Master of Science in Analytics

Course Syllabus

Advanced Linear Algebra for Machine Learning
MSCA 37016
2020 Autumn Pre-Quarter
Location: Online

Danny Ng
lianhuanng@uchicago.edu

Following UChicago policy for the 2020 Autumn quarter, our course will be online via Zoom.

COURSE DESCRIPTION

An advanced linear algebra course focused on the theoretical foundations and applications of linear algebra for machine learning. Upon completion of this course, students will be provided a strong foundation of theoretical linear algebra and linear analysis topics essential for the development of core machine learning and data mining concepts. In addition, various real-life applications of linear algebra for data analytics will be demonstrated.

BOOK REFERENCES

- David Poole "Linear Algebra: A Modern Introduction"
- Gilbert Strang "Introduction to Linear Algebra"
- Lars Eldén "Matrix Methods in Data Mining and Pattern Recognition"
- Dan A Simovici "Linear Algebra Tools for Data Mining"
- Gilbert Strang "Linear Algebra and Learning from Data"
- Stephen Boyd, Lieven Vandenberghe "Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares"

SOFTWARE

R (https://www.r-project.org/)
Python (NumPy and SciPy) (https://www.anaconda.com/products/individual)

EVALUATION

Grades for the course will be based on the following:

- 40% Quizzes (after each class 5x8%)
- 60% Assignments (4x15%)

GRADING SCALE

Pass = 80%–100%
Fail = 0%–79%

TEACHING ASSISTANTS

Teaching assistant: Patrick Butler pnbutler@uchicago.edu Office Hours: TBD
Grader: Patrick Butler pnbutler@uchicago.edu

ATTENDANCE

This course will meet once per week (except for the 4th week when we will meet twice) for 3 hours over 4 weeks according to the course schedule. Your attendance is required and paramount to your success in this course. If you are going to miss a class, please notify and make arrangements with the instructor in advance.

ASSIGNMENTS

There will be 4 weekly assignments (due before the next class) to be submitted online via Canvas https://canvas.uchicago.edu/.

QUIZZES

There will be 5 weekly quizzes (due by the weekend following the class) to be completed online via Canvas https://canvas.uchicago.edu/.

LATE WORK

All assignments must be submitted to the Canvas site for the course on the due date before 11:59 pm. If you turn in an assignment late, 10% credit will be deducted from the total score for each day after the deadline. Assignments turned in more than one week late will not receive credit. In the case of unexpected events, you must contact the instructor before the assignment due date in order to receive a grace period. Students can only receive up to one grace period in the course.

REQUESTING REASONABLE ACCOMODATIONS

Students with disabilities who have been approved for the use of academic accommodations by Student Disability Services (SDS) and need a reasonable accommodation(s) to participate fully in this course should follow the procedures established by SDS for using accommodations. Timely notifications are required in order to ensure that your accommodations can be implemented.

Please follow accommodation implementation instructions provided by the disability liaison in the division after you have completed the SDS procedures for requesting accommodations.

You may want to begin by reading through the information published on this website https://disabilities.uchicago.edu/. Contact SDS at disabilities@uchicago.edu or 773-702-6000 if you are interested in requesting disability accommodations.

ACADEMIC HONESTY & PLAGIARISM

It is contrary to justice, academic integrity, and to the spirit of intellectual inquiry to submit another's statements or ideas of work as one's own. To do so is plagiarism or cheating, offenses punishable under the University's disciplinary system. Because these offenses undercut the distinctive moral and intellectual character of the University, we take them very seriously.

Proper acknowledgment of another's ideas, whether by direct quotation or paraphrase, is expected. In particular, if any written or electronic source is consulted and material is used from that source, directly or indirectly, the source should be identified by author, title, and page number, or by website and date accessed. Any doubts about what constitutes "use" should be addressed to the instructor.

Plagiarism: https://studentmanual.uchicago.edu/academic-policies/academic-honesty-plagiarism/
Copyright: https://www.lib.uchicago.edu/copyrightinfo/

COURSE SCHEDULE

Autumn, Winter, Spring, and Summer quarters are 9 weeks of instruction with the 10th week for assessment or course rescheduling. Refer to the university's academic calendar at www.uchicago.edu/academics/calendar/ for quarterly start and end dates.

Important Note: Changes may occur to the syllabus at the instructor's discretion. When changes are made, students will be notified via Canvas, email, and/or in-class announcement.

Lecture 1: Review of basic linear algebra
Introduction, vector, linear combination, norm, dot product, cosine angle, linear independence, matrix, multiplication, inverse, transpose, trace, determinant, types of matrices e.g. symmetric, orthogonal, *text document similarity, *convolutional neural net CNN

Lecture 2: System of linear equations and vector space
System of linear equations, Gaussian elimination, Gauss-Jordan method, span, vector space, subspace, null space, column space, range, basis, dimension, rank, dimension theorem

Lecture 3: Linear regression and linear transformation
*Linear regression, least squares approximation, orthogonal projection, normal equation, *weighted least squares**,** linear transformation, rotation, *digital image transformation

Lecture 4: Principal component analysis and spectral decomposition
*Principal component analysis PCA, singular value decomposition SVD, covariance matrix, eigenvalue, eigenvector, characteristic polynomial, diagonalization, spectral decomposition

Lecture 5: Convex optimization, Markov chain, and matrix factorization
Gradient, Hessian matrix, positive definiteness, *convex optimization, Markov chain, transition probability matrix, reversible, stationary distribution, *Markov chain Monte Carlo, low-rank matrix approximation, *latent semantic analysis LSA