在 R Markdown 文档中使用中文

谢益辉 邱怡轩 于淼

目录

1	引言	1
2	字体和选项	2
3	R 代码段	2
4	源代码控	4
5	小结	4

1 引言

中文 LaTeX 文档并非难题。当然这句话得站在巨人 CTeX 的肩膀上才能说,它让我们只需要一句

\documentclass{ctexart} % 或者 ctexrep/ctexbook

或者

\usepackage{ctex}

就轻松搞定中文 LaTeX 排版问题。跨平台通用的 LaTeX 编译却是个小难题,主要是没有一种跨平台通用且免费的中文字体。好吧,你们 Windows 用户永远有宋体黑体,你们 Mac 用户有华文字体,而我们苦逼 Linux 用户在编译 LaTeX 文档就没那么简单了¹,是啊,我们有文泉驿,但我们要是用了文泉驿之后把文档发给你们八成不能编译,因为你们没有安装文泉驿。

¹切,傲娇的 Linux 用户怎么会干出找你们复制字体的事情

2

2 字体和选项

LaTeX 包ctex支持若干种字体选项,如果你是ctex 老用户,请注意这里我们要求的最低版本是2.2,你可能需要升级你的LaTeX 包。从版本2.0 开始,ctex 支持根据不同操作系统自动选择中文字体,简直是为人类进步作出了巨大贡献,我们再也不必费尽口舌向用户解释"啊,你用Windows啊,那么你该使用什么字体;啊,你用Mac啊,又该如何如何"。

下面的 YAML 元数据应该能满足多数用户的需求, 主要设置两项参数: 文档类为 ctexart (当然也可以是别的类),输出格式为 rticles::ctex, 其默认 LaTeX 引擎为 XeLaTeX (真的, 别纠结你的旧爱 PDFLaTeX 了)。

documentclass: ctexart
output: rticles::ctex

rticles::ctex 的参数都是普通的 pdf_document 参数,参见文档rmarkdown 包的文档,这里就不赘述了。

Windows 和 Mac 用户应该都已经有自带的中文字体了。Linux 用户可以另外下载 Fandol 字体²,它号称是免费的,不过我们也没太搞清楚它的来头。把那些.otf 字体文件下载下来保存到 ~/.fonts 文件夹下,然后命令行切换到那里,运行 fc-cache -fv。

3 R 代码段

R 代码用 R Markdown 的语法嵌入,即三个反引号开始一段代码 ```{r}和三个反引号 ```结束一段代码:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -17.579 6.7584 -2.601 1.232e-02

speed 3.932 0.4155 9.464 1.490e-12

上面回归方程中的斜率是 3.9324, 完整的回归方程为:

Y = -17.5791 + 3.9324x

画图当然也是木有问题的啦,想画就说嘛,不说我怎么知道你想画呢?

 $^{^2}$ http://ctan.org/tex-archive/fonts/fandol

3 R 代码段 3

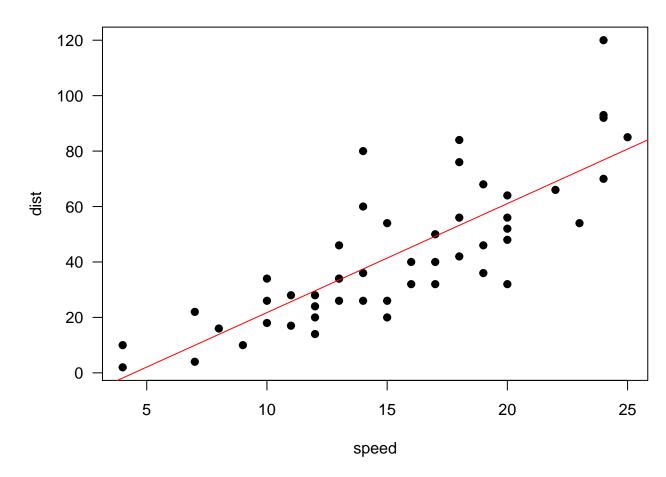


图 1: cars 数据散点图以及回归直线。

4 源代码控 4

请不要问我为什么图浮动到下一页去了,这么初级的 LaTeX 问题问出来信不信我扁你。

4 源代码控

这里提供的 rticles 模板可能由于种种原因不能满足客官的要求, LaTeX 用户就是这样永无止境地调格式(唉,跟 Word 用户到底有啥区别呢)。若真是需要调整,你可以复制一份默认模板去改,如前面所说,本文档的模板是 rticles::ctex_template(),它是一个文本文件。若熟悉 LaTeX 的话一看就明白,只不过里面有些 Pandoc 变量而已;若不熟悉 LaTeX 我们在这里说了也白说,花几天时间好好啃一啃 LaTeX 人门手册吧。

本文档所用的模板是从 Pandoc 默认 LaTeX 模板基础上做了少许改动而来的: https://github.com/yihui/pandoc-templates/blob/ctex/default.latex 具体改动从 GIT 提交消息日志中可以看到, 主要就是去掉了 mathspec 包和 hyperref 包。

5 小结

事实证明我们可以理直气壮地通过 XeLaTeX 将中文 R Markdown 转化为 PDF 文档,麻麻再也不用担心我的论文满屏幕都是反斜杠,朕养完小白鼠之后终于不必先折腾三个小时 LaTeX 再开始写实验报告了:打开 RStudio,菜单 File > New File > R Markdown,然后从模板中选择 CTeX Documents,搞定。