Data Science Math Bridge – Sample Final

Your Data Science Bridge Final Exam is due by end of day on Sunday July 26*. You should post your work in GitHub, and include a link to your GitHub repository in your assignment submission. You are also expected to make a short (3 to 5 minute) presentation during our last Meetup on July 28th.*

Please note:

(1) If you are also taking the SQL bridge and or the R bridge, you may instead design a final project of your own choosing that incorporates what you have learned in both or all three of your bridge courses. For example, you might design and populate tables in SQL, analyze the exported data in R, then include some conditional probability in your analysis. This will require some forethought on your part. You will only be required to present once if you choose to combine finals.

For your final project, you are asked to include a short description of a use case that includes some use of probability, linear algebra, or advanced math. You might include a snippet of R code and/or a very basic demonstration of the math that provides insight about your chosen use case.

My expectations are very modest here—delivering a 10,000 foot view with imperfect understanding is acceptable at this point. Missing the 7/26 deliverable date is not acceptable. I just want you to begin thinking about and researching the connection between data science math and the implementation of data science algorithms and techniques.

Here are some examples. You are encouraged to choose your own example.

- How data science math is used in spam analysis.
- How data science math is used in image recognition.
- How data science math is used in information gain calculations that support feature selection.
- How data science math is used in recommendation engines.
- How data science math is used in Page Rank algorithms.
- How eigenvalues, eigenvectors and determinants are used in Principal Component Analysis. (advanced)
- How data science math is used in topic modeling. (advanced)