Applied Probabilistic Programming & Bayesian Machine Learning

§1-3: Programming, Probably

MIT IAP 2017 Max W. Shen

1 Exercises in Writing Stan Code

Thus far, you have written some simulation code, seen examples of a hierarchical model and a regression model in Stan. If you have done all the exercises, you have played around with editing Stan code as well. Here, we present some example models for you to build from scratch.

As with software engineering, models in Stan are best developed one step at a time, bouncing between adding complexity and running the model to check for expected behavior.

Exercise 1

Implement the following model (with simulation code and Stan code).

$$\begin{aligned} y_i &\sim N(\mu, \sigma) \\ \mu &\sim weakly - informative() \\ \sigma &\sim weakly - informative() \end{aligned} \tag{1}$$

Exercise 2

Implement the following model (with simulation code and Stan code).

You have N-by-D observations that are normally distributed all around the same mean, but with N separate standard deviations. The N standard deviations are normally distributed.

2 Adding Flexibility with the For Loop

The model

```
model {
  for (n in 1:N) {
    y[n] ~ normal(mu, sigma);
  }
}
```

is equivalent to:

```
model {
  y ~ normal(mu, sigma);
}
```

Stan also supports if-else statements, variable declarations in the model block, while loops, break and continue, and more, just like a typical imperative programming language. Refer to Part III: Modeling Language Reference in the Stan reference for more details.

Exercise 3

Implement the following model (with simulation code and Stan code).

You have N-by-D observations and an N-by-L feature matrix. Your observations consist of D samples from a normal distribution with N distinct means and standard deviations. The N distinct means arise from a linear combination of your L basis vectors in your N-by-L feature matrix, and the N distinct standard deviations are normally distributed.

It may be helpful to use a for-loop, as well as either a variable declaration in the model block, or the transformed parameters block mentioned at the end of the previous assignment.

3 Estimating Height

Andrew Gelman, the author of Bayesian Data Analysis 3 and one of the leaders of the Stan development team, considers a straightforward application of Stan here:

http://andrewgelman.com/2014/09/29/general-principles-bayesian-data-analysis-arising-stan-analysis-john-lee-andersons-height/

Note that the Stan code looks different because it's for a much earlier version.

4 Additional Resources

The Stan reference is available here: http://mc-stan.org/documentation/