MATH 412 –Matrix Methods for Data Science and Machine Learning

INSTRUCTOR: Dr. Ahlam Tannouri , *Morgan State University, Mathematics Department, CR 357, 1700 E. Cold Spring Lane, Baltimore, MD 21251 ; E-mail: ahlam.tannouri@morgan.edu*

==========

Matrix Methods for Data Science and Machine Learning

Catalog description: ***Matrix Methods for Data Science and Machine Learning*** *- Three hours; 3 credits. Linear algebra is essential for understanding machine learning algorithms and analyzing large data sets. In this course, students will learn advanced linear algebra topics necessary for organizing information and data, and then using that information and data to solve problems. Prerequisite: MATH 312 with a grade of "C" or better. (OFFERED AS NEEDED).*

Linear algebra is essential for understanding and creating machine learning algorithms and analyzing large data sets. In this course, students will acquire the linear algebra knowledge and skills necessary for organizing information and then using that information to solve problems. Machine learning techniques required to improve the accuracy of predictive models, data science tools and libraries will be introduced. Special topics, applied and abstract are studied. Some topics include Hermitian Matrices, Quadratic Forms, Positive Definite Matrices, Canonical Forms, and Matrix decompositions. Meaningful applications will be introduced to build machine learning algorithms and predictive analytics methodologies.

Course Description: This 3-credit course introduces the students to the advanced linear algebra theory and numerical algorithms behind the models arising in new fields and becoming wide spread in society. Some of the mathematical foundations upon which the basic machine learning concepts are built. The applications of the topics covered in this course will have impactful relevance to practical problems chosen from the sciences, social sciences, business and engineering. These algorithms form the mathematical foundations upon which the basic machine learning concepts are built.

This course is intended to train the next generation of computational science teachers and prepare undergraduate students to pursue a higher education degree in the applied mathematics field and the data science field; Students will develop their expertise to critically analyze the predictive

models built from data and propose novel approaches based on the rigorous mathematical foundations studied in the course to construct better machine learning algorithms.

Culturally responsive teaching leads the instructor to raising important discussion with the students about data collections, biases and the application of machine learning in society such as privacy of individuals, ethics, fairness and social justice.

Recommended Textbook:

Mathematics for Machine Learning, *Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, April 2020, Cambridge University Press, April 2020, ISBN: 9781108455145*
https://mml-book.com . ***(*PDFs of this book freely available after hard book publication.)**
The textbooks will be supplemented with additional notes and readings.

2- Linear_algebra_and_its_applications
[Strang_G.]_Linear_algebra_and_its_applications.PDF

3- Matrix Methods in Data Mining and Pattern Recognition. Lars Elden Lars Eldén, Linköping University Linköping, Sweden

Software: Matlab https://www.mathworks.com/

Preparation:

Basic knowledge of programming and some experience in MATLAB would be preferable, but not required. Students can start preparation by self-studying using the introduction to MATLAB @ **https://www.mathworks.com/help/matlab/getting-started-with-matlab.html.**

Before taking this course, Students should be able to:

- Determine the rank, determinant, eigenvalues and eigenvectors, diagonalization, and different factorizations of a matrix;
  - Find basis and dimension of a vector space and change of basis.
  - Find a basis for the row space, column space and null space of a matrix and find the rank and nullity of a matrix.
  - Compute linear transformations, kernel and range, and inverse linear transformations, and find matrices of general linear transformations.
  - Find the dimension of spaces such as those associated with matrices and linear transformations.
  - Find eigenvalues and eigenvectors and use them in applications.

- Diagonalize, and orthogonally diagonalize symmetric matrices.
- Evaluate the dot product, norm, angle between vectors, and orthogonality of two vectors in R^n.
- Compute inner products on a real vector space and compute angle and orthogonality in inner product spaces.
- Create orthogonal and orthonormal bases: Gram-Schmidt process and use bases and orthonormal bases to solve application problems.

Course Goals:

1. Introduce the students to advanced topics from Computational Linear Algebra.
2. Give interpretation of matrix operations in the context of data
3. Apply Matrix decomposition algorithms to work with data
4. Provide the tools to build a broad mathematical foundation to machine learning.
5. Apply programming skills and use computational linear algebra software to analyze a real-world problem from various fields, make predictions based on data and use results for training and discovery of new results.

Course Learning Outcomes (CLO)

Students completing this course will be able to

1. Use MATLAB to perform matrix computation and decomposition
2. Use MATLAB to plot data and fitting functions
3. Use MATLAB to read and perform operations on images read as data matrices
4. Classify and recognize Hermitian matrices, quadratic forms, positive definite matrices
5. Compute the LU factorization
6. Compute matrix diagonalization
7. Use Gram-Schmidt process to create orthonormal basis.
8. Compute the QR factorization
9. Classify supervised and unsupervised learning algorithms
10. Find the singular value decomposition of a matrix
11. Perform the principal component Analysis
12. Perform linear regression
13. Find the k-mean clusters
14. Define support vector machine.

15. Use a wide variety of computational linear algebra algorithms and apply to data Analysis.
16. Perform evaluation of several algorithms and model selection to apply on data used for projects.

Topic list

- Quick review of vector spaces, subspaces, linear independence, bases, rank, linear transformations.
- Introduction to MATLAB.
- LU decomposition and linear system solving, basic of numerical analysis.
- Norms, inner products, orthogonal bases, Gram-Schmidt orthogonalization, QR factorization
- Projections, least squares problems, data fitting/regression
- Eigenvalues, eigenvectors, diagonalization, positive definite matrices
- **Matrix Decompositions**
  - Cholesky decomposition
  - Eigen decomposition and diagonalization
  - Singular value decomposition
  - Matrix approximation
- Linear regression
- Support vector machines
- Clustering algorithms, k-means
- Dimensionality reduction techniques, SVD/PCA, multi-dimensional scaling
- Applications to statistics & data analysis, web search engines & network problems, information processing (signal & images, error-correcting codes), PageRank algorithm, Recommendation Systems, identification of the foreground in a surveillance video, categorizing documents, the algorithm powering Google's search, reconstructing an image from a CT scan, cryptography, Markov decision processes and more.
- **Applications considered in this course vary based on the interests and experience of students.**

Course Modules

Module # 1: Review Linear Algebra (2 weeks)

Module # 2: Introduction to MATLAB (1 week)

Module # 3: Matrix Decompositions (3 weeks)

Module # 4: Models and Data- Supervised and Unsupervised Learning (2 weeks)

Module # 5: Regression (one week)

Module # 6: Dimensionality Reduction with Principal Component Analysis (2 weeks)

Module # 7: Classification with Support Vector Machine (one week)

Module # 8: Clustering Algorithms (one week)

Module # 9: Review and Applications (2 weeks)

Class Format: based on hybrid Classroom

Under each module posted in the course site, students will find the following items

1- What to do for this module including all deadlines
2- PPT lectures or posted notes
3- Assigned readings
4- Videos taped by the instructor or publicly available with active links displayed
5- Assignment and project
6- Discussion forum
7- Students will be required to do most of readings and practice exercises before the classroom meeting, and watch the videos and the PPT. During class meeting, the instructor will give an overview of the general topic, lecture on the key concept and respond to questions which students are encouraged to submit online before class, lead a discussion forum, and use the technology to highlight and visualize the concepts studied.
8- Students will be asked to prepare short oral presentation on the topic or solution to a problem to present to the class
9- Students can choose to work in small group for the mini-projects and help each other with MATLAB application.
10- Midterm exams and final exam will be supervised in class; some work on exams needs to be run on MATLAB.

==========

This is a three-credit course; you are expected to work at least two hours at home for each hour of lecture. In other words, expect to have 6 hours of preparations and homework each week.

Grading:

MORGAN STATE UNIVERSITY
GROWING THE FUTURE • LEADING THE WORLD

No late assignment will be accepted unless official legitimate excuse documentation is provided justifying any missed deadline.
Students can be called for an oral examination.
Students have the right to discuss grades with the instructor.

- Assignments and mini-projects: five assignments/mini-project each worth 40 points;
- Two midterm exams worth 50 points each
- Final exam:100 points
- Final project 100 points

Grading Scheme: 500 total points=100%

A = 90%-100%

B = 80%-89%

C = 70%-79%

D = 60%-69%

F = 59% and below

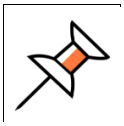
General References:

1. G. Strang: Linear Algebra and Its Applications, 4th Ed., Brooks/Cole, 006.
2. Linear Algebra course is the one offered by MIT Courseware (Prof. Gilbert Strang).
3. Matrix Methods in Data Mining and Pattern Recognition (Fundamentals of Algorithms), by Lars Elden, Published by SIAM ISBN-13: 978-089871669
4. C. Moler: Numerical Computing with MATLAB by Cleve Moler, SIAM 004 (available freely online)
5. https://voices.uchicago.edu/willett/teaching/fall-2019-mathematical-foundations-of-machine-learning/
6. https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/
7. Coding the Matrix: Linear Algebra through Computer Science Applications by Philip Klein, Brown University.
8. http://cs9.stanford.edu/section/cs9-linalg.pdf
9. http://cs9.stanford.edu/notes019fall/cs9-notes1.pdf
10. http://cs9.stanford.edu/notes019fall/cs9-notes3.pdf
11. https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-010/video-lectures/

12. http://aix1.uottawa.ca/~jkhoury/Google.pdf
13. http://aix1.uottawa.ca/~jkhoury/eigenfaces.htm

Recommended BOOK:

14. **Mathematics for Machine Learning,** Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon On, Cambridge University Press, April 2020, ISBN: 9781108455145
15. https://mml-book.com (Links to an external site.). (**PDFs of this book freely available after publication.)**
16. **Linear Algebra and Optimization for Machine Learning**, Charu Aggarwal, Springer International Publishing, 2020, ISBN 978-3-030-40343-0
17. **A Matrix Algebra Approach to Artificial Intelligence,** AUTHOR**:** Xian-Da Zhang, 2020 Springer, ISBN 978-981-15-2769-2



Learner Support Resources

Morgan State Policies & Procedures

1. Acceptable Use Policy ||
2. Academic Integrity ||
3. FERPA ||
4. Privacy Policies ||
5. Accessibility Statements

Morgan State Students Resources

Earl S. Richardson Library Research Support

Disability Support Services

If you need accommodations, please contact the Office of Student Disability Support Services directly. They will, in turn, arrange for accommodations with the course instructor. **For details on accommodation services, please visit the Office of Student Disability Support Services website: http://www.morgan.edu/sdss**
The Morgan State Office of Student Disability Support Services is dedicated to providing support and accommodations.
Call: 443.885.1719.

The on-campus location for the Office of Student Disability Support Services is:
Earl S. Richardson Library, Room 127A
1700 E. Cold Spring Lane
Baltimore, Maryland 21251

Technical Support

- Chat with Canvas Support (Students)
- **Canvas Support Hotline (Students) Call - 1 (443) 873-0119**
- Search the Canvas Guides: Find printed guides and videos - how to use Canvas.
- **Morgan State Service Desk:** servicedesk@morgan.edu - for Non-Canvas related issues. Use this email address to create a ticket, or call 443.885.4357.

Online & Campus-Based Tutoring

Build Rapport

If you find that you have any trouble keeping up with assignments or other aspects of the course, make sure you let your instructor know as early as possible. As you will find, building rapport and effective relationships are key to becoming an effective professional. Make sure that you are proactive in informing your instructor when difficulties arise during the semester so that we can help you find a solution.

Note: Tutoring is strongly advised for students who need help. Don't wait and seek tutoring as soon as possible. If you struggle with the prerequisite content, you need urgent tutoring in order to successfully finish the class.

Smarthinking is 24/7 **online tutoring** available for free to all Morgan students. This service offers tutoring in a variety of subjects, as well as an online writing center. The tutors are live and responsive, not pre-recorded video resources. Smarthinking is available through a link in your Canvas course. Smarthinking Support Contact Information: P: (888) 430-7429, ext. 1| E: support@SMARTHINKING.com

Campus-Based Tutoring Services:
Center for Academic Success & Achievement
Writing Center

Morgan State Counseling Center
The Counseling Center provides a range of psychological and counseling services to meet the mental health and developmental needs of Morgan students. The Counseling Center strives to support the emotional, interpersonal, social, and career development of MSU students.
Any currently enrolled student, graduate or undergraduate, full or part-time, is eligible for Counseling Center services.

Student Life @ Morgan State

MORGAN STATE UNIVERSITY
GROWING THE FUTURE • LEADING THE WORLD

Additional Resources for Students

Technology	Accessibility Statement Link	About the Technology
Internet connection (DSL, LAN, or cable connection desirable)		You will need a broadband internet connection to successfully use the technology in this course.
Canvas Learning Management System	Canvas Accessibility Statement	Canvas is Morgan State University's Learning Management System. All online, hybrid and web enhanced face-to-face courses are delivered using Canvas.
Panopto Video Capture system	Panopto Accessibility Link[i]	Panopto is a video capture tool. It will be used throughout the course to display video content. The Panopto videos will play in your browser, and may also be viewed with assistive technology.
Adobe Reader	Adobe Reader Accessibility Link[ii]	Adobe Reader is software you will use to view PDF documents. The PDF documents are found in each of the Course Modules.
YouTube	YouTube Accessibility Link[iii]	YouTube is a service that displays videos. You will find YouTube videos within many of the course modules. YouTube videos are embedded directly within Blackboard so you do not have to access them outside of Blackboard. However, each YouTube video embedded in Blackboard will also have a link to video located on YouTube itself.
Microsoft Office Word	MS Office Accessibility Link[iv]	You will be required to create Microsoft Word documents in order to complete projects in this course.

Accessibility

- [i]Panopto: http://support.panopto.com/documentation/viewing/accessibility-features
- [ii]Adobe Reader: http://www.adobe.com/accessibility.html
- [iii]YouTube: https://support.google.com/youtube/answer/189278?hl=en
- [iv]MS Office: https://www.microsoft.com/enable/microsoft/mission.aspx
  https://www.microsoft.com/en-us/accessibility

Important Note: Any form of academic dishonesty, including cheating and plagiarism, may be reported to the Office of Student Affairs.

MORGAN STATE UNIVERSITY
GROWING THE FUTURE • LEADING THE WORLD

Right of Revision

Throughout the session, the instructor reserves the right to change any statements, policies or scheduling as necessary due to the dynamic nature of teaching and learning. Students will be informed promptly of any and all changes.

Have a good semester and work hard.