STA380 Project Proposal

Simulating the Bivariate Gaussian Distribution Using MCMC Methods

Anna Ly

annahuynh.ly@utoronto.ca

University of Toronto, Mississauga

The following proposal is an example of using an "original" topic that has not already been suggested in the main project document. I plan to teach the material in this example in the last few weeks of the term. Of course, since I am providing this example (with all raw codes), you are not permitted to use MCMC methods for your project unless you want to do something more advanced and outside the scope of the course. Coloured text represents additional comments.

1 Project Topic

Simulating the Bivariate Gaussian Distribution Using MCMC Methods.
This does not necessarily have to be the same name as your title; I was lazy.

2 Simulation vs. Dataset

The results will be from using pure simulations.

3 Project Details

We will be using 3 introductory MCMC methods to simulate the bivariate Gaussian distribution:

- An independence Metropolis–Hastings algorithm
- A random walk Metropolis–Hastings algorithm
- A Gibbs Sampler

The following outputs will be provided:

- Scatter plot of the samples generated
- Histogram and Q-Q plots to assess marginal normality
- The generated mean and the covariance matrix

Aside: if extra details mostly pertain to the next section, then you can just write them below instead of repeating yourself.

4 User Inputs (Shiny Components)

Users will be able to modify the following:

1. Setting the seed

2. Change the parameters, specifically, for $i \in \{1, 2\}, \mu_i, \sigma_i$ and correlation ρ.

3. Select the information presented:

   - the mean and covariance matrix,
   - histogram and Q-Q plot for marginal normality,
   - and the scatterplot of the samples generated.

4. Be able to choose what information is presented, i.e., they can look at results from just 1 MCMC method, 2 MCMC methods, or all three at once.

5. For the plots, be able to modify the colours.

References

J. Voss. *An introduction to statistical computing: a simulation-based approach.* John Wiley & Sons, 2013.

I purposely omitted other citations, because naturally, you could just copy and paste some of them. I expect all groups to also reference the following:

- R Shiny (use the command `citation("shiny")` in R once you've installed the package).

- The course textbook, or other sources you have used to learn computational statistics methods.

There is no strict requirement when it comes to the citation style. I only request that you use BibTeX citations.