Course Syllabus for SIADS 502 WI21: Math Methods for Data Science

Course Overview and Prerequisites

This course will introduce and in some cases, review mathematical concepts relevant to future work in applied data science. It will cover important concepts in basic and linear algebra, matrix and vector algebra, eigenvectors and eigenvalues, optimization techniques, Bayes rule, and maximum likelihood.

There are no course prerequisites.

Instructor and Course Assistants

Instructor: Prof. Grant Schoenebeck

Teaching Staff: Alex McLeod, Chengyue Qiu

Communication Expectations

- Contacting instructor and course assistant: (email, slack). Most questions should be addressed on slack so that all students can see answers. For email, please use the course email address siads502staff@umich.edu which will reach both instructors and ensure a faster response time.
- Email/Slack response time: typically within 24 hours, occasionally may take 48 hours. If you have not received a response in 48 hours, please repost/resend.
- Grading response time: 5 days after assignments are due
- Office Hours: See Course Schedule below.

Required Textbook

None.

Recommended Textbook

The Data Science Design Manual by Steven X. Skiena (Springer Texts in Computer Science)

All of Statistics: A Concise Course in Statistical Inference by Larry Wasserman (Springer Texts in Statistics)

Technology Requirements (unique to this course)

None.

Learning Outcomes

1. Compute conditional probabilities, correlations, and multiplication of vectors and matrices.

1a. Revisit basic algebra and notations to be used in this course and through the MADS program

1b. Apply basic concepts of vectors, perform basic algebra operations using vectors

1c. Apply basic concepts of matrices, perform basic algebra operations using vectors

1d. Compute basic and conditional probabilities, and understand probability laws

2. Give examples of eigenvalues, eigenvectors, and matrix decomposition (such as Principal Component Analysis).

2a. Demonstrate basic linear algebra and computational properties of matrices including identifying linearity

2b. Recognize how row reduction helps us calculate eigenvalues and eigenvectors

2c. Discuss the concepts of eigenvalues and eigenvectors

2d. Practice concepts of matrix decomposition - particularly principal components

3. Interpret Bayes Rule.

3a. (Re)introduce Bayes Theorem to students and its applications to data science. Give examples of how Bayes Theorem integrates with data science broadly

4. Explain Maximum likelihood estimation in the context of linear regression

4a. Explain the basics of likelihood and maximum likelihood concepts

5. Explain optimization techniques and their connection to mathematics and data science

5a. Describe how optimization works in everyday life, mathematics, and data science

5b. Express the intuition of optimization techniques (gradient descent)

5c. Identify canonical formulations for linear and logistic regression techniques

5d. Express the intuition of loss functions and how they relate to optimization

Course Schedule

- This course begins on **Wednesday, February 3, 2021** and ends on **Tuesday, March 2, 2021**.

- Weekly assignments will be due on **Tuesdays at 11:59 pm (Ann Arbor, Michigan time-Eastern Standard Time - EST, UTC -5) except for the Week 3 writing assignment, which will be due on Friday, February 26 at 11:59pm (Ann Arbor, Michigan time-Eastern Standard Time - EST, UTC -5).**

- **Schedule of Weekly Office Hours via Zoom (Ann Arbor, Michigan time):**

- Mondays 8am-9am with Alex McLeod
- Tuesdays 630pm-730pm with Prof. Grant Schoenebeck
- Wednesdays 430pm-530pm with Prof. Grant Schoenebeck
- Thursdays 8pm-9pm with Alex McLeod
- Fridays 2pm-3pm with Alex McLeod
- Access via Live Events from the course menu.

In **Week 1**, you will review math concepts relevant to data science through (a) a review problem set designed to identify areas of strengths and weaknesses of math fundamentals, (b) lectures revisiting important algebraic concepts, introducing fundamental linear algebra ideas, working with vectors and matrices, and (c) a graded problem set applying these concepts using Python and Jupyter Notebook.

In **Week 2**, you will be introduced to matrix decomposition, row reduction methods for matrices, eigenvalues and eigenvectors, and principal components. This week has two assignments. The first will mathematically apply the week's concepts using Python and Jupyter Notebook in a problem set, while the second will have you read a peer-reviewed article which has incorporated some of the concepts we have covered and write a brief review.

In **Week 3**, we will revisit probability and statistics and apply theories from the course to compute expectation and variance both using analytic techniques and the Monte Carlo method. This week has two assignments. The first will mathematically apply the week's concepts using Python and Jupyter Notebook in a problem set, while the second will have you read a peer-reviewed article which has incorporated some of the concepts we have covered and write a brief review.

In **Week 4**, you will learn about methods of basic optimization techniques. We will cover the general ideas behind optimization, canonical formulations, loss functions, and gradient descent. You will also apply the tools from week 3 to maximum likelihood methods for linear and logistic regression. You will have a graded problem set applying these methods using Python and Jupyter Notebook.

Grading

Course Item	Percentage of Final Grade	Due
Math Review Assignment	0% (but required for course completion)	Tuesday, February 9, 2021
Weekly Practice Set 1	20%	Tuesday, February 9, 2021
Weekly Practice Set 2	20%	Tuesday, February 16, 2021
Writing Assignment 1 (with peer review)	10%	Tuesday, February 16, 2021
Weekly Practice Set 3	20%	Tuesday, February 23, 2021
Writing Assignment 2 (with peer review)	10%	Friday, February 26, 2021
Weekly Practice Set 4	20%	Tuesday, March 2, 2021
Total	**100%**	

Note: All assignments are required to earn credit for this course.

Letter Grades, Course Grades, and Late Submission Policy

Refer to the [MADS Assignment Submission and Grading Policies](#) section of the UMSI Student Handbook (access to Student Orientation course required).

Academic Integrity/Code of Conduct

Refer to the [Academic and Professional Integrity](#) section of the UMSI Student Handbook. (access to Student Orientation course required).

Accommodations

Refer to the [Accommodations for Students with Disabilities](#) section of the UMSI Student Handbook (access to the Student Orientation course required).

Use the [Student Intake Form](#) to begin the process of working with the University's Office of Services for Students with Disabilities.

Accessibility

Refer to the [Screen reader configuration for Jupyter Notebook Content](#) document to learn accessibility tips for Jupyter Notebooks.

Help Desk(s): How to get Help

- Degree program questions or general help - umsimadshelp@umich.edu
- Coursera's Technical Support (24/7) - https://learner.coursera.help/

Library Access

Refer to the [U-M Library's information sheet](#) on accessing library resources from off-campus. For more information regarding library support services, please refer to the [U-M Library Resources](#) section of the UMSI Student Handbook (access to the Student Orientation course required).

Student Mental Health

Refer to the University's [Resources for Stress and Mental Health website](#) for a listing of resources for students.

Student Services

Refer to the [Introduction to UMSI Student Life](#) section of the UMSI Student Handbook (access to the Student Orientation course required).

Technology Tips

- Recommended Technology: This program requires Jupyter Notebook for completion of problem sets and Adobe or other PDF viewer for reading articles.

- Working Offline: While the Coursera platform has an integrated Jupyter Notebook system, you can work offline on your own computer by installing Python 3.5+ and the Jupyter software packages. For more details, consult the [Jupyter Notebook FAQ](#).