

Notes

Devin Judge-Lord

January 23, 2019

Contents

R Markdown	1
Including Plots	2
Writing Math	5
Matrices	5
Example Notes	7

R Markdown

Digital notes are easy to search, read, and share. Markdown is just **plain text** that can be rendered as HTML, PDF, or even Microsoft Word. It allows us to use various syntax, like LaTeX and **R** as needed, either in chunks or inline with the rest of our text. For example, typing “`$y = \alpha + X\beta$`” gives us “ $y = \alpha + X\beta$ ” and typing “`{r nchar("hello world") }`” gives us “11.” See <http://rmarkdown.rstudio.com> and this [cheatsheet](#)

Rstudio’s “Knit” button will call on your LaTeX compiler to make a PDF, but you can also choose HTML and Word from the dropdown menu next to the Knit button. The header of this .Rmd file, called YAML, specifies options for each type of output. I prefer rendering notes in HTML so that I can share them as a web page, but since LaTeX writes to PDF, some LaTeX syntax only works when rendering a pdf.

We embed a chunk of R code like this:

```
data("pressure") # a dataset built into R
head(pressure)
```

```
##    temperature pressure
## 1             0    0.0002
## 2            20    0.0012
## 3            40    0.0060
## 4            60    0.0300
## 5            80    0.0900
## 6           100    0.2700
```

In the .Rmd file, you’ll see that the above chunk of R code begins with “`“`”, then tells the computer that it is `r` code, and then names this chunk “pressure.” Give code chunks informative names! If your chunk produces figures, they will be saved with this name.

If we don’t want our code chunk to be visible, only want the results, we can add `echo = FALSE` to the header:

```
##    temperature pressure
## 1             0    0.0002
## 2            20    0.0012
## 3            40    0.0060
```

```
## 4          60    0.0300
## 5          80    0.0900
## 6         100    0.2700
```

We can set defaults for the whole document (e.g. `knitr::opts_chunk$set(echo = TRUE)`). See the above “setup” chunk in this .Rmd file.

Problem sets and papers will use `echo = FALSE` to display just the output. For notes, we may want to see our code `echo = TRUE` or (if our output is HTML), hide our code.

To display code we don’t want to run, we add `eval = FALSE`.

```
This code will not run.
```

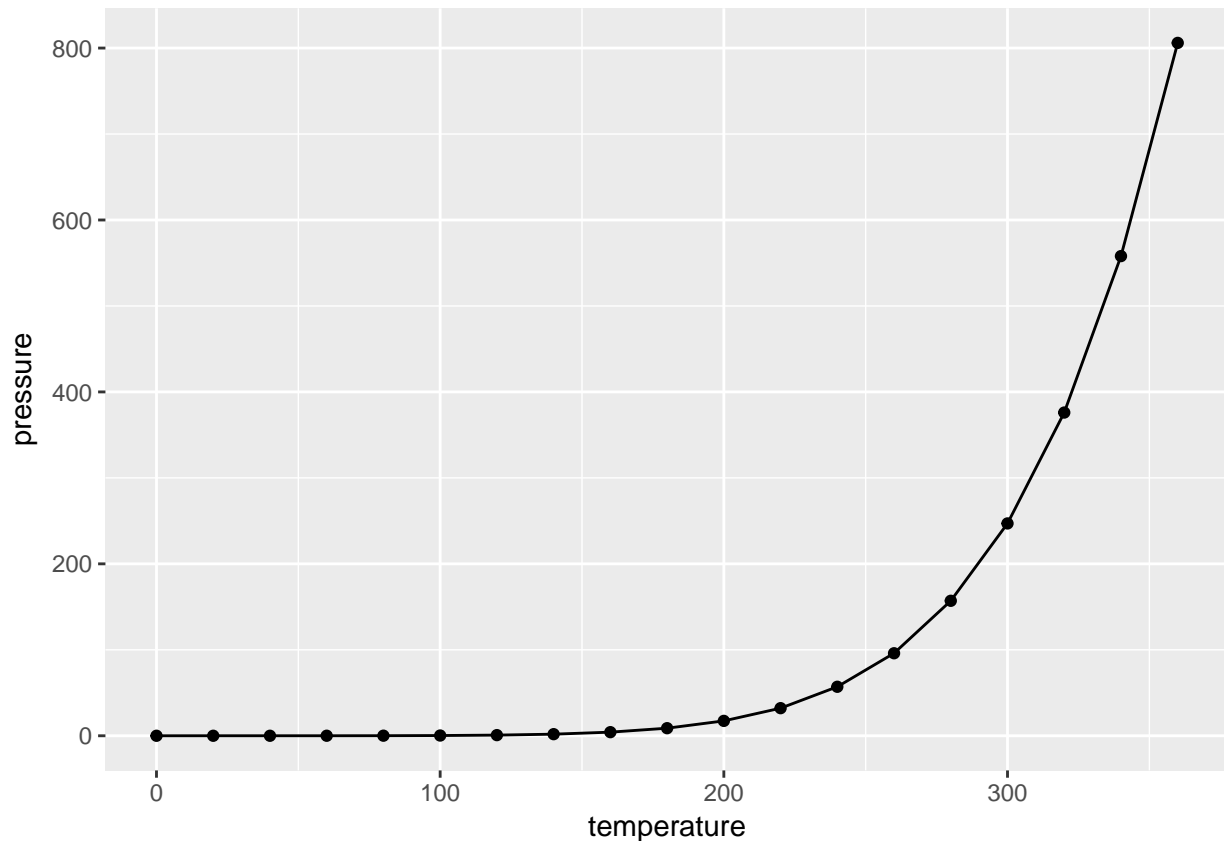
For example, if you want to share code that does not work on your 811 page, use `eval = FALSE`.

It is safer to use `cache = FALSE`, which re-computes all results each time and will thus update if your data have changed. If computing is taking a long time, you can use `cache = TRUE` to knit faster, but this will only run chunks where the *code* has changed.

Including Plots

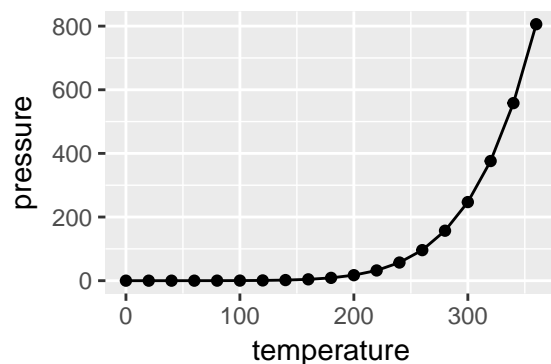
We can also embed plots. For example:

```
library(ggplot2)
## aes() defines aesthetics (e.g. the location, color, and shape of layers on the grid)
## each geom adds a layer
ggplot(data = pressure, aes(x = temperature, y = pressure)) +
  geom_point() +
  geom_line()
```



Note that we needed to load the library `ggplot2` because knitting uses a new **R** session. Your `.Rmd` file must load any required libraries, functions, and data. The upshot is that knitting is fairly reliable and self-contained. It starts fresh, not depending on whatever you may have been doing in **R** previously.

We can adjust the size by adding `fig.width = 3`, `fig.height = 2` to get a 4×3 inch figure.



Let us also center `fig.align='center'` and add a caption, `fig.cap = "Plotting Temperature vs. Pressure"`:

We can link to figures in the text of the document by adding `\\label{fig:pressure}` to the caption. Typing `\\ref{fig:pressure}` will now always correctly reference figure 2 regardless of how many figures are added above it.

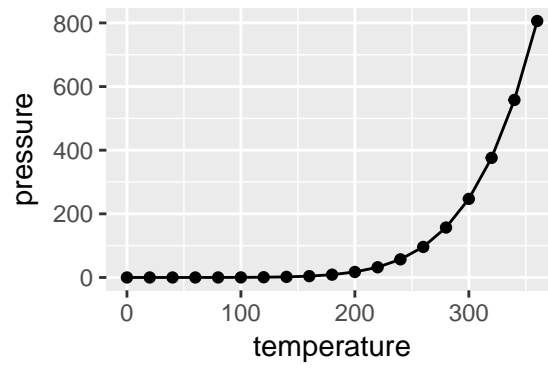


Figure 1: Plotting Temperature vs. Pressure

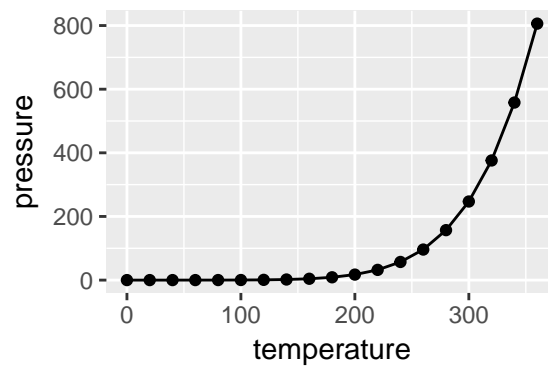


Figure 2: Plotting Temperature vs. Pressure with a label

Writing Math

We can use LaTeX to write nice-looking math. For example, we can write out a regression equation as $y = \alpha + X\beta + \varepsilon$, without having to copy, paste, or insert special symbols. To write out anything math related, we enclose it in dollar signs $\$math\$$. The regression equation above, for example, is $\$y = \alpha + X\beta + \varepsilon\$$. If we want to index something by adding subscripts x_1, x_2 etc, the code is $\$x_{\{1\}}, x_{\{2\}}\$$. Fractions are written such that $\frac{1}{2}$ is $\frac{1}{2}$, and exponents (2^2) are written $2^{\{2\}}$.

The Greek letters have a slash before them which tells Latex to print that as a Greek letter. Similarly 2×2 is written $\$2 \times 2\$$. A useful guide to special characters and formatting in Latex is here <https://users.dickinson.edu/~richesod/latex/latexcheatsheet.pdf>. I regularly google latex symbols.

Matrices

To write a matrix:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix}$$

we write:

```
$$
\left[
\begin{array}{ccc}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 0
\end{array}
\right]
$$
```

where:

```
$$ % start Latex mode
\left[ %creates the left bracket, the "\left" command scales the bracket "["
\begin{array}{ccc} % Creates the "array" (matrix), {ccc} defines the number of columns
1 & 2 & 3 \\ % "&" divides the columns, "\\" creates a new line
4 & 5 & 6 \\
7 & 8 & 0
\end{array} % ends the matrix
\right] %creates the left bracket, the "\right" command scales the bracket "]"
$$ % ends Latex mode
```

Matrix multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1+2+3 \\ 4+5+6 \\ 7+8+0 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \\ 15 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix} = \begin{bmatrix} 1+4+7 & 2+5+8 & 3+6+0 \end{bmatrix} = \begin{bmatrix} 12 & 15 & 9 \end{bmatrix}$$

Example Notes

Linear regression with fixed effects:

$$y_i = \alpha_0 + \alpha_j + x_i\beta + \varepsilon$$

Where α_0 is the intercept,

α_j are fixed effects (effects that do not vary by x , i.e. intercept shifts) for each group j ,

x and y are vectors of observations, and

ε is the error.

OLS estimation

Data:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

OLS equation

$$\hat{\beta} = (X'X)^{-1}(X'y)$$

Logistic link function:

$$Pr(y_i = 1) = \frac{1}{1 + e^{-x_i\beta}}$$

Poisson PMF:

$$\frac{\lambda^k}{k!} e^{-\lambda}$$

Gaussian (normal) PDF:

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$