Frontiers of Statistics: Contraction theory for posterior distributions **Spring 2019**

Lecture 1: February 19

Lecturer: Anirban Bhattacharya & Debdeep Pati *Scribes: scribe-names*

Note: *LaTeX template courtesy of UC Berkeley EECS dept & CMU's convex optimization course taught by Ryan Tibshirani.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

This lecture's notes illustrate some uses of various LaTeX macros. Take a look at this and imitate.

1.1 Some theorems and stuff

We now delve right into the proof.

Lemma 1.1. *This is the first lemma of the lecture.*

Proof. The proof is by induction on . . ..

For fun, we throw in a figure.

Figure 1.1: Example Figure

This is the end of the proof, which is marked with a little box. □

1.1.1 A few items of note

Here is an exercise:

Exercise: Show that P $\neq$ NP.

Theorem 1.2. *This is the first theorem.*

Proof. This is the proof of the first theorem.

Enter proof here.

This concludes the proof. □

1.2 Next topic

Here is a citation, just for fun [CW87].

1.3 Typesetting your math

Here are some additional LaTeX tips about typesetting your math.

1. **Basic math.** Use i for square root of -1. log, exp, sin, cos etc everywhere in math mode, not log, exp, sin, cos. Use $x \mid y$ or $x \,|\, y$ for conditioning, not $x|y$. Always $i = 1, \ldots, n$, not $i = 1, \cdots, n$ or $i = 1 : n$. Cauchy–Schwarz inequality, not Cauchy-Schwarz inequality.

2. **Matrix stuff:** If a symbol isn't defined by default, you can put the text inside mbox or create a macro of your own. For example, $\mathrm{tr}(A)$ or $\det(A)$ rather than $tr(A)$ or $det(A)$. You may alternatively use $|A|$ for the matrix determinant. We have a macro defined for matrix transpose, $A^{\mathrm{T}}A$ looks neater than A^TA. You can alternatively use $A'A$; make sure stick to one choice throughout. Use $\|\cdot\|$ to denote a norm, not $||\cdot||$. Use `mathrm` to distinguish special matrices and have a macro created for it; for example, I_d denotes the $d \times d$ identity matrix, while $\mathrm{J}_d = 11^{\mathrm{T}}$ denotes a $d \times d$ matrix of ones.

3. **Fractions.** Never use fractions inline, only use them in an equation environment. For example, never write $\frac{a}{bc}$ inline, write $a/(bc)$ instead. *Note the bracket, it is confusing to write a/bc and should be avoided, it may be easily interpreted as $(a/b) \times c$.*

4. **Equations.** You may use the `align` environment to write larger equations, and use the double dollar for the ones that fit in one line and doesn't need to be numbered. For example,

$$\widehat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y,$$

and

$$\begin{aligned}
&\int h(\sigma^2, \xi)\, \pi_\xi(\xi)\, d(\sigma^2, \xi) \\
&= \int_\xi \left[\int_{\sigma^2} (\sigma^2)^{-(N/2+1)} e^{-z' M_\xi^{-1} z/(2\sigma^2)} d\sigma^2 \right] |M_\xi|^{-1/2} \pi_\xi(\xi) d\xi \\
&= 2^{N/2}\Gamma(N/2) \int_\xi |M_\xi|^{-1/2}\, (z' M_\xi^{-1} z)^{-N/2}\, \pi_\xi(\xi) d\xi.
\end{aligned}$$

Notice the indentation of the equality signs.

Now, suppose you wanted to number only the last equation and keep the others unnumbered. You can do it using the notag command.

$$\begin{aligned}
&\int h(\sigma^2, \xi)\, \pi_\xi(\xi)\, d(\sigma^2, \xi) \\
&= \int_\xi \left[\int_{\sigma^2} (\sigma^2)^{-(N/2+1)} e^{-z' M_\xi^{-1} z/(2\sigma^2)} d\sigma^2 \right] |M_\xi|^{-1/2} \pi_\xi(\xi) d\xi \\
&= 2^{N/2}\Gamma(N/2) \int_\xi |M_\xi|^{-1/2}\, (z' M_\xi^{-1} z)^{-N/2}\, \pi_\xi(\xi) d\xi. \qquad (1.1)
\end{aligned}$$

Sometimes, you want to assign a single number to a group of equations. That can be done as follows.

$$\begin{aligned} & z \mid \beta, \sigma^2, \xi \sim \mathcal{N}(W\beta, \sigma^2 \mathrm{I}_N), \\ & \beta \mid \sigma^2, \xi \sim \mathcal{N}(0, \xi^{-1}\sigma^2 \mathrm{I}_d), \\ & \sigma^2 \sim \pi_{\sigma^2}(\cdot), \quad \xi \sim \pi_\xi(\cdot). \end{aligned} \tag{1.2}$$

Again, notice the indentations. When you call the equation, use `eqref`. For example, "We introduce the ridge regression model in (1.2)".

5. **Spacings.** Spacings are often used in equations to improve readability. See, for example, the use of quad and slash+comma in the above equations.

6. Use [{ () }] order of brackets whenever possible. Also, make sure that what is inside of a bracket is entirely inside, and if needed, use $\Bigg(\Bigg)$ or $\big(\big)$. For example, instead of

$$b = \frac{1}{2} \log[\frac{a^2}{2}(1 + \sqrt{1 + \frac{1}{a^2}}) + \frac{1}{100}],$$

write

$$b = \frac{1}{2} \log \left\{ \frac{a^2}{2} \left(1 + \sqrt{1 + \frac{1}{a^2}} \, \right) + \frac{1}{4} \right\}.$$

Notice the extra spaces (slash+comma) in a few places to improve readability.

7. Avoid using large square roots whenever possible. For example, instead of

$$\sqrt{\frac{a}{b} \sum_{i=1}^{\infty} v_i},$$

write

$$\left(\frac{a}{b} \, \sum_{i=1}^{\infty} v_i \right)^{1/2}.$$

Again, notice the spacings introduced.

References

[CW87] D. Coppersmith and S. Winograd, "Matrix multiplication via arithmetic progressions," *Proceedings of the 19th ACM Symposium on Theory of Computing*, 1987, pp. 1–6.