

rmdja (2020-10-24 04:38)

rmdja による多様な形式の日本語技術文書の作成

Katagiri, Satoshi (ill-identified)

2020-10-24

目次

第I部 イントロダクション	7
序文	9
本稿の目的	11
R Markdown の現状と問題意識	11
rmdja の利点	12
R 使用経験のないユーザへ	14
第II部 最低限のチュートリアル	15
第1章 クイックスタート	17
第2章 下準備	19
2.1 より丁寧なインストール解説	19
2.2 書籍形式のビルド操作	19
第III部 R Markdown と Bookdown の基本機能	23
このパートの概要	25
第3章 静的なコンテンツの作成	27
3.1 Markdown の基本構文	27
3.1.1 インラインでの書式変更	27
3.1.2 ブロック要素	28
3.2 Markdown を使った図表の挿入	29
3.2.1 コメントアウト	30
3.3 数式	31
3.4 カスタムブロック	32
3.5 脚注	33
第4章 動的なコンテンツの作成	35
4.1 プログラムチャンク	35
4.2 プログラムで数式を生成する	38
4.3 プログラムを使った図の挿入	38

4.4	TODO: 図のレイアウト設定	40
4.5	R プログラムを使った表の装飾	41
第 5 章	相互参照と引用	43
5.1	相互参照	43
5.1.1	図表や式へのアンカーリンク	43
5.1.2	表への相互参照	43
5.1.3	章への相互参照	44
5.1.4	特殊な相互参照	44
5.2	文献引用	45
5.2.1	文献引用のフォーマット設定	46
5.2.2	(TODO) pandoc-citeproc と CSL	47
5.2.3	(WIP) BibLaTeX	47
第 6 章	(WIP) 簡単なレイアウト変更	49
6.1	HTML	49
6.1.1	フォント変更	49
6.2	PDF	49
6.2.1	フォント変更	49
第 7 章	(WIP) rmdja による文書作成支援機能	51
7.0.1	クリエイティブ・コモンズの表記	51
7.0.2	ルビ表記	51
第 IV 部 応用編	53	
このパートについて	55	
第 8 章 様々なグラフィックプログラムの埋め込み	57	
8.1	tikz を使う	57
8.2	Asymptote を使う	57
8.3	(TODO) その他のプログラム	58
8.4	(TODO) その他の R プログラム	58
第 9 章 表のデザイン	61	
9.1	TeX/HTML を出力する関数	61
9.2	kableExtra による表のスタイルのカスタマイズ	62
9.3	formattable パッケージとの併用	64
9.4	huxtable パッケージによる作表	64
9.5	(TODO) gt パッケージによる作表	66
9.6	その他の作表パッケージ	67
第 10 章 文献引用	69	
10.1	(u)pB <small>E</small> T <small>E</small> X を使う	69

第 11 章 (TODO) Web アプレットの挿入	71
11.0.1 TODO: plotly	71
11.0.2 TODO: shiny	71
第 12 章 Python スクリプトの埋め込み	73
12.1 Python のグラフィックに関する制約	75
第 V 部 製本と多様な形式への対応	77
第 13 章 PDF の文書クラス	79
13.1 プレゼンテーション資料の作成	80
13.1.1 主な設定	80
13.2 (WIP) 卒業論文の作成	81
13.3 (WIP) 小説の執筆	83
第 14 章 製本方法の詳細	85
14.1 ファイル構成	85
14.1.1 _output.yml	85
14.1.2 _bookdown.yml	86
第 15 章 出力形式による表現の限界	89
15.1 HTML と PDF で処理を場合分けする	89
15.2 絵文字の出力	90
15.3 画像の保存形式	90
15.4 デフォルトの保存形式	90
第 16 章 製本した文書を配布する	91
16.1 Wep ページのホスティング	91
16.2 TODO: 入稿するには	92
16.2.1 トンボの表示	92
16.2.2 フォントの埋め込み	92
第 VI 部 デバッグ	93
このパートで説明すること	95
第 17 章 製本時のエラーへの対処	97
17.1 エラーがどのタイミングで発生したかを特定する	97
17.2 YAML フロントマターを確認する	97
17.3 PDF 生成時のエラーを確認する	99
付録 A デフォルト値の自動調整	101
A.1 デフォルトのフォント	101
A.2 チャンクのデフォルト設定	102
付録 B PDF の組版に関する細かい話	103

B.1	画像の配置	103
B.2	取り消し線	103
B.3	TODO: しかし英文で書きたい場合	104
付録 C	jecon.bst の紹介	105
付録 D	fontregisterer でグラフ用フォントを自動登録	107
参考文献		109

第I部

イントロダクション

rmdja (2020-10-24 04:38)

序文



注意: 絶賛作りかけ

長大な技術文書や良質な技術文書を作成するには手間がかかる。しかし時間をかけていい文書になるわけではない。無駄な手間を省き、効率よく快適に文書を作成するべきである。

たとえばこういう経験はないだろうか。

- プログラムの解説のため、外部サービスでシンタックスハイライトしてもらったテキストをコピペーストで貼り付ける
- グラフや図解を専用アプリケーションで作成し貼り付ける。修正のたびに貼り付け直す
- 図表に言及する際に「図 1」「表 2」と番号をタイプし、参照先へハイパーアリンクを指定する
- 本文中で引用した参考文献のリストを巻末にコピペーストし、過不足がないか目視で確認
- $\sum_{k=1}^K \int_0^\infty f_k(x) dx$ などといった複雑な数式はプレーンテキストや HTML では表現できないため、画像を生成して貼り付ける
- 冒頭にかっこいいエピグラフを掲載したいので、1時間かけて特別に枠やフォントを作成した
- 市販のワードプロセッサで作成した文書を渡したら、レイアウトが崩れて読めないと言われた

本稿は文書作成者をこのような数え切れないブルシットから解放するのが目的である。

R Markdown (`rmarkdown`) は、R プログラムを埋め込んだ動的なドキュメントから pandoc を利用して PDF や HTML 形式の文書を作成するパッケージであり、数式、図表の挿入、シンタックスハイライトされたプログラムなどを簡単な記述で掲載できる。名前の通り、その基本構文は Markdown である。よって、Markdown と R の知識が最低限あれば (R プログラムが必要ないなら Markdown だけでも) 文書を作成することができる。

`bookdown` パッケージは `rmarkdown` をもとに、ページ数の多い文書を作成し、配布するための機能を拡張したものである。しかし、PDF の出力に関しては欧文を前提としたフォーマッ

トを使用しているため、日本語の適切な表示（組版やフォントの埋め込みなど）のできる文書を作成するには高いハードルが存在した。

本稿では、`rmarkdown` および `bookdown` で日本語文書を作成する際の設定を容易にしたパッケージ `rmdja` を利用した日本語技術文書の作成方法を解説する。現在、書籍、論文、プレゼンテーションの体裁での文書を作成するテンプレートが用意されており、このドキュメント自体も `rmdja` を利用して作成されている。

本稿の目的

R Markdown の現状と問題意識

R Markdown を利用した文書作成方法について、すでにそれなりの数と質の日本語資料が存在する。R Markdown で HTML ファイルのみを作成する場合、日本語であるか欧文であるかはあまり気にする必要はない。HTML であれば既存の資料でも十分に役に立つ。

- kazutan 『R Markdown 再入門』
- kazutan 『R Markdown によるスライド生成』

しかし、PDF を出力する、あるいは HTML と PDF を同時に出力したい、となると、組版に関して細かな設定が必要になるため難易度は一気に上昇する。

実のところ PDF でも日本語を表示する最低限の設定は、YAML フロントマターだけで行える。例えば Atusy 氏が『R Markdown + XeLaTeX で日本語含め好きなフォントを使って PDF を出力する』で紹介しているが、よりシンプルな書き方もできる。

```
output: pdf_document:
  latex_engine: xelatex
  documentclass: bxjsarticle
  classoption:
    - xelatex
    - ja=standard
    - jafont=noto
```

このままでも、とりあえず文字化けすることなく日本語を表示できる。しかし実際に作ってみると、いろいろな障害が立ちはだかり、文書として整ったものにするのは難しい。このままでは参考文献リストの表示も不自然なままである。だがこれ以上のカスタマイズは Atusy 氏がやっているようにテンプレートを修正でしか対処できず、LaTeX に対するそれなりの知識が必要となる。

さらに、同じソースファイルから HTML と PDF を同時に生成すると、また別種の問題が発生する。HTML と PDF は根本的に規格が違うため、様々な場合分け処理が必要であり、それは pandoc だけでは対応しきれない。

HTML 出力に限らない R Markdown の全般的な情報は、既に充実した英語の（公式）ドキュ

メントが多く存在する^{*1}.

- “Dynamic Documents for R • rmarkdown”
- “bookdown demo”
- “bookdown: Authoring Books and Technical Documents with R Markdown”
- “R Markdown Definitive Guide”
- “R Markdown Cookbook”^{*2}

しかしながらこれらを元に1からいろいろな調整を施すのはとても骨が折れるため, rmdja パッケージは日本語文書で HTML や PDF を同時に生成する場合の定番の処理をフォーマットに内蔵することにした.

rmdja の利点

従来 LaTeX や Word, あるいは他の媒体で文書を作成していたユーザにとっても, 文書の内容から面倒な設定を分離するため, 効率的に執筆できる.

Word ユーザにとっては, 以下のような利点がある^{*3}.

- 数十, 数百ページの文書を書いてもクラッシュすることがあまりない
- 輪郭のはっきりしたベクタ画像を簡単に貼り付けられる
- 図表の配置や相互参照を手動で書く必要がない
- 読み手の環境に依存してレイアウトが崩れにくい PDF ファイルを出力できる

ただし, .docx ファイルの出力はできない. 私が Word を持っておらず, R Markdown や pandoc がサポートしていても動作確認のしようがないため.

jupyter は Python のコードチャンクとその結果を簡単に表示できる文書作成ツールである. 出力オプションの少なさ(たとえば長大なコードもそのまま掲載されてしまう)や, IDE として見ても機能が少ないとからあまり使い勝手がよくなかったが, rmdja では **Python** スクリプトの埋め込みにもある程度対応している.

LaTeX のユーザー(シンプルなテキストエディタで書いているユーザも, Overleaf や LyX といった強力なエディタを使用しているユーザ)にとっては, LaTeX とほぼ同じ構文で数式を入力でき, かつ操作を大きく簡略化でき, 実験結果などをソースに埋め込むことができ, 外部プログラムからいちいちコピペする必要がなくなる. ただし, なるべく選択肢は広げておきたいが, なんでもありではかえって余計なことをしがちである. よって既に作成した beamer フォーマットと同様に, XeLaTeX および LuaLaTeXのみの対応を想定している. pLaTeX や upLaTeX には対応していない.

^{*1} 基本的なことがらの多くは上記を読めば分かるのでここでは基本機能をダイジェストで伝えた上で, これらの資料に書いてない応用技を紹介する. YAML のオプションの意味についてはソースコードにコメントを書いた. 以下, 単に, **BKD* と書けば “bookdown: Authoring Books and Technical Documents with R Markdown” (Xie 2020) を, RDG と書けば “R Markdown: The Definitive Guide” (Xie, Allaire, and Gohrehm 2018) を, RCB と書けば “R Markdown Cookbook” (Xie, Dervieux, and Riederer 2020) を指すこととする.

^{*2} 2020/10/19 に書籍としても発売されるらしい.

^{*3} ただし筆者は数年来 Word を使っていなかったため, これらのいくつかは既に改善されているかもしれない.

これまでも R Markdown を使用してきたユーザにとっては、 YAML フロントマターに数十行に渡っていた書いていた日本語表示のための設定の多くがフォーマットのデフォルト値になったため、かなり楽になると思われる。

たとえば、

```
bookdown::pdf_book:
  toc_depth: 3
  toc_appendix: true
  toc_bib: true
  latex_engine: xelatex
  keep_tex: true
  keep_md: true
  citation_package: natbib
  pandoc_args:
    - '--top-level-division=chapter'
    - '--extract-media'
    - '.'
  template: XXXXX.tex.template'
  dev: "cairo_pdf"
  out_width: "100%"
  out_height: "100%"
  quote_footer: ["\\VA{", "}{}"]
  extra_dependencies: gentombow
```

のように書いていたものがこうなる。

```
rmdja::pdf_book_ja:
  keep_tex: true
  keep_md: true
  tombow: true
```

さらにチャンクオプションを書いたり場合によっては.tex ファイルのテンプレートすら調整する必要もあった。それらも rmdja 内で調整している。

さらに、作成した文書は PDF 形式で出力することはもちろん、HTML 形式で様々なサイトで掲載でき^{*4}たり、電子書籍ファイルとしても出力可能である。このような多様な出力形式への対応しているソフトウェアはあまり例を見ない。

^{*4} bookdown 同様に R Markdown で作成した文書をブログ風のフォーマットで出力する blogdown パッケージというのも存在する。

R 使用経験のないユーザへ

R を使わない、あるいはそもそもプログラミングに詳しくない、という人間にもある使用機会がある。たとえば R を普段使わない人間でも bookdown で同人技術書を執筆したという事例がある^{*5}。この事例は主に数式と画像の貼付けのみだから、数式出力に必要な LATEX の知識があればほとんどのことはできてしまう。そして rmdja ではこの事例で言及されている LATEX の設定の多くは自動で制御される。また、小説などはほぼテキストであり、最低限のレイアウトさえ用意すれば数式も、あるいは画像の挿入すらいらないことが多い。rmdja では縦書き文書を PDF で出力する方法も用意している。



印刷用フォーマットおよび縦書き文書フォーマットは現在実験的な導入段階であり、表示の一部に不具合が存在する。

^{*5} <https://teastat.blogspot.com/2019/01/bookdown.html>

第Ⅱ部

最低限のチュートリアル

rmdja (2020-10-24 04:38)

第1章

クイックスタート

`rmdja` パッケージをインストールする。依存している `rmarkdown`, `bookdown`, `knitr` なども同時にインストールされる。

```
01 install.packages("remotes")
02 remotes::install_github("Gedevan-Aleksizde/my_latex_template", repos = NULL, type = "source")
```

まだ RStudio を使っていないなら、RStudio 上で作業することを強く推奨する。さらに、もしも R の操作自体にあまり慣れていないなら、森知晴『卒業論文のための R 入門』などを読むことを薦める。

加えて、以下のパッケージが役に立つので気に入ったらインストールしていただきたい。

```
01 install.packages(c("tidyverse", "ggthemes", "citr", "clipr", "kableExtra"))
```

RStudio を起動し、左上から新規作成を選び、"R Markdown" を選ぶ(図 1.1)。

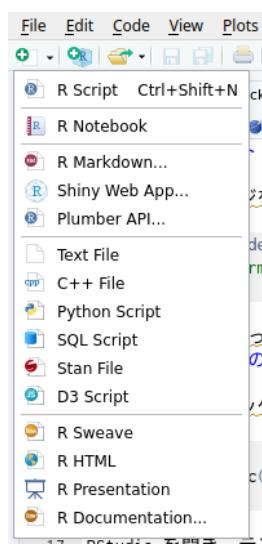


図1.1: 新規作成

"From Template" からテンプレートを選択する (1.2).

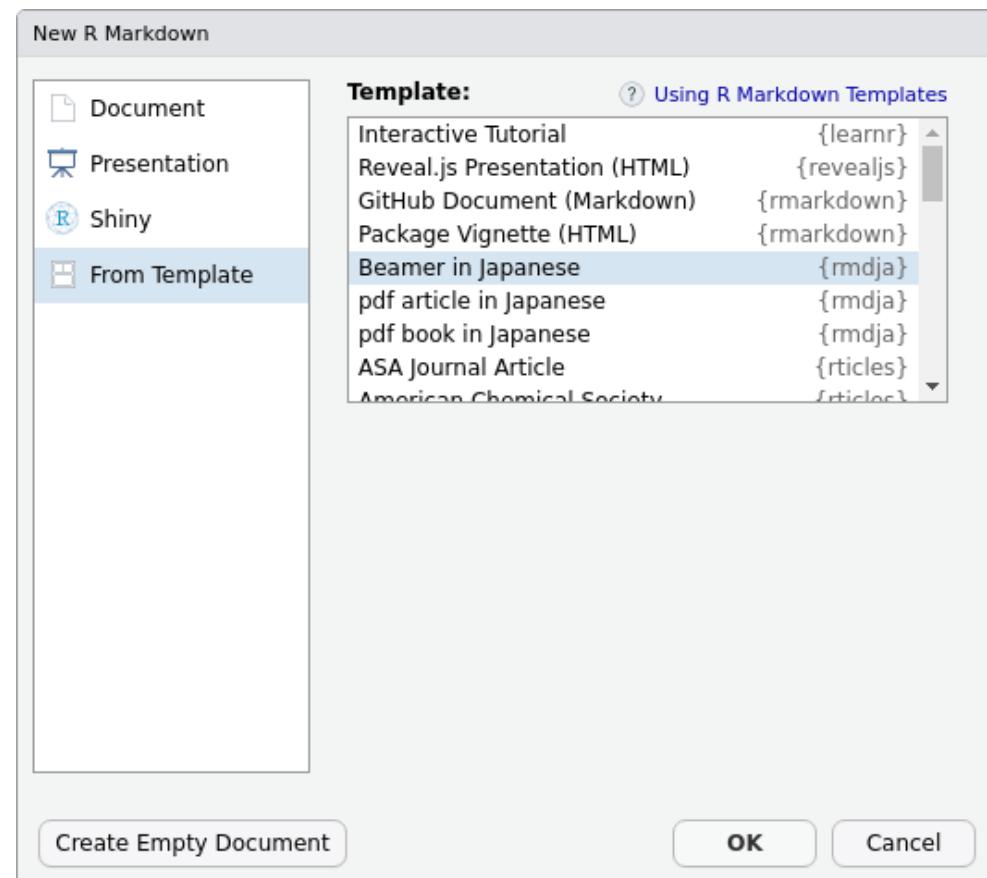


図1.2: R Markdown のテンプレート

現在 (Ver. 0.4.2) 用意されているのは以下の 4 つである。

- プrezentation用スライド形式のテンプレート - Beamer in Japanese
- 論文形式のテンプレート - pdf article in Japanese
- 書籍形式のテンプレート - pdf book in Japanese
- 縦書き文書のテンプレート - pdf vertical writing in Japanese

動作確認として、今回はシンプルな論文形式を選ぶ。ファイルを開いたら、適当な名称で保存し、"knit" ボタンを押すと PDF が作成される。

第 2 章

下準備

以降は順を追って細かい解説をする。

2.1 より丁寧なインストール解説

文書を生成するのに必要なものをインストールする。

このドキュメントは rmdja パッケージに含まれている。よってまずはこれをダウンロードしてほしい。

3種類のテンプレートのうち, pdf book in Japanese のみ、文書のビルドのための下準備が追加で必要になるため、その方法を解説する。それ以外は第 @ref(#quick-start) 節で書いたように “knit” ボタンを押すだけで良い。

最低限のファイルやパッケージで動くほどのデモ用ディレクトリをコピーする。ただし, tidyverse と kableExtra のインストールも必要である。

```
01 file.copy(system.file("resources/examples/bookdown-minimal", package = "rmdja"), "./", recursive = T)
```

2.2 書籍形式のビルド操作

bookdown の文書生成は従来の R Markdown と違い、RStudio の knit ボタンではできない。代わりに、以下の 2通りの方法がある。

1. bookdown::render_book('index.Rmd', format = "bookdown::gitbook")などを呼び出す
2. RStudio の Build ペーンを使う

前者は、以下のように実行する。順に HTML, PDF, epub を出力している

```
01 bookdown::render_book("index.Rmd", "rmdja::gitbook_ja")
02 bookdown::render_book("index.Rmd", "rmdja::pdf_book_ja")
03 bookdown::render_book("index.Rmd", "bookdown::epub_book")
```

コピーしたディレクトリ bookdown-minimal を設定する (図 2.1, 2.2).

Build ページの “Build Book” の三角形を押すと、使用できるフォーマット一覧が表示される。これはスライド、縦書き文書、書籍などといった文書の種類と 1 対 1 で対応しているわけではなく、フォーマット関数に対応している。

- HTML 形式 - rmdja::gitbook_ja
- PDF 形式 - rmdja::pdf_book_ja
- 電子書籍 (EPUB) 形式 - bookdown::epub_book

デフォルトでは “All Formats” にチェックが入っているため、これら 3 種類のファイル形式を一度に生成する。

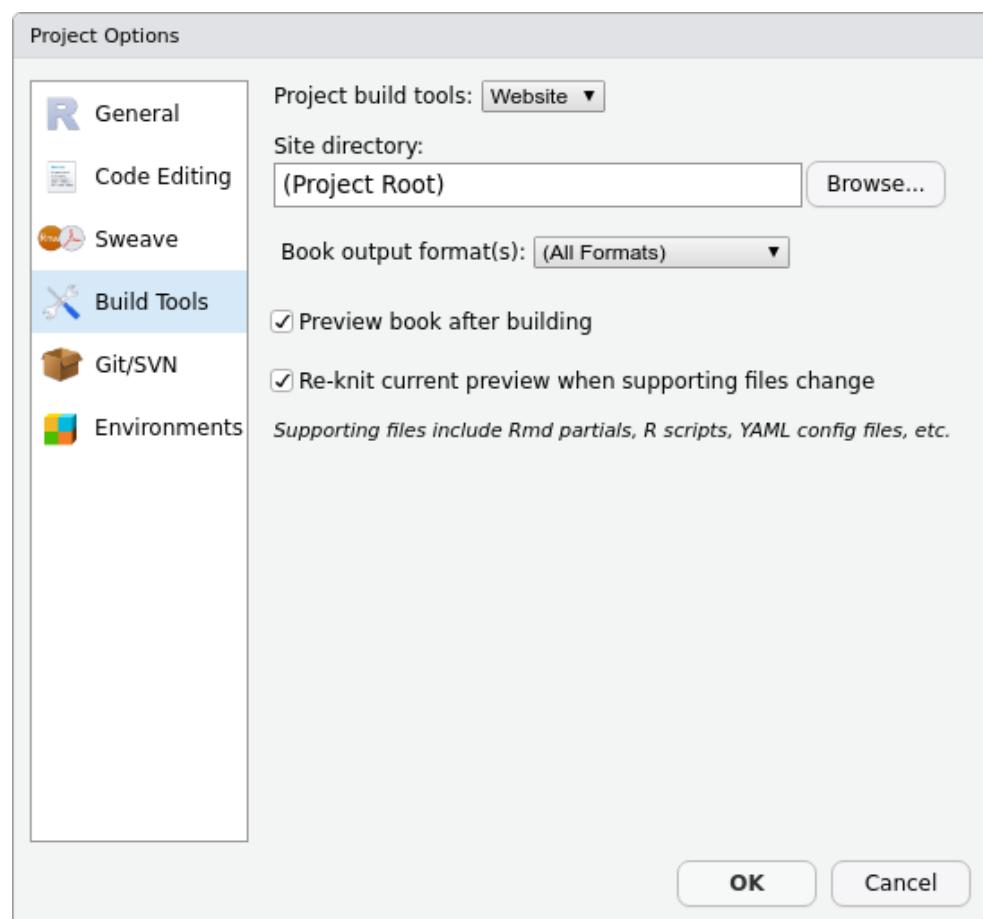


図2.1: Build ページの手動設定

これで _book フォルダに出力がされる。

2.2 書籍形式のビルド操作



図2.2: Build ペーンの手動設定

第 III 部

R Markdown と Bookdown の基 本機能

rmdja (2020-10-24 04:38)

このパートの概要

ここではまず, R Markdown の基本的な機能を紹介する. つまり bookdown 特有のものではなく, R Markdown 全般で使用できる機能も含めて紹介する. これ以降は自己言及的な説明が多いため, この文書を生成しているソースコードと比較しながら確認することをおすすめする. ここで紹介する機能は BKD, RDG, RCB での記述に基づく. これら 3 つのドキュメントを読めば, ほとんどのことは可能になる — rmdja を作る理由になった LaTeX テンプレートの修正以外は — のだが, 本稿の重要な目的の 1 つは複数のファイル形式を両立することであるので, それができない書き方には触れないし, 技術文書の作成にあまり使わないような機能の動作確認はおこなわず, 技術文書作成で頻繁に使われ, 便利と思える機能のみ紹介する.

どちらにしろそのうちこれらを翻訳してくれる人が現れることだろう...たぶん.

第3章

静的なコンテンツの作成

日本語で書かれた資料でごく基本的なことについて、『R Markdown 入門』で一通り紹介されている。やや応用的なことも『R Markdown ユーザーのための Pandoc's Markdown』に書かれている。

また、既に作成している beamer の用例ファイルもどのようなことができるかの参考になるだろう。ただしこちらは PDF のみの出力を前提としているため、一部の機能は HTML で使うことができない。

まずは、単なるマークアップ、つまりプログラミングの複雑な処理を考えなくても良いタイプの構文を紹介する。それらの多くは単なる Markdown のものと同じである。

3.1 Markdown の基本構文

一応基本の Markdown の構文も挙げておく。詳細は (ref:BKDB)“Ch. 2.2 Markdown Syntax”を参照。

3.1.1 インラインでの書式変更

テキストの一部のみ書式を変える

アンダースコアで強調(イタリック)

underline

underscore

** 2つで太字強調

太字強調

太字強調

等幅フォント

'bookdown' と 'rmdja'

bookdown と rmdja

本文中に入力した URL は自動判別され、ハイパーリンクが付けられる。また、[テキスト](URL) という書式で、テキストに対してハイパーリンクを付けることができる。

URL は自動判別される: https://github.com/Gedevan-Aleksizde/my_latex_templates/tree/master/rmdja

[`rmdja` の github リポジトリ](https://github.com/Gedevan-Aleksizde/my_latex_templates/tree/master/rmdja)

URL は自動判別される: https://github.com/Gedevan-Aleksizde/my_latex_templates/tree/master/rmdja

rmdja の github リポジトリ

3.1.2 ブロック要素

以降は行内では使えず、適切に表示するには前後に改行を挟む必要のあるタイプの構文である。

まず、引用ブロックを使えばかっこいいエピグラフを書き放題である。

```
> Нужны новые формы. Новые формы нужны, а если их нет, то лучше не
>
>新しいフォーマットが必要んですよ。新しいフォーマットが、それがないというなら、いっそ何もないほうがいい。
>
> `\\r tufte::quote_footer('--- A. チェーホフ『かもめ』')`
```

Нужны новые формы。Новые формы нужны, а если их нет, то лучше ничего не нужно。

新しいフォーマットが必要んですよ。新しいフォーマットが、それがないというなら、いっそ何もないほうがいい。

— A. チェーホフ『かもめ』

rmdja では、HTML と PDF 両方で同様のデザインの枠で表示するようにしている。

Markdown では # は見出しを意味するが、bookdown にはさらにオプションが用意されている。

見出し名 {-} で、セクション番号のつかない見出しを用意できる。序文、章末の参考文献、付録のセクションに使えるだろう。さらに、bookdown では # (PART) 見出し名で「部」の見

3.2 Markdown を使った図表の挿入

出しを作ることができる。この見出しあはセクションの合間に挟まるが、選択することはできない。文書が長くなったときに、より大きな区切りを付けるのに役に立つだろう。さらに、
(APPENDIX) 見出しあ名 {-} で、以降の見出しあの頭に「補遺 A, B, C, ...」と付番できる。

箇条書きは以下のように書ける。

```
* iris setosa
* iris versicolor
* iris virginica
```

- iris setosa
- iris versicolor
- iris virginica

```
1. iris setosa
2. iris versicolor
3. iris virginica
```

1. iris setosa
2. iris versicolor
3. iris virginica

インデントを使えばネストできる。

- 課長
 - 課長補佐
 - * 課長補佐代理
 - 課長補佐代理心得

3.2 Markdown を使った図表の挿入

markdown は表を記入することもできる。

Table: Markdown 記法の表

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
5.4	3.9	1.7	0.4

表3.1: Markdown 記法の表

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
5.4	3.9	1.7	0.4

画像ファイルも貼り付けられる。



図3.1: Johannes Gutenberg

しかし、キャプションを付けたり、表示位置やサイズを細かく調整したりするためには、後述するように R プログラムを経由して出力したほうが良い。

TODO: md 記法で画像貼り付けたときのサイズ統一

3.2.1 コメントアウト

HTML 式の `<!-- -->` でコメントアウトできる。コメントアウトされた箇所は生成ファイルでもコメントアウトされるのではなく、そもそも出力されなくなる。

3.3 数式

LaTeX 記法で数式を記述できる。HTML ならば Mathjax によってレンダリングされる。数式の記述ルールは少々ややこしい。これは現在の pandoc の仕様で HTML および LaTeX の規格で矛盾なく出力するためやむをえない措置である。

1. 改行をしない行内数式は \$ で囲む、または \(\backslash\)(, \) で囲む。
2. 改行を伴う数式ブロックは \$\$ で囲む、または \[\, , \] で囲む。
3. align, equation 環境等を使う場合は、上記の記号を使わず、直接 LaTeX コマンド `\begin{align} ... \end{align}` を打ち込む。

```
\@ref(eq:binom) は二項分布の確率関数である
\begin{align}
f(k) &= {n \choose k} p^k (1-p)^{n-k} (\#eq:binom)
\end{align}
```

その出力は、以下のようになる。

(3.1) は二項分布の確率関数である

$$f(k) = {n \choose k} p^k (1-p)^{n-k} \quad (3.1)$$

Bookdown では従来の R Markdown でできなかった数式への付番と、本文中での参照アンカーリンクの自動生成が可能となっている（詳細は 5.1 章で）。LaTeX にすでに慣れている読者に注意が必要だが、Bookdown 特有の制約として、付番したい場合は `\label{ID}` ではなく `(\#eq: ID)` を使う。また、PDF (LaTeX) と HTML (Mathjax) の仕様には

1. PDF では `align` は常に数式が付番され、`align*` 等はどうやっても付番されない
2. HTML では `align` でも `align*` であってもラベルを書かなければ付番されず、書けば付番される。

という違いがある。両者で同じ表示にこだわるのなら、付番を取り消す `\notag` を多用することになるだろう。

さらに、bookdown の機能として、LaTeX の「定理」「定義」「証明」などの環境に対応するものが提供されている（参考：BKD Ch. 2.2 Markdown extensions by bookdown）。これらの相互参照も可能である。

例：以下に補題 3.1、定理 3.1 を示す。

補題 3.1 (ボレル-カントリの補題)。 E_1, E_2, \dots をある確率空間の事象とする。これらの事象の確率の和が有限であるなら、それらが無限に多く起こる確率はゼロである。つまり、

$$\sum_{n=1}^{\infty} P(X_n) < \infty \Rightarrow P\left(\lim_{n \rightarrow \infty} \sup X_n\right) = 0,$$

$$\lim_{n \rightarrow \infty} \sup X_n = \bigcap_{n=1}^{\infty} \bigcup_{k \leq n} E_k$$

である。

Proof. 証明は読者の課題とする. □

定理 3.1 (無限の猿定理). 猿がほとんど確実にタイプライタの全てのキーを無限回叩くならば, ほとんど確実にテキストには任意の作品が含まれる.

Proof. 補題 3.1 より自明. □

3.4 カスタムブロック

数式のセクションの定理ブロックの応用で, 独自のブロックセクションを定義することができる. rmdja では BKD Ch. 2.7 Custom blocks で紹介されている例を予め使えるようにしている. それらは `type="..."` で指定できて, 以下の 5 種類がある.

- `rmdcaution`
- `rmdimportant`
- `rmdnote`
- `rmdatip`
- `rmdwarning`

である.



技術書によくある注意を喚起するブロック (`rmdcaution`).



技術書によくある注意を喚起するブロック (`rmdimportant`).



技術書によくある注意を喚起するブロック (`rmdnote`).



技術書によくある注意を喚起するブロック (`rmdatip`).



技術書によくある注意を喚起するブロック (`rmdwarning`).

このブロック内では Markdown の基本構文しか使えず, 引用や相互参照などは使えない. これらをブロック内で使いたい場合は `block` の代わりに `block2` と書く. ただしこちらは pandoc の機能のハックであるため, 将来使えなくなる可能性もある.

3.5 脚注

3.5 脚注

脚注はインラインと、巻末に書く2通りがある。

ここにインラインで脚注^{*1} [脚注の本文]

ここにインラインで脚注^{*1}

本文は巻末に書く^{[^example-1][^example-2]}。

[^example-1]: 脚注の本文その2

[^example-2]: 脚注の本文その2

本文は巻末に書く^{*2*3}。

ここにインラインで脚注^[^ 脚注の本文]

インラインで書くほうがシンプルに見えるが、この記法では間を空けずに連続して脚注を書くことができない。

このように書くと^{[^脚注その1][^脚注その2]}上付きとして認識される

^{*1} 脚注の本文

^{*2} 脚注の本文その2

^{*3} 脚注の本文その2

第4章

動的なコンテンツの作成

4.1 プログラムチャンク

プログラムチャンクは、R Markdown 最大の特徴であり、R のソースコードや、その実行結果を Markdown に挿入できる。さらには **R** 以外の言語の動作も可能である。順番が前後してしまったが、定理などのカスタムブロックは本来はプログラムを入力するためのチャンクブロックであり、それを静的なテキストコンテンツの挿入に流用しているだけである。

以降は R で多くのユーザが頻繁に使うパッケージと、いくつかの技術文書作成に役に立つパッケージをインポートしている前提の説明とする。なお、`rmarkdown`, `bookdown` はチャンク内で特に読み込む必要がない。

```

01 pkgs <- installed.packages()
02 for (p in c("tidyverse", "ggthemes", "equatiomatic", "tufte", "kableExtra")) {
03   if (!p %in% pkgs) install.packages(p)
04 }
05 if (!"rmarkdown" %in% pkgs) remotes::install_github("rstudio/rmarkdown")
06 if (!"bookdown" %in% pkgs) remotes::install_github("rstudio/bookdown")
07 require(tidyverse)

```

要求されたパッケージ `tidyverse` をロード中です

```
-- Attaching packages ----- tidyverse 1.3.0 --
v ggplot2 3.3.2     v purrr    0.3.4
v tibble   3.0.4     v dplyr    1.0.2
v tidyr    1.1.2     v stringr  1.4.0
v readr    1.4.0     vforcats  0.5.0

-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter()    masks stats::filter()
x dplyr::group_rows() masks kableExtra::group_rows()
x dplyr::lag()        masks stats::lag()
```

```
01 require(ggthemes)
```

要求されたパッケージ ggthemes をロード中です

```
01 require(equatiomatic)
```

要求されたパッケージ equatiomatic をロード中です

```
01 require(kableExtra)
```

このように、ログを掲載することもできる。これは再現性を重視する際に重宝するが、一方で単に画像などの出力だけを掲載したい場合もあるだろう。あるいは、プログラムを解説するためにプログラムは掲載するが実行しない、ということも必要になるかもしれない。プログラムと結果の表示/非表示はどちらも簡単に切り替え可能である。そのためには、チャンクオプションを指定する。

- echo: プログラムを掲載するかどうか
- message: プログラム実行結果の標準出力を掲載するかどうか
- warning: プログラム実行結果の警告を掲載するかどうか
- error: プログラム実行結果のエラーを掲載するかどうか
- eval: 文書作成時にプログラムを実行するかどうか
- include: 文書作成時にプログラムを実行し、かつ掲載しないかどうか
- results: 出力をいつもの R の出力風にするか ("markup"), 隠すか ("hide"), 出力を区切らずまとめるか ("hold"), テキストをそのまま出力するか ("asis")。最後はソースコードを動的に生成したい場合などに使う (後述)。

R の論理値は TRUE/FALSE または T/F と書く。

チャンクごとに個別に設定することも、デフォルト値を一括設定することもできる。前者の場合、チャンクオプションは {} 内部にカンマ、で区切って書く。r は R で実行するという意味である。

```
```{r [<name>], [<options>]}
data(cars)
summary(cars)
...```

```

一括設定の場合、以下のようなプログラムでデフォルト値を上書きできる。

```
01 knitr::opts_chunk$set(
02 echo = F,
03 message = T,
04 warnings = F,
```

## 4.1 プログラムチャンク

```
05 error = F
06)
```

なおこのチャンクは eval=F を設定することで、実行されることなくプログラムのみ掲載している。ただし、プログラムのみを掲載するなら、以下のように Markdown の機能でも可能である。こちらの記法は {} がなくなっていることに注意する。

```
```sh
echo Hello, Bookdown
```

```

{ } ブロック内の値にはさらに R プログラムで与えることができる。この使い方は後の章で解説する。

これらのオプションがあるおかげでプログラムとその結果の再現を説明したい場合はソースコードも表示させたり、回帰分析やシミュレーションの結果だけを掲載したい時は結果のみ表示したりできる。これが R Markdown のチャンクの強みである。例えば Jupyter notebook/lab などは従来、コードセルと出力セルを自由に隠すことができなかった。

チャンクに使用できる言語は R だけではない。つまり **Python** なども使用できる（詳細は @ref(#python) 章を参照）。以下で対応しているエンジンの一覧を表示できる。

```
01 names(knitr::knit_engines$get())
```

```
[1] "awk" "bash" "coffee" "gawk" "groovy"
[6] "haskell" "lein" "mysql" "node" "octave"
[11] "perl" "psql" "Rscript" "ruby" "sas"
[16] "scala" "sed" "sh" "stata" "zsh"
[21] "highlight" "Rcpp" "tikz" "dot" "c"
[26] "cc" "fortran" "fortran95" "asy" "cat"
[31] "asis" "stan" "block" "block2" "js"
[36] "css" "sql" "go" "python" "julia"
[41] "sass" "scss" "theorem" "lemma" "corollary"
[46] "proposition" "conjecture" "definition" "example" "exercise"
[51] "proof" "remark" "solution"
```

また、新たにプログラムを追加することもできる。詳細は RDG Ch. 2.7 Other language engines を参考に。

TODO: 他の言語のプログラムを実行する際の注意点

## 4.2 プログラムで数式を生成する

プログラムチャンクは、単にプログラムの計算結果を埋め込むだけでなく、静的なコンテンツを臨機応変に変更して出力させたり、あるいは手作業でやるには煩雑な加工処理を挟んでから表示させるのに役に立つ。

R のプログラムと組み合わせることで回帰分析の結果の数値をコピペすることなく数式で表示することができる。そのためには `equatiomatic` パッケージの `extract_eq()` を使う。

まずは、回帰係数を記号で表現するタイプ。LaTeX 数式をそのまま出力するため、チャックオプションに `results="asis"` を付ける必要があることに注意する。

```
01 data(iris)
02 fit <- lm(Sepal.Length ~ ., data = iris)
03 extract_eq(fit, wrap = T, ital_vars = T, align_env = "aligned")
```

$$\text{Sepal.Length} = \alpha + \beta_1(\text{Sepal.Width}) + \beta_2(\text{Petal.Length}) + \beta_3(\text{Petal.Width}) + \beta_4(\text{Species}_{\text{versicolor}}) + \beta_5(\text{Species}_{\text{virginica}}) + \epsilon$$

さらに `use_coef = T` で係数を推定結果の数値に置き換えた。

```
01 extract_eq(fit, wrap = T, ital_vars = T, use_coef = T, align_env = "aligned")
```

$$\text{Sepal.Length} = 2.17 + 0.5(\text{Sepal.Width}) + 0.83(\text{Petal.Length}) - 0.32(\text{Petal.Width}) - 0.72(\text{Species}_{\text{versicolor}}) - 1.02(\text{Species}_{\text{virginica}}) + \epsilon$$

`equatiomatic` パッケージは現時点では `lm` `glm` に対応しており、`lmer` への対応も進めているようだ。

TODO: この書き方だと PDF で付番できない

## 4.3 プログラムを使った図の挿入

既に Markdown 記法による図表の挿入方法を紹介したが、プログラムチャンクを介して画像を読み込み表示させることもできる。まずは、R のプログラムで既存の画像ファイルを表示させる方法。

```
01 knitr::include_graphics(file.path(img_dir, "Johannes_Gutenberg.jpg"))
```

もちろんのこと既存の画像だけでなく、データを読み込んでヒストグラムや散布図などを描いた結果を画像として掲載することもできる。

## 4.3 プログラムを使った図の挿入



図4.1: Johannes Gutenberg

技術文書や学術論文では、画像の上か下に「図 1: XXXXX」のようなキャプションを付けることが多い。紙の書籍では絵本のように本文と図の順序を厳密に守るより、余白を作らないよう図の掲載位置を調整する必要があるからだ。

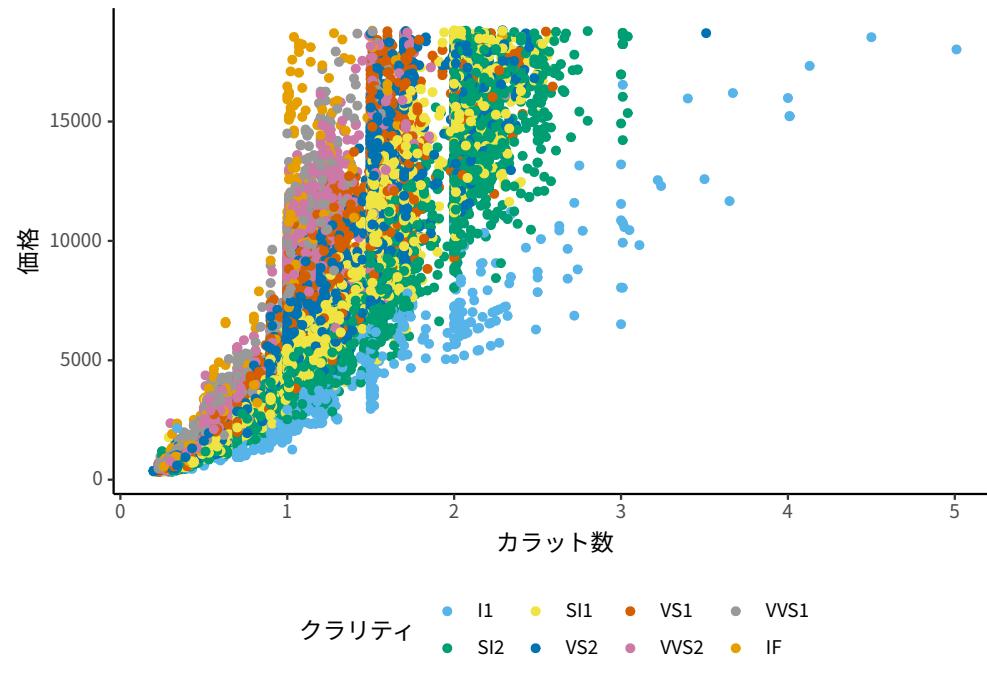
プログラムチャンクにはこのキャプションを入力するオプション `fig.cap` があるため、`plot()` 側でタイトルを付けないほうが良い。例えば `ggplot2` パッケージの関数を使い以下のようなチャンクを書く<sup>\*1</sup>。

```
```{r plot-sample, echo=T, fig.cap="`ggplot2` によるグラフ"}
data("diamonds")
diamonds <- diamonds[sample(1:NROW(diamonds), size =), ]
ggplot(diamonds, aes(x=carat, y=price, color=clarity)) +
  geom_point() +
  labs( x = "カラット数", y = "価格") + scale_color_pander(name = "クラリティ") +
  theme_classic(base_family = "Noto Sans CJK JP") + theme(legend.position = "bottom")
```
```

実際の表示は図 4.2 のようになる。

`ggplot2` 以外のパッケージや言語、たとえば `tikz` や `asymptote`, `DOT` 言語も使用できる。これらは 8 章で紹介する。

<sup>\*1</sup> なお、R ユーザーならば標準グラフィック関数である `plot()` 関数をご存知だろうが、本稿では標準グラフィック関数の使用を推奨しない。標準グラフィック関数のデバイスはもともと日本語フォントを想定しておらず、OS ごとに使用できるフォントも異なるため、品質維持のために使用させない方針とした。工夫すれば標準グラフィック関数でも日本語を適切に出力できるが、`ggplot2` を使用したほうが簡単であることが多いため、標準グラフィック関数の解説書を作る以外では使うべきでない。



#### 4.4 TODO: 図のレイアウト設定

PDFならばフロート設定のため、図が離れた位置に配置されることがある。そのため、「図4.2」のような相互参照を使うと良いだろう。フロートを使うかどうかは、後のセクションで解説する TODO

Rのグラフィックデバイスを使っている限り、通常のRのコンソールと同じコードをチャネル内に書くだけで表示できる。

Rのグラフィックデバイスではないとは、RGLやplotlyなど外部ライブラリに頼ったグラフ作成ツールのことである。判断できない人は、RStudio上で実行して、“Plots”ページに表示されたらRのグラフィックデバイス、“Viewer”ページに表示されたらそうでない、で覚えていただきたい。後者を表示する方法は11章で後述する。Rをこれまで使ったことがなく、それすらも何を言っているのか分からぬ、という場合はggplot2を使ってもらう。

最後の`fig.cap=""`がキャプションである。ただし、どうも日本語キャプションを書いたあとに他のチャネルオプションを指定するとエラーになるようだ。よって`fig.cap=`はオプションの末尾に書くべきである。また、`fig.cap=""`に数式や一部の特殊なテキストを直接入力することができない。この問題は相互参照について解説するセクション5.1で詳細を述べる。

`fig.cap`以外のオプションはおそらく頻繁には変えないため、冒頭でまとめて設定したほうが楽だろう。

## 4.5 R プログラムを使った表の装飾

```
01 knitr::opts_chunk$set(
02 fig.align = "center",
03 fig.width = 6.5,
04 fig.height = 4.5,
05 out.width = "100%",
06 out.height = "100%"
07)
```

なお、これらは rmdja でのデフォルト値であるため、実際にこの値をあえて記述する必要はない。

ここで、fig.width と out.width の違いも述べておく。out.width/out.height は表示する画像サイズの違いで、fig.width/fig.height はプログラムが output した画像の保存サイズである。よって ggplot2 などを使わず画像ファイルを貼り付けるだけの場合は fig.\* は意味をなさない。

## 4.5 R プログラムを使った表の装飾

Markdown 記法を使った表記は既に紹介した。しかしこれは表の数値を全て手動で書かなければならぬ。もちろんこれも R 内のデータを手書きなどせずとも表示できるし、テーブルのデザインもある程度自由に設定できる。

R Markdown のデフォルトでは R のコンソールと同様にテキストとして出力されるが、bookdown では異なるデザインで表示されている。これは knitr, kableExtra パッケージなどで事後処理をかけることで見やすいデザインの表に変換しているからである。

また、この方法はシンプルで使いやすいが数値を手作業で書く必要がある。R はテーブル状のデータ処理に長けているため、このような煩雑さを省くことができる。

```
01 data(iris)
02 kable(
03 head(iris, n = 10),
04 caption = "(ref:kable-cap)"
05)
```

こちらは関数内にキャプションを書く必要があり、チャンクオプションに指定する方法はない（表4.1）。

```
01 data(iris)
02 kable(
03 head(iris, n = 10),
04 booktabs = T,
```

表4.1: knitr::kable() で出力された (PDF ではあまりかっこよくない) 表

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa

表4.2: 奇数行を強調し, PDF では booktabs を利用

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa

```
05 caption = "奇数行を強調し, PDF では booktabs を利用"
06) %>% row_spec(seq(1, 10, by = 2), background = "gray")
```

kable() は R Markdown に必須な knitr パッケージに含まれる関数なので最初に紹介したが、現在は huxtable や gt というより使いやすいパッケージが存在する。より発展的な表のスタイル指定方法については @ref(#advanced-tabulate) 章で話す。

## 第 5 章

# 相互参照と引用

## 5.1 相互参照

### 5.1.1 図表や式へのアンカーリンク

図, 表, 式などに番号を自動で割り当てる, さらにハイパーリンクを付加できる。`\@ref(ID)` を使う。現状では `refstyle` や `prettyref` のように接頭語を自動で付けてくれないが, そのうちなんとかなるかもしれない。

`bookdown` の相互参照は, `LaTeX` の `prettyref.sty` のように, 接頭語: 参照 ID という記法になる。参照 ID は通常, チャンク ID と同じである。既に紹介したように, 数式参照の接頭語は `eq` で, 定理は `thm` である。図表は `fig`, `tab`。その他の接頭語は BKD Ch. 2.2 Markdown extensions by bookdown を参考に。

### 5.1.2 表への相互参照

Markdown 記法で表を書く場合, 以下のように `Table:` の直後にラベルを記入する(表 5.1)。

**Table:** (\#tab:tab-md) Markdown 記法の表

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
5.4	3.9	1.7	0.4

表5.1: Markdown 記法の表

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
5.4	3.9	1.7	0.4

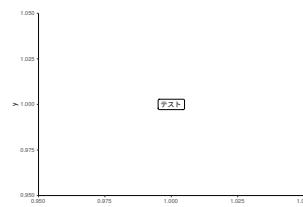
### 5.1.3 章への相互参照

章見出しへの相互参照も可能である。これは Pandoc の機能を利用しているため、接頭辞は不要である。Pandoc の仕様により欧文であればタイトルがそのまま参照 ID となるが、非欧文の文字に対して適用されないため、参照したい章の見出しにの後にスペースを入れて {# 参照 ID} と書く必要がある。

### 5.1.4 特殊な相互参照

チャンクオプションの `fig.cap` などに TeX 数式を書いても正しく表示できない。そのような場合は `ref` 参照を使う。`(ref:figcap1) \coloremoji{[]} $\sum \oint \mathfrak{A}` `\mathscr{B}` `\mathbb{C}` `\coloremoji{[]}` と書くと、図 5.1 のキャプションにも特殊な記号が使える。

なお、複数指定する場合は連続させず、改行で 1 行空けて宣言する必要がある。

図5.1:  $\sum \oint \mathfrak{A} \mathscr{B} \mathbb{C}$ 

この参照は一度しか使えない。

PDF での表示では、図 5.1 のキャプションの外側が文字化けしていることだろう。これは 絵文字出力に関する問題で、別のセクションで解説する。

これはかなり強力で、

1. 定義される前の行にも適用される
2. チャンクオプションだけでなく出力結果にも適用される

## 5.2 文献引用

という仕様である。

TODO: 自己言及的な文章は書かないならこれくらいの認識でいいだろうが、より正確な話はどうするか

## 5.2 文献引用

YAML フロントマターの `bibliography:` に文献管理ファイル (`.bib`, `.json` 等) を指定することで、ファイルに含まれる文献への参照が可能になる。`@` 引用 ID で本文に引用を与えられ、文書に引用した文献の一覧が自動で生成される。また、`citr` パッケージにより、RStudio Addins に文献に対応する引用 ID を取り出して挿入する機能が追加される。

```
01 knitr::include_graphics(file.path(img_dir, "citr.png"))
```

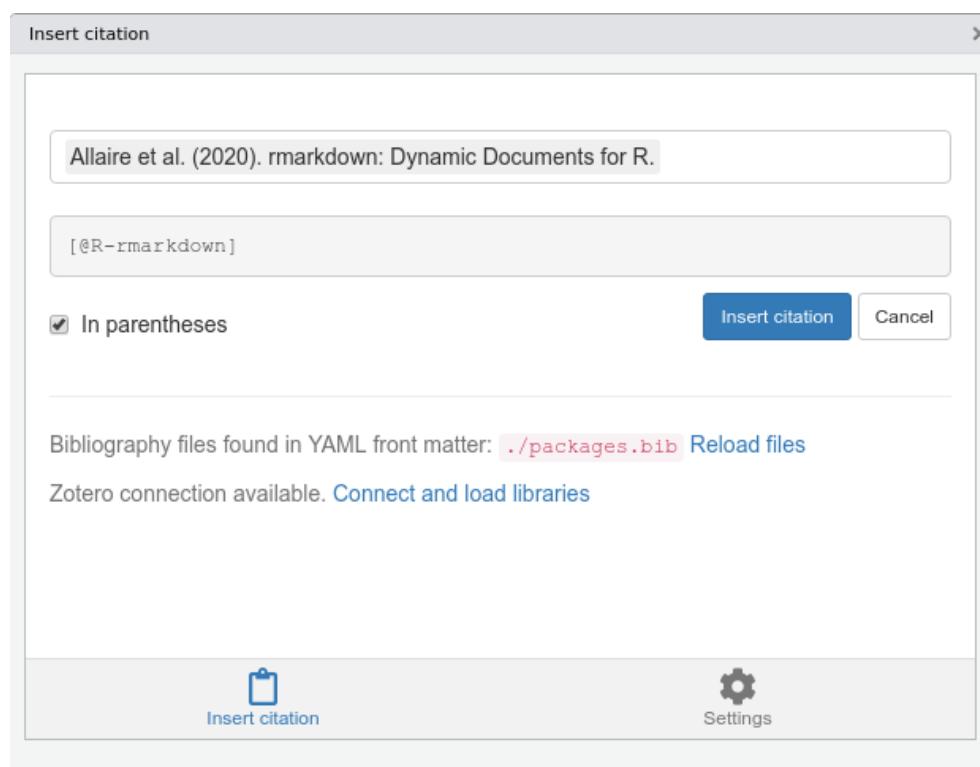


図5.2: `citr` パッケージの例

一方で、この記述がどう反映されるかは文献引用を処理するプログラムによって変化する。さらに厄介なことに、それぞれ全く違うプログラムであるがゆえに、設定に書き方も変わってくる。

R Markdown の文献引用は `pandoc` を経由して処理され、`pandoc` は現時点では `pandoc-citeproc`, `bibtex`, `biblatex` の選択をサポートしている。`pandoc-citeproc` 以外はもともと `LATEX` 用に作られたため、HTML では常に `pandoc-citeproc` で処理される。これはフォーマット引数の `citation_package` で指定できる。

表5.2: 引用プログラムごとの違い

item	HTML 対応	PDF 対応	日本語表示	指定名	文献ファイル	文献スタイル	スタイル
'pandoc-citeproc' (default)	TRUE	TRUE	TRUE	default	.json	.csl	csl:
'biblatex'	FALSE	TRUE	TRUE	biblatex	.bib/.bibtex	なし	bibli
'bibtex'	FALSE	TRUE	FALSE	natbib	.bib	.bst	bibli

- default: pandoc-citeproc を使用する
- biblatex: biblatex を使用する
- natbib: BiBTeX を使用し、本文中の参照には natbib.sty を使う。ただし本来の natbib の引用子オプションは使えない
  - natbiboptions: で number, authoryear などの natbib.sty のオプションを指定できる。

### 5.2.1 文献引用のフォーマット設定

これがややこしい。まず、pandoc-citeproc, bibtex, biblatex はそれぞれ引用文献リストの書式を記述するフォーマットがありそれぞれ拡張子は .csl, .bst, .bbx, である。前者は MS Word のスタイルと同じもので、XML で記述されている<sup>\*1</sup>。一方で .bst は逆ポーランド記法の構文だったりかなりアレである。そして biblatex はこのような専用の書式ファイルを使わず、細かい書式設定はすべて L<sup>A</sup>T<sub>E</sub>X のマクロで調整する想定で作られている。また、スタイルファイルを用意しなくとも指定できるスタイルがいくつか存在する。

bibtex および biblatex に関する設定は biblio-style だが、CSL ファイルの設定は csl を使う。

さらに日本語文献引用が必要な場合の特有の事情として、(1) 対応文字コードの問題 (2) 和文と欧文のスタイルが異なるという問題がある。まず前者について、bibtex は日本語など非 ASCII 文字に対応していない。よって従来 pBiBTeX または jBiBTeX が使用されていた。しかしこれも JIS 規格の範囲までしか対応していないため、一部の珍しい人名漢字などの表示ができなかった<sup>\*2</sup>。そのためさらに upBiBTeX という Unicode 対応版プログラムが作成されている。しかし R Markdown (および pandoc) はそのような日本語特有の事情などしないため、これらを使用することができない。よってオプション選択の範囲では、biblatex か pandoc-citeproc でしか日本語文献を表示できない。これは R Markdown で日本語技術文書を作る際の特にアレな障害の 1 つである。

<sup>\*1</sup> 簡単なカスタマイズなら CSL editor という Web サービスができる。しかしあくまで XML なので、あまり複雑な処理はできないことに注意する。

<sup>\*2</sup> pBiBTeX の UNICODE モードは内部で SJIS に変換しているだけなので、全ての Unicode に対応しているわけではない。

## 5.2 文献引用

表5.3: biblatex の ‘biblio-style’ で指定できるもの一覧

名称	概要
‘numeric’	[1] のような番号
‘alphabetic’	著者名の略称 + 出版年 2 枠
‘authoryear’	著者-出版年形式, natbib の標準と同じ
‘authortitle’	著者名のみ, リストでは出版年は後置され, 引用子では脚注になる
‘verbose’	authortitle と同じだが, 引用子にリスト同様の内容を出力する
‘reading’	個人的なリーディングリスト向け. ファイルやメモ欄も出力する
‘draft’	.bib ファイルの ID で表示. 名前通り下書き
‘debug’	‘.bib’ の全フィールドを表示

### 5.2.2 (TODO) pandoc-citeproc と CSL

### 5.2.3 (WIP) BibLaTeX

BibLaTeX の全てのオプションに対応しているわけではないので詳しいことは [BibLaTeX のドキュメント](#) を読んでいただきたい. 残念ながら, 日本語の情報は非常に乏しい. ここではよく使う style のことだけ言及する.

フロントマターの `biblio-style:` で指定できるのは, インクルード時の `style=` に指定できるものに対応する (表 5.3). つまり, 引用文献の見出しをどうするかである. これは引用リストと本文中の引用子のスタイル両方に影響する.

その他 apa, ieee など特定の学会が使用するスタイルも用意されているが, これらは基本欧文しか想定していないし, アカデミックの事情に詳しい人しかこれらの使用にこだわらないだろうから詳しくは解説しない. これらを含めたそれぞれの出力例は [https://www.overleaf.com/learn/latex/biblatex\\_bibliography\\_styles](https://www.overleaf.com/learn/latex/biblatex_bibliography_styles) に一覧があるのでそちらを参考に.

なお,もちろん引用リストのスタイルと引用子のスタイルを個別にすることはできるが, R Markdown および Pandoc にそのオプションを通す機能はない. (`bibstyle/citestyle` で分けられる, bbx/cbx)

現時点では各分野の学会で日本語文献に対応した BibLaTeX フォーマットを配布しているという情報は見つけられなかった. 参考として私のブログで対応のヒントについて書いた<sup>\*3</sup>.

TODO: その他の非ラテン文字, キリル文字, アラビア文字へプライ文字等は?

TODO: upBibTeX や bibtex で動作しない .bst ファイルの扱い.

<sup>\*3</sup> <https://ill-identified.hatenablog.com/entry/2020/09/20/231335>

なお、普段文献管理ソフトを使っていないが、数本程度の文献を引用リストに載せたい利用者は、`biblatex` の構文を利用して書くのがよいかかもしれない。例えばここに書いてあるように、その場合、デフォルトでは本文の引用は [1], [2] のような番号形式となる。`biblio-style: authoryear` とすることで、`natbib` のような「著者(出版年)」スタイルとなる。

<https://teastat.blogspot.com/2019/01/bookdown.html>

## 第6章

# (WIP) 簡単なレイアウト変更

### 6.1 HTML

#### 6.1.1 フォント変更

HTML は文字通り HTML で出力しているため, CSS の使い方次第でいくらでもデザインを変えることができる.

### 6.2 PDF

#### 6.2.1 フォント変更

PDF を生成する場合, ver 0.3 時点ではデフォルトのフォントを OS に応じて変えている.  
もし変更したい場合は YAML フロントマターの以下の項目を変更する

- `mainfont`: 欧文セリフフォント
- `sansfont`: 欧文サンセリフフォント
- `monofont`: 等幅フォント (コードの表示などに使用)
- `jfontpreset`: 和文フォントのプリセット
- `jmainfont`: 和文メインフォント (一般に明朝体を指定)
- `jsansfont`: 和文セリフフォント (一般にゴシック体を指定)
- `jmonofont`: 和文等幅フォント (コードの表示などに使用)

`jfontpreset` は `zxjafont` または `luatex-jp` によるプリセットで, 3種類のフォントを一括指定できる. 個別指定したフォントはこれを上書きする. 特にこだわりがないなら一括指定で良いが, ソースコードを多く掲載する場合は `M+` や `Ricty` などのフォントを用意すると良いだろう.

さらに, それぞれの項目には `options` と接尾辞のついた項目が用意されている. フォントの相対サイズが合わず不格好な場合は

```
mainfont: Palatinno
mainfontoptions:
- Scale=0.9
```

などと書いて調整できる。

インラインのフォント変更は TODO

## 第 7 章

# (WIP) rmdja による文書作成支援機能

### 7.0.1 クリエイティブ・コモンズの表記

Web 公開する文書ならばクリエイティブ・コモンズの表記をつけたいところだ。公式サイトで毎回発行するのは面倒なので表示する関数を用意にした。ハイパーリンクも付けるようしている。チャンクでは results="asis" オプションが必要になる。また、通常は echo=F を設定すべきだろう。冒頭の表記もこれで作成している。もちろんそれぞれの媒体に対応している。

文言の生成は未対応

### 7.0.2 ルビ表記

ルビはおそらく CJK 言語など一部の言語でしか使われていない(アラビア語とかペルシア語とかの補助記号は詳しく知らないが多分グリフとしてサポートされてるっぽいので無視)ため、ルビ表記も R Markdown ではサポートされていない。そこで簡単にルビを表示できる関数 rmdja::ruby() を用意した。インライン実行で使う。PDF での配置は pxrubrica.sty を利用したグループルビである。よって、1字ごとに配置(モノルビ)にしたいとか、突出指定とか、細かいことは HTML タグや CSS や LaTeX コマンド自分で書く。妥協案として、1字ごとに呼び出す手もある。

グループルビの例: とある科学の超電磁砲, 皇帝ライント, 格館,  
シュテッヒバルムシュロス  
ショウワルツ・ランツェンレイター キネウリワリ メキシコ  
黒色槍騎兵, 喜連瓜破, MEXICO

分割して出力した例: 喜連瓜破, 黒色槍騎兵,

TODO: それ以外にも便利機能を少しづつ増やしていく予定



## 第 IV 部

### 應用編

rmdja (2020-10-24 04:38)

## このパートについて

このパートでは、ここまで紹介した基本機能の応用で、さまざまな R パッケージやその他の外部プログラムの出力を埋め込む方法を紹介する



## 第 8 章

# 様々なグラフィックプログラムの埋め込み

### 8.1 tikz を使う

L<sup>A</sup>T<sub>E</sub>X で使われる tikzdevice を利用して、直接 tikz の記述による画像を埋め込むことができる。チャンクのエンジンを tikz とすることで使用でき、相互参照やキャプション、画像サイズの指定といったチャンクオプションも使える。図 8.1 は tikz で生成した図である。これは HTML でも表示できる。TODO: しかし現状では pdflatex 以外のエンジンに変更できないため、日本語表示が難しい。

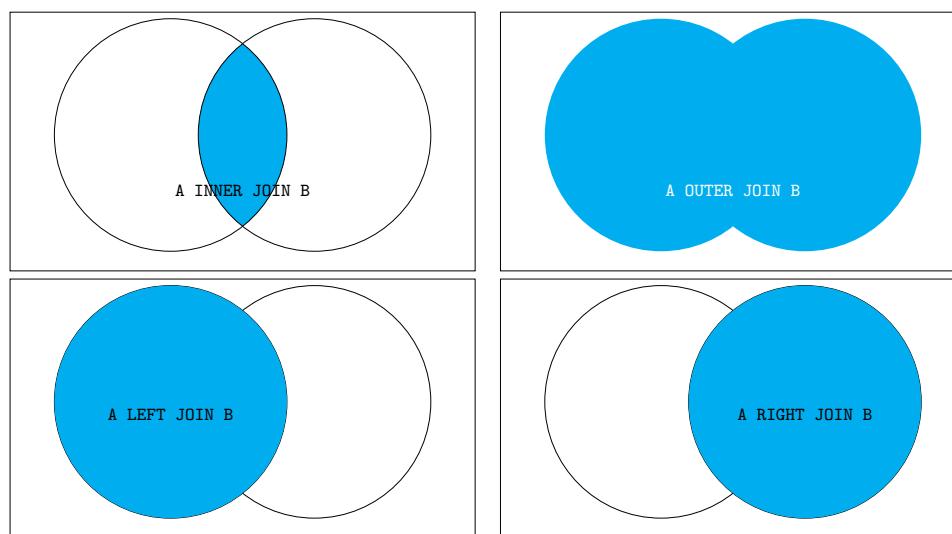


図8.1: tikz を利用した図の表示

### 8.2 Asymptote を使う

同様に、Asymptote のプログラムを埋め込むこともできる。私は Asymptote が分からないので RCB Ch. 15.9 Create graphics with Asymptote と同様のプログラムを書いておく。(図

8.2).

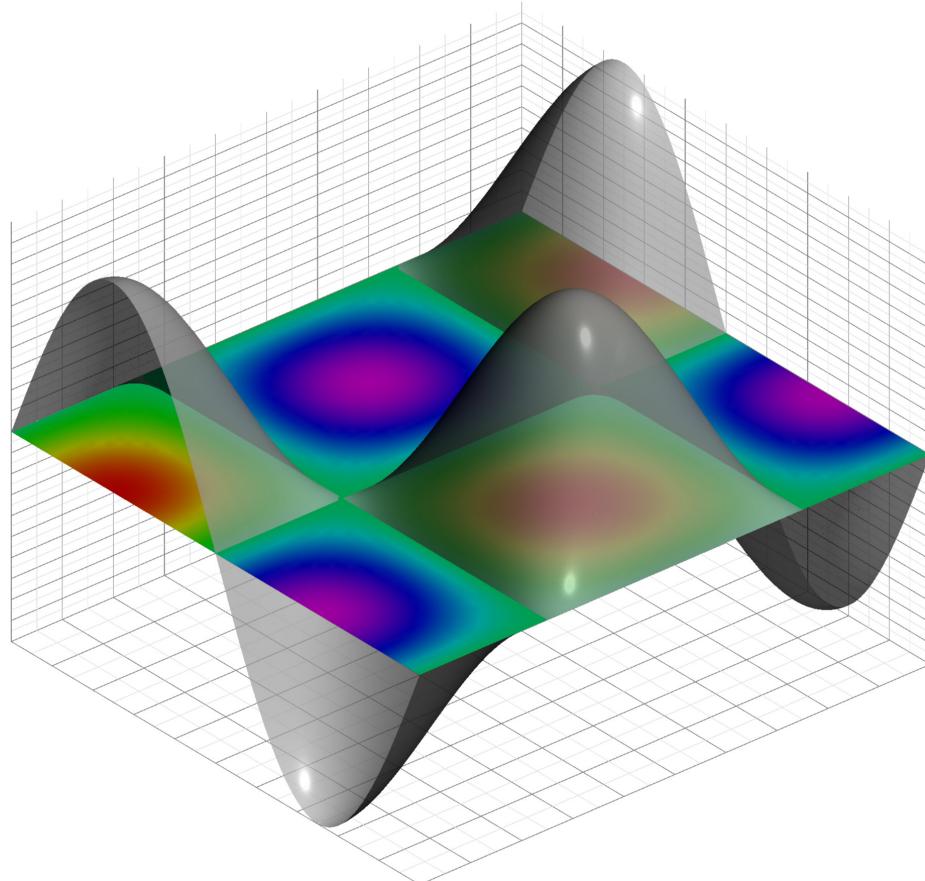


図8.2: Asymptote による画像

### 8.3 (TODO) その他のプログラム

D3.js なども使える

### 8.4 (TODO) その他の R プログラム

なお, DOT 言語は DiagrammeR パッケージを経由して使うこともできる<sup>\*1</sup>(図: 8.3). グラフィカルモデルの記述などはこちらのほうが簡単かもしれない.

```

01 DiagrammeR::grViz("digraph {
02 graph [layout = dot, rankdir = TB]
03
04 node [shape = rectangle]
05 rec1 [label = 'Step 1. 起床する']
```

<sup>\*1</sup> <https://bookdown.org/yihui/rmarkdown-cookbook/diagrams.html>

```
06 rec2 [label = 'Step 2. コードを書く']
07 rec3 [label = 'Step 3. ???']
08 rec4 [label = 'Step 4. 給料をもらう']
09
10 # edge definitions with the node IDs
11 rec1 -> rec2 -> rec3 -> rec4
12 }",
13 height = 500
14)
```

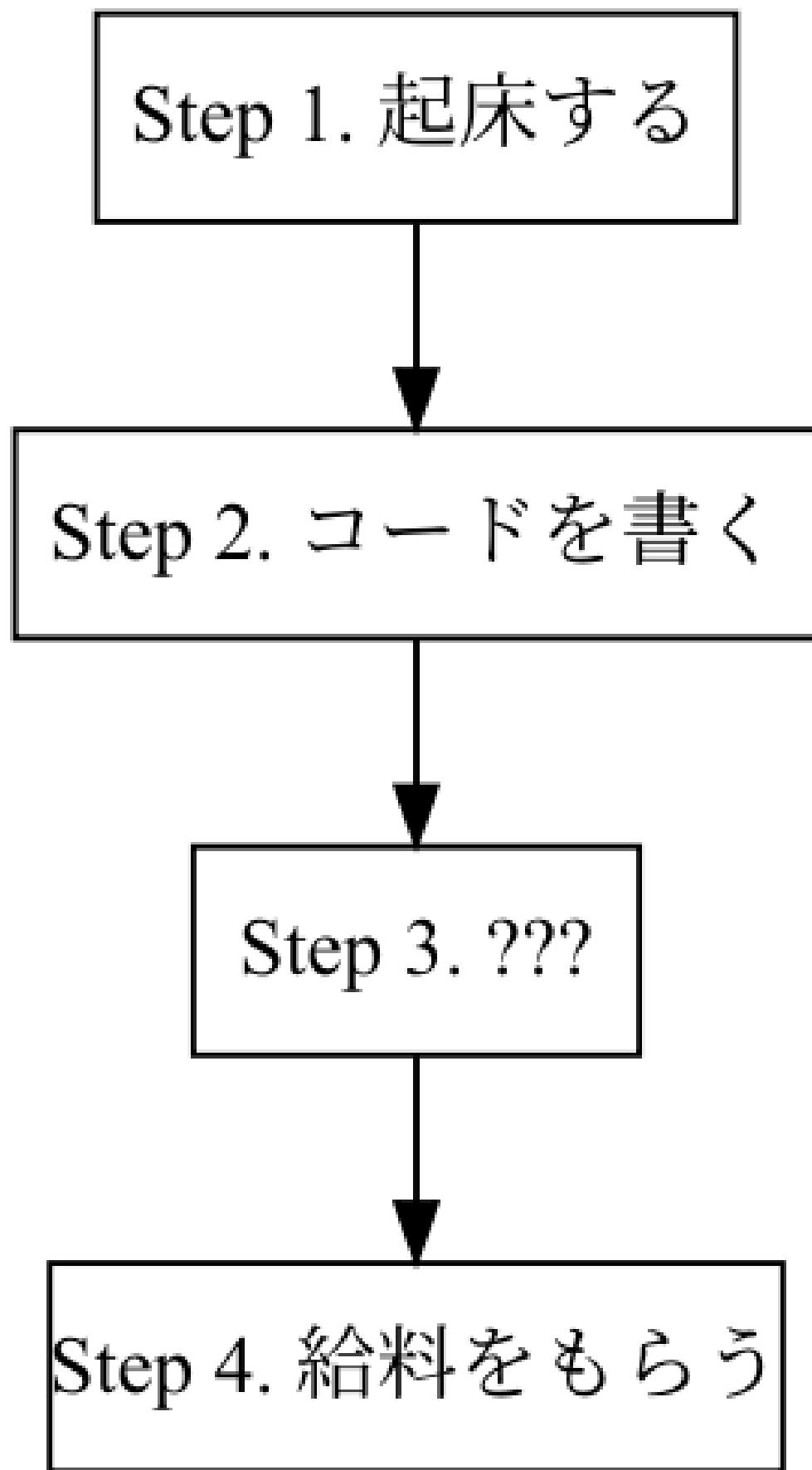


図8.3: DiagrammeR によるグラフィカルモデル (RCB, Ch. 4.15 より)

## 第9章

### 表のデザイン

#### 9.1 TeX/HTML を出力する関数

`stargazer` や `xtable` のように, HTML や `LATEX` 形式で表を出力してくれるパッケージがある. これらは `results='asis'` のチャンクオプションを指定することで関数の出力するテキストをそのまま埋め込むことができる. よって, あとは HTML か `LATEX` かといった出力形式の違いに気をつけければ表示できる(図 ??).

```
01 require(stargazer)
02 stargazer(mtcars,
03 type = if (knitr:::is_latex_output()) "latex" else "html",
04 header = F, title = "(ref:stargazer-title)"
05)
```

表9.1: `stargazer` による表の出力

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
mpg	32	20.091	6.027	10	15.4	22.8	34
cyl	32	6.188	1.786	4	4	8	8
disp	32	230.722	123.939	71	120.8	326	472
hp	32	146.688	68.563	52	96.5	180	335
drat	32	3.597	0.535	2.760	3.080	3.920	4.930
wt	32	3.217	0.978	1.513	2.581	3.610	5.424
qsec	32	17.849	1.787	14.500	16.892	18.900	22.900
vs	32	0.438	0.504	0	0	1	1
am	32	0.406	0.499	0	0	1	1
gear	32	3.688	0.738	3	3	4	5
carb	32	2.812	1.615	1	2	4	8

その他, `Hmisc::latex()`, `stats::xtable()` という古典的な関数がある。後者は名前の通り  $\text{\LaTeX}$  のソースをかなりの自由度で出力できるが、ここまでやるならもう最初から  $\text{\LaTeX}$  で書いたほうがいいのでは、というレベルである。 $\text{\LaTeX}$  に詳しくない場合、かえって難しいかも知れない。`stargazer` や後述する `kableExtra` ができる範囲でやったほうが簡単だし、過剰な装飾の表ができあがることも少ない。

## 9.2 `kableExtra` による表のスタイルのカスタマイズ



`kableExtra` パッケージは `knitr::kable` の拡張であり、様々なスタイルの表を出力できる。そしてそれは HTML でも PDF でも有効である。

まず、(これは `kable()` からある機能だが) `booktabs = T` を指定する。 $\text{\LaTeX}$  のデフォルトの表は罫線が多すぎる。罫線は牢獄の格子を暗示するため過剰な罫線にまみれた表を書いて眺めたりしていると精神が抑圧され有害であると心理学的にも証明されている(嘘)。しかし `booktabs = T` はこれだけで `booktabs.sty` に準拠した見やすいスタイルに変更してくれる。さらに `kableExtra` の機能として、表 9.2 にみられるように条件書式のような装飾が可能である。

```

01 that_cell <- c(rep(F, 7), T)
02 mtcars[1:8, 1:8] %>%
03 kbl(
04 booktabs = T, linesep = "", format = if (knitr:::is_latex_output()) "latex" else "html",
05 caption = "(ref:kableextra-color-cap)"
06) %>%
07 kable_paper(full_width = F) %>%
08 column_spec(2,
09 color = spec_color(mtcars$mpg[1:8]),
10 link = "https://haozhu233.github.io/kableExtra"
11) %>%
12 column_spec(6,
13 color = "white",
14 background = spec_color(mtcars$drat[1:8], end = 0.7),
15 popover = paste("am:", mtcars$am[1:8])
16) %>%
17 column_spec(9,
18 strikeout = that_cell, bold = that_cell,
19 color = c(rep("black", 7), "red")
20)

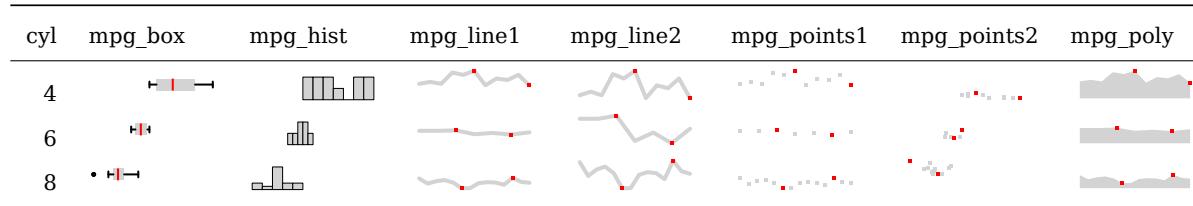
```

## 9.2 kableExtra による表のスタイルのカスタマイズ

表9.2: kabeExtra パッケージを利用した表の作成, 公式ドキュメントの用例より

	mpg	cyl	disp	hp	drat	wt	qsec	vs
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1

表9.3: kabeExtra パッケージによる表内グラフ, 公式ドキュメントの用例より



グラフのインライン挿入も可能である(表 9.3).

```

01 mpg_list <- split(mtcars$mpg, mtcars$cyl)
02 disp_list <- split(mtcars$disp, mtcars$cyl)
03 inline_plot <- data.frame(
04 cyl = c(4, 6, 8), mpg_box = "", mpg_hist = "",
05 mpg_line1 = "", mpg_line2 = "", mpg_points1 = "", mpg_points2 = "", mpg_poly = ""
06)
07 inline_plot %>%
08 kbl(
09 booktabs = T, format = if (knitr:::is_latex_output()) "latex" else "html",
10 caption = "(ref:kableextra-plot-cap)"
11) %>%
12 kable_paper(full_width = FALSE) %>%
13 column_spec(2, image = spec_boxplot(mpg_list)) %>%
14 column_spec(3, image = spec_hist(mpg_list)) %>%
15 column_spec(4, image = spec_plot(mpg_list, same_lim = TRUE)) %>%
16 column_spec(5, image = spec_plot(mpg_list, same_lim = FALSE)) %>%
17 column_spec(6, image = spec_plot(mpg_list, type = "p")) %>%
18 column_spec(7, image = spec_plot(mpg_list, disp_list, type = "p")) %>%
19 column_spec(8, image = spec_plot(mpg_list, polymin = 5))

```

その他細かい使用上の注意をいくつか挙げる。

- `kableExtra::` を参照するのではなく、最初にパッケージをロードしたほうが不具合が起きにくい。
- PDF 出力する場合、多くの `LATEX` パッケージのロードが必要だが、`rmdja` の PDF フォーマットはいずれもテンプレートに組み込んでいるため手動設定は必要ない。
- `knitr::kable()` または `kableExtra::kbl()` の `format` で HTML/tex の出力を決める。現在は判定が自動化されたとのことだが、まれに不具合があるという報告もみられる。よって、どちらも出力したい場合は上記のように `format = knitr::is_latex_output()` で条件分岐させるのが 1 つの手である。
- 表のキャプションは図のようにチャンクオプションに指定するのではなく、`kbl()/kable()` の `caption` 引数に指定する
- キャプション内に Markdown 記法や相互参照など特殊な構文を含めたい場合は、`escape = F` を指定する。

その他、テキストの回り込み、画像の挿入など様々なことが可能である。詳細は公式の解説である “[Create Awesome HTML Table with knitr::kable and kableExtra](#)” および [PDF 版](#) が役に立つ。

### 9.3 formattable パッケージとの併用

`formattable` パッケージは以前からある表を装飾するパッケージである。`kableExtra` との併用も可能だが、`LATEX` に対応しておらず、HTML タグを tex ファイルに吐き出してしまうため動作しない。PDF にも同様に表示するには [StackOverflow](#) で提案されているように、`webshot` 使うなど工夫が必要である。そこまでしてこの装飾にこだわるメリットは薄いと私は考えるので現在この問題に対応する予定はない。かわりに後述する `huxtable` または `gt` を使うべきだと考える。

### 9.4 huxtable パッケージによる作表

`huxtable` は HTML と `LATEX` に対応した作表パッケージであり、公式ドキュメントによると他の類似パッケージと比較して多機能であることを強調している。全体的に `tidyverse` を意識した構文が用意され、`kableExtra` のようにパイプラインを使った記述が扱る。なお `huxtable` のロゴはランダムに生成される。さらに「1行ごとに背景色を変更」「`stargazer` 風の表」などよく使われるスタイルを簡単に設定できるようになっていたり、はては `tidyverse`

## 9.4 huxtable パッケージによる作表

のような表のロング・ワイド変形機能まで備えている。

```

01 require(huxtable)
02 head(mtcars[1:5]) %>%
03 as_huxtable(add_rownames = "Model") %>%
04 set_caption("(ref:huxtable-example-cap)") %>%
05 set_bold(1, everywhere, T) %>%
06 theme_article() %>%
07 map_text_color(
08 everywhere, "mpg", by_colorspace("navy", "red", "yellow")
09) %>%
10 map_background_color(
11 everywhere, "hp", by_quantiles(0.8, c("white", "yellow"))
12) %>%
13 map_italic(everywhere, "Model", by_regex("Merc.*" = T)) %>%
14 set_number_format(col = "drat", value = fmt_percent(digits = 2))

```

表9.4: huxtable パッケージによる作表

<b>Model</b>	<b>mpg</b>	<b>cyl</b>	<b>disp</b>	<b>hp</b>	<b>drat</b>
Mazda RX4	<b>21</b>	6	160	<b>110</b>	390.00%
Mazda RX4 Wag	<b>21</b>	6	160	<b>110</b>	390.00%
Datsun 710	<b>22.8</b>	4	108	93	385.00%
Hornet 4 Drive	<b>21.4</b>	6	258	<b>110</b>	308.00%
Hornet Sportabout	<b>18.7</b>	8	360	<b>175</b>	315.00%
Valiant	<b>18.1</b>	6	225	105	276.00%

そのままでは罫線の設定が `set_top_border()`, `set_bottom_border()`, などしかなく、複雑な条件を指定するのが大変だが、`ggplot2` のテーマ関数のようにスタイルのプリセットが `theme_*`() の名前でいくつか用意されている。例えば上記では `theme_article()` という学術論文風テーマを適用し、表の上下とヘッダにだけ罫線を引いている。条件書式は `map_*`() 関数群で実行できる。また、フォーマットは `set_number_format()` に値を変換するフォーマット関数を与える形で適用できる。こちらはパーセンテージなども正しく表示できる。

また、`huxreg()` は名前の通り回帰分析の結果を表にするなど `stargazer` パッケージに似た機能を提供する。これも同じクラスなので同様にスタイル設定が可能である（表 9.5）。

```

01 lm1 <- lm(mpg ~ cyl, mtcars)
02 lm2 <- lm(mpg ~ cyl + hp, mtcars)
03 glm1 <- glm(I(mpg > 20) ~ cyl, mtcars,
04 family = binomial

```

```

05)
06 huxreg(lm1, lm2, glm1) %>%
07 set_caption("(ref:huxreg-example-cap)") %>%
08 set_text_color(everywhere, "model1", "green") %>%
09 set_text_color(everywhere, "model2", "blue")

```

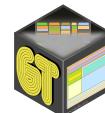
表9.5: huxtable::huxreg() による出力

	(1)	(2)	(3)
(Intercept)	37.885 *** (2.074)	36.908 *** (2.191)	64.400 (17449.775)
cyl	-2.876 *** (0.322)	-2.265 *** (0.576)	-10.781 (2908.296)
hp		-0.019 (0.015)	
N	32	32	32
R2	0.726	0.741	
logLik	-81.653	-80.781	-4.780
AIC	169.306	169.562	13.561

\*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05.

その他の作例は CRAN の公式 vignettes を参考にせよ.

## 9.5 (TODO) gt パッケージによる作表



gt パッケージは RStudio 社によって開発されている。kableExtra や huxtable が行や列を単位にスタイルを設定する設計だったのに対し, gt はデータフレームにヘッダ, 列ラベル, スタブ, 脚注, といったパートに分けて設定して表を作成する。

gt はあまり複雑な条件書式の機能はない。そのかわりヘッダや行ラベルの表現に自由度がある。たとえば “Case Study: gtcars” には

```

01 require(gt)
02 gtcars %>%
03 dplyr::group_by(ctry_origin) %>%
04 dplyr::top_n(2) %>%

```

```

05 dplyr::ungroup() %>%
06 dplyr::filter(ctry_origin != "United Kingdom") %>%
07 dplyr::group_by(ctry_origin) %>%
08 gt() %>%
09 as_latex()

```

mfr	model	year	trim	bdy_style	hp	hp_rpm	trq	trq_rpm	mpg_c	mpg_h	drivetrain	trsmn
<b>United States</b>												
Ford	GT	2017	Base Coupe	coupe	647	6250	550	5900	11	18	rwd	7a
Dodge	Viper	2017	GT Coupe	coupe	645	5000	600	5000	12	19	rwd	6m
<b>Italy</b>												
Ferrari	LaFerrari	2015	Base Coupe	coupe	949	9000	664	6750	12	16	rwd	7a
Lamborghini	Aventador	2015	LP 700-4 Coupe	coupe	700	8250	507	5500	11	18	awd	7a
<b>Japan</b>												
Acura	NSX	2017	Base Coupe	coupe	573	6500	476	2000	21	22	awd	9a
Nissan	GT-R	2016	Premium Coupe	coupe	545	6400	436	3200	16	22	awd	6a
<b>Germany</b>												
BMW	i8	2016	Mega World Coupe	coupe	357	5800	420	3700	28	29	awd	6am
Mercedes-Benz	AMG GT	2016	S Coupe	coupe	503	6250	479	1750	16	22	rwd	7a

よって, `gt` パッケージはこのようにカテゴリや列をグループ化して掲載する表を比較的簡単に作りやすいことがわかる (個人的には単純なマトリクスで表現できないものを表にすることは気が進まないが....).

一方で `gt` の問題点として, 相互参照が面倒だというものがある. TODO

## 9.6 その他の作表パッケージ

`df_print` チャンクオプションで, デフォルトの表示方法を変更することもできる. ここまで `kable` での出力を前提としていたが, 長大な表を掲載したい場合, `paged` オプションがある.

それ以外有名なパッケージとして, `DT`, `flextable` がある. 前者はインタラクティブな表ウィジェットを作成し, 後者はむしろ Word へのエクスポート機能をフィーチャーしており `rmdja` の方向性とは異なる. 加えて既に紹介した `huxtable` または `gt` でだいたいのことはできるため, これ以上の紹介は避ける.

RCB 10.3 Other packages for creating tables も参考にせよ.



## 第 10 章

# 文献引用

### 10.1 (u)pBIBT<sub>E</sub>X を使う

日本語対応 .bst ファイルを使いたい場合は少しトリッキーな操作が必要になる。rmarkdown は tinytex というパッケージでインストールされたスタンダードアローンな処理系で PDF を生成している。冒頭のチャンクで options(tinytex.latexmk.emulation = F) を指定することで、自分のマシンにインストールされている普段使っている処理系に処理させることができる。さらに rmdja では natbib を指定した場合に自動でカレントディレクトリに .latexmkrc をコピーするようにしている。しかしログが残らないなどデバッグしづらいところがあるため、このやり方はやや使いづらい。



## 第 11 章

### (TODO) Web アプレットの挿入

11.0.1 TODO: plotly

11.0.2 TODO: shiny



## 第 12 章

# Python スクリプトの埋め込み

Python スクリプトを埋め込むこともできる。方法は 2 通りあり、都度システムコマンドから呼び出す方法と、`reticulate` パッケージを使うものがある。`reticulate` 登場以前はチャンクごとに呼び出していたため複数のチャンクに分割して記述するのが難しかったが、現在は `reticulate` パッケージを利用することで R と同じような感覚で、あるいは Jupyter のコードセルと同じような感覚で書ける。

`pyenv` を使用する場合、共有ライブラリのインストールが必要なことに注意<sup>\*1</sup>。

初めて使う場合は先に `reticulate` 単体で Python が実行できるか検証してからの方が良い。

```
01 require(reticulate)
02 # Python エンジンを認識しているか確認
03 py_discover_config()
04 repl_python()
05
06 # 以下、Python コマンド実行、`exit` で抜ける
```

現在 (`knitr` 1.18 以降) は R Markdown はデフォルトで `reticulate` を使う。システムコマンド経由にしたい場合はチャンクオプション `python.reticulate=F` を設定する、あるいは `reticulate::eng_python` に変わったエンジンを自作する<sup>\*2</sup>。

あとはチャンクのエンジンに `r` ではなく `python` を指定することで Python コードを埋め込める。

`matplotlib` エンジンで描いたグラフに日本語フォントを埋め込む場合、`matplotlib-japreset` を使えば必要な設定を一括で行う。



`matplotlib-japreset` は現在、Linux 以外での動作保証をしていない

<sup>\*1</sup> 詳しくはこちらを参考に <https://ill-identified.hatenablog.com/entry/2019/11/15/010746>

<sup>\*2</sup> 参考: [https://rstudio.github.io/reticulate/articles/r\\_markdown.html](https://rstudio.github.io/reticulate/articles/r_markdown.html)

```
pip install -U git+https://github.com/Gedevan-Aleksizde/matplotlib-japreset.git@master
```

```
01 from matplotlib_japreset import mplj_cairo
```

Noto Sans CJK JP not found  
 matplotlib-japreset Cairo mode  
 font: Noto Sans CJK JP

```
01 from matplotlib import rcParams

02 from plotnine import *

03 from plotnine.data import mtcars

04 rcParams['font.family']
```

```
['Noto Sans CJK JP']
```

matplotlib-japresetによって主要OSで標準インストールされている日本語フォントが自動的に選ばれる。気に入らない場合はrcParams['font.family']に好きなフォント名を上書きする。

```
01 ggplot(mtcars, aes('wt', 'mpg', color='factor(gear)'))

02) + geom_point() + stat_smooth(method='lm') + facet_wrap(~gear')
```

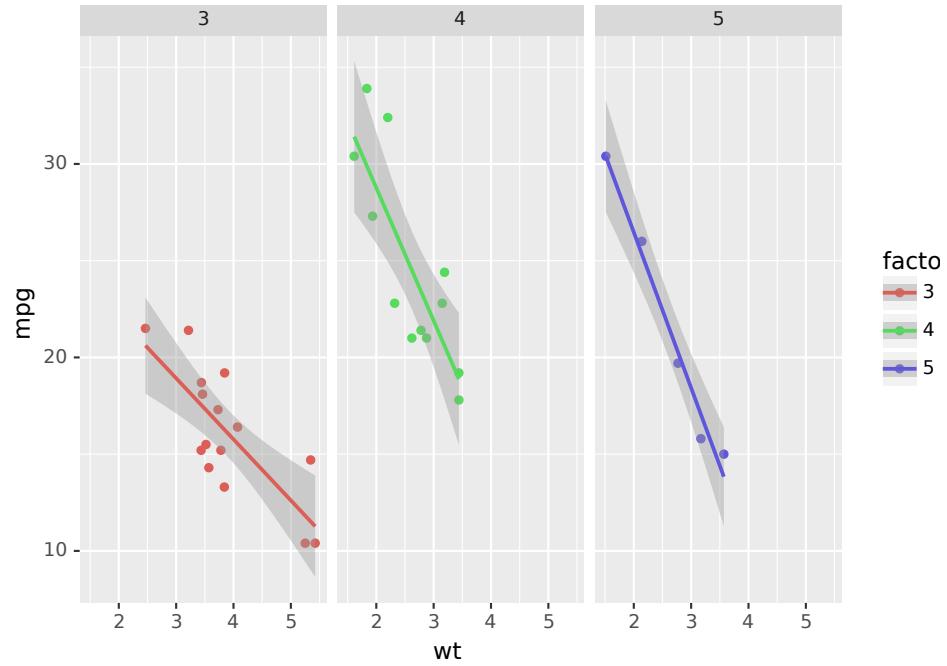


図12.1: Pythonのplotnineによるグラフ表示例

## 12.1 Python のグラフィックに関する制約

matplotlib ベースのグラフィックを出力したい場合、いくつかの制約がある。

- matplotlib > 3.2 では R がクラッシュするため、3.2 を使用する必要
- axes を使用した場合 (subplot などが依存)、`matplotlib.pyplot.show` の呼び出しと、次に別のグラフを呼び出す前の `matplotlib.pyplot.close()` が必要
- `seaborn.FacetGrid` を Cairo デバイスで保存できない (= フォントのサブセット化処理が複雑になる)

よって、現状では matplotlib エンジンでグラフィックを描くときはなるべく `plotnine` を使ったほうがトラブルが少ない。

また `plotly`, `bokeh` などの matplotlib に依存しないモジュールは PDF には対応していないため直接表示できない。一旦画像を保存して、あらためて画像ファイルを埋め込む必要がある。



## 第 V 部

製本と多様な形式への対応

rmdja (2020-10-24 04:38)

## 第 13 章

# PDF の文書クラス

HTML は利用者側が見え方をある程度カスタマイズできる。かつて存在した Evernote Clearly やカスタム CSS を使って、そのぶん PDF は作成者側がよりレイアウトに注意を払うことになるだろう。本稿では文章の区切りを章立てにしている。しかし PDF 数十ページしかない文書を大きな文字サイズの見出しで区切るのは少しものしい感じがする。YAML フロントマターを変更すれば、トップレベルの見出しを変更できる。

`pdf book in Japanese` は “book” ということで書籍の組版をデフォルト設定している。もう少し小規模な文書ならば、レポートや論文記事形式のほうが良いかもしない。例えば、以下のように指定する。

```
documentclass: bxjsreport
```

`documentclass` には LaTeX の文書クラスファイル (.cls) ならなんでも与えることができるが、`XELaTeX` または `LuaETeX` で日本語文書を作成することを想定しているため、以下 2 種類の BXjscls の文書クラス<sup>\*1</sup>の中から選ぶとよい。デフォルトは `bxjsbook` なので、これは明示的に指定する必要はない。

- `bxjsbook`
- `bxjsreport`

このうち、`bxjsbook` が `pdf book in Japanese` のデフォルト設定となっている。`rmdja::texlogo("LaTeX")` の文書クラスは、行間や見出しのレイアウトなどを日本語文書に準じたものにするが、それ以外の細かい調整は `_output.yml` や `_bookdown.yml` の設定を書き換えて調整する。それでも不十分な場合は、.tex ファイルや pandoc テンプレートを直接編集したり、追加のスタイルファイルを読み込んだりするしかない。

---

<sup>\*1</sup> 詳細はここにあるドキュメント参照: <https://www.ctan.org/pkg/bxjscls> 但し、スライド用クラスである `bxjsslide` の使用は想定していない。また、`bxjsarticle` を使う場合は後述の `pdf article in Japanese` テンプレートから作成したほうがよい。さらに `rmdja::texlogo("LuaLaTeX")` を使用するならば `luatex-ja` で提供される日本語文書クラスも指定することができるが、あまりつかったことがないためレイアウトに不備があるかもしれない。以降は PDF ファイルで出力できる各形式についてこまかく解説する。

## 13.1 プrezentation資料の作成

`beamer_presentation_ja` は rmdja の最初期からあったフォーマットで、そもそも当初はこれを作るのが目的だった。このフォーマットは Beamer を使用してプレゼンテーション用スライドを PDF ファイルで作成する。Beamer は `rmdja::texlogo("LaTeX")` の文書クラスの 1 つで、`rmarkdown::beamer_presentation` はこれを利用しているが、例によって日本語表示は想定されていないため、そのためのもうろろの調整込みのラッパーフォーマットである。ただしスライド資料なので組版の禁則処理のような細かい調整は用意していない。`rmdja` ではスライドは PDF 以外の出力は不可能である<sup>\*2</sup>。

通常の文書と違い、デザインを決めるのは主に `theme` である。デフォルトでは `metropolis`<sup>\*3</sup> である。日本語表示のために調整してあるものの、日本語表示と直接関係ない部分はカスタマイズの余地としている。テンプレートには私の好みが反映された調整（プログレスバーの位置調整）が YAML フロントマターに直接書き込まれている。

また、日本語表示と直接関係ないアレンジとして、文献引用を行った場合の参考文献リストの表示が

1. 「参考文献」というセクションタイトルのみのスライドが冒頭に自動で挿入される
2. 引用された文献の数に応じてフレームが自動分割される
3. これらの参考文献フレームでは上部のタイトルが表示されない
4. 文字サイズが脚注サイズに縮小

という設定になっている。通常のプレゼンテーションでは大量の参考文献を読み上げることは少ないという想定で、紙面の限られたスライドに参考文献のみ羅列したスライドでページ数が増えないように考慮したためこうした。これは既に作成した `my_latex_templates` のテンプレートとほぼ同じである。

実際の表示例は `examples` にある。

### 13.1.1 主な設定

フォーマット関数が用意する設定について、Beamer 特有の設定をいくつか紹介する。

- プログラムはデフォルトで非表示 (`echo=F`)
- 出力する画像の大きさ `fig_width`, `fig_height` は `beamer` のデフォルトの大きさに連動している。そして `out_width`, `out_height` はいずれも "100%" にしているため、概ね `beamer` の画面と同じ大きさになる。

<sup>\*2</sup> HTML 形式のスライドはサポート対象外である。日本語文書特有の処理はあまりないということ、普段と違う環境で表示することの多いであろうスライド資料はなるべく環境に依存しない方法で表示すべきと考えているのが理由である。HTML でスライドを作成したい場合、次のページが参考になる：  
[https://kazutan.github.io/SapporoR6/rmd\\_slide.html#/](https://kazutan.github.io/SapporoR6/rmd_slide.html#/)

<sup>\*3</sup> なお `metropolis` テーマ開発者は Fira Sans フォントの使用を想定しており、ビルト時にフォントがないという警告が出ることがあるが無視して良い。(参考: <https://github.com/matze/mtheme/issues/280>)

### 13.2 (WIP) 卒業論文の作成

- プログラムに行番号を表示する `code_rownumber` は `FALSE` にしている
- テーマは `metropolis` を使っているが、昔ながらのテーマも可能である。昔からあるテーマの比較には `Beamer Theme Matrix` というページが便利である。他にも近年登場したテーマがいくつか存在するが、日本語をうまく表示できなかったり `rmdja::texlogo("XeLaTeX")/rmdja::texlogo("LuaLaTeX")` に対応していなかったりするものも多い。他に日本語に対応したテーマとして、`sakuratheme` が存在する。
- `beamer` のアスペクト比はデフォルトで `4:3` であり、YAML フロントマターで指定できる。例えば `16:9` に変更したい場合

```
classoption:
 - aspectratio=169
```

となる。指定可能なのは `3:2, 4:3, 5:4, 14:1, 14:9, 16:9, 16:10` で、上記のようにコロンを抜いて数字のみで指定する。なお `classoption` は `LATEX` で PDF を使う場合に有効になる。

詳細は `beamer` の公式ドキュメントを参考に。

### 13.2 (WIP) 卒業論文の作成

卒業論文... というか学術論文での体裁で PDF ファイルを作成することも可能である。`pdf article in Japanese` という名前のテンプレートで論文形式の PDF ファイルを用意している — HTML 形式で論文提出を要求するという話は聞いたことがないので PDF のみ対応している。

書籍形式との違いは、

- 文書の見出しが「X 章」ではなく「1. YYYY」のようになる（したがって、Rmd ファイルで # で記述した見出しあは、PDF ではセクションタイトルとなる）
- 余白が見開きを想定したものでなくなる

など些細である。実際のところ、文書テンプレートの設定を少しいじっている程度のことしかしていない。テンプレートを開いて確認すればわかるように、

```
output:
 rmdja::pdf_book_ja:
 toc: false
 pandoc_args:
 - '--top-level-division=section'
documentclass: bxjsarticle
```

という設定を追加しているだけである<sup>4</sup>。

---

<sup>4</sup> このテンプレートでは論文形式のフォーマットとして `bxjsarticle` を使用している。LuaLATEXを使用するならば代わりに `ltjsarticle` クラスも使用可能なはずだが、私は使ったことがないので説明を省く。

大学によっては論文の体裁が細かく指定されている場合もあるかもしれない。例えば1ページあたりの行数や、1行あたりの文字数とか。例えば1ページあたり50行、1行あたり40字とする場合、以下のような設定を追加する。ただし、行数は図表の挿入などで変動するし、プロポーショナルフォントや字幅の異なる欧文を多用すれば1行あたりの文字数は多くなりうる。

```
classoptions:
 - 'number-of-lines=50'
 - 'textwidth=40zw'
```

さらに、カラー印刷が許容されない場合もある。ggplot2 は `scale_*_grey()` などでカラーパレットを簡単に変更できる（図 13.1）。

```
01 ggplot(
02 mutate(mtcars, cyl = factor(cyl)),
03 aes(x = mpg, y = wt, color = cyl)
04) +
05 geom_point() +
06 labs(x = "マイル毎米ガロン", y = "重量 (1000 ポンド)") +
07 theme_bw() +
08 scale_color_grey() +
09 scale_fill_grey()
```

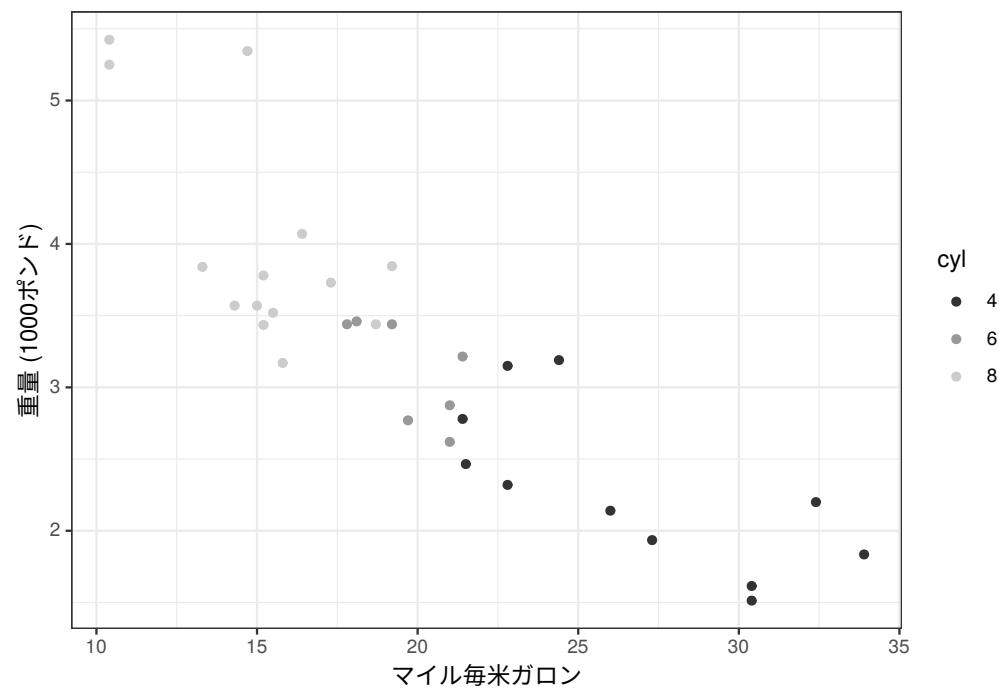


図13.1: グレースケールでのプロット

## 13.3 (WIP) 小説の執筆

作家の京極夏彦は自分の作品を1ページごとに切り取っても作品として成立するようなレイアウトにこだわっているらしいが、すでに説明したように技術文書や学術論文では図表の配置や改行などにあまりこだわりがない。しかし、不可能ではない。HTMLでは難しいが(不可能ではないがHTMLでやるメリットが感じられないで対応する気がない)、PDFではある程度のレイアウトの制御が可能である。ただし、本当に厳格なJIS準拠の組版にこだわるなら、LaTeXを直接編集しなければならない。

rmdjaで用意されている縦書き文書テンプレート pdf vertical writing in Japanese は、`jlreq`を利用して<sup>5</sup>縦書き文書のPDFを作成する(図: 13.2)。HTMLには未対応である。

4

## 第一章 縦書きの例

字下げには全角空白が必要。

雪の結晶の良い顕微鏡写真を撮るには、気温が零度以下になつている必要があることは勿論である。あらかじめ顕微鏡をよく冷しておいて、別に濡れないように冷しておいた硝子板<sup>ガラス</sup>に結晶を受けて、普通の顕微鏡写真を撮るようにして写せばよいのである。気温が零下五度以下であると大分楽であるが、零度に近いような時はまごまごしていると肝心の結晶がとけたり、冷しておいた硝子板に一面に霜がついたりしてなかなか厄介になる。何よりも大切なことは、硝子板に載った沢山の結晶の中どれを撮るかを決める敏速な決断である。まず眼で見て、次に顕微鏡下で写真を撮る価値があるか否かを調べて、決断をして、暗函をかぶせて、さてシャッターを切るまでの時間が、馴れてくれば二十秒位で出来るようになる。この間が五分位に感じられるようになれば大丈夫である。覗きながら、写真を撮るような便利な器械は、特に寒い所では故障が多くて駄目なようと思われる。一冬にただ一度見るか見ない位の珍しい結晶の時に、得てそのような故障が起りやすいようである。

初めの年は廊下の吹きさらしの寒い所を選んで有り合せの顕微鏡で写真を撮ってみたのであるが、結果はなかなか面白かった。北海道は雪の結晶の種類には極めて恵まれているようで、わずか一冬の

この小屋は十勝の吹上温泉の近くに在る。周囲は亭々たる蝦夷松と樺マスの木のように雪に枝を垂れています。白樺の化けたような巨樹が、細い空に向って伸している。これらの出している固体の表面はことごとく黒い樹幹を除いては周囲は、た日の青空のみが鮮かな濃い色彩をれる日は極めて稀れで、冬半年の間、写真を見るような世界になってしまいようである。雪が降り出すと四辻にぱきっとこの世界のみに見られる。このであろうと思われる。もっとも、ヘストックを差し込んで穴を作ると、うな色を呈すること位は誰にも見らうに鮮かな色彩は札幌附近的雪にも、多分積雪の中にも結晶がかなり完全に結晶の面が沢山あるためによるもの

雪の結晶は驚くべく繊細な形をし  
れども比較にならぬ位の美しさを見  
が、丁度開きかけた薔薇の花弁の縁

図13.2: 縦書き文書の出力例

<sup>5</sup> luatex-jaにも縦書き文書クラス `ltjt` シリーズが存在するが、公式ドキュメントにすら詳しい解説がなかったため採用しなかった。



現在、縦書き文書では図のようにゴシック体になってしまうことがある。



エディタは横書きのままである。また、段落改行も Markdown のルールに則して 1 行空けによってなされる。

小説家になろうとかに自動投稿する機能もいまのところない。

## 第 14 章

# 製本方法の詳細

冒頭のチュートリアルで行った製本(ビルド)の仕組みをもう少し詳しく解説する。

bookdown-demo を念頭に置いた解説。rmdja も基本的に同じ。

- `index.Rmd`: デフォルトで最初に読み込まれる Rmd ファイル(名前を変える機能もあるが、現時点では不具合が起こりやすいのでおすすめしない)
- それ以外の Rmd ファイル: 連結して読み込むことが可能
- `_output.yml`: マルチメディア展開のための設定。PDF, HTML, EPUB それぞれの設定を書く
- `_bookdown.yml`: bookdown のレイアウト設定
- その他の設定ファイル: その他製本に必要なもの、画像ファイル、.css ファイル、.bib 等

`_output.yaml`, `_bookdown.yaml` は `index.Rmd` のヘッダに書くこともできるが、長くなりすぎるので分割できる。`bookdown::render_book()` 関数は、ルートディレクトリのこれらを自動で読み込んでくれる。

### 14.1 ファイル構成

これらのファイルの中身を解説する。

#### 14.1.1 `_output.yml`

本来の YAML の `output:` 以下の記述をこの `_output.yml` ファイルに書くことができる。  
`output:` を複数書くと `rmarkdown::render_site()` やビルドツールでそれぞれの形式に一括作成してくれる。

```
output:
 bookdown::gitbook:
 lib_dir: assets
```

```
split_by: section
config:
 toolbar:
 position: static
bookdown::pdf_book:
 keep_tex: yes
bookdown::html_book:
 css: toc.css
documentclass: book
```

詳しくは BKD “Ch. 3 Output Formats” の章を.

#### 14.1.2 \_bookdown.yml

\_bookdown.yml も index.Rmd の YAML ヘッダの bookdown: 以下に対応する内容を書くことができる. 例えはどの Rmd ファイルを読み込むかとか, LaTeX のときだけ, HTML のときだけ読み込むような設定も可能.

<https://ill-identified.hatenablog.com/entry/2020/09/05/202403>

詳しくは, BKD Ch. 4.4 Configuration

Build ページから文書をビルドするには, index.Rmd の YAML ヘッダに site: bookdown::bookdown\_site を書く必要がある. さらに, index.Rmd をプロジェクトディレクトリのルートに置いていない場合は, ツールバーの Build -> Configure Build Tools... から index.Rmd を置いているディレクトリを site ディレクトリとする設定が必要になる(図 14.1, 14.2).

または, bookdown::render\_book("index.Rmd", "rmdja::pdf\_book\_ja") などでも実行できるから, コマンドラインからも実行できる. 同時製本は rmarkdown::render\_site().

#### 14.1 ファイル構成

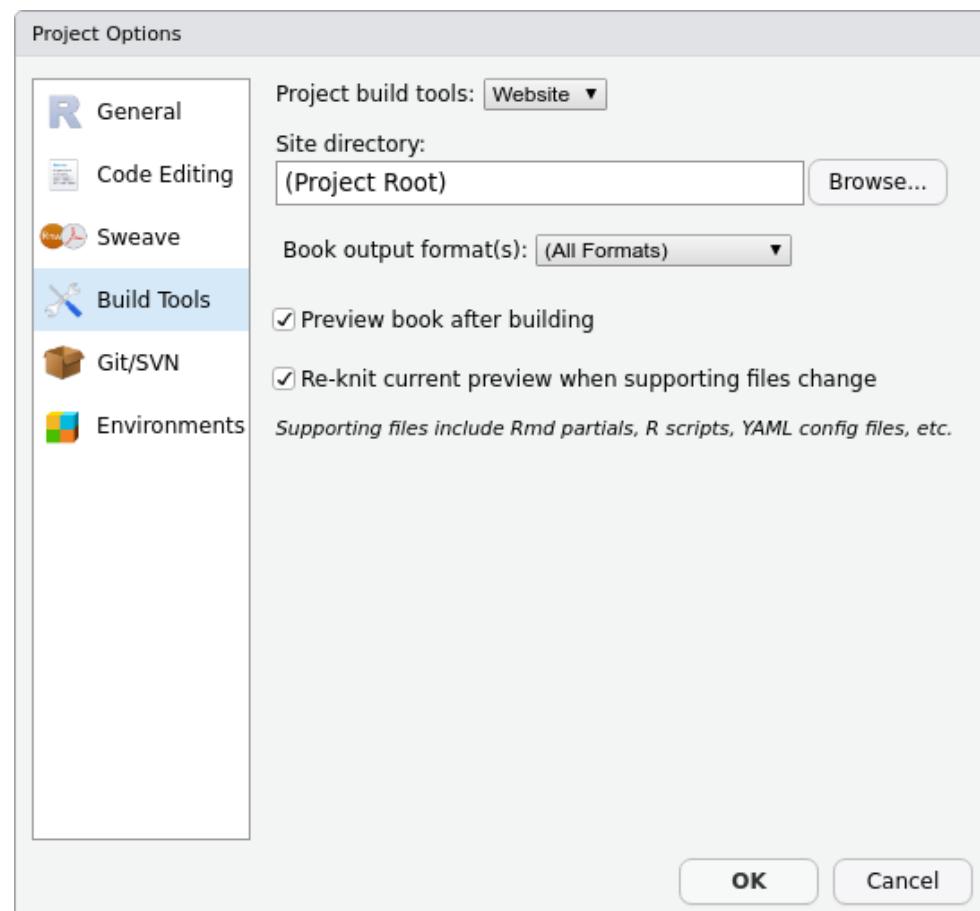


図14.1: Build ペーンの手動設定



図14.2: Build ペーンの手動設定



## 第 15 章

# 出力形式による表現の限界

### 15.1 HTML と PDF で処理を場合分けする

出力方法で言えば、HTML と PDF に大別できる。Rmd は HTML タグも LaTeX コマンドも受け付けるが、それぞれ HTML と PDF に変換する際にしか反映できない。よって、例えば複雑な図表を LaTeX コマンドでじかに Rmd ファイルに書いてしまった場合、HTML では表示されない。

紙媒体と電子媒体では表現できることに差がある。例えば紙はあらゆる環境で同じような見た目になるが、ハイパーリンクは付けられないし、一度出版してしまうと修正は容易ではない。PDF の見た目も読者の環境に依存しにくいが、やはり更新が容易ではない。

bookdown には既に印刷された本の中身を書き換えるする機能はないが、出力ごとに内容を変えることで、PDF にのみ更新履歴を表示することはできる。

knitr::is\_latex\_output(), knitr::is\_html\_output() などは、knit 時にどの媒体への変換処理なのかを判定するのに使える。rmdja::ruby() もこの機能を利用しているし、本文中の LATEX のロゴも HTML と PDF で使い分けている。

また、\_bookdown.yml の設定、rmd\_files は、媒体別に設定することができる。

```
rmd_files:
 html:
 - index.Rmd
 - html-only.Rmd
 latex:
 - index.Rmd
 - latex-only.Rmd
```

## 15.2 絵文字の出力

絵文字を HTML でも PDF でも出力したい場合, `\coloremoji{}` のように絵文字を囲む。ただし, RStudio のエディタは一部のマルチバイト文字の表示に対応していないので予期せぬ不具合に注意する。

現在の主要 Web ブラウザでは、特に設定せずとも Unicode 絵文字をカラー画像に置き換えて表示できるものが多い。しかし PDF 生成時には明示的にフォントを指定するか、画像に置き換える記述が必要である。その実現のため `bxcoloremoji` という LaTeX パッケージ<sup>1</sup>を利用する。このパッケージは CTAN に登録されていないため、別途インストールする必要がある。

## 15.3 画像の保存形式

技術文書での画像の多くはプロットなど単純な図形なので、写真などを掲載するのではなく、PDF で出力する場合はプロット画像も PDF にするのが望ましい。JPG や PNG などのラスタ画像では拡大すると粗くなるが、PDF などのベクタ画像ならば拡大しても粗くならず、かつ単純な図形ならばファイルサイズも小さく済むことが多い。一方で HTML は通常 Web ブラウザで閲覧するため、PDF に対応していないことが多い。HTML でベクタ画像を掲載したい場合は **SVG 形式**で出力する。

R による SVG への出力は、従来組み込みの `SVG()` で行うことが多かったが、近年は新たなパッケージが出ている。有力なのは `svglite` と `rsvg` である。

<https://oku.edu.mie-u.ac.jp/~okumura/stat/svg.html>

`rsvg` のほうが高性能だが、`knitr` で対応しているのは `svglite` なので簡単に使いたいならこちらを推奨する。

## 15.4 デフォルトの保存形式

デフォルトでは、PDF は `cairo_pdf`、HTML では解像度を高めに設定した `PNG` を使用している。これは、件数の多い散布図など、ベクタ形式ではファイルサイズが大きくなりすぎる場合もありうるための判断である。

画像形式を変更したい場合は、チャンクオプションの `dev` で、オプションは `dev.args=list(...)` で変更できる。

<https://bookdown.org/yihui/rmarkdown-cookbook/graphical-device.html>

---

<sup>1</sup> <https://github.com/zr-tex8r/BXcoloremoji>

## 第 16 章

# 製本した文書を配布する

### 16.1 Wep ページのホスティング

HTML ファイルは様々な配布方法がある。もちろん自分でサーバを立てても良い。特に簡単なのは以下の 2 点である。

1. github pages を利用する
2. bookdown.org に投稿する

(1) の詳細は [github.com の公式ドキュメント](#) を見るのが一番良いだろう。

既に bookdown で作成した文書を公開している例は多数ある。例えば既に何度も言及した公式解説サイトはそれじたいが bookdown で作られているし, "R for Data Science" (Wickham and Grolemund 2016)<sup>\*1</sup> は、内容の良さも含め一見に値する。また, "Hands-On Data Visualization: Interactive Storytelling from Spreadsheets to Code" (Dougherty and Ilyankou [forthcoming](#)) という本<sup>\*2</sup>が来年出るらしい。そして面白いことにこれは R Google スプレッドシートとか R 以外の Web 上のサービスの利用法を紹介する文書である。

これらはいずれもソースコードまで公開されている。もちろんここでいうソースコードとは、本文中のプログラムだけでなく文書を生成する Rmd ファイルなども含める。

それ以外にも有名無名の多くのドキュメントが公開されているが、一方で日本語はまだ少ない。内容が豊富で、かつソースコードまで公開されている例として以下が見つかった。

- 『R で計量政治学入門<sup>\*3</sup>』
- 『AI レベルの倫理学<sup>\*4</sup>』

bookdown の機能や見た目を確認することができる。さらに以下 2 つは私が作成したものである。

<sup>\*1</sup> <https://github.com/hadley/r4ds>

<sup>\*2</sup> ソース: <https://github.com/handsondataviz/book>

<sup>\*3</sup> ソース: [https://github.com/shohei-doi/quant\\_polisci](https://github.com/shohei-doi/quant_polisci)

<sup>\*4</sup> ソース: <https://github.com/MToyokura/Ethics-for-A-Level-Japanese>

- ・『三國志で学ぶデータ分析 (Japan.R 2019)』<sup>\*5</sup> (Japan.R 2019 の資料)
- ・『経済学と反事実分析接触篇 Economics and Counterfactual Analysis: A Contact』<sup>\*6</sup>  
(Tokyo.R 第 83 回の資料)

特に私の 2 作品は PDF のレイアウトにも注意を払っているが、当時は kazutan 氏作の `bookdown_ja_template` をさらに改良した kenjimyzk 氏のテンプレートを元にワンオフで作成したフォーマットを使用し、`rmdja` を使用していないため、あまりスマートでない書き方が見られる。

また、HTML 形式の文書には PDF など他のファイル形式のダウンロードリンクを設置することができる。これは `_bookdown.yml` で表示を指定できる。

## 16.2 TODO: 入稿するには

国内の印刷所で PDF 入稿する際のスタンダードは何だろうか？紙媒体でやったことがないで全くわからない。ver. 0.3 時点での対応を紹介する。

### 16.2.1 トンボの表示

`_output.yml` で

```
rmdja::pdf_book_ja:
tombow:true
```

とすると PDF にトンボ (trimming mark) を表示する。これは `gentombow.sty` によるものである。しかし私はこの出力が適切なのか判断することができない。

### 16.2.2 フォントの埋め込み

少なくとも PDF ではフォントを埋め込みそこなったり、Type 3 フォントが設定されないようにしている。ただし Python 等を利用して描いた PDF は個別に設定が必要な場合があり、保証できない。

TODO: <https://teastat.blogspot.com/2019/01/bookdown.html> の記述のうち、まだ対応していないものがある。

---

<sup>\*5</sup> ソース: <https://github.com/Gedevan-Aleksizde/Japan.R2019>

<sup>\*6</sup> ソース: [https://github.com/Gedevan-Aleksizde/20190125\\_tokyor](https://github.com/Gedevan-Aleksizde/20190125_tokyor)

## 第 VI 部

### デバッグ

rmdja (2020-10-24 04:38)

## このパートで説明すること

残念ながら, 現状 bookdown は完全にプログラミング知識のないエンドユーザでも縦横無尽に使用できるかと言うと, まだまだ不安定でそのレベルには達していない. さらに悪いことに, rmarkdown および bookdown は knitr, pandoc, LaTeX といった様々なプログラムを継ぎ接ぎして実装されているため, R の知識だけではエラーが起こった場合や, 意図したとおりの出力が得られないときに原因が分かりにくいことがある. そこで, ここではエラーが出た際にどう対処するかのヒントを書いておく.



## 第 17 章

# 製本時のエラーへの対処

### 17.1 エラーがどのタイミングで発生したかを特定する

R Markdown はさまざまな外部プログラムを利用して、数段階のプロセスを経てソースファイルを変換して文書を作成する複雑なプログラムである。逆に言えば、Rmd ファイルを md ファイルに変換 (knitr による処理) するときにエラーが出たのか、md を各ファイルに変換 pandoc する際に起こったのかをまず特定するのが重要である。そのためには

1. `keep_md: true` を設定する
2. うまくいかないときはキャッシュを削除してから再実行する

という対処法がある。(1) は文字通り中間出力ファイルである .md を残すことを意味する。これが生成されないなら knitr でのエラーだと分かるし、中身を見て不自然な内容になっているのなら Rmd の書き方が knitr に正しく評価されていないことがわかる。

キャッシュも私の経験上よくエラーの原因となっている。以前に実行していたチャンクの結果が更新されていないせいで、knitr の処理の不整合を起こすことがある。`*_files` には出力に必要な画像ファイルが、`*_cache` にはチャンク実行結果のキャッシュが残っている。後者は `knitr::opts_chunk$set(cache = T)` などでキャッシュを残す設定にできるので、F に設定した上でこれらのファイルを削除する。

処理に時間がかかるチャンクがあってキャッシュを作りたい場合は、別途 rds や RData ファイルに結果を保存するという方法もある。しかもしもプログラムの再現性を重視する場合、この方法は望ましくないだろう。しかし残念ながら現状はこうするか、ひたすら長い時間を持つしかない。

TODO: <https://bookdown.org/yihui/rmarkdown-cookbook/cache.html>

### 17.2 YAML フロントマターを確認する

以前『[R] R Markdown の YAML ヘッダでハマったおまえのための記事』というブログ記事にも書いたように、YAML フロントマターは慣れない書き間違えやすいのが現状である。もし自分で変更したのなら、改めて確認すべきだろう。特に、製本直後にすぐに、心当た

りのない R プログラム関係のエラーが出る場合、**チャンクではなく YAML フロントマター** の読み取りに失敗している可能性がある。

以下の**4原則**を覚えておこう。以前は bookdown の話を想定してなかったので、さらに条文を1つ加えた。

1. `output:` 以下はフォーマット関数への引数
2. トップレベルのオプションは pandoc のオプション
3. タイプミスや位置間違えでも動いたり、動かなかったりする
4. `_output.yml` および `_bookdown.yml` を見る。

`output:` には bookdown::gitbook など、bookdown で提供されているフォーマット関数を指定しており、その配下に記入するのはフォーマット関数に与える引数である。よって、関数ヘルプを確認すれば有効な引数を知ることができる。しかし一方で、... が引数になっていることがあるので、タイプミスしてもエラーが出ないことがある。

また、YAML の構文でサポートされている配列は誤評価を引き起こすことがある。

```
output:
 bookdown::gitbook:
 toc_depth: 3
 toc: true
```

```
output:
 bookdown::gitbook:
 - toc_depth: 3
 - toc: true
```

上の例は正しい記法である。一方でハイフン - は YAML では配列を記述するために用意されている。下記の場合、キーワード引数ではなく位置引数のような扱いになるため、`toc` に対して 3 を代入することになり、エラーが発生する。逆に言えば、- を使う場合、キーワードを書かずに値だけを正しい順番で書けば機能する。

インデントしないトップレベルの引数は、基本的に pandoc に与える引数である。これ意味のない引数を与えてエラーを返さないことが多いので、タイプミスに注意する。

しかし、フォーマット関数に `pandoc_args` という構文をサポートしていることや、フォーマット関数で `pandoc` の同名の引数を上書きする仕様のフォーマットもあるため、上記は絶対ではない。これが原因で、「`output:` 以下に書くべきものを間違えてトップレベルに書いたが、意図したとおりに機能した」あるいはその逆が発生することがある。また、`pandoc` の構文ではキーワードにハイフンを使うことができるが、フォーマットは R の関数でもあるため、ハイフンを使はず、アンダースコアで置き換えられる。この違いも書き間違えの原因になる。

### 17.3 PDF 生成時のエラーを確認する

それでもエラーが出る場合、私の経験上ほとんどが生成した .tex ファイルをタイプセットする際にエラーが発生している。.html との両立を考えると、どうしても pandoc が解釈できる構文に限界があるためである。

! LaTeX Error: XXXXX

とか

Error: LaTeX failed to XXXX

といったメッセージが表示されるのをすぐ分かる。さらに丁寧なことに、tinytex のデバッグ方法への [リンク](#) まで表示される

この場合最も重要なのは、以下に尽きる。

1. options(tinytex.verbose = TRUE) を設定する
2. keep\_tex: true を設定する

これは keep\_md と同様に、中間ファイルである .tex を残すことを意味する。

それでも解決しない場合、改めてこのファイルを手動でタイプセットするのも 1 つの方法だ。もしうまくいったり、異なるエラーが出るのなら、環境の違いが問題かもしれない。そして upBibTeX を使うのなら、後者が唯一のデバッグ方法だ。

ここでは rmdja の内部処理を解説する。knitr や rmarkdown の仕様に精通している、自分で細かい設定をしたいユーザ向けの解説である。



## 付録 A

### デフォルト値の自動調整

R Markdown で日本語文書を作成する上での大きな障害の 1 つである、 YAML フロントマターの設定を改善している。 rmdja の文書フォーマットは YAML フロントマターのデフォルト値などを日本語文書に適したものに変更している。さらに、ユーザーを OS ごとのフォントの違いや煩雑で重複だらけの設定から解放するため、内部処理でも動的に設定変更している。もちろんこれらはユーザーによる YAML フロントマターやチャンクオプションの変更で上書きできる。

#### A.1 デフォルトのフォント

PDF 出力時のデフォルトフォントは、生成時に OS を判定して設定している。その設定は表 A.1 のようなルールである。

これらは XeLaTeX ならば zxjafont、LuaLaTeX ならば luatex-jp で用意されているプリセットを使って設定している。使用 OS の判定は R の基本関数による。なお、Noto フォントを選んだのは Ubuntu 18 以降の日本語用フォントだからである。Ubuntu から派生した OS にはプリインストールされていることが多いようだが、Debian, Cent OS, Fedora 等にはおそらくプリインストールされていないので注意。現時点ではフォントが実際にインストールされているかを確認する機能はない。

フォントのプリセットを指定した場合、個別設定は無効になる。さらに、3 種類の和文フォントを全て設定していない場合もデフォルトのプリセットから選ばれる。

OS/エンジン別のデフォルトフォント

engine	Linux	Mac	Windows (>= 8)	Windows (それ以前)
XeLaTeX	Noto	游書体	游書体	MS フォント
LuaLaTeX	Noto	ヒラギノ	游書体	MS フォント

## A.2 チャンクのデフォルト設定

デフォルトのグラフィックデバイスは、HTML では PNG、PDF では cairo\_pdf としている。R でよく描画するような単純な図形はベクタ画像が適しているが、件数のとても多いデータの散布図などはベクタ画像にするとファイルサイズが大きくなるため、そのような画像を適度に「劣化」させてファイルサイズを軽減してくれる cairo\_pdf を標準としている。HTML に関しては、そもそもデフォルトの設定で PDF が表示できない Web ブラウザが多いことから、PNG をデフォルトにした。

また、block, block2, asis などのブロックを echo=F や include=F にするメリットはほぼないため、knitr::opts\_chunk\$set(echo = F, include = F) と一括設定してもこれらは echo=T, include=T のままである。変更したい場合は、チャンクごとに設定することで有効になる。

## 付録 B

# PDF の組版に関する細かい話

ここでは pandoc テンプレート等の設定を解説する。

3種類の和文フォントを個別設定をした場合, X<sub>E</sub>TEX はフォールバックフォントを有効にしている。j\*\*\*\*fontoptions 以下に, FallBack=... というオプションでフォールバックフォントを指定すれば有効になる。

用紙サイズは, デフォルトは a4paper, B5 がよいなら b5paper オプションを classoptions: に指定する。

PDF を印刷所に持ち込んだことがないため詳しいことはわからないが, 『Bookdown による技術系同人誌執筆』で指摘されているようなトンボやノンブルは出力されるように作っている(そしてここで紹介されているような LaTeX のコマンドの多くは rmdja では書く必要がなくなった)。

TODO: PART の扉ページにはまだノンブルが表示されない

### B.1 画像の配置

現在, PDF で画像の配置を固定する方法について何も特別なものを用意していない。単純に自分は必要だとおもったことがないため。固定したい場合は R Markdown や Bookdown のドキュメントを参考にしてほしい。ただし、通常は章や部をまたいで表示されることはない(はず)。

### B.2 取り消し線

LaTeX の各パッケージのバージョンによっては、和文に取り消し線(\sout)を与えるとタイプセット時にエラーが出ることがある。もともと ulem.sty は欧文を前提にしたものなので適当に妥協してほしい。

### B.3 TODO: しかし英文で書きたい場合

rmdja の機能を使いたいが、執筆は英語でしたいと言う場合は最低限以下のような設定変更が必要である。デフォルトは日本語用文書クラスのため、

`documentclass: book / report / article`

など欧文用文書クラスを指定する。

rmdja では和文フォントを参照するので、和文フォントの設定も手動で解除する必要がある。

を指定する。そして各種見出しも英文用に調整する。

TODO

## 付録 C

### jecon.bst の紹介

和文と欧文を使い分けたスタイルファイルとして, jecon.bst がある. jecon.bst の公式ではなく, 私がカスタマイズしたバージョンでも良い. こちらは本来よりも電子媒体としての利用を重視して,

- 参照 URL を表示せず, ハイパーリンクのみにする
- ArXiv ID の表示とハイパーリンク追加

といった変更をしている. 後者は, BibTeX エントリに以下のように `archivePrefix` に `arxiv` と言う値が入っていると, `eprint` の値が ArXiv ID として表示される. これは ArXiv から直接 .bib ファイルを取得したり, Zotero などでインポートすれば必ず入力される項目である.

```
archivePrefix = {arXiv},
eprint = {XXXX.YYYYYY},
...
```

TODO: 現在 jecon.bst の表示も少しおかしいので確認中.



## 付録 D

### fontregisterer でグラフ用フォント を自動登録

Rで描いたグラフに日本語を表示する場合, Linux 系 OS ならばフォント名を適切に設定するだけで表示されるが<sup>3</sup>, Windows や Mac ではフォントをグラフィックデバイスに登録する必要がある. しかし手動登録は面倒なので, インストールされているシステムフォントを全て自動登録するパッケージ, `fontregisterer` を用意している.

```
01 remotes::install_github("Gedevan-Aleksizde/fontregisterer", repos = NULL, type = "source")
```

もちろんこれは R Markdown 以外でも使用できる. このパッケージは読み込まれた時点で登録処理を行うため,

```
01 require(fontregisterer)
```

を最初に実行するだけで良い. 詳しい仕組みの解説は『[おまえはもう R のグラフの日本語表示に悩まない \(各 OS 対応\)](#)』に書いている.



## 参考文献

- Dougherty, Jack and Ilya Ilyankou (forthcoming). "Hands-On Data Visualization Interactive Storytelling from Spreadsheets to Code". URL: <https://handsondataviz.org/>.
- Wickham, Hadley and Garrett Grolemund (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. First edition. Sebastopol, CA: O'Reilly. ISBN: 978-1-4919-1039-9 978-1-4919-1036-8. URL: <https://r4ds.had.co.nz/>.
- Xie, Yihui (2020). *Bookdown: Authoring Books and Technical Documents with r Markdown*. Chapman & Hall. URL: <https://bookdown.org/yihui/bookdown/>.
- Xie, Yihui, JJ. Allaire, and Garrett Grolemund (2018). *R Markdown: The Definitive Guide*. Boca Raton, Florida: Chapman and Hall/CRC. URL: <https://bookdown.org/yihui/rmarkdown/>.
- Xie, Yihui, Christophe Dervieux, and Emily Riederer (2020). *R Markdown Cookbook*. English. S.l.: CRC PRESS. ISBN: 978-0-367-56383-7. URL: <https://bookdown.org/yihui/rmarkdown-cookbook/>.