

目次

まえがき	iii
第1章 クイックスタート	1
1.1 Rについて	1
1.1.1 プログラミング言語としてのR	2
1.1.2 計算環境としてのR	2
1.1.3 コミュニティとしてのR	2
1.1.4 エコシステムとしてのR	3
1.2 なぜRが必要か	3
1.3 Rのインストール	5
1.4 RStudio	8
1.4.1 RStudioのユーザインターフェース	8
1.4.2 RStudio Server	14
1.5 Rの簡単な実行例	14
1.6 まとめ	16
第2章 基本的なオブジェクト	17
2.1 ベクトル	18
2.1.1 数値型ベクトル	18
2.1.2 論理値型ベクトル	20
2.1.3 文字列型ベクトル	21
2.1.4 ベクトルの一部を抽出する	22
2.1.5 名前付きベクトル	25
2.1.6 要素を抽出する	26
2.1.7 ベクトルのクラスを確認する	27

viii 目次

2.1.8	ベクトルを変換する	28
2.1.9	数値型ベクトルのための算術演算子	29
2.2	行列	30
2.2.1	行列を作る	30
2.2.2	行と列に名前をつける	31
2.2.3	行列から部分集合を取得する	31
2.2.4	行列の演算子を使う	33
2.3	配列	34
2.3.1	配列を作成する	35
2.3.2	配列の部分集合をとる	36
2.4	リスト	36
2.4.1	リストを作成する	37
2.4.2	リストから要素を抽出する	37
2.4.3	リストから部分集合をとる	38
2.4.4	名前付きリスト	39
2.4.5	値を代入する	40
2.4.6	リストに関する関数群	41
2.5	データフレーム	42
2.5.1	データフレームを作成する	42
2.5.2	データフレームの行と列に名前をつける	43
2.5.3	データフレームの部分集合をとる	44
2.5.4	値を代入する	47
2.5.5	因子型	49
2.5.6	データフレームのための便利な関数	50
2.5.7	ディスクへのデータの読み込みと書き込み	51
2.6	関数	52
2.6.1	関数を作成する	53
2.6.2	関数を呼び出す	53
2.6.3	動的型付け	54
2.6.4	関数を汎用化する	54
2.6.5	関数の引数のデフォルト値	56
2.7	まとめ	57
第3章 作業スペースの管理		58
3.1	Rの作業ディレクトリ	58

3.1.1	RStudio によるプロジェクトの作成	59
3.1.2	絶対・相対パスの比較	60
3.1.3	プロジェクトファイルの管理	61
3.2	環境の調査	63
3.2.1	既存のシンボルの調査	63
3.2.2	オブジェクトの構造を見る	65
3.2.3	シンボルの削除	68
3.3	グローバルオプションの変更	69
3.3.1	表示される桁数の変更	69
3.3.2	警告レベルの改変	71
3.4	パッケージの管理	73
3.4.1	パッケージに対する理解	74
3.4.2	CRAN からパッケージをインストールする	74
3.4.3	CRAN からパッケージを更新する	76
3.4.4	オンラインレポジトリからパッケージをインストールする	77
3.4.5	パッケージの関数を使う	77
3.4.6	マスキングと名前の衝突	81
3.4.7	パッケージがインストールされているかを確認する	83
3.5	まとめ	84
第4章 基本的な表現式		85
4.1	代入式	85
4.1.1	他の代入演算子	86
4.1.2	非標準的な名前とバッククォートの使用	89
4.2	条件式	91
4.2.1	if を文として使う	91
4.2.2	if を式として使う	95
4.2.3	if をベクトルと一緒に使う	97
4.2.4	ベクトル化された if を使う : ifelse	99
4.2.5	値を分岐させるために switch を使う	100
4.3	ループ式	102
4.3.1	for ループを使う	102
4.3.2	while ループを使う	107
4.4	まとめ	109

第5章 基本的なオブジェクトを扱う	110
5.1 オブジェクト関数を使う	110
5.1.1 オブジェクトの型を調べる	111
5.1.2 データの次元にアクセスする	115
5.2 論理関数を使う	119
5.2.1 論理演算子	119
5.2.2 論理関数	120
5.2.3 欠損値に対処する	123
5.2.4 論理値への型変換	125
5.3 数学関数を使う	125
5.3.1 基本的な関数	125
5.3.2 数値を丸める関数	127
5.3.3 三角関数	127
5.3.4 ハイパボリック関数	128
5.3.5 極限関数	128
5.4 数値解析法を使う	131
5.4.1 求根	131
5.4.2 微積分	135
5.5 統計関数を使う	136
5.5.1 ベクトルからの抽出	137
5.5.2 確率分布を扱う	138
5.5.3 統計量を計算する	140
5.6 apply 族の関数を使う	143
5.6.1 lapply()	144
5.6.2 sapply()	145
5.6.3 vapply()	146
5.6.4 mapply()	147
5.6.5 apply()	148
5.7 まとめ	148
第6章 文字列を扱う	150
6.1 文字列を始める	150
6.1.1 テキストを表示する	150
6.1.2 文字列を連結する	153

6.1.3	テキストを変換する	155
6.1.4	テキストの書式設定をする	159
6.2	日時の書式設定をする	162
6.2.1	テキストを日時にパースする	162
6.2.2	日時を文字列に変換する	166
6.3	正規表現を使用する	167
6.3.1	文字列のパターンを見つける	169
6.3.2	グループを用いてデータを抽出する	170
6.3.3	カスタマイズ可能な方法でデータを読む	172
6.4	まとめ	173
第7章 データを扱う		175
7.1	データを読み書きする	175
7.1.1	テキスト形式のデータファイルの読み書きを行う	175
7.1.2	Excel ワークシートの読み書きを行う	180
7.1.3	ネイティブ形式のデータファイルの読み書きを行う	182
7.1.4	組み込みのデータセットを読み込む	186
7.2	データを可視化する	188
7.2.1	散布図を作成する	188
7.2.2	折れ線グラフを作成する	195
7.2.3	棒グラフを作成する	200
7.2.4	円グラフを作成する	202
7.2.5	ヒストグラムと密度プロットを作成する	202
7.2.6	箱ひげ図を作成する	206
7.3	データを分析する	208
7.3.1	線形モデルを当てはめる	208
7.3.2	決定木を当てはめる	213
7.4	まとめ	215
第8章 Rの内部を覗く		217
8.1	遅延評価を理解する	217
8.2	コピー修正を理解する	222
8.2.1	関数外のオブジェクトに変更を加える	226
8.3	レキシカルスコープについて理解する	228

xii 目次

8.4	環境の動作を理解する	233
8.4.1	環境オブジェクトについて知る	233
8.4.2	環境を作成してつなげる	234
8.4.3	環境をつなげる	235
8.4.4	関数にかかわる環境を理解する	241
8.5	まとめ	243
第9章 メタプログラミング		244
9.1	関数型プログラミングを理解する	244
9.1.1	クロージャを作成して使う	244
9.1.2	高階関数を使う	249
9.2	言語オブジェクトの処理	254
9.2.1	表現式を捕捉して変更を加える	256
9.2.2	表現式を評価する	263
9.2.3	非標準評価について理解する	266
9.3	まとめ	272
第10章 オブジェクト指向プログラミング		274
10.1	オブジェクト指向プログラミングとは	274
10.1.1	クラスとメソッドを理解する	275
10.1.2	継承を理解する	275
10.2	S3 オブジェクトシステム	276
10.2.1	総称関数とメソッドディスパッチについて理解する	276
10.2.2	組み込みクラスとメソッドを使う	279
10.2.3	既存のクラスに総称関数を定義する	286
10.2.4	新しいクラスのオブジェクトを作る	288
10.3	S4 を扱う	299
10.3.1	S4 のクラスを定義する	299
10.3.2	S4 の継承を理解する	304
10.3.3	S4 の総称関数	306
10.3.4	多重ディスパッチを理解する	308
10.4	参照クラスを扱う	311
10.5	R6 を扱う	312
10.6	まとめ	315

第 11 章 データベース操作	316
11.1 リレーショナルデータベースの操作	316
11.1.1 SQLite データベースの作成	317
11.1.2 テーブルおよびテーブル内のフィールドへのアクセス	320
11.1.3 SQL を学ぶ	322
11.1.4 チャンク単位でクエリの結果を取得する	331
11.1.5 データの一貫性を保証するためにトランザクションを用いる	332
11.1.6 データベースへのデータ保存	337
11.2 NoSQL データベースの操作	340
11.2.1 MongoDB の操作	341
11.2.2 Redis の操作	356
11.3 まとめ	361
第 12 章 データ操作	362
12.1 データフレームの基本操作	362
12.1.1 組み込み関数群を用いたデータフレームの操作	363
12.1.2 reshape2 パッケージを用いたデータフレームの変形	370
12.2 sqldf パッケージを用いた SQL によるデータフレームの操作	375
12.3 data.table パッケージを用いたデータ操作	379
12.3.1 インデックスを用いたデータへのアクセス	385
12.3.2 グループ化を用いたデータの集約	387
12.3.3 データテーブルの変形	390
12.3.4 set 系関数による操作	392
12.3.5 データテーブルにおける動的スコープ	393
12.4 dplyr パッケージを用いたデータ操作	398
12.5 rlist パッケージを用いたネストされたデータの操作	405
12.6 まとめ	410
第 13 章 ハイパフォーマンスコンピューティング	412
13.1 コードのパフォーマンス問題を理解する	412
13.1.1 コードのパフォーマンスを測定する	414
13.2 コードのプロファイリング	421
13.2.1 Rprof によるコードのプロファイリング	421

xiv 目次

13.2.2	profvis パッケージによるコードのプロファイリング	424
13.2.3	なぜコードが遅いのかを理解する	426
13.3	コードのパフォーマンスを加速させる	428
13.3.1	組み込み関数を使う	428
13.3.2	ベクトル化を使う	431
13.3.3	バイトコードコンパイラ	433
13.3.4	Intel MKL 版の R ディストリビューションを使う	434
13.3.5	並列処理を使う	435
13.3.6	Rcpp パッケージを使う	443
13.4	まとめ	452
第 14 章 ウェブスクレイピング		453
14.1	ウェブページの内部構造	453
14.2	CSS セレクタを用いたウェブページからのデータ抽出	459
14.3	XPath を用いたデータ抽出	462
14.4	HTML のソース解析によるデータ抽出	466
14.5	まとめ	477
第 15 章 生産性を高める		478
15.1	Markdown 書類を書く	478
15.1.1	Markdown を知る	479
15.1.2	R と Markdown の融合	484
15.1.3	表とグラフを埋め込む	487
15.2	インタラクティブなアプリケーションを作成する	494
15.2.1	Shiny アプリケーションを作成する	494
15.2.2	shinydashboard パッケージを使う	498
15.3	まとめ	502
訳者あとがき		504
索引		506