatory_compliance).
collateral_effect(high, medium_term, portfolio, cycle_aware_collateral_evaluation).

%% ---- MEDIUM TERM: low quality ----
побочный_эффект(низкий, среднесрочный, портфель,
↳ ослабление_управления_рисками_портфеля).
побочный_эффект(низкий, среднесрочный, портфель,
↳ неправильная_оценка_секьюритизаций).
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 ----
побочный_эффект(высокий, долгосрочный, системный, стабильность_финансовой_системы).
побочный_эффект(высокий, долгосрочный, системный,
↳ ограничение_чрезмерного_кредитования).
побочный_эффект(высокий, долгосрочный, системный, кредитование_под
↳ надежное_обеспечение).
побочный_эффект(высокий, долгосрочный, системный,
↳ точное_системное_оценка_обеспечения).
побочный_эффект(высокий, долгосрочный, системный,
↳ смягчение_усиления_эффекта_изъятия_залога).

%% ---- LONG TERM: low quality ----
побочный_эффект(низкий, долгосрочный, система, нестабильность_финансовой_системы).
побочный_эффект(низкий, долгосрочный, система, чрезмерное_кредитное_леверидж).
побочный_эффект(низкий, долгосрочный, система,
↳ кредитование_с_недостаточным_обеспечением).
побочный_эффект(низкий, долгосрочный, система,
↳ системная_неправильная_оценка_обеспечения).
побочный_эффект(низкий, долгосрочный, система, усиление_убытков_от_изъятия_залога).

```

```

побочный_эффект(низкий, долгосрочный, система,
↳ усиление_экономических_потрясений_в_условиях_стресса).

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

collateral_positive_effect(Горизонт, Домен, Эффект) :-
    collateral_effect(высокий, Горизонт, Домен, Эффект).

collateral_negative_effect(Горизонт, Область, Эффект) :-
    collateral_effect(low, Горизонт, Область, Эффект).

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 Временная ориентация деятельности оценщика

Оценщики, как правило, формируют заключения о рыночной стоимости на определенный момент времени — на конкретную дату оценки — ретроспективные (прошлые), текущие или, с определенными ограничениями, прогностические (будущие) — а не вероятностные прогнозы на длительные периоды времени. Эта временная и методологическая ориентация представляет собой одно из основных отличий между регулируемой практикой оценки и работой количественных аналитиков, финансовых аналитиков, инвестиционных аналитиков и экономистов. (Quap и Quigley 1991)

В большинстве заданий по оценке, особенно тех, которые касаются рыночной стоимости в соответствии со стандартами, такими как «Единые стандарты профессиональной оценки (USPAP) или Международных стандартов оценки (IVS), оценщик формулирует согласованную точечную оценку (или, в разрешенных случаях, узкий диапазон), основанную на наблюдаемых рыночных данных, преобладающих условиях и признанных методологиях по дате оценки. USPAP прямо допускает формулировку выводов о стоимости в виде конкретной суммы, диапазона значений или соотношения к эталонному показателю, однако согласованное заключение по-прежнему представляет собой достоверную точечную оценку, привязанную к этой дате. (The Appraisal Foundation 2024; International Valuation Standards Council 2025)

Хотя оценщики неизбежно учитывают обоснованные ожидания участников рынка в отношении тенденций, спроса и предложения, амортизации, ставок капитализации или дисконтирования, а также темпов впитывания, они делают это для обоснования рыночной стоимости на дату оценки, а не для прогнозирования спекулятивных колебаний. Заключения о перспективной стоимости (например, стабилизированная стоимость по завершении проекта) требуют явных исключительных допущений или гипотетических условий и ограничиваются определённой будущей датой оценки, подкреплённой рыночными данными.

В отличие от этого, аналитики, занимающиеся количественным, финансовым и инвестиционным анализом, часто учитывают неопределенность путем составления оценок, охватывающих диапазон возможных результатов в течение различных временных периодов,

с использованием вероятностных распределений, симуляций по методу Монте Карло, анализ сценариев, тестирование чувствительности или показатели «стоимости под риском». Их внимание часто сосредоточено на ожидаемых будущих результатах, доходности с поправкой на риск или инвестиционной стоимости для конкретной стороны в течение различных временных горизонтов, а не на единичной рыночной стоимости на определенную дату. (Hull 2018)

Такая работа также предполагает наличие технических компетенций, которые лицензия по оценке ни не требует, ни не подтверждает. Например, подбор многомерной адаптивной регрессионной сплайновой модели (MARS) требует умения контролировать сложность модели таким образом, чтобы хорошая подгонка внутри выборки (высокий показатель R^2) не достигалась за счёт переобучения, а напротив, переносилась на тестовые данные (высокий коэффициент перекрёстной проверки R^2), что, в свою очередь, предполагает практическое понимание лежащих в основе алгоритмов. (Friedman 1991; Steurer, Hill и Pfeifer 2021) Наряду с такими количественными навыками, эта работа вознаграждает концептуальную способность работать с высокоструктурированными понятиями, различать тонкие нюансы в значении и осваивать незнакомые методы. Это скорее навыки, чем квалификационные документы: лицензированный оценщик может ими обладать, а нелицензированный аналитик — нет, и именно поэтому они сами по себе не определяют, по ту или другую сторону регуляторной границы находится та или иная задача.

Оценщики также могут использовать в ограниченном объеме элементы, ориентированные на будущее, такие как прогнозирование потоков доходов, уровня вакантности, расходов и реверсионной стоимости в анализе дисконтированных денежных потоков, либо оценка перспективной стоимости в контексте девелоперских проектов или судебных разбирательств. Эти компоненты остаются в рамках оценки, если они обоснованы текущими рыночными данными и профессиональными стандартами, которые подчеркивают необходимость обоснованности и избегания неподтвержденных спекуляций. (Diaz и Wolverton 1998; Gallimore 1996; Crosby, Lizieri и McAllister 2010) Современная практика оценки все чаще интегрирует статистические инструменты, анализ сценариев и анализ рисков — особенно при массовой оценке, использовании автоматизированных моделей оценки или работе с портфелями — однако основным результатом, как правило, остается заключение о рыночной стоимости на определенный момент времени. (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 и Neal 2023)

Практику оценки можно охарактеризовать как преимущественно доказательную и ориентированную на точечные оценки «от прошлого к настоящему» (или с ограниченным прогнозным горизонтом), тогда как финансовый и количественный анализ носит более вероятностный и распределительный характер и явно ориентирован на будущее в рамках различных временных горизонтов. (Geltner 1991) Оценщики задают вопрос: «What is the most supportable market value indication as of the specified effective date, given current evidence and defined assumptions?» Финансовые и количественные аналитики же чаще задают вопрос: «What range of outcomes is likely over a future period, with what probabilities, and how should resources be allocated?»

Эта граница не является абсолютной и продолжает меняться по мере усложнения рынка. Тем не менее сохраняются принципиальные различия в целях (рыночная стоимость против инвестиционной стоимости по методу « Δ »), требованиях к доказательствам и нормативной базе.

Однако граница, определяющая сферу деятельности, проходит не по уровню квалификации, а по закону: решающий вопрос заключается просто в том, требуется ли для выполнения данной задачи лицензия оценщика в соответствии с законодательством. Количественные

аналитики, финансовые аналитики, экономисты, инвестиционные аналитики, портфельные менеджеры и специалисты смежных профессий могут на законных основаниях оценивать, прогнозировать, моделировать или анализировать будущую стоимость активов — включая прогнозы, основанные на распределении или диапазонах значений — без лицензии оценщика недвижимости, при условии, что такая работа не представляет собой регулируемую оценку. (Murphy 2012) Ключевые различия заключаются в следующем:

1. типе стоимости (например, точечная оценка рыночной стоимости по сравнению с инвестиционной стоимостью или вероятностным прогнозом);
2. временной фокус (конкретная дата оценки по сравнению с временным периодом или горизонтом);
3. форма вывода (точечное значение или узкий диапазон по сравнению с полным вероятностным распределением);
4. предполагаемое назначение и то, представляется ли работа в качестве оценки; а также
5. участие регулируемых кредитных организаций или правовые контексты, требующие соблюдения стандартов, таких как USPAP.

Данная система в значительной степени сформировалась на основе банковского регулирования, такого как закон FIRREA (1989) в США, целью которого было обеспечение достоверных и независимых заключений о рыночной стоимости при проведении сделок, связанных с федеральными учреждениями. (United States Congress 1989; K'Akumu и Larsen 2024) Многие сложные виды оценки — целевые цены в рамках анализа акций, технико-экономические обоснования, моделирование рисков портфеля, оценка стоимости опционов или прогнозная аналитика — не подпадают под требования лицензирования, поскольку служат инвестиционным, прогнозным или консультационным целям, а не регулируемым функциям оценки.

```
%% 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.

:- модуль(регулирование_объема_оценки,
    [% --- TIER 1: the defining test (regulation) ---
     нормативная_база/1
```

```

, происхождение_нормативной_базы/2
, отличительный_фактор_регулирования/1
, факторы_регулирования_лицензирования/1
, определение_оценки_в_нормативных_актах/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
, нерегулируемая_оценка/1
, регуляторная_асимметрия/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
)].

:- несоединенное_регулирование_нелицензируемой_деятельности/2.
:- несоединенное_регулирование_лицензируемой_деятельности/2.
:- несоединенная_сфера_применения_контраст/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).

регулирующий_дифференцирующий_фактор(тип_значения).
регулирующий_дифференцирующий_фактор(цель_оценки).
регулирующий_различающий_фактор(временной_фокус).
регулирующий_различающий_фактор(форма_результата).

```

```

регулирующий_различающий_фактор(представление_в_виде_оценки).
регулирующий_различающий_фактор(участие_в_сделках,_связанных_с_федеральным_у_
↪ ровнем).

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).

определение_оценки_в_нормативных_актах (официальное_заключение_о_стоимости).
определение_оценки_в_нормативных_актах
↪ (в_отношении_идентифицированного_недвижимого_имущества).
определение_оценки_в_нормативных_актах
↪ (проведенной_в_соответствии_с_признанными_методами).
определение_оценки_в_нормативных_актах(для_целей_компенсации).

regulation_unlicensed_activity(equity_analyst,           целевые_цены_акций).
regulation_unlicensed_activity(investment_bank,          оценки_стоимости_предприятия).
regulation_unlicensed_activity(hedge_fund,
↪ прогнозы_роста_стоимости_коммерческой_недвижимости).
regulation_unlicensed_activity(экономист,
↪ прогнозы_тенденций_на_рынке_жилья).
regulation_unlicensed_activity(страховой_аналитик,
↪ оценки_восстановительной_стоимости).
regulation_unlicensed_activity(застройщик,               технико-экономические_обоснования).
regulation_unlicensed_activity(внутренний_корпоративный_аналитик, оценки_обесценения
↪ активов).
regulation_unlicensed_activity(внутренний_корпоративный_аналитик, оценки_стоимости
↪ портфеля).
regulation_unlicensed_activity(quantitative_analyst,     predictive_valuation_models).
regulation_unlicensed_activity(avm_developer,            algorithmic_residential_price_estimates).

regulation_licensed_activity(licensed_appraiser,
↪ рыночная_стоимость_для_сделки,_связанной_с_федеральными_органами).
regulation_licensed_activity(licensed_appraiser,
↪ официальный_заклучение_о_стоимости_за_вознаграждение).
regulation_licensed_activity(licensed_appraiser,
↪ оценка_для_регулируемого_ипотечного_кредитования).
regulation_licensed_activity(licensed_appraiser,
↪ оценка,_представленная_как_такая_в_соответствии_с_законодательством_штата).

условие_нормативной_правовой_ответственности(представление_себя_в_качестве_лиценз_
↪ ированного_оценщика).
условие_нормативной_правовой_ответственности(подготовка_оценки_в_соответствии_с_з_
↪ аконодательством_штата).
условие_правового_риска_в_связи_с_нормативными_требованиями
↪ (участие_в_сделке,_требующей_оценки,_проведенной_лицензированным_оценщиком).

%% --- 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).

:- динамическая функция 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.
регуляторная_асимметрия(текущая_рыночная_стоимость_для_кредитования,
    ориентированная_на_будущее_оценка_для_инвестиций).
регуляторная_асимметрия(строго_регулируемая_практика_оценки,
    слабо_регулируемая_количественная_практика).
regulatory_asymmetry(ориентация_на_защиту_потребителей_и_кредитование,
    ориентация_на_инвестиционное_консультирование).

regulatory_asymmetry_origin(banking_regulation).
regulatory_asymmetry_origin(consumer_protection).
regulatory_asymmetry_origin(правовые_вопросы_доказательств).
regulatory_asymmetry_origin_disclaimer(
    'Не основано на каком-либо исключительном утверждении о том, что только оценщики
    → обладают экспертными знаниями в области оценки.').

%% =====
↪ =====
%% 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) ---
метод_оценки(оценка).
метод_оценки(количественный_финансовый).

scope_profession(оценщик).
scope_profession(количественный_аналитик).
scope_profession(финансовый_аналитик).
scope_profession(инвестиционный_аналитик).
scope_profession(управляющий_портфелем).
scope_profession(экономист).
scope_profession(инвестиционный_стратег).
scope_profession(финансовый_прогнозист).

scope_orientation(оценка, доказательная).
scope_orientation(количественно-финансовый, вероятностный).

```



```

%%% --- Axis: temporal scope (its OWN viewpoint, distinct from approach) ---
temporal_scope(ретроспективный).
temporal_scope(текущий).
temporal_scope(проспективная точка).
temporal_scope(горизонт будущего).

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(оценка, момент_во_времени).
scope_temporal_focus(финансовый_показатель, распределение_по_горизонту).

%%% --- Axis: output form ---
scope_output_form(оценка, point_estimate).
scope_output_form(оценка, 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(оценка, текущие_рыночные_данные).
scope_inputs_considered(оценка, наблюдаемые_сделки).
scope_inputs_considered(оценка, сложившиеся_рыночные_условия).
scope_inputs_considered(оценка, признанные_методики).
scope_inputs_considered(оценка, рыночные_тенденции).
scope_inputs_considered(оценка, прогнозируемое_соотношение_спроса_и_предложения).
scope_inputs_considered(оценка, амортизация).
scope_inputs_considered(оценка, ставки_капитализации).
scope_inputs_considered(оценка, модели_поглощения).
scope_inputs_considered(оценка, ожидания, уже_учтенные_в_цене).

%%% --- Bounded future-oriented contexts that remain within appraisal ---
scope_future_oriented_context(доходная_недвижимость, будущие_потоки_доходов).
scope_future_oriented_context(доходная_недвижимость, коэффициенты_вакантности).
scope_future_oriented_context(доходная_недвижимость, расходы).
scope_future_oriented_context(доходная_недвижимость,
↪ стоимость_при_возврате_в_первоначальное_состояние).

```



```

scope_future_oriented_context(доходная_недвижимость, dcf_анализ).
область_применения_ориентированного_на_будущее_контекста(застройка,
↪ прогнозируемая_стабилизированная_стоимость).
область_применения_ориентированного_на_будущее_контекста(сохранение,
↪ будущий_ущерб).
область_применения_ориентированного_на_будущее_контекста(экология,
↪ будущий_ущерб).
scope_future_oriented_context(страхование,    будущие_убытки).
scope_future_oriented_context(судебные_разбирательства,    будущие_убытки).
scope_future_oriented_context(право_принудительного_отчуждения,    будущие_убытки).
scope_future_oriented_context(общий,    оставшийся_срок_эксплуатации).
scope_future_oriented_context(общий,    ожидаемые_реакции_рынка).
scope_future_oriented_context(анализ_рынка, направление_ценовой_динамики).

%% --- Professional constraints on any future-oriented component ---
scope_professional_constraint(обоснованность).
scope_professional_constraint(рыночные_данные).
scope_professional_constraint(раскрытие_необычных_допущений_или_гипотетических_усло_
↪ вий).
scope_professional_constraint(отказ_от_необоснованных_спекуляций).

%% --- Contrasts across the descriptive axes: scope_contrast(A,B,Dimension,Distinction) ---
scope_contrast(оценка, количественный_финансовый_показатель, ориентация,
↪ доказательный_подход_против_вероятностного_подхода).
scope_contrast(оценка, количественный_финансовый_показатель, временной_режим,
↪ точечный_момент_времени_против_распределения_по_горизонту).
scope_contrast(оценка, количественный_финансовый_анализ, форма_результата,
↪ точечное_значение_или_узкий_диапазон_vs_вероятностное_распределение).
scope_contrast(оценка, количественный_финансовый_анализ, цель,
↪ рыночная_стоимость_vs_инвестиционная_ценность).
scope_contrast(оценка, количественные_финансовые_показатели, набор_навыков,
↪ набор_инструментов_для_лицензирования_оценки_vs_набор_инструментов_дл_
↪ я_расширенного_статистического_анализа).
scope_contrast(оценка, количественный_финансовый_анализ, типичные_методы,
↪ метод_сравнения_продаж_стоимости_дохода_vs_метод_моделирования_оценки_
↪ _вариативных_опционов).

%% --- The orienting question each population asks (verbatim from the passage) ---
scope_question(оценка,
↪ 'Какой показатель рыночной стоимости является наиболее обоснованным на указанную
↪ дату оценки с учетом имеющихся данных и заданных допущений?').
scope_question(quant_financial,
↪ 'Какой диапазон результатов вероятен в будущем, с какой вероятностью, и как следует
↪ распределять ресурсы?').

%% --- 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 Рыночная стоимость × Коэффициент LTV как приблизительный показатель стоимости при обращении взыскания

В предыдущих разделах было изложено, для чего нужна оценка, к каким результатам она приводит и где пролегают её доктринальные и нормативные границы; в данном разделе рассматривается та противоречивая ситуация, к которой развивалось обсуждение в предыдущих разделах.

Одной из важных практических реальностей, лежащих в основе ипотечного кредитования и анализа залога, является то, что в реальной экономической практике оценка рыночной стоимости, используемая при оформлении ипотечного кредита, зачастую служит неявным показателем ожидаемой возможности возмещения стоимости залога в течение некоторого будущего периода владения, хотя формально такая оценка формулируется как заключение о текущей рыночной стоимости на конкретную дату вступления в силу (Foster и Van Order 1984; Kau и др. 1995; Deng, Quigley и Van Order 2000)

Коэффициент «отношение кредита к стоимости» в размере 80 % не является произвольным. Исторически он сформировался отчасти как буфер риска, призванный защитить кредиторов от:

- умеренное падение рыночных цен,
- расходы, связанные с обращением взыскания на залог,
- скидки при ликвидации,
- расходы на содержание активов,
- судебные издержки,
- начисленные проценты,
- а также неопределенность в течение периода, необходимого для реализации залога после наступления дефолта. (Foote, Gerardi и Willen 2008; Qi и Yang 2009)

Этот запас также объясняет, почему сама сумма кредита — произведение рыночной стоимости и допустимого отношения кредита к стоимости — можно рассматривать как рабочую оценку кредитором возмещаемой стоимости при обращении взыскания: ограничение суммы кредита до определенной доли рыночной стоимости является, по сути, ставкой на то, что выручка от принудительной продажи не опустится ниже остатка по кредиту. Однако определить, насколько значительным должно быть это снижение стоимости, гораздо сложнее, чем позволяет предположить округленная цифра в 80 %. Оценки дисконта при обращении взыскания широко варьируются в зависимости от рынков, методов и периодов; одно из часто цитируемых исследований принудительных продаж оценивает его примерно в 27 % ниже аналогичных рыночных цен (Campbell, Giglio и Pathak 2011), в то время как другие отчёты — с поправкой на состояние недвижимости, местоположение и подбор объектов — указывают на существенно иные значения. (Pennington-Cross 2006; Clauretie и Daneshvary 2009)

Традиционный порог в 80 %, напротив, по сути представляет собой единый общенациональный параметр, применяемый к чрезвычайно неоднородным местным рынкам и последовательным фазам кредитного цикла. Он не может отражать дисконт какого-либо конкретного рынка и поэтому выступает в качестве весьма приблизительной, универсальной оценки, а не точного соотношения между рыночной стоимостью и суммой, взыскиваемой при обращении взыскания. Если и говорить о каком-то соотношении, то оно весьма незначительное: дисконт в размере 20 % находится ниже указанного в отчёте дисконта в 27 %, поэтому максимальная сумма традиционного кредита может превышать валовую выручку от принудительной продажи еще до вычета перечисленных выше затрат на реализацию.

В тех случаях, когда собственный капитал заемщика слишком мал, чтобы служить буфером, его заменяет ипотечное страхование. Это следует рассматривать не как регулирующий механизм, а как компромисс в распределении рисков: сегмент заемщиков обладает

небольшим объёмом активов, но стабильным доходом, а дополнительный риск дефолта и убытков, связанный с предоставлением им кредитов с коэффициентом LTV выше 80% , прогнозируется и принимается на себя страховщиком — частным, посредством частного ипотечного страхования по стандартному кредиту с высоким коэффициентом LTV, или государственным, через Федеральное агентство по жилищному строительству. Однако рынок, на котором такой компромисс является процедурой, является результатом федеральной жилищной политики: государственные предприятия, доминирующие в сфере традиционного кредитования, по уставу лишены права приобретать кредиты с коэффициентом LTV выше 80% без такого усиления кредитоспособности (United States Code 1970) . Условия, при которых страхование прекращается, также являются результатом институциональных решений, а не расчётов рисков, и они резко различаются в разных программах.¹

Таким образом, хотя оценщик официально определяет рыночную стоимость («current market value,»), кредитор с экономической точки зрения интерпретирует эту стоимость с учетом перспективных рисков. По сути, кредитор задает вопрос, близкий к следующему:

«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?»

По сути, это вопрос, ориентированный на будущее.

Это основано на намеренном концептуальном разграничении между:

- заключение оценщика «current market value» , а также
- практическое использование этим заключением кредитором в качестве страхового резерва на случай будущего обесценения залога.

С исторической и правовой точек зрения регулирующие органы стремились не допустить, чтобы оценщики превратились в спекулятивных прогнозистов, поскольку:

- спекулятивное прогнозирование способствовало возникновению прошлых кредитных кризисов,
- суды требовали, чтобы оценки основывались на объективных данных,
- а банковские регуляторы стремились к внедрению стандартизированных и обоснованных методов оценки.

Таким образом, доктрина оценки развивалась в следующем направлении:

- текущая стоимость,
- актуальные рыночные данные,
- установленные даты оценки,
- и ограниченные допущения.

Между тем кредиторы, инвесторы и аналитики, занимающиеся оценкой залогового обеспечения, проводят процедуру использования этих же оценок в рамках вероятностных моделей оценки будущих рисков.

На самом деле, процесс одобрения ипотечного кредита представляет собой комплексную систему, включающую:

1. текущая рыночная оценка,

¹Обычное частное ипотечное страхование, оплачиваемое заёмщиком, автоматически прекращается при коэффициенте LTV 78% и может быть запрошено при коэффициенте 80% в соответствии с Законом о защите домовладельцев 1998 года. (United States Congress 1998) Требования Федерального агентства по жилищному строительству более строгие: для кредитов, выданных 3 июня 2013 года или позднее с первоначальным отношением кредита к стоимости недвижимости выше 90% (например, первоначальный взнос 3,5%), годовая премия взимается в течение всего срока действия кредита, тогда как для кредитов с соотношением суммы кредита к стоимости недвижимости (LTV) на уровне 90 или ниже% она отменяется по истечении одиннадцати лет. (U.S. Department of Housing and Urban Development 2013) Поэтому многие заемщики FHA избавляются от ипотечного страхования только путем рефинансирования в обычный кредит с соотношением суммы кредита к стоимости недвижимости (LTV) около 80% .

2. оценка будущих рисков,
3. вероятность дефолта,
4. сроки обращения взыскания,
5. степень убыточности ликвидации,
6. макроэкономические прогнозы,
7. и анализ стрессовых сценариев портфеля.

Оценщик официально вносит вклад в основном в пункт (1), тогда как кредитор внутренне объединяет этот результат с остальными.

На практике это различие становится не столь чётким. Если кредитор не предоставит кредит, пока не убедится, что залог, вероятно, сохранит достаточную стоимость в будущем, то оценка косвенно выполняет функцию модели защиты будущей стоимости, даже если стандарты оценки формально избегают такого описания.

Эта напряженность стала ещё более заметной в связи с:

- AVMs,
- — аналитика институциональных портфелей,
- — модели секьюритизации,
- — системы оценки на основе машинного обучения,
- — моделирование капитала в соответствии с Базельскими соглашениями,
- — а также крупномасштабное стресс-тестирование ипотечных кредитов.

В современной аналитике залогового обеспечения стоимость недвижимости всё чаще рассматривается с точки зрения вероятностных изменений во времени, а не в виде одномоментных статических оценок. (Basel Committee on Banking Supervision 2017; Steurer, Hill и Pfeifer 2021; Board of Governors of the Federal Reserve System 2021)

Можно утверждать, что традиционное регулирование оценки отражает исторический компромисс:

- Оценщики предоставляют четкие, основанные на фактах ориентиры текущей стоимости,
- в то время как финансовые учреждения берут на себя ответственность за интерпретацию будущих рисков.

Остается ли это разделение полностью обоснованным в современных финансовых системах — это вполне обоснованный вопрос, который вызывает всё более активные дискуссии.

Из этого вытекает практическое следствие для самой практики оценки. Большинство оценщиков не осознают эту роль оценки залога — то, что кредитор использует рыночную стоимость в качестве буфера на будущее, редко заметно в рамках самого задания по оценке — и поэтому не относятся к своим выводам о рыночной стоимости с той осторожностью в отношении будущих тенденций стоимости, которую требует эта роль. Понимание того, как на самом деле используется заключение о текущей стоимости, является первым шагом к устранению этого пробела, и именно эту область данная статья призвана открыть для дальнейшей работы.

```
%% 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(
    'Оценка рыночной стоимости, используемая при оформлении ипотечного кредита, зачастую
    ↪ служит неявным приблизительным показателем ожидаемой возможности возмещения
    ↪ стоимости залога в течение будущего периода владения.').
proxy_thesis(

```

```

'Оценка формально формулируется как заключение о текущей рыночной стоимости по
→ состоянию на конкретную дату оценки.').
proxy_thesis(
    «Кредитор экономически интерпретирует оценку через призму перспективных рисков.»).
proxy_thesis(
    «Существует намеренное концептуальное разграничение между заключением оценщика о
    → текущей рыночной стоимости и использованием этого заключения кредитором в
    → качестве буфера прогнозирования.»).
proxy_thesis(
    'Оценка косвенно функционирует как часть модели защиты будущей стоимости, хотя
    → стандарты оценки формально избегают такого описания.').
proxy_thesis(
    'Остается ли разграничение между доктриной оценки и практикой кредиторов полностью
    → согласованным в современных финансовых системах --- это открытый и все более
    → обсуждаемый вопрос.').

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

proxy_historical_ltv_threshold(80_percent).

прокси-буфер LTV на случай (умеренного падения рынка).
прокси-буфер LTV на случай (расходов, связанных с обращением взыскания).
прокси-резерв LTV на случай (скидок при ликвидации).
прокси-резерв LTV на случай (учетных затрат).
прокси-буфер LTV на случай (судебных расходов).
прокси-буфер LTV на случай (начисленных процентов).
прокси-буфер LTV на случай (неопределенности при реализации).

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

proxy_formal_framing(заключение о текущей рыночной стоимости).
proxy_formal_framing(заключение по состоянию на конкретную дату вступления в силу).
proxy_formal_framing(основанное на актуальных рыночных данных).

proxy_practical_interpretation(оценка рисков с учетом перспективного прогноза).
proxy_practical_interpretation(резерв на будущее обесценение).
proxy_practical_interpretation(исходные данные для оценки возможности взыскания залога в
→ будущем).

proxy_underlying_question(
    'Если данный заемщик в обозримом будущем не выполнит свои обязательства, останется ли
    → залог достаточным для покрытия непогашенной задолженности и убытков, связанных
    → с обращением взыскания?').

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

```

```

%%% =====
↪ =====

proxy_separation_party(оценщик, оценка_текущей_рыночной_стоимости).
proxy_separation_party(кредитор, прогнозный_запас_на_случай_будущего_обесценения).
proxy_separation_party(оценщик, «опорная величина текущей стоимости, основанная на
↪ доказательствах»).
proxy_separation_party(кредитор, «интерпретация будущего риска»).

пробел_между_доктриной_и_практикой_прокси(оценка_текущей_рыночной_стоимости,
неявный_прокси_будущей_стоимости).
пробел_между_доктриной_и_практикой_прокси(мнение_на_дату_вступления_в_силу,
прогноз_на_обозримый_период_держания).
разрыв_между_доктриной_и_практикой_прокси(основанная_на_доказательствах_текуща_
↪ я_базовая_стоимость,
буфер_будущего_риска).
разрыв_между_доктриной_и_практикой_прокси(статическая_одноточечная_оценка,
вероятностная_стоимость_во_времени).

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

proxy_historical_reason(спекулятивное_прогнозирование_способствовало_прошлым_кризис_
↪ ам).
proxy_historical_reason(суды_требовали_наглядных_доказательств).
прокси_историческая_причина(регулирующие_органы_стремились_к_стандартизированн_
↪ ым_обоснованным_практикам).

proxy_doctrine_emphasis(текущая_стоимость).
proxy_doctrine_emphasis(текущие_рыночные_данные).
proxy_doctrine_emphasis(определённые_даты_вступления_в_силу).
proxy_doctrine_emphasis(ограниченные_допущения).

прокси-процедура_практическое_применение(использование_кредиторами_в_вероятностны_
↪ х_моделях_оценки_будущих_рисков).
прокси-процедура_практическое_применение(использование_инвесторами_в_вероятностны_
↪ х_моделях_оценки_будущих_рисков).
прокси-процедура_практическое_применение(использование_аналитиками_обеспечения_в_
↪ вероятностных_моделях_будущих_рисков).

%%% =====
↪ =====
%%% 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(моделирование_капитала_по_Базелю).
proxy_modern_driver(крупномасштабное_стресс-тестирование_ипотечных_кредитов).

```



```

proxy_modern_treatment(вероятностный_подход_во_времени).
proxy_modern_treatment(распределительная_оценка_вместо_точечной).
proxy_modern_treatment(динамический_подход_вместо_статического).

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

ответственность_оценщика_по_доверенности(дисциплинированный_ориентир_на_текущую_
↪ _стоимость).
ответственность_оценщика_по_доверенности(оценка_на_основании_доказательств).
ответственность_оценщика_по_доверенности(заключение_о_текущей_рыночной_стоимост_
↪ _и).

proxy_lender_responsibility(future_risk_interpretation).
proxy_lender_responsibility(default_probability_estimation).
proxy_lender_responsibility(foreclosure_timing_estimation).
прокси_ответственность_кредитора(оценка_серьёзности_ликвидации).
прокси_ответственность_кредитора(анализ_нагрузки_на_портфель).
прокси_ответственность_кредитора(учёт_макроэкономических_ожиданий).

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

proxy_open_question(
    'Сохраняется ли в современных финансовых системах полная согласованность между
    ↪ дисциплинированной оценкой текущей стоимости и интерпретацией будущих рисков со
    ↪ стороны кредитора?«.).
proxy_open_question(
    'Делает ли современный вероятностный анализ залогового обеспечения во времени такое
    ↪ формальное/практическое разделение необоснованным в некоторых секторах?«.).

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

underwriting_component(1, текущая_рыночная_оценка).
underwriting_component(2, оценка_будущего_риска).
underwriting_component(3, вероятность_дефолта).
underwriting_component(4, сроки_взыскания_залога).
underwriting_component(5, степень_ущерба_при_ликвидации).
компонент_андеррайтинга(6, макроэкономические_прогнозы).
компонент_андеррайтинга(7, анализ_стрессовых_условий_портфеля).

ответственность_за_кредитование(текущая_рыночная_оценка, оценщик).
ответственность_за_кредитование(оценка_будущего_риска, кредитор).
ответственность_за_кредитование(вероятность_дефолта, кредитор).

```



```

ответственность_за_кредитование(сроки_взыскания_залога, кредитор).
ответственность_за_кредитование(серьёзность_ликвидации, кредитор).
ответственность_за_кредитование(макроэкономические_прогнозы, кредитор).
ответственность_за_кредитование(анализ_стрессовых_условий_портфеля, кредитор).

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

proxy_practical_implication(
    'Большинство оценщиков не осознают, что рыночная стоимость, через отношение кредита к
    ↳ стоимости, выступает в качестве приблизительного показателя будущей стоимости
    ↳ объекта при обращении взыскания.').
proxy_practical_implication(
    'Прогнозное использование рыночной стоимости кредитором редко становится очевидным в
    ↳ рамках задания на оценку.').
proxy_practical_implication(
    'Поэтому оценщики не подходят к выводам о рыночной стоимости с той осторожностью в
    ↳ отношении будущих тенденций стоимости, которую требует роль оценки залога.').

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(учитывать, как используется заключение о текущей стоимости).
proxy_recommended_practice(с осторожностью подходить к рыночной стоимости в контексте
    ↳ будущих тенденций).

proxy_further_work(устранение_разрыва_между_теорией_и_практикой).
proxy_further_work(подготовка_оценщиков_к_оценке_обеспечения).

```

6 Заключение

Пять отрывков, включенных в приведенные выше базы знаний, описывают одну профессию с пяти взаимодополняющих точек зрения. Рассматриваемый в отдельности, каждый отрывок представляет собой самостоятельное изложение: социальные и общественные интересы, которые они затрагивают, экономические функции, которые поддерживают оценки, их операционная роль в анализе сопутствующих рисков, нормативные границы, отделяющие их от смежных количественных исследований, а также противоречие между их формальной структурой и практическим применением. В совокупности, а также благодаря структуре, явно выраженной в Prolog, они описывают нечто более интересное — профессию, официальная доктрина и экономическая функция которой намеренно разграничены, а внутренняя согласованность которой всё чаще подвергается проверке с помощью аналитических инструментов, применяемых в настоящее время к её выводам.

Первые три базы знаний имеют общую схему: «effect(Quality, Horizon, Domain, Effect)». Каждая из них рассматривает качество оценки как причинный фактор, последствия которого распространяются на краткосрочную, среднесрочную и долгосрочную перспективы, различаясь лишь в том аспекте, через который эти последствия интерпретируются. Экономическая перспектива рассматривает их через призму кредитных рынков, распределения инвестиций, налогообложения и накопления капитала; социальная перспектива — через справедливость, доверие общественности и результаты для сообщества; перспектива обеспечения

— через призму андеррайтинга, управления портфелем и стабильности финансовой системы. Общая схема — не случайность. Она отражает существенное утверждение, проходящее через все три раздела: качество работы по оценке имеет последствия, эти последствия масштабируются во времени, и одни и те же базовые режимы сбоев проявляются по-разному в зависимости от того, чьи интересы выдвигаются на первый план. Общие элементы в аргументативной части предикатов — в частности, терминология, связанная с качеством и временным горизонтом, — позволяют объединить эти три терминологии, благодаря чему один и тот же сбой, такой как систематическая переоценка, можно проследить одновременно через экономическое неправильное ценообразование, социальное неравенство и неверную оценку залогового обеспечения.

Четвертая база знаний меняет плоскость рассмотрения. Она не описывает последствия качества оценки; она описывает, что такое оценка с методологической и правовой точек зрения. Ее предикаты носят таксономический, а не причинно-следственный характер, а ее механизм вывода отвечает на вопросы о юрисдикции, а не о воздействии. Этот сдвиг необходим, поскольку ориентированность первых трех разделов на последствия предполагает определение оцениваемой деятельности, а само это определение является предметом споров. Предикаты сферы деятельности фиксируют то, чем занимаются оценщики — основанные на доказательствах заключения о текущей стоимости с узко ограниченными ориентированными на будущее компонентами, — в то время как предикаты регулирования фиксируют условия, при которых эта деятельность относится к лицензируемой практике. Правила прямой цепочки (`requires_license/1`, `exempt_from_appraisal_license/1`, `regulated_appraisal/1`) предоставляют механизм для определения того, входит ли данная оценка в рамки регулируемой сферы или выходит за ее пределы — вопрос, на который, как предполагают первые три базы знаний, уже дан ответ.

Пятая база знаний выдвигает следующую тезу. В то время как остальные описывают, классифицируют или прослеживают последствия, база знаний о прокси-показателях утверждает, что тщательное доктринальное разграничение между оценкой текущей стоимости и прогнозированием будущей стоимости — разграничение, которое документирует база знаний о регулировании, — находится в противоречии с фактической экономической функцией оценок в ипотечном кредитовании. Норма отношения кредита к стоимости (LTV) в размере 80 %, семь конкретных рисков, от которых она защищает, зафиксированный дословно вопрос кредитора («If this borrower defaults sometime in the foreseeable future, is the collateral likely to remain sufficient . . . »), а также интегрированная семикомпонентная система андеррайтинга — всё это подтверждает одно и то же утверждение: оценка формально представляет собой заключение о текущей стоимости, а с экономической точки зрения — буфер будущей стоимости. Исторический компромисс, приведший к этой схеме — оценщики как дисциплинированные ориентиры текущей стоимости, кредиторы как центр интерпретации будущих рисков — был обоснованным ответом на условия, при которых возникла система лицензирования оценщиков, но современные факторы усиления, перечисленные в отрывке (AVM, системы оценки на основе машинного обучения, базельское моделирование капитала, аналитика секьюритизации, крупномасштабные стресс-тесты), рассматривают стоимость недвижимости с точки зрения вероятности во времени. Эти инструменты объединяют доктринально разделённые уровни в единый вычислительный конвейер, и в результате концептуальное разделение между уровнями становится всё более трудносохраняемым.

Возможно ли такое примирение и на каких условиях — вот вопрос, которым завершается данный отрывок. Базы знаний не дают на него ответа. Однако они предоставляют четкое, поддающееся анализу изложение соображений, имеющих отношение к его решению: последствия качества оценки в различных временных горизонтах и областях, юрисдикционные границы, определяющие, какие последствия подпадают под регулируемую практику, а также структурные причины, по которым эти границы в настоящее время находятся под давлением. Это изложение является фундаментом, на котором можно строить более

конкретные аргументы — о методологии, о реформе стандартов, о взаимосвязи между практикой оценки и более широкой системой количественной оценки.

Примечания по представлению знаний

Стоит особо отметить три структурных паттерна, поскольку они выявились в ходе процесса кодирования, а не на основе какого-либо отдельного фрагмента текста.

Во-первых, система баз знаний не является плоской. Три из пяти баз имеют общую схему; одна определяет юрисдикционные условия, при которых эта схема применяется; одна развивает тезис о разрыве между формальной структурой и экономической функцией. Связующий слой, объединяющий все пять баз, не рассматривает их как пять столбцов одной и той же матрицы. Он рассматривает первые три как описывающие последствия, четвертую — как ограничивающую сферу действия этих описаний, а пятую — как определяющую, где само ограничение подвергается давлению. Эта асимметричная структура является скорее особенностью, чем недостатком: она отражает реальную организацию профессиональной практики, в которой описательные утверждения, утверждения о юрисдикции и критические утверждения занимают действительно разные эпистемические роли.

Во-вторых, префикс «discipline», который различает модули (economic_, social_, collateral_, scope_/ regulation_, proxy_ / underwriting_) играет реальную роль только в тех случаях, когда существует риск коллизии. Предикаты, имеющие прямые аналоги в других модулях, заслуживают префикса; предикаты, уникальные для одной концептуальной области, — нет. Общий словарь значений со стороны аргументов — атомы качества, горизонта и области — именно и позволяет модулям взаимодействовать, а добавление префиксов к этим значениям лишило бы это взаимодействие смысла. Сформировавшийся паттерн носит общий характер: префиксы для предикатов, общие атомы для аргументов, с выборочным, а не универсальным префиксированием в тех случаях, когда предикат является уникальным для одной линзы.

В-третьих, центральное противоречие, выявленное в базе знаний о прокси, тесно связано с регуляторной асимметрией, выявленной в базе знаний о сфере применения/регулировании, но отличается от неё. Эти два явления фиксируются с помощью параллельных двухместных предикатов — «regulatory_asymmetry/2» и «proxy_doctrine_practice_gap/2» — с намеренно разными названиями, поскольку они описывают разные оси одной и той же базовой структуры. Регуляторная асимметрия касается того, кто подлежит регулированию и насколько строго; разрыв между доктриной и практикой касается того, что формально говорится о продукте оценщика, по сравнению с тем, как с ним поступают в экономическом плане. Эти два противоречия усиливают друг друга. Регулирование подталкивает оценщиков к доктрине текущей стоимости; экономическая функция тянет их выводы в сторону использования будущей стоимости. В совокупности складывается картина профессии, официальное самописание которой становится всё более трудно совместимым с аналитическими потоками, в которые попадают её продукты.

Список литературы

- Agarwal, Sumit, Itzhak Ben-David и Vincent Yao (2015). «Collateral Valuation and Borrower Financial Constraints: Evidence from the Residential Real Estate Market». В: Management Science 61.9, с. 2220—2240. DOI: [10.1287/mnsc.2014.2002](https://doi.org/10.1287/mnsc.2014.2002).
- Alexandrov, Alexei, Laurie Goodman и Michael Neal (январь 2023). Reengineering the Appraisal Process: Better Leveraging Both Automated Valuation Models and Manual Appraisals. Тек. отч. 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 (март 2021). Dodd-Frank Act Stress Test 2021: Supervisory Stress Test Methodology. Tex. отч. 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 и Parag Pathak (2011). «Forced Sales and House Prices». B: American Economic Review 101.5, с. 2108—2131. DOI: [10.1257/aer.101.5.2108](https://doi.org/10.1257/aer.101.5.2108).
- Clauretie, Terrence M. и Nasser Daneshvary (2009). «Estimating the House Foreclosure Discount Corrected for Spatial Price Interdependence and Endogeneity of Marketing Time». B: Real Estate Economics 37.1, с. 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 и Patrick M. McAllister (2010). «Means, Motive and Opportunity? Disentangling Client Influence on Performance Measurement Appraisals». B: Journal of Property Research 27.2, с. 181—201. DOI: [10.1080/09599916.2010.499014](https://doi.org/10.1080/09599916.2010.499014).
- Deng, Yongheng, John M. Quigley и Robert Van Order (2000). «Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options». B: Econometrica 68.2, с. 275—307. DOI: [10.1111/1468-0262.00110](https://doi.org/10.1111/1468-0262.00110).
- Diaz Julian, III и Marvin L. Wolverton (1998). «A Longitudinal Examination of the Appraisal Smoothing Hypothesis». B: Real Estate Economics 26.2, с. 349—358. DOI: [10.1111/1540-6229.00749](https://doi.org/10.1111/1540-6229.00749).
- Ebrahim, M. Shahid и Sikandar Hussain (2010). «Financial Development and Asset Valuation: The Special Case of Real Estate». B: Journal of Banking & Finance 34.1, с. 150—162. DOI: [10.1016/j.jbankfin.2009.07.011](https://doi.org/10.1016/j.jbankfin.2009.07.011).
- Eriksen, Michael D. и др. (2019). «The Influence of Contract Prices and Relationships on Appraisal Bias». B: Journal of Urban Economics 111, с. 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 и Paul S. Willen (2008). «Negative Equity and Foreclosure: Theory and Evidence». B: Journal of Urban Economics 64.2, с. 234—245. DOI: [10.1016/j.jue.2008.07.006](https://doi.org/10.1016/j.jue.2008.07.006).
- Foster, Chester и Robert Van Order (1984). «An Option-Based Model of Mortgage Default». B: Housing Finance Review 3.4, с. 351—372.
- Friedman, Jerome H. (1991). «Multivariate Adaptive Regression Splines». B: The Annals of Statistics 19.1, с. 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». B: Journal of Property Research 13.4, с. 261—273. DOI: [10.1080/095999196368781](https://doi.org/10.1080/095999196368781).
- Geltner, David M. (1991). «Smoothing in Appraisal-Based Returns». B: The Journal of Real Estate Finance and Economics 4.3, с. 327—345. DOI: [10.1007/BF00161933](https://doi.org/10.1007/BF00161933).
- Ghent, Andra C., Walter N. Torous и Rossen I. Valkanov (2019). «Commercial Real Estate as an Asset Class». B: Annual Review of Financial Economics 11, с. 153—171. DOI: [10.1146/annurev-financial-110118-123121](https://doi.org/10.1146/annurev-financial-110118-123121).
- Griffin, John M. и Gonzalo Maturana (2016). «Who Facilitated Misreporting in Securitized Loans?» B: The Review of Financial Studies 29.2. Evidence of appraisal inflation and its consequences, с. 384—419. DOI: [10.1093/rfs/hhv130](https://doi.org/10.1093/rfs/hhv130).
- Hendershott, Patric H. и Edward J. Kane (июль 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?» B: Real Estate Economics 23.2, с. 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. 10-е изд. 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. и James E. Larsen (2024). «A Primer on Regulations and the Practice of Residential Property Appraisal». B: 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. и др. (1995). «The Valuation at Origination of Fixed-Rate Mortgages with Default and Prepayment». B: The Journal of Real Estate Finance and Economics 11.1, с. 5—36. DOI: [10.1007/BF01097934](https://doi.org/10.1007/BF01097934).
- LaCour-Little, Michael и Stephen Malpezzi (2003). «Appraisal Quality and Residential Mortgage Default: Evidence from Alaska». B: The Journal of Real Estate Finance and Economics 27.2, с. 211—233. DOI: [10.1023/A:1024728420837](https://doi.org/10.1023/A:1024728420837).
- Lee, Ming-Hsiang, Chien-Wen Peng и Hui-Fen Liao (2020). «An Analysis of Objectivity in the Real Estate Appraisal Process». B: International Real Estate Review 23.4, с. 483—504.
- Mayer, Yanling G. и Frank E. Nothaft (2022). «Appraisal Overvaluation: Evidence of Price Adjustment Bias in Sales Comparisons». B: Real Estate Economics 50.3, с. 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». B: Journal of Real Estate Research 28.2, с. 193—214. DOI: [10.1080/10835547.2006.12091177](https://doi.org/10.1080/10835547.2006.12091177).
- Qi, Min и Xiaolong Yang (2009). «Loss Given Default of High Loan-to-Value Residential Mortgages». B: Journal of Banking & Finance 33.5, с. 788—799. DOI: [10.1016/j.jbankfin.2008.09.010](https://doi.org/10.1016/j.jbankfin.2008.09.010).
- Quan, Daniel C. и John M. Quigley (1991). «Price Formation and the Appraisal Function in Real Estate Markets». B: The Journal of Real Estate Finance and Economics 4.2, с. 127—146. DOI: [10.1007/BF00173120](https://doi.org/10.1007/BF00173120).
- Steurer, Miriam, Robert J. Hill и Norbert Pfeifer (2021). «Metrics for Evaluating the Performance of Machine Learning Based Automated Valuation Models». B: Journal of Property Research 38.2, с. 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%.

Метод остаточных ограничений (RCA): концепция и протокол

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

William Bert Craytor

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

DOI: [10.67330/kgebx46](https://doi.org/10.67330/kgebx46)

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

Эта статья открытого доступа распространяется на условиях [лицензии Creative Commons Attribution 4.0](#)

Аннотация. Существенное ограничение традиционных методов оценки — опора на устаревшие ручные приёмы, такие как анализ парных продаж, которые обходят стороной вклады характеристик объекта в стоимость: в традиционной таблице продаж нет столбца для вкладов в стоимость, есть лишь корректировки. Между тем именно вклады в стоимость дают необходимую основу для математических ограничений, не позволяющих оценщику завысить или занижить стоимость. Метод остаточных ограничений (RCA) развивает традиционный сравнительный подход, объединяя многоэтапный процесс оценки на основе многомерных адаптивных регрессионных сплайнов (MARS) со строгими математическими ограничениями на вклады характеристик — прежде всего неизмеряемых характеристик, то есть латентных переменных вроде эстетики, планировки, состояния и качества, стоимость которых традиционно определяется субъективным суждением оценщика.

Ключевые слова: оценка недвижимости, оценка жилья, метод остаточных ограничений, MARS, сравнительный подход, анализ остатков, латентные переменные, автоматизированная оценка

Содержание

1	Введение	3
2	Сведения об авторе	6
3	Инженер по оценке	9
4	Допущения	10
5	Рабочий процесс RCA	13
5.1	Виды анализа MARS	13
5.2	Рабочий процесс обработки	13
5.3	Рабочий процесс обработки данных	15
5.3.1	Требования к данным для MARS	15
5.3.2	Протоколы: Список закрытых сделок	17

5.4	Структура папок проекта	19
5.4.1	Протоколы: Несколько прогонов MARS (для рынка недвижимости) . .	20
5.5	Основная книга Excel «	22
5.6	Резервное копирование	22
6	Обозначения книг	23
6.1	Описательные измерения	23
6.1.1	Размер проекта (“a”)	24
6.1.2	Параметр версии (“r”)	25
6.1.3	Этапы рабочего процесса (“w”)	25
6.1.4	Листы книги рабочей Excel (“s”)	26
6.1.5	Список объектов недвижимости (“p”) — строки MLSData	26
6.1.6	Переменные MLS (“v”) Столбцы MLSData	27
6.1.7	Пример окончательной нотации	27
6.1.8	Обоснование	27
6.1.9	Матрицы?	28
6.1.10	Первое упрощение	28
6.1.11	Второе упрощение	28
7	Таблица продаж RCA	31
8	Переменные	35
8.1	Онтология «Variable Origin»	35
8.2	Онтология переменных приложений	36
8.3	Члены второй и третьей степени	38
8.3.1	Переменные корректировки	38
8.3.2	Агрегирующие переменные	38
8.3.3	Специальные переменные& Обозначения	39
8.3.4	Разбивка остатков	39
9	Основные доказательства метода RCA	41
10	Модель MARS «	43
10.1	Функции «g» для моделирования переменных и взаимодействий	43
10.2	Индексирование функций g	43
10.3	Размеры графика	44
10.4	Окончательное уравнение модели	45
10.5	Количество диаграмм, необходимых для построения графика модели	45

11	Характеристики данных	48
11.1	Цель программы «Residual»	48
11.2	Оценки состояния и качества в качестве приблизительных регрессоров	49
11.3	Ошибки в данных	49
11.4	Что такое точность?	51
11.5	Рассматриваемый объект оценки как аномалия	53
11.6	Важность CVR2 при работе с непредвиденными данными	53
12	Остатки	55
12.1	Рейтинг & Балл	55
12.2	Кривая CQA-остаточная величина	58
12.3	Характеристики кривой CQA	58
12.3.1	Анализ компонентов кривой CQA	59
12.3.2	Динамика рынка и социальные последствия	59
12.4	Объект оценки «Остаточные величины»	61
12.5	Выводы по оценке RCA	61
12.5.1	Общие соображения	61
12.5.2	Обзор, аудит, жалобы	62
13	Заключение	63

1 Введение

В данной статье представлены логические и математические основы современного подхода к точной оценке рыночной стоимости жилых и иных объектов недвижимости. Этот метод, разработанный на основе 20-летнего опыта применения программного обеспечения «Многомерные адаптивные регрессионные сплайны» (MARS) к жилым домам Северной Калифорнии — особенно в районах залива Сан-Франциско — обеспечивает значительное повышение точности оценки. MARS повышает качество оценки за счет умелого учета нелинейных взаимосвязей между характеристиками объекта недвижимости и его стоимостью, одновременно эффективно выявляя и приоритезируя наиболее значимые переменные. Эта возможность позволяет оценщикам выявлять динамику стоимости в новых рыночных зонах гораздо быстрее, чем это позволяют традиционные методы. MARS составляет основу предлагаемого Метода остаточных ограничений (RCA). Метод RCA получил своё название благодаря строгой методологической основе:

- Он налагает строгие ограничения на вклад в стоимость как измеримых, так и неизмеримых характеристик объекта недвижимости:
 1. Сумма всех вкладов в стоимость, полученных с помощью MARS, включая остаток, должна равняться чистой цене продажи каждого аналога.
 2. Остаток можно разложить на подкомпоненты с заданными значениями при условии, что их сумма равна остатку аналога.

- Метод RCA минимизирует систематическую ошибку при оценке субъективных характеристик — таких как эстетическая привлекательность, состояние и качество — путем привязки их совокупной стоимости к остатку, который точно количественно определяется при построении модели MARS компетентным инженером-оценщиком.

Хотя метод RCA в первую очередь разработан для жилой недвижимости, где субъективные характеристики существенно влияют на рыночную стоимость, он обладает потенциалом для более широкого применения в области оценки активов. Точные оценки имеют решающее значение, учитывая, что рынок жилой недвижимости США оценивается в 45 триллионов. Fannie Mae отмечает, что типичные оценки отклоняются примерно на 5% от фактических цен продажи, хотя эта цифра скрывает систематическую ошибку, поскольку оценщики часто сталкиваются с давлением, вынуждающим их привести оценку в соответствие с ценами продажи. Кроме того, поскольку 60% — 70% оценок проводятся для целей рефинансирования — при отсутствии аналогов по ценам продажи — необходимость в точности, которую обеспечивает RCA, становится очевидной. Согласно анализу Федеральной резервной системы 2012 года, 15% — 20% оценок, проведенных до кризиса 2008 года, отклонялись от цен продажи более чем на $\pm 10\%$; этот диапазон, вероятно, по-прежнему актуален для оценок при рефинансировании сегодня, причем он варьируется в зависимости от типа недвижимости и рынка. Почему важна точность? В случае сделок с низким отношением кредита к стоимости недвижимости (LTV) — например, когда покупатель с хорошей кредитной историей вносит первоначальный взнос в размере \$ 500 000 на дом стоимостью \$ 1 миллион — кредиторы могут отказаться от оценки, исходя из минимального риска обращения взыскания. Однако покупатели, которые обходятся без оценки, рискуют переплатить, обнаружив дефекты только после покупки. Это подчеркивает ценность надежных методов оценки, таких как RCA. Количественная оценка точности оценки по-прежнему представляет собой сложную задачу. Исследователи не располагают прямыми показателями отклонений от «рыночной стоимости» и вынуждены полагаться на косвенные показатели. Традиционные методы оценки, основанные на субъективном суждении и различном уровне компетентности оценщиков, затрудняют анализ. Стратегии включают:

1. Tzioumis (2016) В ходе исследования, охватившего 4 782 оценщиков и 259 436 оценок в 11 штатах, было установлено, что одна треть оценок находилась в пределах 1% от договорной цены, половина — в диапазоне от 1% до 3% выше, а небольшая часть в среднем на 5% выше — что свидетельствует о тенденции к завышению. В ходе исследования были выделены два типа оценщиков: «по рыночной стоимости» (близко к договорной цене) и «выше рыночной стоимости» (постоянно выше).
2. Krivorotov и LaCour-Little (2020) Было установлено, что 90% оценок при покупке соответствовали договорным ценам или превышали их, что подчеркивает «анкерный уклон» — тенденцию к привязке к желаемой стоимости — особенно в контексте рефинансирования.

Эти исследования выявляют предвзятость в практике оценки, усугубляемую ограничениями традиционных инструментов. Если бы оценщики были одинаково компетентны и беспристрастны, их оценки совпадали бы; постоянная завышенная оценка со стороны некоторых групп свидетельствует о системных проблемах.

Приведенные выше исследования, посвященные систематическим отклонениям, а также системная закономерность, на которую они указывают, отражают более глубокую особенность традиционной практики сравнения продаж, для устранения которой и была разработана методика RCA. Традиционная методология опирается на два канала с совершенно разным эпистемическим статусом. В канале измеримых характеристик — общая арендуемая площадь (GLA), количество комнат, площадь участка, возраст, местоположение — от оценщика

ождается, что он будет выводить значения корректировок объективно, посредством анализа аналог пар или локальной линейной регрессии на основе аналогов продаж. Компетентные оценщики выполняют эту работу, и полученные в результате корректировки, по крайней мере, в общих чертах обоснованы с точки зрения данных. Однако даже в этом случае традиционная таблица фиксирует только корректировки; в ней даже нет столбца для вклада в стоимость. Традиционные корректировки также не рассчитываются как разности между вкладами в стоимость — они оцениваются непосредственно на основе сопоставленных пар или аналогичных процедур, без какого-либо расчета базовых вкладов. В RCA, напротив, каждая корректировка представляет собой разницу между вкладами в стоимость объекта оценки и аналога по данной характеристике — той структуре, последствия которой уточняет Доказательство II. Что касается канала субъективных характеристик — состояние, качество строительства, дизайн, функциональная полезность, эстетика — доступные вспомогательные процедуры варьируются от скудных до совершенно недостаточных. Корректировки по субъективным характеристикам могут подкрепляться анализом сопоставленных пар объектов недвижимости, выходящих за пределы непосредственного набора аналогов, ссылкой на признанную практику коллег или ссылкой на документально подтвержденный опыт оценщика на рынке. Все три метода формально соответствуют профессиональным стандартам. Ни один из них не достигает той статистической строгости, которую измеримый канал обеспечивает в ходе процедуры.

В результате возникает процедурная асимметрия между тем, что требуют стандарты, и тем, что может обеспечить данная методология. Стандарты требуют обоснования каждой корректировки, будь то измеримой или субъективной. Методология обеспечивает надежное обоснование в одном случае и целый спектр менее убедительных процедур в другом, начиная от единичного примера сопоставленной пары и заканчивая ссылкой на то, что обычно используют другие оценщики. Честный и компетентный оценщик, тщательно проводящий работу с сопоставленными парами при корректировке с видом, получает значительно более весомое обоснование, чем тот, кто без обоснования ссылается на практику коллег, однако эксперт, сравнивающий две оценки по стандартной форме, не может легко различить их, поскольку в форме фиксируется значение корректировки, а не строгость её вывода. Данная методология допускает такую асимметрию, поскольку более строгих процедур для оценки субъективных характеристик у практикующих оценщиков в целом нет.

Метод RCA улучшает оба канала, но совершенно разными способами. В измеримом канале регрессия MARS учитывает нелинейные зависимости, взаимодействия переменных и структуру базисных функций, которые анализ сопоставленных пар и обычная линейная регрессия не могут отразить. Благодаря построению модели на основе полного набора квалифицированных продаж в рыночной зоне — как правило, нескольких сотен — а не на трех – девяти аналогах традиционной таблицы, вклады в стоимость, полученные с помощью MARS, являются объективными и существенно более точными, чем методология, которую они заменяют. В субъективном канале RCA действует совершенно иначе, чем любые традиционные вспомогательные процедуры. Она вообще не пытается оценивать коэффициенты корректировки для отдельных субъективных характеристик. Вклады всех субъективных характеристик — латентных переменных — в совокупности проявляются в виде остатка: разницы между фактической чистой ценой продажи каждого аналога и его ценой, прогнозируемой по методу MARS, которая рассчитывается на основе данных без участия оценщика. Разложение этого остатка на названные компоненты — дизайн, функциональная полезность, вид, состояние, качество и т. д. — ограничено тем, что их сумма должна равняться общему остатку, и согласно доказательству II (раздел 9) любое разложение, удовлетворяющее этому ограничению, приводит к одному и тому же выводу о стоимости. Единственное суждение, которое остается в рабочем процессе, — это определение места объекта оценки в ранжированном распределении остаточных значений аналогов,

осуществляемое на основе поддающегося проверке ранжирования и явно задокументированное в отчёте. Все еще возможна слабая оценка по методу RCA — неопытный оценщик может создать переобученную модель MARS, разместить объект оценки неправдоподобно или пропустить аномалии в данных — но эта слабость проявляется в аспектах, которые рецензент может выявить путем проверки модели и ранжирования, а не в аспектах, которые невозможно выявить при стандартной рецензии.

Помимо рефинансирования, оценщики обязаны предоставлять владельцам точные оценки, особенно в случае дорогостоящих покупок, где ошибки могут обойтись в десятки или сотни тысяч. Тем не менее многие не хотят тратить \$ 600–\$ 1,000 на оценку, сомневаясь в её полезности. Передовые методы, такие как RCA, хотя и требуют значительных ресурсов на начальном этапе, обещают более высокую точность. С совершенствованием методов и накоплением опыта их эффективность — и распространение — должны возрасти.

2 Сведения об авторе

Автор имеет образование в области математики, информатики и статистики. По профессии он является инженером-программистом, имеющим дополнительный опыт в области оценки и бухгалтерского учета. Он занимается оценкой с перерывами с 2002 года и с 2004 года применяет систему MARS для определения корректировок стоимости и ценовых тенденций на рынке жилой недвижимости.

Оценка с использованием MARS — это «глубокий опыт» в области оценки. MARS генерирует модели, которые более детально, чем любой другой статистический метод, отражают как переменные, влияющие на рыночную стоимость, так и то, как изменения этих значений сказываются на цене продажи. Таким образом, при оценке с использованием MARS оценщик постоянно исследует причинно-следственные связи между значениями переменных и ценами продажи, генерируемыми моделями MARS.

В целях обеспечения прозрачности методология RCA наиболее эффективно применяется к жилым районам, обладающим следующими характеристиками:

1. Сложные районы, в которых объекты недвижимости значительно отличаются друг от друга
 - Старые дома
 - Различный уровень модернизации
 - Многообразие архитектурных стилей и дизайнов
 - Различия в характеристиках объектов недвижимости, таких как общая арендуемая площадь (GLA), площадь участка, высота над уровнем моря и вид из окна.
2. Сложные рыночные условия
 - Быстро меняющиеся рыночные условия
 - Множество различных типов покупателей
3. Надежные источники данных: протокол RCA основан на методах интеллектуального анализа данных и требует большого объема качественных данных. Хотя его можно использовать в разреженных регионах, затраты времени и средств могут оказаться неоправданными.

4. Желательно иметь точные оценки рыночной стоимости. В большинстве случаев RCA должен обеспечивать точность на уровне $\pm 1-2\%$. В отчёте инженера по оценке должны быть указаны условия, необходимые для обеспечения точности итогового заключения о стоимости. Если данных или информации не хватает, либо если в ходе интеллектуального анализа данных не удалось выявить устойчивых закономерностей на рынке, точность оценки снизится, и об этом следует проинформировать клиента. В частности, высокие значения R2 и CVR2 (перекрёстная проверка R2, рассчитанные путем многократного разделения набора данных MLS на случайные обучающие и тестовые подмножества), значимая модель MARS и выраженная закономерность в характеристиках CQA (состояние-качество-привлекательность) объектов недвижимости в ранжированных остатках являются основными средствами оценки точности и надежности вывода о стоимости.

Разработка этих протоколов RCA на протяжении последних более чем 20 лет основана на доступности данных MLSListings Inc. (Саннивейл, Калифорния) и других систем MLS, которые в настоящее время доступны через MLSListings и составляют более 95% опыта автора в области анализа данных. Опыт автора основан на обширной работе по оценке недвижимости в 15 основных округах Северной Калифорнии, а также на дополнительном опыте работы в 10 второстепенных округах.

Методология RCA наиболее эффективна на рынках, отвечающих двум ключевым критериям:

1. Широкая доступность высококачественных данных для оценки
2. Сложные, неоднородные жилые районы, характеризующиеся:
 - Разнообразием архитектурных стилей
 - Различной степенью обновления недвижимости
 - Многолетними тенденциями развития

Опыт округа Калифорнии

Утверждения, приведенные в данной статье, основаны на опыте оценки объектов недвижимости в следующих округах Северной Калифорнии:

Сан-Матео	Санта-Клара	Марин	Напа	Монтерей	Санта-Роза
Аламеда	Сан-Франциско	Контра-Коста	Сакраменто	Эль-Дорадо	Сан-Хоакин
Станислаус	Плейсер	Солано	Мендосино	Лейк	Колуса
Мерсед	Невада	Сан-Бенито	Санта-Круз	Саттер	Юба

Указанные выше районы, по сути, охватывают Кремниевую долину и прилегающие к ней округа, простираясь далее в более сельские районы, включая Центральную долину и предгорья Сьерры. Следует отметить, что автор всегда имел доступ к обширным и относительно точным данным по сравнению с тем, что имеется в распоряжении многих оценщиков в других, менее развитых регионах США.

Используемые данные MLS имеют сертификат RESO и соответствуют стандарту RCF (RESO Common Format). Несмотря на сертификацию RCF, данные, имеющие, как правило, второстепенное значение, по некоторым объектам недвижимости иногда отсутствуют или содержат неточности, что может быть вызвано ошибками при вводе данных в систему MLS либо первоначальным налоговым оценщиком, либо последующим агентом по продажам.

Часто эти ошибки следуют определённой схеме, например, просто не производится ввод данных, если данного параметра не существует. Например, вместо ввода 0 для «мест под навесом», когда навеса нет, данные не вводятся вовсе. В районах, где навесы для автомобилей встречаются редко, аналитик может заменить отсутствующие данные значением 0, чтобы получить полезную корректировку на наличие навесов. Хотя R/earth и другие реализации MARS могут обрабатывать отсутствующие данные, решение о том, как с ними поступать, остается за инженером по оценке (VE).

Данные MLSListings также доступны из систем MLS некоторых других штатов — от Флориды до Гавайев. Это может быть полезно для проверки эффективности работы протоколов и кода в разных штатах.

3 Инженер по оценке

Definition 1: Valuation Engineer

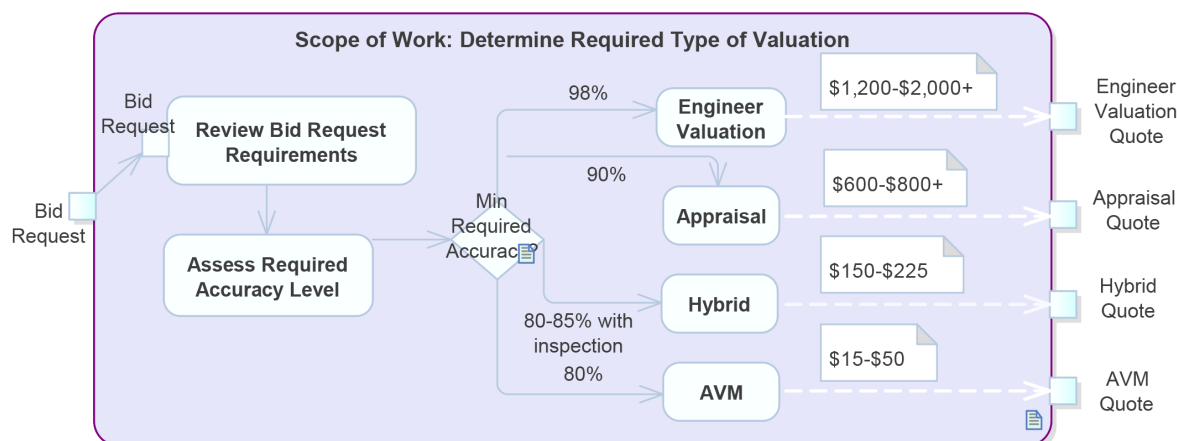
Во избежание двусмысленности, термин «инженер по оценке» (или сокращённо «ВЕ») используется для обозначения специалиста, обладающего квалификацией лицензированного оценщика, а также передовыми знаниями, навыками и опытом в области использования методов интеллектуального анализа данных, статистики, программирования, детальных протоколов и искусственного интеллекта для определения рыночной стоимости сложных объектов недвижимости. Также важно понимать, что оценка — это вид деятельности, требующий наличия той или иной формы лицензии в связи с существующим риском мошенничества.

Кредиторы, как правило, предлагают несколько вариантов оценки:

1. Автоматизированные модели оценки (AVM) — самый экономичный вариант
2. Гибридные оценки — с привлечением нелицензированных инспекторов под контролем лицензированных оценщиков
3. Стандартные оценки — выполняются лицензированными оценщиками

При проведении оценочных работ обычно не учитываются оценки, выполняемые специалистами инженерного уровня. Просто на рынке не так много оценщиков, обладающих необходимыми навыками для их проведения. В них, безусловно, должна быть потребность. При использовании методологии RCA (MARS и ограничительные условия) они могут обеспечить значительно более высокую точность и практически гарантировать достаточно высокую степень объективности.

Для объектов недвижимости, соответствующих нормам, расположенных в новых жилых районах, как правило, достаточно стандартной оценки. Однако для старых домов в районах, где в течение более 30 лет проводились значительные ремонтные работы, рекомендуется проводить инженерную оценку. Выбор между этими методами должен осуществляться в рамках структурированного процесса принятия решений с учетом характеристик объекта недвижимости и рыночных условий.



4 Допущения

Assumption 1: Geographical Area & Property Type

В данной работе основное внимание уделяется односемейным домам и квартирам в кондоминиумах, расположенным в мегаполисах и городских районах Калифорнии. Хотя сравнительный подход применим к другим типам недвижимости и географическим регионам, доходный и затратный подходы, как правило, играют более важную роль при оценке сельскохозяйственных земель, многоквартирных домов и коммерческой недвижимости. В таких случаях метод сравнительных продаж может потребовать модификации или иметь ограниченную применимость.

Assumption 2: Data Storage

Форматы CSV и Excel широко используются в системах MLS и считаются основным способом передачи данных в программы RCA. Хотя для непрерывной передачи данных в базы данных SQL можно использовать брокерские каналы. Для резервного копирования и промежуточного хранения часто используются базы данных SQL, и две из них, играющие важную роль, — это бесплатные системы с открытым исходным кодом SQLite и PostgreSQL. Текстовые или RTF-файлы также используются для фрагментов отчетов, а файлы PNG, JPG и TIFF — для графики. Обратите внимание, что многие также используют SQL Server от Microsoft, однако бесплатная версия «SQL Server Express» имеет значительные ограничения (например, можно использовать только 1 ядро процессора), а другие версии требуют уплаты значительных лицензионных сборов при коммерческом использовании. Базы данных по недвижимости в Калифорнии могут быть очень большими, а операции соединения (join) — весьма трудоемкими, поэтому для выполнения крупных соединений может пригодиться компьютер с 8 и более ядрами и 128 ГБ оперативной памяти. Базы данных SQL также важны, когда необходимо найти все объекты недвижимости в произвольно ограниченной области. Таким образом, если, например, вы использовали кластерный анализ для определения районов, вам, скорее всего, понадобится какая-либо база данных SQL. Однако разные базы данных имеют разные встроенные функции для работы с географической информацией и, следовательно, несовместимы друг с другом. В этой серии статей, посвященных RCA, для сложных задач SQL будет использоваться PostgreSQL, а для простых — SQLite.

Assumption 3: “R/earth” Is Used For MARS processing.

До 2017 года программное обеспечение MARS от Salford Systems было доступным решением по цене \$ 270 в год. Однако после приобретения Salford Systems компанией Minitab MARS был включен в их набор SPM по цене \$ 16 000 в год — цена, недоступная для большинства малых предприятий. К счастью, пакет «earth» для языка программирования R ((Milborrow 2024a)), поддерживаемый Стивеном Милбороу, развился настолько, что теперь соответствует или даже превосходит возможности SPM MARS от Minitab. Пакет earth предлагает дополнительные преимущества: он доступен бесплатно, органично интегрируется с обширной статистической экосистемой R и доступен через Python с помощью библиотеки Pandas, что обеспечивает отличную гибкость для различных рабочих процессов. В обсуждениях, касающихся MARS, предполагается конкретное использование R/earth.

Assumption 4: Notation

Исходные данные из внешних источников, таких как файлы CSV и таблицы Excel, представляются с использованием стандартных латинских символов. Сюда входят данные о недвижимости из службы Multiple Listing Service (MLS), настройки MARS (Multivariate Adaptive Regression Splines) с использованием пакета R/earth, а также данные, специфичные для конкретных проектов. Эта нотация распространяется на соответствующие представления в виде фреймов данных в R или Python.

При обсуждении характеристик объектов недвижимости, а также их различных типов и подмножеств, для повышения точности и аналитической ясности используется нотация на основе греческих букв. Хотя такая математическая нотация может показаться более строгой, она выполняет важную функцию: позволяет отличать абстрактные характеристики объектов от их представлений в виде необработанных данных. По соглашению греческие буквы последовательно обозначают характеристики объектов недвижимости, а латинские — структуры данных в электронных таблицах. Например, символ M представляет матрицу данных MLS, в которой объекты недвижимости расположены в строках, а соответствующие им характеристики — в столбцах.

Assumption 5: Programming Languages

В этой области доминирующими языками являются R и Python с библиотекой Pandas, а языки C или C++ используются в тех случаях, когда важна скорость обработки. Сам R/earth написан на R и C.

Assumption 6: Indexing

Индексация может различаться в разных языках программирования и математике, что может сбивать с толку:

- В математике и R индексация начинается с 1.
- В C, C#, C++ и Python — с 0.

В данной статье, в целях единообразия и простоты, индексация везде начинается с 0 (как в Python или C). Вот как это работает:

- Для списков, матриц или дата-фреймов:
 - Индекс 0 зарезервирован для «объекта оценки» (основного свойства, которое изучается).
 - Пример: в таблице, где строки представляют объекты недвижимости, строка 0 — это объект оценки, а столбец 0 содержит уникальные идентификаторы всех объектов недвижимости.
 - Практический совет: располагайте объекты недвижимости так, чтобы сначала шло объект оценки (строка 0), а затем перечисляйте остальные от самой новой к самой старой дате продажи, пронумеровав их последовательно (0, 1, 2 и т. д.).
- Если «объекта оценки» нет:
 - Если структура данных не содержит объектов недвижимости (например, только переменные или функции), индекс 0 пропускается, и индексирование начинается с 1.
 - Особое примечание для пользователей R: поскольку в R используется индексация, начинающаяся с 1, при написании кода для математических формул из данной статьи в R необходимо прибавить 1 ко всем индексам, чтобы перейти от индексации, начинающейся с 0, к индексации, начинающейся с 1.

5 Рабочий процесс RCA

Понимание рабочего процесса RCA имеет решающее значение для эффективного анализа. Этот процесс состоит из структурированной последовательности аналитических задач и этапов принятия решений

5.1 Виды анализа MARS

Рабочий процесс анализа MARS зависит от того, требуется ли только рыночный анализ или же необходима оценка стоимости конкретного объекта недвижимости. Следует отметить, что для проведения оценки конкретного объекта недвижимости, как правило, требуется анализ его рыночной зоны.

- Рыночный анализ: фокусируется на обширных территориях, таких как округа, города или определенные регионы. Данный подход не затрагивает конкретный объект оценки и, следовательно, не требует вывода о его стоимости. Рабочий процесс сосредоточен на разработке модели, её валидации и формировании вспомогательной аналитики. Процесс завершается после определения и документирования оптимальной модели.
- Анализ отдельного объекта недвижимости включает в себя все элементы рыночного анализа, а также определение стоимости конкретного объекта. Это принципиально отличается от традиционных методов оценки. В традиционных подходах стоимость обычно определяется путем корректировок 3–9 аналогов сделок купли-продажи и расчета средневзвешенного значения. В отличие от этого, RCA определяет стоимость в основном на основе двух компонентов:

1. Прогнозируемая стоимость по модели MARS
2. Расчетная остаточная стоимость конкретного объекта недвижимости (с соответствующим баллом CQA)

Таблица продаж в RCA служит в первую очередь для объяснения различий в стоимости между объектом оценки и аналогами путем иллюстрации вклада отдельных переменных. Ключевым преимуществом RCA является ее стабильность: хотя клиенты могут спорить по поводу выбора аналогов, вывод о стоимости остается неизменным независимо от того, какие аналоги выбраны для представления, поскольку все сделки (будь то 15 или 600) подвергаются корректировкам с использованием одной и той же модели. Хотя более старые сделки автоматически подвергаются корректировкам с учетом текущих рыночных условий, в целом предпочтительны недавние сделки. Эффективность модели в учете ценовых факторов снижается с увеличением временного интервала, поскольку не измеренные рыночные характеристики приобретают все большее значение при более длительных периодах.

5.2 Рабочий процесс обработки

Общий рабочий процесс обработки можно описать следующими этапами:

1. Объем работ.
2. Настройка исходных папок и файлов проекта.
3. Загрузка данных MLS. Информацию о количестве необходимых записей об объектах недвижимости и установке ограничений по датам см. в протоколах 1.1–1.3.

4. Заполните листы настроек: Регрессионные переменные, Взаимодействия, Агрегации, Расчёты.
5. Запустите регрессию MARS (R/earth) для построения ценовой модели.
6. Сгенерируйте вклады в стоимость объекта недвижимости и остатки (чистый остаток и остаток/SF) на основе модели MARS.
7. Проведите ранжирование объектов недвижимости по показателю «остатки/SF».
8. Создайте балы CQA для каждого аналога сделки купли-продажи.
9. Предполагая, что был составлен рейтинг качества объектов недвижимости, который достоверно отражает их общую привлекательность, определите наилучшую позицию рассматриваемого объекта в рейтинге и присвойте ему балл CQA, изучив общие тенденции в рейтинге.
10. Присвойте рассматриваемому объекту соответствующий остаток/SF в расчёте на общую арендуемую площадь (GLA) (объект оценки).
11. Сложите оценочную цену MARS и присвоенный остаток, чтобы получить итоговую оценку стоимости объекта.
12. Разбейте остаток объекта на описательные компоненты и вклады в стоимость.
13. Рассчитайте вклады в стоимость, полученные на основе модели аналогов, корректировки, а также все промежуточные итоги и скорректированные цены продажи.
14. Скорректированные цены продажи будут равны окончательному выводу о стоимости объекта оценки.
15. Выберите 6–12 наиболее схожих объектов недвижимости для таблицы продаж.
16. Разбейте общие остаточные суммы по аналогам на остатки стоимости по компонентам, установив общую остаточную стоимость равной 0 после завершения разбивки.
17. Пересчитайте все промежуточные итоги и скорректированные цены продажи аналогов в таблице продаж. Проверьте значения.
18. Загрузите таблицу продаж в отчёт.

Обратите внимание, что при использовании метода RCA окончательный вывод о стоимости НЕ вытекает напрямую из аналогичных значений, приведенных в таблице продаж. При использовании метода RCA таблица продаж становится лишь инструментом для объяснения того, почему объекты недвижимости, внешне похожие на объект оценки, продаются дороже или дешевле, чем окончательное заключение о стоимости или оценочная стоимость объекта оценки. На это следует смотреть иначе: заключение о стоимости объекта оценки зависит от оценки цены по модели MARS и от остаточной величины, рассчитанной инженером-оценщиком. Однако эти две величины в значительной степени зависят от всех объектов недвижимости, включенных в регрессионный анализ R/earth, который сформировал ранжирование остатков объектов, определившее балл CQA для объекта оценки и, таким образом, определившее его остаточную стоимость посредством присвоения ему места в рейтинге инженером по оценке. Корректировки в таблице продаж на самом деле определяются наложенными математическими ограничениями RCA, хотя разбивка остаточной стоимости в некоторой степени зависит от оценщика, суммы вкладов по разбивке всегда должны в сумме равняться соответствующему остатку. Из доказательства 2 известно, что, хотя фактические значения компонентов остаточной стоимости могут варьироваться в зависимости от

суждения инженера по оценке, их суммы для данного объекта полностью ограничены и, следовательно, не изменят скорректированную цену продажи соответствующей недвижимости, независимо от того, как инженер по оценке изменяет разбивку остаточной стоимости.

5.3 Рабочий процесс обработки данных

Обратите внимание, что в данной статье представлен лишь общий рабочий процесс метода RCA. В других статьях будут подробно рассмотрены различные этапы и подэтапы анализа RCA.

5.3.1 Требования к данным для MARS

Нам необходимо указать дату начала продажных сделок и дату окончания, исходя из даты оценки или оценки, проведенной оценщиками. При этом обычно учитывается предполагаемое количество переменных регрессии и взаимодействий, которые потребуются для каждой степени взаимодействия, превышающей 1.

5.3.2 Протоколы: Список закрытых сделок

Protocol 1.1: Determining the Number of MLS Closed Sales Listings Required

В данном протоколе описывается, как рассчитать рекомендуемое минимальное количество завершённых сделок, которые необходимо провести ввод в регрессионный анализ MARS. Для начала вам потребуется провести приблизительную оценку, запустить MARS для анализа первой, второй и, возможно, третьей степени, чтобы сформировать исходную модель, а затем проверить количество строк в соответствии со следующими правилами, чтобы определить, требуется ли добавить дополнительные строки. В данном протоколе предполагается, что независимые переменные являются независимыми на 100% , что в сфере недвижимости встречается редко; например, общая арендуемая площадь (GLA) и количество комнат обычно сильно коррелируют, а GLA, вероятно, также коррелирует с площадью участка. Это связано с расчётом «степеней свободы (DOF)». Каждая действительно независимая переменная получает одну степень свободы. Однако при наличии зависимостей между так называемыми независимыми переменными точный расчёт DOF требует анализа, выходящего за рамки данной статьи. Просто имейте в виду, что из-за вероятной коллинеарности между переменными приведённые ниже оценки, скорее всего, завышены.

1. Для членов первого порядка (без взаимодействий) хорошо работает правило « ~ 15 » строк на переменную
2. Для каждого допустимого двустороннего взаимодействия следует добавить примерно 15–30 объектов недвижимости
3. Для взаимодействий более высокого порядка (3 и более переменных) каждое должно иметь 30 и более объектов недвижимости.

Однако это не жёсткое правило. Фактическое необходимое количество может зависеть от:

- Отношения сигнал/шум в ваших данных
- Сложности взаимосвязей, которые вы пытаетесь смоделировать
- Изменчивости вашей переменной отклика
- Качество ваших данных

Пример:

Итак, если у вас есть:

- 10 основных переменных
- 5 допустимых двусторонних взаимодействий
- 2 допустимых трёхсторонних взаимодействиях

Вы можете рассчитать минимальное количество продаж следующим образом:

$$\begin{aligned} N &= (10 \text{ переменных} * 15 \text{ строк/переменная}) + \\ &\quad (5 \text{ двусторонних взаимодействий} * 22 \text{ строк/двустороннее взаимодействие}) + \\ &\quad (2 \text{ трёхсторонних взаимодействия} * 30 \text{ строк/трёхстороннее взаимодействие}) \\ &= 320 \text{ строк} \end{aligned}$$

В качестве альтернативы можно поступить следующим образом: после запуска программы R/earth на наборе данных можно использовать количество базисных функций в модели, увеличенное на единицу, в качестве приблизительной оценки степени свободы, а затем умножить это число на 20–30, чтобы получить необходимое количество записей для ввода.

Protocol 1.2: Sales Periods Should Encompass Minimal Change in Buyer Tastes

При сборе исторических данных о недвижимости строгих временных ограничений не существует, если изменения рыночных предпочтений можно отразить с помощью соответствующих переменных. При анализе архитектурной эволюции на вашем рынке обратите внимание на следующие ключевые моменты: изменения доминирующих архитектурных стилей (например, переход от стиля «калифорнийское ранчо» к средиземноморскому) можно смоделировать с помощью:

- Идентификаторы микрорайонов
- Коды районов MLS
- Диапазоны дат строительства

Однако если эти изменения стилей невозможно адекватно учесть с помощью конкретных переменных, вы сталкиваетесь с важным ограничением: хотя «Дата продажи» обычно используется для корректировки с учётом рыночных условий, одновременное использование этой переменной для учета архитектурных предпочтений приведёт к смешению этих двух различных эффектов. В таких случаях предпочтительным подходом является либо:

1. Ограничить сбор данных периодом после появления более нового архитектурного стиля
2. Ограничить выбор аналогов микрорайонами, имеющими одинаковый архитектурный стиль

Это гарантирует, что ваш анализ сохранит согласованность характеристик недвижимости, при этом позволяя должным образом проводить корректировки рыночных условий с течением времени. В качестве альтернативы, по крайней мере в данном случае, если вы можете извлечь информацию об архитектурном стиле из данных, вы можете рассмотреть взаимосвязь между этими двумя переменными. Однако такое извлечение не всегда возможно.

Protocol 1.3: If In An Area Of Few Sales Then Do The Best You Can

При работе с моделированием MARS для данных о продажах для проведения значимого анализа требуется примерно 15 сделок. Если у вас меньше продаж (например, 5 объектов недвижимости), вы можете продублировать каждую запись дважды, чтобы достичь порога в 15 объектов недвижимости. Важные моменты, которые следует учитывать при использовании этого подхода:

1. Используйте только регрессию первой степени, чтобы уменьшить переобучение
2. Игнорируйте или пропускайте перекрёстную проверку, поскольку набор данных слишком ограничен
3. Хотя в данном сценарии метод MARS, вероятно, покажет лучшие результаты, чем методы сопоставленных пар или стандартной линейной регрессии, имейте в виду, что дублирование данных влияет на достоверность оценок неопределённости и статистической значимости

Лучшим долгосрочным решением является расширение набора данных за счёт дополнительных фактических данных о продажах. Однако, если вам приходится работать с ограниченным объёмом данных, дублирование данных представляет собой приемлемое временное решение при условии учёта его статистических ограничений.

5.4 Структура папок проекта

Было бы неплохо обсудить структуру папок проекта. Рекомендуется создать отдельные папки для «Анализа рынка» и «Оценки отдельных объектов недвижимости». Обе эти папки имеют схожую структуру вложенных папок:

```

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
│   │       │           └── ...

```

На этапе 1 обработки будет создана папка «version» с подпапками. В неё будет скопирован файл «MLS_Original.xlsx», хотя эта таблица не должна изменяться, а должна точно отражать данные, загруженные из MLS на определенную дату и время; однако в случае длительного проекта может возникнуть необходимость в новой загрузке данных из MLS. Далее на этом этапе в папку «version» будет скопирована расширенная книга MLS.xlsx с параметрами конфигурации. Может существовать множество папок «version», каждая из которых создается для полного цикла работы MARS.

Затем на следующих этапах будут выполняться добавление столбцов, сортировка и другие

изменения в книге, а по завершении оценки обновленная версия будет сохранена в папке с данными проекта под именем `MLS_Stage(N)`. Повторим: корневая папка `MLS.xlsx` постоянно изменяется, поэтому копия версии представляет собой непрерывную модификацию исходной книги `MLS` с переменными удалениями и добавлениями, а также другими изменениями. Код для обработки на языке `R` или `Python`, скорее всего, будет запускаться с использованием ввода из рабочей книги «`Data MLS.xlsx`». Таким образом, у нас всегда имеется копия всех данных, лежащих в основе каждой версии запуска `RCA`.

Создается столько версий, сколько необходимо, пока не станет невозможно внести дальнейшие улучшения; после чего лучшая версия выбирается в качестве окончательной. Можно изменить название папки, добавив в конец «`_Final`», чтобы это было видно при просмотре каталога. Некоторые старые версии с интересными или конкурирующими моделями, как правило, следует сохранять. Каждая папка версии содержит ряд текстовых файлов, диаграмм и графиков, которые можно встроить в отчёт. Окончательная книга `MLS_Stage5.xlsx` содержит столько листов, сколько необходимо для отчёта. В частности, в ней может быть таблица для переноса данных в таблицу продаж `Fannie Mae`. Она также обязательно будет содержать расчёты или разбивку, лежащие в основе агрегированных данных. Это, возможно, самая трудоёмкая часть подхода `RCA` — поиск идеальной модели в сложной рыночной зоне.

Все необходимое для воссоздания таблиц продаж следует скопировать в папку версии, включая используемый код на языке `R` или `Python` — если в код вносятся частые изменения.

5.4.1 Протоколы: Несколько прогонов `MARS` (для рынка недвижимости)

Protocol 2.1: Overfitting Runs To Surgically Study Market Area Pricing Factors

В ходе первоначальных прогонов для рынка недвижимости переобучение следует использовать стратегически с целью выявления основных факторов, влияющих на цену продажи. Например, рыночная зона, представленная 800 сделками, может содержать лишь 3 объекта недвижимости с мастерскими, и в этих единичных случаях наличие мастерской может оказывать существенное влияние на цену продажи. Прогон только в прямом направлении даст модель, подвергнутую переобучению, показывающую приблизительный вклад этих мастерских в стоимость. Однако, поскольку из 800 сделок таких всего 3, термины, связанные с мастерскими, будут удалены при обратном прогоне и, безусловно, при перекрёстной проверке. Значение, которое проходит только в прямом направлении присваивает мастерским, можно записать и перенести в анализ латентных переменных, чтобы использовать его для любого объекта недвижимости с мастерской, фигурирующего в таблице продаж. Такие переобученные модели также дают подробные оценки для многих других характеристик с небольшим объемом данных, которые были бы исключены в результате обрезки при обратном проходе или в ходе прогонов с перекрёстной проверкой.

Сохраните каждый из этих пробных прогонов в отдельной папке версии: зафиксированные значения вкладов являются доказательствами, на которые `VE` будет ссылаться при последующем разборе остатков. Разумеется, следует понимать, что в конечном итоге `VE` должен опираться на процедуры обрезки и перекрёстной проверки, чтобы сделать окончательную модель устойчивой к объектам недвижимости, с которыми регрессия никогда не сталкивалась — например, к объекту оценки, который появляется на более поздних этапах. Прогон с переобучением служат инструментами для исследования; они никогда не являются окончательной моделью.

Protocol 2.2: Multiple Cross-Validated Runs Under Different earth() Parameters

Рыночные зоны различаются по структуре и размеру, и специалисту по оптимизации (VE) обычно приходится проводить серию экспериментальных прогонов: чтобы понять, какие переменные оказывают наибольшее влияние на цену; чтобы оценить, позволяет ли имеющийся набор измеренных переменных адекватно оценить стоимость, или же необходимо добавить новые данные (расстояние до водоема, высота над уровнем моря, расстояние до торговых центров и т. д.); а также для выявления новых видов взаимосвязей и исключения других как неэффективных или вводящих в заблуждение. В конечном итоге VE стремится получить хороший коэффициент R^2 на обучающих данных и наилучший возможный коэффициент CVR^2 на тестовых данных. Для достижения этой цели может потребоваться сбор новых данных, повышение точности существующих данных или настройка параметров earth() с учётом размера и характера рыночной зоны (например, с учётом доступных степеней свободы — см. Протокол 1.1).

Именно поэтому приведенная выше структура проекта поддерживает множество папок с версиями прогонов: каждая потенциальная модель сохраняется вместе со своими параметрами и показателями R^2/CVR^2 , что позволяет сравнивать динамику результатов по различным прогонам. Версия, у которой показатель R^2 растёт, в то время как CVR^2 остается на прежнем уровне или снижается, отражает скорее шум, чем рыночный сигнал, и именно история версий позволяет это увидеть.

Protocol 2.3: Prolog Expansion of Remarks and Descriptive Feature Columns

Хотя эта функция пока находится на несколько экспериментальной стадии, earthUI теперь поставляется со встроенным движком Prolog (vProlog) и использует грамматику определённых предложений (DCG) для разбора столбцов свободного текста в формате MLS — прежде всего «Public Remarks», но также и замечания агентов, а также разделенные запятыми списки характеристик, такие как характеристики помещений или существующие постройки. Благодаря этому расширению такие фразы, как «собственная солнечная батарея», «гараж на 3 машины» или «мастерская 40x50», превращаются в обычные столбцы характеристик, которые можно использовать в регрессии, а восстановленные таким образом характеристики с небольшим объёмом данных напрямую подаются в анализ латентных переменных (см. последующий подраздел, посвящённый оценкам состояния и качества в качестве приближённых регрессоров; сама характеристика описана в сопутствующей статье earthUI в этом выпуске).

Расширение текста увеличивает набор потенциальных предикторов, что требует проведения нескольких циклов анализа earth/RCA для тестирования: расширенные столбцы необходимо проверять как в составе модели, так и вне её, объединять в производные подсчёты или булевы значения в случаях коллинеарности, а также сверять с CVR^2 , чтобы подтвердить, что они несут рыночный сигнал, а не шум. Каждый такой эксперимент следует рассматривать как отдельную папку версии прогона, точно так же, как в протоколах 2.1 и 2.2. А поскольку сама грамматика разбора находится в стадии разработки, VE должен сверять извлеченные столбцы с исходным текстом, прежде чем полагаться на них в модели.

5.5 Основная книга Excel «

» Поскольку большинство сервисов MLS загружают данные об объектах недвижимости в таблицы Excel в формате CSV, Excel, вероятно, является наиболее удобным способом хранения данных и различных параметров настройки. Для каждого проекта имеется одна основная книга Excel с именем `MLS.xlsx`, которая содержит один лист под названием «`MLSDData`» для хранения списков завершённых сделок по продаже недвижимости, загруженных из MLS, а также несколько листов с настройками и параметрами:

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

Исходные данные хранятся в книге с именем типа «`MLS_Original.xlsx`». Необходимо создать её копию с другим именем, например «`MLS.xlsx`», которая будет загружаться в программу RCA на языке R или Python. Обе эти книги хранятся в папке «`Data`». Теперь рабочая книга «`Data/MLS.xlsx`», вероятно, пройдет множество итераций, в ходе которых будут пробоваться различные комбинации переменных, параметров MARS и изменений других настроек. Каждая итерация создает новую папку для ввода/вывода данных, названную по дате и времени в формате `VER_YYYYMMDDHHMMSS`. В каждой папке будет находиться копия вводной книги `MLS.xlsx`, а также сгенерированные книги Excel для каждого из четырёх этапов обработки (этапы 2–5), а также документы вывода, диаграммы и графики. Обратите внимание, что этап 5, заключительный этап, может содержать ряд других листов для отчёта в зависимости от типа клиента и технического задания.

5.6 Резервное копирование

При записи новых или измененных таблиц Excel в папки версий проекта также можно создавать резервные копии в базах данных SQLite или PostgreSQL, расположенных в папке проекта. Вероятность непреднамеренного повреждения файлов баз данных впоследствии, в процессе проверки или использования в качестве справочных материалов, значительно ниже.

6 Обозначения книг

6.1 Описательные измерения

Для книги «М» существует примерно шесть описательных измерений:

1. Оценки «а»
2. Версии запуска «г»
3. Этапы рабочего процесса «w»
4. Листы «s»
5. Объявление о продаже недвижимости «р»
6. Набор переменных «v»

Итак, отдельную ячейку в конкретной книге М можно описать следующим образом:

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

(Примечание: “*” означает любой допустимый индекс)

или:

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

where:

- n_p = # of rows (depends on sheet)
- n_v = # of columns (depends on sheet)
- n_r = # of run versions
- n_m = # of workflow stages
- n_g = # of analysis projects
- n_h = # of workbook sheets

Следует понимать, что «W» обозначает всю книгу Excel, состоящую из шести основных листов данных, каждый из которых содержит собственную таблицу, которая при запуске соответствующей программы на R или Python загружается в отдельные внутренние датафреймы. И в R, и в Python есть встроенные функции, которые по имени могут определить, к какой строке или столбцу датафрейма относится данное имя. Например, если в $index_v_j$, $v = ("GLA", "LotSize", "BathRms", "BedRms", \dots)$ и $j=0$, $tov_j = "GLA"$. Как R, так и Python внутренне смогут сопоставить «GLA» с столбцом, в котором оно хранится, которым, например, может быть 15-й столбец соответствующего фрейма данных. Поэтому в данной статье мы используем индексацию в наших массивах имен, которые передают имя запущенной программе, которая, в свою очередь, преобразует это имя в фактический индекс столбца или строки для внутреннего фрейма данных. Это немного сложно. Но, если быть точным, это необходимо и наиболее удобно. Речь идет о «разделении обязанностей» между ролью аналитика и ролью разработчика; хотя на самом деле инженер по оценке (VE) может выполнять обе роли. Обратите внимание, что «data frame» в R в Python называется «dataframe».

Также обратите внимание, что лист параметров выполнения MARS не является измерением W , поскольку W относится к версии данных, осуществляемого ввода в MARS, и не связана с параметрами, определяющими, как MARS/earth обрабатывает эти данные.

Для большей ясности представлена сложная многомерная нотация для книги, которая более непосредственно связана с внутренними массивами данных R или Python, в которые загружаются листы книги с данными. Это «тяжелая» нотация. Однако когда мы переходим к доказательствам и т. п., мы избавляемся от значительной части этого «багажа», делая допущения о том, с чем мы работаем. Тем не менее, эта нотация помогает с самого начала получить более полное представление о рабочем процессе и данных.

Начнём с обозначения множества книг, которые мы будем обрабатывать.

6.1.1 Размер проекта (“a”)

После утверждения объема работ (SOW) необходимо создать папку проекта для хранения данных ввода и вывода, включая параметры конфигурации, выходные графики, диаграммы, таблицы и фрагменты отчетов. VE должен идентифицировать все проекты с помощью простого последовательного идентификатора, такого как целое число, чтобы в случае пропажи отчёта это было очевидно. Если проекты отменяются, то в соответствующей папке проекта должен быть единообразный способ указания на то, что проект отменен, но сама папка никогда не должна удаляться. Мы предполагаем, что идентификаторы проектов начинаются с 0 и увеличиваются на 1, хотя возможны и другие варианты. Кроме того, мы предполагаем, что все действительные идентификаторы проектов хранятся в упорядоченном наборе «a», созданном с помощью программы обработки.

Данные проекта по оценке находятся в рабочем листе «Проект» этой книги, но также могут храниться в какой-либо центральной базе данных SQL,

Вот предлагаемые поля проекта:

1. Локальный идентификатор: Короткий целочисленный идентификатор, который генерируется последовательно для компании. Он отлично подойдет в качестве первичного ключа для хранения данных — и может автоматически генерироваться базой данных при создании каждого последующего проекта.
2. Глобальный GUID: При желании вы можете создать GUID (глобальный универсальный идентификатор) с помощью

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

3. Местоположение объекта недвижимости (анализ рынка) или адрес
4. Дата оценки
5. Дата выдачи заказа
6. Дата осмотра
7. Дата составления первоначального отчёта
8. Дата составления первого обновленного отчёта
9. Дата составления второго обновленного отчёта
10. Дата составления третьего обновленного отчёта
11. Имя заказчика
12. Адрес заказчика
13. Название кредитной организации
14. Адрес кредитной организации
15. Имя владельца
16. Адрес владельца

17.

Пример списка идентификаторов проектов:

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

6.1.2 Параметр версии ("r")

На втором этапе рабочего процесса мы выполняем MARS-регрессию над вводом. Вероятно, это потребует от нескольких до множества запусков, поскольку мы будем изменять различные параметры для улучшения модели. При каждом запуске для вывода создается новая папка версии, имя которой формируется на основе даты и времени выполнения. Например,

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

latest_version : r_1 = "ver20241210_083124"

Многообещающие версии, аналог по характеристикам и выводу с конкурентными моделями, следует сохранять до тех пор, пока не будет выбрана окончательная версия, или, возможно, дольше, если они представляют какую-либо полезность. Возможно, вам захочется изменить название папки с версией, чтобы обозначить, что это окончательная версия, добавив слово «Final» в конец названия.

6.1.3 Этапы рабочего процесса ("w")

Рабочий процесс RCA обычно делится на пять отдельных этапов, перечисленных ниже. Каждый этап должен быть успешно завершен перед началом следующего. Специалист по верификации (VE) должен иметь возможность остановить обработку по окончании этапа, чтобы проанализировать результаты перед продолжением работы. Это необходимо, поскольку рыночные зоны являются достаточно сложными и зачастую весьма уникальными. Могут возникнуть нестандартные проблемы, требующие нового подхода. Эксперт по оценке (VE) должен иметь возможность напрямую работать с кодом на языках R или Python, чтобы вносить необходимые изменения и даже пробовать новые подходы. Кроме того, в таких случаях ему может потребоваться экспериментировать и пробовать различные подходы, что потребует многократного запуска того же этапа, над которым он работает. Вместе с тем у VE должна быть возможность запускать все этапы одной командой — с некоторыми оговорками (например, агрегация 3-й степени должна быть уже определена в результате предыдущего прогона этапа 2). Это означает, что по завершении каждого этапа он должен сохранять все свои внутренние данные в файлах или базах данных, чтобы следующий этап можно было запустить так, как будто предыдущий этап все еще находится в памяти, путем повторной загрузки данных предыдущего этапа в память.

Вот этапы:

w = ("stage1", -Настройка
 "stage2", -Обработка MARS
 "stage3", -Итоговая таблица продаж
 "stage4", -Фрагменты отчётов
 "stage5") -Проверка/очистка

stage2 : w₂ = "stage2"

6.1.4 Листы книги рабочей Excel ("s")

В рабочей книге RCA Core Excel имеются следующие листы, каждый из которых сопоставлен с внутренним фреймом данных:

s = ("MLSData",
 "Регрессионные переменные",
 "Расчёты",
 "Взаимодействия",
 "Агрегации 1-й& 3-й степени",
 "Агрегации 2-й степени")

Причина, по которой агрегации 2-й степени выносятся на отдельный лист, заключается в том, что их проще всего определить с помощью симметричной таблицы, в которой показаны все возможные пары взаимодействий, что позволяет указать агрегационную переменную для каждого вклада значений, связанных с данной парой, а также корректировок, которые необходимо агрегировать (сложить). С агрегациями третьей степени мы не смогли бы этого сделать, поскольку для этого потребовалась бы трехмерная симметричная таблица, что невозможно в Excel. Кроме того, таблица размером 20×20×20 содержала бы 8 000 ячеек, с которыми пришлось бы работать. Проще просто перечислить названия всех членов третьей степени в таблице с агрегациями первой степени, например:

Переменная	Агрегировать в
Площадь отделанных помещений над уровнем земли	GLA
Площадь отделанных помещений под уровнем земли	BGLA
Площадь неотределанных помещений	UBA
Площадь участка	LotSize
...	...
GLA_ _ LotSize	GLA
GLA_ _ DateOfSale	GLA
BathRms_ _ DateOfSale	GLA
Lat_ _ Long	Location
Lat_ _ Long_ _ GLA	Location
...	...

6.1.5 Список объектов недвижимости ("p") — строки MLSData

Списки объектов недвижимости составляют строки листа MLSData. В других листах их нет до этапа 4, на котором формируются фрагменты итогового отчёта с окончательным перечнем аналогов. Ниже приведен примерный список идентификаторов объектов:

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

6.1.6 Переменные MLS ("v") Столбцы MLSData

О работе с переменными будет сказано подробнее позже. На данном этапе достаточно отметить, что лист MLSData обычно содержит 30–40 переменных различных типов, однако для регрессии в MARS в качестве независимых переменных обычно выбирается лишь 10–20 из них. Эти переменные также встречаются на других листах, хотя, как правило, в виде списка всех возможных переменных, которые можно использовать для передачи в MARS. Пример списка переменных:

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

6.1.7 Пример окончательной нотации

Давайте применим приведенную выше нотацию в заключительном примере. Используя приведенные выше списки, пусть

$$r_1 \begin{matrix} a_{141} \\ w_1 \end{matrix} W_{s_0}^{p_2} v_0 \quad (3)$$

Обратите внимание, что здесь мы используем нулевую индексацию, то есть считается, что первый элемент списка имеет индекс 0. В Python также используется нулевая индексация, а в R — индексация, начинающаяся с 1.

Итак, из вышесказанного следует, что речь идет о значении ячейки для:

- Переменная v_0 или столбец «GLA»
- Для свойства r_2 или ID «MLS823542»
- Из листа s_0 или «MLSData»
- Для этапа рабочего процесса w_1 или «stage2»
- Из версии r_1 или «ver20241209_151130»
- Из проекта a_{141} или «142»

Приведённую выше запись можно также записать следующим образом:

$$\begin{matrix} \text{"ver20241209_151130"} \\ \text{"stage2"} \end{matrix} \begin{matrix} \text{"142"} \\ W \end{matrix} \begin{matrix} \text{"MLS823542"} \\ \text{"GLA"} \end{matrix} \quad (4)$$

6.1.8 Обоснование

«Почему бы нам просто не использовать индексацию или цифры в обозначениях?» К нашим обозначениям предъявляются два требования:

1. Используются подмножества: например, аналитик часто выбирает подмножества из списка, такие как подмножества переменных. Например, на листе `MLSDData` может содержаться 35–40 переменных, но для конкретного прогона используется только 10–20 из них. От одного прогона к другому `VE` часто меняет перечень переменных, передаваемых в `MARS/earth`.
2. Итерация, например, $i = 1, 2, \dots, n_1$
 Наши алгоритмы, протоколы и доказательства обычно должны иметь возможность проходить по всем элементам упорядоченного списка, от первого до последнего, или выбрать какой-либо конкретный элемент.

Итак, нам нужна буква для обозначения подмножества или списка элементов, и у этой буквы должен быть еще один индекс, с помощью которого мы сможем перебирать элементы списка, чтобы получить строки, используемые для индексации строк или столбцов датафрейма.

6.1.9 Матрицы?

А как насчёт матриц из линейной алгебры? Матрицы работают только с числами — целыми или действительными. Матрицы как структура данных поддерживаются как в `R`, так и в `Python`. Однако многие переменные, с которыми мы работаем в данных `MLS`, называются «строками» или символьными данными. Даже если некоторые переменные представляют собой только числа, зачастую это не действительные числа, а просто имена. Например, коды `MLS` для города могут быть следующими: («660», «661», «662», «663», ..., «672»). Регрессия `Earth` позволяет задавать переменные как «факторные», и в этом случае числа рассматриваются просто как имена, без какой-либо последовательности. В таких случаях построенная модель будет просто предоставлять вклад каждого числа в значение, а не функцию, пробегающую по последовательности. Другие примеры включают «C1», ..., «C6» для параметра «Condition» и «Q1», ..., «Q6» для параметра «quality».

6.1.10 Первое упрощение

Приведенная выше нотация для основной книги проекта `Excel` удобна для связывания всех элементов воедино, но слишком громоздка для использования при объяснении алгоритмов или математических доказательств. При выполнении рутинной работы мы можем исходить из того, что она относится к данному проекту, версии, этапу и листу. Мы просто указываем, с чем работаем, если это не следует из контекста. На самом деле, после того как мы указали, с чем работаем, из приведённого выше нам нужно только следующее:

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

where: n_p = # properties
 n_v = # input variables
 n_s = # input sheets

6.1.11 Второе упрощение

Удобнее переименовать листы в соответствии с конкретными типами данных, которые они поддерживают, вместо использования индексов. В каждом листе строки и столбцы

выполняют разные функции. Таким образом, набор данных W разбивается на массивы данных M , V , A , I и C следующим образом:

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

where: $n_p = \#$ properties
 $n_v = \#$ input variables

Примечание: По умолчанию объект оценки всегда находится первым в списках свойств В таблице Excel «MLSDData» и соответствующем фрейме данных первая строка всегда зарезервирована для объекта оценки. Оно никогда не сортируется вместе с другими объектами недвижимости MLS, расположенными ниже. Оно никогда не вводится в регрессию MARS в качестве независимых данных, а вводится в качестве целевых данных или в качестве зависимой переменной. То же самое относится и к спискам объектов недвижимости. Во всех списках объектов недвижимости первый элемент с индексом 0 всегда считается идентификатором объекта оценки, если не указано иное.

Даже если задача состоит в создании модели для рыночной зоны без каких-либо объектов оценки, первая строка всё равно заполняется фиктивными данными о свойствах. Это значительно упрощает написание программ, определение алгоритмов или проведение доказательств. Итак, если мы используем c для представления списка объектов недвижимости, то c_0 является объектом оценки, а c_i и $i = 1, \dots, n_p$ — аналогами, которые будут введены в регрессионный анализ MARS/earth.

◇ Спецификации переменных регрессии

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

where: $n_v = \#$ available variables
 $n_s = \#$ of specifications

Обратите внимание, $n_s \leq n_v$, что обычно мы выбираем подмножество доступных переменных для ввода в регрессию.

◇ Расчеты

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

where: $n_v = \#$ of variables
 $n_c = \#$ of calc specs

◇ Interactions

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

where: $n_f = \#$ from variables
 $n_t = \#$ to variables

Обратите внимание, что таблица данных I является симметричной, с переменными, перечисленными в алфавитном порядке в строках «from» и столбцах «to». Ячейки помечены цифрой «1», чтобы указать, что взаимодействие между соответствующей парой переменных допускается. В противном случае цифра «0» означает, что взаимодействие не допускается. Для взаимодействия второй степени это относительно просто, так как возможно только одно взаимодействие. В случае регрессии третьей степени у нас есть три парных взаимодействия между всеми тремя переменными, причём каждая пара

взаимодействий должна быть разрешена, чтобы сформировать терм (или переменную) взаимодействия трёх переменных.

◇ Агрегации

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

where: $n_v = \#$ of variables

7 Таблица продаж RCA

Имейте в виду, что протокол RCA требует использования программы на языке R или Python, способной обращаться к R/earth и выполнять все необходимые шаги. Этот процесс требует обработки больших объемов данных, и выполнить эту работу вручную просто невозможно. В то же время, даже если предположить, что программа не содержит ошибок, вам понадобится программное обеспечение, чтобы гарантировать, что необходимая работа будет выполнена точно и без человеческих ошибок.

Кроме того, специалистам по оценке (VE), выполняющим эту работу, необходимо иметь возможность модифицировать такие программы по мере необходимости, особенно когда они сталкиваются с новыми непредвиденными рыночными зонами, районами, объектами недвижимости или другими новыми требованиями и проблемами.

В рамках обработки RCA мы используем три типа таблиц продаж:

1. Таблица продаж RCA Processing, обозначенная в данной статье как «MLSDData», представляет собой очень большой фрейм данных и связанную с ним электронную таблицу, которые могут содержать множество промежуточных сведений, подвергающихся впоследствии процессу агрегирования для получения таблицы продаж RCA, упомянутой ниже. Привести пример здесь было бы невозможно из-за его большого объёма. Скорее всего, инженеры по оценке будут адаптировать эту электронную таблицу под свои предпочтения. В будущем я планирую написать отдельную статью о разработке такой таблицы.
2. Окончательная таблица продаж RCA, используемая в отчетах, не относящихся к GSE, выглядит примерно так: [Table 2](#). Она действительно имитирует, например, форму Fannie Mae, что позволяет читателю почувствовать некоторую знакомость. Однако обратите внимание, что в ней больше столбцов на каждый аналог продажи, чем в таблице Fannie Mae: у каждого аналога есть столбец «Вклад в стоимость» наряду со столбцами «Показатель» и «Корректировка», и у объекта оценки он также присутствует — этих столбцов нет в форме Fannie Mae. Еще более простая форма — это таблица продаж [«Table 1»](#).
3. Это таблица продаж GSE Upload, которая используется для загрузки таблицы продаж в программное обеспечение, такое как Alamode. Однако обратите внимание, что не все программные пакеты для оценки недвижимости предоставляют возможность загрузки данных таблицы продаж из электронной таблицы. По необходимости такие таблицы не могут содержать отдельный столбец для вкладов в оценку, понятия о которых у государственных ипотечных агентств (GSE), таких как Fannie Mae, отсутствует. Стратегия заключается в том, чтобы предоставить им требуемую форму, но подкрепить её таблицей продаж RCA, а также другой сопроводительной документацией, такой как подробная информация о любых агрегированных переменных.

Таблицы продаж RCA Processing не похожи на таблицы продаж Fannie Mae, но, грубо говоря, представляют собой транспонированную версию таблицы продаж Fannie Mae с переставленными строками и столбцами: каждая строка содержит данные по какому-либо объекту недвижимости, а каждый столбец — данные по какой-либо характеристике объекта. Количество строк с данными об объектах может легко достигать сотен или тысяч, а количество столбцов — 200 и более. Первая строка зарезервирована для имен переменных, а вторая — для объекта оценки, даже если мы проводим лишь рыночный анализ. Когда лист импортируется в дата-фрейм, первая строка листа становится характеристикой «names», а вторая строка, относящаяся к объекту оценки, становится первой строкой дата-фрейма.

Столбцы в таблице продаж RCA включают идентификационные столбцы, а затем первичные, внешние и синтетические переменные, которые объединяются и передаются в MARS в качестве независимых переменных. На основе этих переменных MARS делает вывод о том, какие из них оказывают существенное влияние на цену продажи, также известную как зависимая или целевая переменная. При регрессии второй степени мы можем получить пары этих переменных, которые выступают в качестве новых переменных. При регрессии третьей степени мы также можем получить тройки переменных, которые выступают в качестве так называемых независимых переменных, хотя они, конечно же, сильно коррелируют с исходными переменными.

Таблица 1: Пример таблицы RSA: вертикальная компоновка

Subject															Comp 1					Comp 2				
Measure		Value Contribution			Measure		Value Contribution			Adjustment		Measure		Value Contribution			Adjustment							
Sale Price					\$1,150,000							\$1,500,000												
Base Value		\$500,000					\$500,000			\$0.00				\$500,000			\$0							
Measured Features															Measured Features					Measured Features				
Living Area		1,530			\$463,500			1,350		\$382,500			\$81,000		1,700		\$540,000			-\$76,500				
Lot Size		6,000			\$200,000			6,500		\$220,000			-\$20,000		7,000		\$240,000			-\$40,000				
Age		53			\$63,600			57		\$68,400			-\$4,800		51		\$61,200			\$2,400				
Bathrooms		2			\$15,000			3		\$30,000			-\$15,000		3		\$30,000			-\$15,000				
Bedrooms		3			\$6,000			4		\$9,000			-\$3,000		4		\$9,000			-\$3,000				
Total Measured					\$1,248,100.00					\$1,209,900.00			\$38,200				\$1,380,200.00			-\$132,100				
Residual Features															Residual Features					Residual Features				
Residual		CQA 5.8			\$30,000			CQA 2.5		-\$59,900			\$89,900		CQA 8.2		\$119,800			-\$89,800				
Design		Residual Breakdown								Residual Breakdown							Residual Breakdown							
Condition					7,000					-\$15,000			\$22,000				\$35,000			-\$28,000				
Quality					5,000					-\$8,000			\$13,000				\$20,000			-\$15,000				
Func Utility					8,000					-\$10,000			\$18,000				\$30,000			-\$22,000				
Amenities					5,000					-\$12,000			\$17,000				\$15,000			-\$10,000				
Sub-Total					30,000					-\$14,900			\$19,900				\$19,800			-\$14,800				
										-\$9,900			\$89,900				119,800			-\$89,800				
Total								Total							Total									
Tot Value Contribution					\$1,278,100.00					\$1,150,000.00							\$1,500,000.00							
Adjusted Sale Price													\$1,278,100.00							\$1,278,100.00				

8 Переменные

Что касается ввода при запуске MARS, у нас есть только простые переменные, а именно:

1. Характеристики объектов недвижимости, загруженные из таблицы MLS.
2. Производные переменные, созданные оценщиком для уточнения переменных MLS или преобразования существующих переменных. Например, оценщик может преобразовать «Дату продажи» в «DaysOffMarket», равную дате оценки за вычетом «Даты продажи».
3. Оценщик может импортировать данные из сторонних источников, таких как ставки по ипотечным кредитам. Или, если в данных MLS отсутствуют данные о широте, долготе или высоте над уровнем моря, он может импортировать их из других источников. Эти переменные можно назвать «внешними переменными».

На выводе системы MARS мы работаем с ценовой моделью, представляющей собой комбинацию слагаемых, которые либо:

1. Содержат только одну из вводных переменных и предоставляют уравнение, определяющее вклад этой переменной в стоимость. Их называют слагаемыми первой степени или (выходными) переменными. Само слагаемое рассматривается как величина вклада в стоимость и получает название.
2. Содержат две вводные переменные и представляют собой уравнение, определяющее вклад в значение при взаимодействии данных двух переменных. Эти члены можно называть членами второй степени или (выводными) переменными.
3. Содержат три вводные переменные и представляют собой уравнение, определяющее вклад в значение при взаимодействии этих трех переменных. В данной работе эти члены могут называться членами третьей степени или (выводными) переменными.

Таким образом, выводными переменными модели MARS являются вклады в значение, включая вклад остаточной величины. В свою очередь, они рассматриваются как переменные при расчетах в таблице Sales Grid, например при вычислении корректировок или построении графиков.

Существуют различные типы переменных, которые подразделяются на две основные категории: (1) переменные происхождения и (2) переменные применения, т. е. неостаточные и остаточные.

8.1 Онтология «Variable Origin»

1. Переменные первой степени
 - 1.1 Переменные регрессии MLS
 - 1.2 Внешние переменные регрессии
 - 1.3 Производные переменные регрессии

2. Факторные переменные:

Для работы с факторными переменными требуется функция, способная их идентифицировать. Поскольку они задаются в рабочей книге конфигурации проекта Excel, создается функция, которая возвращает значение «true», если переданная ей переменная была классифицирована как факторная.

3. Переменные взаимодействия: они создаются программой MARS при построении модели, если степень взаимодействия установлена на 2 или выше. Однако рекомендуется использовать не более 3 взаимодействующих переменных из-за сложности интерпретации, вероятности переобучения, а также влияния на степень свободы и связанного с этим увеличения количества записей о недвижимости, необходимого для обеспечения надежности модели.
 - 3.1 Переменные второй степени: Они встречаются часто; хорошими примерами могут служить GLA:BathRms, GLA:LotSize, DateOfSale:BathRms, Latitude_ Longitude.
 - 3.2 Переменные третьей степени: менее распространённые взаимодействия, требующие осторожности и указывающие на возможную подгонку.
4. Переменные разбивки остатков

8.2 Онтология переменных приложений

1. Вклад MARS в создание стоимости
 - 1.1 Переменные первого порядка
 - 1.1.1. Переменные регрессии MLS
 - 1.1.2. Внешние переменные регрессии
 - 1.1.3. Производные переменные регрессии
 - 1.2 Переменные взаимодействия
 - 1.2.1. Переменные второй степени
 - 1.2.2. Переменные третьей степени
 - 1.3 Переменные разбивки остатков
 - 1.4 Агрегированные переменные
2. Переменные, сгенерированные инженерами& Значения
 - 2.1 Переменные разбивки остатков
 - 2.1.1. Дизайн/эстетика
 - 2.1.2. Ландшафт
 - 2.1.3. Функциональная полезность
 - 2.1.4. Вид
 - 2.1.5. Состояние
 - 2.1.6. Качество
 - 2.1.7. ...

Эти группы переменных будут обозначены соответствующими символами и более подробно описаны в следующем разделе

Protocol 3.1: Factor Variables

Факторные переменные, или просто факторы, могут встречаться во всех слагаемых, включая взаимодействия. Они важны, поскольку при проведении регрессии MARS и построении графиков с ними необходимо обращаться особым образом.

В то время как переменные с символьными значениями по определению являются факторами, числовые переменные необходимо пометить как факторные, чтобы MARS распознавал их как таковые. Например, районы MLS часто обозначаются номерами. В частности, районы в районе залива Сан-Франциско часто обозначаются уникальными целыми числами. Если эти значения передаются в MARS в качестве факторов, то программа попытается определить, имеют ли определенные районы средний вклад в объем продаж недвижимости в пределах своих границ, не учитывая при этом никаких взаимосвязей с другими районами. Напротив, в случае обычных переменных программа будет предполагать наличие некоей непрерывности в вкладе значений по мере изменения значения переменной. Это важное различие. Кроме того, эти факторные переменные могут встречаться в членах модели вместе с нефакторными переменными, что затрудняет построение графиков. В частности, при построении графиков членов с одной переменной и факторами мы объединяем все такие факторы в одну группу и просто строим гистограмму, показывающую их вклад — положительный или отрицательный.

Если факторы фигурируют в выражениях вместе с другими переменными, они всегда будут иметь только одно возможное значение, поскольку факторные переменные могут принимать только одно значение для данного объекта недвижимости: если значение площади равно 662, то по определению она не может иметь никакого другого значения для данной площади. Если вам это кажется запутанным, то, что ещё хуже, в модели, сгенерированной MARS, нет единого имени переменной для фактора; MARS возьмёт корневое имя факторной переменной, введённой в качестве вводных данных, а затем создаст новые столбцы с именами переменных, образованными путём добавления возможных значений. Таким образом, один столбец в таблице Excel, например `MlsArea`, будет соответствовать нескольким столбцам в массиве данных, созданным для каждого возможного значения, таким как `MlsArea660`, `MlsArea661` и т. д. Переменная, такая как `MlsArea661`, будет иметь значение 1 для объектов недвижимости в зоне 661; в остальных случаях — 0. Затем MARS выполнит регрессионный анализ по этим столбцам, как если бы они представляли булевы переменные. Эти измененные имена будут использоваться в модели, предоставленной MARS.

Вероятно, вам захочется преобразовать эти имена моделей обратно в их исходные названия, чтобы сделать графики более понятными для клиентов, которые предпочитают, чтобы всё было красиво и просто (по крайней мере, настолько просто, насколько это возможно). Вместо того чтобы видеть `MlsArea661=1`, мы будем видеть `MlsArea=661`. Затем мы сможем построить графики членов, просто указав в заголовке графика что-то вроде «Для `MLS Area = 662`». Таким образом, члены степени N с M факторными переменными отображаются как члены степени $(N-M)$, а факторное измерение исчезает в заголовке.

Обратите внимание, что в одном члене могут присутствовать две или более факторных переменных, поэтому, скорее всего, вам понадобится график для каждой возможной пары значений. Если переменная x с тремя различными значениями и переменная y с двумя различными значениями появляются в одном члене вместе с нефакторной переменной z , то вам понадобится $2 \cdot 3 = 6$ графиков. Примером может служить сочетание «`MlsArea`» + «`Architectural Style`»: например, стоимость викторианских домов в Сан-Франциско может варьироваться в зависимости от района города, в котором они расположены. К счастью, в большинстве районов обычно имеется лишь 1–3 архитектурных стиля, достаточно значимых с точки зрения влияния на стоимость, чтобы их можно было использовать в модели (хотя может быть 6 и более различных архитектурных стилей).

8.3 Члены второй и третьей степени

При регрессии второй или третьей степени некоторые члены выводной модели могут включать две или, соответственно, три переменные. Например:

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

Обратите внимание, что мы присваиваем им имена, сортируя имена переменных в алфавитном порядке и объединяя их, возможно, с использованием подчеркивания. То же самое мы делаем с членами третьей степени в тройках имен переменных, сгенерированных моделью. Например, «BathRms_GLA_LotSz».

Эти переменные добавляются к основным вводным переменным на этапе 3, когда программа RCA рассчитывает вклад всех переменных для каждого объекта недвижимости, включая переменные взаимодействия.

8.3.1 Переменные корректировки

Программа RCA также рассчитывает корректировки для всех переменных и объектов недвижимости на этапе 3, после расчета вклада каждого элемента в стоимость, путем вычитания вклада соответствующего аналога объекта недвижимости по данному элементу из вклада объекта оценки по этому же элементу.

8.3.2 Агрегирующие переменные

Агрегирующие переменные определяются для замены подмножеств переменных, объединяя все их значения и корректировки в одну общую переменную, которая будет использоваться в отчёте. Например, мы могли бы разбить как готовую, так и незавершённую жилую площадь на более детализированные переменные:

1. Законно обустроенная площадь над уровнем земли
2. Законно обустроенная площадь ниже уровня земли
3. Незаконно обустроенная площадь над уровнем земли
4. Незаконно обустроенная площадь ниже уровня земли
5. Необустроенная площадь над уровнем земли
6. Необустроенная площадь ниже уровня земли

Да, есть сферы, в которых это действительно имеет смысл.

Регрессия MARS будет проводиться по всем переменным и позволит построить модель, которая определит вклады и корректировки для каждой из них. Однако для отчёта в установленной форме эти различные вклады, вероятно, придется объединить или агрегировать в меньшее количество переменных. Подробности агрегирования следует привести в приложении. Более распространённым примером может служить агрегирование корректировок на взаимодействие в одну из корректировок на основные переменные. Например, мы можем агрегировать корректировку на GLA_LotSize в корректировку на GLA или на LotSize, либо даже распределить их поровну между обеими переменными, как указано в листе «Агрегирование».

8.3.3 Специальные переменные & Обозначения

π_i	Фактическая цена продажи недвижимости $i = 1, 2, \dots, n_p$; для объекта оценки π_0 обозначает вывод о стоимости (см. Доказательство I)
α_i	Скорректированная цена продажи недвижимости $i = 1, 2, \dots, n_p$
ρ_i	Расчетная цена продажи по MARS для недвижимости $i = 0, 1, \dots, n_p$
β	Базовая стоимость модели
$\varphi_{i,j}$	Вклад переменной регрессии $j = 1, 2, \dots, n_r$ в значение для объекта $i = 0, 1, \dots, n_p$
$\delta_{i,j}$	Корректировка для переменной регрессии $j = 1, 2, \dots, n_r$ для объекта $i = 1, 2, \dots, n_p$
ϵ_i	Общая величина остатка для объекта $i = 0, 1, \dots, n_p$
$\epsilon_{i,c}$	Вклад компонента остаточной величины $c, c = 1, 2, \dots, n_c$ для объекта $i = 0, 1, \dots, n_p$
$\xi_{i,c}$	Корректировка для компонента остаточной величины $c, c = 1, 2, \dots, n_c$ для объекта $i = 0, 1, \dots, n_p$

где

$n_p = \#$ of comparable properties

$n_r = \#$ of regression variables

$n_c = \#$ of residual breakdown variables

8.3.4 Разбивка остатков

После построения модели с помощью регрессии MARS система генерирует расчетные цены продажи на основе входных переменных, выбранных системой MARS. Затем это значение вычитается из соответствующей чистой цены продажи, чтобы получить остаток для каждого объекта недвижимости. Теперь, согласно протоколу RCA, после того как оценщик

(VE) определил окончательный набор аналогов (обычно 6–12 объявлений о продаже), ему необходимо проанализировать каждое объявление и разбить рассчитанный остаток на наиболее вероятные характеристики (латентные переменные), которые, по его мнению, влияют на стоимость, игнорируя возможность случайных ошибок в данных. Например, для данного объекта недвижимости у него могут быть следующие переменные разбивки:

1. Дизайн
2. Функциональность
3. Продажа по цене выше/ниже рыночной
4. Состояние
5. Качество

Различные объекты недвижимости, вероятно, будут иметь несколько разных переменных, характеризующих остаточную стоимость, которые необходимо собрать в один упорядоченный набор.

Оценка вклада редких признаков. Некоторые значимые признаки встречаются слишком редко для регрессионного канала: автомастерская, присутствующая в 3 из 780 сделок, не может обеспечить обоснованный коэффициент, и при прямом проходе такая переменная обычно полностью игнорируется. Это остаточные признаки, и их вклад можно оценить в остаточном пространстве: выполните подгонку модели без данного признака, а затем проведите регрессию остатков модели на исключенный столбец (столбцы) методом обычных наименьших квадратов. Для бинарного флага коэффициент представляет собой единовременный вклад (что эквивалентно среднему смещению остатков по сделкам с этим флагом); для числового столбца — вклад на единицу. Когда оцениваются совместно несколько коррелирующих удобств (например, площадь пристройки вместе с флагом «мастерская»), регрессия распределяет кредит между ними, а не учитывает одни и те же доллары дважды в двух строках разбивки. Оценки, основанные на менее чем десяти наблюдениях, носят лишь ориентировочный характер и должны быть подтверждены амортизированной стоимостью или расширенной выборкой (например, флагами, извлеченными из публичных комментариев); само ограничение по остаточной величине ограничивает любое значение, которое в конечном итоге будет присвоено. earthUI (v0.11+) реализует этот метод непосредственно в виде вкладки «Вклад остаточной величины», включая совместную оценку, данные об остаточной величине по каждой сделке и примечание к рабочему файлу.

Обозначения: Используя обозначения из приведенной выше таблицы, получаем:

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

где n_c — количество компонентов разбивки остатка.

Эксперт по оценке (VE) должен самостоятельно определить, как разбить остаток на указанные компоненты. Можно было бы подумать, что это может привести к появлению предвзятости в окончательном заключении о стоимости. Однако это не так, поскольку у нас есть математическое доказательство того, что если сумма корректировок по разбивке равна остаточной стоимости, они не повлияют на заключение о стоимости. Цель разбивки заключается не в установлении стоимости, а в объяснении того, почему остаточная стоимость объекта такова, какова она есть, по сравнению с остаточными стоимостями других объектов недвижимости, проданных по рыночной цене. Единственный момент, в котором субъективный подход инженера может повлиять на окончательный вывод о стоимости, — это оценка остатка стоимости объекта оценки, поскольку, разумеется, для него отсутствует фактическая цена продажи, необходимая для расчета остатка стоимости.

Каждому объекту недвижимости соответствует свой набор вкладов в остаточную стоимость, однако их необходимо объединить в один набор C для всех объектов недвижимости:

$$\text{Letting: } A_i = \{\{\epsilon_{i,c}\}_{c=1}^{n_c}\}$$

$$C = \{x \mid x \in \bigcup_{i=0}^{n_p} A_i\}$$

где $[\epsilon_{i,c}]$ — имя переменной, соответствующей значению $\epsilon_{i,c}$, а A_i — множество таких переменных, и n_c — количество значений в каждом ϵ_i .

Таким образом, если данные по объекту оценки и аналогичные данные по характеристикам расположены в отдельных столбцах, то под всеми вкладами значений переменных регрессионной модели для каждой характеристики мы должны увидеть множество C из n_c соответствующих вкладов значений остатков.

9 Основные доказательства метода RCA

Пусть:

$$n_p = \# \text{ of properties}$$

$$n_r = \# \text{ of regression variables}$$

$$n_c = \# \text{ of residual component variables}$$

Доказательство I: Каждая аналог скорректированная цена продажи равна расчетной цене продажи рассматриваемого объекта.

Для любого аналога i скорректированная цена продажи α_i представляет собой сумму чистой цены продажи π_i и следующих корректировок: разницы между вкладом аналога и объекта оценки в базовую стоимость, регрессионных переменных $j = 1, 2, \dots, n_r$ и остатка. Для объекта оценки, по которому отсутствуют данные о фактической продаже, чистая цена продажи π_0 определяется как итоговая оценка стоимости $\beta + \varphi_{0,1} + \varphi_{0,2} + \dots + \varphi_{0,n_r} + \epsilon_0$. Тогда:

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

Доказательство II: Изменение структуры остаточной ошибки при условии, что сумма вкладов всех компонентов остаточной ошибки равна аналогу остаточной ошибки, не влияет на окончательный вывод протокола RCA относительно значения.

Это почти тривиально. Однако, не задумываясь о математике, об этом легко упустить из виду, поскольку традиционные оценщики сосредотачиваются на корректировках, а не на вкладе в стоимость. Именно вклад в стоимость в данном случае компенсирует друг друга — это упущение, существовавшее на протяжении десятилетий. Для некоторого аналога объекта k , $k > 0$, общая корректировка остатков равна сумме остаточных компонентов разбивки n_c , которая, в свою очередь, равна сумме разниц в вкладах в стоимость, равной общей корректировке остатков объекта оценки минус общая корректировка остатков аналога объекта. И это не зависит от количества созданных остаточных компонентов или от того, что они представляют. До тех пор, пока значение остаточных компонентов равно общему остатку как для объекта оценки, так и для выбранного аналога, ξ_k не изменится.

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

10 Модель MARS «

»

10.1 Функции «g» для моделирования переменных и взаимодействий

Одним из основных принципов метода RCA является то, что «объяснительное рассуждение» имеет первостепенное значение. К сожалению, построение содержательных графиков для различных уравнений вклада значений, особенно при использовании взаимодействий и факторных переменных, может представлять собой сложную задачу. Поэтому считается крайне важным разработать четкий и строгий протокол для создания этих графиков. Данный протокол приводится ниже, однако следует понимать, что он предназначен для разработчиков, пишущих код на языках R или Python для генерации графиков, и может представлять лишь поверхностный интерес для тех, кто не занимается программированием.

Ценовая модель, сгенерированная системой MARS, представляет собой сумму базовой константы и дополнительных слагаемых, каждое из которых включает от 1 до 3 переменных. Для её анализа мы сначала разбиваем модель на отдельные слагаемые. Затем несколько слагаемых, имеющих общий набор переменных и представляющих одно взаимодействие, объединяются в отдельные выражения взаимодействия. Предположим, что максимальная степень регрессии равна 3. Рассмотрим модель регрессии третьей степени, состоящую из базовой константы, n_1 выражений первой степени, n_2 выражений взаимодействия второй степени и n_3 выражений взаимодействия третьей степени. В результате мы получаем четыре отдельные группы выражений, представляющие собой либо однопеременные, либо многопеременные взаимодействия, каждое из которых отображается на графике по-разному в зависимости от двух факторов: (1) включают ли они факторную переменную (категориальную переменную), и (2) количество уникальных переменных, которые они содержат. В частности:

- Выражения первой степени (1 переменная) отображаются в виде двумерных графиков.
- Выражения второй степени (2 переменные) отображаются в виде трехмерных графиков.
- Выражения третьей степени (3 переменные) отображаются в виде набора трехмерных графиков.
- Если выражение содержит факторную переменную, каждая уникальная факторная переменная уменьшает размерность графика на единицу (например, трехмерный график становится двумерным, если одна из переменных является факторной).

Индексация каждой функции g предоставляет информацию, необходимую для определения способа построения её графика, отражая количество уникальных переменных и наличие факторных переменных.

10.2 Индексирование функций g

Каждая функция g индексируется как $f_{j,k} g_k^j$, где:

- $f_{j,k}$ — первый индекс (верхний индекс перед g):
 - j принимает значения от 0 до 3 и определяет группу:
 - * $j = 0$: базовая константа (исключение из обычной системы индексации).

- * $j = 1$: выражения первой степени.
- * $j = 2$: выражения второй степени.
- * $j = 3$: выражения третьей степени.
- k — индекс конкретной функции внутри её группы (например, $k = 1$ для первой функции в группе j).
- Значение $f_{j,k}$ указывает на количество уникальных факторных переменных в функции:
 - * $f_{j,k} = 0$: факторных переменных нет.
 - * $f_{j,k} > 0$: количество факторных переменных (например, 1, 2 или 3).
- Индексы в конце g :
 - Верхний индекс j после g (например, g^j) повторяет номер группы для ясности.
 - Индекс k после g (например, g_k) соответствует k в $f_{j,k}$, идентифицируя функцию внутри её группы.

Хотя индексы f и g совпадают по значению, их роли различаются: $f_{j,k}$ подсчитывает факторные переменные, что имеет решающее значение для принятия решений при построении графиков, в то время как j и k в g определяют её группу и положение.

10.3 Размеры графика

Размер d графика для функции $f_{j,k} g_k^j$ вычисляется следующим образом:

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

где:

- j — количество уникальных переменных (или номер группы, без учета случая с константой).
- $f_{j,k}$ — количество факторных переменных.
- d — количество непрерывных (нефакторных) переменных, которые будут отображаться на графике, при этом максимальное количество составляет 3 (поскольку максимальная степень равна 3). Функция sd непрерывными переменными строится в $d + 1$ измерениях: $d = 1$ даёт двумерный график, $d = 2$ — трёхмерный график (или двумерную тепловую карту), $d = 3$ — набор трёхмерных графиков.

Например:

- Если $j = 3$ (3 переменные) и $f_{j,k} = 0$ (без факторных переменных), то $d = 3 - 0 = 3$, что требует набора 3D-графиков, каждый из которых фиксирует одну из переменных.
- Если $j = 3$ и $f_{j,k} = 1$ (одна факторная переменная), то $d = 3 - 1 = 2$, что требует 3D-графика (или 2D-тепловой карты) для каждого значения факторной переменной, фигурирующей в модели.
- Если $j = 0$ (константа), построение графика, как правило, не требуется, так как это одно значение.

Когда $d = 3$ (например, выражение третьей степени без факторных переменных), для визуализации всей функции может потребоваться несколько 3D-графиков, каждый из которых фиксирует одну переменную на разных значениях, чтобы показать подмножество поведения функции.

10.4 Окончательное уравнение модели

Модель MARS разбита на четыре группы выражений (постоянные, выражения первой, второй и третьей степени). Индексация $f_{j,k}^j g_k^j$ указывает:

- j : группу и количество переменных.
- $f_{j,k}$: количество факторных переменных.
- $d = j - f_{j,k}$: размерность графика.

Эта структура определяет, как визуализируется каждая функция g , обеспечивая соответствие метода построения графика сложности выражения и типам переменных.

Оценка цены продажи $\hat{P}$, полученная с помощью модели MARS, может быть выражена следующим образом:

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

$n_1 = \#$ of first degree terms

$n_2 = \#$ of second degree terms

$n_3 = \#$ of third degree terms

10.5 Количество диаграмм, необходимых для построения графика модели

Сколько графиков потребуется для визуализации всех функций в модели? Ответ зависит от переменных, их взаимодействий, а также от того, являются ли они факторными или нефакторными переменными.

VE может ограничивать переменные и взаимодействия в конфигурационном файле Excel для R/earth, влияя на то, какие взаимодействия допускаются. Практичный подход заключается в том, чтобы позволить R/earth оценить все возможные взаимодействия и выбрать те из них, которые вносят значительный вклад в оценку цены продажи, одновременно проверяя наличие таких проблем, как переобучение или коллинеарность. Некоторые взаимодействия могут быть запрещены, что повлияет на взаимодействия более высокой степени (например, если пара исключена из взаимодействия третьей степени, исключается вся комбинация).

В случае факторных переменных (например, «Зона MLS» с 12 значениями, взаимодействующей с «GLA») построение графиков становится более сложным. Для этого может потребоваться 12 графиков — по одному на каждую зону. Если другой фактор, такой как «Архитектурный стиль» (6 значений), взаимодействует с «Зоной MLS» и «GLA», может потребоваться до $12 \cdot 6 = 60$ графиков. Однако R/earth редко признает все комбинации значимыми; как правило, построение графиков оправдано только для нескольких ключевых взаимодействий (например, определённых зон и стилей), выявленных моделью.

Вот упрощённое руководство по количеству необходимых графиков:

1. Функции одной переменной (без факторов): 1 график.
2. Функции одной переменной (с факторами): 1 гистограмма, объединяющая все значения факторов.

3. Базовая функция: график не требуется.
4. Взаимодействия второй степени (без факторов): 1 график (например, двумерная тепловая карта с цветовой кодировкой цены продажи).
5. Взаимодействия второй степени (1 фактор): по 1 графику на каждое значение фактора в модели (например, 2 значения = 2 графика).
6. Взаимодействия второй степени (2 фактора): 1 график с матрицей комбинаций факторов (например, 3 области $\times$ 2 стиля = 6 ячеек).
7. Взаимодействия 3-й степени (без факторов): несколько графиков (например, 4 графика для количества ванн = 1, 2, 3, 4).
8. Взаимодействия 3-й степени (1 фактор): по одному графику на каждое значение фактора в модели.
9. Взаимодействия третьей степени (2 фактора): по 1 графику на каждую комбинацию значений факторов в модели.
10. Взаимодействия третьей степени (3 фактора): по 1 графику на каждое значение фактора с наименьшим количеством различных значений, с использованием цветовой матрицы для двух остальных.

Для более сложных сценариев могут потребоваться передовые методы построения графиков. Существуют и другие варианты с, вероятно, более сложными графиками.

Пример полной модели, сгенерированной с помощью MARS, можно найти по адресу [Figure 1](#).

11 Характеристики данных

В следующем разделе рассматривается система MARS. Система MARS полностью зависит от получаемых данных, и большая часть этих данных поступает из какой-либо службы MLS (Multiple Listing Service) или от налоговых оценщиков. Однако данные могут проходить через других поставщиков, которые занимаются их обработкой и улучшением, а затем продают с определенной наценкой. В некоторых регионах система достаточно надежна, в других — не столь надежна. В критических случаях инспектор по недвижимости вынужден полагаться на собственные наблюдения и измерения, точность которых может быть поставлена под сомнение.

Инженер по оценке должен понимать, что отражают данные и насколько точно они отражают то, что должны отражать. Это влечет за собой целый ряд вопросов:

- Как определить, что переменная является измерением?
- Если переменная не является измерением, то чем же она может быть?
- Какую пользу может принести регрессионный анализ, если получаемые им данные неточны?

Большая часть данных, которые мы получаем из списков MLS и от налоговых органов для проведения регрессионного анализа, представляет собой числовые данные, отражающие измерения или подсчеты. Часть этих данных имеет логический характер — «Истина» или «Ложь».

11.1 Цель программы «Residual»

Остаток служит косвенным показателем добавленной стоимости, создаваемой переменными, не включенными в регрессионную модель. На языке гедонической литературы эти неизмеренные характеристики называются латентными переменными: переменными, которые влияют на цену продажи, но не измеряются напрямую, в большинстве случаев из-за отсутствия практической возможности объективно их измерить. Их значения приходится выводить косвенно, и на практике значение латентной переменной часто оценивается на основе субъективного суждения — в методе RCA это суждение опирается на ранжирование остаточных значений. Ранжируя объекты недвижимости по их остаточным значениям или остаточным значениям на квадратный фут, мы получаем объективный рейтинг от объектов с наименьшей привлекательностью до объектов с наибольшей. Этот метод ранжирования эффективен при условии, что модель MARS (многомерные адаптивные регрессионные сплайны) разрабатывается квалифицированным аналитиком, глубоко разбирающимся в местном рынке. Ниже приведены некоторые дополнительные требования:

- Качество данных и модели: Правильно построенная регрессионная модель и высококачественные данные являются необходимыми условиями для получения достоверных остатков. В районе залива Сан-Франциско моя цель состоит в разработке модели с коэффициентом R^2 примерно 75–80% и коэффициентом R^2 перекрёстной проверки (CV R^2) 55–70%.
- Избегайте перегонки, ограничивайте R^2 : Инженер по оценке должен стремиться к максимальному значению R^2 около 80 % на большинстве зрелых рынков в районах с высокими ценами, таких как район залива Сан-Франциско, поскольку большая часть конкурентной стоимости жилой недвижимости заключается в ее субъективной привлекательности,

которая не поддается количественному измерению и, следовательно, не может быть напрямую включена в регрессионные модели. Значение R^2 , значительно превышающее 80 %, может указывать на переобученную модель, что требует дополнительного тщательного анализа. Однако следует учесть: существуют рыночные зоны со значениями R^2 выше 90 % без переобучения, а также районы, которые сложно оценить, где максимальное достижимое значение может составлять 50–60 % или ниже.

- Высокие значения R^2 в определенных районах: Однако существуют районы, такие как жилые микрорайоны или другие однородные жилые зоны, где более высокие значения R^2 встречаются на удивление часто из-за однородности объектов недвижимости.

11.2 Оценки состояния и качества в качестве приблизительных регрессоров

Шкалы оценки состояния и качества Fannie Mae (C1–C6 и Q1–Q6) заслуживают особого внимания. Эти шкалы слишком грубы, чтобы рассматривать их как подлинные измеряемые характеристики, а поскольку они присваиваются оценщиком, на практике они представляют собой субъективные оценки — латентные переменные сами по себе. Целесообразно явно различать два уровня: грубую оценку оценщика, используемую в качестве регрессора — «состояние-грубое» (condition-1) и «качество-грубое» (quality-1), с одной стороны, и мелкозернистые латентные переменные, восстановленные из остатка: «состояние-мелкозернистое» (состояние-2) и «качество-мелкозернистое» (качество-2), — с другой. VE может принимать грубозернистые оценки оценщика в качестве приближений и проводить регрессию по ним, получая приблизительные вклады в стоимость, которые полезны в качестве начальных значений для тонкозернистых латентных переменных. В этом случае протокол заключается в том, чтобы вычесть оцененные с помощью регрессии вклады в стоимость этих двух крупнозернистых переменных из оценки модели — удалив их из измеренных характеристик в разделе регрессии таблицы продаж — и добавить их обратно в остаток. Затем они удаляются из остатка и присваиваются соответствующим латентным переменным с тонкой детализацией: `res_condition1` и `res_quality1`. Дополнительная информация, извлеченная из публичных комментариев MLS (`public_remarks`) и связанных списков строк, затем используется для микроуточнения значений этих латентных переменных. В результате вклады значений с высокой степенью детализации заменяют вклады значений с низкой степенью детализации: грубые оценки служат лишь основой для приближенной модели, а значения, в конечном итоге представленные для параметров «состояние» и «качество», являются значениями с высокой степенью детализации, полученными на основе остатков.

11.3 Ошибки в данных

Судя по комментариям в социальных сетях, данные Калифорнийской системы совместного листинга (MLS) в целом отличаются более высоким качеством, чем данные по многим другим регионам США. Как правило, можно ожидать, что данные MLS будут хорошего качества в большинстве развитых мегаполисов, но более низкого качества — в сельских районах.

Хотя в данных по разным причинам иногда встречаются ошибки, при наличии достаточного объема данных эти ошибки, как правило, должны взаимно компенсироваться и оказывать минимальное влияние на средние значения, используемые в ценовых моделях.

Однако в отдельных районах могут возникать систематические искажения. Например, рассмотрим район, где многие дома изначально строились с полноценным вторым этажом, но некоторым покупателям предоставлялась возможность убрать комнаты на втором этаже, чтобы создать сводчатые потолки в одной или нескольких комнатах первого этажа, что

значительно уменьшало общую жилую площадь (GLA). По тем или иным причинам эти изменения впоследствии не были отражены в документации оценщика. В результате в системе MLS может указываться площадь в 2 400 квадратных футов, тогда как фактическая измеренная площадь составляет всего 2 000 квадратных футов. В таких случаях данные MLS следует исправить перед выполнением регрессии MARS (многомерные адаптивные регрессионные сплайны). Важно отметить, что получение этой информации по аналогам может потребовать определенных исследовательских усилий. Несоответствия следует отметить для использования в будущем.

Учитывая вышесказанное, можно обоснованно предположить, что в районах с высококачественной системой MLS погрешности, проникающие в остаток, минимальны по сравнению с влиянием субъективных и неизмеримых характеристик, таких как планировка, состояние и качество. Это является важным фактором при оценке точности метода RCA. Например, если значение коэффициента CQA (R^2) для данной рыночной зоны составляет 80%, то примерно 20% цены приходится на остаток CQA. Погрешность $\pm 10\%$ при отнесении объекта оценки к ранжированию по остатку, следовательно, приводит к отклонению в размере $\pm 10\% \cdot 20\% = \pm 2\%$ error in the final value conclusion. This is how you improve accuracy! In many cases, accurate value estimates well below $\pm 1\%$.

Поэтому рассмотрение остатков в регрессии MARS в качестве косвенных показателей совокупной стоимости немоделируемых переменных является обоснованным при условии, что специалист по оценке обладает достаточным опытом работы на местном рынке, чтобы определить, где и почему возникают системные погрешности измерения. Такой подход гарантирует, что анализ точно отражает фактическую стоимость объектов недвижимости, учитывая ограничения и нюансы данных по недвижимости.

Пример из практики: измерение неизмеренных факторов в Си-Ранч

Когда достижимый уровень « R^2 » в рыночной зоне оказывается ниже того, чего можно было бы ожидать, исходя из кажущейся однородности рынка, данная концепция предполагает выявление и измерение недостающей переменной, а не принятие этого ограничения. Сообщество «Sea Ranch» в округе Сонома наглядно иллюстрирует этот подход.

Конкретный пример реакции «оценивай, но не соглашайся». Сообщество «Sea Ranch» в округе Сонома — это прибрежный жилой комплекс, где расстояние до океана является основным фактором, определяющим стоимость: дом, расположенный в нескольких сот футов от утеса, стоит значительно дороже, чем дом, расположенный на четверть мили вглубь суши, при прочих равных условиях. В данных MLS по этому жилому комплексу расстояние до океана не учитывается в качестве переменной. Первоначальная модель MARS, проведенная на доступных переменных MLS, продемонстрировала более низкую объясняющую способность, чем следовало бы ожидать, учитывая кажущуюся однородность рынка, при этом структура остатков не позволяла достаточно четко упорядочить объекты недвижимости по каким-либо видимым в данных характеристикам — это было сигналом того, что в модели отсутствовал важный фактор, а не того, что рынок действительно был «шумным».

Решением стала прямая съемка. Автор получил карту района, на которой было показано расположение объектов недвижимости относительно береговой линии, и измерил расстояние до океана для нескольких сотен домов. Полученные значения расстояний были добавлены в качестве переменной в набор данных MLS, после чего регрессионный анализ MARS был проведен заново. Новая переменная была признана моделью значимой, объясняющая способность существенно возросла, а ранжирование остатков впоследствии упорядочило объекты недвижимости по признакам, соответствующим остальным субъективным характеристикам — виду, состоянию, качеству — а не по скрытому сигналу расстояния до океана, который

ранее искажал результаты. Вопрос высоты над уровнем моря был решен аналогичным образом в районах, расположенных на склонах холмов, где переменная MLS отсутствует или является недостоверной.

Методологический принцип носит общий характер. Когда достижимый показатель « R^2 » опускается ниже уровня, который должен обеспечиваться видимой однородностью рынка, первая гипотеза заключается в том, что в модели отсутствует важная переменная, и надлежащим ответом является выявление этой переменной и сбор данных непосредственно на местах. Источниками информации служат карты населенных пунктов, данные ГИС, отчеты налоговых инспекторов, отчеты строительных ведомств, спутниковые снимки и результаты выездных осмотров. Затраты на такие измерительные работы реальны, но ограничены; как только переменная собрана для данной рыночной зоны, её можно использовать в будущих заданиях в этой зоне. RCA легко интегрирует новые переменные, поскольку данная методология не зависит от того, какие именно переменные входят в регрессию MARS — структура остатков указывает VE, отражает ли новая переменная реальный сигнал, и вклады переменных подвергаются соответствующим корректировкам.

Альтернативный подход — принятие низкого коэффициента « R^2 » и подготовка отчёта с более широкими диапазонами погрешности — имеет обоснованность в тех случаях, когда отсутствующие факторы действительно невозможно измерить (индивидуальные предпочтения покупателей, кратковременные колебания настроений на рынке, специфические обстоятельства сделок), но не в тех случаях, когда они просто не были учтены. Первым шагом эксперта по оценке (VE) при работе со слабой моделью должно быть установление того, относится ли отсутствующий сигнал к последнему случаю.

11.4 Что такое точность?

Для обеспечения абсолютной точности оценки необходимо:

1. Соответствие рыночной стоимости: гипотетическая сделка купли-продажи объекта оценки на 100% соответствует приведенному определению рыночной стоимости
2. Разработка точных и надежных моделей: Разработана высококачественная модель MARS с коэффициентом R^2 около 75–80 % и коэффициентом CVR^2 примерно 55–70 % в зависимости от местоположения.
3. Качественный анализ ошибок данных: все вероятные ошибки данных по измеряемым переменным проанализированы и устранены, при этом их влияние на остатки минимально.
4. Учет системных отклонений: Любые системные отклонения в данных должны быть учтены и скорректированы.
5. Анализ рангов остатков: Убедитесь, что свойства рангов остатков следуют разумной последовательности характеристик от самого низкого до самого высокого ранга. Аномалии, такие как продажи из наследственной массы, продажи с недостаточной площадью и аукционные продажи, должны быть изучены, объяснены и, возможно, исключены.
6. Определение места объекта оценки: на основе глубокого понимания ранжирования остаточных значений объект недвижимости должен быть точно позиционирован. Необходимо выявить недвижимость с значительно более высокой привлекательностью и другую — со значительно более низкой, а рассматриваемая недвижимость

разместить между ними и присвоить ей балл. Затем следует постараться максимально сузить разрыв в рейтинге между объектами недвижимости с более высокой и более низкой привлекательностью.

7. **Обоснование:** Никогда не жертвуйте обоснованием ради точности. Оценщик должен быть способен объяснить различия в стоимости между объектом оценки и аналогами, по крайней мере, человеку с разумным уровнем интеллекта. Например, хорошую модель MARS с оптимальными показателями R^2 и CVR^2 можно использовать для объяснения вклада в стоимость измеряемых характеристик, в то время как разбивка остатков, выполненная инженером по оценке, может служить для объяснения вклада в стоимость других характеристик.
8. **Переобучение:** Следует избегать переобучения моделей, не используя характеристики объектов недвижимости, которые могут идентифицировать — как по отдельности, так и в сочетании — один или несколько конкретных объектов. В частности, следует избегать использования факторных переменных, значения которых могут указывать на небольшое количество конкретных объектов. Факторная переменная, идентифицирующая районы в значительных наборах данных, является уместной, но та, которая идентифицирует объекты недвижимости с бассейнами в районе, где таких бассейнов очень мало, должна использоваться с осторожностью. Использование характеристик, идентифицирующих уникальные объекты недвижимости, таких как точные ГИС-координаты, номера земельных участков или адреса, категорически запрещено.

Если коэффициент детерминации R^2 составляет 80 %, то можно ожидать, что примерно 20 % значения объекта оценки обусловлено остатком. Однако это лишь приблизительная оценка. Из кривой «CQA-остаток» известно, что большая часть распределения остатков приходится на нижние и верхние крайние значения CQA. Исходя из кривой «CQA-остаток» и крайних значений наклона для низких и высоких значений CQA, можно ожидать, что оценка остатка в этих областях будет менее точной. Однако это компенсируется относительной редкостью объектов недвижимости в этих зонах и более значительной разницей в привлекательности, которую мы наблюдаем между соседними объектами недвижимости. Таким образом, хотя мы и можем говорить о влиянии диапазона значений $\pm 5\%$ в средней зоне, в крайних зонах нас будет больше интересовать несколько более узкий диапазон значений $\pm 2 - \pm 3\%$

Расчет точности RCA, как правило, представляет собой некоторую сложность из-за нелинейности кривой «CQA-остаточная величина». Однако в каждом конкретном случае это возможно, при условии наличия адекватного ранжирования баллов по остаткам. Диапазон баллов CQA составляет от 0,0 до 10,0. При наличии качественной модели MARS ваш рейтинг остаточных значений должен демонстрировать чёткую последовательность от наименьшей до наибольшей привлекательности, и вам не должно составить труда оценить балл CQA с точностью до $\pm 5\%$ от общего диапазона или ± 0.5 . Если в регрессии участвуют 100 объектов недвижимости, это означает, что рассматриваемый объект оценки будет точно отнесён к одной из групп по 10 объектов недвижимости. Итак, обращаясь к данным о продажах в Берлингейме, приведенным в таблице 3, рассчитаем приблизительную точность (или погрешность) для CQA, оцененного VE, равного 7,5, 5,1 и 1,9, где его заключение о стоимости объекта оценки составляет \$ 3 000 000. Обратите внимание, что, поскольку в таблице 3 представлена выборка из нескольких сотен объектов недвижимости, вам потребуется интерполировать остатки там, где это необходимо (при условии линейности). Для ясности: мы предполагаем, что в каждом случае VE с высокой степенью уверенности утверждает, что его рейтинг объекта оценки находится в диапазоне от ± 5 объектов недвижимости в рейтинге остатков, включающем всего 100 объектов недвижимости.

CQA= 7,5±0.5				
CQA	Остаток	Абсолютная погрешность	Погрешность эксперимента	Гипотетическая п
8,00	\$ 257 112			
7,50	\$ 197 113	\$ 59 999		
7,00	\$ 154 615	\$ 42 498	±\$51249	\$ 3 000

CQA=5,1±0.5				
CQA	Остаточная ошибка	Абсолютная разница	Статистическая погрешность	Гипотетич
5,60	\$ 50 911			
5,10	\$ 19 559	\$ 31 352		
4,60	-\$ 28 207	\$ 47 766	±\$39559	\$

CQA= 1,9±0.5				
CQA	Остаток	Абсолютная разница	Ошибка эксперимента	Гипотетическая цена прод
2,40	-\$ 228 411			
1,90	-\$ 276 381	\$ 47 970		
1,40	-\$ 326 938	\$ 50 557	±\$49264	\$ 3 000 000

Наконец, единственный объективный способ проверки точности заключается в разработке строгого протокола (о котором в данной статье говорится лишь отчасти и в общих чертах) и контроле его соблюдения. Это необходимо, поскольку фактические цены продажи могут отклоняться от рыночной стоимости из-за несовершенств рынка. Сравнение оценок и последующих продаж может дать некоторое представление о точности. Однако сложность и нестабильность рынка недвижимости требуют привлечения экспертов-оценщиков или высококвалифицированных инженеров-оценщиков для проведения оценок стоимости на основе установленных стандартов и протоколов. Эти стандарты и протоколы должны быть достаточно строгими, чтобы гарантировать, что компетентные инженеры-оценщики независимо друг от друга придут к практически одинаковым выводам относительно стоимости. В большинстве случаев, когда имеются качественные данные, эта цель должна быть достижимой.

11.5 Рассматриваемый объект оценки как аномалия

Поскольку для оценки рыночной стоимости объекта недвижимости требуется анализ недавних сделок купли-продажи, вполне уместно задаться вопросом: насколько точно недавние сделки отражают различные особенности старого объекта оценки, не прошедшего модернизацию и ремонт. Это особенно актуально для района залива Сан-Франциско, где недавние сделки купли-продажи, как правило, касаются домов, которые были обновлены с целью повышения их продажной цены. Таким образом, среднее качество и состояние аналогов, как правило, превосходят более старые дома, которые довольно долгое время не выставлялись на продажу. Эта разница в стоимости должна быть учтена при оценке остатка объекта оценки, или, другими словами, его балла CQA. Это важный фактор, который необходимо учитывать с точки зрения качества и точности данных при оценке объекта оценки.

11.6 Важность CVR2 при работе с непредвиденными данными

Регрессия MARS основана на многократных итерациях алгоритма R/earth, применяемого к случайно выбранному подмножеству объектов недвижимости, которое, как правило, составляет от 80 % до 90 % набора данных MLS, называемого обучающим набором. Остальные 10

% – 20 % объектов недвижимости служат в качестве тестового набора. На этом этапе на основе обучающего набора данных создаётся модель MARS, которая затем используется для прогнозирования цен продажи объектов недвижимости из тестового набора данных; полученные результаты сравниваются с фактическими ценами продажи. Данная методология признана специалистами по анализу данных наиболее эффективной стратегией для разработки надёжной модели, способной прогнозировать цены продажи объектов недвижимости, данные о которых ещё не были собраны.

Однако основная ценность показателя CVR2 заключается не в том, что он отражает фактическую точность регрессионной модели в отношении реальных продаж в рыночной зоне, а в том, что он позволяет оптимизировать прогнозные возможности модели в отношении гипотетических продаж недвижимости на определенную дату оценки. Данные объекты недвижимости не включены в набор обучающих данных и могут демонстрировать новые или неожиданные комбинации характеристик, отсутствующих как в данных об объектах недвижимости из набора обучающих данных, так и в данных из набора тестовых данных.

Учитывая это, как мы можем оценить цену продажи объекта недвижимости, который не обновлялся в течение последних 15 лет, если предположить, что он был выставлен на продажу, например, за месяц до даты оценки? Модель, демонстрирующая высокий показатель CVR2, с большей вероятностью даст наиболее точную оценку, чем модель, не прошедшая такую многократную перекрёстную проверку. Также крайне важно, чтобы набор данных включал достаточное количество аналогичных объектов недвижимости с различной степенью обновления, ремонта и модернизации.

В конечном итоге, если у нас есть надёжная модель, которая точно прогнозирует вклад различных характеристик объектов недвижимости в нашем наборе данных по аналогичным объектам, мы сможем эффективно определить место нового объекта оценки в рейтинге этих объектов, составленном моделью на основе остатка. Такое определение места позволяет нам присвоить данному объекту соответствующий балл CQA и остаток.

12 Остатки

Как следует из названия «Метод остаточных ограничений», остатки занимают центральное место в этой методологии. Под остатком здесь понимается разница между чистой ценой продажи объекта недвижимости и ценой продажи, прогнозируемой с помощью регрессионного анализа.

Положительный остаток свидетельствует о том, что объект недвижимости был продан по цене, превышающей его оценочную стоимость, рассчитанную на основе типичных измеримых характеристик, таких как общая жилая площадь (GLA) и площадь участка, что позволяет предположить, что он, возможно, обладает более высокой привлекательностью, чем в среднем. Отрицательный остаток указывает на то, что объект недвижимости был продан по цене ниже оценочной, что, возможно, свидетельствует о его меньшей привлекательности или наличии других негативных характеристик, не учтенных в оценке.

Учитывая, что у более просторных домов, скорее всего, остатки выше, для более обоснованного сравнения мы могли бы использовать показатель остатков на квадратный фут, рассчитанный на основе общей арендуемой площади (GLA). Более подробное рассмотрение этого вопроса будет посвящено другой статье.

12.1 Рейтинг & Балл

После построения модели MARS (многомерные адаптивные регрессионные сплайны) для измеряемых переменных для каждого объекта недвижимости рассчитываются остатки. Эти остатки представляют собой разницу между наблюдаемыми чистыми ценами продажи и ценами продажи, прогнозируемыми моделью. Следует создать дополнительный столбец для показателя «Остаток/SF», то есть остатка, деленного на жилую площадь объекта недвижимости, исходя из того, что, как правило, остаток для конкретных объектов недвижимости тем значительнее, чем больше жилая площадь. В некоторых районах могут лучше работать другие статистические показатели, но это тема для другого разговора.

Assumption 7: “Ranking by Residual”

В данной статье под «ранжированием по остаточной стоимости» следует понимать ранжирование по остаточной стоимости, по остаточной стоимости, деленной на SF, либо ранжирование по какой-либо другой функции остатков, которую инженер по оценке сочтет наиболее подходящей для рыночной зоны. Опыт автора в районе залива Сан-Франциско показывает, что показатель «остаточная стоимость/SF» обычно является наиболее надежным, особенно при оценке объекта оценки с необычно малой или большой общей жилой площадью (GLA).

Далее объекты недвижимости сортируются по величине их остатков — от самых отрицательных к самым положительным. Затем для каждого объекта рассчитывается балл «Состояние-Качество-Привлекательность» (CQA) в диапазоне от 0,0 до 10,0. Этот балл рассчитывается следующим образом:

Балл CQA соответствует процентильному рангу остаточной стоимости объекта недвижимости среди всех объектов недвижимости. В частности, балл рассчитывается путём определения процента объектов недвижимости, у которых остатки ниже, чем у данного объекта, и деления этого процента на 10. Например, если остаток объекта недвижимости выше, чем у 50 % объектов, то его балл составит 5,0.

Из-за округления одно свойство может получить балл 10,0, что соответствует собственности с наибольшим остатком. Однако идеальный балл 10,0 теоретически невозможен, поскольку,

по логике вещей, ни одна собственность не может иметь остаток, превышающий остаток 100%-х собственностей, включая само эту собственность. С другой стороны, минимально возможный балл 0,0 возможен, поскольку по крайней мере у одного объекта недвижимости нет других объектов с более низкими остатками. Обратите внимание, что из-за округления балл 0,0 может быть присвоен более чем одному объекту недвижимости, если сравнивается большое количество объектов.

Таким образом, балл CQA служит интуитивно понятным критерием того, насколько характеристики объекта недвижимости соответствуют ожиданиям модели, причем более высокие баллы свидетельствуют о лучшем соответствии прогнозам модели.

Важно признать, что балл CQA можно рассматривать лишь как косвенный индикатор привлекательности аналога недвижимости в том случае, если условия продажи соответствуют его рыночной стоимости. Инженер по оценке несет ответственность за установление того, включает ли остаточная стоимость элементы, завышающие или занижающие рыночную стоимость, путем предоставления исчерпывающего объяснения и тщательного расследования сделки купли-продажи. Хотя значительные продажи по ценам, значительно превышающим или занижающим рыночную стоимость, происходят нечасто, они могут иметь место в таких обстоятельствах, как вынужденная продажа, наследственное дело, недееспособность, сговор и т. д.

В таблице 3 представлен полуреалистичный выдержка из данных MLS вместе с остатками и баллами CQA — как в общем, так и в расчете на квадратный фут. Общее количество объектов недвижимости, проанализированных в рамках данного исследования, составило примерно 600. Следует отметить, что некоторые переменные, включенные в регрессионный анализ, в таблице не отображены. Другими словами, представленные остатки получены после того, как в рамках регрессии MARS было учтено влияние других переменных, таких как дата продажи.

Таблица 3: Аналоги продажи, отсортированные по остаткам

Цена продажи	Расчетная цена продажи	Остаток	CQA	Остаток/SF	CQA SF	GLA	Дата продажи
\$ 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 Кривая CQA-остаточная величина

График зависимости CQA от остаточной величины/SF для того или иного района неизменно приведет к построению характеристики CQA-остаточная величина, представленной по адресу [Figure 2](#). Данный конкретный график отражает данные по городу Берлингем (штат Калифорния) примерно за 2022 год. Аналогичные кривые, как правило, наблюдаются при построении графиков, где по оси абсцисс отложены баллы CQA, а по оси ординат — значения остаточной стоимости; однако я предпочитаю использовать показатель «остаточная стоимость на квадратный фут» (Residual/SF) из-за его значимости для контекстуализации показателя CQA с учётом общей жилой площади (GLA) объекта оценки. Умножив показатель «Остаточная стоимость/SF» на GLA, можно преобразовать показатель CQA объекта недвижимости в соответствующую остаточную стоимость.

Инженер-оценщик должен тщательно проанализировать, оказывает ли размер объекта недвижимости, а именно общая жилая площадь (GLA), существенное влияние на величину остатков. Эта оценка имеет решающее значение, поскольку от нее зависит выбор наиболее точного метода расчета стоимости недвижимости. Например, высококлассная кухонная техника и роскошная сантехника в ванных комнатах могут существенно повысить остаточную стоимость дома. Однако в районах, где размеры и планировка кухонь и ванных комнат, как правило, однородны, эти усовершенствования могут не иметь сильной корреляции с GLA. Их влияние на остаточную стоимость может зависеть не столько от занимаемой ими площади, сколько от качества, которое они представляют.

С другой стороны, дорогостоящие элементы экстерьера и материалы, которые заметно повышают привлекательность недвижимости с улицы, могут иметь более прямую связь с ее размером. В более просторных домах можно реализовать более сложные элементы экстерьера, такие как обширное благоустройство территории, высококачественные кровельные материалы или просторные патио, что естественным образом повышает общую привлекательность недвижимости и, как следствие, ее остаточную стоимость.

Поэтому специалистам по оценке крайне важно учитывать эту динамику при анализе остатков. Они должны не только оценивать внутреннюю стоимость, создаваемую высококачественной отделкой и дополнительными элементами, но и понимать, как эти улучшения взаимодействуют с физическими характеристиками объекта недвижимости. Такой комплексный подход гарантирует, что результаты оценки точно отражают как качественные, так и количественные характеристики объекта недвижимости.

Следовательно, рекомендуется построить графики как кривых «CQA-Residual», так и «CQA-Residual/SF» и проанализировать различия между ними, которые могут указывать на закономерности в конкретной рыночной зоне. Инженер по оценке может поручить своей программе рассчитать оба вида остатков и взять их средневзвешенное значение для расчета остатка, который будет присвоен объекту оценки. Это может иметь значительное влияние, а может и нет.

12.3 Характеристики кривой CQA

Кривая «Состояние-Качество-Привлекательность» (CQA) представляет собой сложную модель, позволяющую понять, как стоимость недвижимости колеблется в зависимости от привлекательности объекта, его состояния и динамики рынка в конкретном районе. Эта модель учитывает взаимодействие экономических и физических факторов, влияющих на рынок недвижимости, отражая все нюансы оценки недвижимости в различных сегментах рынка.

12.3.1 Анализ компонентов кривой CQA

1. Экспоненциальный спад в левой части кривой:

По мере снижения балла CQA с уровня около 2 до 0 объекты недвижимости демонстрируют характеристики быстрого износа. Эта часть кривой хорошо описывается функцией экспоненциального спада, что подчеркивает, насколько быстро объект недвижимости может утратить свою привлекательность и стоимость без надлежащего обслуживания и обновлений. Физическое изнашивание, такое как гниение дерева, выцветание красок и ржавление металла, ускоряет этот процесс, особенно в домах, которые не подвергаются регулярному техническому обслуживанию.

2. Линейный рост в среднем диапазоне:

Основная часть рынка, как правило, относится к этой категории, в которой привлекательность растёт почти линейно при оценках от 2 до 8 баллов. В этот раздел входят объекты недвижимости, которые, как правило, поддерживаются в надлежащем состоянии, но не обязательно отличаются исключительными качествами или выгодным расположением. Он представляет широкий средний сегмент жилья, охватывающий большую часть рынка жилья, где дома предлагаются по конкурентоспособным ценам с учетом их стандартных характеристик и общего состояния.

3. Экспоненциальный рост в правой части кривой:

Здесь кривая растёт экспоненциально в диапазоне баллов от 8 до 10, что отражает сегмент рынка с наиболее высокой привлекательностью. Объекты недвижимости в этом сегменте часто принадлежат состоятельным людям, которые не только вкладывают средства в безупречное содержание объекта, но и в улучшения, значительно повышающие его привлекательность и стоимость. Экспоненциальный рост отражает дефицит таких высококачественных объектов недвижимости и высокий спрос со стороны состоятельных покупателей. При этом следует понимать, что дом по более низкой цене может получить очень высокую оценку с точки зрения привлекательности, в то время как большой и дорогой дом может получить низкую оценку. Эта динамика выявляется благодаря грамотному использованию системы MARS и обычно остается незамеченной большинством оценщиков.

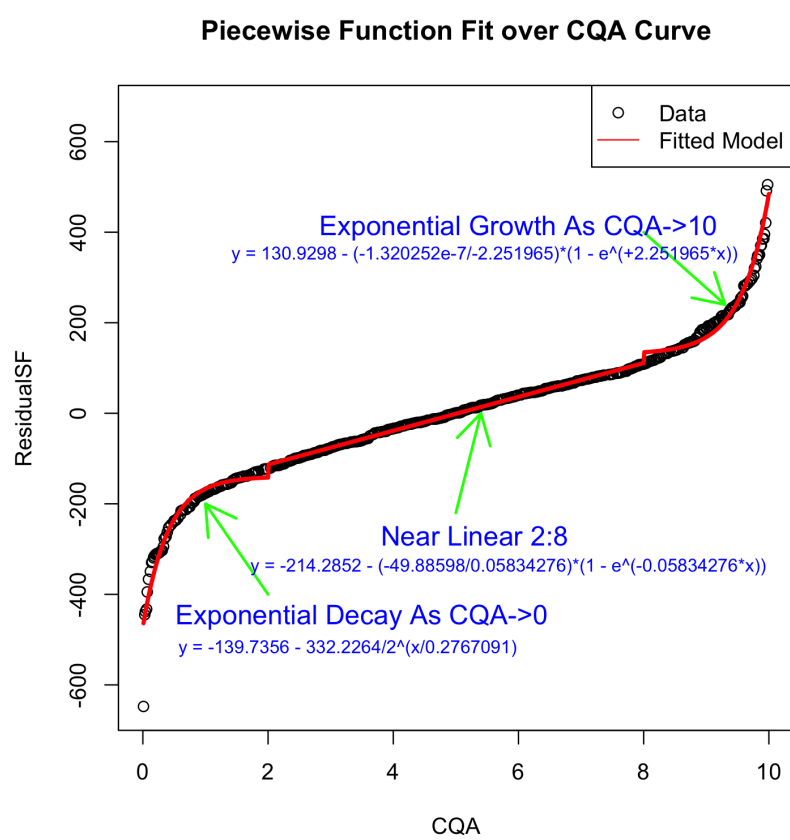
12.3.2 Динамика рынка и социальные последствия

Кривая CQA косвенно отражает ситуацию с относительным благосостоянием и экономическим неравенством в различных регионах. В таких местах, как округ Сан-Матео, разница в стоимости недвижимости между районами вроде Берлинггейма и более состоятельными районами, такими как Хиллсборо, Вудсайд или Атертон, свидетельствует о значительных социально-экономических различиях. Это неравенство сказывается на всех аспектах — от содержания недвижимости до рыночной стоимости — и, в конечном итоге, влияет на демографический состав покупателей и тенденции рынка недвижимости.

Кроме того, история об относительно состоятельном человеке, инвестирующем в высококачественный дом в районе среднего ценового сегмента, где ему комфортно и где он по разным причинам предпочел бы жить, подчеркивает еще один важный аспект экономики недвижимости — концепцию «избыточного строительства». Такие объекты недвижимости, несмотря на свою привлекательность, могут не найти местных покупателей, способных заплатить ту цену, которую они могли бы получить в более высокой рыночной зоне.

Кривая CQA-Остаток является характеристикой для каждой рыночной зоны, и её не следует упускать из виду при проведении анализа.

Рис. 2: CQA — остаточная характеристическая функция



12.4 Объект оценки «Остаточные величины»

В отличие от аналогов, рассматриваемый объект недвижимости не имеет объективно определенной остаточной стоимости; она должна быть оценена инженером-оценщиком (VE). Такая оценка вносит определенную долю субъективности в протокол применения Метода остаточных ограничений (RCA), что является критическим моментом, в котором личное суждение играет ключевую роль.

Для оценки бала «Состояние-Качество-Привлекательность» (CQA) эксперт по оценке (VE) тщательно анализирует объекты недвижимости, отсортированные по рейтингу, выявляя заметные закономерности в рейтинговой таблице. Этот анализ часто включает подробный просмотр фотографий из базы данных MLS, связанных с каждым объектом. Исчерпывающий анализ нецелесообразен при работе с большим набором данных, например, с 600 объектами. Вместо этого VE применяет практическое суждение, чтобы приблизительно определить вероятное положение объекта оценки в общем рейтинге. Объекты недвижимости, требующие капитального ремонта, или «объекты для ремонта», как правило, занимают места в нижней части рейтинга, тогда как объекты с превосходной отделкой и привлекательным внешним видом, как правило, занимают места в верхней части рейтинга.

Затем VE применяет более целенаправленный подход, «перемещаясь» в пределах рейтинга для выявления аналогов, которые максимально соответствуют привлекательности объекта оценки. По мере сужения круга поиска VE уделяет больше внимания сходствам между объектами недвижимости, особенно в таких аспектах, как качество, состояние, эстетика дизайна (включая такие элементы, как покраска, деревянные и каменные отделочные элементы, а также окна), функциональность и другие значимые характеристики.

Крайне важно понимать, что оценки VE в основном основываются на рейтингах, сформированных рынком. Это гарантирует, что субъективные элементы процесса опираются на рыночные реалии, что позволяет добиться справедливой и точной оценки объекта оценки.

12.5 Выводы по оценке RCA

12.5.1 Общие соображения

Определение показателя CQA имеет решающее значение для установления ориентировочной остаточной стоимости объекта недвижимости, поскольку эти факторы неразрывно связаны между собой. Хотя оценка объекта недвижимости возможна и без учета показателя CQA, этот показатель предоставляет ценную информацию, позволяя увидеть, как данный объект позиционируется по сравнению с другими объектами на рынке в рыночной зоне. Например, балл CQA, равный 3,2, указывает на то, что данный объект недвижимости обладает большей привлекательностью, чем 32 % аналогичных объектов в окрестностях. Если имеются возражения против присвоенного бала, стороны спора должны ознакомиться с рейтингами и обосновать любые предлагаемые корректировки на основе фотографических доказательств или других существенных данных. Если система ранжирования надежна, а инженер-оценщик (VE) точно определил место объекта недвижимости в этой системе, маловероятно, что возникнут значительные расхождения.

На этапе подготовки отчета, как правило, отбирается от 6 до 12 актуальных и релевантных аналогов для обоснования вывода об оценке. Проводится анализ различий между этими аналогами и объектом оценки, при этом каждая переменная корректируется соответствующим образом. Затем эти корректировки применяются к чистым ценам продажи аналогов для получения скорректированных цен продажи. Как показано математически, эти скорректированные цены продажи должны сходиться с расчетной ценой продажи объекта

оценки.

12.5.2 Обзор, аудит, жалобы

После определения расчетной остаточной стоимости объекта оценки можно рассчитать корректировки на остаточную стоимость для всех аналогов. Эти корректировки служат ограничительными условиями для дальнейшей ручной декомпозиции, проводимой оценщиком с учетом конкретных переменных CQA, которые считаются уместными. Важно подчеркнуть, что эти подробные разбивки носят пояснительный характер и не влияют на окончательную оценку, как подтверждается в Доказательстве II.

Что касается жалоб, поданных пользователями, не удовлетворёнными выводом об оценке стоимости, традиционному оценщику, как правило, приходится тратить время на объяснение того, почему тот или иной объект недвижимости с более высокой или более низкой ценой не оказывает ожидаемого пользователем влияния на стоимость. При использовании метода RCA все возможные аналогичные сделки (обычно от 100 до 600) оцениваются с точностью до той же стоимости, что и итоговая цена, поэтому не требуется больших усилий, чтобы продемонстрировать корректировки для какого-либо произвольного объекта недвижимости. Весь процесс RCA автоматизирован и объективен, за исключением оценки остаточного коэффициента объекта оценки. Таким образом, еще одним преимуществом метода RCA является то, что критика в отношении несовершенств субъективного суждения оценщика в основном ограничивается именно этой областью — и с ней можно эффективно и рационально справиться. Кроме того, если предположить, что R2 составляет около 80 %, влияние ошибки ± 10 % при оценке субъективного остатка приводит к погрешности ± 2 % в окончательной оценке стоимости.

Во многих случаях объект оценки настолько идеально вписывается в остаточный рейтинг объектов недвижимости, что можно ожидать, что точность будет находиться в пределах ± 1 %.

Сумма расчетного остаточного значения и расчетной цены продажи объекта оценки, полученной с помощью модели MARS, позволяет сформировать окончательный вывод о стоимости, известный как «вывод о стоимости по методу остаточных ограничений» (RCA). Данный процесс обеспечивает комплексный и прозрачный подход к оценке недвижимости, укрепляя доверие к точности и объективности оценки.

13 Заключение

Метод остаточных ограничений представляет собой концептуальную основу, включающую три существенных вклада: свойство структурной антианкорированности, вытекающее из подгонки модели к всему рынку, а не к небольшому аналогу набора данных (Доказательство I); результат инвариантности, позволяющий использовать разбивку остатков в качестве объясняющего уровня без нарушения вывода о значении (Доказательство II и следующее за ним замечание); и эмпирическая закономерность — характеристическая кривая CQA-остаточных значений — которая, как представляется, повторяется во всех рыночных зонах в характерной экспоненциально-линейно-экспоненциальной форме. Ограничения данной концептуальной рамки столь же существенны: декомпозиция остаточных величин является инвариантной, но не идентифицирующей; этап размещения CQA основан на субъективной оценке VE, которая ещё не была подтверждена в ходе формального исследования межэкспертной согласованности, а утверждения о точности требуют сравнительного анализа с альтернативными подходами.

В данной статье представлены концептуальная основа и протокол. То, чего в ней пока нет, но что входит в более обширную исследовательскую программу, реализуемую в настоящее время, — это исполняемый конвейер с возможностью отслеживания аудита, который обеспечивает реализацию протокола в программном обеспечении с обоснованными настройками по умолчанию. Эта конвейерная система — пакет `earthUI` на CRAN для этапа базисных функций MARS, пакет `glmnetUI` для регуляризованной структурной модели на основе этих базисных функций и пакет `mgcvUI` для пространственного сглаживания — является объектом оценки сопутствующих статей в этом выпуске и служит мостом от RCA как метода к RCA как конечному результату. Разрыв между теорией и практикой в регулировании оценки (Craytor, 2026, настоящий выпуск) объясняет, почему этот «мост» важен: выводы оценки используются в экономике в качестве прогнозных данных о рисках, и система, которая делает заключение о текущей стоимости поддающимся аудиту и воспроизводимым, является необходимым условием для любого добросовестного подхода к такому прогнозному использованию. RCA — это один из шагов в рамках этого более широкого проекта. Это не весь проект, и в данной статье не утверждается, что это так.

Список литературы

- Ambrose, Brent W. и др. (2025). «Do Appraiser and Borrower Race Affect Mortgage Collateral Valuation?» B: 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». B: Valuation Engineer Journal 1.1.
- Friedman, Jerome H. (1991). «Multivariate Adaptive Regression Splines». B: The Annals of Statistics 19.1, с. 1–141.
- Hastie, Trevor, Robert Tibshirani и 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 и др. (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 и Michael LaCour-Little (2020). «AVM versus Appraisal-Based Underwriting in Refinance Mortgages: The Trade-off Between Noise and Bias». B: SSRN Electronic Journal. DOI: [10.2139/ssrn.3603327](https://doi.org/10.2139/ssrn.3603327).
- Kuhn, Max и Kjell Johnson (2016). Applied Predictive Modeling. Corrected at 5th printing. New York: Springer. ISBN: 978-1-4614-6848-6.

- Lundberg, Scott M. и Su-In Lee (2017). «A Unified Approach to Interpreting Model Predictions». B: Advances in Neural Information Processing Systems 30 (NeurIPS 2017), с. 4765—4774.
- Milborrow, Stephen (2024a). Earth: Multivariate Adaptive Regression Splines. R Package. URL: <https://CRAN.R-project.org/package=earth>.
- (2024b). Notes on the Earth Package. Tex. отч. URL: <http://www.milbo.org/doc/earth-notes.pdf>.
- (2024c). Variance Models in Earth. Tex. отч. URL: <http://www.milbo.org/doc/earth-varmod.pdf>.
- Perry, Andre M., Jonathan Rothwell и David Harshbarger (2018). The Devaluation of Assets in Black Neighborhoods. Tex. отч. Brookings Institution.
- Perry, Andre M., Jonathan Rothwell, J. Williamson и др. (2024). How Racial Bias in Appraisals Affects the Devaluation of Homes in Majority-Black Neighborhoods. Tex. отч. Brookings Institution.
- Property Appraisal and Valuation Equity Task Force (2022). Action Plan to Advance Property Appraisal and Valuation Equity. Tex. отч. U.S. Department of Housing и Urban Development.
- Shapley, Lloyd S. (1953). «A Value for n -Person Games». B: Contributions to the Theory of Games, Volume II. Под ред. Harold W. Kuhn и Albert W. Tucker. Annals of Mathematics Studies 28. Princeton: Princeton University Press, с. 307—317.
- Tzioumis, Konstantinos (2016). «Appraisers and Valuation Bias: An Empirical Analysis». B: Real Estate Economics 45.3, с. 679—712. DOI: [10.1111/1540-6229.12133](https://doi.org/10.1111/1540-6229.12133).
- Williamson, Jake и Mark Palim (2022). Appraising the Appraisal: A Closer Look at Divergent Appraisal Values for Black and White Borrowers Refinancing Their Home. Tex. отч. 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: интерактивный интерфейс для пакета earth (MARS)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

William Bert Craytor

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

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

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

Эта статья открытого доступа распространяется на условиях [лицензии Creative Commons Attribution 4.0](#)

Аннотация. earthUI — графический интерфейс пользователя для пакета R earth, реализующего многомерные адаптивные регрессионные сплайны (MARS). Он предлагает три режима работы — общее прогнозное моделирование, оценку недвижимости и анализ рыночной зоны — и проводит пользователя через импорт данных, настройку модели, подгонку, графики g-функций и загружаемые отчёты. В статье описаны требования к формату данных, рабочий процесс моделирования, форматы вывода и полный справочник возможностей earthUI.

Ключевые слова: MARS, многомерные адаптивные регрессионные сплайны, earth, R, графический интерфейс, оценка недвижимости

Publisher's Note (July 2026): This article describes earthUI 0.10.0. Shortly before press, earthUI 0.11.0 absorbed the companion applications glmnetUI and mgcvUI (described in the following two articles), consolidating the MARS, elastic-net, and GAM workflows in a single application. The interface and workflow described here are unchanged by the consolidation.

Содержание

1	Введение	4
1.1	Что такое earthUI?	4
1.2	Один набор, три процедуры	4
1.3	Три режима работы	4
1.4	Недвижимость — специальные материалы	5
1.5	Установка earthUI и vProlog	5
1.6	Начало работы	6
2	Требования к данным для ввода MLS	7
2.1	Формат файла & Структура	7
2.2	Обязательные столбцы для оценки Режим	7
2.3	Особые правила именования столбец	9
2.4	Качество данных& Полнота данных	9
2.5	Размещение строк объекта оценки	10
3	Режим общего назначения	11
3.1	Обзор	11
3.2	Рабочий процесс с боковой панелью	11
3.3	Вкладки главной панели	12
3.4	Сохранение настроек	12
3.5	Темный режим	12
3.6	Язык и регион & Региональные настройки	13

3.6.1	Поддерживаемые страны	13
3.6.2	Paper Override	13
3.6.3	Сохранение настроек по умолчанию	14
4	Режим оценки	15
4.1	Обработка строк объекта оценки	15
4.2	Дата оценки & Возраст, с которого разрешена продажа	15
4.3	Специальный столбец Обозначения	15
4.4	Остаточный вклад (оценка редких признаков)	16
4.5	Обзор корректировок RCA	17
5	Режим анализа рыночных зон	18
5.1	Отличия от режима оценки Режим «	18
5.2	Когда следует использовать режим «Рынок»	18
6	Отбор переменных	19
6.1	Целевые переменные	19
6.2	Таблица предикторов	19
6.3	Определение типов данных& Переопределения	20
6.4	Флаги факторов и линейных моделей	20
6.5	Специальные типы столбцов См.	20
6.6	Несколько целей	21
7	Выбор параметров	22
7.1	Основные параметры	22
7.2	«Forward Pass»	23
7.3	Допустимые взаимодействия	24
7.4	Обрезка	24
7.5	Перекрёстная проверка	24
7.6	Фильтрация подмножеств	24
7.7	Рекомендуемые значения параметров	25
7.7.1	Параметры передачи вперед	25
7.7.2	Параметры обрезки	26
7.7.3	Перекрёстная проверка Параметры	27
7.7.4	Модель дисперсии	27
7.7.5	Файл вывода Earth	28
7.8	Настройки по умолчанию	28
8	Подгонка модели Земли	29
8.1	Кнопка «Fit»	29
8.2	Подгонка модальной кривой & Trase Вывод	29
8.3	Журнал тренировок	29
9	Загрузка данных	30
9.1	Столбцы вывода	30
9.2	Упорядочивание столбцов & Excel Формат	31
9.3	Результаты CQA: балл	31
9.4	Сортировка & Строка объекта оценки	31
9.5	mgcvUI Auto-Export	31
9.6	glmnetUI: импорт	32
9.6.1	Основные преимущества	32
9.6.2	Важно: столбец «Вес» должен соответствовать	32

9.6.3	Рабочий процесс	32
10	Расчеты RCA& Скачивание	34
10.1	Открытие диалогового окна RCA	34
10.2	Интерполяция баллов CQA	34
10.3	Столбцы вывода	34
10.4	Вкладка «Корректировки RCA»	35
11	Таблица сравнения продаж	36
11.1	Обзор & Цель	36
11.2	Модальное окно выбора аналогов	36
11.3	Макет таблицы	36
11.4	Цена продажи, уступки & Net Цена продажи	37
11.5	Сгруппированные строки (Местоположение, Объект, Возраст)	37
11.6	Строки с переменными модели	37
11.7	CQA Остаточная величина & Остаточная характеристика Строки	38
11.8	Скорректированная цена продажи Формула	38
11.9	Защита листа & Редактируемые ячейки	38
11.10	Работа с таблицей в Excel	38
12	Скачивание отчетов	40
12.1	Создать отчет Quarto	40
12.2	Конвертировать отчет Quarto	40
12.3	Содержание отчёта	40
13	Модули постобработки: эластичная сеть и GAM	42
13.1	Почему стоит поддерживать Fits?	42
13.2	Запуск и общее состояние	42
13.3	Процедура «эластичная сеть»	42
13.4	Процедура GAM	42
13.5	Практический алгоритм действий	43
14	Демо-набор данных: Оценка _ 1.csv	44
14.1	Описание	44
14.2	Столбцы	45
14.3	Рекомендуемое руководство по быстрому запуску	46
15	Экспериментальная часть: Извлечение текста из Prolog & Вывод Правила	47
15.1	Описание функции	47
15.2	Требования: vProlog (Trealla), не SWI-Prolog	48
15.3	Включение и запуск	48
15.4	Правила вывода	48
15.5	Файл активных данных с версиями	49
15.6	Программный интерфейс & Набор файлов	49
15.7	Связанный эксперимент: «Климатический регион» Исходные распределения:	49
15.7.1	За пределами Калифорнии: пакеты юрисдикций и модули рыночных зон	50

1 Введение

1.1 Что такое earthUI?

earthUI — это графический интерфейс пользователя для пакета R «earth», реализующего метод многомерных адаптивных регрессионных сплайнов (MARS) (Friedman 1991). В качестве ядра моделирования используется пакет earthСтивена Милбороу (Stephen Milborrow) (Milborrow 2024), подробно описанный на сайтах Milborrow (6.r.[a]) и Milborrow (6.r.[b]). Он работает как локальное приложение Shiny — без входа в систему, без сервера и без учетных записей. Вы запускаете его из R, импортируете набор данных (CSV или Excel), настраиваете модель и выполняете её подгонку в интерактивном режиме.

Приложение обеспечивает полный рабочий процесс: импорт данных, настройку переменных, подгонку модели, графики диагностики, оценку значимости переменных, уравнения модели, а также возможность скачивания отчетов в форматах HTML, Word или PDF.

1.2 Один набор, три процедуры

Начиная с текущей версии, пакет earthUI также содержит два модуля постобработки, которые в предыдущих версиях представляли собой отдельные пакеты glmnetUI и mgcvUI. При установке earthUI устанавливаются все три процедуры; каждая из них запускается как отдельное локальное приложение Shiny:

- `earthUI::launch()` — приложение earth (MARS), описанное в данной статье, на порту 7878. `earth()` — это центральный механизм оценки.
- `earthUI::launch_glmnet()` — модуль постобработки эластичной сети (порт 7879), который импортирует базисные функции модели earth и переоценивает их с использованием регуляризации Lasso/эластичной сети (см. раздел «Импорт glmnetUI» ниже).
- `earthUI::launch_mgcv()` — модуль постобработки GAM (порт 7880), который использует положения узлов модели Earth в качестве начальных точек для процесса сглаживания слагаемых mgcv (см. раздел «mgcvUI Auto-Export» ниже).

Эти три процедуры используют одно дерево проекта, одну базу данных настроек и одну копию вспомогательной инфраструктуры, поэтому проект, открытый в одном приложении, сразу становится доступным для остальных. Процедуры glmnet и mgcv служат подтверждением выводов о значении модели Земли, а не конкурирующими рабочими процессами: предполагаемый алгоритм заключается в том, чтобы сначала провести подгонку и уточнение модели Земли, а затем запустить модули на том же проекте, когда требуется подтверждение результатов.

1.3 Три режима работы

earthUI организует работу в рамках проектов — именованных объектов, которые объединяют в одном фиксированном месте вводные файлы, выводы и настройки. Каждый проект имеет назначение, которое выбирается в мастере создания нового проекта при его создании. Существует три режима:

- Общий — Регрессия по методу Earth для любого типа совокупности или набора данных. Это режим по умолчанию. Он обеспечивает полный рабочий процесс моделирования MARS без каких-либо дополнений, специфичных для конкретной области.
- Для оценки — Регрессия Earth, адаптированная для оценки недвижимости. Добавляет функции, специфичные для оценки отдельных объектов недвижимости, включая работу с объектом оценки, специальные обозначения столбцов, Метод остаточных ограничений (RCA) и .

- Анализ рыночной зоны — регрессия Earth, адаптированная для исследований рыночных зон. Добавляет функции для анализа групп объектов недвижимости на определённом рынке вместо оценки отдельного объекта оценки.

Во всех трёх режимах основной движок моделирования одинаков — вы всегда проводите подгонку модели Земли (MARS). Настройка «Цель» определяет, какие дополнительные инструменты и элементы интерфейса будут доступны.

Цель задается для каждого проекта отдельно. Цель проекта — наряду с указанием страны, штата, округа и города — фиксируется при создании проекта. Переход к другому проекту сбрасывает earthUI в чистое состояние по умолчанию (импортированные данные, результаты моделирования, вкладки, настройки переменных и параметры earth — всё это очищается), а затем восстанавливает сохранённую конфигурацию данного проекта для его цели.

1.4 Недвижимость — специальные материалы

При выборе пункта «Для оценки» или «Анализ рыночной зоны» earthUI активирует ряд функций, предназначенных для анализа рынка недвижимости:

- Обозначения специальных столбцов — Каждому предиктору можно присвоить специальную роль, такую как « contract_date », « dom », « concessions », « latitude », « longitude », « living_area », « lot_size », « actual_age », « effective_age », « area », « site_dimensions » или « display_only ». Эти обозначения определяют, как столбец будет обрабатываться во время подгонки модели, вывода результатов и формирования таблицы продаж. Полную справочную информацию см. в главе 6.
- Округление широты и долготы — Столбцы, обозначенные как широта или долгота, автоматически округляются до 3 знаков после запятой, чтобы предотвратить переобучение.
- Столбец «Срок продажи» — Когда столбец обозначен как contract_date и указана дата оценки, earthUI вычисляет столбец « sale_age » (количество дней между датой продажи и датой оценки) и использует его в качестве предиктора.
- Расчеты RCA (только при оценке) — В режиме оценки, после подгонки модели earthUI может вычислить результаты метода Residual Constraint (RCA), который выводит корректировки для каждого аналога, сводки чистых/валовых корректировок и скорректированную цену продажи объекта оценки.
- Таблица сравнения продаж (только для оценки) — Генерирует многолистовую книгу Excel с формулами для корректировок аналогов, разбивкой остаточных характеристик и скорректированными ценами продаж. См. главу 11.

Совет: Режим «Общее назначение» подходит для любых наборов данных — финансовых, научных, инженерных или связанных с недвижимостью. Режимы «Оценка» и «Рынок» просто предоставляют дополнительные удобные функции для специалистов в сфере недвижимости.

1.5 Установка earthUI и vProlog

Пакет earthUI опубликован на CRAN, поэтому для его стандартной установки в R достаточно одной строки:

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

В этой статье описывается текущий выпуск. Если на момент чтения этой статьи на CRAN по-прежнему доступна более старая версия, версию для разработчиков можно установить с GitHub (обратите внимание на subdir — пакет находится в папке pkg/ репозитория):

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

vProlog (необязательно). Экспериментальный механизм обработки комментариев (см. раздел «Извлечение текста из Prolog» в конце данной статьи) использует пакет vProlog — автономный встроенный движок Prolog. Он намеренно не является обязательной зависимостью: earthUI устанавливается и работает без него, а функции Prolog просто остаются отключенными до тех пор, пока он не будет установлен. vProlog распространяется через репозиторий r-universe автора, в котором доступны готовые бинарные файлы для Windows и macOS (компиляторный инструментарий не требуется):

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

Установка системы Prolog (SWI-Prolog или любой другой) не требуется — движок Trealla встроен непосредственно в сам пакет. Все остальное, описанное в этой статье, работает и без vProlog.

1.6 Начало работы

Чтобы использовать earthUI:

- 1 . Установите пакет из CRAN, как описано выше.
- 2 . Запустите приложение : выполните команду `earthUI::launch()` в R, или из командной строки выполните `Rscript -e 'earthUI::launch()'` . Приложение откроется в вашем веб-браузере на порту 7878. Вы также можете получить доступ к приложению напрямую, перейдя по адресу `http://localhost:7878` в своем браузере. (Модули постобработки запускаются таким же образом: `launch_glmnet()` на порту 7879, `launch_mgcv()` на порте 7880.)
- 3 . Создайте или откройте проект в разделе 1 боковой панели. Мастер создания нового проекта запрашивает название проекта, цель (Общее, Для оценки или «Анализ рыночной зоны»), местоположение (страна, штат, округ, город), а также, по желанию, исходный файл данных, который копируется в папку проекта «in/».
- 4 . Импортируйте данные в разделе 2, где перечислены файлы CSV и Excel, находящиеся в папке «in/» активного проекта. Файлы, добавленные вручную через Finder или Проводник, появятся после нажатия кнопки «Обновить».
- 5 . Настройка переменных — выберите целевые переменные и предикторы, укажите типы данных и назначьте специальные роли столбцам.
- 6 . Настройка параметров Земли — degree, prgune, фильтры подмножеств и другие аргументы `earth()` .
- 7 . Подгонка модели — нажмите «Fit Earth Model» и просмотрите результаты на главной панели.
- 8 . Экспорт — загрузите прогнозы в формате Excel, сгенерируйте наборы отчетов Quarto или (в режиме оценки) рассчитайте корректировки RCA и таблицы сравнения продаж.

Настройки сохраняются на стороне сервера в базе данных проекта с указанием проекта и назначения с помощью кнопок «Сохранить текущие настройки как стандартные» в разделах 3 и 4 (см. главу 3 «Сохранение настроек»).

2 Требования к данным для ввода MLS

В рамках рабочих процессов по оценке недвижимости и анализу рынка исходные данные обычно получают из экспорта из системы Multiple Listing Service (MLS). В этой главе описывается ожидаемая структура файла и столбцы, которые может использовать earthUI.

2.1 Формат файла & Структура

earthUI поддерживает файлы в форматах CSV и Excel (.xlsx, .xls). При импорте названия столбцов автоматически преобразуются в формат `snake_case` — например, «Living SqFt» превращается в «`living_sqft`», «Contract Date» — в «`contract_date`», а «Цена продажи» — в «`sale_price`». Такая нормализация обеспечивает единообразные ссылки на столбцы на протяжении всего рабочего процесса. Разделитель CSV и знак десятичной запятой, используемые при импорте, определяются настройками локали (см. главу 3, «& региональных настроек»).

Ваш файл с данными должен представлять собой плоскую таблицу, в которой каждая строка соответствует одному объекту недвижимости, а каждый столбец — одной характеристике. Первая строка файла должна содержать заголовки столбцов.

2.2 Обязательные столбцы для оценки Режим

Хотя earthUI поддерживает работу с любым набором столбцов, для полноценного рабочего процесса оценки (корректировки по методу RCA + таблица сравнительных продаж) рекомендуется включить следующие столбцы в экспорт из MLS:

Колонка «	» Специальный тип	Назначение
Цена продажи	(целевая)	Переменная отклика для модели
Дата заключения договора	contract_date	Используется для расчета sale_age (количество дней с даты оценки)
Дата размещения объявления	listing_date	Используется вместе с датой заключения договора для расчета DOM, если колонка DOM отсутствует
Количество дней на рынке	dom	Количество дней на рынке; отображается в таблице продаж
Уступки	concessions	Уступки при продаже; чистая цена продажи (Net SP) = цена продажи – Уступки в таблице продаж
Жилая площадь (SF)	living_area	Позволяет рассчитывать остатки на квадратный фут (residual_sf , cqa_sf)
Площадь участка	lot_size	Сгруппировано в строке «Площадь участка Размеры» в таблице продаж
Размеры участка	site_dimensions	Сгруппировано с площадью участка в таблице продаж
Широта	latitude	Округлено до 3 знаков после запятой; используется для расчета близости и группы «Местоположение»
Долгота	longitude	Округлено до 3 знаков после запятой; используется для расчета близости и группы «Местоположение»
Идентификатор района	area	Сгруппировано в строке «Местоположение: Долгота Широта Район» в таблице продаж
Фактический возраст	actual_age	Сгруппировано в строке «Фактический возраст Эффективный возраст» в таблице продаж
Эффективный возраст	effective_age	Сгруппировано с фактическим возрастом в таблице продаж
Номер участка	(нет)	Отображается в строке APN таблицы продаж
Идентификатор объявления	(нет)	Отображается в строке APN таблицы продаж
Адрес	display_only	Отображается в строке «Адрес» таблицы продаж; исключено из модели

Названия столбцов в таблице могут быть на иностранном языке — «специальные» названия указаны на английском языке, чтобы программа R могла обрабатывать их особым образом. В противном случае указанные названия столбцов будут отображаться в регрессионных моделях, графиках и (при проведении оценок) в промежуточной таблице продаж.

Не все столбцы являются обязательными. earthUI адаптируется — если какой-либо столбец отсутствует, соответствующий показатель просто опускается. Например, если столбец «concessions» не указан, строка «Net SP» в таблице «Sales» Таблица отображает цену продажи без вычета скидок. Однако для моделей ценообразования в сфере недвижимости настоятельно рекомендуется использовать определённые столбцы для достижения приемлемой подгонки:

- 1 . Срок с момента продажи (Sale Age) — количество дней между датой заключения договора купли-продажи и датой оценки или анализа. Если используется многолетняя история продаж, особенно за периоды свыше 5 лет, показатель « sale_age » часто играет ключевую роль в оценке цены продажи. Фактически этот показатель зачастую настолько важен, что без него earthUI не может предложить никакой модели вообще.
- 2 . Жилая площадь — также известна под такими названиями, как «Living Sqft», «GLA» (общая жилая площадь) и т. д. Это ещё один основной фактор, определяющий цену продажи.

- 3 . Общее количество ванных комнат — общее количество полных, четвертных, половинных и 3/4 ванных комнат. Например, две полные ванные комнаты и одна половинная будут иметь значение 2,5.
- 4 . Гаражные места или площадь гаража — количество гаражных мест или площадь гаража в квадратных футах.
- 5 . Площадь участка — площадь земельного участка, обычно в квадратных футах или акрах.
- 6 . Долгота, широта и , если доступно, идентификатор района . Корректировки по этим параметрам будут объединены в одну корректировку по местоположению в таблице продаж.

2.3 Особые правила именования столбцов

earthUI идентифицирует столбцы по их специальным обозначениям типов, а не по названиям столбцов. При экспорте в формат MLS вы можете называть столбцы как угодно — главное, чтобы в таблице «Настройка переменных» (глава 6) был назначен правильный специальный тип.

Например, ваша система MLS может экспортировать данные о жилой площади как «GLA», «Living SqFt», «liv _ area» или «gross _ living _ area». После импорта (когда название преобразуется в формат snake_case) вам достаточно просто указать его как «living_area» в раскрывающемся списке «Special». После этого earthUI будет использовать его для расчёта остаточной стоимости на квадратный фут и группировки в таблице продаж, независимо от его исходного названия.

2.4 Качество данных& Полнота данных

- Отсутствующие значения (NA) : Строки, содержащие значения NA в любом столбце предикторов или целевом столбце, автоматически удаляются перед подгонкой модели. Параметр « na.action » внутренне всегда устанавливается в значение « na.fail », при этом удаление значений NA происходит до вызова функции « earth() ».
- Столбцы с датами : Даты должны быть представлены в формате, который R может обработать (например, 2025-06-15 , 06/15/2025 , June 15, 2025). earthUI автоматически определяет столбцы с датами, если успешно анализируется не менее 80% значений. . Двухзначные годы (например, 06/15/25) интерпретируются как ближайший соответствующий год, поэтому недавние экспорты из MLS обрабатываются правильно; четырёхзначные годы всегда принимаются буквально.
- Числовые столбцы : цена продажи, жилая площадь, площадь участка, уступки и подобные поля должны быть числовыми. Если ваши экспортные данные из MLS содержат цены с символами валют или разделителями тысяч (например, «\$ 350,000» или «350.000,00»), вам, возможно, потребуется очистить их перед импортом.
- Столбцы факторов : категориальные переменные, такие как идентификатор района, стиль или состояние, должны содержать управляемое количество уникальных значений. earthUI автоматически определяет факторы, но вы можете переопределить это определение.

Совет: после импорта проверьте столбец «T/F/NA» в таблице предикторов — в нём указано количество значений «true», «false» и «NA» для каждого столбца. Столбцы, в которых более 30 % пропущенных значений, отмечены красным цветом. Рекомендуется исключить столбцы с большим количеством пропущенных значений или очистить данные перед импортом.

2.5 Размещение строк объекта оценки

В режиме оценки в первой строке должна быть указана оценка объекта. Все остальные строки — это аналоги сделок купли-продажи. Цена продажи объекта оценки может быть оставлена пустой (NA) или установлена на любое значение — earthUI в любом случае будет рассматривать её как NA во время подгонки.

В режиме «Анализ рыночной зоны» размещение объекта в строке 1 является необязательным, и ни одна строка не исключается автоматически. Чтобы исключить строку с объектом (или любую другую строку) из подбора, присвойте ей нулевое значение в соответствующем столбце весов или воспользуйтесь фильтром подмножества.

В режиме «Общий» специальной обработки строк не предусмотрено — все строки обрабатываются одинаково.

3 Режим общего назначения

3.1 Обзор

Режим «Общего назначения» является режимом по умолчанию при запуске earthUI. Он обеспечивает полный рабочий процесс моделирования MARS для любых наборов данных — не только для недвижимости. Вы можете использовать earthUI для работы с научными данными, финансового анализа, инженерных исследований или решения любых регрессионных задач, в которых необходимо изучить нелинейные зависимости и взаимодействия между переменными.

В общем режиме интерфейс не отображает функции, специфичные для рынка недвижимости (специальные столбцы, срок продажи, округление координат, RCA). Боковая панель упрощена, чтобы сосредоточить внимание на отборе переменных, настройке параметров, подгонке модели и экспорте.

3.2 Рабочий процесс с боковой панелью

Боковая панель состоит из пронумерованных сворачиваемых разделов, которые помогут вам пройти весь путь от выбора проекта до экспорта:

1. Проект — выберите существующий проект или создайте новый с помощью мастера «Новый проект», в котором указываются цель, страна и административное местоположение (все эти параметры блокируются после создания). Активный проект определяет, откуда считываются файлы ввода и куда записываются все результаты — отдельного поля для указания папки вывода нет.

2. Импорт данных — выберите файл CSV или Excel из папки in/ текущего проекта; ссылка «Обновить» запускает повторное сканирование папки на наличие файлов, добавленных вручную. Для файлов Excel с несколькими листами появляется окно выбора листа. Названия столбцов автоматически преобразуются в регистр «snake» _.

3. Настройка переменных — выбор целевой переменной (поддерживается несколько целевых переменных), дата оценки и таблица предикторов с флажками «Включить», «Фактор», «Линейный» и «Скрытый». Подробную информацию см. в главе 6.

4. Параметры Earth Call — все аргументы функции 'earth()': degree, штраф, nk, метод прореживания, перекрёстная проверка и др. Полный справочник по параметрам см. в главе 7.

Выполнение обработки Prolog — необязательный раздел без номера, расположенный между разделами 4 и 5; он отключен, если экспериментальная функция Prolog не включена в настройках. См. главу «15».

5. Подгонка модели Земли — поле для ввода случайного начального значения и кнопка, запускающая модель в асинхронном режиме. Подробности о подгонке см. в главе 8.

6. Скачать остатки по оцениваемой целевой переменной (переменным) & — (в режимах оценки и рыночном режиме функция называется «Скачать остатки по оцениваемым ценам продажи & ») Экспортирует прогнозы, остатки, оценки CQA и вклады по g-функциям в файл Excel. См. главу 9.

- 7–8. Корректировки RCA и таблица цен — только в режиме оценки. См. главы 10 и 11.

9. Сгенерировать отчёт Quarto — записывает автономный набор отчёта Quarto (исходный код .qmd и ресурсы) в папку вывода проекта. Имеет номер 7 вне режима оценки. См. главу 12.

10. Конвертация отчёта Quarto — преобразует любой файл формата «.qmd» в HTML, Word или PDF. См. главу 12.

3.3 Вкладки главной панели

После подгонки на главной панели отображается десять вкладок:

Вкладка «	» Содержание
Данные	Предварительный просмотр импортированных данных. В режиме оценки данные разбиты на таблицы «Объект оценки» и «Аналоги».
Уравнение	Уравнение построенной модели, отображаемое в формате LaTeX/MathJax, с указанием каждой базисной функции и её коэффициента.
Сводка	Ключевые показатели (R^2 , GRSq, GCV, RSS, $CV R^2$) и таблица коэффициентов.
Важность переменных	Гистограмма и таблица с ранжированными баллами важности предикторов.
Вклад	Графики вклада g-функции. Выберите g-функцию из раскрывающегося списка; для одномерных предикторов отображаются линейные графики, для двумерных взаимодействий (степень ≥ 2) — 3D-перспективные и контурные графики.
Корреляция	Тепловая карта корреляций предикторов.
Диагностика	Три графика: «Остатки против прогнозируемых значений», «Нормальный Q-Q» и «Фактические против прогнозируемых значений».
Корректировки RCA	(только для оценки) Гистограммы показателей «Корректировка по остаткам»%, «Чистая корректировка»% и «Валовая корректировка»% по аналогам, заполняемые после завершения шага 7 («Расчет корректировок RCA»).
Остаточный вклад	Оценивает вклад редких или исключенных признаков в рынок путем совместной регрессии остатков модели по столбцам, которые модель никогда не использовала (см. subsection 4.4).
ANOVA	Таблица декомпозиции ANOVA, показывающая вклад каждой базисной функции в RSS.
Вывод Earth	Необработанный текстовый вывод из <code>'earth::summary()'</code> , включая детали прохода обрезки.

3.4 Сохранение настроек

earthUI сохраняет вашу конфигурацию на стороне сервера в базе данных проекта (projects.sqlite в корневом каталоге regProj) с ключом, определяемым проектом и целью. Сохранение осуществляется с помощью кнопок: кнопка «Сохранить текущее как по умолчанию» в разделе 3 сохраняет цель, дату оценки и конфигурацию предиктора, а кнопка в разделе 4 — параметры вызова earth и матрицу «Допустимые взаимодействия». Каждая кнопка обновляет только свою половину, поэтому они никогда не перезаписывают друг друга, а результаты подгонки не сохраняются автоматически — вы можете свободно экспериментировать, не перезаписывая сохраненные настройки по умолчанию. Несколько файлов ввода проекта (небольшой тестовый фрагмент, полный набор данных) используют одну конфигурацию, и поскольку настройки хранятся в дереве проекта, они переносятся вместе с ним при синхронизации, создании резервной копии или совместном использовании папки regProj.

3.5 Темный режим

Нажмите на значок луны/солнца в правом верхнем углу, чтобы переключаться между светлой и темной темами. Настройки темы сохраняются в локальном хранилище и сохраняются

между сеансами.

3.6 Язык и регион & Региональные настройки

earthUI поддерживает международные стандарты форматов чисел, дат и файлов CSV посредством системы локалей, основанной на странах. В раскрывающемся списке «Страна» в меню «Настройки» (значок шестеренки) на верхней панели меню можно выбрать одну из 31 предустановки для поддерживаемых стран. Каждая предустановка настраивает:

- Разделитель CSV — запятая (,) для США/Великобритании/Японии или точка с запятой (;) для большинства стран Европы, где запятая используется в качестве десятичного знака.
- Десятичный знак — точка (.) или запятая (,).
- Разделитель тысяч — запятая (США/Великобритания/Япония), точка (Германия/Италия/Испания), пробел (Финляндия/Франция/Польша/Прибалтика/Украина/Россия) или апостроф (Швейцария).
- Формат даты — ММ/ДД/ГГГГ (США), ДД/ММ/ГГГГ (большая часть Европы), или ГГГГ-ММ-ДД (Швеция/Литва/Япония/Канада).
- Формат бумаги — Letter (США/Канада/Мексика) или A4 (везде else).

3.6.1 Поддерживаемые страны

Страна	CSV Сентябрь	Десятичный формат	Тысячи	Дата	Бумага
США, Канада (CSV)	,	.	,	MDY/YMD	Формат «Letter»
Великобритания, Ирландия, Австралия, Новая Зеландия	,	.	,	DMY	Формат A4
Германия, Австрия, Бельгия, Нидерланды, Италия, Испания, Португалия, Дания, Бразилия	;	,	.	DMY	Формат A4
Франция, Польша, Чехия, Норвегия, Латвия, Эстония	;	,	(пробел)	DMY	A4
Финляндия, Швеция	;	,	(пробел)	DMY/YMD	A4
Швейцария	;	,	,	DMY	A4
Украина, Россия	;	,	(пробел)	DMY	A4
Турция	;	,	.	DMY	A4
Япония, Южная Корея	,	.	,	YMD	A4
Мексика	,	.	,	DMY	Letter
Литва	;	,	(пробел)	YMD	A4

3.6.2 Paper Override

Под полем выбора страны находится раскрывающийся список «Формат бумаги» (Letter или A4), позволяющий изменить размер страницы отчёта без смены страны. Разделитель

CSV, знак десятичной запятой и порядок форматирования даты соответствуют настройкам выбранной страны; при смене страны настройка формата бумаги сбрасывается на значения по умолчанию для этой страны.

3.6.3 Сохранение настроек по умолчанию

Нажмите «Сохранить» в меню «Настройки», чтобы сохранить настройки региональных параметров глобально. Эти настройки по умолчанию будут применяться ко всем будущим сеансам, независимо от того, какой проект вы откроете. Настройки для отдельных проектов (целевая переменная, предикторы, параметры) сохраняются отдельно в базе данных проекта; настройки региональных параметров по умолчанию сохраняются для всех проектов в базе данных SQLite на уровне пользователя. Этот двухуровневый подход разработан для таких организаций, как аудиторские фирмы, которые работают с данными из разных стран — установите наиболее часто используемую страну в качестве значения по умолчанию, а затем при необходимости вносите корректировки.

Совет: Формат чисел на осях графика и в подписях к наклонам автоматически адаптируется к вашей языковой настройке. В немецкой языковой настройке для разделителя тысяч используются точки (200.000), в финской — пробелы (200 000), а в швейцарской — апострофы (200'000). Символы валют не отображаются — earthUI не зависит от валюты.

4 Режим оценки

При выборе цели «Для оценки» earthUI автоматически настраивается на оценку одного объекта недвижимости. Все функции, описанные в главе 3, остаются доступными; в данной главе рассматриваются только дополнительные возможности, специфичные для оценки.

4.1 Обработка строк объекта оценки

В режиме оценки первая строка вашего набора данных представляет собой объект оценки, а все остальные строки — аналогичные сделки. Ваш вводный файл должен быть организован соответствующим образом (см. главу 2). Цена продажи объекта может быть оставлена пустой или установлена на любое значение — earthUI автоматически рассматривает ее как NA во время подгонки модели.

После импорта вкладка «Данные» разделяется на два раздела: «Объект оценки недвижимости» (строка 1) и «Аналоги» (строки 2 и далее). Строка 1 всегда исключается из подгонки модели — об этом свидетельствует уведомление «Пропуск строки 1 (объект оценки). Подгонка модели на N строках». После подгонки модель по-прежнему генерирует прогнозы для строки объекта оценки, которые отображаются в выводе как «`est_<target>`».

4.2 Дата оценки & Возраст, с которого разрешена продажа

Поле «Дата оценки» отображается в разделе «Настройка переменных» во всех режимах работы (по умолчанию установлена сегодняшняя дата) — не только в режимах оценки, но и в режимах общего и рыночного анализа — однако оно используется только в режимах оценки и рыночного анализа. Если в раскрывающемся списке «Special column» («Специальный столбец») вы укажете столбец «`contract_date`» (Дата вступления в силу), earthUI вычислит столбец «`sale_age`» (Количество дней) — количество целых дней между датой заключения договора каждой сделки и датой оценки. Этот столбец заменяет исходный столбец с датами в качестве предикторов.

При первом нажатии кнопки «Подгонка» после указания даты заключения договора earthUI создает столбец «`sale_age`» и выдает уведомление о необходимости нажать кнопку «Подгонка» еще раз, чтобы включить его.

4.3 Специальный столбец Обозначения

В режимах оценки и рынка для каждого предиктора в таблице «Настройка переменных» появляется специальный раскрывающийся список. Полный список специальных типов и их эффектов:

Специальный тип	Эффект
no	По умолчанию — без специальной обработки
contract_date	Запускает автоматический sale_age расчет с датой оценки
listing_date	Используется вместе с датой заключения договора для расчета DOM, если столбец dom отсутствует
dom	Обозначает столбец «Дни на рынке»; отображается в таблице продаж
concessions	Обозначает скидки при продаже; используется в формуле чистой продажной цены (цена продажи — скидки) в таблице продаж
latitude	Значения округляются до 3 знаков после запятой; используются для расчёта близости по формуле Хаверсина и определения группы местоположения в таблице продаж
longitude	Значения округлены до 3 знаков после запятой; используются для расчёта близости по формуле Хаверсина и определения группы местоположения в таблице продаж
area	Сгруппированы в строке «Местоположение: Долг. Шир. Площадь» в таблице продаж
living_area	Включает остатки на квадратный фут (residual_sf, cqa_sf) в выводе
lot_size	Сгруппированы в строке «Размер участка Размеры» в таблице продаж
site_dimensions	Сгруппировано по площади участка в таблице продаж
actual_age	Сгруппировано в строке «Фактический возраст Эффективный возраст» в таблице продаж
effective_age	Сгруппировано по фактическому возрасту в таблице продаж
display_only	Столбец включен в экспортные данные, но исключен из подгонки модели (допускается несколько)
remarks	(Экспериментально) Колонка с комментариями в виде свободного текста, разбитая на столбцы признаков с помощью опционального шага Prolog (допускается несколько); см. главу 15
list	(Экспериментально) Колонка, разделенная запятыми, разбитая на столбцы «one-hot» TRUE/FALSE с помощью опционального шага Prolog (допускается несколько); см. главу 15

Для каждого специального типа допускается только один столбец (за исключением display_only, remarks и list). Присвоение специального типа второму столбцу автоматически удаляет его из первого. Рядом с именем переменной появляется небольшой синий значок, указывающий присвоенный ей специальный тип.

4.4 Остаточный вклад (оценка редких признаков)

Характеристики, слишком редкие для регрессионного канала — мастерская, представленная лишь в нескольких сделках, — относятся к слою остаточных величин RCA. На вкладке «Вкладка остаточных величин» оценивается их вклад в рынок путем регрессии остатков подгоняемой модели по столбцам, которые модель никогда не использовала, с помощью метода обычных малейших квадратов. Выбор нескольких столбцов позволяет оценить их совместно, благодаря чему коррелирующие характеристики (например, outbuilding_sqft плюс pr_has_workshop) распределяют вклад между собой, а не учитываются дважды. Двоичные флаги дают суммарные оценки и перечисляют остатки по каждой отмеченной сделке в качестве доказательства; числовые столбцы дают вклад на единицу; в отчёте указана стандартная ошибка в размере одной SE ($t \geq 2$ в качестве обычного критерия значимости). Термины, подкреплённые менее чем десятью наблюдениями, автоматически получают отметку «caution» в примечании сгенерированного рабочего файла, где в одном

абзаце для отчёта документируются метод, оценки и предостережения. Из списка кандидатов исключаются только те предикторы, которые фактически использовались моделью — редкий флаг, предложенный при подгонке, но так и не выбранный при прямом проходе, остаётся приемлемым, поскольку модель не учла ни одной из его значений.

4.5 Обзор корректировок RCA

Кнопка «Рассчитать корректировки RCA» & Download (боковая панель, раздел 7, отображается только в режиме оценки после подгонки аналогов) позволяет рассчитать корректировки, основанные на рыночных данных, для каждого аналога по отношению к объекту оценки. Полный рабочий процесс RCA описан в главе 10.

После расчёта корректировок RCA становится доступной кнопка «Сгенерировать таблицу продаж» & (раздел 8 боковой панели). Порядок действий с таблицей продаж описан в главе 11.

Совет: Балл CQA, присваиваемый объекту оценки, определяет, какая доля остаточного распределения относится именно к этому объекту. Балл 5,00 означает, что объект оценки находится на медиане среди аналогов.

5 Режим анализа рыночных зон

Если в качестве цели выбрать «Анализ рыночной зоны», earthUI предоставляет те же функции, характерные для рынка недвижимости, что и в режиме оценки (специальные столбцы, срок с момента продажи, округление координат), но ориентирован на анализ группы объектов недвижимости, а не на оценку отдельного объекта.

5.1 Отличия от режима оценки Режим «

»

- Отсутствие автоматического исключения объектов оценки — в отличие от режима оценки, в режиме рынка строка 1 не исключается из подгонки автоматически. Все строки включаются в подгонку; чтобы исключить строку объекта оценки из подгонки, присвойте ей нулевое значение в соответствующем столбце весов (как в демонстрационном наборе данных) или используйте фильтр подмножества.
- Отсутствие разделов RCA и «Таблица продаж» — шаги «Рассчитать корректировки RCA & Скачать» и «Сгенерировать таблицу продаж& Скачать» недоступны. Режим «Market» ориентирован на подгонку модели и вывод результатов, а не на корректировки по каждому аналогу.
- Нумерация боковых панелей — без шагов RCA и «Sales Grid» раздел «Generate Quarto Report» имеет номер 7 вместо 9 (раздел «Convert Quarto Report» по-прежнему имеет номер 10).

5.2 Когда следует использовать режим «Рынок»

Режим «Анализ рыночной зоны» подходит в следующих случаях:

- Построение регрессионной модели для микрорайона или рыночной зоны с целью понимания факторов, влияющих на стоимость
- Анализ того, как такие переменные, как площадь, возраст, площадь участка и расположение, влияют на цены продажи в группе объектов недвижимости
- Подготовка материалов для анализа рыночной конъюнктуры или определения границ района
- Работа с набором данных, в котором рассматриваемый объект оценки находится в первой строке, но вы хотите иметь возможность включать или исключать его из модели

Совет: Режим «Market» также полезен для общей регрессионного анализа на рынке недвижимости, когда требуются специальные характеристики столбцов (срок продажи, округление координат), но не нужен рабочий процесс корректировки RCA.

6 Отбор переменных

В разделе 3 боковой панели — «Настройка переменных» — вы выбираете, какие столбцы будут включены в модель и как с ними будет проводиться работа.

6.1 Целевые переменные

В раскрывающемся списке «Целевая переменная (ответ)» в верхней части раздела 3 перечислены все столбцы вашего набора данных. Выберите один столбец для стандартной модели с одним ответом или несколько столбцов для модели с несколькими ответами. При выборе нескольких целевых переменных модель подгоняется для всех ответов одновременно с использованием метода «`earth(cbind(y1, y2, ...) ~ .)`».

Столбцы, выбранные в качестве целевых, автоматически исключаются из списка предикторов.

6.2 Таблица предикторов

Ниже селектора целей в таблице перечислены все оставшиеся столбцы со следующими полями:

Столбец	Описание
Переменная	Название столбца (полное название, отображаемое в всплывающей подсказке, если оно сокращено), перед которым стоит переключатель $\times / +$ (см. ниже). В режимах оценки/рынка синий значок указывает на назначенный специальный тип.
Тип	Раскрывающийся список типов данных: <code>numeric</code> , <code>integer</code> , <code>character</code> , <code>logical</code> , <code>factor</code> , <code>Date</code> , <code>POSIXct</code>
Включить	Флажок — включить этот столбец в качестве предиктора в модель
Фактор	Флажок — рассматривать этот столбец как категориальную переменную
Линейный	Флажок — принудительно использовать только линейный ввод (без функций типа « <code>hinge</code> »)
Латентная	Флажок — полностью исключить этот столбец из модели (исключить из <code>earth</code> и из данных, экспортируемых в <code>glmnetUI</code> / <code>mgcvUI</code>), зарезервировав его для работы с латентными переменными
Специальный	Раскрывающийся список (только для оценки/рынка) — см. справочник по типам столбцов «Специальные»
T/F/NA	Справочник по специальным типам столбцов ниже Три значения на столбец: для логических значений — <code>TRUE/FALSE/NA</code> ; для числовых значений — $\neq 0 / = 0 / NA$. Отображается красным, если отсутствует более 30% значений

В подсказке над таблицей дано объяснение сокращений: «`Type` = тип данных столбца, `Inc` = включить в качестве предиктора, `Factor` = рассматривать как категориальную переменную, `Linear` = только линейная модель (без переходных членов), `Special` = роль столбца (например, дата заключения договора `_`)».

«Мягкое отключение» вместо удаления. Переключатель «X» рядом с каждым именем переменной делает строку серым цветом и исключает столбец из модели, не удаляя его из данных; при переключении на зеленый значок «+» столбец снова включается. Столбцы никогда не удаляются.

6.3 Определение типов данных & Переопределения

earthUI автоматически определяет типы данных при импорте. Распознаются столбцы с числовыми, целочисленными, логическими, факторными и датными значениями. Символьные столбцы, которые выглядят как даты (не менее 80 % значений соответствуют распространенным форматам дат), классифицируются как Date.

Вы можете переопределить любое распознавание, изменив значение в раскрывающемся списке «Тип». При изменении типа столбца на «character» или «factor» флажок «Factor» автоматически устанавливается. Изменение типа влияет на то, как столбец передается в функцию «earth()».

6.4 Флаги факторов и линейных моделей

- «Фактор» — Если этот флажок установлен, столбец рассматривается как категориальная переменная. Это подходит для таких столбцов, как `style`, `area_id` или `grade`, которые представляют дискретные группы, а не непрерывные измерения. Столбцы-факторы включаются в модель в качестве индикаторных (фиктивных) переменных.
- Линейный — Если этот флажок установлен, столбец принудительно включается в модель линейно — для него не создаются функции типа «hinge» (кусочно-линейные). Это полезно, когда известно, что переменная имеет строго линейную зависимость от целевой переменной.

6.5 Специальные типы столбцов См.

В режимах «Оценка» и «Рынок» в раскрывающемся списке «Special» доступны следующие параметры. Каждый тип можно назначить не более чем одному столбцу (за исключением «display_only», «remarks» и «list», которые допускают назначение нескольким столбцам):

Дата & . Типы времени:

- `contract_date` — Запускает автоматический расчет показателя «sale_age». Исходный столбец даты заменяется столбцом с целыми числами, отражающим количество дней между датой продажи и датой оценки.
- `listing_date` — Используется в качестве резервного варианта для расчета показателя «Days on Market» (DOM = дата заключения договора – дата размещения объявления), если явно не указан столбец «dom».
- `dom` — Идентифицирует столбец «Дни на рынке». Отображается в строках «APN» и «Дата продажи» таблицы продаж.

Типы денежных единиц:

- `concessions` — Обозначает скидки при продаже (кредиты продавца, поощрения для покупателя и т. д.). Используется в таблице продаж для расчета чистой цены продажи: Чистая цена продажи = Цена продажи – Скидки.

Размер & . Типы мест:

- `latitude` — Значения автоматически округляются до 3 знаков после запятой для предотвращения переобучения. Используются для расчета близости по формуле Хаверсина (расстояние от объекта оценки до каждого аналога) и сгруппированы в строке «Местоположение» таблицы продаж.

- `longitude` — Округление такое же, как для широты. Используется для расчета близости и в группе «Местоположение».
- `area` — Обычно служит идентификатором района или территории. Группируется вместе с широтой и долготой в строке «Местоположение: Долгота | Широта | Площадь» таблицы продаж.
- `living_area` — Позволяет производить расчёты остатков на квадратный фут (`residual_sf` и `sqa_sf`) в выводе данных.
- `lot_size` — Сгруппировано в строке «Размеры участка» таблицы продаж.
- `site_dimensions` — Сгруппировано с площадью участка в таблице продаж (например, «75x120»).

Возрастные группы:

- `actual_age` — Сгруппировано в строке «Фактический возраст | Эффективный возраст» таблицы продаж.
- `effective_age` — Сгруппировано по фактическому возрасту в таблице продаж.

Типы дисплеев:

- `display_only` — Этот столбец включается в экспортные файлы Excel, но полностью исключается из процесса подгонки модели. Используйте эту настройку для полей с адресами, номерами MLS, идентификаторами земельных участков или другими справочными данными, которые не должны выступать в качестве предикторов. Эту настройку можно применить к нескольким столбцам.

Типы извлечения текста (экспериментальные):

- `remarks` — Столбец с произвольным текстом (например, `public_remarks`), который будет разбит на столбцы с характеристиками с помощью дополнительного шага Prolog. Не выполняет никаких действий, пока данная функция не будет включена в настройках. Допускается указание нескольких столбцов. См. главу 15.
- `list` — Столбец с разделителем запятой (например, `existing_structures`), который будет разбит на однозначные `<col>_<value>` столбцы TRUE/FALSE в рамках того же этапа. Допускается использование нескольких столбцов. См. главу 15.

6.6 Несколько целей

Выбор более одной целевой переменной позволяет провести подгонку модели Земли с несколькими откликами. При выборе нескольких целевых переменных:

- Модель прогнозирует все ответы одновременно, используя общие базисные функции для всех целей
- Кнопка «wp» (веса ответов) становится активной, что позволяет присвоить числовой вес каждому ответу
- Модели дисперсии (`varmod.method`) отключены (не поддерживаются для многовариантного прогнозирования)
- На вкладках «Результаты» отображаются показатели по каждому ответу, уравнения и графики диагностики

Совет: Столбцы, помеченные как «`display_only`» (не используемые для прогнозирования), остаются в наборе данных и отображаются при экспорте в Excel, но полностью исключаются из таблицы предикторов. Используйте эту функцию для столбцов с идентификаторами, адресами или другими справочными полями.

7 Выбор параметров

Раздел 4 боковой панели — «Параметры Earth Call» — предоставляет доступ ко всем аргументам, поддерживаемым функцией «earth()». У каждого параметра есть синий значок справки (?) с всплывающей подсказкой. Параметры сгруппированы в сворачиваемые подразделы.

7.1 Основные параметры

Параметр	По умолчанию	Описание
подмножество	(пустое)	Выражение фильтрации строк. См. раздел «Фильтрация подмножеств» ниже.
веса	NULL xml-ph-0028@de. NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. веса NULL xml-ph-0028@de. Фильтрация» ниже. веса NULL	Селектор столбцов для весов случаев (строк). В списке указаны только числовые столбцы.
wp	NULL	Веса отклика для многоцелевых моделей. Кнопка открывает диалоговое окно с одним числовым полем ввода для каждой цели (по умолчанию по 1,0 для каждой).
keepxy	off	Сохранить x, y, подмножество и веса в объекте модели.
trace	0	Уровень трассировки для вывода результатов подгонки (от 0 до 5).
glm	none	Дополнительное семейство GLM: gaussian, binomial или poisson.

Параметр	По умолчанию	Описание
degree	1	Максимальный порядок взаимодействия. Установка значения $\text{degree} \geq 2$ автоматически включает перекрёстную проверку и отображает матрицу «Допустимые взаимодействия»
штраф	2	Штраф GCV на узел. Более высокие значения дают более простые модели.

7.2 «Forward Pass»

Параметр	По умолчанию	Описание
nk	auto	Максимальное количество членов до обрезки. Автоматически заполняется рекомендуемым значением для загруженных данных (см. раздел «Рекомендуемые значения параметров» ниже).
thresh	0.001	Пороговое значение для прямого шага. Меньшие значения допускают большее количество членов.
minspan	auto	Минимальный интервал между узлами. Заполняется автоматически рекомендуемым значением. Отрицательные значения устанавливают максимальное количество узлов на одного предиктора.
endspan	auto	Конечный интервал — минимальное расстояние от узла до края данных. Заполняется автоматически рекомендуемым значением.
newvar.штраф	0	Штраф за введение новой переменной (поощряет повторное использование существующих предикторов).
fast.k	20	Количество родительских членов, которые следует учитывать в алгоритме fast MARS.
fast.beta	1	Регулирует коэффициент старения в алгоритме fast MARS.

7.3 Допустимые взаимодействия

Если значение параметра «degree» установлено на 2 или выше, под основными параметрами появляется таблица взаимодействий. Это верхнетреугольная таблица из флажков, по одному для каждой пары предикторов. Установленный флажок означает, что между двумя предикторами допускается взаимодействие; снятие флажка запрещает данное взаимодействие.

При нажатии на название предиктора (заголовок строки или столбца) включаются или отключаются все взаимодействия для этой переменной. Флажки «Разрешить все» и «Снять все» в верхней части окна позволяют управлять настройками для всех элементов сразу. В матрице используются фиксированные заголовки, поэтому заголовки столбцов и строк остаются видимыми при прокрутке.

В информационном уведомлении содержится следующее напоминание: «Термины взаимодействия повышают риск переобучения. Включена перекрёстная проверка (10 блоков).»

7.4 Обрезка

Параметр	По умолчанию	Описание
pmethod	backward	Метод обрезки: backward , none , exhaustive , forward , seqrep , или cv .
nprune	NULL	Максимальное количество слагаемых после обрезки. Оставьте пустым для автоматического выбора.

7.5 Перекрёстная проверка

Параметр	По умолчанию	Описание
nfold	10	Количество блоков перекрёстной проверки. Установите значение 0, чтобы отключить перекрёстную проверку.
ncross	20	Количество повторений перекрёстной проверки.
stratify	on	Стратифицировать выборки перекрёстной проверки, чтобы каждая складка имела схожее распределение ответов.

Когда степень $\geq$ равна 2, автоматически включается перекрёстная проверка (nfold установлено на 10) для предотвращения переобучения из-за взаимодействий.

7.6 Фильтрация подмножеств

В поле ввода «Подмножество» можно указать выражение на языке R, которое определяет, какие строки будут использоваться для подгонки модели. Выражение можно ввести непосредственно (например, `sale_age < 365 & area_id == 460`) или создать его визуально с помощью кнопки «Построить фильтр» (...).

Диалоговое окно «Создать фильтр» — Нажмите «Создать фильтр ...», чтобы открыть пошаговое диалоговое окно. В каждой строке условия есть раскрывающийся список столбцов, оператор (<, >, <=, >=, ==, !=) и поле ввода значения, которое адаптируется к типу столбца: числовое поле для чисел, выбор даты для дат и раскрывающийся список уникальных значений для столбцов с символами или факторами. Условия объединяются с помощью операторов AND (&) или OR (|). В нижней части окна в режиме предварительного просмотра отображается выражение и количество соответствующих строк. Нажмите «Применить», чтобы вставить выражение в поле ввода текста.

В ручных выражениях для столбцов даты необходимо использовать обертки `as.Date("...")` или `as.POSIXct("...")`. Диалоговое окно «Создать фильтр» обрабатывает это автоматически.

Фильтрация по подмножествам является неразрушающей — исключенные строки остаются в наборе данных и получают прогнозы при экспорте в Excel.

Совет: Используйте скобки для группировки условий «ИЛИ» при их объединении с условиями «И». Без скобок выражение «A & B | C» будет интерпретироваться как «(A & B) | C», что может не соответствовать вашим намерениям.

7.7 Рекомендуемые значения параметров

В разделе 4 earthUI отображает рекомендуемые значения под ключевыми параметрами. Эти рекомендации динамически обновляются в зависимости от количества строк, используемых для подгонки (n), и выбранных предикторов (p). Формулы получены на основе статьи Фридмана о методе MARS, внутренних алгоритмов earth и эмпирических испытаний. Поля `nk`, `minspan` и `endspan` заполняются рекомендуемыми значениями при загрузке данных; изменения, внесенные вручную, сохраняются.

7.7.1 Параметры передачи вперед

`nk` (максимальное количество термов до обрезки):

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

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

Значение по умолчанию для параметра Earth, $\min(200, \max(20, 2p)) + 1$, не учитывает размер набора данных и может оказаться слишком малым для больших наборов данных, что ограничивает прямой проход, не позволяя ему адекватно проанализировать все предикторы. Параметр $\lfloor n/10 \rfloor$ допускает примерно один член на каждые 10 наблюдений — это стандартное эмпирическое правило, позволяющее избежать чрезмерной параметризации. Ограничение в 100 предотвращает снижение эффективности.

`minspan` (минимальное количество наблюдений между узлами):

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

endspan (минимальное количество наблюдений от границ данных):

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

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

При автоматическом расчете minspan в Earth используется уравнение 43 Фридмана: $\lfloor (-\ln(-\ln 0.95) + \ln(p \cdot n)) / (2.5 \ln 2) \rfloor$, которое масштабируется как $\ln(np)$ и растёт слишком медленно для больших наборов данных. По адресу $n=1500$ оно даёт лишь ≈ 7 , что соответствует примерно 245 возможным местам расположения узлов на каждый непрерывный предиктор — это слишком много, что позволяет прямому проходу проводить подгонку под шум. Рекомендуемая формула нацелена на примерно 100 кандидатов на каждый предиктор для больших значений n , в то время как для малых значений n (где она хорошо протестирована) используется автоматический расчёт earth. Endspan устанавливается немного больше, чем minspan, чтобы обеспечить дополнительную защиту границ, что помогает, когда данные имеют тонкие хвосты (что часто встречается в сфере недвижимости).

штраф (штраф GCV за каждый узел):

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

Это стандартное значение в Earth — 2 для аддитивных моделей, 3 для моделей взаимодействия (более высокий штраф компенсирует большее пространство поиска).

newvar.penalty (штраф за добавление нового предиктора):

Рекомендуемое значение: 0,1, если предикторы коррелируют между собой (например, жилая площадь с количеством спален, количеством ванных комнат, площадью участка). Это приводит к тому, что при пошаговом прохождении модель будет отдавать предпочтение повторному использованию предикторов, уже присутствующих в модели, а не введению коррелирующих альтернативных вариантов. Механизм: улучшение RSS для каждого нового предиктора умножается на $1/(1 + \text{penalty})$. При значении newvar.penalty = 0.1 новая переменная должна улучшать RSS как минимум на 10 % больше, чем очередной узел существующей переменной. Это позволяет получать более простые модели с меньшим количеством предикторов без ущерба для окончательных оценок коэффициентов (штраф снимается после отбора). Оставьте значение 0, если предикторы не коррелируют. Обратите внимание, что earth не поддерживает ненулевое значение newvar.penalty вместе с весами по случаям — earthUI отключает ввод (принудительно устанавливая значение 0) при указании столбца с весами.

7.7.2 Параметры обрезки

pmethod: Рекомендуется: backward. Основанная на GCV обрезка по методу «backward» по умолчанию является детерминированной — одни и те же данные всегда дают одну и ту же

модель, независимо от случайных начальных значений. Это важно для воспроизводимости результатов, особенно при проведении оценок. Метод «cv» использует перекрёстную проверку для обрезки, что влечет за собой зависимость от начальных значений.

`prune`: Рекомендуется: оставить пустым (NULL). Позвольте GCV выбрать оптимальное количество термов. Установка значения `prune` накладывает жесткое ограничение, которое перебивает решение GCV.

7.7.3 Перекрёстная проверка Параметры

`nfold` (блоки перекрестных проверок): рекомендуемое значение в зависимости от диапазона размера выборки:

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

Увеличение количества блоков позволяет обучать каждый блок на большем объеме данных (что дает менее смещенную оценку), но приводит к уменьшению размера выделенного тестового блока. В таблице показано, как с ростом выборки количество блоков постепенно увеличивается до стандартного 10-кратного, при этом каждый тестовый блок остается достаточно большим (примерно 50+ строк) для получения стабильной метрики по каждому блоку, а при небольших выборках количество блоков уменьшается до 5-блоков для небольших выборок — где оценка CVR^2 по-прежнему содержит шум, поэтому следует увеличить `ncross` (например, до 15–20), чтобы усреднить дисперсию. Ограничение в 10 отражает то, что каждый дополнительный блок представляет собой полную подгонку по данным, приносящую пренебрежимо малую выгоду для диагностики. При использовании `pmethod = "backward"` перекрёстная проверка не влияет на модель — она лишь вычисляет диагностический CVR^2 и предоставляет остатки для модели дисперсии.

`ncross` (количество повторений CV): рекомендуемое значение в зависимости от диапазона размера выборки:

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

Повторы выравнивают случайные отклонения, связанные с разбиением на блоки, что способствует стабилизации CVR^2 ; для больших выборок требуется меньше повторов, поскольку однократная перекрёстная проверка уже обеспечивает более стабильные результаты. Остатки от повторной перекрёстной проверки также используются в модели дисперсии.

7.7.4 Модель дисперсии

`varmod.method`: Рекомендуется: `lm` — модель Милборроу по умолчанию: робастная линейная модель дисперсии для интервалов предсказания (разброс оценивается по шумным остаткам,

поэтому простая модель обеспечивает лучшую обобщаемость). Если lm не сходится на ваших данных, попробуйте rlm или $const$ (гомоскедастичность, всегда сходится), либо воспользуйтесь $earth$ (также обрабатывает нелинейный разброс). Для этого требуются $nfold > 1$ и $ncross > 1$.

7.7.5 Файл вывода Earth

После каждой успешной подгонки $earthUI$ записывает файл « $\langle filename \rangle_earth_output_ \langle timestamp \rangle.txt$ » в папку вывода. Этот файл содержит термины модели, сводные статистические показатели (R^2 GRSq, CVRSq), сведения о модели дисперсии и журнал трассировки. Для каждой подгонки создается отдельный файл, что обеспечивает накопленную историю данных для сравнения конфигураций параметров.

7.8 Настройки по умолчанию

Радио-кнопка в верхней части раздела 4 определяет, какие настройки по умолчанию будут загружены:

- Использовать последние настройки для файла ввода — восстанавливает настройки, которые вы использовали в последний раз с этим конкретным файлом
- Использовать настройки по умолчанию — применяет ваши сохраненные пользовательские настройки по умолчанию
- Настройки по умолчанию Earth — сбрасывает все параметры к заводским значениям

Нажмите кнопку «Сохранить текущие значения в качестве значений по умолчанию» внизу раздела 4, чтобы сохранить текущие параметры Земли и матрицу «Допустимые взаимодействия»; кнопка «Сопоставить» в разделе 3 позволяет сохранить цель, дату оценки и конфигурацию предикторов. Обе настройки сохраняются для каждого проекта и цели в базе данных проекта.

8 Подгонка модели Земли

8.1 Кнопка «Fit»

В разделе 5 боковой панели расположена кнопка «Fit Earth Model» вместе с полем «Random seed». Это начальное значение обеспечивает воспроизводимость распределения блоков при перекрёстной проверке — в раскрывающемся списке отображаются последние 10 использованных начальных значений, а использованное значение фиксируется на вкладке «Earth Output» после подгонки. Перед подгонкой earthUI проверяет вашу конфигурацию — необходимо выбрать целевую переменную и включить как минимум один предиктор. В режимах «Оценка» и «Рынок» столбцы «Широта» и «Долгота» округляются, а срок реализации рассчитывается на основе даты оценки и даты заключения договора (если указана).

8.2 Подгонка модальной кривой & Trace Вывод

При нажатии кнопки «Подгонка» появляется тёмное модальное окно, в котором отображается:

- Заголовок — «Подгонка модели Земли» с таймером, отсчитывающим секунды, и кнопкой «Прервать», которая отменяет фоновую подгонку
- Журнал трассировки — прокручиваемая область текста с моноширинным шрифтом, отображающая вывод в режиме реального времени от функции `earth()` : информацию о наборе данных, ход прямого прогона, блоки перекрёстной проверки и время завершения
- Цветовая кодировка — линии по умолчанию — зелёные, блоки перекрёстной проверки — жёлтые, ошибки — красные

earthUI выполняет подгонку моделей асинхронно с помощью `callr::r_bg()`, который запускает вычисления в отдельном процессе R. Это позволяет приложению оставаться отзывчивым во время длительных процессов подгонки. Наблюдатель с периодичностью опроса 300 мс считывает вывод из фонового процесса и передает его в журнал трассировки.

По завершении подгонки появляется сообщение «Готово за X,X секунд», в модальном окне появляется кнопка закрытия (X), а фоновый слой удаляется, чтобы вкладки с результатами могли отображаться по требованию. Затем модальное окно автоматически закрывается, как только завершится отображение видимых выводов — либо вы можете закрыть его раньше.

В случае успешного выполнения на кнопке «Подгонка» появляется зелёная галочка. Если происходит ошибка, сообщение об ошибке отображается в журнале трассировки и появляется уведомление.

8.3 Журнал тренировок

Каждый запуск (успешный или неудачный) записывает файл журнала в вашу папку вывода:

- Имя файла : `<datafile>_earth_log_<YYYYMMDD_HHMMSS>.txt`
- Содержимое : временная метка, весь вывод трассировки процесса подгонки и любые сообщения об ошибках
- Расположение : папка вывода активного проекта

Совет: Если `callr` не установлен, earthUI переключается на синхронную подгонку с отображением простой полосы прогресса. Установите `callr`, чтобы получить полный набор возможностей по выводу трассировок.

9 Загрузка данных

После подгонки раздел боковой панели № 6 — под названием «Скачать оценочные целевые переменные & и остатки» в общем режиме и «Скачать оценочные цены продажи & и остатки» в режимах оценки и рынка — загружает файл Excel с прогнозами и диагностикой. Этот вывод используется на этапе 7 (RCA) для присвоения рейтинга CQA (Condition-Quality-Appeal — состояние-качество-привлекательность) объекту оценки. Вывод отсортирован по residual_sf и cqa_sf, чтобы помочь вам оценить, какое место занимает объект оценки в рейтинге. Если модель хорошего качества, то объекты недвижимости должны быть ранжированы от наименее привлекательных до наиболее привлекательных на основе остаточных признаков, которые не вошли в регрессионную модель. Среднее значение должно быть приблизительно равно 0, нижняя половина — отрицательные значения, а верхняя половина — положительные значения. Вы должны найти дома самого низкого качества, или «объекты для ремонта», в нижней части рейтинга, а самые привлекательные дома — в верхней. Обычно бывают исключения в виде аномалий, таких как обращение взыскания, продажи с убытком, продажи в рамках наследственного процесса и срочные продажи, необходимые в связи со сменой работы или по другим причинам. Изучение таких аномалий обычно позволяет выявить вескую причину ценовой аномалии.

Надпись на кнопке меняется в зависимости от режима работы:

- Общие сведения/Рынок: Скачать вывод (Excel)
- Оценка: Скачать промежуточный вывод (Excel)

Формат имени файла: <datafile>_modified_<YYYYMMDD_NHMMSS>.xlsx.

9.1 Столбцы вывода

Для каждой целевой переменной добавляются следующие столбцы:

Столбец	Описание
est_<target>	Прогноз по модели, например, est_sale_price (1 десятичное место)
residual	Разница между фактическим и прогнозным значением (1 десятичное место)
cqa	Оценка состояния, качества и привлекательности (2 десятичных места, диапазон 0–10)
residual_sf	Остаток, деленный на жилую площадь (если указано, 1 знак десятичной)
cqa_sf	Показатель CQA, рассчитанный на основе ранжирования по остатку_sf (2 знака десятичной)
<variable>_contribution	Значение вклада по каждой g-функции (1 dp, по одному столбцу на g-функцию)
basis	Вклад значения пересечения, одинаковый для всех объектов недвижимости (1 dp)
calc_residual	Фактическое значение минус (базис + все вклады) — столбец верификации (1 dp)

В моделях с несколькими целями добавляется суффикс «_<i>», чтобы различать столбец для каждой цели.

9.2 Упорядочивание столбцов & Excel Формат

Столбцы ранжирования размещаются в самом левом положении в выводе в следующем порядке: `residual_sf`, `cqa_sf`, `residual`, `cqa`. Это упрощает просмотр отсортированного списка и определение места, которое занимает объект оценки в распределении CQA.

К этим столбцам применен формат чисел в Excel:

Столбец	Формат	Пример
<code>residual_sf</code>	Числовой, 2 знака после запятой	12,34
<code>cqa_sf</code>	Числовой, 2 знака после запятой	7,25
<code>residual</code>	Числовой, 0 знаков после запятой	15 200
<code>cqa</code>	Числовой, 2 знака после запятой	6,80

9.3 Результаты CQA: балл

балл

Балл CQA (Condition-Quality-Appeal) ранжирует остаток каждой строки по отношению ко всем остальным остаткам. Для данной строки балл CQA представляет собой процент строк с остатком меньшего значения, умноженный на 10. В результате получается шкала от 0 до 10, где:

- Высокий показатель CQA ($\approx 9-10$) — объект недвижимости был продан значительно дороже, чем прогнозировала модель (большой положительный остаток)
- Низкий показатель CQA ($\approx 0-1$) — недвижимость была продана значительно дешевле, чем прогнозировалось (большой отрицательный остаток)
- $CQA \approx 5$ — остаток близок к медиане всех остатков

Если обозначен столбец « `living_area` », сайт `cqa_sf` предоставляет аналогичный рейтинг, основанный на остатках стоимости на квадратный фут.

9.4 Сортировка & Строка объекта оценки

В режимах «Оценка» и «Рынок» аналогичные строки сортируются по показателю « `residual_sf` » в порядке убывания (или по показателю « `residual` », если жилая площадь не указана). Строка объекта оценки (строка 1) остается на первой позиции и не подвергается сортировке. В общем режиме строки экспортируются в исходном порядке.

Такая сортировка в сочетании с крайними левыми столбцами рейтинга позволяет оценить быстро просматривать аналоги от тех, стоимость которых была завышена в наибольшей степени, до тех, стоимость которых была занижена в наибольшей степени, и определить, где находится присвоенный объекту оценки балл CQA в рамках данного распределения.

Перед вызовом функции « `predict()` » уровни факторов в данных прогнозирования приводятся в соответствие с данными обучения. Строки с неизвестными уровнями факторов будут давать прогнозы с значением «NA».

Совет: Столбец « `calc_residual` » всегда должен равняться значению « `residual` » (с учетом округления). Если эти значения значительно различаются, это указывает на проблему в вычислениях, которую стоит изучить.

9.5 mgcvUI Auto-Export

При каждой успешной подгонке с помощью $\text{degree} \leq 2$, `earthUI` автоматически сохраняет полный объект результата в виде файла `.rds` в папку вывода проекта. Имя файла соответствует шаблону `<datafile>_earthUI_result_<YYYYMMDD_HHMMSS>.rds`.

Этот файл можно загрузить с помощью модуля постобработки `mgcv` — входящего в пакет `earthUI` и запускаемого командой «`earthUI::launch_mgcv()`» — с помощью команды «`readRDS()`». Процедура GAM использует координаты узлов и базисные функции модели Земли в качестве исходных точек для сглаживающих слагаемых GAM, что обеспечивает плавный переход от моделирования по методу MARS к моделированию по методу GAM.

Модели с $\text{degree} > 2$ пропускаются, поскольку `mgcvUI` поддерживает только парные взаимодействия. При запуске `earthUI` в режиме Trilogy (из пользовательского интерфейса Trilogy) появляется дополнительный раздел «Lock Model» «Вывод» позволяет зафиксировать конкретную подгонку `.rds` — вместе с её базисом предикторов и общими настройками CQA — в качестве канонической модели `earth`, которую импортируют `glmnetUI` и `mgcvUI`.

9.6 glmnetUI: импорт

Тот же файл `.rds`, сохраненный для процедуры GAM, также можно импортировать в модуль постобработки эластичной сети — он также входит в пакет `earthUI` и запускается с помощью команды `earthUI::launch_glmnet()`. В приложении `glmnet` перейдите к разделу 2: «Импорт из `earthUI`» и с помощью кнопки «Обзор» выберите файл `.rds`.

`glmnetUI` использует стандартный подход «`earth::model.matrix()`»: базисные функции модели `earth` (шарниры, взаимодействия и линейные члены) становятся столбцами матрицы дизайна `glmnet`. Затем `glmnet` выполняет регуляризованную регрессию по этим базисным столбцам, выбирая и сокращая их с помощью `lasso`/эластичной сети. Таким образом, создается адаптивное построение базиса в `earth` с регуляризацией `glmnet`.

9.6.1 Основные преимущества

- Интерпретируемые взаимодействия : взаимодействия на основе «шарниров» в модели `Earth` (включая трёхсторонние члены) имеют чёткое геометрическое значение, в отличие от встроенных взаимодействий типа «векторное произведение» в `glmnet`, которые трудно объяснить в суде или при проведении аудита.
- Автоматическая нелинейность : функции шарниров `earth` позволяют учитывать нелинейные зависимости, которые стандартная модель `glmnet` с исходными предикторами не смогла бы учесть.
- Отбор переменных : `glmnet` может дополнительно прореживать базис `earth`, исключая термы шарниров, которые не вносят вклада.

9.6.2 Важно: столбец «Вес» должен соответствовать

Warning

Если `earthUI` использует столбец весов (например, для исключения строки с весом = 0), тот же самый столбец весов необходимо указать в настройках переменных `glmnetUI`. Если настройки весов различаются, количество строк не совпадёт, и базис `earth` не сможет быть добавлен к матрице модели `glmnet`.

9.6.3 Рабочий процесс

- 1 . Построить модель Земли в `earthUI` (любая степень).
- 2 . Результирующий файл «`.rds`» автоматически сохраняется в папке «вывод» проекта.
- 3 . В `glmnetUI` откройте раздел 2: «Import from `earthUI`» и выберите файл «`.rds`».
- 4 . Убедитесь, что в обоих приложениях используются один и тот же файл данных и один и тот же столбец весов используются в обоих приложениях.

- 5 . Настройте параметры glmnet в разделе 5 (матрица взаимодействий автоматически отключается при активном импорте из earthUI).
- 6 . Нажмите «Fit Model» (Подгонка модели). Столбцы базиса earth появятся на вкладках «Equation» (Уравнение), «Coefficients» (Коэффициенты) и «Contributions» (Вклады).

10 Расчеты RCA& Скачивание

Кнопка «Рассчитать корректировки по методу RCA» (&) появляется в секции 7 боковой панели и отображается только в режиме оценки после того, как модель была подгонена. Метод остаточных ограничений позволяет получить основанные на рыночных данных корректировки для каждого аналога относительно объекта оценки.

10.1 Открытие диалогового окна RCA

При нажатии на кнопку открывается небольшое модальное диалоговое окно со следующим содержанием:

- Тип балла — переключатели для выбора «CQA» или «CQA по SF» (опция «SF» доступна только в том случае, если указан столбец «living_area»). Если вы выберете «CQA на SF», то на основе оценки CQA, присвоенной объекту оценки, его остаточный балл будет равен жилой площади, умноженной на остаточную площадь, соответствующую заданной оценке CQA_SF.
- Значение CQA — ввод оценки CQA объекта оценки (0,00–10,00). Этот балл отражает ваше мнение о том, где объект оценки находится в распределении качества аналогов.
- Кнопка «Generate» — вычисляет вывод RCA и сохраняет его в формате Excel. Кнопка остается неактивной до тех пор, пока не будет введено допустимое значение балла.

10.2 Интерполяция баллов CQA

При нажатии кнопки «Generate» earthUI интерполирует остаточную величину объекта оценки на основе аналогичных пар CQA/остаточных величин:

1. Баллы CQA и остатки аналогов отсортированы по CQA в порядке возрастания
2. Линейная интерполяция (stats::approx()) преобразует введенное вами значение CQA значение на остаток
3. При использовании показателя CQA на квадратный фут остаток на квадратный фут преобразуется обратно в общий остаток путем умножения на жилую площадь объекта оценки
4. Расчетная оценка объекта оценки вычисляется следующим образом: прогноз модели + интерполированный остаток

10.3 Столбцы вывода

Файл RCA в формате Excel (<datafile>_adjusted_<YYYYMMDD_NHMMSS>.xlsx) содержит все промежуточные столбцы вывода, а также:

Колонка	Описание
subject_value	Прогноз модели + интерполированное отклонение (только строка 1)
subject_cqa	Введенный вами балл CQA (только строка 1)
<variable>_adjustment	Вклад испытуемого за вычетом вклада контрольной группы (по g-функции)
residual_adjustment	Отклонение испытуемого за вычетом отклонения контрольной группы
net_adjustments	Сумма всех поправок (вклад + отклонение)
gross_adjustments	Сумма абсолютных значений всех корректировок
residual_pct	Корректировка остатка стоимости как доля от цены продажи

Колонка	Описание
net_adj_pct	Чистые корректировки как доля от цены продажи
gross_adj_pct	Валовые корректировки как доля от цены продажи
adjusted_sale_price	Цена продажи + чистые корректировки

В столбцах «Корректировки» для каждого аналога указано, в какой степени модель характеризует разницу с каждой переменной. «adjusted_sale_price» — это цена продажи аналога после применения всех корректировок, вычисленных по модели, — набор скорректированных цен продажи, которые должны группироваться вокруг оценочной стоимости объекта оценки.

10.4 Вкладка «Корректировки RCA»

После вычисления значения RCA на вкладке «RCA Adjustments» на главной панели отображаются гистограммы распределений показателей Residual Adj %, Чистая корректировка % и Валовая корректировка % по аналогам (по одному набору на каждый целевой объект для моделей с несколькими целевыми объектами), что позволяет быстро визуально оценить, насколько плотно сгруппированы корректировки.

Совет: Балл CQA равный 5,00 означает, что объект оценки находится на медиане среди аналогов. Баллы выше 5 указывают на то, что качество объекта выше среднего; ниже 5 — ниже среднего.

11 Таблица сравнения продаж

11.1 Обзор & Цель

Таблица сравнения продаж представляет собой рабочую книгу Excel, формируемую в режиме оценки (раздел 8 боковой панели) после расчета корректировок RCA. В ней объект оценки представлен наряду с выбранными аналогами в формате структурированной таблицы, с использованием формул Excel, которые автоматически рассчитывают корректировки и скорректированную цену продажи для каждого аналога.

Эта таблица предназначена для рабочего файла оценщика — она объединяет корректировки, полученные на основе регрессионного анализа из модели «Земля», с редактируемыми ячейками, в которых оценщик может распределить остаток CQA по конкретным характеристикам объектов недвижимости. Имя вывода — SalesGrid_<YYYYMMDD_HHMMSS>.xlsx.

11.2 Модальное окно выбора аналогов

При нажатии кнопки «Сгенерировать таблицу продаж & Скачать» открывается модальное диалоговое окно, в котором вы выбираете, какие аналоги включить:

- Рекомендуемые аналоги отмечены по умолчанию. Это аналоги с валовой корректировкой ниже 25% от цены продажи, отсортированные по сроку продажи в порядке возрастания (сначала самые последние продажи) и с ограничением в 30.
- Дополнительные аналоги перечислены ниже; по умолчанию они не отмечены. Их процент валовой корректировки составляет 25% или более, но они по-прежнему доступны для выбора (отображается до 50 объектов, отсортированных по проценту валовой корректировки в порядке возрастания).
- Можно выбрать не более 30 аналогов, что даст до 10 листов (по 3 аналога на лист).

После подтверждения выбора выбранные аналоги сортируются по проценту валовой корректировки в порядке возрастания, благодаря чему наиболее похожие аналоги отображаются на листе 1.

11.3 Макет таблицы

Каждый лист имеет макет из 20 столбцов, в которых размещаются данные по объекту оценки и 3 аналога. Столбцы для каждого объекта:

Сущность	Столбцы (5 на каждую сущность)
Объект оценки	Метка, Фактическое значение 1, Фактическое значение 2, Фактическое значение 3, Вклад в стоимость (VC)
Каждый аналог	Фактическое значение 1, Фактическое значение 2, Фактическое значение 3, Вклад в стоимость (VC), Корректировка

Строки сверху вниз:

1. Название — название листа (например, «Промежуточные аналоги продаж Таблица — Лист 1 из 3»)
2. Заголовки — «Объект оценки» и заголовки столбцов аналогов с адресом
3. Адрес — полный адрес объекта недвижимости
4. APN | MLS # | DOM | Subj.Prox — номер участка, ID объявления, количество дней на рынке и расстояние по формуле Хаверсина (миль) от объекта оценки
5. Цена продажи | Уступки | Чистая цена продажи — формула расчета, уступки и формула цены продажи

- 6 . Заголовок строк переменных регрессии
- 7 . БАЗОВОЕ ЗНАЧЕНИЕ — пересечение с осью абсцисс модели
- 8 . Дата продажи | Вне рынка | На рынке — дата заключения договора, срок продажи и количество дней на рынке
- 9 . Сгруппированные строки (условные) — Местоположение, Площадь участка и/или Возрастные группы
- 10 . Строки с переменными модели — по одной строке на каждый предиктор (за исключением сгруппированных переменных)
- 11 . Пустая разделительная строка
- 12 . Строка заголовка остаточных характеристик
- 13 . CQA |Остаток — балл CQA + оставшаяся формула остатка
- 14 . Строки остаточных характеристик — именованные + пустые строки для записи оценщика ввод
- 15 . Общая стоимость VC / Чистая корректировка
- 16 . Чистая корректировка %
- 17 . Валовая корректировка %
- 18 . Скорректированная цена продажи — строка формулы
- 19 . Нижний колонтитул с указанием авторских прав

11.4 Цена продажи, уступки & Net Цена продажи

В строке 5 для каждого аналога указаны три значения:

- Цена продажи — цена продажи аналога из данных
- Льготы — значение из столбца, обозначенного как concessions (0, если не обозначено)
- Чистая цена продажи — формула Excel: Цена продажи– Льготы

В столбце «Объект оценки» для цены продажи указано «N/A» (поскольку она неизвестна) и, при наличии, сумма льгот.

11.5 Сгруппированные строки (Местоположение, Объект, Возраст)

Когда заданы определенные специальные типы и эти переменные появляются в модели, earthUI создает сгруппированные строки, объединяющие связанные переменные:

Долг.| Шир.| Площадь — Отображается, когда в модели присутствует любой из следующих элементов: longitude, latitude или area. Показывает фактические значения для каждой составляющей переменной, совокупный вклад в значение (сумма отдельных значений VC), а также для аналогов — корректировку (— совокупного значения VC аналога по сравнению с совокупным значением VC аналога). Оформлено с светло-голубым фоном в ячейках VC.

Размеры объекта «|» — Отображаются, когда в модели присутствуют элементы «lot_size» или «site_dimensions». Имеют ту же структуру, что и группа «Location».

Фактический возраст | Эффективный возраст — Отображается, когда в модели присутствуют actual_age или effective_age. Имеет ту же структуру.

Переменные, используемые в сгруппированных строках, исключаются из строк с переменными отдельных моделей, расположенных ниже, что позволяет избежать двойного учета в итоговых значениях корректировок.

11.6 Строки с переменными модели

Ниже сгруппированных строк (или непосредственно под строкой «BASE VALUE», если сгруппированных строк нет) отображается по одной строке для каждого оставшегося предиктора модели:

- Фактические значения для объекта оценки и каждого аналога (фактические значения данных)

- Вклад значения (VC) — вклад g-функции в прогнозируемое значение для данной строки
- Корректировка (только для аналогов) — VC объекта оценки минус VC аналога

11.7 CQA | Остаточная величина & Остаточная характеристика Строки

В строке «Остаточная величина CQA» (|) отображается балл CQA для каждого объекта недвижимости, а также содержится формула для расчета оставшейся части остатка — той части, которая ещё не была отнесена к конкретным характеристикам. Формула выглядит следующим образом:

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

Ниже расположены строки с остаточными характеристиками — 5 именованных строк («Местоположение», «Вид/привлекательность», «Состояние», «Качество», «Прочее») плюс 6 пустых строк. Это ячейки ввода, в которые оценщик вводит вклады в стоимость для характеристик, не учтенных моделью. По мере того как оценщик вводит значения, значение формулы «Остаточная величина» автоматически уменьшается.

Для каждого аналога в столбце «Корректировка» содержится формула: характеристика объекта оценки VC — характеристика аналога VC.

11.8 Скорректированная цена продажи Формула

Строка «Скорректированная цена продажи» содержит формулы Excel:

- Объект оценки : Сумма всех ячеек «Вклад в стоимость» от «БАЗОВОЙ СТОИМОСТИ» до последней строки с остаточными характеристиками. Это отражает общий прогноз модели для данного объекта оценки плюс любые остаточные характеристики, заданные оценщиком.
- Сравнимые объекты : Чистая цена продажи (Net SP) + сумма всех ячеек «Корректировка» (Adjustment) от первой группированной или переменной строки до последней строки с остаточными характеристиками. Это даёт цену продажи аналога после всех корректировок, полученных в результате регрессионного анализа и введенных оценщиком.

11.9 Защита листа & Редактируемые ячейки

Каждый лист защищен от случайного изменения формул и данных. Защита блокирует:

- Все ячейки с формулами (чистая прибыль, оставшаяся сумма остатка, корректировки, скорректированная цена продажи и т. д.)
- Все ячейки с данными (адреса, цены продажи, фактические значения, вклады в стоимость по модели)
- Метки и заголовки

Единственными разблокированными (доступными для редактирования) ячейками являются поля ввода вклада остаточных признаков — ячейки в строке «Остаточные значения» (|) раздела «CQA», в которые оценщик вводит данные по разбивке. Они выделены светло-желтым фоном, что указывает на возможность их редактирования.

11.10 Работа с таблицей в Excel

1. Откройте файл в Excel или другом совместимом приложении для работы с электронными таблицами.

- 2 . Просмотрите корректировки регрессии — корректировки, полученные на основе модели, уже введены и заблокированы.
- 3 . Распределите остаточную величину — в строках с остаточными характеристиками (желтые ячейки) введите свою оценку того, какая часть оставшейся остаточной величины приходится на каждую характеристику (местоположение, вид, состояние, качество и т. д.). Формула «Оставшаяся остаточная величина» будет уменьшаться по мере распределения.
- 4 . Проверьте «Скорректированную цену продажи» — эта формула автоматически обновляется по мере ввода значений остаточных характеристик.
- 5 . Несколько листов — если вы выбрали более 3 аналогов, переключайтесь между листами. В левом столбце каждого листа указан один и тот же объект оценки с 3 различными аналогами.

Совет: Цель состоит в том, чтобы распределить остаток СQA по строкам признаков до тех пор, пока показатель «Остаточный остаток» не станет близок к нулю. Это демонстрирует, что можно учесть разницу между регрессионной оценкой каждого аналога и его фактической ценой продажи.

12 Скачивание отчетов

После подгонки модели формирование отчёта представляет собой двухэтапный рабочий процесс в Quarto: «Генерация отчёта Quarto» (раздел 9 в боковой панели в режиме оценки, 7 — в остальных случаях) создаёт автономный набор отчётов, а «Конвертация отчёта Quarto» (раздел 10) преобразует его — или любой другой файл Quarto — в конечные форматы.

12.1 Создать отчёт Quarto

Нажатие кнопки «Generate Quarto Report» приводит к созданию автономного набора Quarto в папке вывода проекта по адресу `<base>_qmd/`, содержащего заполненный исходный файл `.qmd`, все предварительно сгенерированные графические ресурсы (PNG и PDF), данные отчёта и шаблон справочного документа в формате Word. На этом этапе рендеринг не происходит — набор можно редактировать вручную, объединять с другими исходными файлами Quarto с помощью операторов `include` или просматривать в автономном режиме с помощью программы `quarto preview`.

12.2 Конвертировать отчёт Quarto

В разделе 10 представлен выбор файлов (по умолчанию — последний сгенерированный набор текущего проекта, но вы можете выбрать любой файл `.qmd`) и флажки для выбора формата:

- HTML — файл `.html` с отображением математических формул с помощью KaTeX, встроенными изображениями и оглавлением. Использует тему Flatly Bootstrap и шрифт Roboto Condensed.
- Word — файл `.docx`, подходящий для редактирования и распространения. Включает оглавление, нумерацию страниц и использует пользовательский шаблон ссылок для единообразного стилизового оформления.
- PDF — набран с помощью LuaLaTeX для получения вывода профессионального качества. Размер бумаги соответствует настройкам локали (Letter или A4). Включает страницы в альбомной ориентации для больших матриц взаимодействия и использует шрифт Roboto Condensed и Latin Modern Math для уравнений. Вариант в формате PDF предлагается только при обнаружении дистрибутива LaTeX.

Нажмите «Конвертировать», чтобы выполнить рендеринг выбранных форматов. Рендеринг выполняется в фоновом режиме — в модальном диалоговом окне отображается прошедшее время и ход выполнения Quarto, при этом приложение продолжает реагировать на действия пользователя. По завершении процесса появляется уведомление с указанием местоположения файла.

Все операции (подгонка модели, загрузка данных, расчеты RCA, создание таблицы продаж и формирование отчетов) регистрируются с указанием меток времени начала и окончания, а также продолжительности выполнения в файле `<datafile>_earthui_log.txt`, расположенном в папке вывода, для целей устранения неполадок и мониторинга производительности.

12.3 Содержание отчёта

Каждый сгенерированный отчёт содержит следующие разделы:

1. Описание набора данных — количество наблюдений, целевая переменная (переменные), используемые предикторы, категориальные предикторы
2. Спецификация модели — степень, статус перекрёстной проверки, количество членов, предикторы в модели

- 3 . Допустимые взаимодействия (только для степени ≥ 2) — матрица взаимодействий с галочками, отформатированная для данного формата вывода
- 4 . Сводка результатов — R^2 , GRSq, GCV, RSS и CV R^2 (если была включена перекрёстная проверка)
- 5 . Уравнение модели — полное уравнение модели Земли в нотации LaTeX нотация
- 6 . Коэффициенты & Базисные функции — таблица всех слагаемых, их коэффициентов и базисных функций, которые их определяют
- 7 . Важность переменных — гистограмма и таблица с ранжированием оценок важности предикторов
- 8 . Вклады g-функций — графики для каждой g-функции: линейные для одномерных членов, 3D-перспективные и контурные графики для двумерных взаимодействий
- 9 . Матрица корреляций — тепловая карта корреляций предикторов
- 10 . Диагностика — графики «Остатки против подгоночных значений», «Нормальный Q-Q» и «Фактические против прогнозируемых
- 11 . Разложение ANOVA — таблица, показывающая вклад каждой базисной функции в RSS
- 12 . Вывод Earth — исходный текстовый вывод из `earth::summary()` , включая детали прохода обрезки

В случае моделей с несколькими ответами разделы 4–8 и 10 повторяются для каждой целевой переменной.

Совет: Для создания отчетов требуется пакет R «quarto» и установленная среда Quarto. Для вывода в формате PDF также необходимо установить дистрибутив LaTeX (например, с сайта `tinytex::install_tinytex()`) — в противном случае флажок «PDF» будет скрыт.

13 Модули постобработки: эластичная сеть и GAM

Пакет earthUI содержит два модуля постобработки — процедуру «эластичная сеть», основанную на glmnet, и процедуру GAM, основанную на mgcv. Каждая из них подробно рассматривается в отдельной статье в этом выпуске; в данном разделе объясняется, какое место они занимают в рабочем процессе earthUI и как эти три процедуры взаимодействуют друг с другом.

13.1 Почему стоит поддержать Fits?

Модель «Earth» выдает итоговое значение. Модули постобработки предназначены для его подтверждения: две дополнительные оценки одного и того же объекта оценки, полученные с помощью методологически разных движков и подогнанные к одинаковым проектным данным. Когда переоценка базы «Earth» с помощью метода «эластичная сеть» и независимое сглаживание оценки GAM приводят к практически одинаковым значениям, вывод оценщика больше не опирается на допущения одной единственной модели — что является значимым преимуществом в контексте экспертизы и судебных разбирательств. Когда они расходятся, само это расхождение носит характер диагностики и, как правило, указывает на недостаток данных, чрезмерно гибкий термин или выпадающий аналог, который стоит пересмотреть. Ни один из модулей не имеет приоритета над моделью Earth; они являются вспомогательными доказательствами, намеренно представленными без присвоения весов.

13.2 Запуск и общее состояние

Каждая процедура представляет собой отдельное локальное приложение Shiny:

- earthUI::launch_glmnet() — приложение эластичной сети, на порту 7879.
- earthUI::launch_mgcv() — приложение GAM, на порту 7880.

Все три приложения обращаются к одному и тому же дереву проекта и одной и той же базе данных настроек, поэтому проект, открытый вами в earthUI, немедленно становится доступным в модулях — вместе с переносимыми полями, которые earthUI сохраняет для него: «Ответ» (цель), «Дата оценки», тип и значение RCA «subject-CQA», а также обозначение жилой площади. Вам не нужно вводить их заново; модули считывают их, поэтому одни и те же значения передаются во все три подгонки.

13.3 Процедура «эластичная сеть»

Процедура glmnet импортирует базисные функции модели «земля» — её шарниры, взаимодействия и линейные слагаемые — в качестве столбцов матрицы плана эксперимента и заново оценивает их с применением регуляризации Lasso/эластичная сеть (см. раздел «glmnetUI Import» выше для ознакомления с механизмом работы и предостережением относительно столбца весов). В результате сохраняются интерпретируемые и геометрически значимые члены модели earth, в то время как регуляризатор уменьшает или исключает те, которые не оправдывают себя. См. статью о glmnetUI в этом выпуске для полного описания.

13.4 Процедура GAM

Процедура mgcv действует в обратном направлении: вместо повторного использования базиса модели «Earth» она проводит подгонку новых сглаженных членов с штрафными коэффициентами, используя координаты узлов модели «Earth» в качестве начальных точек (см. раздел «mgcvUI Auto-Export» выше). Поскольку GAM заново выводит функциональные формы в рамках иной парадигмы сглаживания, её согласие с моделью Земли является

подлинным независимым подтверждением, а не просто повторением. Поддерживаются модели Земли с $\text{degree} \leq 2$. Полное описание см. в статье о `mgcvUI` в этом выпуске.

13.5 Практический алгоритм действий

- 1 . Сначала подгоните и уточните модель «earth» в `earthUI`, пока не будете ею удовлетворены. Каждая успешная подгонка ($\text{degree} \leq 2$) автоматически экспортирует файл `.rds`, который используют модули.
- 2 . Запустите процедуры `glmnet` и `mgcv` для того же проекта, импортируйте результат `earth` и выполните подгонку. Работайте с одной процедурой за раз — подбор занимает много времени, а процедуры взаимодействуют через общий проект, поэтому разумно работать последовательно, а не параллельно.
- 3 . Свободно повторяйте итерации. Каждое приложение сохраняет свои результаты при переключении между ними, и ничто не заставляет вас создавать отчёты по ходу работы.
- 4 . Только когда все три результата вас устроят, вернитесь к каждой процедуре и сгенерируйте для неё отчёт Quarto. Каждый метод регистрирует свою подгонку и вывод о значении в записи «трилогии» проекта, благодаря чему все три отчёта можно отследить как относящиеся к одному сравнительному запуску.

Совет: Модули не являются обязательными на любом уровне: собственный рабочий процесс `earthUI` — подгонка персонала, анализ причин и следствий (RCA), таблицы продаж, отчёты — является полноценным и без них. Прибегайте к ним, когда выводы требуют подтверждения: в случае сложной темы, узкого сегмента рынка или задания, которое, вероятно, будет подвергнуто проверке.

14 Демо-набор данных: Оценка _ 1.csv

earthUI содержит демонстрационный набор данных MLS для ознакомления с рабочим процессом оценки. Загрузите его программным способом с помощью:

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

Либо в приложении Shiny скопируйте его в папку « in/» проекта (или прикрепите в качестве исходного файла данных при создании проекта) и выберите его в разделе 2 («Импорт данных»).

14.1 Описание

Файл содержит данные о 1 501 сделке по продаже жилой недвижимости, а также об одном объекте оценки в первой строке (всего 1 502 строки), полученные в результате моделирования экспорта данных из системы MLS. Данные отражают продажи односемейных домов на рынке, охватывающем несколько регионов, с различными размерами, возрастом и местоположением объектов оценки.

Это не реальные данные, но они основаны на реальном районе в Северной Калифорнии. Вся идентифицирующая информация была изменена или удалена.

14.2 Столбцы

Столбец	Тип	Специальный тип	Описание
вес	числовой	—	Вес наблюдения (0 = исключить из подгонки)
id	числовой	отображение_ только	Внутренний идентификатор записи
объект недвижимости_ id	числовой	отображение_ только	Идентификатор объекта недвижимости в системе MLS
объявление_ id	символьный	отображение_ только	Номер объявления в системе MLS
участок_ номер	символьный	отображение_ только	Номер участка, присвоенный оценщиком округа (APN)
улица_ адрес	символьный	отображение_ только	Адрес объекта недвижимости
город_ название (рассчитывает_)			
—	числовой	—	Количество дней от даты заключения договора до даты оценки (рассчитано заранее)
coe_	Дата	отображение_ только	Дата закрытия эскроу
listing_ статус	символ	отображение_ только	Статус объявления (например, «Продано»)
цена_ аренда	числовой	(целевая)	Цена продажи — переменная отклика
аренда	числовой	—	Ежемесячная арендная плата (для моделей с несколькими целями)
список_ цена	числовой	отображение_ только	Цена объявления
исходная_ список_ цена	числовой	отображение_ только	Исходная цена объявления
жилая_ кв. футов	числовой	жилая_ площадь	Общая жилая площадь в квадратных футах
спальни_ всего	целое число	—	Количество спален
ванные_ всего	числовой	—	Общее количество ванных комнат (например, 2,5 = 2 полные + 1 половина)
участок_ размер	числовой	участок_ размер	Площадь участка в квадратных футах
площадь_ id	целое число	площадь	Идентификатор площади в системе MLS
площадь_ текст	символ	отображение_ только	Название площади
возраст	числовой	фактический_ возраст	Возраст объекта недвижимости в годах
год_ постройки	целое число	отображение_ только	Год постройки
широта	числовой	широта	Широта (округленная до 3 знаков после запятой для модели)
долгота	числовой	долгота	Долгота (округлённая до 3 знаков после запятой для модели)
широтаб	числовой	отображение_ только	Широта с полной точностью
долготаб	числовой	отображение_ только	Долгота с полной точностью
гараж_ парковочных мест	целое число	—	Количество парковочных мест в гараже
с плавающей запятой_ количество	целое число	—	Количество каминов
количество_ этажей_	числовой	—	Количество этажей
style	character	—	Архитектурный стиль
view	character	—	Тип вида (например, «Район», «Холмы»)
days_ on_ рынок	integer	dom	Количество дней на рынке
объявление_ date	Дата	объявление_ date	Дата размещения объявления

14.3 Рекомендуемое руководство по быстрому запуску

1. Запустите earthUI: `earthUI::launch()`
2. Создайте проект с целью «Для оценки» в мастере «New Project», указав файл `Appraisal_1.csv` в качестве исходного файла данных
3. В разделе 2 («Import Data») выберите « `Appraisal_1.csv` »
4. В качестве целевого файла выберите `sale_price`
5. Назначьте специальные типы, как показано в таблице выше
6. Включите предикторы: `sale_age` , `living_sqft` , `baths_total` , `lot_size` , `area_id` (в качестве фактора), `age` , `latitude` , `longitude` , `garage_spaces`
7. Установите степень равной 1, нажмите «Fit Earth Model» (Подгонка модели Земли)
8. Загрузите промежуточный вывод (Шаг 6), просмотрите рейтинг CQA
9. Вычислите корректировки RCA (Шаг 7) с баллом CQA~ 5,00
10. Сгенерируйте таблицу сравнения продаж (Шаг 8)

Совет: Демонстрационный набор данных — отличный способ ознакомиться со всеми возможностями earthUI перед импортом ваших собственных данных MLS. Указанные выше рекомендуемые предикторы, как правило, позволяют построить модель с коэффициентом R^2 выше 0,85.

15 Экспериментальная часть: Извлечение текста из Prolog & Вывод Правила

В версии 0.10.0 введена дополнительная и сугубо экспериментальная функция извлечения готовых к использованию в модели характеристик из описательных текстовых столбцов экспорта MLS — прежде всего из столбца «Публичные примечания» в столбец помещаются данные о характеристиках комнат или существующих конструкциях, а также замечания агента и разделенные запятыми списки характеристик, такие как характеристики комнат или существующие конструкции. В качестве механизма извлечения используется Prolog, а в частности — грамматики определённых предложений (DCG): правила, распознающие фразы типа «готово к заселению», «гараж на 3 машины» или «собственная солнечная батарея» в тексте примечаний, написаны в виде грамматических правил Prolog.

Эта функция пока находится в стадии разработки. Сегодня её можно частично использовать для извлечения информации — входящий в набор грамматика уже распознаёт полезный набор признаков, — однако для её адаптации к лексике вашей собственной системы MLS требуется умение читать и писать на языке Prolog, в частности, с использованием DCG. Если это не ваш случай, пропустите эту главу полностью: функция по умолчанию отключена, и остальные части earthUI работают точно так же и без неё.

Warning

Это экспериментальная функция, находящаяся в стадии активной разработки. Её грамматика извлечения не является ни полной, ни проверенной для оценки в производственных условиях — извлечённые столбцы необходимо сверять с исходным текстом, прежде чем использовать их в модели. Для эффективного использования этой функции требуются практические знания языка Prolog и DCG. Все её действия носят аддитивный и обратимый характер; если что-то работает некорректно, отключите переключатель в настройках и продолжайте работу в обычном режиме.

15.1 Описание функции

Объявления в системе MLS содержат большой объем информации, влияющей на стоимость, в виде свободного текста и в полях списка с разделителями. Этот необязательный шаг преобразует этот текст в обычные столбцы, которые можно использовать в регрессионном анализе, с помощью двух специальных ролей столбцов (глава 6):

- `remarks` Столбцы анализируются грамматикой Prolog DCG с формированием одного столбца-особенности на каждый извлеченный атрибут. Префикс столбца создается из первых букв имени исходного столбца: из выражения « `public_remarks` » получаются столбцы « `pr_*` », из выражения « `agent_remark` » получаются столбцы « `ag_*` » и т. д.
- Столбцы « `list` » (например, « `existing_structures = "Shed, Workshop"` ») разбиваются на столбцы «one-hot» `<col>_<value>` с значениями TRUE/FALSE. Эта часть реализована исключительно на языке R и не требует использования Prolog.

Набор грамматики (`inst/prolog/mls_remarks.pl`) в настоящее время извлекает, помимо прочего: уровень состояния с шкалой от 1 до 5 `condition_score`; флаги и типы видов из окон; флаги ремонта кухни и ванной; паркет и другие виды напольных покрытий; список и количество улучшений; места в гараже и площадь гаража; бассейн, спа и другие элементы внешней инфраструктуры; количество каминов; хозяйственные постройки (сарай, амбар, мастерская, касита, гостевой домик, питомник и т. д.) с указанием размеров, полученных из таких данных, как «40x50»; стоянки и хранилища для автофургонов; владение солнечными батареями по сравнению с арендой, аккумуляторные батареи; площадь участка (включая

пересчёт акров в квадратные футы); цены в контексте продажи, связанные с конкретным объектом оценки (снижение цены, взносы в ТСЖ, аренда или выкуп солнечных батарей, налог Мелло-Рус); характеристики местоположения; а также балл настроения по ключевым моментам и проблемам в формате « $[-1, 1]$ ».

15.2 Требования: vProlog (Trealla), не SWI-Prolog

Анализ комментариев осуществляется с помощью пакета vProlog R — автономного движка Trealla Prolog. Его нет на CRAN; установите его из репозитория r-universe автора, как показано в разделе «Установка earthUI и vProlog» в начале этой статьи (готовые бинарные файлы для Windows и macOS, исходный код по адресу github.com/wcraytor/vProlog). Системная установка Prolog (SWI-Prolog или иная) не требуется и не используется. Если vProlog отсутствует, разбор комментариев пропускается с выводом соответствующего уведомления, в то время как разделение на фрагменты list все равно выполняется. Вспомогательная функция vProlog_available() предоставляет отчёт о том, имеются ли как движок, так и входящий в набор набор грамматики.

15.3 Включение и запуск

1. Откройте «Настройки» (значок шестеренки) и установите флажок «Включить обработку столбцов Prolog / списка». По умолчанию этот переключатель выключен, сопровождается примечанием «расширенная функция — для пользователей, владеющих Prolog», настройка сохраняется для каждого проекта и отображается в диалоговом окне «Информация о проекте».
2. В таблице предикторов установите для соответствующих столбцов специальную роль «remarks» или «list» (каждая роль может присваиваться нескольким столбцам).
3. Воспользуйтесь разделом «Выполнить обработку Prolog», расположенным между разделами 4 и 5 на боковой панели. Он предлагает два шага:

1. Развернуть списки «&» — запускает грамматику DCG для каждого столбца «remarks» и выполняет «one-hot»-разбиение каждого столбца «list». Развертывание происходит инкрементально: обрабатываются только те столбцы, которые были отмечены после последнего запуска, а кнопка отключается, как только все отмеченные столбцы уже развернуты. Индикатор выполнения отслеживает ход разбора.

2. Выполнить правила деривации — применяются ваши собственные правила Prolog derive/2 к развернутым столбцам (см. ниже). Перед выполнением правила проходят синтаксическую проверку и проверку в режиме «сухого запуска».

15.4 Правила вывода

Для каждой строки данных earthUI проверяет столбцы этой строки на соответствие фактам cell(Name, Value) и собирает все решения ваших правил derive(Col, Val) — каждый отдельный Col становится новым столбцом, и для каждого столбца принимается первое найденное значение, поэтому за ним может следовать резервное условие. Вспомогательные предикаты (count_true/2, count_false/2, count_true_prefix/2, cell_or/3, is_na/1) поддерживают агрегацию и обработку значений NA, а факты drop/1 удаляют столбцы после выполнения правил. Например:

```
% 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).
```

Правила хранятся в файле `<project>_rules.pl` для каждого проекта, с начальным значением подразумеваемого шаблона, и редактируются с помощью функции «Редактировать правила деривации» ... в плавающем редакторе кода, который можно перетаскивать и который также можно закрепить на боковой панели. Кнопка «Проверить» проверяет правила (проверка синтаксиса, проверка загрузки и пробный запуск на первой строке ваших данных) без их выполнения.

15.5 Файл активных данных с версиями

После выполнения любого из этих шагов расширенная таблица данных записывается в папку ввода проекта как `<original_name>_<YYMMDD_NHMMSS>.xlsx` и становится активным файлом данных проекта — при повторном открытии проекта загружаются расширенные данные, поэтому расширения никогда не пересчитываются, и именно этот файл используется программами `glmnetUI` и `mgcvUI`. Исходные столбцы с комментариями автоматически исключаются из модели; сгенерированные столбцы не включаются автоматически в качестве предикторов. В результате разбора текста может быть сгенерировано много столбцов, поэтому те из них, которые вы (пока) не хотите использовать в регрессии, пометьте как «Latent» или отключите с помощью флага «x» (глава 6).

15.6 Программный интерфейс & Набор файлов

Весь конвейер экспортирован для пакетного использования: `execute_prolog_processing(df, remarks_cols, list_cols, rules_text)` координирует проходы, построенные на основе `remarks_to_columns()`, `split_list_column()`, `apply_user_rules()` и `validate_rules()`. Исходные коды Prolog, включённые в набор по адресу `inst/prolog/`, представляют собой: `mls_remarks.pl` (грамматика DCG и агрегация признаков), `mls_analyze.pl` (драйвер ввода-вывода файлов для пакетного синтаксического анализа), `mls_rules.pl` (вспомогательные предикаты, доступные для правил деривации) и `climate_regions_ca.pl` (ниже).

15.7 Связанный эксперимент: «Климатический регион» Исходные распределения:

Дополнительная экспериментальная установка позволяет изучить характеристики с небольшим объемом данных — те, которые регрессионная модель не может оценить, поскольку они встречаются лишь в нескольких аналогах (например, болотный охладитель). Три вспомогательные программы на языке R — `climate_region_for()`, `market_area_profile()` и `climate_feature_priors()` — классифицируют местоположение объекта (округ + город) по одному из семи признанных покупателями климатических регионов Калифорнии (практическое упрощение 16 климатических зон Калифорнийской энергетической комиссии 16 климатических зон для зданий) и возвращают порядковые априорные значения вклада для систем ОВКВ и средств охлаждения (отопление, кондиционер, болотный охладитель, вентилятор для всего дома, фильтрация) по регионам. Модуль Prolog `climate_regions_ca.pl`

отражает ту же классификацию для использования в правилах вывода; обе стороны генерируются из одного канонического исходного кода на языке R, поэтому они не могут расходиться. Априорные распределения предназначены для обоснования оценки характеристик и служат в качестве проверки правдоподобия оцениваемых коэффициентов — отнесение к регионам (особенно в случае округов, охватывающих несколько регионов) должно проверяться оценщиком.

15.7.1 За пределами Калифорнии: пакеты юрисдикций и модули рыночных зон

Калифорния — это первая юрисдикция, а не граница. earthUI используется во всем мире, и исходные данные о климате в настоящее время реорганизуются в виде пакетов по юрисдикциям: автономных таблиц фактов, сформированных на основе данных (определения зон, сопоставления административных единиц, априорные данные об объектах), привязанные к кодам стран и административных уровней проектной системы, благодаря чему один пакет может описывать штат США, федеральную землю Германии или небольшую страну в целом. Сопутствующий слой модулей рыночных зон фиксирует местную терминологию и уточнённые исходные данные для конкретной рыночной зоны.

Оба предназначены для создания в полевых условиях инженерами по оценке и оценщиками на их родном языке — и, если смотреть реалистично, с помощью моделей LLM: в мире не так много программистов на Prolog, поэтому полевые артефакты намеренно представляют собой таблицы, а не программы. Формат `rack` представляет собой документированную схему с полным примером (сам штат Калифорния), специально для того, чтобы такой помощник, как Claude, мог составить корректный файл `rack` на основе описания рынка простым языком; earthUI генерирует базовый код на Prolog из файла `rack`, и перед тем, как файл `rack` можно будет активировать, он должен пройти (проверки схемы, охвата, референциальной целостности и пробного запуска) должны пройти успешно, прежде чем пакет можно будет активировать. Каждый `rack`, созданный специалистами на местах, содержит обязательную информацию о происхождении — имя автора, дату и цитируемый авторитетный источник — так что за знаниями стоит ответственный специалист, а не инструмент составления, и в отчётах можно указывать точный `rack` и версию, использованные при составлении.

Совет: Практический подход: разверните примечания, объедините детализированные флаги `pr_*` в несколько производных значений или булевых переменных с помощью правил `derive/2`, `drop/1` исходные флаги, которые вы объединили, и оставьте в качестве потенциальных предикторов только производные столбцы. Это позволяет сократить список предикторов и упростить управление коллинеарностью.

References

- 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: интерактивный интерфейс для пакета glmnet (лассо и эластичная сеть)

Русский перевод, подготовленный с использованием искусственного интеллекта и
прошедший редакционную проверку; английская версия является основной.

William Bert Craytor

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

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

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

Эта статья открытого доступа распространяется на условиях [лицензии Creative Commons Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Аннотация. glmnetUI — графический интерфейс пользователя для пакета R glmnet, который подгоняет регуляризованные обобщённые линейные модели с использованием штрафов лассо, гребневой регрессии и эластичной сети. Он предлагает три режима работы — общее прогнозное моделирование, оценку недвижимости и анализ рыночной зоны — и проводит пользователя через импорт данных, настройку модели, подгонку с перекрёстной проверкой, траектории коэффициентов и загружаемые отчёты. В статье описаны требования к формату данных, рабочий процесс моделирования, форматы вывода и полный справочник возможностей glmnetUI.

Ключевые слова: лассо, гребневая регрессия, эластичная сеть, регуляризованная регрессия, glmnet, R, графический интерфейс, оценка недвижимости

Yayıncı notu (Temmuz 2026): Baskıdan kısa süre önce glmnetUI, earthUI (sürüm 0.11.0 ve sonrası) içinde birleştirildi; earthUI burada betimlenen esnek ağ iş akışını artık MARS ve GAM iş akışlarının yanında sunmaktadır. glmnetUI, son yayımlanmış hâliyle (0.5.0) <https://github.com/wcraytor/glmnetUI> adresinde erişilebilir olmayı sürdürmektedir; geliştirme earthUI'de devam etmektedir. Bu makale, yayımlandığı hâliyle glmnetUI için geçerlidir.

Содержание

1	Введение	3
1.1	История и предыстория	3
1.1.1	The Lasso	3
1.1.2	Гребневая регрессия и эластичная сеть	4
1.1.3	Пакет glmnet	4
1.1.4	Сравнение с MARS (Земля) и GAM (mgcv)	4
1.1.5	Рекомендуемый рабочий процесс	6
1.2	Что такое glmnetUI?	6
1.3	Три режима работы	6
1.4	Недвижимость — специальные материалы	7
1.5	Начало работы	7
2	Требования к данным для ввода MLS	9
2.1	Формат файла & Структура	9
2.2	Обязательные столбцы для оценки Режим	9
2.3	Особые правила именования столбца	10
2.4	Качество данных & Полнота данных	10
2.5	Размещение объекта оценки в строке	11

3	Режим общего назначения	12
3.1	Обзор	12
3.1.1	Пропустить первую строку	12
3.1.2	Настройки в зависимости от цели	12
3.2	Рабочий процесс с боковой панелью	12
3.3	Проекты	13
3.4	Импорт из earthUI (Необязательно)	14
3.5	Вкладки главной панели	14
3.6	Сохранение настроек	15
3.7	Темный режим	16
3.8	Язык и регион & Региональные настройки	16
3.8.1	Сохранение настроек по умолчанию	16
4	Режим оценки	17
4.1	Обработка строк объекта оценки	17
4.2	Дата оценки & Возраст, с которого разрешена продажа	17
4.3	Специальный столбец Обозначения	17
4.4	Обзор корректировок RCA	17
5	Режим анализа рыночной зоны	18
5.1	Отличия от режима оценки	18
5.2	Когда следует использовать режим «Рынок»	18
6	Отбор переменных	19
6.1	Целевая переменная	19
6.2	Таблица предикторов	19
6.3	Определение типов данных& Переопределения	19
6.4	Ожидаемые признаки& Обеспечение соблюдения	19
6.5	Специальные типы столбцов Ссылка	20
7	Выбор параметров	22
7.1	Alpha (параметр микширования)	22
7.2	Lambda Selection	22
7.3	Семья	23
7.4	Стандартизировать	23
7.5	Ограничения знаков	23
7.6	«Расслабленное лассо»	23
7.7	Начальные значения для генератора случайных чисел	23
7.8	Матрица взаимодействий	24
7.9	Блок от основного эффекта (только взаимодействие)	24
7.10	Расширенные параметры	25
7.11	Настройки по умолчанию	25
8	Подгонка модели	26
8.1	Кнопка «Fit»	26
8.2	Что происходит во время подгонки	26
9	Вкладки «Результаты»	27
9.1	Предварительный просмотр данных	27
9.2	Уравнение	27
9.3	Корреляция	27
9.4	Резюме	27

9.5	Коэффициенты	28
9.6	Важность переменных	28
9.7	Материалы	28
9.8	ANOVA	28
9.9	Диагностика	28
9.10	Вывод Glmnet	29
9.11	Отчёт «	29
10	Загрузка данных	30
10.1	Столбцы вывода	30
10.2	Результаты CQA: балл	30
11	Расчеты RCA& Скачивание	31
11.1	Открытие диалогового окна RCA	31
11.2	Интерполяция баллов CQA	31
11.3	Столбцы вывода	31
12	Таблица сравнения продаж	32
12.1	Диалоговое окно выбора аналогов	32
12.2	Макет книги	32
12.3	Защита листа & Остатки ячейки	32
13	Скачивание отчетов	33
13.1	«Двухэтапный кварто» Рабочий процесс	33
13.2	Форматы отчётов	33
13.3	Содержание отчёта	33
14	Сравнение с earthUI	35
14.1	Основные различия	35
14.2	Общие функции	35
14.3	Когда использовать какой	36
15	Демо-набор данных: Оценка _ 1.csv	37
15.1	Описание	37
15.2	Столбцы	38
15.3	Рекомендуемое руководство по быстрому запуску	38
16	Системные требования& Устранение неполадок	40
16.1	Поддерживаемые платформы	40
16.2	Дополнительные зависимости	40
16.3	Примечания по платформе	40
16.4	Плавная деградация	40
16.5	Устранение неполадок в программе «	40

1 Введение

1.1 История и предыстория

1.1.1 The Lasso

Метод Lasso (Least Absolute Сжатие и Отбор) был предложен Робертом Тибширани в 1996 году. Название представляет собой намеренную игру слов, связанную с ковбойским лассо

— верёвкой, которая ограничивает и отбирает. Тибширани выбрал его потому, что метод буквально «закидывает лассо» на коэффициенты, ограничивая их абсолютные значения и приводя некоторые из них к точно нулю, тем самым отбирая те переменные, которые останутся в модели.

Идея Тибширани заключалась в том, чтобы добавить штрафное условие « L_1 » (сумма абсолютных значений коэффициентов) к целевой функции метода наименьших квадратов. В отличие от штрафа L_2 , используемого в гребневой регрессии (Hoerl & Kennard, 1970), который сжимает коэффициенты в сторону нуля, но никогда не достигает его, штраф L_1 дает разреженные решения, в которых многие коэффициенты равны точно нулю. Это делает Lasso одновременно методом регуляризации и методом отбора переменных.

1.1.2 Гребневая регрессия и эластичная сеть

Гребневая регрессия была предложена Артуром Хёрлом и Робертом Кеннардом в 1970 году. В ней добавляется штраф по методу « L_2 » (сумма квадратов коэффициентов), который пропорционально уменьшает все коэффициенты, но сохраняет все переменные в модели. Ридж-регрессия эффективна, когда предикторы коррелируют между собой (мультиколлинеарность), но не упрощает модель.

Метод «эластичной сети» (эластичная сеть), предложенный Хуи Цзоу и Тревором Хасти в 2005 году, сочетает в себе оба вида штрафов. Параметр смешивания α регулирует соотношение: $\alpha = 1$ — это чистый Lasso, $\alpha = 0$ — чистый Ridge, а промежуточные значения дают компромиссный вариант. Эластичная сеть преодолевает ограничение метода Lasso: когда предикторы сильно коррелируют, Lasso, как правило, выбирает один и игнорирует остальные, тогда как эластичная сеть объединяет коррелирующие предикторы в группы.

1.1.3 Пакет glmnet

Пакет R «glmnet» был разработан Джеромом Фридманом, Тревором Хэсти и Робертом Тибширани в Стэнфордском университете. Он реализует регуляризацию по методу «эластичная сеть» для обобщенных линейных моделей с использованием чрезвычайно эффективного алгоритма координатного спуска (Friedman, Hastie и Tibshirani 2010), который впоследствии был расширен на модели пропорциональных рисков Кокса (Simon и др. 2011) и на все семейства обобщенных линейных моделей (Tay, Narasimhan и Hastie 2023). Пакет поддерживает гауссовы (линейные), биномиальные (логистические), пуассоновские и другие семейства обобщенных линейных моделей.

К ключевым особенностям «glmnet» относятся:

- Метод регуляризации : подгонка модели по таблице значений λ за один проход, что значительно быстрее, чем отдельная подгонка для каждого значения λ .
- Перекрёстная проверка : `cv.glmnet()` автоматически выбирает оптимальную степень регуляризации с помощью k -блока перекрёстной проверки.
- Границы коэффициентов : поддерживает верхние и нижние границы для отдельных коэффициентов (ограничения по знаку).
- Ослабленное Lasso : повторно подгоняет модель по выбранным переменным с уменьшенным штрафом или без него, что снижает смещение.

1.1.4 Сравнение с MARS (Земля) и GAM (mgcv)

Приложение «Earth» и его два модуля постобработки (glmnet и mgcv) используют разные механизмы моделирования. В приведённой ниже таблице приведены основные различия:

Особенности	glmnet (эластичная сеть)	earth (MARS)	mgcv (GAM)
Тип модели	Линейная (с регуляризацией)	Кусочно-линейная (адаптивные сплайны)	Сглаживание нелинейности (сплайны с штрафной пенализацией)
Нелинейность	Только через взаимодействия или разложение в базис	Автоматически через шарнирные функции с узлами, определяемыми данными	Автоматически через функции сглаживания (тонкая пластина, кубическая и т. д.)
Отбор переменных	Встроенный с помощью штрафа L_1 (Lasso)	Встроенный с помощью пошагового метода «вперед/назад» с обрезкой по GCV	Опционально с помощью сжатия с двойным штрафом (select=TRUE)
Взаимодействия	Попарные векторные произведения (сложно интерпретировать)	Произведения на основе шарниров (интерпретируемые, до степени 3)	Сглаженные тензорные произведения (te() , ti())
Интерпретируемость	Коэффициенты представляют собой линейные веса; их легко объяснить по отдельности	g-функции показывают кусочно-линейные частичные эффекты по каждой переменной	Сглаженные кривые частичных эффектов; очень интуитивно понятны
Лучше всего подходит для	Высокоразмерных данных, отбор переменных, небольшие выборки	Автоматическое обнаружение нелинейности и взаимодействий	Сглаживание нелинейных зависимостей, гибкое моделирование
Корректировки RCA	Линейные корректировки по каждой переменной	Кусочно-линейные корректировки с помощью g-функций	Сглаживание корректировок с помощью частичных эффектов
Контроль переобучения	Лямбда, проверенная с помощью перекрёстной проверки	Обрезка на основе GCV	Оценка параметров сглаживания REML/GCV

Метод MARS (Multivariate Adaptive Regression Splines) был предложен Джеромом Фридманом в 1991 году. Он позволяет строить кусочно-линейные модели путем адаптивного выбора функций-шарниров $\max(0, x - k)$ и положений их узлов на основе данных. Реализация в R представлена в пакете earth , разработанном Стивеном Милбороу. MARS отлично справляется с автоматическим обнаружением нелинейных зависимостей и взаимодействий, не требуя от пользователя их указания. Его g-функции (сгруппированные базисные слагаемые по переменным) обеспечивают наглядные кривые частичных эффектов.

Обобщенные аддитивные модели (GAM) были формализованы Хасты и Тибширани (1990). Реализация на языке R mgcv, разработанная Саймоном Вудом, предоставляет сплайны с штрафной регрессией и автоматическим выбором степени сглаживания по методам REML или GCV. GAM генерируют плавные кривые частичных эффектов, которые интуитивно понятны для визуализации и объяснения. Когда earthUI экспортирует положения

узлов в mgcvUI, узлы, полученные с помощью earth, служат отправными точками для сглаживающих слагаемых GAM, сочетая адаптивное размещение узлов MARS с сглаживанием GAM.

1.1.5 Рекомендуемый рабочий процесс

Для оценки недвижимости и аналогичных задач, требующих понятных и обоснованных моделей:

1. Начните с earthUI, чтобы выявить важные переменные, нелинейные зависимости и взаимодействия в ваших данных. Автоматический выбор базиса и g-функции earthUI позволяют получить отличные первоначальные выводы.
2. Уточните результаты с помощью mgcvUI, если вам нужны более плавные частичные эффекты. Импортируйте узлы из earthUI в mgcvUI для плавного перехода от кусочно-линейных моделей к моделям сглаживания.
3. Используйте glmnetUI, если вам требуется агрессивный отбор переменных (Lasso), когда количество предикторов приближается к числу наблюдений или превышает его, либо если вам нужна чисто линейная модель с регуляризацией для обеспечения обоснованности результатов.

1.2 Что такое glmnetUI?

glmnetUI — это графический интерфейс пользователя для пакета R «glmnet». Он работает как локальное приложение Shiny — без входа в систему, без сервера и без учетных записей. Вы запускаете его из R, импортируете набор данных (CSV или Excel), настраиваете модель и выполняете ее подгонку в интерактивном режиме.

Начиная с текущей версии, glmnetUI больше не является отдельно устанавливаемым пакетом: он поставляется в составе пакета earthUI в качестве модуля постобработки для эластичной сети, запускаемого с помощью команды `earthUI::launch_glmnet()`. Само приложение — а также всё, что описано в данной статье, — осталось без изменений; отличаются лишь процедура установки и команда запуска. Эти три процедуры (earth, glmnet, mgcv) используют одно дерево проекта, одну базу данных настроек и одну копию вспомогательной инфраструктуры.

Приложение обеспечивает полный рабочий процесс: импорт данных, настройку переменных, подгонку модели, графики диагностики, оценку значимости переменных, уравнения модели, а также возможность скачивания отчетов в форматах Word, PDF или HTML.

Регрессия с эластичной сетью сочетает в себе два метода регуляризации:

- Гребневая регрессия (альфа = 0): пропорционально сжимает все коэффициенты к нулю. Подходит, если вы считаете, что все переменные являются значимыми.
- Регрессия Lasso (альфа = 1): может сжимать коэффициенты точно до нуля, выполняя автоматический отбор переменных. Позволяет получать более простые и более интерпретируемые модели.
- Эластичная сеть ($0 < \text{альфа} < 1$): сочетание обоих подходов. Полезно, когда предикторы коррелируют между собой.

1.3 Три режима работы

Каждый проект glmnetUI имеет назначение, выбранное при его создании (раздел 1, «Проект»). Назначение активного проекта определяет, какие инструменты и элементы интерфейса отображаются; чтобы работать в другом режиме, создайте или откройте проект с соответствующим назначением. Существуют три режима:

- Общие настройки — Регрессия эластичной сети для любого типа совокупности или набора данных. Это режим по умолчанию. Он обеспечивает полный рабочий процесс регуляризованной регрессии без каких-либо дополнений, специфичных для конкретной области.
- Для оценки — Регрессия эластичной сети, адаптированная для оценки недвижимости. Добавляет функции, специфичные для оценки отдельных объектов недвижимости, включая обработку объекта оценки, специальные обозначения столбцов и Метод остаточных ограничений (RCA).
- Анализ рыночной зоны — Регрессия эластичная сеть, адаптированная для исследований рыночных зон. Добавляет функции для анализа групп объектов недвижимости на определённом рынке с возможностью исключения объекта оценки из ряда.

Во всех трёх режимах основной механизм моделирования одинаков — вы всегда проводите подгонку модели «эластичная сеть» (glmnet). Настройка «Цель» определяет, какие дополнительные инструменты и элементы интерфейса будут доступны.

1.4 Недвижимость — специальные материалы

При выборе опции «Для оценки» или «Анализ рыночной зоны» glmnetUI активирует ряд функций, предназначенных для анализа рынка недвижимости:

- Обозначения специальных столбцов — Каждому предиктору можно присвоить специальную роль, такую как «contract_date», «dom», «concessions», «latitude», «longitude», «living_area», «lot_size», «actual_age», «effective_age», «area», «site_dimensions» или «display_only». Эти обозначения определяют, как столбец будет обрабатываться во время подгонки модели и при выводе результатов.
- Округление широты и долготы — Столбцы, обозначенные как широта или долгота, автоматически округляются до 3 знаков после запятой, чтобы предотвратить переобучение.
- Столбец «Срок продажи» — Когда столбец обозначен как contract_date и указана дата оценки, glmnetUI вычисляет столбец «sale_age» (количество дней между датой продажи и датой оценки) и использует его в качестве предиктора.
- Расчеты RCA (только для оценки) — В режиме оценки, после подгонки модели glmnetUI может вычислить результаты метода Residual Constraint (RCA), который генерирует корректировки для каждого аналога, сводные данные по чистым и валовым корректировкам, а также скорректированную цену продажи объекта оценки.

Совет: Режим «Общее назначение» подходит для любых наборов данных — финансовых, научных, инженерных или связанных с недвижимостью. Режимы «Оценка» и «Рынок» просто предоставляют дополнительные удобные функции для специалистов в сфере недвижимости.

1.5 Начало работы

Чтобы воспользоваться glmnetUI:

1. Установка earthUI в R: `install.packages("earthUI")` — процедура glmnet включена. (См. раздел «Установка earthUI и vProlog» в статье об earthUI в этом выпуске, где описаны версия для разработчиков версии и дополнительных компонентов.)
2. Запустите приложение : выполните команду `earthUI::launch_glmnet()` в R. Приложение откроется в вашем веб-браузере на порту 7879. Вы также можете получить доступ к приложению напрямую, перейдя по адресу <http://localhost:7879> в своём браузере. Приложение запоминает последний использованный режим назначения и автоматически восстанавливает его.

- 3 . Создайте или откройте проект в разделе 1 боковой панели. В проекте задаются цель и место работы (страна, штат, округ, город), а также папки ввода и вывода. Проекты хранятся в общем корневом каталоге regProj и доступны для других приложений earthUI и mgcvUI. Чтобы переключить режимы, откройте проект с другим назначением.
- 4 . Импортируйте данные в разделе 2, выбрав файл CSV или Excel из папки вводных данных активного проекта. Добавьте файлы в эту папку, затем нажмите «Обновить», чтобы отобразить их список.
- 5 . Импорт из earthUI (необязательно) — раздел 3 позволяет импортировать файл результатов earthUI .rds для использования базисных функций earth с регуляризацией glmnet. Пропустите этот шаг, если у вас нет результатов earthUI.
- 6 . Настройка переменных — выберите целевую переменную и предикторы, укажите типы данных, ожидаемые знаки и назначьте специальные роли столбцам.
- 7 . Настройка параметров glmnet — альфа, лямбда, семейство, ограничения по знаку, взаимодействия и другие параметры.
- 8 . Подгонка модели — нажмите «Fit Glmnet Model» и просмотрите результаты на главной панели.
- 9 . Экспорт — скачайте прогнозы в формате Excel, сгенерируйте и конвертируйте отчёт Quarto или (в режиме оценки) рассчитайте корректировки RCA и таблицу сравнения продаж.

Настройки автоматически сохраняются в локальном хранилище вашего браузера и восстанавливаются при повторной загрузке того же файла ввода.

2 Требования к данным для ввода MLS

В рамках рабочих процессов по оценке недвижимости и анализу рынка исходные данные обычно получают из экспорта из системы Multiple Listing Service (MLS). В этой главе описывается ожидаемая структура файла и столбцы, которые может использовать glmnetUI.

2.1 Формат файла & Структура

glmnetUI поддерживает файлы в форматах CSV и Excel (.xlsx, .xls). При импорте названия столбцов автоматически преобразуются в формат snake_case — например, «Living SqFt» превращается в «living_sqft», «Contract Date» — в «contract_date», а «Цена продажи» — в «sale_price». Такая нормализация обеспечивает единообразные ссылки на столбцы на протяжении всего рабочего процесса. Разделитель CSV и знак десятичной запятой, используемые при импорте, определяются настройками локали (см. главу 3, «& ные региональные настройки»).

Ваш файл с данными должен представлять собой плоскую таблицу, в которой на каждую характеристику приходится одна строка, а на каждую характеристику — один столбец. Первая строка файла должна содержать заголовки столбцов.

2.2 Обязательные столбцы для оценки Режим

Хотя glmnetUI работает с любым набором столбцов, для полноценного проведения оценки рекомендуется использовать следующие столбцы:

Столбец	Специальный тип	Назначение
Цена продажи	(целевая)	Переменная отклика для модели
Дата заключения договора	contract_date	Используется для расчета sale_age (количество дней с даты оценки)
Дата размещения объявления	listing_date	Используется вместе с датой заключения договора для расчета DOM, если столбец DOM отсутствует
Дни на рынке	dom	Дни на рынке; отображается в экспортируемых данных
Уступки	concessions	Уступки при продаже; чистая цена продажи (Net SP) = цена продажи – Уступки
Жилая площадь (SF)	living_area	Позволяет рассчитывать остатки на квадратный фут (residual_sf, cqa_sf)
Размер участка	lot_size	Столбец «Размер участка»
Размеры участка	site_dimensions	Группируется по размеру участка
Широта	latitude	Округляется до 3 знаков после запятой
Долгота	longitude	Округляется до 3 знаков после запятой
Идентификатор района	area	Идентификатор рыночного района / микрорайона
Фактический возраст	actual_age	Возраст объекта недвижимости
Эффективный возраст	effective_age	Эффективный возраст объекта недвижимости
Адрес	display_only	Отображается в экспортируемых данных; исключается из модели

Названия столбцов в таблице могут быть на иностранном языке — «специальные» названия указываются на английском языке, чтобы программа R могла обрабатывать их особым образом. В противном случае указанные названия столбцов будут отображаться в регрессионных моделях, графиках и, при проведении оценок, в выводах.

Не все столбцы являются обязательными. glmnetUI адаптируется — если какой-либо столбец отсутствует, соответствующая характеристика просто опускается. Однако для моделей оценки стоимости недвижимости настоятельно рекомендуется использовать определенные столбцы, чтобы добиться приемлемой качества подгонки:

- 1 . Срок с момента продажи — количество дней между датой заключения договора купли-продажи и датой оценки или анализа. Если используется многолетняя история продаж, особенно за периоды свыше 5 лет, показатель « sale_age » часто играет ключевую роль в оценке цены продажи.
- 2 . Жилая площадь — также известна под такими названиями, как «Living Sqft», «GLA» (общая жилая площадь) и т. д. Это ещё один основной фактор, определяющий цену продажи.
- 3 . Общее количество ванных комнат — общее количество полных, четвертных, половинных и 3/4 ванных комнат. Например, две полные ванные комнаты и одна половинная ванная комната будут иметь значение 2,5.
- 4 . Гаражные места или площадь гаража — количество гаражных мест или площадь гаража в квадратных футах.
- 5 . Площадь участка (недвижимость) — площадь земельного участка, обычно в квадратных футах или акрах.
- 6 . Долгота, широта и , если доступно, идентификатор района . Эти географические переменные помогают модели учитывать географические различия в ценах.

2.3 Особые правила именования столбец

glmnetUI идентифицирует столбцы по их специальным обозначениям типов, а не по названиям столбцов. В экспорте в формат MLS вы можете называть столбцы как угодно — главное, чтобы в таблице «Настройка переменных» (глава 6) был назначен правильный специальный тип.

Например, ваша система MLS может экспортировать данные о жилой площади как «GLA», «Living SqFt», «liv _ area» или «gross _ living _ area». После импорта (когда она преобразуется в формат snake_case) вам достаточно просто указать её как « living_area » в раскрывающемся списке «Special». glmnetUI будет затем использовать её для расчёта остатков на квадратный фут независимо от её исходного названия.

2.4 Качество данных& Полнота данных

- Отсутствующие значения (NA) : Строки, содержащие значения NA в любом столбце предиктора или целевого столбца, автоматически удаляются перед подгонкой модели.
- Столбцы с датами : Символьные столбцы, соответствующие распространенным форматам даты, автоматически анализируются при импорте. При отсутствии 4-значного года приоритет отдается 2-значным форматам года.
- Числовые столбцы : Цена продажи, жилая площадь, площадь участка, уступки и аналогичные поля должны быть числовыми. Если ваша система MLS экспортирует цены с символами валют или разделителями тысяч (например, «\$ 350,000» или «350.000,00»), вам может потребоваться очистить их перед импортом.
- Столбцы-факторы : текстовые столбцы, столбцы TRUE/FALSE и столбцы R-факторов автоматически обрабатываются как категориальные — регрессия может обрабатывать их только как факторы. Числовой столбец (например, числовое значение площади или код стиля) становится фактором только при установке флажка «Factor» ; флажок

glmnetUI никогда не определяет категориальность на основе диапазона значений числового колонка, поэтому дискретные числовые предикторы, такие как `bath_count`, остаются непрерывными. Роль «Special» не зависит от выбора фактора.

Совет: После импорта проверьте столбец «NA» в таблице предикторов. Столбцы с большим количеством пропущенных значений могут привести к исключению строк. Рекомендуется исключить столбцы с высоким процентом пропущенных значений или очистить данные перед импортом.

2.5 Размещение объекта оценки в строке

В режиме оценки в первой строке должна быть указана недвижимость, которая подлежит оценке. Все остальные строки представляют собой аналоги. Строка с объектом оценки исключается из процесса подгонки модели. После подгонки модели она по-прежнему генерирует прогнозы для этой строки.

В режиме «Анализ рыночной зоны» размещение объекта оценки в строке 1 является необязательным.

В режиме «Общий» специальной обработки строк не предусмотрено — все строки обрабатываются одинаково.

3 Режим общего назначения

3.1 Обзор

Режим «General Purpose» (Общего назначения) является режимом по умолчанию при запуске glmnetUI. Он обеспечивает полный рабочий процесс регрессии с использованием метода эластичной сети для любых наборов данных — не только для данных о недвижимости. Вы можете использовать glmnetUI для работы с научными данными, финансового анализа, инженерных исследований или решения любых задач регрессии.

В общем режиме в интерфейсе отсутствуют функции, специфичные для рынка недвижимости (специальные столбцы, срок продажи, округление координат, RCA). Боковая панель упрощена, чтобы сосредоточить внимание на отборе переменных, настройке параметров, подгонке модели и экспорте.

3.1.1 Пропустить первую строку

В верхней части боковой панели появляется флажок «Пропустить первую строку», как только открывается проект типа «Общий» или «Рынок». Если флажок установлен, первая строка исключается из процесса подгонки модели. Это полезно, когда первая строка содержит целевое или эталонное наблюдение, которое не должно влиять на модель. В режиме «Оценка» строка 1 (объект оценки) всегда исключается автоматически.

3.1.2 Настройки в зависимости от цели

Настройки (выбор предикторов, параметры модели, взаимодействия) сохраняются отдельно для каждой комбинации входного файла и режима назначения. При переключении между режимами «Общий», «Оценка» и «Рынок» настройки каждого режима сохраняются независимо друг от друга.

3.2 Рабочий процесс с боковой панелью

Боковая панель состоит из пронумерованных сворачиваемых разделов, которые помогают вам пройти весь путь от выбора проекта до экспорта отчёта. Большинство разделов отображаются только при открытии активного проекта, а некоторые из них зависят от цели проекта. Папка для вывода данных определяется на основе активного проекта — отдельного поля для указания папки вывода данных нет.

1. Проект — создание или открытие проекта. В проекте задаются цель и место проведения работ (страна, штат, округ, город), а также папка ввода (`<os>_in/`) и папка вывода. Проекты хранятся в общей корневой папке `regProj` (по умолчанию `~/regProj`, значение можно переопределить с помощью переменной среды `REGPROJ_ROOT`) и используются совместно с родственными приложениями `earthUI` и `mgcvUI`, поэтому модели всех трёх приложений находятся в одной папке. Нажмите «Закрыть проект», чтобы переключиться.

2. Импорт данных — Выберите файл CSV или Excel из папки ввода текущего проекта. Добавьте файлы в папку с помощью Finder/Проводника, затем нажмите «Обновить», чтобы отобразить их список. Для файлов Excel с несколькими листами появится окно выбора листа. Названия столбцов автоматически преобразуются в формат «snake _».

3. Импорт из `earthUI` (необязательно) — Импортируйте файл `earthUI` с результатами в формате `.rds`, чтобы использовать базисные функции «hinge» из `earthUI` в сочетании с регуляризацией «эластичной сети» из `glmnet`. Кнопка «Обзор» оформлена в стиле, соответствующем окну выбора файлов из раздела 2. Этот шаг является необязательным — пропустите его, если у вас нет результата `earthUI` для импорта.

4. Настройка переменных — селектор целевой переменной, таблица предикторов с флажками «Включить», «Фактор» и «Принудительно», а также раскрывающиеся списки «Знак». Столбец «Специальное» (для указания дат заключения договоров, координат и т. д.) отображается только в режимах «Оценка» и «Рынок». Подробную информацию см. в главе 6.

5. Параметры вызова glmnet — все настройки модели: alpha, lambda, family, standardize, sign constraints, relaxed lasso, начальное значение секретного ключа, матрица взаимодействий, а также расширенные параметры (lambda.min.ratio, nlambdas, метрика потерь CV, порог сходимости, максимальное количество итераций, пересечение). Полный справочник по параметрам см. в главе 7.

6. Подгонка модели Glmnet — кнопка, запускающая модель. Результаты отображаются на вкладках справа.

7. Скачать результаты — экспортирует прогнозы, остатки, балы CQA и вклады по каждой переменной в виде файла Excel в папку с выводами проекта. Доступно во всех режимах после подгонки модели.

8. Расчет корректировок RCA & Скачать — Вычисляет корректировки RCA по каждой переменной для каждого аналога и интерполирует остаток по объекту оценки с помощью CQA. Отображается только в режимах «Оценка» и «Рынок» после подгонки модели. См. главу 11.

9. Сгенерировать таблицу продаж & Скачать — Генерирует промежуточную таблицу сравнения продаж, в которой аналоги ранжируются по проценту валовой корректировки. Отображается только в режиме оценки после вычисления корректировок RCA. См. главу 12.

10. Сгенерировать отчет Quarto — создает автономный набор отчета Quarto (.qmd) — со всеми графиками и таблицами — в папку вывода проекта. Функция становится доступной после подгонки модели.

11. Конвертировать отчет Quarto — преобразует сгенерированный набор .qmd в формат HTML, Word или PDF. Для создания PDF-файла требуется установка LaTeX; если она не обнаружена, опция PDF будет скрыта. См. главу 13.

В режимах «Рынок» и «Общий» скрытые этапы, предназначенные исключительно для оценки, пропускаются, и два раздела отчёта перенумеровываются соответствующим образом.

3.3 Проекты

glmnetUI организует работу в рамках проектов, расположенных в общей корневой папке regProj (в том же дереве, что и earthUI и mgcvUI). Каждый проект имеет собственную папку для вводных данных (<os>_in) и отдельные для каждого движка папки для вывода (например, <os>_out_glmnet).

Чтобы создать проект, нажмите кнопку «+ Новый» (...) в разделе 1 и укажите:

- Цель — «Общая», «Оценка» или «Рыночная зона».
- Местоположение — политико-административная иерархия проекта. Набор уровней, которые необходимо указать, зависит от выбранной вами страны, поэтому иностранные системы наименований поддерживаются автоматически — в проекте в США используются штат/округ/город, во французском — регион/департамент/коммуна, в японском — префектура/город или район/округ и т. д.
- Название проекта — короткое название, длиной не более 8 символов (буквы, цифры, _ и -).

Приложение не использует введенное вами имя дословно: оно формирует уникальное имя папки с отметкой времени, добавляя в начало коды местоположения и дату-время создания в формате <country>_<admin level codes>_<YYYYMMDD-HHMMSS>_<your name> —

например, проект в США с именем « burl» становится « us_ca_081_burlin_20260604-131500_burl». Ограничение в 8 символов применяется только к той части, которую вы вводите. Префикс с отметкой времени добавляется только к вновь созданным проектам; проекты, уже находящиеся на диске, сохраняют свои исходные имена и остаются полностью доступными.

3.4 Импорт из earthUI (Необязательно)

Раздел 3 боковой панели — «Импорт из earthUI» — позволяет импортировать файл результатов earthUI .rds. Благодаря этому glmnetUI может использовать базисные функции earth типа «hinge» (кусочно-линейные слагаемые) в сочетании с регуляризацией «эластичной сети» из glmnet, объединяя автоматическое обнаружение нелинейности earth с отбором переменных и ограничениями на коэффициенты glmnet.

При импорте результатов earthUI подгонка модели glmnet определяется предикторами модели earth — а не флажками «Включить», — поскольку glmnet выполняет подгонку на основе базисного разложения earth. Таблица «Конфигурация переменных» явно отражает это: поля «Включить» и «Тип» заблокированы, строки для предикторов, не использованных моделью earth, отключены, а в примечании поясняется, что набор предикторов берется из модели earth. Поля «Сила» и «Знак» остаются доступными для редактирования для предикторов earth, поскольку они по-прежнему применяются к базису earth во время подгонки.

Кнопка «Обзор» оформлена в стиле, соответствующем выбору файлов из раздела 2. Этот шаг является полностью необязательным — если у вас нет файла результатов earthUI, просто пропустите раздел 3 и перейдите к разделу 4.

3.5 Вкладки главной панели

После подгонки на главной панели появляются следующие вкладки:

Вкладка	Содержание
Предварительный просмотр данных	Предварительный просмотр импортированных данных. В режиме оценки данные разбиты на таблицы «Объект оценки» и «Аналоги».
Уравнение	Уравнение построенной модели, отображаемое в формате LaTeX/MathJax, с указанием каждого ненулевого коэффициента и членов взаимодействия.
Корреляция	Тепловая карта корреляций числовых предикторов (до подгонки модели).
Сводка	Ключевые метрики (R^2 , Adj R^2 , GR ^{2} , CV R^2 , RMSE, MAE) и метрики оценки (COD, PRD, Median Ratio).
Коэффициенты	Таблица ненулевых коэффициентов с предупреждениями о нарушении знака.
Важность переменных	Гистограмма и таблица ранжирования важности предикторов ($ \beta \times sd(x)$). Интерактивный Plotly-ховер, если доступен.
Вклады	Графики частичных эффектов по каждой переменной с метками наклона. Точечный график + линия подгонки для числовых предикторов, бокс-плоты для факторов, 3D-поверхность/точечный график/тепловая карта для взаимодействий.
ANOVA	Таблица декомпозиции дисперсии: SS по каждой переменной, % SS модели, коэффициент.
Диагностика	Четыре графика: путь коэффициента, ошибка CV, фактические значения против прогнозируемых, остатки против подгонки.
Результаты Glmnet	Исходные результаты модели: использованное начальное значение случайного числа, вывод модели, выбранное значение lambda, lambda.min/1se, gamma (если ослаблено) и полный вектор коэффициентов.
Отчёт	Настройки отчёта и поля экспорта.
Корректировки RCA	Результаты анализа RCA (режим оценки, после вычислений).

3.6 Сохранение настроек

glmnetUI автоматически сохраняет вашу конфигурацию в локальном хранилище браузера, используя в качестве ключа имя файла для ввода. При повторной загрузке того же файла все настройки восстанавливаются: выбор целевой переменной, флажки предикторов, типы данных, ожидаемые знаки, параметры glmnet и матрица взаимодействий. Режим назначения, использовавшийся последним, также сохраняется глобально и восстанавливается при повторном запуске приложения. Доступны три варианта:

- Использовать последние настройки для файла ввода — восстановить сохраненные настройки для этого конкретного файла
- Использовать настройки по умолчанию — применить сохраненные глобальные настройки по умолчанию
- Настройки glmnet по умолчанию — восстановить заводские настройки по умолчанию ($\alpha = 1$, CV lambda.1se, 10 блоков, гауссово распределение, стандартизация включена)

Кнопка «Сохранить текущие настройки в качестве стандартных» позволяет сохранить все текущие настройки в качестве глобальных настроек по умолчанию.

3.7 Темный режим

Нажмите на значок луны/солнца в правом верхнем углу, чтобы переключаться между светлой и тёмной темами. Настройки темы сохраняются в локальном хранилище и сохраняются между сеансами. Все элементы интерфейса (таблицы, графики, карточки, кнопки) адаптируются к выбранной теме с помощью переменных CSS.

3.8 Язык и регион & Региональные настройки

glmnetUI поддерживает международные стандарты формата чисел и CSV-файлов благодаря системе локалей, основанной на странах. В раскрывающемся списке «Настройки» в заголовке окна доступны переключатели «Страна» и «Формат» для 31 поддерживаемой страны. Каждая предустановка настраивает:

- Разделитель CSV — запятая (,) для США, Великобритании и Японии или точка с запятой (;) для большинства стран Европы, где запятая используется в качестве десятичного знака.
- Десятичный знак — точка (.) или запятая (,).
- Разделитель тысяч — запятая (США/Великобритания/Япония), точка (Германия/Италия/Испания), пробел (Финляндия/Франция/Польша/Прибалтика/Украина/Россия) или апостроф (Швейцария).
- Формат бумаги — Letter (США/Канада/Мексика) или A4 (везде else).

3.8.1 Сохранение настроек по умолчанию

Нажмите «Сохранить как значение по умолчанию», чтобы сохранить настройки локали глобально. Эти настройки по умолчанию будут применяться ко всем будущим сеансам независимо от того, какой файл данных вы загрузите.

Совет: формат чисел на осях графика и в метках наклона автоматически адаптируется к вашей языковой настройке. В немецкой языковой настройке для разделителя тысяч используются точки (200.000), в финской — пробелы (200 000), а в швейцарской — апострофы (200'000). Символы валют не отображаются — glmnetUI не зависит от валюты.

4 Режим оценки

При выборе цели «Для оценки» glmnetUI настраивается на оценку одного объекта недвижимости. Все функции, описанные в главе 3, остаются доступными; в данной главе рассматриваются только дополнительные возможности, специфичные для оценки.

4.1 Обработка строк объекта оценки

В режиме оценки первая строка вашего набора данных представляет собой объект оценки, а все остальные строки — аналогичные сделки. Ваш вводный файл должен быть организован соответствующим образом (см. главу 2). Цена продажи объекта может быть оставлена пустой или установлена на любое значение — glmnetUI автоматически рассматривает ее как NA во время подгонки модели.

После импорта вкладка «Предварительный просмотр данных» разделяется на два раздела: «Объект оценки» (строка 1) и «Сопоставимые сделки» (строки 2 и далее). Строка 1 всегда исключается из подгонки модели. После подгонки модели модель по-прежнему генерирует прогнозы для строки, относящейся к данному объекту оценки.

4.2 Дата оценки & Возраст, с которого разрешена продажа

В режимах «Оценка» и «Рынок» в разделе «Настройка переменных» появляется поле «Дата оценки» (по умолчанию установлена сегодняшняя дата). Если вы выберете для столбца значение «contract_date» в раскрывающемся списке «Специальный столбец», glmnetUI рассчитает столбец «sale_age» — количество целых дней между датой заключения каждого договора купли-продажи и датой оценки. Этот столбец добавляется в качестве предиктора.

При изменении даты оценки показатель «sale_age» автоматически пересчитывается.

4.3 Специальный столбец Обозначения

В режимах оценки и рынка для каждого предиктора в таблице «Конфигурация переменных» появляется специальный раскрывающийся список. Полную справку по специальным типам и их эффектам см. в главе 6.

4.4 Обзор корректировок RCA

Кнопка «Рассчитать корректировки RCA» & (боковая панель, раздел 8, отображается в режимах «Оценка» и «Рынок» после подгонки) позволяет рассчитать корректировки, основанные на рыночных данных, для каждого аналога по отношению к объекту оценки. Полный рабочий процесс RCA описан в главе 11.

Совет: Балл CQA, присваиваемый объекту оценки, определяет, какая доля остатков распределения относится именно к этому объекту. Балл 5,00 означает, что объект оценки находится на медиане среди аналогов.

5 Режим анализа рыночной зоны

Если в качестве цели выбрать «Анализ рыночной зоны», glmnetUI предоставляет те же функции, характерные для рынка недвижимости, что и в режиме оценки (специальные столбцы, срок с момента продажи, округление координат), но ориентирован на анализ группы объектов недвижимости, а не на оценку отдельного объекта.

5.1 Отличия от режима оценки

- Обработка строк с объектами оценки осуществляется гибко — в режиме рынка строка 1 не рассматривается автоматически как объект оценки.
- Отсутствует таблица сравнительных продаж — шаг «Расчет корректировок RCA & Скачать» (раздел 8) доступен в рыночном режиме, но шаг «Таблица продаж» (раздел 9) доступен только в режиме оценки.

5.2 Когда следует использовать режим «Рынок»

Режим «Анализ рыночной зоны» подходит в следующих случаях:

- Построение регрессионной модели для района или рыночной зоны с целью понимания факторов, влияющих на стоимость
- Анализ того, как такие переменные, как площадь, возраст, площадь участка и расположение, влияют на цены продажи в группе объектов недвижимости
- Подготовка материалов для анализа рыночной конъюнктуры или
- Работа с набором данных, в котором требуются специальные свойства столбцов (срок продажи, округление координат), но не требуется рабочий процесс корректировки RCA

Совет: Режим «Market» также полезен для общей регрессионного анализа рынка недвижимости, когда требуются специальные характеристики столбцов, но не нужен рабочий процесс корректировки RCA.

6 Отбор переменных

В разделе 4 боковой панели — «Настройка переменных» — вы можете выбрать, какие столбцы будут включены в модель и как с ними будет производиться работа.

6.1 Целевая переменная

В раскрывающемся списке «Целевая (реакционная) переменная» в верхней части раздела 4 перечислены все столбцы вашего набора данных. Выберите один столбец в качестве целевой переменной (например, «Цена продажи»). Целевой столбец автоматически исключается из списка предикторов.

6.2 Таблица предикторов

Ниже селектора целей в таблице перечислены все оставшиеся столбцы со следующими полями:

Столбец	Описание
Переменная	Название столбца
Тип	Раскрывающийся список типов данных: numeric , integer , character , Date , POSIXct
Включить	Флажок — включить этот столбец в качестве предиктора в модель
Фактор	Флажок — рассматривать переменную как категориальный фактор (создает фиктивные столбцы в матрице модели)
Флажок «Force»	— принудительно включить в модель (коэффициент штрафа = 0, никогда не исключается методом Lasso)
«Special»	Раскрывающийся список (только для оценки/рынка) — см. справочник по специальным типам столбцов ниже
«Sign»	Знак ожидаемого коэффициента: положительный, отрицательный или любой
«NAs»	Количество отсутствующих значений

6.3 Определение типов данных& Переопределения

glmnetUI автоматически определяет типы данных при импорте. Распознаются числовые, целочисленные и даты столбцы. Символьные столбцы, которые выглядят как даты (типичные форматы дат), классифицируются как «Date».

Вы можете переопределить любое распознавание, изменив значение в раскрывающемся списке «Тип». Изменение типа влияет на то, как столбец кодируется в матрице модели.

Столбцы «Текст», «TRUE/FALSE» и «R-фактор» автоматически рассматриваются как категориальные. Числовой столбец становится фактором только в том случае, если установить флажок «Фактор» — glmnetUI никогда не определяет категориальность на основе диапазона значений числового столбца, поэтому дискретные числовые предикторы, такие как количество ванн, остаются непрерывными, если вы не укажете иное. Настройка «Фактор» сохраняется в localStorage.

6.4 Ожидаемые признаки& Обеспечение соблюдения

Для каждого предиктора в раскрывающемся списке «Знак» указывается, ожидается ли, что его коэффициент будет положительным, отрицательным или может принимать значения

любого знака. После подгонки модели:

- Коэффициенты с неожиданными знаками помечаются в таблице «Коэффициенты»
- Предупреждения о знаках отображаются на вкладке «Сводка»

Если в параметрах `glmnet` (глава 7) установлен флажок «Enforce Sign Constraints», модель вынуждена вычислять коэффициенты, соответствующие ожидаемым знакам. Это реализовано с помощью функций `upper.limits` и `lower.limits` в пакете `glmnet`.

6.5 Специальные типы столбцов Ссылка

В режимах «Оценка» и «Рынок» в раскрывающемся списке «Специальное» доступны следующие параметры:

Взвешивание:

- `weight` — Столбец с весами наблюдений (доступна во всех режимах; допускается только одна; строки с весом = 0 исключаются из подгонки)

Дата& . Типы времени:

- `contract_date` — Запускает автоматический расчет показателя «`sale_age`» с даты оценки
- `listing_date` — Используется в качестве резервного варианта для расчета показателя «Days on рынке»
- `dom` — Определяет столбец «Дни на рынке»
- `sale_age` — Указывает существующий столбец «Срок продажи», благодаря чему любой столбец может служить показателем срока продажи вместо того, чтобы `glmnetUI` вычислял его на основе `contract_date`

Типы денежных единиц:

- `concessions` — Определяет условия скидок при продаже (кредиты продавца, поощрения для покупателя и т. д.)

Типы транзакций:

- `sale_type` — Обозначает столбец с типом продажи/транзакции

Размер& . Типы мест:

- `latitude` — Значения автоматически округляются до 3 знаков после запятой
- `longitude` — Округление аналогично широте
- `area` — Идентификатор рыночной зоны или района (например, `area_id`, район, рынок_area)
- `living_area` — Включает расчёт остаточной стоимости на квадратный фут (`residual_sf` и `cqa_sf`)
- `lot_size` — Столбец «Размер участка»
- `site_dimensions` — Группируется по площади участка (например, «75x120»)

Возрастные группы:

- `actual_age` — Столбец «Возраст объекта недвижимости»
- `effective_age` — Фактический возраст объекта недвижимости

Типы дисплеев:

- `display_only` — Столбец включается в экспортные файлы Excel, но полностью исключается из процесса подгонки модели. Используйте эту настройку для полей адреса, номеров MLS, идентификаторов земельных участков или других справочных данных. Эту настройку можно применить к нескольким столбцам.

Совет: Столбцы, помеченные как «display_only» (не используемые в модели), остаются в наборе данных и отображаются при экспорте в Excel, но полностью исключаются из таблицы предикторов. Используйте эту функцию для столбцов с идентификаторами, адресами или другими справочными полями.

7 Выбор параметров

Раздел 5 боковой панели — «Параметры вызова glmnet» — предоставляет доступ ко всем параметрам настройки модели эластичной сети. У каждого параметра есть синий значок справки (?) с всплывающей подсказкой, содержащей пояснение.

7.1 Alpha (параметр микширования)

Параметр alpha определяет тип регуляризации:

Значение альфа	Тип	Поведение
0	Ридж	Пропорционально уменьшает все коэффициенты; никогда не устанавливает их значение равным нулю
1	Лассо	Может устанавливать коэффициенты равными нулю (отбор переменных)
от 0 до 1	Эластичная сеть	Сочетание методов «Ридж» и «Лассо»

Доступны два метода выбора альфа-значений:

- Фиксированное значение — Установите одно значение альфа с помощью ползунка (по умолчанию: 1,0)
- Поиск по таблице — Проведите тестирование нескольких значений альфа и выберите то, которое дает наименьшую ошибку CV. Настройте диапазон и количество значений.

7.2 Lambda Selection

Параметр lambda определяет степень регуляризации:

Метод	Описание
Перекры́стная проверка	Рекомендуется. Проводит тестирование множества значений lambda с использованием k-кратной перекры́стной проверки.
Руководство	Укажите конкретное значение lambda, если у вас есть на то основание.

При использовании перекры́стной проверки доступны два варианта выбора лямбды:

- `lambda.1se` (по умолчанию) — Наибольшее значение lambda в пределах 1 стандартной ошибки от минимального значения. Позволяет получить более простые и лучше регуляризованные модели. Рекомендуется для работ по оценке.
- `lambda.min` — Значение лямбды, минимизирующее ошибку перекры́стной проверки. Позволяет получить модель, наилучшим образом подходящую к данным, но потенциально более сложную.

Параметр «Количество блоков CV» определяет, как данные будут разделены для перекры́стной проверки (по умолчанию: 10 блоков).

7.3 Семья

Выберите семейство распределений для вашей зависимой переменной:

Семейство	Пример применения
gaussian	Непрерывные значения (например, цена продажи). Наиболее распространенный случай.
binomial	Бинарные исходы (например, продано/не продано).
poisson	Данные, выраженные в виде количества (например, количество продаж).

7.4 Стандартизировать

Если этот флажок установлен (по умолчанию), перед подгонкой все предикторы нормируются так, чтобы их среднее значение равнялось 0, а стандартное отклонение — 1. Это гарантирует, что штрафная функция будет одинаково учитывать все предикторы независимо от их исходного масштаба. Коэффициенты возвращаются в исходном масштабе. Обычно этот флажок следует оставлять установленным.

7.5 Ограничения знаков

Если включена опция «Принудительное соблюдение ограничений по знакам», коэффициенты ограничиваются таким образом, чтобы они соответствовали ожидаемым знакам, заданным в таблице переменных (глава 6):

- Знак «положительный» принудительно устанавливает коэффициент ≥ 0
- Знак «отрицательный» принудительно устанавливает коэффициент ≤ 0
- Переменные, для которых установлено значение «либо», не имеют ограничений

Это реализуется с помощью параметров `upper.limits` и `lower.limits` в файле `glmnet()`.

7.6 «Расслабленное лассо»

Обычный метод Lasso применяет одинаковый штраф как при отборе переменных, так и при оценке коэффициентов, что может привести к чрезмерному сжатию важных коэффициентов в сторону нуля. Метод Relaxed Lasso разделяет эти этапы:

1. Сначала выберите переменные с помощью инструмента «Лассо» (штраф определяет, какие коэффициенты останутся в модели)
2. Затем выполните повторную подгонку оставшихся переменных с меньшим штрафом или без него для получения менее смещенных оценок

Ползунок «Gamma» регулирует степень смягчения:

Гамма-эффект	Эффект
0	Полностью релаксированная подгонка (переподгонка по OLS, без какого-либо сжатия)
1	Без релаксации (так же, как в обычном <code>glmnet</code>)
Межгрупповой	Сочетает релаксированную и штрафную подгонки

При перекрёстной проверке коэффициент гамма выбирается автоматически для обеспечения оптимальной производительности.

7.7 Начальные значения для генератора случайных чисел

Ввод текста «Random seed», предварительно заполненный случайным целым числом, обеспечивает воспроизводимость распределения блоков при перекрёстной проверке — перед

подгонкой вызывается функция `set.seed()`, а использованное начальное значение отображается на вкладке «Glmnet Output» и в сообщении о статусе подгонки. После каждой подгонки автоматически генерируется новое начальное значение (с использованием системного времени, а не генератора случайных чисел R). Последние 5 начальных значений, использованных для текущего файла, отображаются в виде активных ссылок для удобного доступа; история начальных значений сбрасывается при загрузке нового файла.

7.8 Матрица взаимодействий

Интерактивная верхнетреугольная матрица позволяет управлять тем, какие пары переменных могут взаимодействовать. В каждой ячейке находится флажок — если он установлен, это означает, что член взаимодействия $x1:x2$ включен в матрицу модели.

- «Разрешить всё» — Установить флажки для всех пар взаимодействий
- «Снять всё» — Снять флажки со всех пар (по умолчанию в новом файле все флажки сняты)
- Щелкните по имени переменной — Включить/выключить все взаимодействия для этой переменной

Настройки взаимодействия сохраняются в локальном хранилище браузера (`localStorage`) для каждого файла ввода и восстанавливаются при повторной загрузке того же файла.

Если какое-либо взаимодействие включено, экспортированный отчет содержит матрицу «Включенные взаимодействия» (предикторы по обеим осям, галочка в тех случаях, где был включен термин продукта « $x1:x2$ »), отображаемую в альбомной ориентации и повторяющую матрицу «Допустимые взаимодействия» из `earthUI`. Обратите внимание, что коэффициент включенного термина все же мог быть сведен к нулю с помощью инструмента «Лассо».

Совет: При большом количестве предикторов метод Lasso автоматически обнуляет ненужные термины взаимодействия. Разумным подходом будет начать с параметра «Allow All» и позволить модели самостоятельно выбрать нужные термины.

Warning

Включение взаимодействий в модели `glmnet` приведет к корректировкам, которые может быть сложно интерпретировать и объяснить в суде или в ходе аудита. Используйте взаимодействия с осторожностью и убедитесь, что полученные корректировки являются обоснованными.

7.9 Блок от основного эффекта (только взаимодействие)

Щелкните правой кнопкой мыши по имени переменной в матрице взаимодействий, чтобы заблокировать её включение в модель в качестве основного эффекта. После имени переменной появится выделенный жирным шрифтом индикатор «1». При блокировке:

- Переменная удаляется из слагаемых основной формулы
- Переменная по-прежнему может фигурировать в слагаемых взаимодействия ($x1:x2$)
- Снова щелкните правой кнопкой мыши, чтобы разблокировать
- Состояние сохраняется в `localStorage` для каждого файла ввода

Типичный пример применения: при моделировании временных корректировок, которые должны зависеть от размера недвижимости, следует заблокировать влияние «`sale_age`» в качестве основного эффекта и допустить его только во взаимодействии с «`living_area`». В результате получается взаимодействующий член «`sale_age:living_area`» вместо постоянного коэффициента «`sale_age`» — темпы роста стоимости с течением времени варьируются в зависимости от размера дома.

7.10 Расширенные параметры

Сворачиваемый раздел «Advanced» в нижней части страницы «Call Parameters» в glmnet открывает доступ к дополнительным настройкам. Для них используются разумные значения по умолчанию, но они отображаются на экране, чтобы все настройки модели можно было задокументировать для судебного разбирательства или аудита.

- Коэффициент Lambda Min — Отношение наименьшего значения лямбды к наибольшему в траектории регуляризации. По умолчанию: 0,00001. Для небольших наборов данных (<50 наблюдений) более высокое значение (например, 0,01) может улучшить результаты.
- Количество значений лямбды — Количество значений лямбды в траектории. По умолчанию: 100.
- Метрика потерь при перекрёстной проверке — Функция потерь для перекрёстной проверки: MSE (по умолчанию для гауссового распределения), MAE (устойчивая к выбросам) или Deviance.
- Порог сходимости — Порог сходимости метода координированного спуска. По умолчанию: 1e-07.
- Максимальное количество итераций — Максимальное количество итераций метода координированного спуска. Значение по умолчанию: 100 000. Увеличьте, если появляются предупреждения о сходимости.
- Включить подгонку — включить член подгонки (константу). Значение по умолчанию: включено. Член подгонки является «базовым значением» в оценке RCA.

7.11 Настройки по умолчанию

Существует три параметра, определяющих порядок инициализации параметров при загрузке файла:

- Использовать последние настройки для файла ввода — восстановить точные настройки, использовавшиеся в последний раз с этим файлом
- Использовать настройки по умолчанию — применить сохраненные глобальные настройки по умолчанию
- Настройки по умолчанию glmnet — сбросить до заводских настроек ($\alpha = 1$, CV lambda.1se, 10 блоков, гауссово распределение, стандартизация включена)

Нажмите «Сохранить текущие настройки в качестве стандартных», чтобы сохранить все текущие настройки в качестве глобальных настроек по умолчанию для будущих файлов.

8 Подгонка модели

8.1 Кнопка «Fit»

В разделе 6 боковой панели находится кнопка «Fit Glmnet Model» (Построить модель Glmnet). При нажатии на эту кнопку модель будет запущена с текущими настройками.

8.2 Что происходит во время подгонки

При нажатии кнопки «Подгонка» glmnetUI:

- 1 . Подготовка данных — удаление строк с значениями «NA», применение весов (если указаны), исключение строки «объект оценки» (в режиме оценки) и кодирование факторных переменных в виде фиктивных столбцов.
- 2 . Постраивает матрицу модели — создает матрицу дизайна, включая все допустимые термы взаимодействия из матрицы взаимодействий.
- 3 . Применяет ограничения — если включено принудительное соблюдение знаков, устанавливает `upper.limits` и `lower.limits` для коэффициентов.
- 4 . Выполняет перекрёстную проверку — если выбрана CV, вызывает `cv.glmnet()` для нахождения оптимального значения `lambda`. Если включен метод «Relaxed Lasso», вызывается функция `cv.glmnet(..., relax = TRUE)` .
- 5 . Подгоняет окончательную модель — сохраняет подгоняемую модель, коэффициенты, прогнозы и остатки.
- 6 . Обновляет все вкладки результатов — «Сводка», «Коэффициенты», «Уравнение», «Важность переменных», «Вклады», «ANOVA» и «Диагностика» заполняются данными.

В сообщении о состоянии под кнопкой «Fit» отображается количество наблюдений, исключенных строк и выбранное значение лямбды.

9 Вкладки «Результаты»

9.1 Предварительный просмотр данных

Отображает импортированные данные в виде интерактивной таблицы DataTable. В режимах «Оценка» и «Рынок» предварительный просмотр разбит на две таблицы: «Объект оценки» (строка 1) и «Сопоставимые сделки» (строки 2 и далее).

Каждая ячейка отображается в одной строке и обрезается до ширины столбца, поэтому широкие поля свободного ввода (например, примечания к объектам недвижимости) не растягивают строку вниз по странице. Дважды щелкните любую ячейку, чтобы открыть всплывающее окно с её полным, необрезанным содержимым.

9.2 Уравнение

Отображает уравнение модели, созданной в процессе подгонки, сгенерированное в LaTeX с помощью MathJax. Показывает все ненулевые коэффициенты с надлежащим математическим форматированием:

- Положительные коэффициенты обозначаются знаком «+»
- Отрицательные коэффициенты обозначаются знаком «−»
- Слагаемые взаимодействия отображаются как « $x_1 \times x_2$ »
- Числа форматируются с запятыми и соответствующим количеством знаков после запятой

9.3 Корреляция

Матрица тепловой карты, отображающая корреляции Пирсона между всеми числовыми предикторами и зависимой переменной. Доступна сразу после импорта данных — не требуется подгонка модели.

- Цветовой градиент: синий (−1) → белый (0) → красный (+1)
- Значения корреляции отображаются в каждой ячейке
- Размер текста и осей адаптируется в зависимости от количества переменных
- Полезно для выявления мультиколлинеарности перед подгонкой

9.4 Резюме

Статистика подгонки модели, отображаемая в виде карточек:

- R^2 — доля объясненной дисперсии
- Adj R^2 — скорректированное значение R^2 с учетом количества предикторов
- GR R^2 — обобщенное R^2 , полученное на основе коэффициента отклонения из glmnet при выбранном значении лямбды
- CV R^2 — перекрёстная проверка R^2 (при использовании CV)
- RMSE — среднеквадратичная ошибка
- MAE — средняя абсолютная ошибка

В режимах оценки и рынка отображаются дополнительные карточки:

- Медианное соотношение — медиана отношения (прогнозное / фактическое)
- COD — коэффициент дисперсии (чем ниже, тем лучше)
- PRD — ценовой дифференциал

Предупреждение о переобучении появляется, когда при обучении R^2 превышает CV R^2 более чем на 0,1.

Под карточками для удобства расположена таблица коэффициентов с указанием знаков, которая дублирует вкладку «Коэффициенты».

9.5 Коэффициенты

Таблица всех ненулевых коэффициентов, в которой указано:

- Имя переменной
- Значение коэффициента
- Ожидаемый знак (из конфигурации переменной)
- Флаг предупреждения о знаке, если направление коэффициента не соответствует ожиданиям

9.6 Важность переменных

Величина стандартизированного коэффициента: $|\beta| \times \text{sd}(x)$ для каждого предиктора, агрегированная по фиктивным столбцам для факторных переменных. Два режима отображения:

- Интерактивный Plotly (при наличии установленного пакета plotly): Горизонтальная гистограмма с всплывающими подсказками, отображающими точные значения значимости, коэффициента и относительной доли.
- Статический ggplot2 (резервный вариант): Горизонтальная гистограмма с подписями на осях, отформатированными с запятыми.

Под диаграммой в таблице DataTable приведены значения переменных, их значимость, коэффициенты и относительные значения %.

9.7 Материалы

Графики частичных эффектов по каждой переменной, отражающие вклад каждого предиктора в прогнозы модели. Выберите переменную из раскрывающегося списка.

Для числовых предикторов: диаграмма рассеяния, отражающая зависимость значений переменной по оси x от её вклада ($\beta \times x$), с красным отрезком линии и подписью, указывающей наклон. Метка наклона показывает маргинальный эффект на единицу (например, +1,234.56/unit) с адаптивными единицами для переменных с небольшим диапазоном значений, таких как широта/долгота (например, +0.12/0.001). Подзаголовок показывает значение пересечения с осью (базис).

Для факторных переменных: бокс-плат, отражающий распределение значений вклада по уровням факторов.

Для взаимодействий между двумя переменными: три наложенные друг на друга визуализации — трехмерный график поверхности (интерактивный Plotly, если доступен; в противном случае — статический persp()) совместного вклада, диаграмма рассеяния с точками, окрашенными в зависимости от вклада, и тепловая карта среднего вклада по всем диапазонам значений обеих переменных.

9.8 ANOVA

Таблица декомпозиции дисперсии, в которой для каждого предиктора указано:

- SS — Сумма квадратов: $\sum(\text{contribution}_j - \overline{\text{contribution}_j})^2$
- %модели SS — Доля от общей суммы квадратов модели
- Коэффициент — Доминирующий коэффициент для данного предиктора

Включает строку пересечения и строку «TOTAL MODEL». Термы взаимодействия (содержащие «:») отображаются в виде отдельных групп.

9.9 Диагностика

Четыре графика для диагностики с крупным шрифтом (базовый размер 16pt), делениями по осям с шагом 15 и метками в формате с запятой (без научной нотации):

- 1 . Траектория коэффициентов — Показывает, как изменяются все коэффициенты при увеличении $\log(\lambda)$. Линии окрашены в цвета, соответствующие переменным. Помогает визуализировать траекторию регуляризации и отбор переменных.
- 2 . Ошибка перекрёстной проверки — Ошибка перекрёстной проверки как функция $\log(\lambda)$. Точки отображают среднюю ошибку перекрёстной проверки; полосы погрешности показывают ± 1 стандартную ошибку. Пунктирные вертикальные линии обозначают $\lambda_{\min}$ (синий) и λ_{1se} (зеленый).
- 3 . Фактические значения против прогнозируемых — Диаграмма рассеяния наблюдаемых значений против прогнозируемых. Точки должны группироваться вокруг 45-градусной пунктирной линии отсчета. Более широкое рассеяние указывает на большую ошибку прогнозирования.
- 4 . Остатки по сравнению с результатами подгонки — Диаграмма рассеяния остатков по сравнению с результатами подгонки. Должно наблюдаться случайное рассеяние вокруг нулевой линии (пунктирной). Определённые закономерности указывают на некорректную спецификацию модели (например, нелинейность, гетероскедастичность).

На каждом графике есть кнопка «Скачать PNG» для сохранения изображения с разрешением 150 DPI.

9.10 Вывод Glmnet

Исходный вывод модели, отображающий: использованное начальное значение генератора случайных чисел, вывод модели, выбранное значение λ , $\lambda_{\min}/\lambda_{1se}$, γ (если включен режим Relaxed Lasso) и полный вектор коэффициентов.

9.11 Отчёт «

»

Настройте поля отчета (имя оценщика, адрес объекта недвижимости, дату отчета, номер дела) и экспортируйте его в формат Word (.docx) или PDF прямо с этой вкладки. Полную информацию о рабочем процессе подготовки отчетов на базе Quarto (HTML, Word или PDF) см. в главе 13.

10 Загрузка данных

После подгонки кнопка «Скачать вывод (Excel)» (боковая панель, раздел 7, доступна во всех режимах) позволяет экспортировать файл Excel с прогнозами и диагностикой в папку вывода проекта. Заголовок раздела звучит как «Скачать расчетные цены продажи & остатки» в режимах «Оценка» и «Рынок» и как «Скачать расчетные целевые переменные & остатки» в режиме «Общий».

10.1 Столбцы вывода

Колонка	Описание
est_<target>	Прогноз по модели (например, est_sale_price)
residual	Фактическое значение— Прогнозируемое значение
cqa	Оценка «Состояние-Качество-Привлекательность» (по шкале от 0 до 10)
residual_sf	Остаточная площадь / жилая площадь (если указано living_area)
cqa_sf	Показатель CQA, рассчитанный на основе рейтинга по residual_sf
<var>_contribution	Вклад каждой переменной в прогноз
basis	Значение пересечения (одинаковое для всех строк)

После успешного завершения на кнопке загрузки появляется белая галочка.

10.2 Результаты CQA: балл

балл

CQA оценивает остаток каждой строки по сравнению со всеми остальными по шкале от 0 до 10:

- Высокий показатель CQA ($\sim 9-10$) : продан значительно дороже прогнозируемой цены — вероятно, отличное состояние/качество/привлекательность
- Низкий показатель CQA ($\sim 0-1$) : продан значительно дешевле прогнозируемой цены — вероятно, неудовлетворительное состояние или продажа в связи с финансовыми затруднениями
- Показатель CQA ~ 5 : близок к медиане

В режимах «Оценка» и «Рынок» строки сортируются по параметру « residual_sf » в порядке убывания, если указан столбец « living_area ».

11 Расчеты RCA& Скачивание

Рабочий процесс RCA (Метод остаточных ограничений) доступен в режимах «Оценка» и «Анализ рыночной зоны» после подгонки модели. Последующая таблица сравнительных продаж (глава 12) предназначена исключительно для оценки.

11.1 Открытие диалогового окна RCA

Нажмите кнопку «Рассчитать корректировки RCA» & «Скачать» в разделе 8 боковой панели. Выберите:

- Тип оценки CQA или «CQA на SF ». Если выбран вариант «CQA на SF», остаточный балл объекта оценки нормируется с учётом жилой площади.
- Введите балл CQA объекта оценки (0,00–9,99, по умолчанию 5,00)

11.2 Интерполяция баллов CQA

- 1 . Баллы CQA и остатки для аналогов отсортированы
- 2 . Линейная интерполяция сопоставляет значение CQA с остатком
- 3 . Значение для объекта оценки = прогноз модели + интерполированный остаток

11.3 Столбцы вывода

Колонка	Описание
subject_value »	Прогноз по модели + интерполированные остатки
<var>_adjustment	Вклад объекта оценки– Вклад «comp»
residual_adjustment	Остатки объекта оценки– Остатки «comp»
net_adjustments	Сумма всех корректировок
gross_adjustments	Сумма абсолютных корректировок
adjusted_sale_price	Цена продажи «comp» + чистые корректировки

После успешного выполнения вычислений на кнопке появляется белая галочка.

12 Таблица сравнения продаж

Таблица сравнения продаж доступна только в режиме оценки, после расчета корректировок RCA (шаг 8). Она формирует отформатированную книгу Excel, подходящую для включения в отчёты об оценке.

12.1 Диалоговое окно выбора аналогов

При нажатии кнопки «Сгенерировать таблицу продаж» & «Скачать» в боковой панели в разделе 9 открывается модальное диалоговое окно со списком всех аналогов из вывода анализа RCA. Аналоги разделены на две группы:

- Рекомендуемые аналоги (уже отмечены): Валовая корректировка < 25% от цены продажи, отсортированные по дате продажи (сначала самые последние).
- Другие аналоги (по умолчанию не отмечены): Остальные аналоги, отсортированные по дате продажи.

Выберите до 30 аналогов, а затем нажмите «Сформировать таблицу продаж», чтобы создать книгу.

12.2 Макет книги

Сгенерированная книга содержит до 10 листов, на каждом из которых размещены 3 аналога рядом друг с другом:

- Строки заголовков : адреса объекта оценки и покупателя, цены продажи, даты продажи и количество дней на рынке
- Сгруппированные строки : местоположение (широта, долгота, площадь), размер участка (площадь участка, габариты участка) и возраст (фактический возраст, эффективный возраст) с заголовками групп и подпунктами
- Строки с переменными модели : по одной строке на каждый ненулевой коэффициент модели, показывающие вклад объекта оценки, вклад аналогов и корректировки
- Строки с остаточными характеристиками : пустые, незаблокированные ячейки для ввода оценщиком характеристик, не учтенных моделью (например, состояние, качество, вид)
- Сводные строки : чистая корректировка%, валовая корректировка% и скорректированная цена продажи с активными формулами Excel

12.3 Защита листа & Остатки ячейки

Каждый лист защищен паролем для предотвращения случайных изменений, однако ячейки с остаточными значениями характеристик явно разблокированы. Это позволяет оценщикам вводить значения для характеристик, не включенных в модель, при этом сохраняя расчеты корректировок, основанные на формулах.

13 Скачивание отчетов

13.1 «Двухэтапный кварто» Рабочий процесс

Экспорт отчёта из боковой панели представляет собой двухэтапный рабочий процесс в Quarto:

1. «Генерация отчёта Quarto» (раздел 10 боковой панели) — создаёт самодостаточный набор Quarto (исходный файл .qmd , а также графики, report_data.rds и reference.docx) в папке вывода glmnet проекта.
2. «Convert Quarto отчёт» (раздел 11 в боковой панели) — преобразует файл .qmd в выбранные вами форматы вывода. По умолчанию указывается путь к самому последнему созданному набору, но вы можете выбрать любой .qmd .

13.2 Форматы отчётов

На этапе «Преобразование» доступны три формата:

- HTML — Автономный HTML-документ. Работает на всех платформах без дополнительного программного обеспечения, кроме pandoc. Рекомендуемый вариант по умолчанию.
- Word (.docx) — Отформатированный документ. Подходит для редактирования и распространения. Работает на всех платформах.
- PDF — Сгенерирован с помощью Quarto на основе LaTeX. Размер бумаги соответствует настройкам локали (Letter или A4). Требуется установка LaTeX (см. приложение «Системные требования»). Если LaTeX не обнаружен, опция PDF автоматически скрывается из списка флажков форматов.

Отчёты формируются с помощью Quarto, если это возможно, а в противном случае — с помощью rmarkdown. На вкладке «Отчёт» также доступны кнопки «Экспорт в Word» и «Экспорт в PDF».

13.3 Содержание отчёта

Отчёты включают всё содержимое вкладок:

- Описание набора данных (имя файла, наблюдения, зависимая переменная, предикторы, цель)
- Спецификация модели (альфа, лямбда, семейство, стандартизация, ослабленные ограничения, ненулевые коэффициенты)
- Сводка результатов (R^2 , скорректированное R^2 , обобщённое R^2 , CV R^2 , RMSE, MAE)
- Уравнение модели (с отображением в LaTeX)
- Таблица коэффициентов
- Важность переменных (гистограмма и таблица)
- Графики вклада по каждой переменной, соответствующие вкладке «Вклады»:
 - Числовые предикторы : диаграмма рассеяния с линией линейной подгонки
 - Факторные переменные : бокс-плот по уровням
 - Взаимодействия : диаграмма рассеяния с цветовой маркировкой по вкладу, тепловая карта среднего вклада и статический снимок 3D-поверхности (в перспективе)
- Матрица включённых взаимодействий (когда включено любое взаимодействие) — предикторы на обеих осях, с галочкой там, где был включён слагаемое $x_1:x_2$, отображается в альбомной ориентации

- Тепловая карта матрицы корреляций
- Таблица декомпозиции ANOVA
- Диагностические графики:
 - Путь коэффициентов
 - Ошибка перекрёстной проверки
 - Фактические и прогнозируемые значения
 - Остатки и результаты подгонки
 - Нормальный график Q-Q
- Вывод модели

Во всех подписях к осям и легендах цветов используются числа с форматом с запятыми (без научной нотации).

После успешного создания отчёта на кнопке отчёта появляется белая галочка, а во время оформления отчёта отображается таймер, показывающий время, прошедшее с начала процесса.

14 Сравнение с earthUI

Приложение «Earth» и его модуль постобработки glmnet представляют собой единый набор инструментов для регрессионного моделирования. Они используют общий формат данных, специальные типы столбцов, рабочий процесс RCA и демонстрационные наборы данных, но построены на разных механизмах моделирования.

14.1 Основные различия

Функции	glmnetUI	earthUI
Методы	Эластичная сеть (glmnet)	MARS/Earth (earth)
Отбор переменных	Автоматический отбор с использованием штрафа Lasso	Автоматический отбор с использованием метода «forward/backward pruning»
Зависимости коэффициентов	Только линейные эффекты	Кусочно-линейные (функции-шарниры)
Взаимодействия	Попарные через матрицу взаимодействий	До 3-й степени
Знаки коэффициентов	Можно задать с помощью ограничений	Невозможно задать напрямую
Смягчённая подгонка	Доступна смягчённая Lasso-подгонка	Неприменимо
Настройка альфа-коэффициента	Доступен поиск по таблице	Неприменимо (нет параметра смешивания)
Таблица сравнения продаж	Доступно (шаг 9)	Доступно (шаг 8)
Графики g-функций	Не применимо	Визуализация сгруппированных слагаемых
Форматы отчетов	HTML, Word, PDF	HTML, Word, PDF

14.2 Общие функции

Оба инструмента предоставляют:

- Импорт из CSV и Excel с помощью библиотеки snake_ Преобразование регистра в столбцах
- Поддержка локалей 31 страны с учетом разделителя в CSV, знака десятичной запятой и формата бумаги
- Специальные обозначения столбцов (контракт_ дата, жилая_ площадь, площадь, широта, долгота и т. д.)
- Автоматический расчёт срока действия договора_ Расчёт срока действия договора с даты оценки
- Переключение между тёмным и светлым режимами
- Сохранение настроек с помощью localStorage
- Тепловая карта корреляций, значимость переменных, вклады, ANOVA, диагностика
- Рабочий процесс оценки CQA и корректировки RCA
- Таблица сравнения продаж с выбором аналогов для сравнения и отформатированным выводом в Excel
- Совместимость с теми же демонстрационными наборами данных (Оценка_ 1.csv)

14.3 Когда использовать какой

- Используйте `glmnetUI` , если вам нужна простая, интерпретируемая линейная модель с автоматическим отбором переменных. Модель эластичной сети особенно хорошо подходит в тех случаях, когда у вас много предикторов и вы хотите, чтобы модель сама выбрала наиболее значимые из них.
- Используйте `earthUI` , если вы ожидаете нелинейных зависимостей (например, влияние жилой площади на изменение цены при разных размерах) или вам нужны функции-шарниры для учета пороговых эффектов. Модели `earth` также могут учитывать взаимодействия высших порядков (2-й и 3-й степени). `earthUI` дополнительно предлагает экспериментальную функцию извлечения комментариев на основе Prolog/DCG, которой нет аналога в `glmnetUI`.
- Используйте обе модели для сравнения линейных и нелинейных моделей на одних и тех же данных. Если обе модели дают схожие значения R^2 , более простая модель `glmnet` может быть предпочтительнее с точки зрения интерпретируемости.

15 Демо-набор данных: Оценка _ 1.csv

15.1 Описание

Демонстрационный набор данных входит в состав glmnetUI (и используется совместно с earthUI). Найти его можно следующим образом:

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

Файл содержит 1 502 сделки по продаже жилой недвижимости (плюс 1 объект оценки, указанный в 1-й строке) из смоделированного экспорта данных MLS. Данные отражают продажи односемейных домов на рынке, охватывающем несколько районов, с различными размерами, возрастом и местоположением объектов недвижимости.

Это не реальные данные, но они основаны на реальном районе в Северной Калифорнии. Вся идентифицирующая информация была изменена или удалена.

15.2 Столбцы

Колонка	Тип	Специальный тип	Описание
weight	числовой	вес	Вес наблюдения (0 = исключить из подбора)
id	числовой	display_only	Внутренний идентификатор записи
property_id	числовой	display_only	Идентификатор объекта недвижимости в системе MLS
listing_id	символьный	display_only	Номер объявления в системе MLS
parcel_number	символьный	display_only	Номер участка по оценке округа (APN)
street_address	символьный	display_only	Адрес объекта недвижимости
city_name	символьный	display_only	Город
postal_code	символьный	display_only	Почтовый индекс
county_name	символьный	display_only	Округ
contract_date	Дата	contract_date	Дата заключения договора купли-продажи (рассчитывается sale_age)
sale_age	числовой	—	Количество дней от даты заключения договора до даты оценки
sale_price	числовой	(целевая переменная)	Цена продажи — переменная ответа
living_sqft	числовой	living_area	Общая жилая площадь в квадратных футах
beds_total	целое число	—	Количество спален
baths_total	числовой	—	Общее количество ванных комнат (например, 2,5 = 2 полные + 1 половина)
lot_size	числовой	lot_size	Площадь участка в квадратных футах
area_id	целое число	area	Идентификатор района в системе MLS
age	числовой	actual_age	Возраст объекта недвижимости в годах
latitude	числовой	latitude	Широта (округленная до 3 знаков после запятой для модели)
longitude	числовой	longitude	Долгота (округленная до 3 знаков после запятой для модели)
garage_spaces	целое число	—	Количество парковочных мест
days_on_market	целое число	dom	Количество дней на рынке
listing_date	Дата	listing_date	Дата размещения объявления
sale_concessions	числовой тип	concessions	Уступки продавца

15.3 Рекомендуемое руководство по быстрому запуску

- 1 . Запустите процедуру glmnet: earthUI::launch_glmnet()
- 2 . Создайте новый проект (раздел 1) с целью «For оценка»
- 3 . Поместите файл Appraisal_1.csv в папку вводных данных проекта и выберите его в разделе 2 (при необходимости нажмите «Refresh»)
- 4 . Выберите в качестве цели « sale_price
- 5 . Назначьте специальные типы, как показано в таблице выше
- 6 . Включите предикторы: sale_age , living_sqft , baths_total , lot_size , area_id (в качестве фактора), age , latitude , longitude , garage_spaces
- 7 . Оставьте альфа = 1 (Lasso), CV с lambda.1se
- 8 . Нажмите «Fit Glmnet Model» (Построить модель Glmnet)

- 9 . Просмотрите вкладки «Summary» (Сводка), «Коэффициенты» (Коэффициенты) и «Диагностика» (Диагностика)
- 10 . Скачайте вывод (Шаг 7), просмотрите рейтинг CQA
- 11 . Рассчитайте корректировки RCA (Шаг 8) с баллом CQA ~5,00
- 12 . Сгенерируйте таблицу сравнения продаж (Шаг 9) с рекомендуемыми аналогами

16 Системные требования& Устранение неполадок

16.1 Поддерживаемые платформы

glmnetUI работает на macOS, Windows и Linux. Настоятельно рекомендуется использовать RStudio Desktop (версия 2023.06 и выше) — в него входит набор pandoc, необходимый для создания отчётов.

16.2 Дополнительные зависимости

- LaTeX (TinyTeX, MiKTeX или MacTeX): только для создания отчетов в формате PDF. Если программа не обнаружена, опция PDF автоматически скрывается. Установите с помощью: `tinytex::install_tinytex()`
- sysfonts + showtext : шрифт Roboto Condensed для графиков. При недоступности или отсутствии подключения к сети автоматически используется системный шрифт sans-serif.
- earth ($\geq 5.3.0$): Требуется только для конвейера импорта earthUI .

16.3 Примечания по платформе

macOS: наименьшее количество проблем. Установите TinyTeX для работы с PDF: `tinytex::install_tinytex()`

Windows: Хорошо работает с RStudio. На корпоративных/заблокированных компьютерах могут действовать ограничения на использование временных каталогов — приложение выдаст предупреждение в консоли, но продолжит работу без сохранения настроек.

Linux: Возможно, потребуются системные библиотеки: `sudo apt install libcurl4-openssl-dev libssl-dev libxml2-dev libsqlite3-dev libfontconfig-dev`

16.4 Плавная деградация

Отсутствующий компонент	Поведение
Отсутствует LaTeX	Параметр PDF скрыт. HTML и Word по-прежнему доступны.
Отсутствует интернет / ошибка шрифтов	Используется системный шрифт без засечек. Сообщение в консоли зарегистрировано.
Отсутствует RSQLite / файловая система только для чтения	Настройки не сохраняются. Приложение работает нормально.
Временный каталог недоступен для записи	Сбой генерации отчёта с явной ошибкой. Укажите TMPDIR в качестве пути к каталогу, доступному для записи.

16.5 Устранение неполадок в программе «

»

«Опция PDF недоступна»: выполните команду `tinytex::install_tinytex()` в R, перезапустите приложение.

«Настройки не сохраняются»: проверьте права доступа к каталогу пользовательских данных (macOS: `~/Library/Application Support/R/glmnetUI/`, Linux: `~/.local/share/R/glmnetUI/`, Windows: `%APPDATA%/R/data/glmnetUI/`).

Порт 7879 уже занят : запустите команду ‘ lsof -ti:7879 | xargs kill ‘ (macOS/Linux) или ‘ Stop-Process -Id (Get-NetTCPConnection -LocalPort 7879).OwningProcess ‘ (Windows PowerShell).

Список литературы

- Friedman, Jerome, Trevor Hastie и Robert Tibshirani (2010). «Regularization Paths for Generalized Linear Models via Coordinate Descent». В: Journal of Statistical Software 33.1, с. 1—22. DOI: [10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- Simon, Noah и др. (2011). «Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent». В: Journal of Statistical Software 39.5, с. 1—13. DOI: [10.18637/jss.v039.i05](https://doi.org/10.18637/jss.v039.i05).
- Tay, J. Kenneth, Balasubramanian Narasimhan и Trevor Hastie (2023). «Elastic Net Regularization Paths for All Generalized Linear Models». В: Journal of Statistical Software 106.1, с. 1—31. DOI: [10.18637/jss.v106.i01](https://doi.org/10.18637/jss.v106.i01).

mgcvUI: интерактивный интерфейс для пакета mgcv (обобщённые аддитивные модели)

Русский перевод, подготовленный с использованием искусственного интеллекта и
прошедший редакционную проверку; английская версия является основной.

William Bert Craytor

Том 1, Выпуск 1 · Лето 2026

Опубликовано: 1 июля 2026

DOI: [10.67330/xtv1pa25](https://doi.org/10.67330/xtv1pa25)

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

Эта статья открытого доступа распространяется на условиях [лицензии Creative Commons Attribution 4.0](#)

Аннотация. mgcvUI — графический интерфейс пользователя для пакета R mgcv, который подгоняет обобщённые аддитивные модели (GAM) с помощью штрафных регрессионных сплайнов с автоматическим выбором степени сглаживания. Он предлагает три режима работы — общее прогнозное моделирование, оценку недвижимости и анализ рыночной зоны — и проводит пользователя через импорт данных, настройку сглаживающих членов, подгонку модели, графики частных эффектов и загружаемые отчёты. В статье описаны требования к формату данных, рабочий процесс моделирования, форматы вывода и полный справочник возможностей mgcvUI.

Ключевые слова: GAM, обобщённые аддитивные модели, сглаживающие члены, штрафные регрессионные сплайны, mgcv, R, графический интерфейс, оценка недвижимости

Publisher's Note (July 2026): Shortly before press, mgcvUI was consolidated into earthUI (version 0.11.0 and later), which now provides the GAM workflow described here alongside the MARS and elastic-net workflows. mgcvUI remains available in its final released form (0.3.0) at <https://github.com/wcraytor/mgcvUI>; future development continues in earthUI. This article applies to mgcvUI as released.

Содержание

1	Введение	4
1.1	История и предыстория	4
1.1.1	Обобщённые аддитивные модели	4
1.1.2	Пакет mgcv	4
1.1.3	Сравнение с MARS (Земля) и эластичной сетью (glmnet)	5
1.1.4	earthUI-mgcvUI Конвейер	6
1.1.5	Рекомендуемый рабочий процесс	6
1.2	Что такое mgcvUI?	6
1.3	Три режима работы	7
1.4	Недвижимость — специальные материалы	7
1.5	Начало работы	8
2	Требования к данным для ввода MLS	9
2.1	Формат файла & Структура	9
2.2	Обязательные столбцы для оценки Режим	9
2.3	Особые правила именования столбца	10
2.4	Качество данных & Полнота данных	10
2.5	Размещение объекта оценки в строке	11

3	Режим общего назначения	12
3.1	Обзор	12
3.2	Рабочий процесс с использованием боковой панели	12
3.3	Импорт из earthUI (Необязательно)	13
3.4	Вкладки главной панели	13
3.5	Сохранение настроек	14
3.6	Темный режим	14
3.7	Язык и регион & Региональные настройки	14
4	Режим оценки	16
4.1	Обработка строк объекта оценки	16
4.2	Дата оценки & Возраст, с которого разрешена продажа	16
4.3	Специальный столбец Обозначения	16
4.4	Обзор корректировок RCA	16
5	Режим анализа рыночной зоны	17
5.1	Отличия от режима оценки	17
5.2	Когда следует использовать режим «Рынок»	17
6	Отбор переменных	18
6.1	Целевая переменная	18
6.2	Преобразование отклика	18
6.3	Таблица предикторов	18
6.4	Определение типов данных & Переопределения	19
6.5	Специальные типы столбцов См.	19
7	Выбор параметров	21
7.1	Предустановки параметров	21
7.2	Семья	21
7.3	Метод	21
7.4	Gamma (коэффициент сглаживания)	22
7.5	Перекрёстная проверка	22
7.6	Выбрать (Сжатие)	22
7.7	Базовые данные по умолчанию	22
7.8	Значение k по умолчанию (базисная размерность)	23
7.9	Тип тензора	23
7.10	Допустимые взаимодействия	23
7.11	Взаимодействия «фактор сглаживание»	24
7.12	Расширенные параметры	24
8	Подгонка модели	25
8.1	Кнопка «Fit»	25
8.2	Что происходит во время подгонки	25
8.3	Общая информация Подгонка & Ход работы	25
9	Вкладки «Результаты»	26
9.1	Данные	26
9.2	Уравнение	26
9.3	Резюме	26
9.4	Важность переменных	27
9.5	Графики вкладов	27
9.5.1	Univariate Smooth Terms — $s(x)$	27

9.5.2	Фактор-сглаживание Термины — $s(x, by = factor)$	27
9.5.3	Interaction Terms — $ti(x_1, x_2)$ and $te(x_1, x_2)$	27
9.5.4	Параметрические члены — линейные и факторные	27
9.6	Корреляция	28
9.7	Диагностика	28
9.8	Корректировки RCA	28
9.9	Аноа	28
9.10	Вывод Mgcsv:	28
9.11	Проверка знаков	28
9.12	Concurvity	29
9.13	Проверка исходных данных	29
10	Загрузка данных	30
10.1	Столбцы вывода	30
10.2	Результаты CQA: балл	30
11	Расчеты RCA& Скачивание	31
11.1	Открытие диалогового окна RCA	31
11.2	Интерполяция баллов CQA	31
11.3	Столбцы вывода	31
12	Таблица сравнения продаж	32
12.1	Диалоговое окно выбора аналогов	32
12.2	Макет книги	32
12.3	Защита листа & Остатки ячейки	32
13	Скачивание отчетов	33
13.1	Шаг 1: Создание отчёта Quarto Отчёт «	33
13.2	Шаг 2: Конвертация Quarto Отчёт «	33
13.3	Содержание отчёта	33
14	Экспорт функций модели	35
14.1	Функции R	35
14.2	Функции Python	35
14.3	Заголовочный файл C++	35
14.4	База данных SQLite	35
15	Сравнение с earthUI и glmnetUI	37
15.1	Основные различия	37
15.2	Общие функции	37
15.3	Рекомендуемый рабочий процесс	38
16	Демо-набор данных: Оценка _ 1.csv	39
16.1	Описание	39
16.2	Столбцы	39
16.3	Рекомендуемое руководство по быстрому запуску	40
17	Системные требования& Устранение неполадок	41
17.1	Поддерживаемые платформы	41
17.2	Версия R:	41
17.3	Необходимые пакеты R:	41
17.4	Дополнительные зависимости	42
17.5	Примечания по конкретным платформам	42

17.5.1 macOS	42
17.5.2 Windows	42
17.5.3 Linux	42
17.6 Плавная деградация	43
17.7 Устранение неполадок в программе «	43
17.7.1 «Параметр PDF недоступен»	44
17.7.2 «Настройки не сохранятся»	44
17.7.3 «График диагностики недоступен»	44
17.7.4 “Отчёт об ошибке: количество элементов, которые необходимо заменить ...”	44
17.7.5 Порт 7880 уже занят	44
17.7.6 Ошибка установки пакета в Linux	44

1 Введение

1.1 История и предыстория

1.1.1 Обобщенные аддитивные модели

Обобщённые аддитивные модели (GAM) были формализованы Тревором Хэсти и Робертом Тибширани в их монографии 1990 года. GAM расширяют линейную модель путём замены каждого линейного члена $\beta_j x_j$ на функцию сглаживания $f_j(x_j)$, в результате чего модель принимает следующий вид:

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_p(x_p) + \varepsilon$$

Каждая функция f_j оценивается на основе данных с помощью регрессионного сплайна с штрафным термином. Процесс сглаживания позволяет модели отражать нелинейные зависимости без необходимости предварительного указания пользователем формы функции. Штрафной термин регулирует степень «волнистости» каждой сглаженной функции, предотвращая переобучение и одновременно обеспечивая достаточную гибкость для отслеживания реальных закономерностей.

1.1.2 Пакет mgcv

Пакет R «mgcv» (Mixed GAM Computation Vehicle) был разработан Саймоном Вудом из Университета Бата (Wood 2017). Впервые выпущенный в 2001 году, он в настоящее время является стандартной реализацией GAM в R и поставляется с каждой установкой R. Лежащие в его основе методы описаны в серии статей, посвящённых стабильной оценке нескольких параметров сглаживания (Wood 2004), сплайнам регрессии «тонкой пластины» (Wood 2003), быстрой стабильной оценке REML (Wood 2011), а также общему выбору сглаженных моделей (Wood, Pya и Säfken 2016). Ключевые особенности включают:

- Автоматический выбор степени сглаживания : параметр сглаживания для каждого члена оценивается в ходе подгонки модели с использованием методов REML (ограниченное максимальное правдоподобие), GCV (обобщённая перекрёстная проверка) или ML (максимальное правдоподобие).
- Несколько типов базисов : сплайны регрессии тонкой пластины (tp), кубические регрессионные сплайны (cr), P-сплайны (ps) и B-сплайны (bs), каждый из которых обладает своими свойствами.
- Сглаживание с помощью тензорного произведения : $te()$ и $ti()$ позволяют обеспечить плавное взаимодействие между переменными без «проклятия размерности».

- Отбор переменных : опция « `select = TRUE` » добавляет дополнительный штраф, который может свести сглаживающие члены к нулю, эффективно осуществляя отбор переменных.
- Диагностика конкурвности : измеряет аналог мультиколлинеарности для сглаживающих членов.

1.1.3 Сравнение с MARS (Земля) и эластичной сетью (glmnet)

Приложение «Earth» и его два модуля постобработки (glmnet и mgcv) используют разные механизмы моделирования. В приведённой ниже таблице приведены основные различия:

Функции	mgcv (GAM)	earth (MARS)	glmnet (эластичная сеть)
Тип модели	Сглаживаемая нелинейная (штрафные сплайны)	Кусочно-линейная (адаптивные сплайны)	Линейная (с регуляризацией)
Нелинейность	Автоматически с помощью сглаживаемых функций (тонкая пластина, кубическая и т. д.)	Автоматически с помощью функций-шарниров с узлами, определяемыми данными	Только через взаимодействия или разложение в базис
Отбор переменных	Опционально с помощью сжатия с двойным штрафом (<code>select=TRUE</code>)	Встроенно с помощью пошагового метода «вперед/назад» с обрезкой по GCV	Встроенно с помощью штрафа « L_1 » (Lasso)
Взаимодействия	Сглаживание тензорным произведением (<code>te()</code> , <code>ti()</code>)	Произведения на основе шарниров (интерпретируемые, до 3-й степени)	Парные векторные произведения (сложно интерпретировать)
Интерпретируемость	Сглаженные кривые частичных эффектов; очень интуитивно понятно	g-функции показывают кусочно-линейные частичные эффекты по каждой переменной	Коэффициенты представляют собой линейные веса; их легко объяснить по отдельности
Лучше всего подходит для	Сглаженные нелинейные зависимости, гибкое моделирование	Автоматическое обнаружение нелинейности и взаимодействий	Высокомерные данные, отбор переменных, небольшие выборки
Корректировки RCA	Сглаживание с помощью частичных эффектов	Кусочно-линейные корректировки с помощью g-функций	Линейные корректировки по каждой переменной
Контроль переобучения	Оценка параметров сглаживания методами REML/GCV	Обрезка на основе GCV	Лямбда, проверенная с помощью перекрёстной проверки

Метод MARS (Multivariate Adaptive Regression Splines) был предложен Джеромом Фридманом в 1991 году. Он позволяет строить кусочно-линейные модели путем адаптивного выбора функций-шарниров и положений их узлов на основе данных. Реализация этого метода в языке R представлена в пакете earth, разработанном Стивеном Милборроу.

Регрессия с эластичной сетью, реализованная в пакете «glmnet» Фридмана, Хасты и Тибширани (2010), сочетает в себе штрафные функции Lasso (L_1) и Ridge (L_2) для одновременного отбора переменных и сжатия коэффициентов.

1.1.4 earthUI–mgcvUI Конвейер

Эффективный рабочий процесс сочетает в себе автоматический поиск узлов в earthUI с сглаживанием оценки в mgcvUI. Когда earthUI экспортирует файл результатов .rds, а mgcvUI импортирует его:

1. Earth обнаруживает узлы — MARS находит основанные на данных положения шарниров (точки изменения) для каждой переменной.
2. mgcvUI использует их в качестве якорей — Узлы earth становятся основой для кубических регрессионных сплайнов (основа cr) в GAM, давая сглаживаемым членам преимущество благодаря положениям узлов, которые отражают реальную структуру данных.
3. mgcv уточняет подгонку — Оценка сплайна с штрафом проводит сглаживание кусочно-линейной подгонки earth, создавая гладкие кривые частичных эффектов, сохраняя при этом общую форму, обнаруженную earth.

Данный процесс объединяет эксплоративное моделирование (earth) с подтверждающим моделированием (GAM), сочетая в себе лучшие черты обоих подходов.

1.1.5 Рекомендуемый рабочий процесс

Для оценки недвижимости и аналогичных задач, требующих понятных и обоснованных моделей:

1. Начните с earthUI, чтобы выявить важные переменные, нелинейные зависимости и взаимодействия в ваших данных. Автоматический выбор базиса и g-функции earthUI позволяют получить отличные первоначальные выводы.
2. Уточните результаты с помощью mgcvUI, если вам нужны более плавные частичные эффекты. Импортируйте узлы из earthUI в mgcvUI для плавного перехода от кусочно-линейных моделей к моделям сглаживания.
3. Используйте glmnetUI, если вам требуется агрессивный отбор переменных (Lasso), когда количество предикторов приближается к числу наблюдений или превышает его, либо если вам нужна чисто линейная модель с регуляризацией для обоснованности результатов.

1.2 Что такое mgcvUI?

mgcvUI — это графический интерфейс пользователя для пакета R «mgcv». Он работает как локальное приложение Shiny — без входа в систему, без сервера и без учетных записей. Вы запускаете его из R, импортируете набор данных (CSV или Excel), настраиваете модель и в интерактивном режиме выполняете подгонку GAM.

Начиная с текущей версии, mgcvUI больше не является отдельно устанавливаемым пакетом: он поставляется в составе пакета earthUI в качестве модуля постобработки GAM, запускаемого командой `earthUI::launch_mgcv()`. Само приложение — а также всё, что описано в данной статье, — осталось без изменений; отличаются лишь установка и команда запуска. Три процедуры (earth, glmnet, mgcv) используют одно дерево проекта, одну базу данных настроек и одну копию вспомогательной инфраструктуры.

Приложение обеспечивает полный рабочий процесс: импорт данных, настройку переменных с возможностью сглаживания членов, подгонку модели с обработкой в фоновом

режиме, графики для диагностики, кривые сглаженных частичных эффектов, а также отчёты, доступные для скачивания в форматах Word, PDF или HTML.

Обобщённые аддитивные модели заменяют линейную зависимость βx на функцию сглаживания $f(x)$ для каждого предиктора. Это означает:

- Нет необходимости задавать функциональные формы — данные сами определяют форму каждой зависимости.
- Сглаживание частичных эффектов — вклад каждого предиктора представляет собой сглаженную кривую, а не прямую линию, что позволяет легко увидеть, как переменная влияет на зависимую переменную при различных значениях.
- Автоматический выбор степени сглаживания — `mgcv` определяет, насколько извилистой должна быть каждая кривая, с помощью REML или GCV, обеспечивая баланс между подгонкой и переобучением.

1.3 Три режима работы

У каждого проекта есть цель, которая выбирается при его создании в разделе 1 боковой панели («Проект») и впоследствии фиксируется. Цель активного проекта определяет, какие инструменты и элементы интерфейса отображаются; чтобы работать в другом режиме, создайте или откройте проект с соответствующей целью. Существуют три режима:

- Общие — Регрессия GAM для любого типа совокупности или набора данных. Это режим по умолчанию. Он обеспечивает полный рабочий процесс GAM без каких-либо дополнений, специфичных для конкретной области.
- Для оценки — Регрессия GAM, адаптированная для оценки недвижимости. Добавляет функции, специфичные для оценки отдельных объектов недвижимости, включая обработку объекта оценки, специальные обозначения столбцов и Метод остаточных ограничений (RCA).
- Анализ рыночной зоны — Регрессия GAM, адаптированная для исследований рыночных зон. Добавляет функции для анализа групп объектов недвижимости в определённом рынке с опциональной обработкой строк, относящихся к конкретному объекту.

Во всех трёх режимах основной механизм моделирования одинаков — вы всегда выполняете подгонку модели GAM с помощью функции «`mgcv::gam()`». Настройка «Purpose» определяет, какие дополнительные инструменты и элементы интерфейса будут доступны.

1.4 Недвижимость — специальные материалы

При выборе опции «Для оценки» или «Анализ рыночной зоны» программа `mgcvUI` активирует ряд функций, предназначенных для анализа рынка недвижимости:

- Обозначения специальных столбцов — Каждому предиктору может быть присвоена специальная роль, такая как «`contract_date`», «`dom`», «`concessions`», «`latitude`», «`longitude`», «`living_area`», «`lot_size`», «`actual_age`», «`effective_age`», «`area`», «`site_dimensions`» или «`display_only`». Эти обозначения определяют, как столбец будет обрабатываться во время подгонки модели и при выводе результатов.
- Столбец «Age» (Срок с момента продажи) — Если столбец обозначен как «`contract_date`» (Срок с момента продажи) и указана «дата оценки» (Дата вступления в силу), `mgcvUI` вычисляет столбец «`sale_age`» (Срок с момента продажи) (количество дней между датой продажи и датой оценки) и использует его в качестве предиктора.
- Вычисления RCA (только при оценке) — В режиме оценки, после подгонки модели `mgcvUI` может вычислить вывод по методу остаточных ограничений (RCA), который даёт корректировки по каждому аналогу, сводные данные по чистым/валовым кор-

ректировкам, а также скорректированную цену продажи недвижимости, подлежащей оценке.

Совет: Режим «Общее назначение» подходит для любых наборов данных — финансовых, научных, инженерных или связанных с недвижимостью. Режимы «Оценка» и «Рынок» просто предоставляют дополнительные удобные функции для специалистов в сфере недвижимости.

1.5 Начало работы

Чтобы воспользоваться mgcvUI:

- 1 . Установка earthUI в R: `install.packages("earthUI")` — процедура mgcv уже включена. (См. раздел «Установка earthUI и vProlog» в статье об earthUI в этом выпуске, где описаны версии для разработки версии и дополнительных компонентов.)
- 2 . Запустите приложение : выполните команду `earthUI::launch_mgcv()` в R. Приложение откроется в вашем веб-браузере на порту 7880.
- 3 . Создайте или откройте проект в разделе 1 боковой панели. Проект определяет цель (Общий, Для оценки или Анализ рыночной зоны) и место проведения работ (страна, штат, округ, город), а также определяет папки ввода и вывода. Проекты хранятся в общем корневом каталоге regProj и доступны для других приложений earthUI и glmnetUI. Открытие проекта с другой целью позволяет переключаться между режимами.
- 4 . Импортируйте данные из раздела 2, выбрав файл CSV или Excel из папки вводных данных активного проекта. Добавьте файлы в эту папку, затем нажмите «Обновить», чтобы отобразить их список.
- 5 . Импорт из earthUI (необязательно) — Раздел 3 позволяет импортировать файл результатов earthUI .rds для определения начальных значений в GAM с помощью узлов, обнаруженных earthUI. Пропустите этот шаг, если у вас нет результатов earthUI .
- 6 . Настройка переменных — выберите целевую переменную и предикторы, установите типы данных, назначьте специальные роли столбцам и пометьте переменные как факторные или линейные.
- 7 . Настройка параметров mgcv — семейство, метод, гамма, тип базиса, k, тип тензора, взаимодействия и другие параметры.
- 8 . Постройте модель — нажмите «Fit Mgcv GAM Model» и просмотрите результаты на главной панели.
- 9 . Экспорт — скачайте прогнозы в формате Excel, сгенерируйте и конвертируйте отчёт Quarto или (в режиме оценки) рассчитайте корректировки RCA и таблицу продаж.

Настройки автоматически сохраняются в базе данных SQLite и восстанавливаются при повторной загрузке того же файла ввода.

2 Требования к данным для ввода MLS

В рамках рабочих процессов по оценке недвижимости и анализу рынка вводные данные обычно получают из экспорта из системы Multiple Listing Service (MLS). В этой главе описывается ожидаемая структура файла и столбцы, которые может использовать mgcvUI.

2.1 Формат файла & Структура

mgcvUI поддерживает файлы в форматах CSV и Excel (.xlsx, .xls). При импорте названия столбцов автоматически преобразуются в формат snake_case — например, «Living SqFt» превращается в «living_sqft», «Contract Date» — в «contract_date», а «Цена продажи» — в «sale_price». Такая нормализация обеспечивает единообразные ссылки на столбцы на протяжении всего рабочего процесса. Разделитель CSV и знак десятичной запятой, используемые при импорте, определяются настройками локали (см. главу 3, «& ные региональные настройки»).

Ваш файл с данными должен представлять собой плоскую таблицу, в которой каждая строка соответствует одному объекту недвижимости, а каждый столбец — одной характеристике. Первая строка файла должна содержать заголовки столбцов.

2.2 Обязательные столбцы для оценки Режим

Хотя mgcvUI работает с любым набором столбцов, для полноценного рабочего процесса оценки рекомендуется использовать следующие столбцы:

Столбец	Специальный тип	Назначение
Цена продажи	(целевая)	Переменная отклика для модели
Дата заключения договора	contract_date	Используется для расчета sale_age (количество дней с даты оценки)
Дата размещения объявления	listing_date	Используется вместе с датой заключения договора для расчета DOM, если столбец DOM отсутствует
Дни на рынке	dom	Дни на рынке; отображается в экспортируемых данных
Уступки	concessions	Уступки при продаже; чистая цена продажи (Net SP) = цена продажи – Уступки
Жилая площадь (SF)	living_area	Позволяет рассчитывать остатки на квадратный фут (residual_sf, cqa_sf)
Размер участка	lot_size	Столбец «Размер участка»
Размеры участка	site_dimensions	Группируется по размеру участка
Широта	latitude	Округлено до 3 знаков после запятой
Долгота	longitude	Округлено до 3 знаков после запятой
Идентификатор района	area	Идентификатор рыночного района / микрорайона
Фактический возраст	actual_age	Возраст объекта недвижимости
Эффективный возраст	effective_age	Эффективный возраст объекта недвижимости
Адрес	display_only	Отображается в экспортируемых данных; исключается из модели

Названия столбцов в таблице могут быть на иностранном языке — «специальные» названия должны быть на английском языке, чтобы программа R могла обрабатывать их особым образом. В противном случае указанные названия столбцов будут отображаться в регрессионных моделях, графиках и, при проведении оценок, в выводах.

Не все столбцы являются обязательными. mgcvUI адаптируется — если какой-либо столбец отсутствует, соответствующий объект просто не учитывается. Однако для моделей ценообразования на рынке недвижимости настоятельно рекомендуется использовать определенные столбцы:

- 1 . Срок продажи — количество дней между датой заключения договора купли-продажи и датой оценки. Если используется многолетняя история продаж, `sale_age` часто играет ключевую роль.
- 2 . Жилая площадь — также известна под такими названиями, как «Living Sqft», «GLA» (общая жилая площадь).
- 3 . Общее количество ванных комнат — общее количество полных, четвертных, половинных и 3/4 ванных комнат.
- 4 . Гаражные места или площадь гаража — количество гаражных мест или площадь в квадратных футах.
- 5 . Площадь участка — площадь недвижимости.
- 6 . Долгота, широта и , при наличии , идентификатор района для определения географических различий в ценах.

2.3 Особые правила именования столбец

mgcvUI идентифицирует столбцы по их специальным обозначениям типов, а не по названиям столбцов. При экспорте в формат MLS вы можете называть столбцы как угодно — главное, чтобы в таблице «Настройка переменных» (глава 6) был назначен правильный специальный тип.

2.4 Качество данных& Полнота данных

- Отсутствующие значения (NA) : Строки, содержащие значения NA в любом столбце предиктора или целевой переменной, автоматически удаляются перед подгонкой модели. mgcvUI формирует подмножество данных, включающее только столбцы, используемые в модели, поэтому значения NA в неиспользуемых столбцах не приводят к удалению строк.
- Числовые столбцы : Цена продажи, жилая площадь, площадь участка и подобные поля должны быть числовыми.
- Факторные столбцы : Текстовые, TRUE/FALSE и R-факторные столбцы автоматически рассматриваются как категориальные — они могут входить в модель только в качестве факторов. Числовой столбец (например, числовое значение площади или код стиля) становится фактором только при установке флажка «Фактор» ; mgcvUI никогда не выводит категориальность из диапазона значений числового столбца, поэтому дискретные числовые предикторы, такие как `bath_count` , остаются непрерывными (сглаженными). Роль «Special» не зависит от решения о факторах .

Совет: После импорта проверьте столбец «NA» в таблице предикторов. Столбцы с большим количеством пропущенных значений могут привести к исключению строк. Столбцы, в которых количество пропущенных значений превышает $\geq 50\%$, помечаются красным цветом и автоматически исключаются.

2.5 Размещение объекта оценки в строке

В режиме оценки в первой строке должна быть указана недвижимость, которая подлежит оценке. Все остальные строки представляют собой аналоги сделок. Строка с объектом оценки исключается из процесса подгонки модели. После подгонки модели она по-прежнему генерирует прогнозы для этой строки.

В режиме «Анализ рыночной зоны» размещение объекта оценки в строке 1 является необязательным.

В режиме «Общий» специальной обработки строк не предусмотрено — все строки обрабатываются одинаково.

3 Режим общего назначения

3.1 Обзор

Режим «General Purpose» (Общего назначения) является режимом по умолчанию при запуске mgcvUI. Он обеспечивает полный рабочий процесс GAM для любых наборов данных — не только для данных по недвижимости. Вы можете использовать mgcvUI для работы с научными данными, финансового анализа, инженерных исследований или решения любых задач регрессионного анализа, в которых предполагается сглаживание нелинейных зависимостей.

В общем режиме интерфейс не отображает функции, специфичные для рынка недвижимости (специальные столбцы, срок продажи, RCA). Боковая панель упрощена, чтобы сосредоточить внимание на отборе переменных, настройке параметров, подгонке модели и экспорте.

3.2 Рабочий процесс с использованием боковой панели

Боковая панель состоит из пронумерованных сворачиваемых разделов, которые помогают вам пройти весь путь от выбора проекта до экспорта отчёта. Большинство разделов отображаются только при открытии активного проекта, а некоторые из них зависят от цели проекта. Папка для вывода данных определяется на основе активного проекта — отдельного поля для указания папки вывода данных нет.

1. Проект — создание или открытие проекта. В проекте задаются цель и место проведения работ (страна, штат, округ, город), а также папка ввода (`<os>_in/`) и папка вывода. Проекты хранятся в общем корневом каталоге regProj (по умолчанию `~/regProj`, можно переопределить с помощью переменной среды REGPROJ_ROOT или раскрывающегося списка «Настройки») и совместно используются с родственными приложениями earthUI и glmnetUI, поэтому модели из всех трёх приложений хранятся в одном месте. Нажмите «Закрыть проект», чтобы переключиться.

2. Импорт данных — Выберите файл CSV или Excel из папки ввода текущего проекта. Добавьте файлы в папку с помощью Finder/Проводника, затем нажмите «Обновить», чтобы отобразить их список. Для файлов Excel с несколькими листами появится окно выбора листа. Названия столбцов автоматически преобразуются в формат «snake _».

3. Импорт из earthUI (необязательно) — Импортируйте файл результатов earthUI .rds, чтобы установить начальные значения для GAM с помощью координат узлов, обнаруженных earthUI (а также для отбора переменных). Этот шаг является необязательным — пропустите его, если у вас нет результатов earthUI для импорта.

4. Настройка переменных — выбор целевой переменной, преобразование отклика (без преобразования, логарифмическое, логарифмическое с основанием 10), таблица предикторов с флажками «Включить», «Фактор» и «Линейный». Столбец «Специальное» отображается только в режимах «Оценка» и «Рынок». Подробную информацию см. в главе 6.

5. Параметры вызова MgcV — Все настройки модели: предустановки параметров, семейство, метод, гамма, перекрёстная проверка, выбор, тип базиса, k, тип тензора, матрица взаимодействия и расширенные параметры. Полное справочное руководство по параметрам см. в главе 7.

6. Подгонка модели MgcV GAM — кнопка, запускающая модель. Результаты отображаются на вкладках справа.

7. Скачать результаты — экспортирует прогнозы, остатки, балы CQA и вклады по каждой переменной в виде файла Excel в папку вывода проекта. Доступно во всех режимах после подгонки модели.

8. Расчет корректировок RCA & Скачать — Вычисляет корректировки RCA для каждой переменной по каждому аналогу и интерполирует остаток по объекту оценки с помощью CQA. Доступно только в режимах «Оценка» и «Рынок», после подгонки модели.

9. Создать таблицу продаж & Скачать — Создает таблицу сравнения продаж, предлагая аналогичные объекты с учетом процента валовой корректировки. Отображается только в режиме оценки.

10. Создать отчет Quarto — создает автономный набор отчета Quarto (.qmd) — со всеми графиками и таблицами — в папке вывода проекта. Функция становится доступной после подгонки модели.

11. Конвертировать отчет Quarto — преобразует сгенерированный набор .qmd в формат HTML, Word или PDF. Для создания PDF-файла требуется установка LaTeX; если она не обнаружена, опция PDF будет скрыта. См. главу 13.

В других режимах шаги, предназначенные для конкретных задач, опущены (а оставшиеся шаги перенумерованы): в режиме «Рынок» отсутствует шаг «Таблица продаж», а в режиме «Общий» отсутствуют оба шага — «RCA» и «Таблица продаж».

3.3 Импорт из earthUI (Необязательно)

Раздел 3 боковой панели — «Импорт из earthUI» — позволяет импортировать файл результатов earthUI .rds. Это активирует конвейер earth-mgcv: позиции узлов, определяемые данными earth, становятся опорными точками для членов процесса сглаживания GAM.

При импорте результата из earthUI mgcvUI:

- Отображает сводку, содержащую целевую переменную, коэффициент R-квадрат, количество членов и положения узлов для каждой переменной.
- Автоматически выбирает предустановку параметров «Earth Pipeline» (кубическая регрессионная сплайн-основа, гамма = 1,4).
- Устанавливает целевую переменную в соответствии с целевой переменной «earth».
- Предварительно устанавливает флажок «Include» для предикторных переменных earth.
- Предварительно устанавливает флажки «Factor» для категориальных переменных earth и «Linear» для линейных предикторов earth.
- Предварительно заполняет взаимодействия, обнаруженные earth (выделены желтым цветом в матрице взаимодействий, заблокированы для ручного изменения).
- Предоставляет кнопку «Экспорт узлов в CSV» для загрузки узлов earth в виде файла CSV.

Затем в таблице «Конфигурация переменных» отображается примечание, объясняющее, что отбор переменных («Включить», «Линейная», «Фактор») и узлы сплайна имеют начальное значение из модели Земли. В отличие от glmnetUI, где подгонка glmnet привязана к базису earth, mgcvUI проводит подгонку GAM на основе вашего текущего выбора — earth служит лишь отправной точкой, и все параметры остаются доступными для редактирования.

Кнопка «Очистить» удаляет импортированные данные о заземлении и возвращает устройство в автономный режим.

3.4 Вкладки главной панели

После импорта данных на главной панели отображаются следующие вкладки (вкладки, зависящие от модели, появляются после подгонки):

Вкладка «	» Содержание
Данные	Интерактивная таблица DataTable с импортированными данными. В режимах «Оценка» и «Рынок» разбита на таблицы «Объект оценки» и «Аналоги».
Уравнение	Полная формула GAM, отображенная с помощью MathJax, а также таблица определений функций сглаживания.
Резюме	Ключевые показатели (R^2 , $CV R^2$, Dev. Explained, AIC, n, Кривая сглаживания (сглаженная) и таблицы сглаженных/параметрических термов.
Важность переменных	Горизонтальная гистограмма F-статистики (сглаживание) и значений $ t $ (параметрическая), а также таблица с ранжированием.
Вклад	Интерактивные кривые частичных эффектов с доверительными лентами и маркерами EarthKnot, построенные с помощью Plotly.
Корреляция	Тепловая карта корреляций числовых предикторов (доступна до подгонки модели).
Диагностика	Остатки по сравнению с моделью, Q-Q-диаграмма, гистограмма, зависимость ответа от модели (через gratia). Разброс фактических и прогнозируемых значений.
Корректировки RCA	Гистограммы анализа RCA (режим оценки/рынка, после вычислений).
ANOVA	Таблица ANOVA с объединенными параметрическими и членами сглаживания.
Вывод Mgcsv	Исходный текст: временные параметры, формула, сводка модели, семейство, конкурвитность.
Проверка знаков	Проверка согласованности направлений Earth и GAM (при импорте данных из Earth).
Конкурвитность	Матрицы общей и парной конкурвитности с цветовой кодировкой.
Проверка базиса	gam.check() тесты адекватности базисных измерений.

3.5 Сохранение настроек

mgcsvUI автоматически сохраняет вашу конфигурацию в базе данных SQLite, с ключом, соответствующим имени файла ввода. При повторной загрузке того же файла все настройки восстанавливаются: выбор целевой переменной, флажки предикторов, типы данных, параметры mgcsv и матрица взаимодействия. Цель является свойством активного проекта, а не глобальной настройкой, и последний использованный файл ввода запоминается для каждого проекта и автоматически загружается при повторном открытии проекта.

3.6 Темный режим

Нажмите на значок в виде луны/солнца в правом верхнем углу, чтобы переключаться между темами «Nord Light» и «Nord Dark». Настройки темы сохраняются в localStorage и сохраняются между сессиями. Все элементы пользовательского интерфейса адаптируются к выбранной теме.

3.7 Язык и регион & Региональные настройки

mgcsvUI поддерживает международные стандарты формата чисел и CSV-файлов благодаря системе локалей, основанной на странах. В раскрывающемся списке «Настройки» в заголовке

окна доступны переключатели «Страна» и «Формат бумаги» для более 30 поддерживаемых стран. Каждая предустановка настраивает:

- Разделитель в файлах CSV — запятая (,) для США/Великобритании/Японии или точка с запятой (;) для большинства стран Европы.
- Десятичный знак — точка (.) или запятая (,).
- Разделитель тысяч — запятая (США/Великобритания/Япония), точка (Германия/Италия/Испания), пробел (Финляндия/Франция/Польша/страны Балтии) или апостроф (Швейцария).
- Формат бумаги — Letter (США/Канада/Мексика) или A4 (везде else).

Нажмите «Сохранить», чтобы глобально сохранить настройки локали. В разделе «Импорт данных» также имеется переключатель «Локализация» для отдельных файлов, предназначенный для разовых импортов, в которых используются иные соглашения. В том же раскрывающемся списке «Настройки» находится поле «Корневая папка regProj», которое позволяет изменить расположение общего дерева проектов, используемого mgcvUI, earthUI и glmnetUI.

Совет: Формат чисел на осях графика и в метках наклона автоматически адаптируется к вашей языковой локали. В немецкой локали для разделителя тысяч используется точка (200.000), в финской — пробел (200 000), а в швейцарской — апостроф (200'000). Символы валют не отображаются — mgcvUI не зависит от валюты.

4 Режим оценки

Если цель активного проекта указана как «Для оценки», mgcvUI настраивается на оценку одного объекта недвижимости. Все функции, описанные в главе 3, остаются доступными; в данной главе рассматриваются только дополнительные возможности, специфичные для оценки.

4.1 Обработка строк объекта оценки

В режиме оценки первая строка вашего набора данных представляет собой объект оценки, а все остальные строки — аналог сделки. Ваш вводный файл должен быть организован соответствующим образом (см. главу 2). Цена продажи объекта может быть оставлена пустой или установлена на любое значение — mgcvUI автоматически рассматривает её как NA во время подгонки модели.

После подгонки модель генерирует прогнозы для соответствующей строки объекта оценки, что служит основой для рабочего процесса корректировки RCA.

4.2 Дата оценки & Возраст, с которого разрешена продажа

В режимах «Оценка» и «Рынок» в разделе «Настройка переменных» появляется поле «Дата оценки» (по умолчанию установлена сегодняшняя дата). Если вы выберете для столбца значение «contract_date» в раскрывающемся списке «Special», mgcvUI рассчитает столбец «sale_age» — количество целых дней между датой заключения договора по каждой сделке и датой оценки. Этот столбец добавляется в качестве предиктора.

При изменении даты оценки показатель «sale_age» автоматически пересчитывается.

4.3 Специальный столбец Обозначения

В режимах оценки и рынка для каждого предиктора в таблице «Конфигурация переменных» появляется специальный раскрывающийся список. Полную справку по специальным типам и их эффектам см. в главе 6.

4.4 Обзор корректировок RCA

Кнопка «Рассчитать корректировки RCA» (&) (отображается в режимах оценки и рыночного анализа после подгонки аналогов) позволяет рассчитать рыночные корректировки для каждого аналога по отношению к объекту оценки. Полный рабочий процесс RCA описан в главе 11.

Совет: Балл CQA, присваиваемый объекту оценки, определяет, какая доля остатков распределения относится именно к этому объекту оценки. Балл 5,00 означает, что объект оценки находится на медиане по сравнению с аналогами.

5 Режим анализа рыночной зоны

Если целью активного проекта является «Анализ рыночной зоны», mgcvUI предоставляет те же функции, характерные для сферы недвижимости, что и в режиме оценки (специальные столбцы, срок с момента продажи), но ориентирован на анализ группы объектов недвижимости, а не на оценку отдельного объекта оценки.

5.1 Отличия от режима оценки

- Обработка строк с объектами оценки осуществляется гибко — в режиме «Рынок» строка 1 не рассматривается автоматически как объект оценки.
- Отсутствует таблица продаж — этап «Сгенерировать таблицу продаж& Скачать» недоступен.

5.2 Когда следует использовать режим «Рынок»

Режим «Анализ рыночной зоны» подходит в следующих случаях:

- Построение модели GAM для района или рыночной зоны с целью понимания нелинейных факторов, определяющих стоимость
- Анализ того, как такие переменные, как площадь, возраст, площадь участка и расположение, влияют на цены продажи в группе объектов недвижимости
- Подготовка материалов для анализа рыночной конъюнктуры или района
- Работа с набором данных, в котором требуются специальные характеристики столбцов (срок продажи), но не нужен рабочий процесс с таблицей продаж

Совет: Режим «Рынок» также полезен для общей регрессионного анализа на рынке недвижимости, когда требуются специальные характеристики столбцов, но не нужен рабочий процесс оценки отдельного объекта недвижимости.

6 Отбор переменных

В разделе 4 боковой панели — «Настройка переменных» — вы выбираете, какие столбцы будут включены в модель и как с ними будет проводиться работа.

6.1 Целевая переменная

В раскрывающемся списке «Целевая (реакционная) переменная» в верхней части раздела 4 перечислены все числовые столбцы вашего набора данных. Выберите один столбец в качестве реакционной переменной (например, «Цена продажи»). Целевой столбец автоматически исключается из списка предикторов.

6.2 Преобразование отклика

Под селектором цели в раскрывающемся списке «Преобразование ответа» доступны три варианта:

- Нет (сырые данные) — Использовать ответную переменную в исходном виде. Подходит, когда зависимость между предикторами и ответной переменной является приблизительно аддитивной в исходной шкале.
- Логарифм (естественный) — Применить преобразование по естественному логарифму перед подгонкой. При обратном преобразовании это делает корректировки пропорциональными (%), а не абсолютными (\$). Особенно полезно для цен продажи, где корректировка \$ 10 000 означает совершенно разные вещи при разных ценовых уровнях.
- Log10 — Применить логарифмическое преобразование с основанием 10.

Совет: Логарифмическое преобразование делает корректировки по времени пропорциональными (%), а не абсолютными (\$). Для моделей рынка недвижимости, охватывающих широкий диапазон цен, логарифмическое преобразование часто позволяет получить остатки с более предсказуемым поведением.

При выборе логарифмического преобразования значения ≤ 0 в ответе автоматически отфильтровываются. Все выводные данные (прогнозы, вклады, остатки) автоматически преобразуются обратно в исходную шкалу.

6.3 Таблица предикторов

Ниже селектора целей в прокручиваемой таблице перечислены все остальные столбцы. Для столбцов с флажками в целях компактности используются повернутые вертикальные заголовки. Порядок столбцов следующий:

Столбец	Описание
Variable	Название столбца (моноширинный шрифт, усечено с использованием многоточия). Ячейка с рамкой.
Type	Раскрывающийся список типов данных: числовой, целочисленный, символьный, факторный, логический, Дата, POSIXct. Раскрывающийся список с рамкой.
Include	Флажок — включить этот столбец в модель в качестве предиктора. Вертикальный заголовок.
Factor	Флажок — рассматривать как категориальный фактор (создает отдельный коэффициент для каждого уровня). Вертикальный заголовок.
Linear	Флажок — принудительно использовать линейную зависимость (без сглаживания, только βx). Вертикальный заголовок.
Special	Раскрывающийся список (только для оценки/рынка) — см. Справочник по специальным типам столбцов ниже. Раскрывающийся список с рамкой.
NAs	Количество и процент отсутствующих значений, обозначенных цветом: красный ($\geq 50\%$), оранжевый ($\geq 20\%$), серый (> 0). Ячейка с рамкой.

«Фактор» vs. «Сглаживание» vs. «Линейный»: по умолчанию для включенных числовых переменных создается член сглаживания $f(x)$. Установка флажка «Фактор» приводит к созданию категориального члена (по одному коэффициенту на каждый уровень). Установка флажка «Линейный» создаёт простой линейный член βx вместо процесса сглаживания. Переменные, отмеченные как «Фактор» и «Линейный», обрабатываются как «Фактор».

6.4 Определение типов данных & Переопределения

mgcvUI автоматически определяет типы данных при импорте. Распознаются столбцы с числовыми, целочисленными, логическими, факторными и датными значениями. Вы можете переопределить любое определение, изменив значение в раскрывающемся списке «Тип». Изменение типов влияет на то, как столбец обрабатывается в модели.

6.5 Специальные типы столбцов См.

В режимах «Оценка» и «Рынок» в раскрывающемся списке «Специальное» доступны следующие параметры:

Взвешивание:

- weight — Столбец с весами наблюдений (допускается только одна; строки с весом = 0 исключаются из подгонки)

Дата & . Типы времени:

- contract_date — Запускает автоматический расчет показателя «sale_age» с даты оценки
- sale_age — Обозначает заранее рассчитанный столбец «Срок продажи» (количество дней с даты оценки), используемый без изменений
- listing_date — Используется в качестве резервного варианта для расчета показателя «Days on Market»
- dom — Обозначает столбец «Days on Market»

Тип записи:

- sale_type — Определяет столбец типа записи, который позволяет отличить данный объект оценки от аналогов сделок

Типы денежных единиц:

- `concessions` — Определяет условия скидок при продаже (кредиты продавца, поощрения для покупателя и т. д.)

Размер& . Типы мест:

- `latitude` — Значения автоматически округляются до 3 знаков после запятой
- `longitude` — Применяется тот же метод округления, что и для широты
- `area` — Идентификатор рыночной зоны или района
- `living_area` — Включает расчет остаточной стоимости на квадратный фут (`residual_sf` и `cqa_sf`)
- `lot_size` — Столбец «Площадь участка»
- `site_dimensions` — Группируется с площадью участка

Возрастные группы:

- `actual_age` — Столбец «Возраст объекта недвижимости»
- `effective_age` — Фактический возраст объекта недвижимости

Типы дисплеев:

- `display_only` — Столбец включается в экспортные файлы Excel, но полностью исключается из процесса подгонки модели. Используйте его для полей адреса, номеров MLS, идентификаторов земельных участков или других справочных данных.

Совет: Столбцы, помеченные как «`display_only`» (не используемые в модели), остаются в наборе данных и отображаются при экспорте в Excel, но полностью исключаются из таблицы предикторов. Используйте эту функцию для столбцов с идентификаторами, адресами или другими справочными полями.

7 Выбор параметров

Раздел 5 боковой панели — «Параметры вызова MgcV» — предоставляет доступ ко всем параметрам настройки GAM. У каждого параметра есть синий значок справки (?) с всплывающей подсказкой.

7.1 Предустановки параметров

В раскрывающемся списке в верхней части экрана доступны две предустановки, которые задают оптимальные настройки по умолчанию для типичных рабочих процессов:

Предустановки	Базис	Настройки
Автономный режим (поиск)	tp (тонкая пластина)	k = auto (10), gamma = 1.2, select = TRUE
Earth Pipeline (уточнение)	cr (кубическая регрессия)	k = auto (10), gamma = 1,4, select = TRUE

Пресет «Earth Pipeline» выбирается автоматически при импорте узлов earthUI. Для интеграции узлов earthUI требуется базис кубического регрессионного сплайна (cr), поскольку он позволяет указывать точные координаты узлов.

7.2 Семья

Выберите семейство распределений для вашей зависимой переменной:

Family	Use Case
gaussian	Непрерывные значения (например, цена продажи). Наиболее распространенный случай.
Gamma	Положительные непрерывные значения с дисперсией, пропорциональной среднему значению.
poisson	Данные, выраженные в количестве (например, количество продаж).
binomial	Бинарные исходы.
inverse.gaussian	Данные с сильным правосторонним асимметричным распределением.

7.3 Метод

Метод оценки параметра сглаживания:

Метод	Описание
REML	Ограниченное максимальное правдоподобие (по умолчанию, рекомендуется). Наиболее устойчиво к переобучению.
GCV.Cp	Обобщённая перекрёстная проверка. Имеет тенденцию выбирать несколько более сложные модели.

Метод	Описание
ML	Метод максимального правдоподобия. Аналогичен REML, но может приводить к недостаточному сглаживанию.
P-REML	Вариант REML по Пирсону.
P-ML	Вариант ML по Пирсону.
GACV.Cp	Обобщённая приближённая перекрёстная проверка.

7.4 Gamma (коэффициент сглаживания)

Параметр «gamma» умножает количество эффективных степеней свободы в критерии сглаживания, способствуя получению более гладких (менее «волнистых») подгонок. Значение по умолчанию — 1,2.

Значение	Эффект
1,0	Стандартное сглаживание
1,2–1,4	Немного более плавное, чем по умолчанию (рекомендуется)
2,0+	Значительно более плавное — подходит для данных с шумом или небольших выборок

Более высокие значения параметра gamma предотвращают переобучение за счет более сурового штрафа за сложность модели. В пресете «Earth Pipeline» используется значение 1,4 для дополнительного сглаживания при доработке узлов модели Земли.

7.5 Перекрёстная проверка

Если этот флажок установлен (по умолчанию), mgcvUI после подгонки модели выполняет 10-кратную перекрёстную проверку для вычисления коэффициента R-квадрат по результатам перекрёстной проверки. Это позволяет получить объективную оценку прогностической способности модели на вневывборочной выборке.

7.6 Выбрать (Сжатие)

Если этот флажок установлен (по умолчанию), к каждому члену, подвергающемуся сглаживанию, добавляется дополнительный штраф, который может привести к его сведению к нулю. Это позволяет осуществлять автоматический отбор переменных — незначимые члены, подвергающиеся сглаживанию, фактически удаляются из модели. Реализовано с помощью функции ‘select = TRUE’ в файле ‘gam()’.

7.7 Базовые данные по умолчанию

Тип базисных функций сплайна для сглаживаемых слагаемых:

Базис	Полное имя	Описание
tp	Сплайн регрессии тонкой пластины	По умолчанию. Изотропный (размещение узлов не требуется). Хороший универсальный выбор.
cr	Кубический регрессионный сплайн	Позволяет явно задавать расположение узлов. Требуется для интегрирования по узелкам на поверхности Земли.

Базис	Полное имя	Описание
ps	P-сплайн	Равномерно распределённая В-сплайн-основа с штрафом за разность. Вычислительно эффективна.
bs	B-сплайн	Гибкая В-сплайн-основа.

7.8 Значение k по умолчанию (базисная размерность)

Параметр «basis dimension» k определяет максимальную «волнистость» каждого сглаживания. Значение 0 (по умолчанию) означает «автоматический режим» — mgcvUI использует $k = 10$ или количество «земных узлов», в зависимости от того, что более уместно.

- Небольшое значение k (3–5) : Очень плавные кривые, практически линейные.
- Среднее значение k (8–15) : Умеренная гибкость (как правило, достаточная).
- Большое значение k (20+) : Очень высокая гибкость — существует риск переобучения без сильной регуляризации.

Совет: Показатель «Эффективные степени свободы» (EDF) на вкладке «Сводка» (Summary) показывает, сколько степеней свободы фактически использует каждая функция сглаживания. Если значение EDF близко к $k - 1$, базис, возможно, слишком мал — увеличьте k . Если значение EDF значительно меньше k , функция сглаживания подвергается сильному штрафу.

7.9 Тип тензора

Для взаимодействий между непрерывными переменными доступны два типа тензорного произведения: а именно:

Тип	Функция	Описание
ti	ti(x1, x2)	Тензорное взаимодействие — моделирует только эффект взаимодействия, при этом основные эффекты обрабатываются отдельно с помощью одномерных $s()$ слагаемых. Предпочтительно с точки зрения интерпретируемости.
te	te(x1, x2)	Тензорное произведение — моделирует весь совокупный эффект, включая основные эффекты. Более сложно разложить для корректировок RCA.

7.10 Допустимые взаимодействия

В сворачиваемом разделе «Допустимые взаимодействия» отображается верхнетреугольная матрица флажков для всех включенных сглаживаний (нелинейных, нефакторных) предикторов. Каждый флажок включает тензорное произведение сглаживания между соответствующей парой переменных.

- «Разрешить всё» / «Снять всё» — Кнопки массового переключения.
- Щелкните по имени переменной — Включить или отключить все взаимодействия для этой переменной.

- Щелкните правой кнопкой мыши по имени переменной — «Исключить из основного эффекта» — помечает переменную для использования только в взаимодействиях (без одномерного сглаживания).

При импорте узлов earthUI взаимодействия, обнаруженные earth, предварительно проверяются и блокируются (выделяются желтым фоном). Если в используемой модели earth установлен параметр `degree = 1` (без взаимодействий), отображается информационное сообщение.

Warning

Включение тензорных взаимодействий существенно повышает сложность модели. Используйте взаимодействия только в том случае, если у вас есть предметные знания или данные наземных наблюдений, подтверждающие их обоснованность. Убедитесь, что полученные корректировки являются обоснованными и объяснимыми.

7.11 Взаимодействия «фактор сглаживание»

При включении в модель как факторных, так и переменных сглаживания появляется отдельная матрица взаимодействий «Фактор × Переменная сглаживания». Каждый установленный флажок создает член $s(x, by = factor)$ — отдельную кривую сглаживания для каждого уровня фактора.

7.12 Расширенные параметры

При раскрытии раздела «Дополнительно» появляются дополнительные настройки:

- Оптимизатор — Алгоритм сглаживания оценки параметров: `outer/newton` (по умолчанию), `outer/bfgs` или `efs` (расширенный метод Феллнера-Шалла).
- Масштаб — Параметр масштаба. 0 (по умолчанию) = оценка по данным. Установите известное значение для моделей с фиксированным масштабом.
- Метод дискретных ковариат — Быстрая дискретизированная подгонка для больших наборов данных (`discrete = TRUE`).
- Потоки — Количество параллельных потоков (1–32). Действует только при значении `discrete = TRUE`.

8 Подгонка модели

8.1 Кнопка «Fit»

В разделе 6 боковой панели находится кнопка «Fit Mgcsv GAM Model ». При нажатии на нее модель запускается с текущими настройками.

8.2 Что происходит во время подгонки

При нажатии кнопки «Подгонка» в mgcvUI:

- 1 . Проверяет конфигурацию — проверяет наличие как минимум одного предиктора и не менее 10 полных строк. Выдает предупреждение, если менее 50% строк являются полными (с указанием столбцов, в которых содержится наибольшее количество значений NA).
- 2 . Подготавливает данные — формирует подмножества, содержащие только столбцы, значимые для модели, удаляет строки с значениями «NA», применяет веса (если они указаны), исключает строку объекта оценки (в режиме оценки), применяет преобразования отклика.
- 3 . Создает формулу GAM — формирует сглаживающие члены $s(x, bs="cr k=10)$ для числовых предикторов, факторные члены для категориальных переменных, линейные члены для переменных, помеченных как «Линейные», и тензорные произведения для выбранных взаимодействий. Узлы Earth встраиваются в качестве явных позиций узлов в базисных членах типа « cr ».
- 4 . Обработывает крайние случаи — удаляет константные переменные, преобразует бинарные предикторы в факторы, ограничивает количество « k » до числа уникальных значений минус 1, дополняет «earth knots» квантилями данных, когда требуется больше узлов.
- 5 . Подгонка GAM — вызывает функцию « mgcv::gam() » с построенной формулой, семейством, методом, гамма-параметром, параметром «select» и другими параметрами.
- 6 . Запускает перекрёстную проверку (если она включена) — 10-кратная перекрёстная проверка с вычислением R-квадрата вне выборки.
- 7 . Обновляет все вкладки с результатами — «Сводка», «Уравнение», графики вклада, «Диагностика» и все остальные вкладки заполняются данными.

8.3 Общая информация Подгонка& Ход работы

Если доступен пакет callr, подгонка модели выполняется в фоновом процессе, благодаря чему приложение продолжает реагировать на запросы пользователя. На экране отображается накладка с индикатором хода выполнения в стиле терминала с темной темой:

- Таймер, отображающий прошедшее время в секундах/минутах.
- Журнал трассировки с цветовой кодировкой : синий — для статуса, желтый — для хода сглаживания блока CV, красный — для ошибок, зеленый — для результатов.
- Кнопка «Прервать » для остановки процесса подгонки в случае необходимости.
- Кнопка «Заккрыть », появляющаяся после завершения.

Если функция callr недоступна, подгонка выполняется синхронно (приложение зависает до завершения подгонки).

После успешной подгонки на кнопке «Fit» появляется белая галочка, а в строке состояния отображается следующая информация: «R-sq = X, CV R-sq = X, Dev = X%, AIC = X, n = X».

9 Вкладки «Результаты»

9.1 Данные

Отображает импортированные данные в виде интерактивной таблицы DataTable с горизонтальной прокруткой и 15 строками на страницу. В режимах «Оценка» и «Рынок» вкладка разделена на таблицу «Объект оценки» (строка 1) и таблицу «Аналоги сделок» (остальные строки); в режиме «Общий» отображается единая таблица.

Каждая ячейка отображается в одной строке и обрезается до ширины столбца, поэтому широкие поля свободного ввода (например, примечания к объектам недвижимости) не растягивают строку вниз по странице. Один щелчок выделяет строку; двойной щелчок по любой ячейке открывает всплывающее окно с её полным, необрезанным содержимым.

9.2 Уравнение

Отображает подогнанную модель GAM в двух частях:

Образец уравнения (отображение с помощью MathJax):

- В левой части показана зависимость с соответствующим преобразованием ($\log$, $\log_{10}$) и функцией связи.
- В правой части показаны значение пересечения, члены сглаживаемые при сглаживании в виде $f_i(\text{variable})$, факторные члены, сгруппированные по переменным с указанием числа уровней, а также линейные члены со значениями коэффициентов.
- Семейство и метод отображаются над уравнением.

Таблица определений функций сглаживания:

- Строка формулы члена (например, $s(\text{living_sqft}, \text{bs} = "cr k=8")$)
- Имя переменной
- Тип (s , te , ti или линейный)
- Тип базисной функции
- Значение k
- Координаты узелков на поверхности Земли (если импортированы)

9.3 Резюме

В верхней части экрана отображаются шесть метрических карточек:

- R-квадрат — доля объясненного отклонения
- CV R-квадрат — R-квадрат перекрёстной проверки (если включена перекрёстная проверка)
- Dev. Explained% — процент объяснённого отклонения
- AIC — информационный критерий Акайке
- n — количество использованных наблюдений
- Smoother — количество сглаживающих слагаемых

Под карточками:

- Таблица сглаженных членов — Член, EDF (эффективные степени свободы), Ref.df, F-статистика, p-значение
- Таблица параметрических членов — Член, оценка, стандартная ошибка, t-значение, p-значение

9.4 Важность переменных

Горизонтальная гистограмма, отражающая относительную значимость каждого члена модели:

- Сглаживание членов : ранжированы по F-статистике (синие столбцы)
- Параметрические члены : ранжированы по значению $|t|$ (зеленые столбцы)

Под диаграммой в таблице DataTable приведены столбцы «Срок», «Тип», «EDF», «Статистика» и «p-значение», отсортированные по значениям статистического показателя в порядке убывания.

9.5 Графики вкладов

Каждый член модели, влияющий на прогнозируемое значение, имеет интерактивный график в Plotly, отображаемый в карточке с рамкой. На вкладке представлены графики для всех типов членов:

9.5.1 Univariate Smooth Terms — $s(x)$

- Синяя линия с лентой доверия 95%
- Серые точки, отображающие частичные остатки (для моделей без логарифмического преобразования)
- Красные пунктирные вертикальные линии в точках расположения «земных узлов» (если «земные узлы» импортированы)
- Всплывающая подсказка при наведении курсора : значение переменной, вклад (\$), 95% CI, коэффициент (наклон на единицу)

В случае моделей с логарифмическим преобразованием ось Y преобразуется обратно в вклады в долларах с помощью формулы $\bar{y} \times (e^{f(x)} - 1)$, что позволяет непосредственно интерпретировать кривые в долларовом выражении.

Для переменных широты и долготы наклон указывается с шагом 0,001 градуса, а не с шагом в один градус, поскольку такой шаг является более значимым географическим шагом.

9.5.2 Фактор-сглаживание Термины — $s(x, \text{by} = \text{factor})$

По одной цветной линии на каждый уровень фактора, наложенные на один график. Полезно для того, чтобы увидеть, как влияние предиктора варьируется между группами (например, между разными районами).

9.5.3 Interaction Terms — $ti(x_1, x_2)$ and $te(x_1, x_2)$

Тепловая карта, отображающая двумерную поверхность вкладов. Цветовая шкала: красный (отрицательный) → белый (нулевой) → синий (положительный). При наведении курсора отображаются как значения переменных, так и величина вклада.

9.5.4 Параметрические члены — линейные и факторные

- Числовые линейные слагаемые : диаграмма рассеяния точек данных с наложенной линией подгонки
- Факторные слагаемые : гистограмма, показывающая средний вклад по каждому уровню фактора

9.6 Корреляция

Матрица тепловой карты, отображающая корреляции Пирсона между всеми числовыми предикторами. Доступна сразу после импорта данных — подгонка модели не требуется.

- Цветовой градиент: синий (−1) → белый (0) → красный (+1)
- Значения корреляции, выведенные в каждую ячейку
- Полезно для выявления конкурвивности (сглаживания мультиколлинеарности) перед подгонкой

9.7 Диагностика

Два набора диагностических визуализаций:

Диагностическая панель (с сайта `gratia::appraise`): четыре графика — остатки по сравнению с прогнозируемыми значениями, Q-Q-график, гистограмма остатков и зависимость реакции от прогнозируемых значений.

Диаграмма рассеяния «фактические значения против прогнозируемых»: наблюдаемые значения против прогнозируемых с пунктирной линией отсчета под углом 45 градусов. Для моделей с логарифмическим преобразованием значения преобразуются обратно в исходную шкалу.

9.8 Корректировки RCA

Три гистограммы, отображаемые после расчёта RCA (только в режиме «оценка/рынок»):

- Корректировка остаточной стоимости% — Распределение корректировок остаточной стоимости в процентах от цены продажи
- Чистая корректировка% — Распределение чистых корректировок
- Валовая корректировка% — Распределение валовых корректировок

В подписи к каждой гистограмме указаны среднее значение, медиана и стандартное отклонение, а также нанесены пунктирные линии, обозначающие среднее значение и медиану.

9.9 Anova

Таблица комбинированного ANOVA, объединяющая параметрические и члены сглаживания из сводки GAM. В ней указаны член, тип (параметрический или сглаживание) и все соответствующие статистические показатели.

9.10 Вывод Mgcvc:

Вывод необработанного текста:

- Подгонка времени
- Формула GAM
- Вывод результатов моделирования
- Полный вывод результатов по методу «`summary(model)`»
- Семейство и функция связи
- Общие показатели конкурвивности

9.11 Проверка знаков

При импорте узлов `earthUI` на этой вкладке сравнивается направление каждой переменной в модели Earth с её направлением в GAM:

- `earth _` направление : возрастающее, убывающее или смешанное (на основе знаков функции шарнира)

- `gam _` направление : возрастающее, убывающее или смешанное (оценивается по гладкой производной на таблице из 100 точек)
- согласованность : TRUE/FALSE
- предупреждение : описание любого несоответствия

В строке состояния отображается сообщение «Все направления согласуются» или количество переменных с несоответствиями.

9.12 Concurvity

Конкурвита — это «сглаживающийся» аналог мультиколлинеарности. На этой вкладке отображаются:

- Таблица общей кривизны : строки для показателей «наихудший», «наблюдаемый» и «оценочный», столбцы — для каждого члена сглаживания. Цветовая кодировка: белый ($< 0,5$), жёлтый ($0,5-0,8$), красный ($> 0,8$).
- Таблица парной конкурвити : матрица парных значений для наихудшего случая с такой же цветовой кодировкой.

Значения, превышающие $\sim 0,8$ в строке «worst», указывают на то, что данное сглаживание может быть хорошо аппроксимировано другими сглаживаниями, что свидетельствует о избыточности.

9.13 Проверка исходных данных

Вывод, полученный с помощью команды `mgcv::gam.check()`: проверка адекватности размерности базиса для каждого члена сглаживания. Если показатель адекватности базиса (k) члена сглаживания слишком мал, эта проверка помечает его низким p -значением, что указывает на необходимость увеличить коэффициент сглаживания (k).

10 Загрузка данных

После подгонки кнопка «Скачать результаты (Excel)» (боковая панель, раздел 7) позволяет экспортировать файл Excel с прогнозами и данными для диагностики в папку вывода mgcv активного проекта.

10.1 Столбцы вывода

Столбец	Описание
basis	Значение переменной «пересечение» после обратного преобразования (одинаковое для всех строк)
<var>_contribution	Вклад каждого сглаживаемого члена в прогноз. Для логарифмических моделей пропорционально распределяется по значениям в долларах.
est_<target>	Прогноз модели (например, est_sale_price), обратно преобразованный из логарифмического масштаба, если применимо
residual	Фактическое значение— прогнозируемое значение
calc_residual	Верификация: фактическое значение— сумма вкладов (обратно преобразованная)
cqa	Оценка «Состояние-Качество-Привлекательность» (CQA) (шкала от 0 до 10)
residual_sf	Остаточная / жилая площадь (если обозначено как « living_area »)
cqa_sf	Показатель CQA, рассчитанный на основе рейтинга по адресу residual_sf

В режиме оценки в строке 1 (объект оценки) значения sf для остаточных величин/cqa/cqa _ установлены как NA. Строки сортируются по значениям sf остаточных величин _ в порядке убывания, если указан столбец жилой площади _.

После успешного завершения на кнопке загрузки появляется белая галочка.

10.2 Результаты CQA: балл

балл

CQA оценивает остаток каждой строки по сравнению со всеми остальными по шкале от 0 до 10:

- Высокий показатель CQA ($\sim 9-10$) : продан значительно дороже прогнозируемой цены — вероятно, отличное состояние/качество/привлекательность
- Низкий показатель CQA ($\sim 0-1$) : продан значительно дешевле прогнозируемой цены — вероятно, неудовлетворительное состояние или продажа в связи с финансовыми затруднениями
- Показатель CQA ~ 5 : близок к медиане

11 Расчеты RCA& Скачивание

Рабочий процесс Метода остаточных ограничений доступен в режимах «Оценка» и «Анализ рыночной зоны» после подгонки модели. Объект оценки должен находиться в первой строке.

11.1 Открытие диалогового окна RCA

Нажмите кнопку «Рассчитать корректировки RCA» & «Скачать» в разделе 8 боковой панели. Появится модальное диалоговое окно со следующим содержанием:

- Тип балла : CQA или CQA на SF (если указано « living_area »). Вариант «на SF» масштабирует остаточный балл объекта оценки с учетом жилой площади.
- Балл CQA объекта : Принимает значения от 0,00 до 10,00. Это поле изначально пустое; после импорта модели earthUI тип бала и значение, использованные в earthUI, переносятся из общих настроек проекта.

Нажмите «Generate» для вычисления и загрузки. Файл Excel сохраняется в папке вывода mgcv активного проекта (<os>_out_mgcgv).

11.2 Интерполяция баллов CQA

- 1 . Балы CQA и остатки для аналогов отсортированы
- 2 . Линейная интерполяция (approx()) сопоставляет значение CQA с остатком
- 3 . Значение для объекта оценки = прогноз модели + интерполированный остаток
- 4 . Строки с весом 0 обрабатываются так, как если бы у них был тот же остаток, что и у объекта оценки

11.3 Столбцы вывода

Колонка	Описание
subject_value	Прогноз по модели + интерполированное отклонение
subject_cqa	Оценка CQA, введенная пользователем
<var>_adjustment	Вклад субъекта– Вклад компании (в долларах)
residual_adjustment	Остаток субъекта– Остаток компании
net_adjustments	Сумма всех корректировок
gross_adjustments	Сумма абсолютных корректировок
residual_adj_pct	Корректировка остатка как% от цены продажи
net_adj_pct	Чистые корректировки как% от цены продажи
gross_adj_pct	Валовые корректировки как% от цены продажи
adjusted_sale_price	Фактическая цена продажи + чистые корректировки

После успешного выполнения вычислений на кнопке появляется белая галочка.

12 Таблица сравнения продаж

Таблица сравнения продаж доступна только в режиме оценки, после расчёта корректировок RCA. Она формирует отформатированную книгу Excel, подходящую для включения в отчёты об оценке.

12.1 Диалоговое окно выбора аналогов

При нажатии кнопки «Сгенерировать таблицу продаж» & «Скачать» открывается модальное диалоговое окно со списком всех аналогов. Аналоги разделены на две группы:

- Рекомендуемые аналоги (уже отмечены): Валовая корректировка < 25% от цены продажи, отсортированные по дате продажи (сначала самые последние). Максимум 30 аналогов.
- Другие аналоги (по умолчанию не отмечены): Остальные аналоги, отсортированные по проценту валовой корректировки.

Выберите аналоги, а затем нажмите «Создать таблицу продаж», чтобы создать книгу.

Обязательные специальные типы : `contract_date` и `living_area` должны быть указаны. Рекомендуемые специальные типы : `latitude` , `longitude` и `lot_size` — таблица будет работать и без них, но некоторые поля останутся пустыми.

12.2 Макет книги

Сгенерированная книга содержит несколько листов, на каждом из которых размещены по 3 аналога рядом друг с другом:

- Строки заголовков : адреса объекта оценки и покупателя, цены продажи, даты продажи и количество дней на рынке
- Сгруппированные строки : местоположение (широта, долгота, площадь), размер участка (площадь участка, габариты участка) и возраст (фактический возраст, эффективный возраст) с заголовками групп и подпунктами
- Строки с переменными модели : по одной строке на каждый сглаживающий член, показывающие вклад объекта оценки, вклад аналогов и суммы корректировок
- Строки с остаточными характеристиками : пустые, разблокированные ячейки для ввода оценщиком для ввода характеристик, не учтенных моделью (например, состояние, качество, вид)
- Сводные строки : чистая корректировка% , валовая корректировка% и скорректированная цена продажи с активными формулами Excel

12.3 Защита листа & Остатки ячейки

Каждый лист защищен от случайного редактирования, однако ячейки со значениями остаточных характеристик явно разблокированы. Это позволяет оценщикам вводить значения для характеристик, отсутствующих в модели, при этом сохраняя расчеты корректировок, основанные на формулах.

13 Скачивание отчетов

Создание отчетов представляет собой двухэтапный рабочий процесс в Quarto, заменивший одноэтапный процесс создания отчетов, использовавшийся в предыдущих версиях.

13.1 Шаг 1: Создание отчёта Quarto Отчёт «

»

Кнопка «Generate Quarto Отчёт» (доступная после того, как модель прошла подгонку) сохраняет автономный набор «.qmd» в папку вывода mgsu активного проекта. Набор содержит исходный код Quarto, предварительно сгенерированные файлы графиков (PNG и PDF), файл report_data.rds с данными, включающий все таблицы и метаданные модели, а также шаблон стиля reference.docx для вывода в Word. Поскольку набор является автономным, вы можете редактировать исходный код .qmd или объединять его с другими документами Quarto перед выводом.

Графики предварительно рендерятся при генерации набора; если генерация отдельного графика завершается сбоем, записывается изображение-заполнитель, а остальная часть набора генерируется в обычном режиме.

13.2 Шаг 2: Конвертация Quarto Отчёт «

»

Раздел «Convert Quarto Report» преобразует файл «.qmd» в один или несколько форматов вывода с помощью «quarto::quarto_render()». По умолчанию используется последний созданный набор, но кнопка «Browse» позволяет выбрать любой файл «.qmd». Доступны три формата:

- HTML — Автономный HTML-документ. Работает на всех платформах без дополнительного программного обеспечения. Рекомендуемый вариант по умолчанию.
- Word (.docx) — Оформляется с использованием входящего в набор шаблона стиля «reference.docx». Подходит для редактирования и распространения.
- PDF — Требуется установка LaTeX (см. «Системные требования» ниже). Если LaTeX не обнаружен, опция PDF автоматически скрывается из списка флажков форматов.

Устаревшие одношаговые рендереры — render_gam_report() и export_gam_docx() — по-прежнему доступны для программного использования.

13.3 Содержание отчёта

Сгенерированный отчёт содержит:

- Набор данных & Модель : файл данных, размер выборки, семейство, метод, а также информация о подгонке
- Сводка результатов : R^2 , CV R^2 , объясненное отклонение, AIC и связанные показатели подгонки
- Уравнение модели : уравнение модели с использованием нотации f_i для сглаживающих слагаемых и явных коэффициентов для линейных слагаемых, а также определения сглаживающих функций. Длинные уравнения разбиваются на несколько строк в PDF.
- Сводные таблицы : процесс сглаживания (EDF, F, p-значение) и параметрические члены (оценка, стандартная ошибка, t, p-значение)
- Включенные взаимодействия : Матрица тензорных и факторно-сглаженных взаимодействий, включенных в модель (отображается только при наличии взаимодействий)
- Важность переменных : Гистограмма, ранжированная по F-статистике (процесс сглаживания) или по коэффициенту «|t|» (параметрические члены)

- Сглаженные эффекты (все типы членов):
 - Одномерные графики с сглаживанием кривых (статические версии интерактивных графиков, созданные с помощью ggplot2)
 - Многолинейные графики «фактор × сглаживание кривой»
 - Тепловые карты взаимодействий (термины *ti* и *te*)
 - Графики параметрических слагаемых (точечный график для числовых, гистограмма для факторов)
- Тепловая карта корреляций предикторов
- Диагностика : диагностика остатков и рассеяние фактических значений по отношению к прогнозируемым
- mgcv Сводка : Полный вывод результатов подгонки модели
- Приложение с обозначениями : Непронумерованный заключительный раздел, объясняющий обозначения функций вклада, использованные в отчёте

После успешного создания отчёта на кнопке «Создать отчёт Quarto» появляется белая галочка.

14 Экспорт функций модели

mgcvUI позволяет экспортировать функции, подвергшиеся сглаживанию при подгонке GAM-модели, в виде автономного кода на нескольких языках программирования. Каждая функция, подвергшаяся сглаживанию, вычисляется на таблице из 200 точек и экспортируется в виде таблицы поиска с линейной интерполяцией, что позволяет использовать модель вне среды R без какой-либо зависимости от пакета mgcv.

В текущей версии экспорт функций осуществляется программным способом. Экспортируемые вспомогательные функции принимают объект результатов подгонки (созданный с помощью `fit_gam()`) и записывают файлы, описанные ниже:

- `generate_r_code()`, `generate_python_code()` и `generate_cpp_code()` — исходные файлы для каждого языка
- `export_functions_sqlite()` — база данных SQLite с таблицами поиска
- `export_functions_zip()` — формирует набор из любой комбинации четырёх форматов в zip-архив

14.1 Функции R

Файл: `gam_functions.R`

- Комментарии заголовка с переменной отклика, R^2 , $CV R^2$ и формулой.
- Константа `INTERCEPT`.
- По одному `approxfun()` на каждую сглаживаемую переменную с заранее вычисленными векторами x/y .
- Коэффициенты линейных слагаемых в виде именованных констант.
- Использование: сложить пересечение, каждый `g_var(value)` и каждый `coef_var * value`.

14.2 Функции Python

Файл: `gam_functions.py`

- `numpy import`.
- Константа `INTERCEPT`.
- Для каждой переменной: массивы `_x_var` и `_fx_var`, функция `g_var(x)` с использованием `np.interp()`.
- Коэффициенты линейных слагаемых в виде констант `COEF_*`.

14.3 Заголовочный файл C++

Файл: `gam_functions.hpp`

- Только заголовок с использованием `#pragma once`.
- Пространство имён `gam_functions`.
- Вспомогательная функция `interp()` с использованием `std::lower_bound`.
- Для каждой переменной: встроенная функция со статическими массивами констант.
- Коэффициенты линейных слагаемых в виде `constexpr double`.

14.4 База данных SQLite

Файл: `gam_functions.sqlite`

- Таблица `smooth_grids`: столбцы (`variable`, `x`, `fx`) — таблицы поиска по 200 точкам.
- Таблица `model_info`: отклик, r^2 в квадрате, `cv_rsqr`, отклонение, объясненная доля, `aic`, `n` — количество наблюдений, семейство, метод, формула.

- Таблица `linear_terms` : переменная, коэффициент.
- Таблица `intercept` : значение пересечения.

15 Сравнение с earthUI и glmnetUI

Приложение «Earth» и его модули постобработки представляют собой единый набор инструментов для регрессионного моделирования. Они используют общий формат данных, специальные типы столбцов, рабочий процесс RCA и демонстрационные наборы данных, но при этом используют разные механизмы моделирования.

15.1 Основные различия

Функции	mgcvUI	earthUI	glmnetUI
Методы	GAM (mgcv)	MARS/Earth (earth)	эластичная сеть (glmnet)
Взаимосвязи	Сглаживаемые нелинейные	Кусочно-линейные	Линейные
Отбор переменных	select=TRUE Сжатие	Пошаговое прореживание	Штраф Lasso
Взаимодействия	Тензорные произведения (te/ti)	Шарнирные произведения (степень 2–3)	Парные произведения
Частичные эффекты	Гладкие кривые с доверительным интервалом	g-функции (кусочные)	Линейные наклоны
Конвейер Earth	Импорт узлов из Earth	Источник узлов	Импорт базиса шарниров
Экспорт кода	R, Python, C++, SQLite	R, Python, C++, SQLite	—
Форматы отчетов	Word, PDF, HTML	Word, PDF, HTML	Word, PDF, HTML

Стоит отметить одну отличительную особенность, присущую исключительно earthUI: earthUI включает экспериментальную функцию Prolog/DCG, которая извлекает структурированные характеристики объектов недвижимости из текстовых комментариев MLS.

15.2 Общие функции

Все три инструмента предоставляют:

- Общее дерево проектов regProj (по умолчанию ~/regProj) — проекты, файлы ввода и выходы для каждого метода используются совместно всеми приложениями одного уровня
- Импорт из CSV и Excel с помощью snake_ Преобразование столбцов с учетом регистра
- Поддержка локалей более 30 стран с указанием разделителя в CSV, знака десятичной запятой и формата бумаги
- Специальные обозначения столбцов (контракт_ дата, жилая площадь_ , площадь, широта, долгота и т. д.)
- Автоматический расчет срока действия продажи_ расчета срока действия с даты оценки
- Переключение между темным и светлым режимами (тема Nord)
- Сохранение настроек (SQLite и/или localStorage)
- Тепловая карта корреляции, важность переменных, диагностика
- Оценка CQA и рабочий процесс корректировки RCA

- Таблица сравнения продаж с выбором аналогов для сравнения и отформатированным выводом в Excel
- Совместимость с теми же демонстрационными наборами данных (Appraisal_1.csv)

15.3 Рекомендуемый рабочий процесс

- Начните с earthUI , чтобы выявить важные переменные, нелинейные зависимости и взаимодействия. Автоматическое построение модели в earthUI — лучшая отправная точка.
- Если вам нужны более плавные частичные эффекты сглаживания, доработайте модель с помощью mgcvUI . Импортируйте узлы из earthUI для плавного перехода.
- Используйте glmnetUI, если вам нужна чисто линейная модель с агрессивным отбором переменных или ограничениями на знаки коэффициентов.
- Сравните все три подхода на одних и тех же данных, чтобы оценить, оправдана ли дополнительная сложность нелинейных моделей.

16 Демо-набор данных: Оценка _ 1.csv

16.1 Описание

Демо-набор данных используется как в earthUI, так и в glmnetUI. Он находится в папке «demo_mls/» исходного кода mgcvUI, либо, если установлен earthUI, в:

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

Файл содержит примерно 1 500 записей о продажах жилой недвижимости, полученных в результате симулированного экспорта данных из системы MLS, причем рассматриваемый объект оценки находится в первой строке, а аналогичные сделки — в остальных строках. Данные отражают продажи односемейных домов на рынке, охватывающем несколько районов, с различными параметрами площади, возраста и местоположения объектов недвижимости.

Это не реальные данные, но они основаны на реальном районе в Северной Калифорнии. Вся идентифицирующая информация была изменена или удалена.

16.2 Столбцы

Колонка	Тип	Специальный тип	Описание
sale_type	символ	sale_type	Код типа записи, позволяющий отличить данный объект от аналог продаж
weight	числовой	weight	Вес наблюдения (0 = исключить из моделирования)
id	числовой	display_only	Внутренний идентификатор записи
property_id	числовой	display_only	Идентификатор объекта недвижимости в системе MLS
listing_id	символ	display_only	Номер объявления в системе MLS
parcel_number	символ	display_only	Номер участка по оценке округа (APN)
street_address	символ	display_only	Адрес объекта недвижимости
city_name	символ	display_only	Город
postal_code	символ	display_only	Почтовый индекс
county_name	символ	display_only	Округ
contract_date	Дата	contract_date	Дата заключения договора купли- (рассчитывается sale_age)
sale_age	числовой	—	Количество дней от даты заключения договора до даты оценки (рассчитано заранее)
coe_date	Дата	display_only	Дата завершения эскроу
listing_status	символьный	display_only	Статус объявления (например, «Продано»)
sale_price	числовой	(целевой)	Цена продажи — переменная ответа
rent	числовой	—	Ежемесячная арендная плата (для моделей с несколькими целевыми переменными)
list_price	числовой	display_only	Цена объявления
original_list_price	числовой	display_only	Первоначальная цена объявления
living_sqft	числовой	living_area	Общая жилая площадь в квадратных футах
ppsf	числовой	display_only	Цена за квадратный фут
beds_total	целое число	—	Количество спален
baths_total	числовое	—	Общее количество ванных комнат (например, 2,5 = 2 полные + 1 туалет)
lot_size	числовое	lot_size	Площадь участка в квадратных футах
area_id	целое число	area	Идентификатор района в системе MLS
area_text	символ	display_only	Название района
age	числовое	actual_age	Возраст объекта недвижимости в годах

Колонка	Тип	Специальный тип	Описание
year_built	целое число	display_only	Год постройки
latitude	числовое	latitude	Широта (округлённая до 3 знаков после запятой для модели)
longitude	числовой	longitude	Долгота (округлённая до 3 знаков после запятой для модели)
latitude6	числовой	display_only	Широта с полной точностью
longitude6	числовой	display_only	Долгота с полной точностью
garage_spaces	целое число	—	Количество гаражных мест
fp_count	целое число	—	Количество каминов
no_of_stories	число	—	Количество этажей
style	символ	—	Архитектурный стиль
view	символ	—	Тип вида (например, «Окрестности», «Холмы»)
days_on_market	целое число	dom	Количество дней на рынке
listing_date	Дата	listing_date	Дата размещения объявления
sale_concessions	число	concessions	Уступки продавца

16.3 Рекомендуемое руководство по быстрому запуску

1. Запустите процедуру `mgcv: earthUI::launch_mgcv()`
2. Создайте новый проект (раздел 1) с целью «For оценка»
3. Поместите файл `Appraisal_1.csv` в папку вводных данных проекта и выберите его в разделе 2 (при необходимости нажмите «Refresh»)
4. В качестве целевого показателя выберите « sale_price »
5. При желании установите для параметра «Response Transform» значение «Log (natural)» для пропорциональных корректировок
6. Назначьте специальные типы, как показано в таблице выше
7. Включить предикторы: `sale_age` , `living_sqft` , `baths_total` , `lot_size` , `area_id` (в качестве фактора), `age` , `latitude` , `longitude` , `garage_spaces`
8. Сохранить настройки по умолчанию: `REML`, `gamma = 1.2`, `tp basis`, `k = auto`, `select = TRUE`
9. Нажмите «Подгонка MgcV GAM Model»
10. Просмотрите вкладки «Summary», вкладки «Contribution» (графики сглаживания) и «Диагностика»
11. Скачайте вывод (шаг 7), просмотрите рейтинг CQA
12. Рассчитайте корректировки RCA (шаг 8) с баллом CQA ~5,00
13. Сгенерируйте таблицу сравнения продаж (шаг 9) с рекомендуемыми аналогами

17 Системные требования& Устранение неполадок

17.1 Поддерживаемые платформы

mgcvUI работает на macOS, Windows и Linux. Приложение разработано и в основном протестировано на macOS с использованием RStudio. Примечания, касающиеся конкретных платформ, приведены ниже.

17.2 Версия R:

Требуется $R \geq$ версии 4.1.0. Настоятельно рекомендуется использовать RStudio Desktop (версия 2023.06 или более поздняя) — в него входит набор пакетов Quarto и pandoc (необходимые для визуализации отчетов), а также он предоставляет удобную среду для запуска приложения.

17.3 Необходимые пакеты R:

Следующие пакеты устанавливаются автоматически в качестве зависимостей:

Package	Purpose
shiny	Фреймворк для веб-приложений
mgcv	Подгонка моделей GAM
valengrCore	Общий модуль оценки (специальные роли, перенос настроек из earthUI)
ggplot2, plotly	Статические и интерактивные графики вклада
gratia	Извлечение и диагностика сглаживаемых членов
DT	Интерактивные таблицы данных
quarto, rmarkdown	Создание и визуализация отчётов с помощью Quarto
readxl	Импорт файлов Excel
shinyFiles	Браузеры файлов и папок
RSQLite, DBI	Настройки и базы данных проектов (SQLite)
jsonlite	Сериализация состояния переменных

17.4 Дополнительные зависимости

Component	When Needed
LaTeX (TinyTeX, MiKTeX или MacTeX)	Только для генерации отчетов в формате PDF. Если установка LaTeX не обнаружена, опция PDF автоматически скрывается из раскрывающегося списка форматов. Установить с помощью: <code>tinytex::install_tinytex()</code>
sysfonts + showtext	Шрифт Roboto Condensed для графиков ggplot2. Если он недоступен или отсутствует, приложение автоматически переключается на системный шрифт sans-serif.
earth ($\geq 5.3.0$)	Требуется только для конвейера импорта earthUI. Не требуется для автономного использования.
callr	Фоновая подгонка модели (обеспечивает отзывчивость приложения). Без него подгонка выполняется синхронно.
writexl, openxlsx	Загрузка файлов Excel (вывод моделирования, RCA, таблица продаж). Приложение выдает запрос, если они отсутствуют.
bslib	Тематическое оформление Bootstrap 5 (Nord Light/Dark).
officer	Устаревший программный экспорт в Word (<code>export_gam_docx()</code>).

17.5 Примечания по конкретным платформам

17.5.1 macOS

Минимальное количество проблем. Для системных библиотек рекомендуется использовать Homebrew. Если требуются отчёты в формате PDF:

```
# In R:
install.packages("tinytex")
tinytex::install_tinytex()
```

17.5.2 Windows

Хорошо работает с RStudio Desktop. Важные замечания:

- Корпоративные/заблокированные компьютеры : если доступ к временному каталогу или каталогу пользовательских данных ограничен, сохранение настроек и формирование отчётов могут завершиться сбоем. Приложение отобразит предупреждение в консоли R, но продолжит работу без сохранения настроек.
- LaTeX : Установите TinyTeX из R (та же команда, что и для macOS выше) или установите MiKTeX с сайта miktex.org .
- Длинные пути к файлам : В Windows действует ограничение на длину пути в 260 символов. Старайтесь, чтобы путь к рабочему каталогу был коротким.

17.5.3 Linux

Наибольшие различия наблюдаются между различными дистрибутивами. Пользователям Ubuntu/Debian могут потребоваться системные библиотеки для компиляции пакетов R:

```
sudo apt install libcurl4-openssl-dev libssl-dev \
libxml2-dev libsqlite3-dev libfontconfig1-dev
```

Для просмотра отчетов в формате PDF: `tinytex::install_tinytex()` или установите `texlive-full` через системный менеджер пакетов.

На серверах без монитора (без дисплея) интерактивные графики Plotly отображаются в отчётах в виде статических запасных вариантов. Для работы самого приложения Shiny требуется подключение к веб-браузеру .

17.6 Плавная деградация

`mgcvUI` разработан таким образом, что может работать даже при отсутствии дополнительных компонентов:

Отсутствующий компонент	Поведение
Отсутствует LaTeX	Опция PDF скрыта в раскрывающемся списке форматов отчетов. Форматы HTML и Word по-прежнему доступны.
Отсутствует интернет / ошибка шрифтов	Шрифт Roboto Condensed заменен системным шрифтом sans-serif. В консоль занесено сообщение.
Отсутствует RSQLite / файловая система только для чтения	Настройки не сохраняются между сеансами. Отображается предупреждение в консоли. В остальном приложение работает нормально.
<code>gratia::appraise</code> завершается с ошибкой	На вкладке «Диагностика» и в отчетах отображается сообщение-заполнитель. В качестве альтернативы используйте команду <code>mgcv::gam.check()</code> в консоли.
Временный каталог недоступен для записи	Сгенерировать отчёт не удастся с явным сообщением об ошибке. Устраните проблему, установив переменную среды <code>TMPDIR</code> в путь к каталогу, доступному для записи.

17.7 Устранение неполадок в программе «

»

17.7.1 «Параметр PDF недоступен»

Установка LaTeX не обнаружена. Выполнить в R:

```
install.packages("tinytex")
tinytex::install_tinytex()
```

Затем перезапустите приложение. Появится опция «PDF».

17.7.2 «Настройки не сохраняются»

Не удаётся создать каталог базы данных SQLite или выполнить запись в него. Проверьте права доступа к:

- macOS/Linux : `~/Library/Application Support/R/mgcvUI/` или `~/local/share/R/mgcvUI/`
- Windows : `%APPDATA%/R/data/mgcvUI/`

17.7.3 «График диагностики недоступен»

Функция `gratia::appraise()` иногда завершается с ошибкой при работе со сложными моделями (множество тензорных взаимодействий, нестандартные семейства). Это известная проблема в исходном коде. Обходной путь: запустите команду `mgcv::gam.check(model)` в консоли R, чтобы сразу увидеть графики диагностики.

17.7.4 “Отчёт об ошибке: количество элементов, которые необходимо заменить ...”

Это была известная проблема в более ранних версиях, связанная с пакетом `gratia`. Убедитесь, что у вас установлена последняя версия `mgcvUI`. Текущая версия перехватывает эту ошибку и генерирует отчёт с заполнителем в разделе диагностики.

17.7.5 Порт 7880 уже занят

Если приложение не запускается из-за того, что порт 7880 занят (в результате предыдущего сеанса), выполните следующую команду:

```
# macOS/Linux:
lsof -ti:7880 | xargs kill
```

```
# Windows (PowerShell):
Stop-Process -Id (Get-NetTCPConnection -LocalPort 7880).OwningProcess
```

17.7.6 Ошибка установки пакета в Linux

Отсутствуют системные библиотеки. Установите заголовочные файлы для разработки, перечисленные в разделе «Linux» выше, а затем повторите попытку выполнить команду `install.packages()`.

Список литературы

- Wood, S. N. (2003). «Thin-plate regression splines». B: Journal of the Royal Statistical Society (B) 65.1, с. 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». B: Journal of the American Statistical Association 99.467, с. 673—686. DOI: [10.1198/016214504000000980](https://doi.org/10.1198/016214504000000980).

- Wood, S. N. (2011). «Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models». В: Journal of the Royal Statistical Society (B) 73.1, с. 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. 2-е изд. Chapman и Hall/CRC.
- Wood, S. N., N. Pya и B. Säfken (2016). «Smoothing parameter and model selection for general smooth models (with discussion)». В: Journal of the American Statistical Association 111, с. 1548—1575. DOI: [10.1080/01621459.2016.1180986](https://doi.org/10.1080/01621459.2016.1180986).

Обзор

Foundations • The Valuation Engineer

Берт Крейтор, редактор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/82vssv81](https://doi.org/10.67330/82vssv81)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Издание «Foundations» публикует точные, пригодные для цитирования изложения концепций, лежащих в основе современной методологии оценки. В статьях рассматриваются термины, с которыми сталкиваются практикующие оценщики — иногда под другими названиями, иногда даже не осознавая их значения — и помещаются в их формальный контекст: откуда они взялись, что означают с технической точки зрения, где встречаются на практике и почему они важны для обоснованности оценки.

В первом выпуске определены шесть основных понятий гедонической модели, представленных таким образом, что каждое из них опирается на терминологию предыдущих. Последовательность начинается с «Неоднородного блага» (свойства недвижимости, которое и делает методологию оценки вообще необходимой), проходит через «Пространство характеристик» и «Гедоническую функцию цены» (формальный аппарат, который систематизирует набор характеристик, представленных объектом недвижимости), далее через *неявную цену* и *гедоническую регрессию* (как маргинальный вклад каждой характеристики выявляется на основе рыночных данных), и, наконец, до *латентной переменной* (что на самом деле представляет собой невыявленный остаточный сигнал в любой гедонической модели).

Один наглядный пример — восемь аналогов из района Пасифика, штат Калифорния — проходит красной нитью через все шесть статей и служит педагогической нитью данного номера.

Почему именно эти шесть, именно в таком порядке

Выбор этих шести понятий не является произвольным. Каждое из них играет ключевую роль в современной гедонической методологии, каждое действительно встречается в повседневной практике оценки, и каждое настолько часто неправильно понимается или остается без названия, что его формулировка в виде цитируемого определения имеет практическую ценность.

Порядок определяется концептуальной зависимостью. Понятие «неоднородное благо» является исходным: оно обозначает свойство недвижимости (а также поддержанных автомобилей, рынков труда и ряда других товаров), которое обуславливает необходимость всей гедонической модели. Пространство характеристик — это формальный аппарат для представления неоднородных благ в виде векторов; без него невозможно дать точные выводы. Гедоническая функция цены — это объект равновесия, который отображает точки в пространстве характеристик на цены. Неявные цены — это её частные производные — предельные вклады, которые можно интерпретировать как утверждения при каждой корректировке в таблице сравнения продаж, независимо от того, называет ли их так оценщик или нет. Гедоническая регрессия — это эмпирическая оценка, позволяющая восстановить неявные цены на основе наблюдаемых сделок; именно здесь теория переходит в анализ данных. Латентные переменные начинают действовать там, где начинаются остатки регрессии: это ненаблюдаемые характеристики, влияние которых сохраняется в остатке в виде структурированного сигнала, а не в виде шума, который можно игнорировать.

Читатель, следующий этой последовательности, должен проследить непрерывную линию размышлений от вопроса «зачем вообще нужна методология оценки?» до «как следует трактовать остатки в модели, подстроенной для подгонки?» — это путь, который, по опыту автора, большинство оценщиков проходят только по частям в течение многих лет практики.

Пример кода:

Один набор данных, включающий восемь аналогов жилой недвижимости в городе Пасифика (Калифорния), используется во всех шести заявках:

Адрес	Площадь ()	Площадь под застройку (GLA, sf)	Участок (sf)	Вид	Состояние	Цена продажи
A		1,650	9,000	да	хорошее	\$ 1,425,000
B		1,650	7,200	нет	хорошо	\$ 1,280,000
C		1,850	7,500	нет	хорошо	\$ 1,407,500
D		1,450	7,000	нет	хорошо	\$ 1,155,000
E		1,700	8,000	да	отлично	\$ 1,470,000
F		1,550	7,100	нет	средне	\$ 1,177,500
G		1,700	10,500	нет	хорошо	\$ 1,392,500
H		1,600	7,300	нет	хорошо	\$ 1,312,500

Пример намеренно выбран достаточно небольшим, чтобы его можно было проанализировать вручную, и достаточно насыщенным, чтобы проработать каждую концепцию. В записи 001 рассматривается разница в \$145,000 между аналогами A и B как совокупный вклад двух характеристик (вид и участок), взаимосвязанных в рамках одного наблюдения за парной сделкой. В записи 002 все восемь аналогов представлены в виде точек в четырёхмерном пространстве характеристик и показано, что наивные метрики расстояния приводят к их неверной ранжировке. В записи 003 предлагается функциональная форма гедонической функции цены в этом районе. В записи 004 аналитически восстанавливаются неявные цены и исходная разницу «A против B» на вклад от вида и участка в размере \$ 127,039, при этом \$ 18,000 остаются неучтёнными. В записи 005 модель формально подгоняется в R и предоставляется полный вывод из `lm()`, включая стандартные ошибки и остатки. В записи 006 рассматривается единственное остаточное значение, которое модель не может объяснить — отклонение +\$ 46,455 по `Comp H` — и устанавливается связь между ним и концепцией латентных переменных, лежащей в основе Метода остаточных ограничений, разработанного автором в других работах.

Набор данных доступен в репозитории данного выпуска по адресу `shared/data/pacifica_comps.csv` для читателей, желающих воспроизвести какие-либо из расчётов.

Как читать этот выпуск

Читателю, не знакомому с гедонической методологией, рекомендуется читать шесть статей в порядке. Взаимосвязь между ними реальна: каждая статья опирается на терминологию, установленную в предыдущих, и пропуск каких-либо статей может привести к путанице, которую сами статьи не могут предвидеть.

Читатель, уже знакомый с гедонической регрессией — к числу которых относятся большинство оценщиков, прошедших курс углублённой статистики, работавших с системой автоматической оценки недвижимости (AVM) или знакомых с литературой по массовой оценке, —

может использовать данную статью в качестве повторения материала и справочника. Статьи написаны таким образом, чтобы их можно было цитировать по отдельности; статья, опубликованная в другом разделе журнала, в которой упоминается «*неявная цена*», может сослаться на статью № 004 по DOI, а не повторно излагать данное понятие.

Читателю, интересующемуся конкретно методологией Метода остаточных ограничений, следует сначала ознакомиться со статьёй № 006, чтобы понять концепцию латентных переменных, а затем вернуться к статьям № 003–005, посвящённым гедонический аппарат, который расширяет RCA, а затем к записям 001–002, где изложена базовая терминология, на которой основано всё остальное.

Читатель, интересующийся вопросом о том, *что делает корректировку оценки обоснованной*, обнаружит, что раздел «Почему это важно для обоснованности» в каждой статье, прочитанный по порядку, представляет собой краткое эссе о методологической обоснованности оценки, основанной на корректировках. Это эссе нигде в данном выпуске не представлено в явной форме, но оно является неявным объектом оценки каждой статьи.

Редакционная позиция

«*Foundations*» — это не учебный раздел. Каждая статья достаточно кратка, чтобы прочитать её за один присест; она написана для читателей, уже хорошо знакомых с практикой оценки, и призвана дать точное определение понятия, а не исчерпывающее изложение. Цель заключается в том, чтобы обеспечить надёжный ориентир, на который могут ссылаться другие статьи в журнале.

Статьи проходят редакционную экспертизу и могут сопровождаться приглашенными комментариями от эксперта в данной области. Они публикуются по лицензии Creative Commons Attribution и получают DOI как через Crossref (в качестве канонической ссылки), так и через Zenodo (в качестве версионированного архива). Каждая статья может со временем уточняться — незначительные изменения обновляют версию в Zenodo, а существенные изменения обновляют оба DOI с редакционной заметкой, описывающей внесенные изменения. DOI Crossref всегда указывает на самую последнюю версию; доступ к конкретным историческим версиям осуществляется через Zenodo.

В будущих выпусках словарный запас *рубрики «Основы»* будет расширен за счет смежных областей: пространственной зависимости, условного математического ожидания, эндогенности, смещения из-за пропущенных переменных, идентификации, а также формального аппарата, позволяющего отличать прокси-переменные от латентных переменных. Приглашаем вас присылать предложения по концепциям, заслуживающим рассмотрения; их можно направлять редактору.

Шесть материалов в этом выпуске:

1. **Неоднородное благо** — почему вообще необходима методология оценки.
2. **Пространство характеристик** — векторное представление атрибутов набора.
3. **Гедоническая функция цены** — равновесное отображение от характеристик к ценам.
4. **Неявная цена** — предельная рыночная цена каждой характеристики.
5. **Гедоническая регрессия** — оценка функции на основе наблюдаемых сделок с полным примером в R.
6. **Латентная переменная** — восстановление структурированного сигнала из остатков.

Список литературы.

Корт, А. Т. (1939). «Гедонические ценовые индексы на примере автомобильной промышленности». В сборнике «Динамика спроса на автомобили». General Motors Corporation.

Ланкастер, К. Дж. (1966). «Новый подход к теории потребительского поведения». «Журнал политической экономики», 74(2), 132–157.

Розен, С. (1974). «Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции». «Журнал политической экономики», 82(1), 34–55.

Крейтор, Б. (2025). Метод остаточных ограничений: концептуальная основа & Протокол. Zenodo. doi:10.5281/zenodo.1478

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (000-overview.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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).

```

Неоднородное благо

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/slscsv65](https://doi.org/10.67330/slscsv65)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. Однородный товар — это такой товар, при покупке которого любой конкретный экземпляр подходит: бочка нефти марки WTI, бушель желтой кукурузы № 2, киловатт-час электроэнергии из сети, акция конкретной компании. Стандарты классификации или технические характеристики, лежащие в основе этих товаров, существуют именно для того, чтобы отдельные единицы были взаимозаменяемыми, так что цены относятся к категории, а не к конкретной единице. Неоднородное благо — это такое благо, при котором важна конкретная единица: именно *этот* дом, а не *тот*; именно *этот* подержанный автомобиль, а не тот, что стоит рядом.

Две особенности определяют гетерогенность товара в техническом смысле. Во-первых, покупатели обращают внимание на ряд различных характеристик, которые варьируются в зависимости от отдельных единиц товара на рынке. Во-вторых, эти характеристики объединены в набор в силу технологии производства данного товара; их невозможно приобрести по отдельности. Невозможно приобрести 200 квадратных футов общей жилой площади отдельно. Невозможно приобрести школьный округ независимо от земельного участка. Эти характеристики поставляются в комплексе, и рынок устанавливает цену на целые наборы, а не на сами характеристики.

Недвижимость является типичным примером неоднородного блага и демонстрирует данное свойство в необычайно яркой степени. Характеристики местоположения абсолютно невоспроизводимы: два земельных участка не могут занимать одно и то же место, поэтому каждый объект недвижимости обладает по крайней мере одной характеристикой (своими координатами), которая отличает его от всех остальных объектов на рынке. Физические, юридические и экологические характеристики накладывают дополнительные измерения вариативности на вектор, который в целом является уникальным для каждого объекта недвижимости.

Где оценщик встречается с этим понятием. Каждая корректировка в сравнительном подходе является косвенным признанием того, что оцениваемый объект и аналоги представляют собой неоднородные благосы. Если бы объекты недвижимости были однородными, цены продажи уже были бы непосредственно сопоставимы; никакой таблицы не потребовалось бы. Таблица представляет собой попытку сопоставить цены различных наборов объектов путем корректировки с учётом различий в их составе. Стоит отметить, чего таблица *не* делает: она никогда не разлагает цену продажи на вклад отдельных характеристик в её стоимость. В традиционной таблице нет столбца для вклада каждой характеристики в стоимость — есть только корректировки, оцениваемые напрямую, по каждой характеристике в отдельности.

Затратный подход учитывает неоднородность путем реконструкции набора из его составляющих — фундамента, каркаса, отделки, амортизации — и оценки стоимости каждой из них

в отдельности. Доходный подход решает эту проблему путем абстрагирования набора в единый профиль денежных потоков, при этом ставки капитализации и доходности учитывают неоднородность на уровне набора в виде рыночного множителя. Все три подхода к оценке являются методологическими ответами на тот основополагающий факт, что недвижимость является неоднородным благом. Они отличаются тем, как они учитывают этот факт, а не тем, сталкиваются ли они с ним.

Оценщики также неофициально сталкиваются с гетерогенностью всякий раз, когда они размышляют о выборе аналогов, анализе парных продаж или сравнении «яблоко к яблоку». Стремление найти аналоги, схожие по как можно большему числу характеристик, является реакцией специалиста на неоднородность: минимизировать совокупное расстояние между объектом оценки и аналогом, при этом ненаблюдаемая составляющая этого совокупного расстояния в меньшей степени влияет на разницу в ценах продажи.

Почему это важно для обоснованности. Неоднородность объектов недвижимости является причиной того, что методология оценки должна существовать как самостоятельная дисциплина. Если бы объекты недвижимости были взаимозаменяемыми, рыночная стоимость можно было бы непосредственно определить по объявленным ценам; оценка свелась бы к простому поиску в справочнике. Поскольку это не так, стоимость необходимо выводить на основе цен на схожие, но не идентичные объекты, и каждый этап вывода представляет собой методологический выбор, который необходимо обосновать.

Формальное признание неоднородности позволяет прояснить ряд практических вопросов, от которых зависит обоснованность USPAP и «Правила стандартов». Подбор аналогов — это выбор объектов в пространстве характеристик, достаточно близких к объекту оценки, чтобы полученные на их основе неявные цены были обобщаемыми. Обоснованность корректировки зависит от того, можно ли подтвердить базовую неявную цену рыночными данными. Согласование между подходами отражает суждение оценщика о том, какой подход наиболее достоверно учитывает неоднородность *данного* объекта недвижимости. Каждый из этих вопросов обоснованности становится более острым, как только в качестве базового условия указывается «неоднородное благо».

Разобранный пример оценки. Две недавние сделки по продаже недвижимости в одном и том же районе Пасифика, штат Калифорния:

Характеристики	Комплект А	Комплект В
Общая жилая площадь (sf)	1,650	1,650
Спальни / Ванные комнаты	3 / 2	3 / 2
Площадь участка (кв. футов)	9,000	7,200
Вид	Да	Нет
Цена продажи	\$ 1,425,000	\$ 1,280,000

В рамках модели однородных благ это одно и то же объект недвижимости: одинаковая общая жилая площадь, одинаковое количество комнат, один и тот же район. В рамках модели неоднородных благ это отдельные наборы характеристик, различающиеся по двум параметрам (участок и вид из окна), которые покупатели оценивают по отдельности. Разница в цене в \$145,000 является *совокупным* вкладом этих двух характеристик — двух неявных цен, переплетенных в одной паре.

Для индивидуального вычисления неявных цен требуется больше пар, в идеале — с изменением одной характеристики за раз. Три или четыре дополнительные сделки, охватывающие

комбинации «с видом/без вида» и «большой участок/малый участок», позволяют начать определять неявные цены по отдельности. При достаточном разнообразии среди достаточного числа аналогов гедоническая регрессия формально выполняет разграничение; при ограниченном разнообразии оценщик вынужден прибегать к рассуждениям на основе парных продаж с интервальным диапазоном, надежность которых явно зависит от того, насколько близко наборы друг к другу находятся в пространстве характеристик.

Этот пример наглядно показывает, почему неоднородность является основным условием: даже два «одинаковых» дома не являются одинаковыми, и методологический аппарат оценки существует именно для того, чтобы учитывать различия между ними.

Список литературы.

Ланкастер, К. Дж. (1966). Новый подход к теории потребителя. *Journal of Political Economy*, 74(2), 132–157.

Розен, С. (1974). Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции. «*Журнал политической экономики*», 82(1), 34–55.

Перекрёстные ссылки: characteristics space; hedonic price function; implicit price.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (001-heterogeneous-good.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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).
```

Характеристики Пространство

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/3uymh010](https://doi.org/10.67330/3uymh010)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. В предыдущем разделе было установлено, что неоднородное благо представляет собой отдельный набор характеристик. Естественный следующий вопрос: как нам следует формально описывать и сравнивать такие наборы? Ответ, предложенный Ланкастером (1966), заключается в том, чтобы рассматривать каждый товар как точку в векторном пространстве, осями которого являются характеристики, важные для покупателей.

В этой модели жилой объект в Пасифике — это уже не «дом в Парк-Пасифика», а вектор с координатами, характеризующими общую жилую площадь, площадь участка, вид из окна, состояние, школьный округ, расстояние до океана и так далее. Два объекта недвижимости, у которых совпадают все эти значения характеристик, будут занимать одну и ту же точку в пространстве характеристик; объекты, отличающиеся по какой-либо характеристике, занимают разные точки. Геометрия этого пространства кодирует степень сходства или различия между двумя объектами таким образом, что учитываются все аспекты, по которым покупатели проводят различия.

Количество измерений зависит от того, какие характеристики являются значимыми для исследуемого рынка. В случае типичной оценки жилой недвижимости соответствующее пространство может иметь около дюжины измерений: физические (общая арендуемая площадь, участок, количество спален, ванных комнат, возраст, качество, состояние), связанные с местоположением (школьный округ, вид из окна, близость к объектам инфраструктуры или факторам, снижающим привлекательность района) и юридические/экономические (зонирование, сервитуты). В случае оценки коммерческой недвижимости значимые измерения смещаются в сторону характеристик, связанных с денежными потоками (структура аренды, качество арендаторов, степень заполняемости), но лежащая в основе геометрическая идея остается той же.

Полезность пространства характеристик заключается в том, что оно преобразует качественную интуицию специалиста («эти аналоги похожи на объект оценки») в количественное утверждение («эти аналоги близки к объекту оценки в пространстве характеристик»). Как только расстояние определено, каждое решение о сравнении с точки зрения продаж можно рассматривать как утверждение о положении в этом пространстве.

Где оценщик встречается с этим понятием. Оценщики постоянно работают в пространстве характеристик, как правило, даже не осознавая этого. Каждая строка таблицы сравнительных объектов представляет собой ось координат; каждый аналог — вектор значений координат; каждая корректировка — попытка переместиться вдоль одной оси, удерживая остальные в неизменном положении.

Восемь аналогов в районе Пасифика, которые мы будем использовать на протяжении всего этого выпуска, наглядно иллюстрируют данную конструкцию. Каждый аналог представляет

собой точку в четырёхмерном пространстве характеристик, определённом общей арендуемой площадью (GLA), площадью участка, видом из окна и состоянием объекта:

Площадь ()	Площадь застройки (GLA, sf)	Площадь участка (sf)	Вид	Состояние	Цена продажи
Адрес	1,650	9,000	да	хорошее	\$ 1,425,000
B	1,650	7,200	нет	хорошо	\$ 1,280,000
C	1,850	7,500	нет	хорошо	\$ 1,407,500
D	1,450	7,000	нет	хорошо	\$ 1,155,000
E	1,700	8,000	да	отлично	\$ 1,470,000
F	1,550	7,100	нет	средне	\$ 1,177,500
G	1,700	10,500	нет	хорошо	\$ 1,392,500
H	1,600	7,300	нет	хорошо	\$ 1,312,500

Аналоги A и B из предыдущего материала представляют собой первые две строки; остальные шесть дополняют этот пример различными способами, которые мы будем использовать на протяжении всей остальной части выпуска.

Каждый аналог представляет собой точку в $\mathbb{R}^4$:

- Аналог A находится по адресу (1650, 9000, 1, 2) .
- Аналог B находится по адресу (1650, 7200, 0, 2) .
- Аналог G находится по адресу (1700, 10500, 0, 2) .

Координаты вида и состояния кодируются в числовом виде (вид — $\{0, 1\}$, состояние — $\{1, 2, 3\}$); цена *не* является координатой в пространстве характеристик, а представляет собой функцию положения внутри него.

Почему это важно для обоснованности. Концепция пространства характеристик позволяет более четко определить, что на самом деле представляет собой выбор аналогов. Аналог — это не просто «похожий объект», а точка в соответствующей области пространства характеристик, расположенная рядом с объектом оценки. Чем ближе аналог находится к объекту оценки в этом пространстве, тем меньше требуется корректировок и тем более надежным становится полученное значение .

Такое переосмысление выявляет ряд практических вопросов, которые неформальная таблица оставляет за кадром:

- *Какие параметры имеют значение?* Выбор осей — это решение, принимаемое в процессе моделирования. Включение «вида» в качестве оси имеет смысл только в том случае, если рынок отдельно учитывает его в цене; включение «ориентации по фэн-шуй» имеет смысл только в том случае, если какой-либо сегмент рынка учитывает этот фактор при ценообразовании. Обоснованная оценка требует обоснования выбора осей, а не только значений, зафиксированных на них.
- *Какая метрика расстояния применяется?* Два аналога могут различаться по цене на \$ 200,000, поскольку они находятся далеко друг от друга в пространстве характеристик, но для вычисления расстояния требуется масштабирование по каждой оси. Непреобразованное евклидово расстояние рассматривает один квадратный фут общей арендуемой площади (GLA) как эквивалент одного квадратного фута участка, чего не делает рынок.

- *Находятся ли аналоги в той области, где ценовая функция ведет себя предсказуемо?* Если объект оценки находится в разреженном углу пространства характеристик, то рядом с ним нет ни одного аналога, и оценка стоимости неизбежно носит экстраполяционный характер.

Разобраный пример оценки. Восемь аналогов в Пасифике занимают относительно компактную область в четырёхмерном пространстве характеристик. Рассмотрим вопрос о том, какой из аналогов «ближе всего» к гипотетическим характеристикам объекта оценки по адресу (1,700, 7,500, 1, 2) — 1,700 кв. футов, участок площадью 7,500 кв. футов, вид из окна, хорошее состояние.

Если мы наивно воспользуемся евклидовым расстоянием с немасштабированными осями, то параметр «участок» будет доминировать, поскольку его значения (7,000–10,500) числово намного больше, чем значения параметров «вид» (0–1) или «состояние» (1–3). Расстояния до объекта оценки:

Аналог	Расстояние (немасштабированное евклидово)	Примечания
Аналог	1,501	Аналог Вид совпадает; участок доминирует аналог
Аналог	308	Аналог Вид отличается; участок находится рядом аналог
Аналог	230	Аналог Вид отличается; участок находится рядом аналог
D	559	вид отличается; общая площадь не указана
E	502	вид совпадает; состояние отличается
F	422	вид и состояние отличаются
G	3,000	вид отличается; участок находится далеко
H	224	вид отличается; ближе всех

Если судить по чистой дистанции, Comp H является ближайшим соседом. Однако у Comp H нет вида из окна, а у объекта оценки он есть. Наивная метрика ввела нас в заблуждение: числовой масштаб параметра «размер участка» затмил категориальный параметр «вид из окна», который на самом деле имеет большее значение для цены.

Обоснованная метрика расстояния должна масштабировать каждую ось в соответствии с её неявной ценой — грубо говоря, в зависимости от того, сколько долларов стоит одна единица изменения по этой оси. Как только оси масштабируются с учётом неявной цены, вид (стоимостью ~\$100,000) перестаёт доминировать над участком (~\$ 25 за квадратный фут). Объект A, который соответствует объекту оценки по площади, таким образом становится подходящим основным аналогом.

Именно это и обеспечивает механизм неявных цен, разработанный в следующих записях: принципиальное масштабирование пространства характеристик, благодаря которому «близость» превращается из геометрического предположения в вычисление, основанное на рыночных данных.

Список литературы.

- Ланкастер, К. Дж. (1966). Новый подход к теории потребителя. *Journal of Political Economy* , 74(2), 132–157.
- Розен, С. (1974). Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции. *«Журнал политической экономики»* , 82(1), 34–55.

Перекрытые ссылки: heterogeneous good; hedonic price function; implicit price.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (002-characteristics-space.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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).
```

Гедоническая функция цены

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/9x1a0489](https://doi.org/10.67330/9x1a0489)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. Теперь у нас есть два компонента: неоднородное благо (набор характеристик) и пространство характеристик (векторное пространство, в котором каждый набор является точкой). Гедоническая функция цены объединяет их. Это функция $p(\mathbf{z})$, которая присваивает цену каждой точке $\mathbf{z}$ в пространстве характеристик, указывая, сколько рынок платит за набор с данной конкретной комбинацией характеристик.

Ключевой вывод Розена (1974) заключается в том, что функция « $p(\mathbf{z})$ » не определяется каким-либо отдельным покупателем или продавцом. Она возникает в результате рыночного равновесия. Каждый покупатель имеет *функцию ставки*, отражающую максимальную сумму, которую он готов заплатить за любой набор товаров, с учетом своих предпочтений и дохода. У каждого продавца есть *функция предложения*, отражающая минимальную цену, которую он готов принять с учётом своих издержек и альтернативных вариантов. Гедоническая функция цены представляет собой оболочку этих предложений и спросов — границу, вдоль которой покупатели и продавцы фактически заключают сделки.

В сфере недвижимости « $p(\mathbf{z})$ » — это ответ на вопрос: если бы вы указали полный набор характеристик, какую цену установил бы рынок за него? Этот ответ не является произвольным. Он отражает совокупные предпочтения всех покупателей на рынке, а также совокупные издержки и альтернативы всех продавцов, опосредованные равновесным соотношением между покупателями и наборами характеристик.

Гедоническая функция цены является центральным теоретическим объектом гедонического анализа. Неявные цены представляют собой её частные производные. Гедоническая регрессия — это её эмпирическая оценка. Обоснованность любой корректировки на сопоставимость продаж в конечном счёте сводится к утверждению, касающемуся этой функции.

Где оценщик встречается с этим понятием. Оценщики постоянно оперируют концепцией « $p(\mathbf{z})$ », причем, как правило, не называя её прямо. Ряд практических действий представляет собой неявные манипуляции с гедонической функцией цены:

Корректировки в сравнительном подходе к оценке объектов, аналогичных по продажам. Когда оценщик корректирует стоимость объекта «Comp В» в сторону повышения на \$ 100,000, чтобы учесть тот факт, что объект оценки имеет вид из окна, а объект сравнения В — нет, оценщик выдвигает утверждение о $p(\mathbf{z})$: а именно, что при неизменности всех остальных характеристик функция увеличивается на \$ 100,000 по мере того, как характеристика «вид из окна» изменяется от 0 до 1. Это утверждение о локальном наклоне p в измерении «вид».

Согласование подходов. Затратный подход позволяет определить ориентировочную стоимость «снизу», суммируя стоимость компонентов воспроизводимого набора. Доходный подход определяет стоимость на основе денежных потоков, которые в состоянии равновесия отражают

цену наборов, генерирующих аналогичные денежные потоки. Сравнительный подход позволяет получить цену непосредственно из данных о совершенных сделках с аналогичными наборами. На хорошо функционирующем рынке все три подхода должны сходиться в одной точке на $p(\mathbf{z})$. Когда они расходятся, оценщик должен определить, какой подход наиболее достоверно оценил функцию для объекта оценки с учётом его положения в пространстве характеристик.

Анализ рынка и корректировки с учетом времени. Гедоническая функция цены смещается со временем по мере изменения спроса, предложения, условий финансирования и ожиданий. Годовая ставка роста стоимости в 6 процентов представляет собой утверждение о том, как в целом сместилась функция « $p(\mathbf{z})$ »; вывод о том, что стоимость объектов недвижимости с видом растёт быстрее, чем объектов без вида, является утверждением о том, что функция сместилась неравномерно по пространству характеристик.

Почему это важно для обоснованности. Явное обозначение гедонической функции цены позволяет прояснить ряд основополагающих вопросов, касающихся обоснованности этой концепции:

Существует ли такая функция? Модель Розена предполагает наличие конкурентного рынка с достаточным количеством сделок, при котором функция равновесия имеет смысл. На рынках с низким объемом торгов, на быстро меняющихся рынках или на субрынках с нетипичными группами покупателей эта функция может быть плохо определена или нестабильна. Оценщик, работающий на таком рынке, должен осознавать, что механизм вывода основан на неустойчивом теоретическом фундаменте.

Является ли эта функция одинаковой везде? Модель Розена дает по одной функции на каждый сегмент рынка, а не одну общую функцию. Пасифика и Сан-Франциско — это отдельные рынки с разными гедоническими функциями; односемейные дома и кондоминиумы представляют собой отдельные субрынки даже в пределах одного города. Оценщик, смешивающий сегменты, неявно предполагает, что одна функция охватывает оба сегмента, что, как правило, неверно.

Какую функциональную форму принимает функция « p »? Эта функция не обязательно должна быть линейной, логарифмически-линейной или иметь какую-либо другую конкретную форму. Розен показал лишь то, что p является огибающей ценовых предложений и спроса; он не уточнял её форму. Эмпирически функция « p » демонстрирует нелинейности (предельная стоимость дополнительной спальни часто не является постоянной в зависимости от общей жилой площади), взаимодействия (надбавка за вид из окна зависит от ориентации участка) и пороговые значения (кухня, не достигающая определенного уровня качества, может привести к скидке, несоразмерной её оценке балла состояния). Обоснованная оценка требует либо функциональной формы, подтверждаемой рынком, либо гибкого оценщика, способного восстановить форму на основе данных.

Разобранный пример оценки. Для исследуемого района Пасифика мы будем работать с простой аддитивной функциональной формой, которая будет более подробно рассмотрена в следующей записи:

$$p(\mathbf{z}) = \beta_0 + \beta_{\text{GLA}} z_{\text{GLA}} + \beta_{\text{lot}} z_{\text{lot}} + \beta_{\text{view}} z_{\text{view}} + \beta_{\text{cond}} z_{\text{cond}}.$$

Эта форма основана на двух сильных допущениях: что вклад каждой характеристики линейно зависит от её уровня, и что характеристики не взаимодействуют между собой. Оба допущения поддаются проверке и часто нарушаются; здесь мы принимаем их в качестве отправной точки.

Для восьми аналогов, приведённых в примере, рассматриваемом в этом выпуске, в ходе нашего анализа, который будет представлен в следующих записях, мы оценим значения β , при которых функция будет выглядеть примерно так:

Параметр	Оценочное значение
β_0 (пересечение)	\$ 80,777
β_{GLA} (на кв. фут)	\$ 575
β_{lot} (на кв. фут)	\$ 22,70
β_{view} (да/нет)	\$ 86,179
β_{cond} (на ступень)	\$ 49,824

При подстановке в аддитивную форму эту функцию позволяет спрогнозировать цену для любой точки в четырёхмерном пространстве характеристик. Подстановка координат объекта Comp A даёт прогнозируемую цену, близкую к его наблюдаемому значению \$ 1,425,000; подстановка гипотетического объекта по адресу (1700, 7500, 1, 2) даёт оценку стоимости данного объекта недвижимости.

Суть этой статьи заключается не в этих конкретных цифрах, а в концептуальном подходе: *функция существует, в принципе, ещё до того, как мы её оцениваем*. Функция является свойством рынка. Наша оценка — это попытка восстановить её. Восстановление может быть частичным, смещённым или содержать шумы, и следующие три записи посвящены тому, как проводить такое восстановление ответственно.

Ссылки.

Корт, А. Т. (1939). Гедонические ценовые индексы на примере автомобильной промышленности. В сборнике «Динамика спроса на автомобили». General Motors Corporation.

Розен, С. (1974). «Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции». *Journal of Political Economy*, 82(1), 34–55.

Трипплетт, Дж. Э. (2006). *Справочник по гедоническим индексам и корректировкам качества в индексах цен*. Издательство ОЭСР.

Перекрёстные ссылки: heterogeneous good; characteristics space; implicit price; hedonic regression.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (003-hedonic-price-function.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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.").
defensibility_question(segment_specificity,
    "One function per market segment, not one function globally. Conflating segments is usually
    ⇨ wrong.").
defensibility_question(functional_form,
    "Rosen showed only that p is an envelope, not its shape. Empirically nonlinear with
    ⇨ interactions and thresholds; estimation requires defending the form.").

% =====
% WORKED EXAMPLE: ADDITIVE LINEAR SPECIFICATION
% =====
%
% For the Pacifica neighborhood, the article adopts the additive linear
% specification:
% 
$$p(z) = \beta_0 + \beta_{\text{gla}} * z_{\text{gla}} + \beta_{\text{lot}} * z_{\text{lot}} + \beta_{\text{view}} * z_{\text{view}} + \beta_{\text{cond}} * z_{\text{cond}}$$

%
% The next entry (004) takes the coefficients as implicit prices; entry

```

```

% 005 estimates them by hedonic regression.

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

specification_form(additive_linear).
specification_assumption(linearity_in_levels,
    "Contribution of each characteristic is linear in its level.").
specification_assumption(no_interactions,
    "Characteristics do not interact in the price function.").
specification_caveat(
    "Both assumptions are testable and frequently violated; adopted here as a starting point.").

% Coefficient values reported in the article (estimated in entry 005,
% asserted here so this entry's worked example can be queried).
coefficient(intercept,      80777.07).
coefficient(beta_gla,       574.93).
coefficient(beta_lot,       22.70).
coefficient(beta_view,      86179.03).
coefficient(beta_cond,      49824.33).

%! predicted_price(?Property, -Predicted) is det.
%
% The fitted additive-linear hedonic price function applied to a
% property's characteristics. Symbolic view/cond are mapped to {0,1}
% and {1,2,3} respectively.

predicted_price(P, Predicted) :-
    coord(P, gla_sqft, GLA),
    coord(P, lot_sqft, Lot),
    coord(P, view, View),
    coord(P, cond_numeric, Cond),
    coefficient(intercept, B0),
    coefficient(beta_gla, Bg),
    coefficient(beta_lot, B1),
    coefficient(beta_view, Bv),
    coefficient(beta_cond, Bc),
    Predicted is B0 + Bg * GLA + B1 * Lot + Bv * View + Bc * Cond.

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
    ).

% =====
% 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).
```

Неявная цена

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/md6wke17](https://doi.org/10.67330/md6wke17)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. Гедоническая функция цены $p(\mathbf{z})$ присваивает цену каждому набору. Однако большинство практических вопросов в области оценки касаются не цены полного набора, а цены конкретного *компонента* набора. Сколько стоит вид из окна? На сколько увеличивает цену дополнительная сотня квадратных футов? Насколько влияет разница между «хорошим» и «отличным» состоянием?

На эти вопросы отвечает понятие «неявной цены». Этот термин был введен Розеном (1974), который показал, что наблюдаемые цены дифференцированных товаров в совокупности с их характеристиками «определяют набор неявных или «гедонических» цен»; Ланкастер (1966) сформулировал лежащую в основе эту концепцию, согласно которой товары представляют собой набор характеристик. Для каждой характеристики z_i неявная цена — это скорость, с которой $p(\mathbf{z})$ изменяется при изменении z_i при фиксированных значениях всех остальных характеристик:

$$\pi_i(\mathbf{z}) = \frac{\partial p(\mathbf{z})}{\partial z_i}.$$

Неявная цена измеряется в долларах на единицу характеристики — долларах на квадратный фут общей арендуемой площади (GLA), долларах на квадратный фут участка, долларах на единицу шкалы состояния. Для бинарной характеристики, такой как вид из окна, неявная цена представляет собой полную надбавку, уплачиваемую за переход от отсутствия вида (0) к наличию вида (1).

Слово «неявный» отражает ключевой факт: эти цены *не* объявляются. Никто не предлагает продать «один квадратный фут общей арендуемой площади» за \$ 575. Цена неявно заложена в объявленных ценах на равновесные наборы товаров и услуг, а гедонический подход служит для её вычисления.

Неявные цены — это выводы модели гедонистического анализа, имеющие значение для оценки. Корректировка в таблице сравнительных продаж, по сути, представляет собой оценку неявной цены, умноженную на разницу в характеристиках. Каждая корректировка — это предположение о значении π_i в соответствующей точке пространства характеристик.

Где оценщик встречается с этим понятием. Таблица корректировок на основе сравнения продаж представляет собой таблицу применения неявных цен, даже если оценщики не называют её именно так. Запись в таблице, которая проводит корректировку аналога на «+\$ 10,000 за 200 sf меньшую GLA, чем у объекта оценки», утверждает, что неявная цена GLA в соответствующей точке пространства характеристик составляет \$ 50 за квадратный фут. Запись в

таблице, корректирующая показатель на «\$ 100,000 за худший вид», утверждает, что неявная цена за характеристику «вид» составляет \$ 100,000. Вся обоснованность сравнительного подхода к анализу продаж зависит от того, можно ли подтвердить эти неявные цены на рынке. Однако на практике оценщики рассчитывают корректировки напрямую — на основе сопоставимых пар, опыта или практики коллег — без какого-либо вычисления вкладов в стоимость; в традиционной таблице фиксируются только сами корректировки. Интерпретация неявных цен представляет собой толкование того, что утверждают эти корректировки, а не описание того, как они выводятся.

Такое обозначение также поясняет, почему анализ парных продаж является основополагающим методом вывода корректировок. Два аналога, которые отличаются ровно по одной характеристике, дают прямую оценку неявной цены этой характеристики: разница в цене, деленная на разницу в характеристиках. Надежность анализа парных продаж зависит от того, действительно ли два аналога достаточно близки в пространстве характеристик, чтобы неявная цена, вычисленная на их основе, могла быть обобщена на объект оценки.

Три свойства неявных цен, имеющие практическое значение:

Они носят локальный характер. Неявная цена общей арендуемой площади (GLA) при площади 1,650-sf не обязательно совпадает с ценой при площади 2,500-sf. Гедоническая функция цены, как правило, нелинейна, поэтому её частные производные зависят от местоположения в пространстве характеристик. Корректировка, полученная на основе анализа парных продаж для площади 1,500-sf, не обязательно применима к объекту оценки площадью 3,000-sf.

Они носят условный характер. Неявная цена вида в субрынке недвижимости на побережье отличается от цены в субрынке, расположенном в глубине суши. Частная производная вычисляется при фиксированных остальных характеристиках, однако значение этой производной зависит от того, где именно фиксируются остальные характеристики.

Они зависят от рынка. Неявные цены — это равновесные цены, поэтому они отражают предпочтения совокупности покупателей на момент совершения сделок. Изменение рыночной конъюнктуры может повлиять на неявные цены, даже если сами характеристики остаются неизменными. Корректировки с учетом времени — это неявное признание того, что вчерашние неявные цены могут не соответствовать сегодняшним.

Почему это важно для обоснованности. Фрейминг с использованием неявных цен выявляет вопросы об обоснованности корректировок, которые могут оставаться незамеченными при использовании неформальной таблицы:

Где вы находитесь в пространстве характеристик? Неявная цена обоснована только в том случае, если она оценивается вблизи местоположения объекта оценки в пространстве характеристик. Надбавка за вид, рассчитанная на основе аналогов, расположенных вдали от побережья, и примененная к объекту на берегу моря, представляет собой экстраполяцию на область пространства характеристик, в которой поведение функции неизвестно.

Какова локальная геометрия? Если две характеристики взаимодействуют — например, если вид из окна ценится выше на более крупном участке — то рассмотрение их отдельно как аддитивных величин является неверной спецификацией. Неявная цена вида из окна зависит от площади участка, и корректировка, игнорирующая это взаимодействие, приводит к недооценке фактической структуры функции.

Как была восстановлена неявная цена? Анализ парных продаж, коэффициент регрессии, отраслевое эмпирическое правило и «профессиональное суждение» — все они указывают на одну и ту же величину — неявную цену — но с совершенно разным эпистемическим статусом. Обоснованность оценщика зависит от того, какой источник послужил основой для неявной цены, использованной при каждой корректировке.

Разобранный пример оценки. Для примера модели «Pacifica» с аддитивной спецификацией, принятой в предыдущем записи, неявные цены представляют собой просто коэффициенты линейной функции:

$$\pi_{\text{GLA}} = \beta_{\text{GLA}}, \quad \pi_{\text{lot}} = \beta_{\text{lot}}, \quad \pi_{\text{view}} = \beta_{\text{view}}, \quad \pi_{\text{cond}} = \beta_{\text{cond}}.$$

С использованием коэффициентов, оцененных в записи 003:

Характеристика	Неявная цена	Единицы измерения
Общая арендуемая площадь (GLA)	\$ 575	за кв. фут
Участок	\$ 22,70	за кв. фут
Вид из окна	\$ 86,179	за «да/нет»
Состояние	\$ 49,824	за ступень

Эти неявные цены предполагают анализ парных продаж. Вернёмся к разнице между «Comp A» и « » «Comp B» из записи № 001:

- Конкурс A имеет страницу ($\Delta z_{\text{view}} = +1$); вклад +\$86,179.
- У дома A участок больше на 1,800 квадратных футов ($\Delta z_{\text{lot}} = +1,800$); contribution $1,800 \times \$22,70 = +\$40,860$.
- Все остальные характеристики идентичны.
- Прогнозируемая разница в цене: $\$86,179 + \$40,860 = \$127,039$.

Наблюдаемая разница в цене составила \$ 145,000. На долю декомпозиции по неявным ценам приходится около \$ 127,000; оставшиеся \$18,000 отражают извлеченную из выборки изменчивость в небольшом наборе данных (и, в более общем плане, те виды ненаблюдаемых факторов, которые будут учтены в слагаемом латентной переменной).

Следует отметить два практических момента. Во-первых, допущение о линейности играет реальную роль: при нелинейном $p(\mathbf{z})$ одна и та же физическая разница между участками может иметь большую ценность для небольшого участка, чем для большого, и приведенное выше аддитивное разложение будет искаженным. Во-вторых, неявная цена \$ 22,70 за квадратный фут участка представляет собой предельную стоимость дополнительного квадратного фута, а не общую стоимость участка. Общий вклад участка площадью 9,000 квадратных футов в цену набора составляет $9,000 \times \$22,70 = \$204,300$, но добавление ещё одного квадратного фута повышает цену набора всего на \$ 22,70.

Это различие — предельная и общая стоимость () — является источником значительной путаницы в повседневном языке оценки и представляет собой одно из наиболее важных концептуальных достижений, связанных с явным обозначением неявной цены.

Список литературы.

- Розен, С. (1974). «Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции». *Journal of Political Economy*, 82(1), 34–55.
- Ланкастер, К. Дж. (1966). Новый подход к теории потребителя. *«Журнал политической экономики»*, 74(2), 132–157.
- Фриман, А. М., III (1979). Гедонические цены, стоимость недвижимости и оценка экологических выгод: обзор проблем. *Scandinavian Journal of Economics*, 81(2), 154–173.
- Тейлор, Л. О. (2003). Гедонический метод. В *«Введение в нерыночную оценку»* (с. 331–393). Kluwer Academic Publishers.

Палмквист, Р. Б. (2005). Модели стоимости недвижимости. В «Справочнике по экологической экономике» (Т. 2, с. 763–819). North-Holland.

Трипплетт, Дж. Э. (2006). *Справочник по гедоническим индексам и корректировкам на качество в ценовых индексах*. Издательство ОЭСР.

Перекрёстные ссылки: heterogeneous good; characteristics space; hedonic price function; hedonic regression.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (004-implicit-price.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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).

```

Гедоническая регрессия

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

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

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. В предыдущих разделах гедоническая функция цены и её неявные цены рассматривались как теоретические объекты. Эта функция существует в принципе, но на практике мы наблюдаем лишь конечное множество сделок: торгуемые наборы товаров и цены, по которым они торгуются. Гедоническая регрессия служит эмпирическим мостом между этими двумя аспектами.

Схема проста. Каждая сделка предоставляет пару « $(\mathbf{z}_i, p_i)$ »: вектор характеристик торгуемого набора и его цену продажи. Совокупность этих пар представляет собой выборку из (ненаблюдаемой) гедонической функции цены. В ходе регрессионного анализа к этой выборке подбирается функциональная форма, при этом оцениваются параметры, характеризующие функцию. После оценки параметров частные производные подгоняемой функции позволяют получить оценки неявных цен.

Прежде чем запустить регрессионный анализ, необходимо выбрать один из трёх вариантов моделирования:

1. *Какие характеристики включаются в модель?* Этот выбор определяет размерность оцениваемой функции. Исключение значимой характеристики приводит к тому, что её влияние переносится на коэффициенты коррелирующих включённых характеристик (систематическая ошибка, связанная с пропущенной переменной).
2. *Какую функциональную форму выбрать?* Линейную, логарифмически-линейную, логарифмически-логарифмическую, полиномиальную, с расширенной базой, полупараметрическую, основанную на машинном обучении. Этот выбор ограничивает форму, которую может принимать оцениваемая функция, и какие закономерности неявных цен она может воспроизвести.
3. *Какой оценщик?* OLS, обобщённые наименьшие квадраты, робастные M-оценщики, квантильная регрессия, пространственные эконометрические оценщики, регрессия с штрафными терминами. Выбор определяет, что именно оптимизирует оценщик и какие допущения лежат в основе его выводов.

Каждый выбор имеет методологическое значение и влечет за собой определенные последствия с точки зрения обоснованности.

Где оценщик встречается с этим понятием. Гедоническая регрессия лежит в основе массовой оценки и является движущей силой большинства автоматизированных моделей оценки (AVM). Когда оценщик округа ежегодно переоценивает 100,000 земельных участков, эту работу выполняет гедоническая регрессия на основе недавних сделок купли-продажи. Когда Zillow

или Redfin выдают автоматизированную оценку, за этим стоит модель гедонического типа (часто представляющая собой ансамбль деревьев, а не OLS, но концептуально все равно являющаяся гедоническим оценщиком). Когда AVM кредитора на вторичном рынке одобряет или отклоняет объект недвижимости для рефинансирования, задействуется тот же механизм.

При оценке отдельных объектов недвижимости всё чаще используется и гедоническая регрессия, даже если окончательная оценочная стоимость определяется на основе сравнения с аналогичными сделками. Неявные цены, рассчитанные на основе регрессии по выборке объектов в данном районе, позволяют обосновать величину корректировок гораздо более обоснованно, чем интуиция оценщика или процентные значения, основанные на эмпирических правилах. Коэффициенты регрессии становятся доказательством, а не мнением, при применении корректировок в таблице оценок.

Почему это важно для обоснованности. Оценки, полученные с помощью регрессионного анализа, отличаются по степени обоснованности от анализа парных продаж. Преимущества: регрессионный анализ использует все имеющиеся данные о продажах, а не по две единицы за раз; он позволяет выделить влияние одной характеристики, одновременно контролируя множество других; кроме того, он даёт стандартные ошибки, которые количественно оценивают неопределённость каждой неявной цены. Недостатки: регрессия навязывает функциональную форму, которая может не соответствовать истинной ценовой функции; она требует большего объема данных, чем анализ парных продаж; и её результаты могут быть чувствительны к выбору выборки и спецификации.

При оценке на основе регрессионного анализа неоднократно возникают следующие вопросы, касающиеся обоснованности:

Является ли выборка адекватной? Регрессия оценивает гедоническую функцию в том виде, в каком она действует в рыночном сегменте и периоде времени, охватываемых выборкой. Применение регрессии, построенной на данных о продажах недвижимости, расположенной в глубине суши, к объекту оценки, расположенному на берегу моря, является методологической ошибкой, независимо от того, насколько обоснованы статистические результаты.

Является ли данная спецификация уместной? Линейная спецификация — это серьёзное допущение. Если реальная функция демонстрирует убывающую доходность от GLA выше некоторого порогового значения, линейная подгонка приведёт к неверной оценке неявных цен в этом диапазоне. Диагностика, графики остатков и альтернативные спецификации являются частью обоснованного рабочего процесса.

Что означает стандартная ошибка? Стандартная ошибка отражает вариабельность выборки при принятой спецификации модели. Она не учитывает неверную спецификацию модели, систематическую ошибку, связанную с пропущенными переменными, или отбор выборки. Небольшая стандартная ошибка при неверной спецификации модели более опасна, чем большая стандартная ошибка при правильной спецификации.

Совпадает ли полученный результат с результатами других подходов? Обоснованная обоснованная неявная цена, рассчитанная на основе регрессионного анализа, должна согласовываться с выводами анализа парных сделок, с данными о стоимости компонентов, полученными по затратному подходу, а также с мнением оценщиков, хорошо знакомых с рынком. Регрессионный коэффициент, противоречащий всем трем факторам, является поводом для дополнительного изучения, а не для публикации.

Разобраный пример оценки. Теперь мы можем применить полный набор методов гедонической регрессии к восьми аналогам Pacifica. Набор данных доступен по адресу `shared/data/pacifica_comps`.

Код на R:

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

Вывод:

```
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
```

Интерпретация. Модель, подгнанная для описания гедонической функции цены, описывает ее следующим образом:

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

Каждый коэффициент представляет собой расчетную неявную цену соответствующей характеристики. Показатель GLA вносит вклад в размере \$ 575 за квадратный фут (с наименьшей стандартной ошибкой и наибольшей статистической значимостью); вид вносит вклад в размере около \$ 86,000 в стоимость объекта недвижимости в Пасифике; дополнительная единица оценки состояния (например, переход от «хорошего» к «отличному») стоит примерно \$ 50,000.

Сразу бросаются в глаза несколько моментов, на которые следует обратить внимание:

Выборка очень мала. Восемь наблюдений и четыре предикторы оставляют три степени свободы остатков. Стандартные ошибки соответственно велики, и значения p для «лота», «вида» и «состояния» не достигают принятого порога в 0,05, несмотря на их экономическую обоснованность. Это следствие небольшого объема выборки, а не доказательство того, что эти характеристики не имеют значения. При наличии 80 аналогов, а не 8, значения p значительно уменьшились бы.

Общая подгонка модели высокая. Скорректированный коэффициент $R^2 = 0.93$ и значение p в 0,012 указывают на то, что включенные характеристики в совокупности объясняют большую часть колебаний цен в выборке. Сочетание сильной совокупной подгонки со слабой индивидуальной значимостью является классическим признаком мультиколлинеарности, которая в данном случае носит структурный характер: GLA, участок и состояние объекта являются ковариатами в выборке.

Один остаток выделяется. Максимальный остаток составляет +\$46,455, что соответствует компании Н. Остальные семь остатков находятся в диапазоне от $\pm \$18,000$. Что-то в компании

Comp H не учитывается функцией подгонки — модель систематически занижает её цену на величину, почти в три раза превышающую следующий самый большой остаток.

Это отклонение от среднего значения не является шумом. Оно свидетельствует о том, что ненаблюдаемая характеристика Comp H влияет на его цену таким образом, который не может быть учтён с помощью включённых в модель переменных. В разделе, посвящённом латентным переменным, рассматривается именно этот вопрос.

Проверка на соответствие логике парных продаж. Анализ парных продаж, приведённый в записи № 004, предсказывал разницу в \$ 127,039 между аналогами А и В, исходя исключительно из данных о просмотрах и лотах. Неявные цены в регрессии идентичны (поскольку они являются источником данного расчёта). Регрессия обобщает логику парных продаж на все восемь аналогов одновременно, но при данном размере выборки не даёт ничего принципиально нового; она подтверждает результаты парного анализа, одновременно предоставляя отчёт о стандартных ошибках, чего не может сделать метод парных продаж.

Список литературы.

Корт, А. Т. (1939). «Гедонические ценовые индексы на примере автомобилей». В сборнике «Динамика спроса на автомобили». General Motors Corporation.

Гриличес, З. (1961). Гедонические ценовые индексы для автомобилей: эконометрический анализ изменения качества. В сборнике «Статистика цен федерального правительства». Национальное бюро экономических исследований.

Розен, С. (1974). Гедонические цены и неявные рынки: дифференциация продукции в условиях чистой конкуренции. «*Journal of Political Economy*», 82(1), 34–55.

Трипплетт, Дж. Э. (2006). *Справочник по гедоническим индексам и корректировкам качества в ценовых индексах*. Издательство ОЭСР.

Перекрёстные ссылки: heterogeneous good; characteristics space; hedonic price function; implicit price; latent variable.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (005-hedonic-regression.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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).
```

Латентная переменная

Foundations • The Valuation Engineer

Берт Крейтор • 8 августа 2026 г.

Версия 1.0.0

DOI: [10.67330/11n3g168](https://doi.org/10.67330/11n3g168)

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Формальное определение. 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.

Интуитивное представление. В первых пяти статьях вектор характеристик рассматривался так, как будто его можно полностью измерить: общая площадь, участок, вид из окна, состояние и т. д. — все эти характеристики представлены в виде столбцов данных. На практике ни один вектор характеристик никогда не бывает полным. Некоторые характеристики отсутствуют в данных, поскольку их сложно измерить (качество внутренней отделки, планировка, микролокация и микроориентация), либо отсутствуют в стандартных источниках данных (записи MLS не отражают качество освещения, звукоизоляцию или недавние несанкционированные улучшения), либо потому что они ещё не были признаны значимыми (особенности района, которые, как оказывается, позволяют прогнозировать цену только после того, как их выявляет анализ).

Эти ненаблюдаемые, но значимые характеристики являются *латентными переменными*. Они существуют, влияют на цену, но не представлены в столбцах нашей регрессии. Их влияние не исчезает из-за того, что они не измеряются; оно накапливается в остатках модели, где с точки зрения включённых переменных оно выглядит как «шум», но на самом деле представляет собой структурированный сигнал с точки зрения лежащего в основе процесса генерации данных.

Латентные переменные служат связующим звеном между идеализированной гедонической моделью и неоднозначной реальностью данных оценки. Их явное учёт является необходимым условием для объективного анализа неопределённости модели, структуры остаточных величин и пределов того, что можно объяснить с помощью наблюдаемых характеристик.

Где оценщик встречается с этим понятием. Оценщики сталкиваются с латентными переменными в трёх типичных ситуациях.

Корректировки по качеству и состоянию. «Качество» и «состояние» как категории оценки сами по себе представляют собой попытку свести многомерный латентный конструкт к единой порядковой шкале. Оценка состояния СЗ в UAD 3.6 призвана обобщить отложенный ремонт недвижимости, износ отделки, возраст инженерных систем и внешний вид — от четырёх до десяти базовых характеристик, сведённых к одному числу. Это одноцифровое обобщение является прокси-показателем скрытого конструкта; мастерство оценщика заключается в том, насколько точно этот прокси-показатель отражает суть дела в каждом конкретном случае.

Необъяснимое отклонение в таблице. Когда три скорректированных аналога указывают на стоимость, но их значения расходятся больше, чем ожидалось, это расхождение несёт информационную ценность. Редко когда это полностью объясняется выборочным шумом; чаще всего оно отражает какую-то характеристику, которую оценщик ещё не выявил или не учёл при корректировке. Исследование причины расхождения остаточных значений — «почему этот

аналог указывает на цену \$ 50,000 выше, чем остальные?» — является неформальной версией вывода о латентных переменных, применяемой оценщиком.

Недавнее улучшение, которое никто не может увидеть. В объекте недвижимости могли быть проведены ремонт кухни, замена крыши или построенное с разрешением дополнительное жилое помещение (ADU), что пока не отражено в данных MLS или оценщика. С точки зрения оценщика, это улучшение можно увидеть при личном осмотре, но оно отсутствует в данных. С точки зрения регрессионного анализа, это улучшение является скрытым. Оценщик должен сопоставить результаты личного наблюдения с показателями, основанными на данных, и оба источника должны сходиться в оценке истинного вклада этой скрытой характеристики.

Почему это важно для обоснованности. Латентные переменные вводят ряд соображений относительно обоснованности, которые аддитивная линейная модель, приведенная в записи 005, напрямую не учитывает.

Систематическая ошибка, связанная с пропущенной переменной. Если латентная переменная коррелирует с наблюдаемой переменной, её влияние проникает в коэффициент наблюдаемой переменной. Если недавно проданные дома в Пасифике, как правило, имеют отремонтированные кухни (скрытая характеристика), а дата заключения договора купли-продажи дома присутствует в данных, то коэффициент срока продажи (также известный как «*sale_age*» или «*days since sold*») будет включать в себя часть эффекта, связанного с ремонтом кухни. Оцененный неявный вклад цены «*sale_age*» становится смещенным — он отражает «*sale_age*» *плюс* часть неизмеренного статуса ремонта кухни. Обоснованная оценка учитывает, что коэффициенты наблюдаемых переменных зависят от того, что отсутствует в модели.

Интерпретация остатков. В стандартной теории регрессии остатки рассматриваются как гауссовый шум с нулевым средним. В подходе, основанном на латентных переменных, остатки рассматриваются как смесь шума и сигнала: часть представляет собой вариацию выборки, а часть — совокупный эффект ненаблюдаемых характеристик. Разграничение того и другого имеет существенное методологическое значение. Остаток, коррелирующий с местоположением, датой продажи, личностью оценщика или любой другой наблюдаемой переменной, указывает на то, что соответствующая латентная переменная можно частично восстановить.

Выбор модели. Добавление дополнительных наблюдаемых переменных — это очевидное средство борьбы с систематической ошибкой, связанной с пропущенными переменными, однако латентные переменные, имеющие наибольшее значение, зачастую являются именно теми, которые сложнее всего включить в модель (условия мелкомасштабного уровня, качество недавних улучшений, микроэффекты района). Иногда обоснованным решением является признание того, что модель никогда не сможет их учесть, и отнесение необъяснимой дисперсии к остатку, а не принудительная подгонка дополнительных переменных.

Разобранный пример оценки. Регрессия в записи 005 дала необычный остаток для Compr N: модель занизила прогноз цены Compr N на \$ 46,455, в то время как остальные семь остатков уложились в диапазон от $\pm \$18,000$. Наблюдаемые характеристики объекта Compr N — общая арендуемая площадь (GLA) 1,600, участок 7,300, без вида, состояние «хорошее» — прочно помещают его в середину выборки; ничто в его наблюдаемом профиле не предсказывает необычно высокую цену. И тем не менее он был продан за \$ 46,455 больше, чем, согласно модели, он должен был бы стоить.

Естественным предположением является то, что у компании N имеется скрытая характеристика, которая влияет на её цену, но не отражается в данных. При формировании набора данных для данного вопроса ситуация именно такова: в компании Compr N недавно был проведен ремонт кухни, который добавил примерно \$60,000 к стоимости набора услуг. Ремонт не входил в список переменных регрессии, поэтому этот вклад накоплен в остатке, а не в коэффи-

циенте. Оцененное значение остатка (\$ 46,455) немного меньше, чем истинный скрытый вклад (\$ 60,000), поскольку регрессия распределяет небольшие доли скрытого эффекта по другим коэффициентам, смещая их в сторону увеличения в коррелирующих измерениях.

Из этого вытекают три практические последствия:

Рассматривайте остаток, выпадающий за пределы диапазона, как доказательство, а не как ошибку. Первым надлежащим ответом на остаток по Compr H является не исключение его как выпадающего значения, а выяснение, что именно в данном объекте недвижимости объясняет недооценку. В данном случае очный осмотр или поиск в архивах разрешений по строительству позволили бы выявить факт реконструкции. Несостоятельность модели имеет ценность для диагностики.

Следует учитывать, что остальные коэффициенты имеют небольшую систематическую погрешность. Коэффициенты «вид», «состояние» и «участок» поглотили небольшую часть эффекта реконструкции за счет корреляций с наблюдаемыми характеристиками объекта сравнения H. Эта погрешность невелика на сайте $n = 8$ и уменьшилась бы еще больше при увеличении количества аналогов и ковариат, но она не равна нулю.

Остатки в целом обладают восстанавливаемой структурой. В этом и заключается основной тезис методологии «Метода остаточных ограничений» (RCA), разработанной автором (Craytor, 2025): в данных по оценке жилой недвижимости остатки систематически не являются независимым и идентично распределенным гауссовым шумом. Они отражают структурированный вклад ненаблюдаемых характеристик — мелкозернистого состояния, мелкозернистого качества, функциональной полезности и других латентных конструкторов — которые можно частично восстановить, рассматривая само остаток как сигнал латентной переменной, подлежащий моделированию, а не отбрасыванию.

Подход RCA к оценке состояния и качества делает эту двухуровневую структуру применимой на практике. Порядковые оценки оценщиков — те упрощённые прокси-показатели, о которых говорилось выше и которые можно обозначить как «состояние-1» (состояние-грубое) и качество-1 (качество-грубое) — настолько приблизительны, что их вполне можно рассматривать как латентные переменные сами по себе, но их всё же можно ввести в качестве регрессоров для получения приблизительных вкладов неявной цены. Затем протокол перемещает эти оцененные вклады из измеренных характеристик в разделе регрессии таблицы продаж обратно в остаток, откуда они перераспределяются в тонкозернистые латентные переменные состояние-2 (состояние-тонкозернистое) и качество-2 (качество-тонкозернистое) и уточняются с учетом дополнительных данных, таких как публичные комментарии в системе MLS. Рейтинг грубой детализации обеспечивает начальное присвоение значений; латентная переменная тонкой детализации несет в себе восстанавливаемую структуру. В конечном итоге вклад значений тонкой детализации заменяет вклад значений грубой детализации.

Приведенная здесь иллюстрация с использованием одной компоненты представляет собой простейший случай интуиции RCA. При наличии большего объема данных и более сложной модели структуру остаточной величины можно разбить на распознаваемые скрытые факторы, каждый из которых вносит вклад в рыночную цену набора ценных бумаг таким образом, который недоступен для стандартной гедонической спецификации. Латентные переменные — это не помеха, которую нужно устранить; они представляют собой новую границу обоснованного моделирования оценки.

Список литературы.

Спирман, К. (1904). «Общий интеллект»: объективное определение и измерение. *American Journal of Psychology*, 15(2), 201–292.

Боллен, К. А. (2002). Латентные переменные в психологии и социальных науках. *Ежегодный обзор психологии*, 53, 605–634.

Крейтор, Б. (2025). Метод остаточных ограничений: структура& Протокол. Zenodo.[doi:10.5281/zenodo.14787917](https://doi.org/10.5281/zenodo.14787917)

Перекрёстные ссылки: characteristics space; hedonic price function; implicit price; hedonic regression.

Приложение: формализация на Prolog

Следующий машиночитаемый вспомогательный файл (006-latent-variable.pl) формализует эту статью в базе знаний журнала на Prolog. Он также публикуется рядом с этой статьёй как отдельный просматриваемый файл (galley).

```
%% 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\\pacificacomps').

% 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).

```


Недавние судебные дела — США

Недавние судебные решения • «The Valuation Engineer»

Уильям Берт Крейтор • Том 1, выпуск 1 • Лето 2026 г.

DOI:10.67330/fz34f711

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Краткое изложение дел по юрисдикциям

Калифорния

«The Retail Property Trust» против « » — Дело № 1, рассмотренное Комиссией по обжалованию налоговых оценок округа Ориндж. Апелляционный суд штата Калифорния, решение вынесено в апреле 2026 года.

Апелляционный суд рассмотрел вопрос о возможности предоставления льготы по налогу на недвижимость в соответствии со статьёй 170 Налогового кодекса штата Калифорния в связи с ограничениями доступа, введёнными правительственными постановлениями, изданными во время пандемии COVID-19. Истец — торговый центр — утверждал, что ограничение доступа во время пандемии представляло собой «ущерб» недвижимости в понимании статьи 170, что дает ему право на снижение налога.

Последствия для оценщиков. Данная интерпретация понятия «ущерб» в рамках статьи 170 определит будущие иски о предоставлении льгот, основанные на вызванном внешними факторами снижении полезности недвижимости, независимо от того, является ли оно следствием распоряжений в сфере общественного здравоохранения, инфраструктурных проектов или иных государственных мер, влияющих на доступ. Оценщикам, которым предлагается высказать мнение о влиянии ограничений доступа на стоимость коммерческой недвижимости, следует изучить аргументацию суда относительно того, является ли ограничение полезности — в отличие от физического ущерба — ущербом, подлежащим компенсации для целей освобождения от уплаты налога.

Колорадо

Cheroutes против Rocket Mortgage & Solidifi. США Окружной суд, округ Колорадо; дело находится на рассмотрении, обновления по ходу дела ожидаются в 2026 году.

В иске утверждается наличие предвзятости при оценке в нарушение Закона о справедливом жилищном обеспечении в связи с проведенной в 2021 году оценкой дуплекса в Денвере, стоимость которого была определена в размере примерно \$ 640,000 — что на \$ 220,000 ниже оценочной стоимости предыдущего года. Ранее федеральный судья отклонил ходатайство ответчиков об отклонении иска, постановив, что компания Rocket Mortgage и компания по управлению оценкой Solidifi должны отвечать на предъявленные исковые требования, а не отказываться от ответственности на том основании, что кредиторы и компании по управлению оценкой не могут нести ответственность за действия независимых оценщиков, работающих по гонорару. В 2026 году по делу проводится сбор доказательств.

Последствия для оценщиков. Данное дело является основным действующим прецедентом для проверки ответственности АМС и кредиторов в соответствии с Законом о справедливом жилищном обеспечении в связи с деятельностью оценщиков, работающих за вознаграждение. Если суд в конечном итоге признает АМС и кредитора ответственными по существу дела, это существенно расширит круг сторон, подверженных искам о нарушении принципа справедливого жилищного обеспечения, возникающим в связи с оценками, и вероятно, повлияет на договоры о привлечении оценщиков, заключаемые АМС.

Флорида

Групповой иск во Флориде, оспаривающий комиссионные сборы, взимаемые компаниями АМС с заемщиков, оформивших ипотечные кредиты на жилье. Подан 2 января 2026 года.

В иске оспариваются комиссии компании по управлению оценкой, взимаемые с заемщиков по ипотечным кредитам на жилье, при этом утверждается, что разница между вознаграждением, выплачиваемым оценщику, и комиссией, взимаемой с заемщика при оформлении сделки, превышает пределы, установленные федеральным законодательством, и не раскрывается надлежащим образом. Данное дело находится на стыке требования закона Додда-Франка о «принятых и разумных» комиссионных (12 U.S.C. § 1639e) и систем раскрытия информации для потребителей в соответствии с Законом о честности в кредитовании и Законом о процедурах расчетов по сделкам с недвижимостью.

Последствия для оценщиков. Данное судебное разбирательство ставит под сомнение структуру вознаграждения компаний по управлению активами (АМС), и в случае признания иска коллективным и вынесения решения по существу дела может привести к пересмотру практики раскрытия информации во всей отрасли. Исход дела может опосредованно повлиять на условия установления гонораров для оценщиков, работающих через АМС, поскольку разница между гонорарами, выплачиваемыми АМС, и гонорарами, взимаемыми с заемщиков, станет предметом судебного рассмотрения.

Грузия

Дела о природоохранных сервитутах в «Грин Рок». Налоговый суд Соединенных Штатов, решения, рассмотренные в издании «*Tax Notes*» 2 марта 2026 года.

Налоговый суд рассмотрел каждый из критериев теста «наиболее эффективное использование» при оценке природоохранных сервитутов, установленных на соседних участках (участок «Грин-Рок» и «Грин-Рок-Саут») в провинции Пьемонт штата Джорджия. Эксперты истцов утверждали, что наиболее эффективное использование объектов недвижимости до предоставления сервитута было осуществлено путем эксплуатации гранитного карьера, и использовали анализ дисконтированных денежных потоков для оценки справедливой рыночной стоимости до предоставления сервитута в размере примерно \$ 13,75 млн по сравнению со стоимостью после предоставления сервитута в размере \$ 180,000.

Суд выразил сомнения по каждому из критериев данного теста. Что касается юридической допустимости, суд не был убежден в том, что изменение зонирования с целью разрешения эксплуатации карьера было разумно вероятным, отметив, в частности, что изменение классификации зонирования для участка Green Rock было получено вскоре после того, как была запрошена новая оценка природоохранного сервитута — последовательность событий, кото-

рую суд счел труднодоверенной. Что касается физической возможности, суд счел геологические доказательства заявителей недостаточными для подтверждения того, что на участке на Харман-роуд можно было бы развернуть рентабельный карьер, несмотря на общеизвестное наличие пригодного для добычи гранита в провинции Пьемонт. Суд далее постановил, что даже если бы были допущены правовая допустимость и физическая возможность, это не имело бы решающего значения без выполнения остальных условий анализа.

Значение для оценщиков. Данное заключение представляет собой наглядный пример строгого последовательного применения четырехкомпонентного критерия «наиболее эффективного использования». Оно закрепляет позицию, установленную в делах *«Ranch Springs, LLC» против комиссара по налогам штата*, 164 Т.С. № 6 (2025), и *«Beaverdam Creek Holdings, LLC» против комиссара по налогам штата*, Т.С. Мемо. 2025-53, согласно которой Налоговый суд не будет принимать во внимание спекулятивные утверждения о наиболее эффективном использовании даже в тех случаях, когда общая осуществимость предлагаемого использования подтверждается документально в прилегающем регионе. Оценщики, готовящие оценки сервитутов, должны разрабатывать доказательства, специфичные для конкретного участка, по каждому из критериев, а не полагаться на региональную или подтвержденную аналогиями осуществимость, и должны быть особенно осторожны в отношении временных совпадений — таких как изменения зонирования, полученные одновременно с заказом на оценку, — которые могут вызвать скептицизм со стороны суда.

Мэриленд

Приобретение городом Балтимор ипподрома «Пимлико». Процедура принудительного отчуждения недвижимости, проводится в 2025–2026 гг.

Приобретение городом Балтимором ипподрома «Пимлико» на основании права принудительного отчуждения вызывает вопросы о справедливой компенсации, касающиеся взаимосвязи между стоимостью действующего предприятия и стоимостью недвижимости в случае объектов специального назначения. Поскольку операционная деятельность «Пимлико» функционально неотделима от его материальной базы, традиционный анализ аналогов сделок с недвижимостью имеет ограниченную полезность, а оценка по доходному подходу в значительной степени зависит от предположения о продолжении проведения скачек.

Последствия для оценщиков. За ходом данного разбирательства стоит следить любому оценщику, которому поручено провести оценку объекта одноцелевого назначения, где деятельность предприятия и недвижимость функционально неразделимы, например, таких объектов, как ипподромы, киностудии, тематические парки, пристани для яхт, а также определенных объектов в сфере здравоохранения или гостеприимства. Методология, сформировавшаяся в результате судебных разбирательств, вероятно, повлияет на то, как при расчете справедливой компенсации будут учитываться приросты стоимости действующего предприятия в будущих делах об отчуждении.

Иски в соответствии с Федеральным законом о справедливом жилищном обеспечении против оценщиков

Бракстон против « » Bank of America. . Подано в апреле 2026 года.

Владелица жилья подала иск против Bank of America и оценщика из штата Вирджиния, обвиняя их в расовой дискриминации, что является нарушением Закона о справедливом жилищном

обеспечении. Истица утверждает, что она и её арендаторы присутствовали при проведении оценки и что поведение оценщика свидетельствовало о расовой враждебности. Данное дело продолжает тенденцию, сложившуюся в ходе предыдущих судебных разбирательств по делам о предвзятости при оценке, включая дело *«Коннолли/Мотт против loanDepot»* в округе Мэриленд и дело *«Остин против Miller»* в округе Марин, штат Калифорния, в которых отдельные истцы предъявляли иски на основании Закона о справедливом жилищном обеспечении против оценщиков, компаний по управлению оценкой (АМС) и кредиторов.

Последствия для оценщиков. Неуклонный рост числа исков о предвзятости при оценке, подаваемых против отдельных оценщиков, в сочетании с сокращением масштабов федерального регулирования в сфере борьбы с дискриминацией по признаку неравномерного воздействия указывает на то, что сфера рисков ответственности изменилась: теперь основными источниками ответственности за предвзятость при оценке являются частные истцы и государственные органы, обеспечивающие соблюдение законов о справедливом жилищном обеспечении, а не федеральные агентства. Оценщикам следует пересмотреть свои процедуры ведения документации, чтобы гарантировать, что выбор аналогов, вывод по корректировкам и сверка итоговой стоимости могут быть независимо обоснованы в случае последующей проверки, а также следует помнить, что заявления или поведение во время самого осмотра зачастую становятся фактической основой для исков истцов.

Недавние судебные дела в ЕС

Недавние судебные решения • «The Valuation Engineer»

Уильям Берт Крейтор • Том 1, выпуск 1 • Лето 2026 г.

DOI:10.67330/r1yn1y02

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Объем и источники

Данный обзор составлен на основе общедоступных первоисточников, где это возможно — в первую очередь, официальной базы данных судебной практики Суда Европейского Союза (curia.europa.eu), базы данных HUDOC по делам Европейского суда по правам человека, опубликованных решений Федерального финансового суда (bundesfinanzhof.de), Федерального конституционного суда Германии (bverfg.de), L; например, egifrance для французской судебной практики, а также Верховного трибунал (палата по земельным вопросам) Соединённого Королевства — а также из вторичных комментариев в профессиональных и академических публикациях. Решения на национальных языках были прочитаны в оригинале и описаны на английском языке; названия дел сохраняют оригинальное написание и диакритические знаки.

Базы данных по правовым исследованиям, доступные только по подписке (juris.de, beck-online, Westlaw UK, Lamy, Aranzadi), не используются. Решения национальных судов, находящихся на уровне ниже Верховного суда, анализируются выборочно, при этом приоритет отдается решениям, имеющим явное прецедентное или методологическое значение. Ссылки на судебные дела, приведенные в настоящем документе, следует сверять с официальным отчетом суда, прежде чем полагаться на них в практической деятельности или при дальнейшей публикации.

В данном базовом выпуске в качестве основных национальных юрисдикций рассматриваются Германия, Франция и Соединённое Королевство. Италия, Испания, Нидерланды и страны Северной Европы не рассматриваются отдельно в настоящем выпуске, но будут включены в последующие квартальные выпуски, если это будет оправдано. Швейцария, хотя и не входит в Европейский Союз, будет рассматриваться в будущих выпусках ввиду сильного методологического влияния швейцарских стандартов оценки на европейскую практику в целом.

Европейский нормативно-правовой контекст

Два события в сфере регулирования на уровне Европейского Союза существенно влияют на условия деятельности оценщиков, хотя ни одно из них само по себе не является судебным решением:

Закон ЕС об ИИ и автоматизированные модели оценки. Регламент (ЕС) 2024/1689 (Закон ЕС об ИИ) вступил в силу в 2024 году, при этом положения, применимые к системам ИИ с высоким уровнем риска, вступают в силу поэтапно до августа 2026 года. 19 мая 2026 года Европейская комиссия опубликовала проект руководящих принципов по классификации систем ИИ с высоким уровнем риска для консультаций с заинтересованными сторонами.

В Приложении III к Закону об ИИ оценка кредитоспособности в отношении физических лиц включена в число восьми категорий случаев использования, предположительно сопряженных с высоким риском, и в проекте руководящих принципов Комиссии эта категория подробно разъясняется. Поскольку автоматизированные модели оценки (AVM), используемые при выдаче ипотечных кредитов, мониторинге ипотечного портфеля и страховом андеррайтинге, функционально неотделимы от оценки кредитоспособности по соответствующим кредитам, AVM, эксплуатируемые кредитными учреждениями или страховщиками в пределах Европейского Союза, в большинстве случаев будут классифицироваться как системы искусственного интеллекта с высоким уровнем риска. Последствия в плане соблюдения нормативных требований включают обязательные системы управления рисками, механизмы контроля за управлением данными, автоматическую регистрацию с хранением журналов не менее шести месяцев, механизмы человеческого надзора, позволяющие операторам понимать и отменять решения системы, а также обязательства по мониторингу после вывода на рынок.

Последствия для оценщиков. Европейским оценщикам следует ожидать, что рабочие процессы, основанные на AVM и используемые их клиентами-кредиторами, будут объектом оценки с значительно ужесточенными требованиями к документации и обоснованности результатов. В тех случаях, когда оценщики предоставляют вводные данные для моделей, данные для калибровки или проводят экспертную проверку выводов AVM, эти виды деятельности сами по себе становятся элементами политики кредитора в области соблюдения нормативных требований в отношении ИИ с высоким уровнем риска.

Директива об энергетической эффективности зданий (переработанная редакция). Директива (ЕС) 2024/1275, переработанная редакция Директивы об энергетической эффективности зданий (EPBD), была опубликована в Официальном журнале 8 мая 2024 года и вступила в силу 28 мая 2024 года. Срок транспонирования для государств-членов — 29 мая 2026 года. Среди прочих положений переработанная EPBD требует от государств-членов уделять приоритетное внимание реновации 16 43 минимальные стандарты энергоэффективности (MEPS); требует строительства зданий с нулевыми выбросами для новых объектов общественного назначения с 2028 года и для всего нового строительства с 2030 года; а также предписывает проведение оценки углеродного следа на протяжении всего срока эксплуатации для новых зданий с 2030 года.

Последствия для оценщиков. Переработанная директива EPBD повлияет на оценку недвижимости по нескольким каналам: минимальные пороги энергоэффективности создают новую категорию устаревания, обусловленного необходимостью соблюдения нормативных требований; рейтинг сертификата энергетической эффективности (EPC) станет существенным фактором, определяющим цену как жилого, так и коммерческого фонда; а затраты на реконструкцию, необходимые для достижения соответствия требованиям, должны быть отражены в затратном подходе и при вычете остаточного срока экономической эксплуатации. Оценщикам также следует учитывать, что кредиторы, объектом оценки которых являются лица, подпадающие под действие Директивы об ипотечном кредитовании, будут все чаще требовать четкого учета рейтингов EPC и связанных с MEPS требований к капитальным затратам при оценке залогового имущества.

Примечания к последующим частям

В последующих ежеквартальных обзорах, основанных на данном базовом отчете, основное внимание будет уделяться изменениям: новым решениям, разрешению нерешенных вопросов, указанных в настоящем документе (в частности, рассмотрению Федеральным конституционным судом Германии дела № 1 *BvR 472/26*), а также законодательство государств-членов, транспонирующее переработанную Директиву по энергетической эффективности зданий (EPBD) в национальное право. Швейцарская правовая база, анализ решений верховных судов Италии и Испании, а также события в рамках деятельности *Конституционного совета* Франции будут добавлены по мере поступления соответствующих материалов.

Краткое изложение дел по юрисдикциям

Суд Европейского союза

***GR Real* (Дело C-351/23).** Суд Европейского союза, решение от 24 июня 2025 года.

Суд Европейского Союза вынес решение по предварительному запросу, поступившему от словацкого суда в связи с внесудебной процедурой обращения взыскания по ипотеке, в ходе которой потребители лишились своего семейного дома в результате аукциона, хотя судебное разбирательство по оспариванию потенциально недобросовестных условий договора еще не было завершено. Суд постановил, что гарантии защиты прав потребителей, предусмотренные Директивы 93/13/ЕЭС требуют, чтобы национальные процессуальные нормы допускали эффективную судебную проверку справедливости договорных условий до того, как ипотечное жилье может быть продано в рамках внесудебного обращения взыскания, и что процедуры, позволяющие окончательно осуществить продажу до завершения такой проверки, несовместимы с Директивой.

Последствия для оценщиков. Данное решение не является непосредственно делом об оценке, однако оно изменяет процессуальную среду, в которой заказываются и используются оценки ипотечного обеспечения. В ряде государствах-членах, где действуют внесудебные процедуры принудительного исполнения (в частности, в Словакии, Чешской Республике, Венгрии и странах Балтии), практическим следствием этого является замедление процесса принудительных аукционов и увеличение срока, в течение которого оценка залога, проведенная кредитором, используется в качестве основания. Оценщики, готовящие отчёты об оценке залога для целей принудительного взыскания в этих юрисдикциях, должны учитывать, что отчёты об оценке могут использоваться в течение более длительных периодов, чем ранее, и могут подвергаться проверке на соответствие рыночным изменениям, произошедшим в период между датой первоначальной оценки и окончательной продажей.

Европейский суд по правам человека

Снос здания в Баку (Азербайджан), решение от июля 2025 года. Европейский суд по правам человека, статья 1 Протокола № 1.

Европейский суд по правам человека установил нарушение права на мирное владение имуществом в связи со сносом и изъятием здания в Баку, которые были санкционированы после того, как власти сослались на риск землетрясения. Страсбургский суд установил, что национальные власти при определении размера компенсации не привели достаточных обоснований для отклонения отчёта об оценке, представленного заявителями, и необоснованно приняли во

внимание только сумму, указанную в экспертном заключении, заказанном самим судом. Суд присудил в общей сложности € 75 000 в качестве компенсации материального ущерба, €3 000 в качестве компенсации морального ущерба и € 1 055 в качестве судебных издержек.

Последствия для оценщиков. Данное решение укрепляет линию судебной практики по статье 1 Протокола № 1, требующую, чтобы национальные суды мотивированно рассматривали противоречивые доказательства по оценке в рамках процедур изъятия, а отклонение отчёта об оценке заявителя в пользу экспертного заключения, заказанного судом, само по себе должно быть мотивированным. Для европейских оценщиков, работающих в сфере изъятия и справедливой компенсации, данное дело подчеркивает, что документальная обоснованность отчёта об оценке — четкое обоснование, подтверждающее выбор аналогов, выбор сопоставимых объектов, вывод корректировок и согласование окончательной стоимости — сама по себе является вопросом, затрагивающим основные права, когда отчёт представляется в поддержку иска против государства.

Франция

Кассационный суд, Торговая палата, 19 июня 2024 г., №°22-24.169. (Обсуждается в профессиональных комментариях вплоть до 2025 года.)

Кассационный суд постановил в рамках споров о рыночной стоимости (*valeur vénale*) в контексте налога на наследство, что продажа унаследованной недвижимости, произошедшая после наступления налогооблагающего события (смерти *наследодателя*), не может использоваться наследниками для пересмотра в сторону понижения первоначально заявленной «*valeur vénale*», даже если цена продажи, состоявшейся после наступления события, значительно ниже, а сама продажа произошла в течение нескольких месяцев после смерти. В данном решении применяется принцип, согласно которому «*valeur vénale*» для целей налога на наследство определяется по состоянию на дату налогооблагающего события и не корректируется с учетом последующих сделок с той же недвижимостью.

Последствия для оценщиков. Данное решение подтверждает строгую приверженность принципу «даты оценки» в практике начисления налога на наследство во Франции и ограничивает значимость сделок купли-продажи, заключенных после даты оценки, в качестве ретроспективных доказательств стоимости. Оценщики, готовящие оценки для целей налога на наследство во Франции, должны подготовить исчерпывающие данные о сопоставимых сделках, основанные на сделках, *совершённых на дату оценки или до неё*, и не должны представлять сделки купли-продажи самого объекта оценки, заключённые после даты оценки, в качестве подтверждающих доказательств. Аналогичная аргументация, вероятно, будет применима и к налогу на дарение (*droits de donation*), и к налогу на имущество (*IFI*).

Кассационный суд, 3-я^е-я гражданская палата, 8 января 2026 г., №°23-22.803. Досудебные (внесудебные) экспертные отчёты.

Третья гражданская палата Кассационного суда вынесла постановление, пересматривающее правовую значимость отчётов, подготовленных в порядке внесудебного урегулирования, в гражданском судопроизводстве Франции. Данное решение является следствием более широкой процессуальной реформы, осуществлённой декретом «» от 18 июля 2025 года, в соответствии с которым медиация и внесудебное урегулирование споров стали нормой по умолчанию в гражданском процессе. Суд разъяснил, что экспертные отчёты, подготовленные по договорной инициативе одной из сторон, тем не менее могут быть приняты в качестве доказательств и им

может быть придан вес при условии соблюдения принципа состязательности и после оценки судом независимости эксперта и его методологии.

Последствия для оценщиков. Данное решение укрепляет доказательную силу *экспертных заключений по недвижимости*, подготовленных вне рамок системы судебных экспертов, назначаемых судом (*experts judiciaires*), при условии, что в отчёте соблюдены методологические стандарты профессии и задокументирована независимость эксперта. Для французских оценщиков, занимающихся спорными оценками в рамках наследственных, разводных или коммерческих споров, данное решение создает более прочную основу для того, чтобы отчёты, подготовленные по заказу клиента, могли цитироваться и использоваться в последующих судебных разбирательствах. Соответственно, особую важность приобретает тщательное документирование независимости (отсутствие подчиненности, экономической заинтересованности или семейных связей с клиентом) и применение методов, предписанных профессиональными стандартами.

Германия

Федеральный финансовый суд, постановления от 12 ноября 2025 г. (II R 25/24, II R 31/24, II R 3/25). Трилогия по *федеральной модели налога на недвижимость* — конституционность реформированной системы оценки недвижимости для целей налогообложения. Вторая палата Федерального финансового суда Германии (*Bundesfinanzhof*, BFH) в ходе рассмотрения трёх параллельных апелляционных дел постановила, что положения метода оценки на основе капитализации дохода (*Ertragswertverfahren*), применяемого в рамках так называемой «*федеральной модели*» (*Bundesmodell*) для оценки жилой недвижимости в одиннадцати федеральных землях Германии в целях реформированного налога на недвижимость (*Grundsteuer*), вступающего в силу с 1 января 2025 года, соответствуют Конституции. Истцами выступили владельцы квартир в Северной Рейн-Вестфалии, Саксонии и Берлине, которые оспорили свои налоговые оценки на том основании, что установленные правила оценки нарушают общий принцип равенства (статья 3, пункт 1 Основного закона) и приводят к систематически неточным результатам оценки.

Федеральный финансовый суд (BFH) постановил, что законодатель создал систему оценки, *концептуально ориентированную на рыночную стоимость* (*konzeptionell einer Verkehrswertorientierung folgt*) и призванную приблизить, в среднем по всем оцениваемым объектам недвижимости, «объективно реалистичную стоимость недвижимости» в пределах диапазона справедливой рыночной стоимости (*gemeiner Wert*). Суд пришел к выводу, что законодатель не превысил допустимые рамки при сопоставлении целей системы оценки с неизбежными для такой системы неравенствами, и что значительный вес, приданный предотвращению повторного «отставания в оценке» посредством автоматизированного периодического обновления, является конституционно допустимым. Федеральный финансовый суд (BFH) также подчеркнул, что налогоплательщики сохраняют возможность, установленную ранее в постановлениях BFH «*Beschlüsse*» от 27 мая 2024 г. (II B 78/23 и II B 79/23), доказать более низкую справедливую рыночную стоимость (*gemeiner Wert*) посредством квалифицированной оценки в соответствии с §198 BewG, при этом более низкая стоимость должна быть подтверждена отчётом об оценке, составленным государственно уполномоченным оценщиком недвижимости с применением признанных методов оценки.

Конституционная жалоба (*Verfassungsbeschwerde*), поданная при поддержке организаций «*Bund der Steuerzahler Deutschland*» и «*Haus & Grund Deutschland*», в настоящее время находится на

рассмотрении в Федеральном конституционном суде (*Bundesverfassungsgericht*) под номером дела 1 BvR 472/26.

Последствия для оценщиков. Эта трилогия представляет собой наиболее значимую судебную практику Германии в области оценки недвижимости за рассматриваемый период. Для оценщиков основным практическим следствием является возросший спрос на экспертизы рыночной стоимости (§ 198 BewG) по «более низкой стоимости» (*Verkehrswertgutachten zum Nachweis des niedrigeren gemeinen Werts*) в тех случаях, когда *Grundsteuerwert*, определенный в соответствии с *Bundesmodell*, превышает фактическую рыночную стоимость объекта оценки. Явное одобрение этого подхода со стороны Федерального финансового суда (BFH) в сочетании с его общей приверженностью «*Бундесмоделю*» означает, что отчет об оценке теперь является действующим механизмом для индивидуального освобождения налогоплательщиков от систематической завышенной оценки. Оценщикам следует ожидать устойчивого роста количества заказов на такие отчеты и обеспечить соблюдение методологии ImmoWertV 2021 и требований п§а 198 BewG. Находящееся на рассмотрении конституционное жалоба в Федеральном конституционном суде (1 BvR 472/26) будет иметь важное значение для отслеживания любых дальнейших изменений в данной нормативной базе и, как ожидается, станет основным объектом оценки следующего квартального выпуска.

Федеральный финансовый суд, решение от 7 октября 2025 г., IX R 26/24. Распределение покупной цены между земельным участком и зданием (*Kaufpreisaufteilung*). В деле о налоге на доходы, касающемся амортизации (*Absetzung für Abnutzung*, AfA) недвижимости, сдаваемой в аренду, Девятый сенат Федерального налогового суда рассмотрел методику распределения договорной покупной цены между подлежащей амортизации земельной составляющей (*Grund und Boden*) и подлежащей амортизации строительной составляющей (*Gebäude*). Истцы оспаривали использование судом низшей инстанции стандартных значений стоимости земельных участков (*Bodenrichtwerte*) на том основании, что исходная база данных не была достоверно задокументирована в отношении конкретного местоположения объекта оценки.

Последствия для оценщиков. Данное решение продолжает линию судебной практики Федерального финансового суда (BFH) (в частности, вышеупомянутое решение BFH II R 25/24 от 12 ноября 2025 года, а также более раннее решение BFH IX R 12/21 от 20 сентября 2022 года), подтверждающую, что ImmoWertV является источником признанных принципов оценки для распределения покупной цены. Данное дело также иллюстрирует важность базы данных, лежащей в основе *Bodenrichtwerte*: если *Gutachterausschüsse* (экспертные оценочные комитеты) не располагают достаточными данными о сопоставимых сделках купли-продажи для конкретного местоположения, *Bodenrichtwert* может быть оспорен по причинам, связанным с качеством данных. Для оценщиков, готовящих заключения по распределению покупной цены или по базе амортизации, это означает, что документация, подтверждающая базу данных *Bodenrichtwert*, включая лежащий в ее основе сборник данных о покупных ценах (*Kaufpreissammlung*), становится неотъемлемой частью обоснованного анализа.

Великобритания

Мехруниса Аликхан Ахмед Хан против Городского совета Вулверхэмптона. Верхний трибунал (палата по земельным вопросам), 2025 г.

Верхний трибунал (палата по земельным вопросам) подтвердил, что надлежащей датой оцен-

ки для определения размера компенсации за принудительное отчуждение является дата, когда право собственности на недвижимость переходит к органу, осуществляющему отчуждение, в соответствии с общим заявлением о переходе права собственности (GVD). Трибунал также рассмотрел вопрос о взаимосвязи между оценками, произведенными в момент события, и ретроспективными оценками: в рассматриваемом деле первоначальная оценка местного органа власти, произведенная сразу после GVD, составила £ 150 000, в то время как вторая оценка, подготовленная в 2025 году в ходе разбирательства в трибунале со ссылкой на предполагаемую дату перехода права собственности через три месяца после уведомления собственника, составила £ 120 000. Несмотря на то что трибунал подтвердил законную дату оценки, он отдал предпочтение цифре, рассчитанной ближе по времени к дате вынесения решения о принудительном отчуждении, что привело к выплате истцу более высокой суммы.

Выводы для оценщиков. Данное решение носит методологический характер и содержит полезные рекомендации для британских оценщиков, работающих в сфере принудительного отчуждения. Хотя юридической датой оценки является дата перехода права собственности, предпочтение, которое трибунал отдаёт оценке, произведённой в момент события, по сравнению с ретроспективной оценкой, подготовленной спустя годы, но с той же юридической датой оценки, показывает, что рыночные данные, полученные в момент события, имеют практический вес даже в тех случаях, когда они формально не были бы допустимы на саму дату перехода права собственности. Оценщики должны обеспечить, чтобы отчёты об оценке, подготовленные вблизи даты вступления в силу, сопровождались достаточными подтверждающими доказательствами, чтобы служить в качестве основного документа, предвидя, что в ходе последующих разбирательств в трибунале может быть отдано предпочтение отчёту, составленному на момент оценки, а не ретроспективной реконструкции.

Недавние судебные дела в Латинской Америке

Недавние судебные решения • «The Valuation Engineer»

Уильям Берт Крейтор • Том 1, выпуск 1 • Лето 2026 г.

DOI:10.67330/ywgzghg47

Русский перевод, подготовленный с использованием искусственного интеллекта и прошедший редакционную проверку; английская версия является основной.

Объем и источники

Данный обзор основан на общедоступных первоисточниках, где это возможно — в частности, на базе данных дел Межамериканского суда по правам человека (corteidh.or.cr), Верховного суда Бразилии (STF) и Верховного суда (STJ) (portal.stf.jus.br, stj.jus.br), мексиканского «*Semanario Judicial de la Federación*» (sjfsemanal.scjn.gob.mx), аргентинского «*Corte Suprema de Justicia de la Nación*» (csjn.gov.ar), чилийский портал *Poder Judicial* (poderjudicial.cl) и *Diario Constitucional*, а также Колумбийский *Конституционный суд* (corteconstitucional.gov.co) — а также из вторичных комментариев в профессиональных и академических изданиях. Решения на испанском и португальском языках были прочитаны в оригинале и описаны на английском языке; названия дел сохраняют оригинальное написание и диакритические знаки.

Не используются базы данных по правовым исследованиям, доступные только по подписке (vLex, La Ley, Thomson Reuters Proview Latin America) не используются. Решения национальных судов ниже уровня верховного суда анализируются выборочно, при этом приоритет отдается решениям, имеющим явное прецедентное или методологическое значение. Ссылки на судебные дела, приведенные в настоящем документе, следует сверять с официальным отчетом суда, прежде чем полагаться на них в практической деятельности или для дальнейшей публикации.

В данном базовом выпуске в качестве основных национальных юрисдикций рассматриваются Аргентина, Бразилия, Чили, Колумбия и Мексика. Перу, Венесуэла, Эквадор, Боливия, Парагвай, Уругвай, а также юрисдикции Центральной Америки и Карибского бассейна не рассматриваются отдельно в данном выпуске, но будут включены в последующие квартальные выпуски по мере необходимости. Работа по оценке в рамках дел о возвращении недвижимости в Венесуэле и связанных с санкциями дел на Кубе будет рассмотрена в одном из будущих выпусков ввиду их методологической специфики и значимости для трансграничной практики.

Региональный контекст

В отличие от исследования, посвященного Европейскому союзу, данное исследование, посвященное Латинской Америке, не имеет прямого аналога на наднациональном уровне регулирования, представленном Законом ЕС об искусственном интеллекте и переработанной Директивой об энергетической эффективности зданий. Регулирование оценки недвижимости в Латинской Америке в подавляющем большинстве случаев носит национальный характер; основными наднациональными факторами влияния являются судебная практика Межамериканского суда по правам человека в области основных прав в соответствии со статьёй 21 Американской

конвенции (право на собственность) и принятие Международных стандартов финансовой отчётности (в частности, МСФО 13 «Оценка по справедливой стоимости») большинством региональных юрисдикций для целей финансового отчёта. Тем не менее, две методологические темы прослеживаются во всей региональной судебной практике и заслуживают упоминания в данном контексте:

Оценка общинной недвижимости коренных народов. Наиболее характерной методологической проблемой в латиноамериканской практике оценки является определение стоимости общинной недвижимости коренных народов и лиц африканского происхождения в контексте реституции, изъятия и выплаты компенсаций. В судебной практике Межамериканского суда последовательно отстаивается позиция, что общинная недвижимость коренных народов находится под конституционной защитой наравне с индивидуальной частной недвижимостью, однако методы оценки такой недвижимости — в условиях, когда, как правило, отсутствуют аналоги сделок на рыночных условиях, когда недвижимость находится в общинном и неотчуждаемом владении, а культурное и духовное значение территории является неотъемлемой частью ее стоимости — остаются недостаточно разработанными. По состоянию на конец 2025 года двадцать четыре решения Межамериканского суда в пользу коренных народов оставались неисполненными или были выполнены лишь частично, причем в ряде случаев речь шла о нерешенных вопросах расчета компенсации, зависящих от неурегулированной методологии оценки.

Инфляция, валюта и индексация при выплате компенсации за изъятие. В ряде региональных юрисдикций, в частности в Аргентине, требуется применение методологий оценки и расчётов компенсации, позволяющих учитывать высокий уровень инфляции, оценки в иностранной валюте и сложные режимы индексации. Верховный суд Аргентины постановил, что компенсация за изъятие, первоначально выраженная в иностранной валюте, не может быть конвертирована в местную валюту по искусственному курсу, и что оценка должна отражать реальную экономическую стоимость недвижимости на момент выплаты компенсации. Аналогичные проблемы возникают в Венесуэле (в связи с реституцией и оценками, связанными с санкциями) и периодически в Бразилии в периоды высокой инфляции. Для оценщиков, работающих в различных региональных юрисдикциях, эти соображения, связанные с инфляцией и валютой, являются неотъемлемой частью методологии оценки, что не имеет аналогов ни в Европе, ни в США.

Примечания к последующим частям

В последующих ежеквартальных отчетах, основанных на данном базовом сценарии, основное внимание будет уделяться следующим изменениям: новым решениям Межамериканского суда, разрешению нерешенных вопросов, указанных в данном документе, а также ходу реализации Колумбией «Jurisdicción Agraria y Rural» в целях начала его функционирования в 2027 году, а также динамика в государствах-членах в рамках практики оценки по справедливой стоимости в соответствии с МСФО 13. Перу, Венесуэла, Эквадор и юрисдикции Центральной Америки будут добавлены по мере поступления существенной информации. Особое внимание будет уделено методологическому развитию оценки общинной собственности коренных народов, где материально-правовые постановления Межамериканского суда опередили практическую методологию оценки, доступную оценщикам и органам, ответственным за выплату компенсаций.

Краткое изложение дел по юрисдикциям

Межамериканский суд по правам человека

«Майя-Кекчи Агуа-Калиенте» против Гватемалы и «Сообщество гарифуна Сан-Хуана и его члены» против Гондураса. Межамериканский суд по правам человека, решения 2024 года, рассмотренные в региональной судебной практике до 2025 года.

В двух вынесенных одновременно решениях Межамериканский суд высказался по вопросу об обязательствах Гватемалы и Гондураса по демаркации, оформлению прав собственности и предотвращению лишения прав на исконные территории, принадлежащие общинам коренных народов и афро-потомков. Община майя-кекчи Агуа-Калиенте получила окончательное право собственности на свою территорию лишь после сорока пяти лет административных и судебных разбирательств; община гарифуна Сан-Хуан в Гондурасе сталкивалась с продолжающимся лишением прибрежных земель, унаследованных от предков. В обоих решениях рассматривались обязательства государства в соответствии со статьёй 21 (право на недвижимость) Американской конвенции по признанию, демаркации и защите общинной недвижимости коренных народов. Региональные обозреватели отмечают, что рамки реституции, установленные Межамериканским судом в этих решениях, являются несколько более ограничительными, чем в более ранних решениях, особенно в отношении «невиновных третьих лиц» (*terceros inocentes*), которые приобрели пересекающиеся права, не зная о претензиях коренных народов.

Последствия для оценщиков. Данные решения подтверждают, что общинная недвижимость коренных народов и лиц африканского происхождения находится под конституционной защитой в рамках всей Межамериканской системы и что оценка такой недвижимости в контексте реституции или компенсации не может основываться на обычных аналогах сравнительных сделок между независимыми сторонами (которые, как правило, отсутствуют) без методологического дополнения. Оценщики, привлекаемые в рамках процедур реституции земель коренных народов, должны учитывать, что оценка должна включать три отдельных компонента: базовую рыночную стоимость земельных участков с аналогичными физическими характеристиками в данном регионе; альтернативные издержки и улучшения, характерные для использования коренными народами; и неэкономическое культурное и духовное значение территории, которое в судебной практике Межамериканского суда рассматривается как подлежащий компенсации аспект права собственности. Методология оценки третьего компонента по-прежнему недостаточно разработана и остается объектом постоянного рассмотрения в судебной практике Межамериканского суда.

Мапуче / «Совет всех земель» против Чили. Межамериканский суд по правам человека, решение от 2024 года (рассматривается в судебной практике за 2025 год).

Межамериканский суд признал чилийское государство виновным в нарушении прав человека в отношении 135 представителей народа мапуче в связи с акциями по отстаиванию территориальных прав, проводившимися организацией «Consejo de Todas las Tierras» (1989–1992 гг.), а также в связи с применением антитеррористического законодательства к территориальным претензиям коренных народов. Данное решение развивает доктрину, сформулированную в деле «Норин Катриман и другие против Чили» (2014 г.), и прямо признает право коренных народов на самоопределение в связи с территориальными претензиями.

Последствия для оценщиков. Хотя основные положения данного дела касаются скорее уголовно-процессуальных гарантий, чем непосредственно методологии оценки, признание права корен-

ных народов на самоопределение в отношении территории создает процессуальный контекст, который влияет на работу по оценке в рамках реституции территорий коренных народов Чили и в связанных с этим контекстах. Оценщики, привлеченные в таких ситуациях, должны обеспечить, чтобы методология предусматривала надлежащие процессы консультаций и признавала право общины на самоопределение в отношении того, как определяются ценности её территории.

Аргентина

Аргентинская доктрина по вопросам Трибунала по оценкам (Tribunal de Tasaciones de la Nación). Судебная практика, обобщенная по состоянию на 2025 год.

Верховный суд Аргентины (CSJN) и нижестоящие федеральные суды продолжают развивать доктрину, касающуюся оценок, проводимых Национальным судом по оценкам (Tribunal de Tasaciones de la Nación, TTN) в делах о изъятии и *обратном изъятии (expropiación)*. Согласно устоявшейся доктрине CSJN (постановления 313:82, неоднократно применявшейся в период 2024–2025 гг.) гласит, что оценки TTN имеют презумпцию веса в силу технической компетентности членов трибунала, доказательной базы, на которой они основаны, и единообразия практики; отклонение от оценки TTN требует доказательства явной ошибки или упущения. Федеральные суды, применявшие эту доктрину в 2024–2025 годах — включая Федеральный суд Обера а в деле «*Coca против Entidad Binacional Yacyret*» а (июнь 2024 г., *экспроприация по вопросам обратного применения*, возникающим в связи с гидроэлектростанцией Yacyret а) — продолжать опираться на оценки TTN в качестве основополагающего критерия для выплаты компенсации.

Последствия для оценщиков. Для аргентинских оценщиков практическим следствием является то, что частные отчёты об оценке, представленные в опровержение оценок TTN, сталкиваются со значительным бременем доказывания. Чтобы отменить или изменить оценку TTN, частный отчёт должен продемонстрировать методологическую ошибку или упущение в работе TTN, а не просто альтернативный вывод об оценке. Оценщики, готовящие такие отчёты, должны сосредоточить методологическую критику на подборе аналогов сделок TTN, выводах по корректировкам и учете условий местного рынка, а не только на величине разницы в стоимости.

Оценка стоимости в иностранной валюте и паритет возмещения. Доктрина Верховного суда Аргентины, согласно которой компенсация за изъятие, изначально выраженная в иностранной валюте, не может быть конвертирована в аргентинские песо по искусственному паритету «один к одному», остается в силе до 2025 года. Эта доктрина защищает экспроприированных собственников от компенсаций, которые вследствие инфляции или валютной политики не отражают реальную экономическую стоимость недвижимости на момент выплаты компенсации. Этот принцип периодически подтверждался в 2024–2025 годах в связи с затянувшимися судебными разбирательствами по делам об изъятии.

Выводы для оценщиков. Аргентинские оценщики, готовящие оценки по делам о изъятии, охватывающие периоды значительной инфляции или изменения валютной политики, должны четко фиксировать валютную основу оценки и учитывать надлежащую методику индексации за период между датой оценки и предполагаемой датой выплаты. Отчёты, в которых не учитываются инфляция и валютный риск, вряд ли будут убедительными для третейских судов, применяющих доктрину компенсации CSJN.

Бразилия

Верховный федеральный суд, дело № 922.144 (тема № 865, имеющая общее значение). Верховный федеральный суд Бразилии, постановление от 2024 года, имеющее непрерывное применение в 2025 году.

Верховный Федеральный Суд постановил, что выплата компенсации в рамках процедуры *принудительного отчуждения с ао* соответствует Конституции и совместима с федеральной системой *предварительного orio* режима, установленной статьёй 100 Федеральной Конституции. В частности, разница между суммой, первоначально внесенной государственным органом (в целях *imiss ao provis oria na posse*), и суммой, окончательно установленной в результате судебной экспертизы, должна выплачиваться через систему *precat orio*, а не в виде немедленного дополнительного взноса. Данное решение обеспечивает баланс между конституционным требованием о *справедливой, pr evia e em dinheiro* компенсации (статья 5, XXIV) и требованиями упорядоченной выплаты в рамках системы *precat orio*.

Последствия для оценщиков. Данное решение имеет прямые практические последствия для оценки первоначального взноса, проводимой оценщиками, действующими по поручению органов, осуществляющих изъятие. Поскольку любая заниженная оценка первоначального депозита будет впоследствии компенсирована через систему *предварительной оценки (orio)* (что может повлечь за собой существенную задержку), у истцов существует структурный стимул оспаривать первоначальные оценки как заниженные. Оценщики, готовящие оценки первоначального депозита, должны с особой тщательностью документировать методологию, чтобы защитить оценку от последующей судебной экспертизы, предвидя, что разница между первоначальным депозитом и окончательным судебным решением будет объектом оценки и подлежать начислению компенсационных процентов (*juros compensatórios*) в соответствии с судебной практикой STF судебной практики и положений Декрета-закона № 3.365/1941.

Верховный суд Бразилии (Superior Tribunal de Justiça), дело № REsp 2.129.162 (повторное рассмотрение вопроса). Верховный суд Бразилии, решение от мая 2025 года.

Верховный судебный суд (STJ) разрешил повторяющийся спор, касающийся гонораров адвокатов в случаях, когда орган, осуществляющий изъятие, отказывается от иска *об изъятии с ао*. Суд постановил, что гонорары в случаях отказа от иска должны определяться в соответствии со статьёй 27, § 1, Декрета-закона № 3.365/1941, а не общих процессуальных норм *Гражданского процессуального кодекса (C)*. Решение было вынесено по иску компании Cemig (Companhia Energética de Minas Gerais) о предоставлении административного сервитута для линии электропередачи.

Последствия для оценщиков. Данное решение носит скорее процессуальный, чем методологический характер, однако оно имеет последствия для структуры участия оценщиков в делах, связанных с административным сервитутом и *экспроприацией с ао*. Оценщики, привлекаемые в качестве судебных экспертов (*peritos judiciais*) в таких разбирательствах, должны обеспечить, чтобы договоренности об оплате были зафиксированы в документах со ссылкой на применимые нормы декрета-закона, а не на общие процессуальные правила.

Верховный суд (Superior Tribunal de Justiça), доктрина о компенсационных процентах (*juros compensat orios*). Продолжение применения в 2025 году.

Верховный суд штата (STJ) придерживается своей доктрины, согласно которой компенса-

онные проценты (*juros compensatórios*) в делах об изъятии начисляются только после вынесения судебного решения о праве собственности на изъятая недвижимость. Проценты начисляются на разницу между первоначальным залогом и окончательной суммой компенсации по исторически установленной ставке в двенадцать процентов годовых (Súmula 618 Верховного Федерального Суда), и служат для компенсации владельцу изъятного имущества досрочной утраты владения до момента выплаты окончательной компенсации .

Последствия для оценщиков. Бразильские оценщики должны учитывать, что выявленный ими в ходе судебной экспертизы разрыв в оценке — между суммой первоначального депозита, внесенного государственным органом, и оцененной рыночной стоимостью — становится основой для расчета компенсационных процентов. Таким образом, экономическое значение этого разрыва больше, чем можно было бы предположить, исходя из номинальной разницы, особенно в случае длительных судебных разбирательств.

Чили

Верховный суд, дело №º , 31.423-2025. Решение Верховного суда Чили от сентября 2025 года.

Верховный суд Чили отклонил кассационную жалобу, поданную девелоперской компанией (*inmobiliaria*) против Государственного казначейства (*Fisco de Chile*) в рамках специального разбирательства по делу о изъятии на основании статьи 12 Декрета-закона №º 2.186. Истец утверждал, что оценка, установленная Комиссией экспертов (*Comisión de Peritos*), не отражала недвижимость наиболее эффективное использование (*mayor y mejor uso*) и стоимость существующих построек и дополнительных работ, а также дополнительно требовал возмещения упущенной выгоды (*lucro cesante*) на основании действующего договора аренды, который принес бы будущий доход от аренды. Суд постановил, что аргументы истца сводились к несогласию с оценкой доказательств судом низшей инстанции, а не к фактическому нарушению правовых норм, регулирующих оценку, и что аргументация суда низшей инстанции — которая основывалась на доказательствах Комиссии экспертной комиссии в совокупности с документальными и свидетельскими доказательствами — было надлежащим образом обосновано.

Выводы для оценщиков. Данное решение демонстрирует значительную степень доверия, которую доктрина Верховного суда Чили оказывает оценке, проведенной Экспертной комиссией в рамках процедур изъятия, а также ограниченные возможности кассационного пересмотра выводов по оценке исключительно на основании методологических разногласий. Оценщики, готовящие контраргументы по оценке в чилийских делах о изъятии, должны сосредоточиться на доказуемых методологических ошибках в работе Комиссии, а не на обосновании альтернативной оценки, и должны иметь в виду, что иски о возмещении упущенной выгоды, основанные на существующих договорах аренды, подпадают под существенные требования к доказательствам.

Верховный суд (дело о изъятии в Кокимбо). Решение Верховного суда Чили от 17 июля 2025 года.

Верховный суд Чили утвердил решение апелляционной инстанции, которым сумма компенсации по делу о изъятии в Кокимбо была увеличена до 2 969,75 UF (на основании отчёта частного оценщика) против заявленной государством суммы в размере 2 349,07 UF (на основании зарегистрированных сделок купли-продажи, внесенных в реестр недвижимости *Conservador de*

Bienes Ra). Суд подчеркнул, что опора апелляционного суда на отчёт частного оценщика, рассмотренный в совокупности с показаниями свидетелей и дополнительным отчётом об оценке соседнего объекта недвижимости, являлась допустимым применением *sana crítica* (обоснованного взвешивания доказательств) и что кассационное рассмотрение не распространяется на пересмотр такой оценки доказательств.

Выводы для оценщиков. Решение по делу «Кокимбо» является аналогом решения по делу «Рол Н°» № 31.423-2025: если в случае оспаривания застройщиком превалирует оценка Комиссии, то более высокая оценка частного оценщика также может превалировать, если она подкреплена надлежащим методологическим обоснованием и подтверждающими доказательствами. Практический урок для чилийских оценщиков заключается в том, что методология и документальное обоснование частной оценки имеют решающее значение, и что одних только данных о зарегистрированных сделках продажи недостаточно, чтобы опровергнуть хорошо обоснованную частную оценку, в которой используются дополнительные аналогии и контекстуальные доказательства.

Конституционный суд, октябрь 2025 г. Оценка *исов* в соответствии с Налоговым кодексом.

Конституционный суд Чили отклонил жалобу о неконституционности положений Налогового кодекса (*Código Tributario*), регулирующих оценку недвижимости (*bienes raíces*) для целей уплаты территориального налога (*impuesto territorial*). Судьи, высказавшие особое мнение, утверждали, что оспариваемые нормы предусматривают недостаточные механизмы, позволяющие стороне, в отношении которой принимается решение, оспорить оценку, что ставит под угрозу конституционную гарантию рациональной и справедливой процедуры.

Последствия для оценщиков. Данное решение сохраняет существующую систему административной оценки для целей налогообложения недвижимости в Чили, однако в особом мнении указано на уязвимое место в действующей системе, которое может стать предметом пересмотра в ходе будущих конституционных или административно-правовых споров. Чилийские оценщики, готовящие отчеты об оценке в связи со спорами по налогу на недвижимость, должны следить за изменениями в политической обстановке после вынесения данного решения, поскольку представляются вероятными административные реформы, направленные на улучшение механизмов обжалования налогоплательщиками.

Колумбия

Конституционный суд, утверждение юрисдикции по аграрным и сельским вопросам. Решение Конституционного суда Колумбии от 2025 года.

Конституционный суд Колумбии одобрил конституционную экспертизу законодательного акта о создании новой судебной системы «*Jurisdicción Agraria y Rural*» — специализированной судебной системы по вопросам сельскохозяйственных земель в рамках судебной ветви власти Колумбии. Суд одобрил ключевые статьи, определяющие цель реформы, её интеграцию в судебную власть, создание специализированных трибуналов и судов, а также бюджетные аспекты. Новая юрисдикция должна начать работу в 2027 году и предназначена для разрешения споров о реституции земель, споров об аграрной недвижимости, а также вопросов, возникающих в связи с реализацией мирного соглашения 2016 года. На момент утверждения в действующей административной и судебной системе находятся на рассмотрении аграрные споры, затрагивающие более 61 000 гектаров.

Последствия для оценщиков. Новая юрисдикция станет основной инстанцией для рассмотрения споров по оценке, возникающих в связи с процедурами реституции земель и вопросами, касающимися сельской недвижимости в Колумбии. Колумбийские оценщики, работающие в сфере реституции (в том числе те, кто работает в Unidad de Restitución de Tierras и Instituto Geográfico Agust) в Кодацци) должны учитывать, что новые суды будут формировать собственную судебную практику в области оценки и что методологические стандарты могут измениться по сравнению с нынешней административно ориентированной системой. Практикующим юристам следует следить за началом их работы в 2027 году и за развитием первоначальной судебной практики этих судов.

Существующая система оценки на основе IGAC для возврата земель. Продолжение применения в 2025 году.

Практика реституции земель в Колумбии по-прежнему опирается на документ «*aval uo comercial*», подготовленный Институтом географии и картографии (Instituto Geográfico Agust) в Кодацци (IGAC), в качестве основного основания для выплаты компенсации в контексте реституции и замещения, дополненного документами «*aval uos*», подготовленными земельными биржами (Lonjas de Propiedad Raíz) в случае необходимости. Данная система изложена в постановлении Конституционного суда SU-034 от 2018 года (о процессуальных требованиях к соблюдению) и в последующих постановлениях, включая T-120 от 2024 года (о взаимосвязи между оценкой IGAC и индивидуальными предпочтениями бенефициаров в случаях компенсации в натуральной форме).

Последствия для оценщиков. Для оценщиков, работающих в сфере возврата земель в Колумбии, институциональная структура ставит оценки IGAC в центр процесса расчета компенсации, при этом Lonja de Propiedad Raíz aval uos выступает в качестве дополнительного или механизма разрешения споров. Создаваемые в настоящее время суды Jurisdiccí по Agraria y Rural может изменить эту институциональную архитектуру, и относительный вес оценок IGAC по сравнению с оценками Lonja является одним из вопросов, которые, вероятно, будут рассматриваться новыми судами.

Мексика

Верховный суд Мексики (Suprema Corte de Justicia de la Nación) по вопросу об обязательствах по налогу на доходы (ISR) в связи с авансом (aval) uo — ценовой разницей. Решение Верховного суда Мексики будет оглашено в начале 2026 года.

Пленум Верховного суда Мексики (SCJN) признал конституционными статьи 125, 130(IV) и 132 Закона о подоходном налоге (ISR), которые предусматривают, что в случае, если аванс uo при приобретении недвижимости превышает договорную покупную цену более чем на десять процентов, разница рассматривается как доход, подлежащий налогообложению у покупателя, и, следовательно, является объектом оценки для налогообложения по налогу на доходы физических лиц (ISR). Суд постановил, что данное правило не нарушает конституционных принципов соразмерности и справедливости в области налогообложения, мотивируя это тем, что приобретение недвижимости по цене, существенно ниже рыночной стоимости, представляет собой объективное увеличение имущества покупателя и, следовательно, по праву является налогооблагаемой экономической выгодой.

Последствия для оценщиков. Данное решение имеет существенные практические последствия для сделок с недвижимостью в Мексике, а также для оценки недвижимости, обеспечивающей

их проведение. Оценки (*aval uos*), проводимые брокерами (*corredores p ublicos*) или другими уполномоченными специалистами, теперь имеют прямые налоговые последствия для покупателей в тех случаях, когда сумма оценки (*aval uo*) превышает договорную цену более чем на десять процентов. Оценщики должны быть готовы к более тщательной проверке методологии при составлении оценочных отчетов (*aval uos*) для сделок и к росту спроса на оценочные отчеты (*aval uos*), в которых основа сопоставимых продаж документируется с особой строгостью. Данное решение также косвенно повышает значимость выбора методологии при выборе между альтернативными подходами к оценке, поскольку этот выбор теперь может повлиять на налоговые обязательства покупателя.

Верховный суд Мексики (Suprema Corte de Justicia de la Nación), рассмотрение ходатайства об ампаро (Amparo) в порядке пересмотра (revisión) по делу № 391/2024. Вторая палата Верховного суда Мексики, постановление от 25 июня 2025 года.

Вторая палата Верховного суда Мексики рассмотрела конституционную жалобу горнодобывающей компании на указ, затрагивающий правовые основы изъятия горных участков и связанных с ними вспомогательных прав (временное пользование, сервитуты, преимущества при предоставлении концессий на водопользование, а также положения о продлении аренды). Суд постановил, что Указ не нарушает принцип неретроактивности, поскольку затронутые положения касались лишь ожиданий будущих прав, а не приобретенных прав.

Последствия для оценщиков. Данное решение имеет важное значение для оценки прав на добычу полезных ископаемых и горнодобывающих объектов в Мексике. То, что суд квалифицировал права на расширение горнодобывающих участков и связанные с ними положения как «простые ожидания», а не приобретенные права, влияет на оценку горнодобывающих концессий и анализ регуляторного риска при их оценке. Оценщики, готовящие отчеты об оценке горнодобывающих объектов недвижимости или связанных с ними прав в Мексике, должны четко учитывать аспект регуляторного риска при оценке и избегать трактовки условных или зависящих от продления прав значений так, как если бы они были полностью закрепленными.