A Math Reference: Sample Modeling Equations to Borrow

Miratrix

2019-09-20

Introduction

This document has a bunch of mathematical equations we use in the class. It is a good reference for how to write your own math equations in your life moving forward. Generally, people write math equations using something called Latex. Latex (or Tex) is a way of writing documents where mixed in with the writing of what you want to say are commands describing how you want your document to look. This is a very powerful thing: there are Tex editors that allow you to write entire articles, books, reports, poetry, or whatever with extreme control over the typesetting used. It creates beautifully typeset documents that are easy to distinguish from those written in, say, MS Word due to how they adjust whitespace on the page. That being said, it can be a lot to jump in to.

Enter R Markdown. R Markdown is a useful and easy way to take advantage of this syntax without the overhead of writing entire documents in Latex, even if you don't have any R code in your document. Inside R Markdown you can write math equations, and then when you render the report it not only runs all the R code, it formats all the math for you as well! You can even have R Markdown render to MS Word to give you a word doc with all your math equations ready to go.

Using this document

You are reading, probably, the pdf version of this document. But really you should open the Rmd file and cut and paste the relevant equations into your own work, and then modify as necessary.

Overview of Using Latex

For math in your writing, you denote the beginning and the end of a math equation in your text using “$”s, one at the start and one at the stop. E.g., “$ math stuff $”. Most greek letters are written as their names with a backslash “\” just before it. E.g., “\alpha”.

So if I want to write an alpha, I write “`$\alpha$`” and get α.

I can do subscripts by using an underscore. E.g., “`$\alpha_j$`” gives α_j.

Some useful greek letters

Here are some useful greek letters and symbols

Letter	Name
α	\alpha
β	\beta
δ, Δ	\delta, \Delta
ϵ	\epsilon
σ, Σ	\sigma, \Sigma

Letter	Name
ρ	\rho
μ	\mu (Meew!)
τ	\tau
$\times$	\times
$\sim$	\sim

See many more symbols at, e.g., https://www.caam.rice.edu/~heinken/latex/symbols.pdf. This was found by searching "tex symbols" on Google.

Equations on lines by themselves

To write an equation on a line by itself, surround the math stuff with the double "$$" and "$$". E.g., if we write:

```
$$ Y = a X + b $$
```

We get

$$Y = aX + b$$

If we want multiple lines, we use a double backslash ("\\") to denote a line break (even if we have a line break we have to do this—we have to explicitly tell the program converting our raw text to nice formatted text where the line breaks are). E.g., if we write

```
$$
Y = 10 X + 2 \\
Y - 5 = 3 X^2 + 5
$$
```

we get:

$$Y = 10X + 2Y - 5 = 3X^2 + 5$$

Notice how these equations are not quite lined up nicely. It automatically centers each line. We can make it nicer using the `\begin{aligned}` and `\end{aligned}` commands to get our equations to line up. We then, inside the begin-end block of math, line things up by the `&` symbols on each row of our equation:

```
$$
\begin{aligned}
Y &= 10 X + 2 \\
Y - 5 &= 3 X^2 + 5
\end{aligned}
$$
```

and

$$\begin{aligned}
Y &= 10X + 2 \\
Y - 5 &= 3X^2 + 5
\end{aligned}$$

Also consider:

```
$$
\begin{aligned}
a + b + c + d &= c \\
 d &= e + f + g + h
\end{aligned}
$$
```

giving

$$\begin{aligned} a + b + c + d = c \\ d = e + f + g + h \end{aligned}$$

Normal text in equations

If you put words in your equations, they get all italliced and weird:

$$5 + mydog = 10$$

You can fix using the "\mbox{}" command as so:

$$5 + \mbox{my dog} = 10$$

Random Intercept Model

Our canonical Random Intercept model is as follows. First, our Level 1 model:

$$\begin{aligned} &y_{ij} = \alpha_j + \beta_1 ses_{ij} + \epsilon_{ij} \\ &\epsilon_{ij} \sim N(0, \sigma_y^2) \end{aligned}$$

Our Level 2 model:

$$\begin{aligned} &\alpha_j = \gamma_0 + \gamma_1 sector_j + u_j \\ &u_j \sim N(0, \sigma_\alpha^2) \end{aligned}$$

The Gelman and Hill bracket notation looks like this:

$$\begin{aligned} &y_i = \alpha_{j[i]} + \beta_1 ses_i + \epsilon_i \\ &\epsilon_i \sim N(0, \sigma_y^2) \\ &\alpha_j = \gamma_0 + \gamma_1 sector_j + u_j \\ &u_j \sim N(0, \sigma_\alpha^2) \end{aligned}$$

The reduced form would look like this:

$$y_i = \gamma_0 + \gamma_1 sector_{j[i]} + \beta_1 ses_i + u_{j[i]} + \epsilon_i$$

with

$$\epsilon_i \sim N(0, \sigma_y^2), \mbox{ and } u_j \sim N(0, \sigma_\alpha^2)$$

If we want to be really prissy, we can write down the i.i.d. aspect of our random effects like this

$$\epsilon_i \stackrel{i.i.d}{\sim} N(0, \sigma_y^2), \mbox{ and } u_j \stackrel{i.i.d}{\sim} N(0, \sigma_\alpha^2)$$

The `\stackrel{}{}` command takes two bits of latex, each in the curly braces, and stacks them on top of each other.

Random Slope Model

The canonical random slope model for HS&B with ses at level 1 and sector at level 2

Level 1 models:

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j} ses_{ij} + \epsilon_{ij} \\
\epsilon_{ij} &\sim N(0, \sigma_y^2)
\end{aligned}
$$

Level 2 models:

$$
\begin{aligned}
\beta_{0j} &= \gamma_{00} + \gamma_{01} sector_j + u_{0j} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} sector_j + u_{1j}
\end{aligned}
$$

with

$$
\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \tau_{00} & \tau_{01} \\ & \tau_{11} \end{pmatrix} \right]
$$

The derivation of the reduced form is:

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \beta_{1j} ses_{ij} + \epsilon_{ij} \\
&= (\gamma_{00} + \gamma_{01} sector_j + u_{0j}) + (\gamma_{10} + \gamma_{11} sector_j + u_{1j}) ses_{ij} + \epsilon_{ij} \\
&= \gamma_{00} + \gamma_{01} sector_j + u_{0j} + \gamma_{10} ses_{ij} + \gamma_{11} sector_j ses_{ij} + u_{1j} ses_{ij} + \epsilon_{ij} \\
&= \gamma_{00} + \gamma_{01} sector_j + \gamma_{10} ses_{ij} + \gamma_{11} sector_j ses_{ij} + (u_{0j} + u_{1j} ses_{ij} + \epsilon_{ij})
\end{aligned}
$$

Commentary: There are various and competing ways of writing the covariance matrix for the random effects. The τ_{**} notation is easy and expands to any size matrix (if we, for example, have more than one random slope). But all the τ_{**} are variances, not standard deviations, and we often like to talk about random effect variation in terms of standard deviation. We can thus use something like this, instead:

$$
\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \sigma_0^2 & \rho \sigma_0 \sigma_1 \\ & \sigma_1^2 \end{pmatrix} \right]
$$

Here we have a correlation of random effects, ρ, instead of a covariance τ_{01}. And we can talk about the standard deviation of, e.g., the random intercepts, as σ_0 rather than $\sqrt{\tau_{00}}$. Different ways of writing the same mathematical thing are called different parameterizations; they are equivilent, but are more or less clear for different contexts. Unfortunately, this means there is no one right answer for how to write down a mathematical equation!

Summations and fancy stuff

Fractions are as follows:

$$
cor(A, B) = \frac{cov(A, B)}{\sigma_A \sigma_B}
$$

For reference, you can do summations and whatnot as follows:

$$
Var(Y_i) = \frac{1}{n-1} \sum_{i=1}^{n} \left(Y_i - \bar{Y} \right)^2
$$

And if you have fractions you can have big brackets with “\left(” and “\right)” as follows:

$$
X = \left(\frac{1}{2} + y \right)
$$