

mgcvUI: mgcv (一般化加法モデル) パッケージのための 対話型インターフェース

日本語訳。人工知能の支援を得て翻訳し、編集部が校閲した。英語版が正本である。

William Bert Craytor

第 1 巻第 1 号 · 2026 年夏号

発行日: 2026 年 7 月 1 日

DOI: [10.67330/xtv1pa25](https://doi.org/10.67330/xtv1pa25)

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

本論文はオープンアクセス論文であり、[クリエイティブ・コモンズ表示 4.0 国際ライセンス \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) の条件に従って配布される。

要旨. mgcvUI は、R mgcv パッケージのためのグラフィカルユーザインターフェース (graphical user interface) である。mgcv は、自動的な平滑度選択 (smoothness selection) を備えた罰則付き回帰スプライン (penalized regression spline) によって一般化加法モデル (GAM, generalized additive model) を当てはめる。mgcvUI は、一般的な予測モデリング、不動産の鑑定評価、市場地域分析 (market-area analysis) という三つの用途モードを提供し、データの取り込み、平滑化項 (smooth term) の指定、モデルの当てはめ、診断プロットと効果プロットの確認、そしてレポートのダウンロードに至るまで利用者を導く。本論文は、mgcvUI のデータ形式要件、モデリングの作業手順、出力画面、および全機能のリファレンスを文書化する。

キーワード: 一般化加法モデル, GAM, 罰則付きスプライン, 平滑化, mgcv, R, グラフィカルユーザインターフェース, 不動産評価

発行者注記 (2026 年 7 月): 印刷直前に、mgcvUI は earthUI (バージョン 0.11.0 以降) に統合され、本論文で解説する GAM の作業手順は、MARS およびエラスティックネットの作業手順とともに earthUI から提供されるようになった。mgcvUI は最終リリースの形態 (0.3.0) で <https://github.com/wcraytor/mgcvUI> に引き続き公開されており、今後の開発は earthUI において続けられる。本論文は、リリースされた状態の mgcvUI を対象としている。

目次

1	序論	4
1.1	沿革と背景	4
1.1.1	一般化加法モデル	4
1.1.2	mgcv パッケージ	4
1.1.3	MARS (earth) およびエラスティックネット (glmnet) との比較	5
1.1.4	earthUI-mgcvUI パイプライン	5
1.1.5	推奨される作業手順	6
1.2	mgcvUI とは何か	6
1.3	三つの用途モード	6
1.4	不動産に固有の機能	7
1.5	使い始める	7
2	MLS 入力データの要件	9
2.1	ファイル形式と構造	9
2.2	鑑定評価モードに必要な列	9
2.3	特殊列の命名規約	10
2.4	データの品質と完全性	10
2.5	対象行の配置	10

3	General 用途モード	11
3.1	概観	11
3.2	サイドバーの作業手順	11
3.3	earthUI からの取り込み (任意)	12
3.4	主パネルのタブ	12
3.5	設定の保存	13
3.6	ダークモード	13
3.7	ロケールと地域の設定	13
4	鑑定評価モード	15
4.1	対象行の取扱い	15
4.2	基準日と経過日数	15
4.3	特殊列の指定	15
4.4	RCA 補正の概観	15
5	市場地域分析モード	16
5.1	鑑定評価モードとの相違点	16
5.2	市場モードを用いる場面	16
6	変数の選択	17
6.1	目的変数	17
6.2	応答変換	17
6.3	予測変数表	17
6.4	データ型の検出と上書き	18
6.5	特殊列型のリファレンス	18
7	パラメータの選択	19
7.1	パラメータのプリセット	19
7.2	Family (分布族)	19
7.3	Method (推定法)	19
7.4	Gamma (平滑化乗数)	20
7.5	交差検証	20
7.6	Select (縮小)	20
7.7	既定の基底	20
7.8	既定の k (基底次元)	20
7.9	テンソルの型	21
7.10	許可する交互作用	21
7.11	因子別平滑化項の交互作用	21
7.12	高度なパラメータ	21
8	モデルの当てはめ	22
8.1	当てはめボタン	22
8.2	当てはめの際に起こること	22
8.3	背景での当てはめと進捗	22
9	結果のタブ	23
9.1	Data	23
9.2	Equation	23
9.3	Summary	23
9.4	Variable Importance	23
9.5	寄与プロット	24
9.5.1	一変量の平滑化項 — $s(x)$	24

9.5.2	因子別平滑化項 — $s(x, \text{by} = \text{factor})$	24
9.5.3	交互作用項 — $ti(x_1, x_2)$ および $te(x_1, x_2)$	24
9.5.4	パラメトリック項 — 線形および因子	24
9.6	Correlation	24
9.7	Diagnostics	24
9.8	RCA Adjustments	24
9.9	Anova	25
9.10	Mgev Output	25
9.11	Sign Check	25
9.12	Concurvity	25
9.13	Basis Check	25
10	データのダウンロード	26
10.1	出力の列	26
10.2	CQA スコア	26
11	RCA の計算とダウンロード	27
11.1	RCA ダイアログを開く	27
11.2	CQA スコアの内挿	27
11.3	出力の列	27
12	取引事例比較表	28
12.1	取引事例選択ダイアログ	28
12.2	ワークブックの構成	28
12.3	シートの保護と残余的なセル	28
13	レポートのダウンロード	29
13.1	第 1 段階: Quarto レポートの生成	29
13.2	第 2 段階: Quarto レポートの変換	29
13.3	レポートの内容	29
14	モデル関数の書き出し	31
14.1	R 関数	31
14.2	Python 関数	31
14.3	C++ ヘッダ	31
14.4	SQLite データベース	31
15	earthUI および glmnetUI との比較	32
15.1	主要な相違点	32
15.2	共通する機能	32
15.3	推奨される作業手順	32
16	デモ用データセット: Appraisal_1.csv	34
16.1	説明	34
16.2	列	34
16.3	推奨される手早い始め方	35
17	システム要件と問題解決	36
17.1	対応プラットフォーム	36
17.2	R のバージョン	36
17.3	必要な R パッケージ	36
17.4	任意の依存関係	37
17.5	プラットフォーム固有の注意事項	37

17.5.1	macOS	37
17.5.2	Windows	37
17.5.3	Linux	37
17.6	機能の穏やかな縮退	38
17.7	問題解決	38
17.7.1	“PDF option is not available”	38
17.7.2	“Settings will not persist”	38
17.7.3	“Diagnostics plot unavailable”	39
17.7.4	“Report error: number of items to replace...”	39
17.7.5	ポート 7880 が使用中	39
17.7.6	Linux でパッケージの導入が失敗する	39

1 序論

1.1 沿革と背景

1.1.1 一般化加法モデル

一般化加法モデル (GAM, generalized additive models) は、Trevor Hastie と Robert Tibshirani が 1990 年の著書において定式化した。GAM は、線形モデルの各線形項 $\beta_j x_j$ を平滑関数 $f_j(x_j)$ で置き換えることにより線形モデルを拡張する。すなわちモデルは次の形になる。

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_p(x_p) + \varepsilon$$

各 f_j は、罰則付き回帰スプラインを用いてデータから推定される。平滑関数により、利用者があらかじめ関数形を指定することなく、非線形の関係を捉えることが可能になる。罰則は各平滑化項のうねり具合を制御し、実在するパターンを追跡するのに十分な柔軟性を保ちながら過適合を防ぐ。

1.1.2 mgcv パッケージ

R パッケージ mgcv (Mixed GAM Computation Vehicle) は、バース大学の Simon Wood が開発した (Wood 2017)。2001 年に最初に公開され、現在では GAM の標準的な R 実装であり、すべての R インストールに同梱されている。その基礎となる方法は、安定した複数平滑化パラメータの推定 (Wood 2004)、薄板回帰スプライン (Wood 2003)、高速かつ安定な REML 推定 (Wood 2011)、および一般の平滑モデル選択 (Wood, Pya, and Säfken 2016) を扱う一連の論文に記述されている。主要な機能は以下のとおりである。

- **自動的な平滑度選択:** 各項の平滑化パラメータは、REML (制限付き最尤法, Restricted Maximum Likelihood)、GCV (一般化交差検証, Generalized Cross-Validation)、または ML (最尤法, Maximum Likelihood) を用いて、モデルの当てはめの一部として推定される。
- **複数の基底の型:** 薄板回帰スプライン (tp)、三次回帰スプライン (cr)、P スプライン (ps)、B スプライン (bs) があり、それぞれ性質が異なる。
- **テンソル積平滑化項:** `te()` と `ti()` により、次元の呪いを招くことなく変数間の平滑な交互作用を扱える。
- **変数選択:** `select = TRUE` オプションは追加の罰則を課し、平滑化項をゼロまで縮小しうるため、事実上の変数選択となる。
- **コンカービティ (concurvity) 診断:** 平滑化項における多重共線性の類似物を測る。

1.1.3 MARS (earth) およびエラスティックネット (glmnet) との比較

earth アプリケーションとその二つの後処理モジュール (Post Processing Module) (glmnet と mgcv) は、異なるモデリングエンジンを用いる。以下の表に主要な相違点をまとめる。

項目	mgcv (GAM)	earth (MARS)	glmnet (エラスティックネット)
モデルの型	平滑な非線形 (罰則付きスプライン)	区分線形 (適応スプライン)	線形 (正則化つき)
非線形性	平滑関数 (薄板、三次など) により自動的に	データ駆動のノットをもつヒンジ関数により自動的に	交互作用または基底展開を通じてのみ
変数選択	二重罰則による縮小で任意に選択可 (select=TRUE)	GCV による枝刈りを伴う前進/後退逐次法として内蔵	L_1 罰則 (ラッソ) として内蔵
交互作用	テンソル積平滑化項 (te(), ti())	ヒンジ関数の積 (解積可能、次数 3 まで)	対ごとの交差積 (解積が困難)
解釈可能性	平滑な偏効果曲線であり、きわめて直観的	g 関数が変数ごとの区分線形な偏効果を示す	係数が線形の重みであり、個別に説明しやすい
適する用途	平滑な非線形関係、柔軟なモデリング	非線形性と交互作用の自動的な検出	高次元データ、変数選択、小標本
RCA 補正	偏効果による平滑な補正	g 関数による区分線形な補正	変数ごとの線形な補正
過適合の制御	REML/GCV による平滑化パラメータの推定	GCV に基づく枝刈り	交差検証されたラムダ

MARS (多変量適応回帰スプライン, Multivariate Adaptive Regression Splines) は、1991 年に Jerome Friedman が導入した。これは、ヒンジ関数とそのノット位置をデータから適応的に選択することにより区分線形モデルを構築する。R 実装は Stephen Milborrow による earth パッケージである。

エラスティックネット回帰は、Friedman、Hastie、Tibshirani (2010) による glmnet パッケージに実装されており、ラッソ (L_1) とリッジ (L_2) の罰則を組み合わせ、変数選択と係数の縮小を同時に行う。

1.1.4 earthUI–mgcvUI パイプライン

earthUI の自動的なノット発見と mgcvUI の平滑推定を組み合わせると、強力な作業手順が得られる。earthUI が .rds 結果ファイルを書き出し、mgcvUI がそれを取り込むと、次のことが起こる。

1. **earth がノットを発見する** — MARS は各変数についてデータ駆動のヒンジ位置 (変化点) を見いだす。
2. **mgcvUI がそれを錨として用いる** — earth のノットは GAM における三次回帰スプライン (cr 基底) の基礎となり、データの実際の構造を反映したノット位置から平滑化項を出発させる。
3. **mgcv が当てはめを洗練する** — 罰則付きスプライン推定が earth の区分線形な当てはめを平滑化し、earth が発見した全体的な形状を保ちながら平滑な偏効果曲線を生成する。

このパイプラインは探索的なモデリング (earth) と検証的なモデリング (GAM) とを橋渡しし、両者の長所を併せもつ。

1.1.5 推奨される作業手順

不動産の鑑定評価をはじめ、解釈可能で防御可能なモデルを要する用途では、次の順序を推奨する。

1. **まず earthUI から始める。** データの中の重要な変数、非線形の関係、交互作用を発見するためである。earth の自動的な基底選択と g 関数は、優れた初期の洞察を与える。
2. **より平滑な偏効果が欲しければ mgcvUI で洗練する。** earthUI のノットを mgcvUI に取り込めば、区分線形なモデルから平滑なモデルへ滞りなく移行できる。
3. **glmnetUI を用いるのは、積極的な変数選択（ラッソ）が必要なとき、予測変数の数が観測数に近づくか上回るとき、あるいは防御可能性のために正則化された純粋な線形モデルが欲しいときである。**

1.2 mgcvUI とは何か

mgcvUI は、R の mgcv パッケージのためのグラフィカルユーザーインターフェースである。ローカルの Shiny アプリケーションとして動作し、ログインもサーバもアカウントも存在しない。R から起動し、データセット（CSV または Excel）を取り込み、モデルを構成して、対話的に GAM の当てはめを行う。

現行リリースの時点で、mgcvUI はもはや別個にインストールされるパッケージではない。earthUI パッケージの内部に GAM 後処理モジュールとして同梱されており、`earthUI::launch_mgcv()` で起動する。アプリケーション自体は—そして本論文で述べる事柄はすべて—変わっていない。異なるのはインストールと起動コマンドだけである。三つのルーチン (earth、glmnet、mgcv) は、一つのプロジェクトツリー、一つの設定データベース、そして支援基盤の一つの複製を共有する。

本アプリケーションは、データの取り込み、平滑化項の指定を含む変数の構成、背景処理によるモデルの当てはめ、診断プロット、平滑な偏効果曲線、そして Word、PDF、HTML 形式でダウンロードできるレポートに至る、完全な作業手順を提供する。

一般化加法モデルは、各予測変数について線形の関係 βx を平滑関数 $f(x)$ で置き換える。これが意味するところは次のとおりである。

- **関数形を指定する必要がない** — 各関係の形状はデータが定める。
- **平滑な偏効果** — 各予測変数の寄与は直線ではなく平滑な曲線となるため、ある変数がある値の違いに応じて応答にどう影響するかを容易に見て取れる。
- **自動的な平滑度選択** — mgcv は、当てはめと過適合とを釣り合わせながら、各曲線をどれだけうねらせるべきかを REML または GCV によって定める。

1.3 三つの用途モード

すべてのプロジェクトは**用途** (purpose) をもつ。これはサイドバーの第 1 節 (**Project**) でプロジェクトを作成する際に選択し、以後は固定される。有効なプロジェクトの用途によって、どのツールと インターフェース要素が現れるかが決まる。別のモードで作業したいときは、その用途をもつプロジェクトを作成するか開けばよい。モードは三つである。

- **General** (一般) — あらゆる種類の母集団またはデータセットに対する GAM 回帰。既定のモードである。領域固有の追加機能をもたない、完全な GAM の作業手順を提供する。
- **For Appraisal** (鑑定評価用) — 不動産の鑑定評価に合わせた GAM 回帰。対象不動産の取扱い、特殊列の指定、残差制約アプローチ (RCA, Residual Constraint Approach) など、単一不動産の評価に固有の機能を追加する。
- **Market Area Analysis** (市場地域分析) — 市場地域の研究に合わせた GAM 回帰。画定された市場内の不動産群を分析するための機能を追加し、対象行の取扱いも任意に利用できる。

三つのモードのいずれにおいても、中核のモデリングエンジンは同一であり、利用者は常に `mgcv::gam()` によって GAM を当てはめている。用途の設定は、どの追加ツールとインターフェース要素が利用可能になるかを制御するにすぎない。

1.4 不動産に固有の機能

For Appraisal または Market Area Analysis を選択すると、mgcvUI は不動産分析のために設計されたいくつかの機能を有効にする。

- **特殊列の指定** — 各予測変数には、`contract_date`、`dom`、`concessions`、`latitude`、`longitude`、`living_area`、`lot_size`、`actual_age`、`effective_age`、`area`、`site_dimensions`、`display_only` といった特殊な役割を付与できる。これらの指定は、当てはめと出力において当該列がどのように扱われるかを制御する。
- **Sale Age 列** — ある列が `contract_date` として指定され、かつ基準日 (Effective Date) が与えられている場合、mgcvUI は `sale_age` 列 (売買日と基準日との間の日数) を計算し、これを予測変数として代置する。
- **RCA 計算 (鑑定評価モード専用)** — 鑑定評価モードでは、モデルを当てはめた後に mgcvUI が残差制約アプローチ (RCA) の出力を計算できる。これは取引事例ごとの補正、純補正／総補正の要約、および対象不動産の補正後売買価格を生成する。

ヒント: General 用途モードは、金融、科学、工学、不動産のいずれのデータセットにも通用する。鑑定評価モードと市場モードは、不動産の専門家のための便宜的な機能を加えるにすぎない。

1.5 使い始める

mgcvUI を使うには次のようにする。

1. R で **earthUI をインストールする**: `install.packages("earthUI")` — mgcv ルーチンはこれに含まれている。(開発版と任意の構成要素については、本号の earthUI 論文中の “Installing earthUI and vProlog” を参照せよ。)
2. **アプリケーションを起動する**: R で `earthUI::launch_mgcv()` を実行する。アプリケーションはポート 7880 でウェブブラウザに開く。
3. サイドバーの第 1 節で**プロジェクトを作成または開く**。プロジェクトは用途 (General、For Appraisal、Market Area Analysis) と作業地 (国、州、郡、市)、およびその入出力フォルダを定める。プロジェクトは共有された `regProj` ルートの下に保存され、姉妹アプリケーションである earthUI および glmnetUI と共有される。モードを切り替えるには、異なる用途のプロジェクトを開けばよい。
4. 第 2 節で、有効なプロジェクトの入力フォルダから CSV または Excel ファイルを選んで**データを取り込む**。当該フォルダにファイルを追加した後、**Refresh** をクリックして一覧に表示させる。
5. **earthUI から取り込む (任意)** — 第 3 節では、earthUI の結果 `.rds` ファイルを取り込み、earth が発見したノットで GAM を初期化できる。earthUI の結果がない場合はこの手順を飛ばしてよい。
6. **変数を構成する** — 目的変数と予測変数を選び、データ型を設定し、特殊列の役割を割り当て、変数を因子または線形として指定する。
7. **mgcv のパラメータを設定する** — `family`、`method`、`gamma`、基底の型、`k`、テンソルの型、交互作用、その他のオプションである。
8. **モデルを当てはめる** — “Fit Mgcvm GAM Model” をクリックし、主パネルで結果を確認する。
9. **書き出す** — 予測値を Excel としてダウンロードし、Quarto レポートを生成して変換し、あるいは (鑑定評価モードでは) RCA 補正と取引事例比較表を計算する。

設定は SQLite データベースに自動的に保存され、同じ入力ファイルを再度読み込んだときに復元される。

2 MLS 入力データの要件

不動産の鑑定評価および市場分析の作業手順では、入力データは通常、マルチプル・リスティング・サービス (MLS) からの書き出しである。本章では、想定されるファイル構造と、mgcvUI が利用できる列について述べる。

2.1 ファイル形式と構造

mgcvUI は CSV および Excel (.xlsx, .xls) ファイルを受け付ける。取り込みの際、列名は自動的に snake_case に変換される。たとえば “Living SqFt” は living_sqft に、“Contract Date” は contract_date に、“Sale Price” は sale_price になる。この正規化によって、作業手順の全体を通じて列の参照が一貫したものになる。取り込みの際に用いられる CSV の区切り記号と小数点記号は、ロケール設定によって定まる（第 3 章 “Locale & Regional Settings” を参照）。

データファイルは、1 不動産につき 1 行、1 属性につき 1 列の平坦な表であるべきである。ファイルの第 1 行には列見出しが必要である。

2.2 鑑定評価モードで必要な列

mgcvUI はどのような列の組み合わせでも動作するが、鑑定評価の完全な作業手順は、以下の列があることによって効果を発揮する。

列	特殊型	用途
Sale Price	(目的変数)	モデルの応答変数
Contract Date	contract_date	sale_age (基準日からの日数) の計算に用いる
Listing Date	listing_date	DOM 列がない場合、契約日と併せて DOM を計算するのに用いる
Days on Market	dom	市場滞留日数。書き出しに表示される
Concessions	concessions	売買時の譲歩。純売買価格 = 売買価格 - 譲歩
Living Area (SF)	living_area	平方フィート当たり残差 (residual_sf, cqa_sf) を可能にする
Lot Size	lot_size	敷地規模の列
Site Dimensions	site_dimensions	敷地規模と同じ群に扱われる
Latitude	latitude	小数点以下 3 桁に丸められる
Longitude	longitude	小数点以下 3 桁に丸められる
Area ID	area	市場地域／近隣の識別子
Actual Age	actual_age	不動産の実際経過年数
Effective Age	effective_age	不動産の有効経過年数
Address	display_only	書き出しに表示され、モデルからは除外される

表計算ソフトの列名は外国語であってもよい。“特殊” な名称が英語であるのは、R のプログラムがそれらに特別な扱いを与えられるようにするためである。それ以外の場合、与えられた列名がそのまま回帰モデル、グラフ、そして鑑定評価を行うならば出力レポートに現れる。

すべての列が必須なわけではない。mgcvUI は適応的であり、ある列が欠けていれば、対応する機能が単に省かれるだけである。ただし、不動産の価格モデルについては、いくつかの列を強く推奨する。

1. **Sale Age** — 契約売買日と基準日との間の日数。複数年にわたる売買履歴を用いる場合、sale_age はしばしば中心的な役割を果たす。
2. **Living Area** — “Living Sqft”、“GLA” (総居住面積) といった名称でも呼ばれる。

3. **Total Bath Count** — フル、4 分の 1、2 分の 1、4 分の 3 の各浴室を合計した数。
4. **Garage Bays** または **Garage Area** — 車庫の台数または面積。
5. **Lot Size** — 当該不動産の土地面積。
6. 地理的な価格変動のための **Longitude**、**Latitude**、および利用できるならば **Area ID**。

2.3 特殊列の命名規約

mgcvUI は列を、その列名ではなく**特殊型の指定**によって識別する。MLS の書き出しにおいて列にはどのような名前を付けてもよい。重要なのは、変数構成表（第 6 章）において正しい特殊型を割り当てることである。

2.4 データの品質と完全性

- **欠測値 (NA)**: 予測変数または目的変数の列に NA を含む行は、当てはめの前に自動的に除去される。mgcvUI はデータをモデルで用いる列だけに絞込むため、使用しない列の NA は行の脱落を引き起こさない。
- **数値列**: 売買価格、居住面積、敷地規模などの項目は数値でなければならない。
- **因子列**: テキスト列、TRUE/FALSE 列、および R の factor 列は自動的に質的変数として扱われ、因子としてのみモデルに入りうる。**数値**の列（たとえば数値の地域コードや様式コード）が因子になるのは、その **Factor** チェックボックスを選択したときだけである。mgcvUI は数値の値域から質的性を推測することは決してないため、**bath_count** のような離散的な数値の予測変数は連続量（平滑化される）のままとなる。Special の役割は因子の判断とは独立である。

ヒント: 取り込み後は予測変数表の NA 列を確認せよ。欠測値の多い列は行の脱落を招きうる。NA が $\geq 50\%$ の列は赤で強調され、自動的に除外される。

2.5 対象行の配置

Appraisal（鑑定評価）モードでは、**第 1 行が対象不動産でなければならない**。残りの行はすべて比較対象となる取引事例である。対象行はモデルの当てはめから除外される。当てはめの後も、モデルは対象行についての予測値を生成する。

Market Area Analysis（市場地域分析）モードでは、対象を第 1 行に置くことは任意である。

General（一般）モードでは特別な行の扱いはなく、すべての行が同等に扱われる。

3 General 用途モード

3.1 概観

General Purpose (一般用途) モードは、mgcvUI を起動したときの既定である。これは不動産に限らず、あらゆるデータセットに対して完全な GAM の作業手順を提供する。科学データ、財務分析、工学的研究、あるいは平滑な非線形関係が見込まれるあらゆる回帰問題に mgcvUI を用いることができる。

General モードでは、インターフェースは不動産に固有の機能 (特殊列、経過日数、RCA) を省く。サイドバーは、変数選択、パラメータ構成、モデルの当てはめ、書き出しに焦点を絞って簡素化される。

3.2 サイドバーの作業手順

サイドバーは、番号を付された折りたたみ可能な節に整理されており、プロジェクトの選択からレポートの書き出しまで利用者を導く。ほとんどの節は有効なプロジェクトを開いてはじめて現れ、いくつかはプロジェクトの用途に固有である。出力フォルダは有効なプロジェクトから導かれるので、出力フォルダの独立した入力欄は存在しない。

1. Project — プロジェクトを作成または開く。プロジェクトは用途と作業地 (国、州、郡、市)、およびその入力フォルダ (`<os>_in/`) と出力フォルダを定める。プロジェクトは共有された `regProj` ルート (既定は `~/regProj`。REGPROJ_ROOT 環境変数または Settings のドロップダウンで変更できる) の下に保存され、姉妹アプリケーションである earthUI および glmnetUI と共有されるため、三つすべてのモデルが並んで存在する。切り替えるには **Close Project** をクリックする。

2. Import Data — 有効なプロジェクトの入力フォルダから CSV または Excel ファイルを選ぶ。Finder / エクスプローラでフォルダにファイルを追加し、**Refresh** をクリックして一覧に表示させる。複数のシートをもつ Excel ファイルでは、シート選択欄が現れる。列名は自動的に `snake_case` に変換される。

3. Import from earthUI (optional) — earthUI の結果 `.rds` ファイルを取り込み、earth が発見したノット位置で GAM を初期化する (併せて変数の選択も初期化される)。この手順は任意であり、取り込むべき earthUI の結果がなければ飛ばしてよい。

4. Variable Configuration — 目的変数の選択欄、応答変換 (`none`, `log`, `log10`)、および Include、Factor、Linear のチェックボックスをもつ予測変数表である。Special 列は鑑定評価モードと市場モードでのみ現れる。詳細は第 6 章を参照せよ。

5. Mgcvc Call Parameters — モデル構成のすべてである。パラメータのプリセット、family、method、gamma、交差検証、select、基底の型、k、テンソルの型、交互作用の行列、および高度なパラメータを含む。パラメータの完全なリファレンスは第 7 章を参照せよ。

6. Fit Mgcvc GAM Model — モデルを実行するボタンである。結果は右側のタブに現れる。

7. Download Output — 予測値、残差、CQA スコア、および変数ごとの寄与を、Excel ファイルとしてプロジェクトの出力フォルダに書き出す。モデルを当てはめた後は、いずれの用途モードでも利用できる。

8. Calculate RCA Adjustments & Download — 各取引事例について変数ごとの RCA 補正を計算し、対象の残差を CQA によって内挿する。鑑定評価モードと市場モードで、当てはめの後にのみ現れる。

9. Generate Sales Grid & Download — 取引事例比較表を生成し、総補正率によって取引事例を推奨する。鑑定評価モードでのみ現れる。

10. Generate Quarto Report — すべてのプロットと表を含む自己完結的な Quarto (`.qmd`) レポート一式を、プロジェクトの出力フォルダに書き出す。モデルを当てはめた後に利用できる。

11. Convert Quarto Report — 生成された .qmd 一式を HTML、Word、または PDF に変換する。PDF には LaTeX の導入が必要であり、検出されない場合は PDF の選択肢は表示されない。第 13 章を参照せよ。

用途に固有の手順は他のモードでは省かれ（残りの手順は番号が振り直される）。市場モードでは取引事例比較表の手順がなく、General モードでは RCA と取引事例比較表の双方の手順がない。

3.3 earthUI からの取り込み（任意）

サイドバーの第 3 節 — **Import from earthUI** — では、earthUI の結果 .rds ファイルを取り込むことができる。これにより earth-mgcv パイプラインが可能になる。すなわち、earth のデータ駆動のノット位置が GAM の平滑化項の錨点となるのである。

earthUI の結果を取り込むと、mgcvUI は次のことを行う。

- 目的変数、決定係数、項の数、変数ごとのノット位置を示す要約を表示する。
- “Earth Pipeline” パラメータプリセット（三次回帰スプライン基底、 $\gamma = 1.4$ ）を自動的に選択する。
- 目的変数を earth の目的変数に一致させる。
- earth の予測変数について Include をあらかじめ選択する。
- earth の質的変数について Factor を、earth の線形の予測変数について Linear をあらかじめ選択する。
- earth が検出した交互作用をあらかじめ設定する（交互作用の行列で黄色に強調され、手動での変更は禁じられる）。
- earth のノットを CSV ファイルとしてダウンロードする **Export Knots CSV** ボタンを提供する。

すると変数構成表には、変数の選択 (Include, Linear, Factor) とスプラインのノットが earth モデルから初期化された旨の注記が現れる。glmnetUI では glmnet の当てはめが earth の基底に束縛されるのに対し、mgcvUI は利用者の現在の選択から GAM を当てはめる。earth は出発点を与えるにすぎず、どの制御も編集可能なままである。

Clear ボタンは、取り込んだ earth のデータを削除し、単独動作のモードに戻す。

3.4 主パネルのタブ

データを取り込むと、主パネルは以下のタブを提供する（モデルに依存するタブは当てはめの後に内容が満たされる）。

タブ	内容
Data	取り込んだデータの対話的な DataTable。鑑定評価モードと市場モードでは、対象不動産の表と比較取引事例の表に分かれる。
Equation	MathJax で描画された GAM の完全な式と、平滑関数の定義の表。
Summary	主要な指標 (R^2 、 $CV R^2$ 、Dev. Explained、AIC、n、平滑化項の数) と、平滑化項／パラメトリック項の表。
Variable Importance	F 統計量 (平滑化項) と $ t $ 値 (パラメトリック項) の横棒グラフ、および順位付けした表。
Contribution	信頼帯と earth のノット標識を伴う、対話的な plotly の平滑な偏効果曲線。
Correlation	数値の予測変数間の相関のヒートマップ (当てはめの前でも利用できる)。
Diagnostics	残差対当てはめ値、Q-Q プロット、ヒストグラム、応答対当てはめ値 (gratia による)。実測値対予測値の散布図。
RCA Adjustments	RCA 分析のヒストグラム (鑑定評価／市場モードで、計算の後)。
Anova	パラメトリック項と平滑化項を統合した分散分析表。
MgcV Output	生のテキスト。所要時間、式、モデルの要約、family、コンカービティ。
Sign Check	earth と GAM の方向の整合性の検査 (earth を取り込んだ場合)。
Concurvity	全体および対ごとのコンカービティ行列を色分けして表示。
Basis Check	<code>gam.check()</code> による基底次元の妥当性の検定。

3.5 設定の保存

mgcvUI は、入力ファイル名を鍵として構成を SQLite データベースに自動的に保存する。同じファイルを再度読み込むと、目的変数の選択、予測変数のチェックボックス、データ型、mgcv のパラメータ、交互作用の行列といった設定がすべて復元される。用途は大域的な設定ではなく有効なプロジェクトの属性であり、最後に用いた入力ファイルはプロジェクトごとに記憶され、プロジェクトを再び開いたときに自動的に読み込まれる。

3.6 ダークモード

右上隅の月／太陽のアイコンをクリックすると、Nord Light と Nord Dark のテーマを切り替えられる。テーマの設定は localStorage に保存され、セッションをまたいで持続する。すべての UI 要素が選択されたテーマに適応する。

3.7 ロケールと地域の設定

mgcvUI は、国に基づくロケールの仕組みによって、数値と CSV の書式に関する国際的な慣行に対応する。タイトルバーの **Settings** ドロップダウンは、30 を超える対応国について **Country** と **Paper** の選択欄を提供する。各プリセットは次の項目を設定する。

- **CSV の区切り記号** — 米国／英国／日本ではコンマ (,)、欧州の大半ではセミコロン (;)。
- **小数点記号** — ピリオド (.) またはコンマ (,)。
- **桁区切り記号** — コンマ (米国／英国／日本)、ピリオド (ドイツ／イタリア／スペイン)、空白 (フィンランド／フランス／ポーランド／バルト諸国)、アポストロフィ (スイス)。
- **用紙寸法** — Letter (米国／カナダ／メキシコ) または A4 (それ以外のすべて)。

Save をクリックするとロケールの設定が大域的に保存される。Import Data の節にも、異なる慣行を用いる単発の取り込みのために、ファイルごとの **Locale** 選択欄がある。同じ

Settings ドロップダウンには **regProj root folder** の欄もあり、mgcvUI、earthUI、glmnetUI が共有するプロジェクトツリーの位置を変更できる。

ヒント: プロットの軸と傾きのラベルにおける数値の書式は、ロケールに自動的に適応する。ドイツのロケールは桁区切りにピリオドを用い (200.000)、フィンランドは空白を (200 000)、スイスはアポストロフィを用いる (200'000)。通貨記号は表示されない。mgcvUI は通貨に依存しないからである。

4 鑑定評価モード

有効なプロジェクトの用途が **For Appraisal** である場合、mgcvUI は単一不動産の評価に向けて自らを構成する。第 3 章で述べた機能はすべて利用できる。本章では鑑定評価に固有の追加事項のみを扱う。

4.1 対象行の取扱い

鑑定評価モードでは、データセットの第 1 行が対象不動産であり、残りの行はすべて比較取引事例である。入力ファイルはそうのように整理されていなければならない（第 2 章を参照）。対象の売買価格は空欄のままだでも、任意の値を入れてもよい。mgcvUI は当てはめの際にこれを自動的に NA として扱う。

当てはめの後、モデルは対象行についての予測値を生成する。これが RCA 補正の作業手順の基礎となる。

4.2 基準日と経過日数

鑑定評価モードと市場モードでは、変数構成の節に **Effective Date**（基準日）の欄が現れる（既定は当日の日付）。Special ドロップダウンである列を `contract_date` と指定すると、mgcvUI は `sale_age` 列 — 各売買の契約日と基準日との間の整数日数 — を計算する。この列が予測変数として追加される。

基準日を変更すると、`sale_age` は自動的に再計算される。

4.3 特殊列の指定

鑑定評価モードと市場モードでは、変数構成表の各予測変数に対して **Special** ドロップダウンが現れる。特殊型とその効果の完全なリファレンスは第 6 章を参照せよ。

4.4 RCA 補正の概観

Calculate RCA Adjustments & Download ボタン（鑑定評価モードと市場モードで当てはめの後に表示される）は、対象に対する各取引事例の市場から導かれた補正を計算する。RCA の完全な作業手順は第 11 章に述べる。

ヒント: 対象不動産に与える *CQA* スコアは、残差分布のどれだけを対象不動産に帰属させるかを制御する。5.00 のスコアは、対象不動産を取引事例の中央値に置く。

5 市場地域分析モード

有効なプロジェクトの用途が **Market Area Analysis** である場合、mgcvUI は鑑定評価モードと同じ不動産に固有の機能（特殊列、経過日数）を提供するが、単一の対象を評価するのではなく不動産群を分析することに向けられている。

5.1 鑑定評価モードとの相違点

- 対象行の取扱いが柔軟である — 市場モードでは、第 1 行が自動的に対象不動産として扱われることはない。
- 取引事例比較表がない — “Generate Sales Grid & Download” の手順は利用できない。

5.2 市場モードを用いる場面

市場地域分析モードが適するのは、次のような場合である。

- 非線形の価値要因を理解するために、近隣または市場地域について GAM モデルを構築するとき
- 面積、築年数、敷地規模、位置といった変数が不動産群の売買価格にどう影響するかを分析するとき
- 市況分析または近隣の画定のための裏づけを準備するとき
- 特殊列の機能（経過日数）は使いたいが取引事例比較表の作業手順は必要ないデータセットを扱うとき

ヒント: 市場モードは、特殊列の機能は使いたい但し単一不動産の評価の作業手順は必要ないという、一般的な不動産回帰にも有用である。

6 変数の選択

サイドバーの第 4 節 — **Variable Configuration** — は、どの列をモデルに参加させ、それらをどう扱うかを選ぶ場所である。

6.1 目的変数

第 4 節の最上部にある **Target (response) variable** ドロップダウンには、データセット中のすべての数値列が並ぶ。応答変数として一つの列（たとえば売買価格）を選ぶ。目的変数の列は予測変数の一覧から自動的に除外される。

6.2 応答変換

目的変数の選択欄の下にある **Response Transform** ドロップダウンは、三つの選択肢を提供する。

- **None (raw)** — 応答をそのまま用いる。予測変数と応答との関係が元の尺度でほぼ加法的である場合に適する。
- **Log (natural)** — 当てはめの前に自然対数変換を施す。逆変換したとき、補正は絶対額 (\$) ではなく比率 (%) になる。売買価格にはとりわけ有用である。\$10,000 の補正は価格水準が異なればまったく異なる意味をもつからである。
- **Log10** — 底 10 の対数変換を施す。

ヒント: 対数変換は時点修正を絶対額 (\$) ではなく比率 (%) にする。広い価格帯にわたる不動産モデルでは、対数変換がしばしばより素直な残差をもたらす。

対数変換を選択した場合、応答の値のうち ≤ 0 のものは自動的に除去される。出力（予測値、寄与、残差）はすべて自動的に元の尺度へ逆変換される。

6.3 予測変数表

目的変数の選択欄の下には、残りのすべての列を並べたスクロール可能な表がある。チェックボックスの列は、簡潔さのために縦向きに回転した見出しを用いる。列の順序は次のとおりである。

列	説明
Variable	列名（等幅フォント、長い場合は省略記号で切り詰め）。枠線付きのセル。
Type	データ型のドロップダウン。numeric, integer, character, factor, logical, Date, POSIXct。枠線付きのドロップダウン。
Include	チェックボックス — この列を予測変数としてモデルに含める。縦向きの見出し。
Factor	チェックボックス — 質的な因子として扱う（水準ごとに別個の係数を作る）。縦向きの見出し。
Linear	チェックボックス — 線形関係を強制する（平滑化項ではなく βx のみ）。縦向きの見出し。
Special	ドロップダウン（鑑定評価／市場モードのみ） — 後述の特殊列型リファレンスを参照。枠線付きのドロップダウン。
NAs	欠測値の数と割合を色分けして表示。赤 ($\geq 50\%$)、橙 ($\geq 20\%$)、灰 (> 0)。枠線付きのセル。

因子・平滑化項・線形の別: 既定では、含められた数値変数は平滑化項 $f(x)$ を得る。**Factor** を選択すると質的な項（水準ごとに一つの係数）が作られる。**Linear** を選択すると、

平滑化項の代わりに単純な線形項 βx が作られる。Factor と Linear の双方を指定した変数は Factor として扱われる。

6.4 データ型の検出と上書き

mgcvUI は取り込みの際にデータ型を自動的に検出する。numeric、integer、logical、factor、および日付の列が認識される。**Type** ドロップダウンを変更すれば、どの検出結果も上書きできる。型を変更すると、その列がモデルの中でどう扱われるかが変わる。

6.5 特殊列型のリファレンス

鑑定評価モードと市場モードでは、**Special** ドロップダウンが以下の選択肢を提供する。

重み付け:

- **weight** — 観測の重みの列（一つのみ指定可。重みが 0 の行は当てはめから除外される）

日付と時刻の型:

- **contract_date** — 基準日からの **sale_age** の自動計算を起動する
- **sale_age** — あらかじめ計算された経過日数の列（基準日からの日数）を示し、そのまま用いられる
- **listing_date** — 市場滞留日数を計算する際の代替として用いられる
- **dom** — 市場滞留日数の列を示す

レコードの種別:

- **sale_type** — 対象と比較取引事例とを区別するレコード種別の列を示す

金額の型:

- **concessions** — 売買時の譲歩（売主の与信、買主への誘因など）を示す

規模と位置の型:

- **latitude** — 値が自動的に小数点以下 3 桁に丸められる
- **longitude** — 緯度と同じ丸めの扱いを受ける
- **area** — 市場地域または近隣の識別子
- **living_area** — 平方フィート当たり残差の計算 (**residual_sf** および **cqa_sf**) を可能にする
- **lot_size** — 敷地規模の列
- **site_dimensions** — 敷地規模と同じ群に扱われる

築年数の型:

- **actual_age** — 不動産の実際経過年数の列
- **effective_age** — 不動産の有効経過年数

表示の型:

- **display_only** — 当該列は Excel への書き出しには含まれるが、モデルの当てはめからは完全に除外される。住所の項目、MLS 番号、地番 ID、その他の参照用データに用いる。

ヒント: **display_only** と指定された列はデータセットに残り、*Excel* への書き出しにも現れるが、予測変数表からは完全に除外される。*ID* 列、住所、その他の参照用の項目に用いるとよい。

7 パラメータの選択

サイドバーの第 5 節 — **Mgcv Call Parameters** — は、GAM の構成オプションのすべてへの入口である。各パラメータには青い補助アイコン (?) が付いており、ツールチップで説明が表示される。

7.1 パラメータのプリセット

最上部のドロップダウンは、よくある作業手順に対して妥当な既定値を設定する二つのプリセットを提供する。

プリセット	基底	設定
Standalone (discovery)	tp (薄板)	k = auto (10), gamma = 1.2, select = TRUE
Earth Pipeline (refinement)	cr (三次回帰)	k = auto (10), gamma = 1.4, select = TRUE

earthUI のノットを取り込むと、**Earth Pipeline** プリセットが自動的に選択される。earth のノットを統合するには三次回帰スプライン (cr) 基底が必要である。この基底のみが正確なノット位置の指定を許すからである。

7.2 Family (分布族)

応答変数の分布族を選ぶ。

Family	用途
gaussian	連続的な応答 (たとえば売買価格)。最も一般的である。
Gamma	分散が平均に比例する正の連続的な応答。
poisson	計数データ (たとえば売買件数)。
binomial	二値の結果。
inverse.gaussian	強く右に歪んだ正のデータ。

7.3 Method (推定法)

平滑化パラメータの推定法である。

Method	説明
REML	制限付き最尤法 (既定、推奨)。過適合に対して最も頑健である。
GCV.Cp	一般化交差検証。やや複雑なモデルを選ぶ傾向がある。
ML	最尤法。REML に似るが、平滑化が不足しうる。
P-REML	ピアソン型の REML の変種。
P-ML	ピアソン型の ML の変種。
GACV.Cp	一般化近似交差検証。

7.4 Gamma (平滑化乗数)

gamma パラメータは平滑化の規準における有効自由度に乘じられ、より平滑（うねりの少ない）当てはめを促す。既定は 1.2 である。

値	効果
1.0	標準的な平滑化
1.2–1.4	既定よりやや平滑（推奨）
2.0+	はるかに平滑。雑音の多いデータや小標本に適する

gamma の値を大きくすると複雑さへの罰則が重くなり、過適合を防ぐ。Earth Pipeline プリセットは、earth のノットを洗練する際の追加的な平滑化のために 1.4 を用いる。

7.5 交差検証

これを選択すると（既定）、mgcvUI は当てはめの後に 10 分割交差検証を実行して CV 決定係数を計算する。これは標本外の予測力についての率直な推定値を与える。

7.6 Select (縮小)

これを選択すると（既定）、各平滑化項に追加の罰則が課され、その項をゼロまで縮小しうる。これにより自動的な変数選択が可能になり、重要でない平滑化項は事実上モデルから取り除かれる。gam() における select = TRUE として実装されている。

7.7 既定の基底

平滑化項に用いるスプライン基底関数の型である。

基底	正式名称	説明
tp	薄板回帰スプライン	既定。等方的（ノット位置の指定が不要）。汎用の選択として良好である。
cr	三次回帰スプライン	ノット位置の明示的な指定を許す。earth のノットの統合に必要である。
ps	P スプライン	差分罰則をもつ等間隔の B スプライン基底。計算効率が高い。
bs	B スプライン	柔軟な B スプライン基底。

7.8 既定の k (基底次元)

基底次元 k は各平滑化項の最大のうねり具合を制御する。値 0（既定）は“自動”を意味し、mgcvUI は $k = 10$ または earth のノットの数のうち適切なほうを用いる。

- 小さい k (3–5): きわめて平滑な曲線で、ほぼ線形になる。
- 中程度の k (8–15): 適度な柔軟性（通常はこれで足りる）。
- 大きい k (20 以上): きわめて柔軟である。強い正則化がなければ過適合の危険がある。

ヒント: Summary タブの有効自由度 (EDF) は、各平滑化項が実際に何自由度を用いているかを示す。EDF が $k - 1$ に近ければ基底が小さすぎるおそれがあるので、 k を増やすとよい。EDF が k よりはるかに小さければ、その平滑化項は十分に罰則が効いている。

7.9 テンソルの型

連続変数どうしの交互作用については、二つのテンソル積の型が利用できる。

型	関数	説明
ti	ti(x1, x2)	テンソル交互作用。交互作用の効果のみをモデル化し、主効果は一変量の $s()$ 項が別途扱う。解釈可能性の点で望ましい。
te	te(x1, x2)	テンソル積。主効果を含む同時効果の全体をモデル化する。RCA 補正のために分解することは難しい。

7.10 許可する交互作用

折りたたみ可能な **Allowed Interactions** の節は、含められたすべての平滑（非線形かつ非因子）予測変数について、上三角のチェックボックス行列を表示する。各チェックボックスは、対応する変数対の間のテンソル積平滑化項を有効にする。

- **Allow All / Clear All** — 一括切替のボタン。
- **変数名をクリック** — その変数についてのすべての交互作用を切り替える。
- **変数名を右クリック** — “Block from main effect” — その変数を交互作用専用（一変量の平滑化項をもたない）として指定する。

earthUI のノットを取り込むと、earth が検出した交互作用があらかじめ選択され、変更が禁じられる（黄色の背景で強調される）。earth モデルが $\text{degree} = 1$ （交互作用なし）を用いていた場合は、その旨の情報が表示される。

Warning

テンソル交互作用を有効にすると、モデルの複雑さは大幅に増す。交互作用は、それを裏づける領域知識または earth が検出した証拠があるときに用いよ。結果として得られる補正が合理的で説明可能であることを検証されたい。

7.11 因子別平滑化項の交互作用

因子変数と平滑変数の双方が含まれている場合、別個の **Factor-by-Smooth Interactions** 行列が現れる。各チェックボックスは $s(x, \text{by} = \text{factor})$ 項を作る。これは因子の水準ごとに別個の平滑曲線を与えるものである。

7.12 高度なパラメータ

折りたたみ可能な “Advanced” の節では、追加の設定を扱える。

- **Optimizer** — 平滑化パラメータ推定のアルゴリズム。outer/newton（既定）、outer/bfgs、または efs（拡張 Fellner-Schall 法）。
- **Scale** — 尺度パラメータ。0（既定）はデータからの推定を意味する。尺度が固定されたモデルでは既知の値を設定する。
- **Discrete covariate method** — 大きなデータセットのための離散化による高速な当てはめ ($\text{discrete} = \text{TRUE}$)。
- **Threads** — 並列スレッドの数 (1–32)。discrete = TRUE のときにのみ効果がある。

8 モデルの当てはめ

8.1 当てはめボタン

サイドバーの第 6 節には **Fit Mgcgv GAM Model** ボタンがある。これをクリックすると、現在の構成でモデルが実行される。

8.2 当てはめの際に起こること

Fit をクリックすると、mgcvUI は次のことを行う。

1. **構成を検証する** — 少なくとも一つの予測変数と、少なくとも 10 行の完全な行があるかを確認する。完全な行が 50% に満たない場合は警告する (NA の大半を生じさせている列を列挙する)。
2. **データを準備する** — モデルに関係する列だけに絞り込み、NA を含む行を除去し、(指定されていれば) 重みを適用し、(鑑定評価モードでは) 対象行を除外し、応答変換を施す。
3. **GAM の式を構築する** — 数値の予測変数について平滑化項 $s(x, bs="cr", k=10)$ を、質的変数について因子項を、Linear と指定された変数について線形項を、選択された交互作用についてテンソル積を構成する。earth のノットは cr 基底の項に明示的なノット位置として埋め込まれる。
4. **境界的な場合を処理する** — 定数の変数を除去し、二値の予測変数を因子に変換し、 k を一意な値の数から 1 を引いた値で上限し、より多くのノットが必要な場合は earth のノットをデータの分位点で補う。
5. **GAM を当てはめる** — 構成した式、family、method、gamma、select、その他のパラメータとともに `mgcv::gam()` を呼び出す。
6. **交差検証を実行する** (有効な場合) — 10 分割の交差検証によって標本外の決定係数を計算する。
7. **すべての結果タブを更新する** — Summary、Equation、寄与プロット、Diagnostics、その他すべてのタブの内容が満たされる。

8.3 背景での当てはめと進捗

callr パッケージが利用できる場合、モデルの当てはめは背景のプロセスで実行されるため、アプリケーションは応答性を保つ。暗色テーマの端末風の進捗表示が次の情報を示す。

- 経過した秒／分を表示するタイマー。
- 色分けされた追跡ログ: 状態は青、交差検証の分割の進捗は黄、誤りは赤、結果は緑である。
- 必要に応じて当てはめの処理を打ち切る **Abort** ボタン。
- 完了後に現れる **Close** ボタン。

callr が利用できない場合、当てはめは同期的に実行される (当てはめが完了するまでアプリケーションは応答しなくなる)。

当てはめに成功すると、Fit ボタンに白いチェック印が現れ、状態行に “R-sq = X, CV R-sq = X, Dev = X%, AIC = X, n = X.” と表示される。

9 結果のタブ

9.1 Data

取り込んだデータを、横スクロールが可能で 1 ページ 15 行の対話的な DataTable として表示する。鑑定評価モードと市場モードでは、このタブは **Subject Property** の表（第 1 行）と **Comparable Sales** の表（残りの行）に分かれる。General モードでは単一の表が表示される。

各セルは 1 行に収められ、列幅に合わせて切り詰められるので、幅の広い自由記述の項目（不動産の備考など）が行をページの下方向へ引き伸ばすことはない。シングルクリックで行を選択し、**任意のセルをダブルクリック**すると、切り詰められていない完全な内容がポップアップで表示される。

9.2 Equation

当てはめられた GAM モデルを二つの部分に分けて表示する。

モデルの式（MathJax で描画される）：

- 左辺には、適切な変換 ($\log$, $\log_{10}$) とリンク関数を伴う応答が示される。
- 右辺には、切片の値、 $f_i(\text{variable})$ としての平滑化項、変数ごとにまとめられ水準数を付した因子項、および係数の値を伴う線形項が示される。
- family と method は式の上に表示される。

Smooth Function Definitions の表:

- 項の式の文字列（たとえば `s(living_sqft, bs="cr", k=8)`）
- 変数名
- 型 (s, te, ti, または linear)
- 基底関数の型
- k の値
- earth のノット位置（取り込んだ場合）

9.3 Summary

上部に六つの指標カードが並ぶ。

- **R-squared** — 説明された逸脱度の割合
- **CV R-squared** — 交差検証による決定係数（交差検証が有効な場合）
- **Dev. Explained %** — 説明された逸脱度の百分率
- **AIC** — 赤池情報量規準
- **n** — 用いられた観測の数
- **Smooths** — 平滑化項の数

カードの下には次の表がある。

- **Smooth Terms** の表 — 項、EDF（有効自由度）、Ref.df、F 統計量、p 値
- **Parametric Terms** の表 — 項、推定値、標準誤差、t 値、p 値

9.4 Variable Importance

モデルの各項の相対的な重要度を示す横棒グラフである。

- **平滑化項**: F 統計量で順位付け（青の棒）
- **パラメトリック項**: $|t|$ 値で順位付け（緑の棒）

グラフの下に DataTable には、項、型、EDF、統計量、p 値が、統計量の降順で並べて示される。

9.5 寄与プロット

予測値に寄与するモデルの各項には、対話的な **plotly** の寄与プロットがあり、枠線付きのカードに表示される。このタブはすべての項の型についてプロットを示す。

9.5.1 一変量の平滑化項 — $s(x)$

- 95% 信頼帯を伴う青い線
- 偏残差を示す灰色の散布点（対数変換していないモデルの場合）
- earth のノット位置における赤い破線の垂直線（earth を取り込んだ場合）
- ホバー時のツールチップ: 変数の値、寄与 (\$)、95% 信頼区間、変化率（単位当たりの傾き）

対数変換したモデルでは、y 軸は $\bar{y} \times (e^{f(x)} - 1)$ という式によって金額の寄与へ逆変換されるため、曲線を直接に金額として解釈できる。

緯度・経度の変数については、傾きは 0.001 度当たりで示される。

9.5.2 因子別平滑化項 — $s(x, \text{by} = \text{factor})$

因子の水準ごとに 1 本の色分けされた線を、単一のプロットに重ねて描く。ある予測変数の効果が群（たとえば異なる近隣）によってどう変わるかを見るのに有用である。

9.5.3 交互作用項 — $ti(x_1, x_2)$ および $te(x_1, x_2)$

2 次元の寄与曲面を示すヒートマップである。色の尺度は赤（負）→ 白（ゼロ）→ 青（正）である。ホバーすると双方の変数の値と寄与の量が表示される。

9.5.4 パラメトリック項 — 線形および因子

- 数値の線形項: データ点の散布図に当てはめられた直線を重ねる
- 因子項: 因子の水準ごとの平均寄与を示す棒グラフ

9.6 Correlation

すべての数値の予測変数間のピアソン相関を示すヒートマップ行列である。データを取り込んだ直後から利用でき、モデルの当てはめを必要としない。

- 色の勾配: 青 (-1) → 白 (0) → 赤 (+1)
- 各セルに相関の値が印字される
- 当てはめの前にコンカービティ（多重共線性の平滑版の類似物）を見いだすのに有用である

9.7 Diagnostics

二組の診断用の図がある。

診断パネル (`gratia::appraise` による): 四つのプロット — 残差対当てはめ値、Q-Q プロット、残差のヒストグラム、応答対当てはめ値である。

実測値対予測値の散布図: 観測値と予測値を、45 度の破線の基準線とともに示す。対数変換したモデルでは、値は元の尺度へ逆変換される。

9.8 RCA Adjustments

RCA の計算の後に表示される三つのヒストグラムである（鑑定評価／市場モード専用）。

- **Residual Adjustment %** — 売買価格に対する百分率としての残差補正の分布
- **Net Adjustment %** — 純補正の分布

- **Gross Adjustment %** — 総補正の分布

各ヒストグラムは副題に平均、中央値、標準偏差を示し、平均と中央値には破線の基準線を引く。

9.9 Anova

GAM の要約からパラメトリック項と平滑化項を統合した分散分析表である。項、型（パラメトリックまたは平滑）、および関連するすべての統計量を示す。

9.10 Mgcvc Output

次の内容を示す生のテキスト出力である。

- 当てはめの所要時間
- GAM の式
- モデルの print 出力
- `summary(model)` の完全な出力
- `family` とリンク関数
- 全体のコンカービティの尺度

9.11 Sign Check

earthUI のノットを取り込んだ場合、このタブは earth モデルにおける各変数の方向と GAM におけるその方向とを比較する。

- **earth_direction**: 増加、減少、または混在（ヒンジ関数の符号に基づく）
- **gam_direction**: 増加、減少、または混在（100 点の格子上での平滑化項の導関数から推定）
- **consistent**: TRUE/FALSE
- **warning**: 不一致があればその記述

状態行には “All directions consistent”、または不一致のある変数の数が表示される。

9.12 Concurvity

コンカービティは多重共線性の平滑版の類似物である。このタブは次を示す。

- **Overall Concurvity** の表: worst / observed / estimate の各尺度を行に、平滑化項ごとを列にとる。色分けは白 (<0.5)、黄 (0.5–0.8)、赤 (>0.8) である。
- **Pairwise Concurvity** の表: 同じ色分けによる最悪の場合の対ごとの行列。

“worst” の行の値が ~0.8 を超えていれば、その平滑化項は他の平滑化項によってよく近似できることを意味し、冗長であることを示唆する。

9.13 Basis Check

`mgcv::gam.check()` の出力であり、各平滑化項について基底次元の妥当性を検定する。ある平滑化項の k が小さすぎる場合、この検定は低い p 値でそれを指摘し、 k を増やすべきことを示唆する。

10 データのダウンロード

当てはめの後、**Download Output (Excel)** ボタン（サイドバー第 7 節）は、予測値と診断を含む Excel ファイルを、有効なプロジェクトの mgcv 出力フォルダに書き出す。

10.1 出力の列

列	説明
basis	逆変換した切片の値（全行で同一）
<var>_contribution	平滑化項ごとの予測への寄与。対数モデルでは金額へ比例配分される。
est_<target>	モデルの予測値（たとえば est_sale_price）。該当する場合は対数尺度から逆変換される
residual	実測値 − 予測値
calc_residual	検証用: 実測値 − 寄与の総和（逆変換後）
cqa	Condition-Quality-Appeal (CQA) スコア（0-10 の尺度）
residual_sf	残差／居住面積（living_area が指定されている場合）
cqa_sf	residual_sf による順位付けから算出した CQA

鑑定評価モードでは、第 1 行(対象)の residual/cqa/cqa_sf は NA に設定される。living_area 列が指定されている場合、行は residual_sf の降順に並べられる。

正常に完了すると、ダウンロードボタンに白いチェック印が現れる。

10.2 CQA スコア

CQA は、各行の残差を他のすべての行と比べて 0-10 の尺度で順位付けする。

- **高い CQA (〜9-10):** 予測よりはるかに高く売れた — 状態・品質・訴求力が優れている可能性が高い
- **低い CQA (〜0-1):** 予測よりはるかに低く売れた — 状態が劣るか、困窮売買である可能性が高い
- **CQA が 〜5:** 中央値の近くにある

11 RCA の計算とダウンロード

RCA（残差制約アプローチ, Residual Constraint Approach）の作業手順は、モデルを当てはめた後、**Appraisal モード**と **Market Area Analysis モード**で利用できる。対象不動産は第 1 行になければならない。

11.1 RCA ダイアログを開く

サイドバー第 8 節の **Calculate RCA Adjustments & Download** ボタンをクリックする。次の項目をもつモーダルダイアログが現れる。

- **Score type:** CQA または CQA per SF (living_area が指定されている場合)。per-SF の選択肢は、対象の残差を居住面積で尺度調整する。
- **Subject CQA Score:** 0.00 から 10.00 までの値を受け付ける。この欄は空欄で始まる。earthUI のモデルを取り込んでいる場合、earthUI で用いたスコアの型と値が、共有されたプロジェクト設定から引き継がれる。

Generate をクリックすると計算とダウンロードが行われる。Excel ファイルは有効なプロジェクトの mgcv 出力フォルダ (<os>_out_mgcv) に書き出される。

11.2 CQA スコアの内挿

1. 取引事例の CQA スコアと残差が並べ替えられる
2. 線形内挿 (approx()) が、与えられた CQA の値を残差に対応づける
3. 対象の価値 = モデルの予測値 + 内挿された残差
4. 重みが 0 の行は、対象と同じ残差をもつものとして扱われる

11.3 出力の列

列	説明
subject_value	モデルの予測値 + 内挿された残差
subject_cqa	利用者が入力した CQA スコア
<var>_adjustment	対象の寄与 - 取引事例の寄与 (金額)
residual_adjustment	対象の残差 - 取引事例の残差
net_adjustments	すべての補正の総和
gross_adjustments	補正の絶対値の総和
residual_adj_pct	売買価格に対する残差補正の%
net_adj_pct	売買価格に対する純補正の%
gross_adj_pct	売買価格に対する総補正の%
adjusted_sale_price	実際の売買価格 + 純補正

計算が正常に終わると、ボタンに白いチェック印が現れる。

12 取引事例比較表

取引事例比較表は、RCA 補正を計算した後、**Appraisal** モードでのみ利用できる。鑑定評価報告書に組み込むのに適した、書式の整った Excel のワークブックを生成する。

12.1 取引事例選択ダイアログ

Generate Sales Grid & Download ボタンをクリックすると、すべての比較取引事例を列挙するモーダルダイアログが開く。取引事例は二つの群に分けられる。

- **Recommended comps**（あらかじめ選択済み）：総補正が売買価格の $< 25\%$ であるもので、経過日数の順（最も新しいものから）に並ぶ。最大 30 件である。
- **Other comps**（既定では未選択）：残りの取引事例で、総補正率の順に並ぶ。

取引事例を選び、**Generate Sales Grid** をクリックしてワークブックを作成する。

必須の特殊型: `contract_date` と `living_area` を指定していなければならない。**推奨される特殊型**: `latitude`、`longitude`、`lot_size` — これらがなくとも比較表は機能するが、いくつかの項目は空欄になる。

12.2 ワークブックの構成

生成されるワークブックは複数のシートを含み、各シートには 3 件の取引事例が並べて収められる。

- **見出しの行**: 対象と取引事例の住所、売買価格、売買日、市場滞留日数
- **群にまとめられた行**: 位置（緯度、経度、地域）、敷地規模（敷地面積、敷地寸法）、築年数（実際経過年数、有効経過年数）を、群の見出しと下位項目とともに示す
- **モデル変数の行**: 平滑化項ごとに 1 行を設け、対象の寄与、取引事例の寄与、補正額を示す
- **残余的な特性の行**: モデルが捉えていない特性（状態、品質、眺望など）について鑑定評価者が入力するための、空欄で施錠されていないセル
- **要約の行**: 純補正%、総補正%、補正後売買価格を、Excel の生きた数式とともに示す

12.3 シートの保護と残余的なセル

各シートは誤った編集を防ぐために保護されているが、残余的な特性の値のセルは明示的に施錠を解いてある。これにより、鑑定評価者は数式に基づく補正の計算を保ちながら、モデルにない特性の値を入力できる。

13 レポートのダウンロード

レポートの生成は二段階の Quarto の作業手順であり、以前のリリースの一段階のレポートに代わるものである。

13.1 第 1 段階: Quarto レポートの生成

Generate Quarto Report ボタン (モデルを当てはめた後に利用できる) は、自己完結的な .qmd 一式を有効なプロジェクトの mgcv 出力フォルダに書き出す。この一式には、Quarto のソース、あらかじめ描画されたプロットのファイル (PNG および PDF)、すべての表とモデルのメタデータを取めた `report_data.rds` のペイロード、そして Word 出力のためのスタイルテンプレート `reference.docx` が含まれる。一式は自己完結的であるから、描画の前に .qmd のソースを編集したり、他の Quarto 文書と組み合わせたりできる。

プロットは一式の生成時にあらかじめ描画される。個々のプロットが失敗した場合は代替の画像が書き出され、一式の残りは通常どおり生成される。

13.2 第 2 段階: Quarto レポートの変換

Convert Quarto Report の節は、`quarto::quarto_render()` によって .qmd ファイルを一つ以上の出力形式に描画する。既定では最後に生成された一式が対象となるが、**Browse** ボタンによって任意の .qmd ファイルを選択できる。三つの形式が利用できる。

- **HTML** — 自己完結的な HTML 文書。追加のソフトウェアなしにすべてのプラットフォームで機能する。既定として推奨される。
- **Word (.docx)** — 同梱の `reference.docx` スタイルテンプレートで描画される。編集と配布に適する。
- **PDF** — **LaTeX の導入を必要とする** (後述のシステム要件を参照)。LaTeX が検出されない場合、PDF の選択肢は形式のチェックボックスから自動的に隠される。

旧来の一段階の描画関数 `render_gam_report()` と `export_gam_docx()` は、プログラムからの利用のために引き続き提供されている。

13.3 レポートの内容

生成されるレポートには次のものが含まれる。

- **データセットとモデル:** データファイル、標本の大きさ、family、method、および当てはめに関する情報
- **結果の要約:** R^2 、 $CV R^2$ 、説明された逸脱度、AIC、および関連する当てはめの指標
- **モデルの式:** 平滑化項を f_i の記法で、線形項を明示的な係数で表したモデルの式、および平滑関数の定義。長い式は PDF では複数行に折り返される。
- **要約の表:** 平滑化項 (EDF, F, p 値) およびパラメトリック項 (推定値、標準誤差、t、p 値)
- **有効な交互作用:** モデルに含まれるテンソル項と因子別平滑化項の交互作用の行列 (交互作用が存在する場合にのみ表示される)
- **変数重要度:** F 統計量 (平滑化項) または $|t|$ (パラメトリック項) で順位付けした棒グラフ
- **平滑な効果** (すべての項の型について) :
 - 一変量の平滑化項のプロット (対話的プロットの静的な ggplot2 版)
 - 因子別平滑化項の多線プロット
 - 交互作用のヒートマップ (t_i 項および t_e 項)
 - パラメトリック項のプロット (数値は散布図、因子は棒グラフ)
- **予測変数間の相関のヒートマップ**

- **診断:** 残差の診断、および実測値対予測値の散布図
- **mgcv の要約:** 当てはめたモデルの完全なテキスト出力
- **記法の付録:** レポートで用いる価値寄与関数の記法を説明する、番号を付さない末尾の節

生成が正常に終わると、Generate Quarto Report ボタンに白いチェック印が現れる。

14 モデル関数の書き出し

mgcvUI は、当てはめた GAM の平滑関数を、複数の言語の自立したコードとして書き出すことができる。各平滑化項は 200 点の格子上で評価され、線形内挿を伴う参照表として書き出されるため、mgcv に一切依存することなく R 以外の環境でモデルを利用できる。

現行リリースでは、関数の書き出しはプログラムから行う。書き出しの補助関数は、(fit_gam() が生成した) 当てはめ結果のオブジェクトを受け取り、以下に述べるファイルを書き出す。

- generate_r_code()、generate_python_code()、generate_cpp_code() — 言語ごとのソースファイル
- export_functions_sqlite() — SQLite の参照表データベース
- export_functions_zip() — 四つの形式の任意の組み合わせを zip 書庫にまとめる

14.1 R 関数

ファイル: gam_functions.R

- 応答変数、 R^2 、 $CV R^2$ 、および式を記した冒頭のコメント。
- INTERCEPT 定数。
- 平滑変数ごとに、あらかじめ計算された x/y のベクトルをもつ approxfun()。
- 線形項の係数を名前付きの定数として。
- 用法: 切片、各 g_var(value)、各 coef_var * value を合計する。

14.2 Python 関数

ファイル: gam_functions.py

- numpy の読み込み。
- INTERCEPT 定数。
- 変数ごとに _x_var と _fx_var の配列、および np.interp() を用いる g_var(x) 関数。
- 線形項の係数を COEF_* 定数として。

14.3 C++ ヘッダ

ファイル: gam_functions.hpp

- #pragma once を用いるヘッダのみの実装。
- gam_functions 名前空間。
- std::lower_bound を用いる interp() 補助関数。
- 変数ごとに、静的な const 配列をもつインライン関数。
- 線形項の係数を constexpr double として。

14.4 SQLite データベース

ファイル: gam_functions.sqlite

- 表 smooth_grids: 列は (variable, x, fx) であり、200 点の参照表である。
- 表 model_info: response, r_squared, cv_rsqr, dev_explained, aic, n_obs, family, method, formula。
- 表 linear_terms: variable, coefficient。
- 表 intercept: 切片の値。

15 earthUI および glmnetUI との比較

earth アプリケーションとその後処理モジュールは、回帰モデリングのための一つのツール群である。これらは同じデータ形式、特殊列型、RCA の作業手順、デモ用データセットを共有するが、モデリングエンジンは異なる。

15.1 主要な相違点

項目	mgcvUI	earthUI	glmnetUI
手法	GAM (mgcv)	MARS/Earth (earth)	エラスティックネット (glmnet)
関係	平滑な非線形	区分線形	線形
変数選択	select=TRUE による縮小	逐次的な枝刈り	ラッソ罰則
交互作用	テンソル積 (te/ti)	ヒンジ関数の積 (次数 2-3)	対ごとの積
偏効果	信頼区間を伴う平滑曲線	g 関数 (区分的)	線形の傾き
earth パイプライン	earth からノットを取り込む	ノットの供給元	ヒンジ基底を取り込む
コードの書き出し	R, Python, C++, SQLite	R, Python, C++, SQLite	—
レポートの形式	Word, PDF, HTML	Word, PDF, HTML	Word, PDF, HTML

earthUI のみがある特徴として一つ挙げておくべきものがある。earthUI には、自由記述の MLS 備考から構造化された不動産の属性を抽出する、実験的な Prolog/DCG の機能が含まれている。

15.2 共通する機能

三つのツールはいずれも次のものを提供する。

- 共有された regProj プロジェクトツリー (既定は ~/regProj)。プロジェクト、入力ファイル、手法ごとの出力が姉妹アプリケーション間で共有される
- snake_case への列名変換を伴う CSV および Excel の取り込み
- CSV の区切り記号、小数点記号、用紙寸法を含む 30 を超える国のロケールへの対応
- 特殊列の指定 (contract_date, living_area, area, latitude, longitude など)
- 基準日からの sale_age の自動計算
- ダーク／ライトモードの切替 (Nord テーマ)
- 設定の保存 (SQLite および／または localStorage)
- 相関のヒートマップ、変数重要度、診断
- CQA の採点と RCA 補正の作業手順
- 取引事例の選択と書式の整った Excel 出力を伴う取引事例比較表
- 同じデモ用データセット (Appraisal_1.csv) との互換性

15.3 推奨される作業手順

- まず earthUI から始める。重要な変数、非線形の関係、交互作用を発見するためである。earth の自動的なモデル構築は最良の出発点を与える。
- より平滑な偏効果が欲しければ mgcvUI で洗練する。earthUI のノットを取り込めば滞りなく移行できる。

- **glmnetUI** を用いるのは、積極的な変数選択または係数の符号制約を伴う純粋な線形モデルが必要なときである。
- **三つすべてを同じデータで比較する**。非線形モデルの追加的な複雑さが正当化されるかを評価するためである。

16 デモ用データセット: Appraisal_1.csv

16.1 説明

デモ用データセットは earthUI および glmnetUI と共有されている。これは mgcvUI のソースの demo_mls/ フォルダにある。earthUI が導入されている場合は次のようにして得られる。

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

このファイルは、模擬的な MLS の書き出しによるおよそ 1,500 件の住宅の売買記録を含み、第 1 行が対象不動産、残りの行が比較取引事例である。データは、規模、築年数、位置がさまざまな複数地域の市場における戸建住宅の売買を表している。

これは実在のデータではないが、北カリフォルニアの現実的な近隣に基づいている。識別に関わる情報はすべて改変または除去されている。

16.2 列

列	型	特殊型	説明
sale_type	character	sale_type	対象と比較取引事例とを区別するレコード種別のコード
weight	numeric	weight	観測の重み (0 = 当てはめから除外)
id	numeric	display_only	内部のレコード ID
property_id	numeric	display_only	MLS の不動産識別子
listing_id	character	display_only	MLS の登録番号
parcel_number	character	display_only	郡評価官の地番 (APN)
street_address	character	display_only	不動産の住所
city_name	character	display_only	市
postal_code	character	display_only	郵便番号
county_name	character	display_only	郡
contract_date	Date	contract_date	売買契約日 (sale_age を計算する)
sale_age	numeric	—	契約日から基準日までの日数 (計算済み)
coe_date	Date	display_only	エスクロー完了日
listing_status	character	display_only	登録の状態 (たとえば “Sold”)
sale_price	numeric	(目的変数)	売買価格 — 応答変数
rent	numeric	—	月額賃料 (複数の目的変数のモデル用)
list_price	numeric	display_only	登録価格
original_list_price	numeric	display_only	当初の登録価格
living_sqft	numeric	living_area	総居住面積 (平方フィート)
ppsf	numeric	display_only	平方フィート当たり価格
beds_total	integer	—	寝室の数
baths_total	numeric	—	浴室の合計数 (たとえば 2.5 = フル 2 + ハーフ 1)
lot_size	numeric	lot_size	敷地面積 (平方フィート)
area_id	integer	area	MLS の地域識別子
area_text	character	display_only	地域名
age	numeric	actual_age	不動産の経過年数
year_built	integer	display_only	建築年
latitude	numeric	latitude	緯度 (モデル用に小数点以下 3 桁に丸められる)
longitude	numeric	longitude	経度 (モデル用に小数点以下 3 桁に丸められる)
latitude6	numeric	display_only	完全な精度の緯度
longitude6	numeric	display_only	完全な精度の経度
garage_spaces	integer	—	車庫の台数
fp_count	integer	—	暖炉の数
no_of_stories	numeric	—	階数
style	character	—	建築様式
view	character	—	眺望の種別 (たとえば “Neighborhood”、 “Hills”)
days_on_market	integer	dom	市場滞留日数

列	型	特殊型	説明
listing_date	Date	listing_date	登録日
sale_concessions	numeric	concessions	売主の譲歩

16.3 推奨される手早い始め方

1. mgcv ルーチンを起動する: `earthUI::launch_mgcv()`
2. 用途を **For Appraisal** として新しいプロジェクトを作成する (第 1 節)
3. `Appraisal_1.csv` をプロジェクトの入力フォルダに置き、第 2 節でそれを選択する (必要なら **Refresh** をクリックする)
4. 目的変数として `sale_price` を選ぶ
5. 比率による補正が欲しければ、任意で Response Transform を **Log (natural)** に設定する
6. 上の表に示したとおりに特殊型を割り当てる
7. 予測変数として次を含める: `sale_age`、`living_sqft`、`baths_total`、`lot_size`、`area_id` (因子として)、`age`、`latitude`、`longitude`、`garage_spaces`
8. 既定のままにする: REML、`gamma = 1.2`、`tp` 基底、`k = auto`、`select = TRUE`
9. **Fit MgcV GAM Model** をクリックする
10. Summary、Contribution (平滑化項のプロット)、Diagnostics の各タブを確認する
11. 出力をダウンロードし (第 7 段階)、CQA の順位付けを確認する
12. CQA スコアを ~5.00 として RCA 補正を計算する (第 8 段階)
13. 推奨される取引事例で取引事例比較表を生成する (第 9 段階)

17 システム要件と問題解決

17.1 対応プラットフォーム

mgcvUI は **macOS**、**Windows**、**Linux** で動作する。本アプリケーションは macOS と RStudio で開発され、主にその環境で試験されている。プラットフォームに固有の注意事項は以下に示す。

17.2 R のバージョン

R $\geq 4.1.0$ が必要である。RStudio Desktop (2023.06 以降) を強く推奨する。これは **Quarto** と **pandoc** (レポートの描画に必要) を同梱しており、アプリケーションを起動するのに便利な環境を提供するからである。

17.3 必要な R パッケージ

以下のパッケージは依存関係として自動的に導入される。

パッケージ	用途
shiny	ウェブアプリケーションの枠組み
mgcv	GAM モデルの当てはめ
valengrCore	鑑定評価の共有基盤 (特殊な役割、earthUI からの引継ぎ)
ggplot2, plotly	静的および対話的な寄与プロット
gratia	平滑化項の抽出と診断
DT	対話的なデータ表
quarto, rmarkdown	Quarto レポートの生成と描画
readxl	Excel ファイルの取り込み
shinyFiles	ファイルとフォルダの選択画面
RSQLite, DBI	設定とプロジェクトのデータベース (SQLite)
jsonlite	変数の状態の直列化

17.4 任意の依存関係

構成要素	必要になる場面
LaTeX (TinyTeX, MiKTeX, または MacTeX)	PDF レポートの生成のみ。LaTeX の導入が検出されない場合、PDF の選択肢は形式のドロップダウンから自動的に隠される。導入は次のとおり: <code>tinytex::install_tinytex()</code>
sysfonts + showtext	ggplot2 のプロットに用いる Roboto Condensed フォント。利用できないかオフラインの場合、アプリケーションは自動的にシステムのサンセリフフォントに退避する。
earth ($\geq 5.3.0$)	earthUI の取り込みパイプラインにのみ必要。単独での利用には不要である。
callr	背景でのモデルの当てはめ（アプリケーションの応答性を保つ）。これがなければ当てはめは同期的に実行される。
writexl, openxlsx	Excel のダウンロード（モデル出力、RCA、取引事例比較表）。欠けている場合、アプリケーションが導入を促す。
bslib	Bootstrap 5 のテーマ設定（Nord Light/Dark）。
officer	旧来のプログラムからの Word 書き出し（ <code>export_gam_docx()</code> ）。

17.5 プラットフォーム固有の注意事項

17.5.1 macOS

問題が最も少ない。システムライブラリには Homebrew を推奨する。PDF レポートが必要な場合は次のようにする。

```
# In R:
install.packages("tinytex")
tinytex::install_tinytex()
```

17.5.2 Windows

RStudio Desktop で良好に動作する。主な注意事項は次のとおりである。

- **企業内の制限された端末:** 一時ディレクトリまたは利用者データのディレクトリが制限されている場合、設定の保存とレポートの生成が失敗しうる。アプリケーションは R のコンソールに警告を表示するが、保存機能を伴わずに動作を続ける。
- **LaTeX:** R から TinyTeX を導入する（上の macOS と同じコマンド）か、miktex.org から MiKTeX を導入する。
- **長いファイルパス:** Windows には 260 文字のパス長の制限がある。作業ディレクトリのパスは短く保つこと。

17.5.3 Linux

ディストリビューションによる差が最も大きい。Ubuntu / Debian の利用者は、R パッケージをコンパイルする前にシステムライブラリを必要とする場合がある。

```
sudo apt install libcurl4-openssl-dev libssl-dev \
  libxml2-dev libsqlite3-dev libfontconfig1-dev
```

PDF レポートについては、`tinytex::install_tinytex()` を実行するか、システムのパッケージ管理から `texlive-full` を導入する。

表示装置をもたないサーバでは、`plotly` の対話的プロットはレポートにおいて静的な代替として描画される。Shiny アプリケーション自体はウェブブラウザからの接続を必要とする。

17.6 機能の穏やかな縮退

`mgcvUI` は、任意の構成要素が欠けていても動作するように設計されている。

欠けている構成要素	挙動
LaTeX がない	レポート形式のドロップダウンから PDF の選択肢が隠される。HTML と Word は引き続き利用できる。
インターネットがない／フォントの取得に失敗	Roboto Condensed フォントがシステムのサンセリフに置き換えられる。コンソールにメッセージが記録される。
SQLite がない／ファイルシステムが読み取り専用	設定はセッション間で保存されない。コンソールに警告が表示される。それ以外は通常どおり動作する。
<code>gratia::appraise</code> が失敗	Diagnostics タブとレポートに代替のメッセージが表示される。代わりにコンソールで <code>mgcv::gam.check()</code> を用いる。
一時ディレクトリに書き込めない	レポートの生成が明確な誤りのメッセージとともに失敗する。TMPDIR 環境変数を書き込み可能な場所に設定して解決する。

17.7 問題解決

17.7.1 “PDF option is not available”

LaTeX の導入が検出されていない。R で次を実行する。

```
install.packages("tinytex")
tinytex::install_tinytex()
```

その後アプリケーションを再起動する。PDF の選択肢が現れる。

17.7.2 “Settings will not persist”

SQLite データベースのディレクトリを作成できないか、書き込めない。次の場所の権限を確認せよ。

- macOS/Linux: `~/Library/Application Support/R/mgcvUI/` または `~/.local/share/R/mgcvUI/`
- Windows: `%APPDATA%/R/data/mgcvUI/`

17.7.3 “Diagnostics plot unavailable”

`gratia::appraise()` 関数は、複雑なモデル（多数のテンソル交互作用、標準的でない分布族）に対して時折失敗する。これは上流の既知の問題である。回避策として、R のコンソールで `mgcv::gam.check(model)` を実行すれば診断プロットを直接に見ることができる。

17.7.4 “Report error: number of items to replace...”

これは `gratia` パッケージに関係する、以前のバージョンにおける既知の問題であった。最新版の `mgcvUI` が導入されていることを確認せよ。現行版はこの誤りを捕捉し、診断の節を代替の表示に置き換えてレポートを生成する。

17.7.5 ポート 7880 が使用中

(以前のセッションの影響で) ポート 7880 が占有されているためにアプリケーションが起動しない場合は、次を実行する。

```
# macOS/Linux:
```

```
lsof -ti:7880 | xargs kill
```

```
# Windows (PowerShell):
```

```
Stop-Process -Id (Get-NetTCPConnection -LocalPort 7880).OwningProcess
```

17.7.6 Linux でパッケージの導入が失敗する

システムライブラリが欠けている。上の Linux の節に挙げた開発用ヘッダを導入し、`install.packages()` を再度試みよ。

参考文献

- Wood, S. N. (2003). “Thin-plate regression splines”. In: *Journal of the Royal Statistical Society (B)* 65.1, pp. 95–114. DOI: [10.1111/1467-9868.00374](https://doi.org/10.1111/1467-9868.00374).
- (2004). “Stable and efficient multiple smoothing parameter estimation for generalized additive models”. In: *Journal of the American Statistical Association* 99.467, pp. 673–686. DOI: [10.1198/016214504000000980](https://doi.org/10.1198/016214504000000980).
- (2011). “Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models”. In: *Journal of the Royal Statistical Society (B)* 73.1, pp. 3–36. DOI: [10.1111/j.1467-9868.2010.00749.x](https://doi.org/10.1111/j.1467-9868.2010.00749.x).
- (2017). *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman and Hall/CRC.
- Wood, S. N., N. Pya, and B. Säfken (2016). “Smoothing parameter and model selection for general smooth models (with discussion)”. In: *Journal of the American Statistical Association* 111, pp. 1548–1575. DOI: [10.1080/01621459.2016.1180986](https://doi.org/10.1080/01621459.2016.1180986).