Northeastern University, Department of Mathematics

MATH 4570: Matrix Methods in Data Analysis and Machine Learning

Instructor: He Wang

Email: he.wang@northeastern.edu

Class time and location See **Canvas**: https://canvas.northeastern.edu/

Office Hours/ TA office hours See **Canvas** for the Zoom link, date and time.

Primary Textbooks:

Here is a list of textbooks as references: (None of them are required.)

- Supplementary *lecture notes* will be available on the course webpage.
- *Linear Algebra and Learning from Data.*, Gilbert Strang, Wellesley-Cambridge Press
- *An Introduction to Statistical Learning: with Applications in R* (Springer Texts in Statistics) by Daniela Witten, Trevor Hastie, Robert Tibshirani, Gareth M. James
- *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* - by Aurélien Géron (Good for machine learning by Python)
- *Finite-dimensional linear algebra*, Mark S. Gockenbach, CRC Press. (Good for advanced linear algebra)
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* by Trevor Hastie, Robert Tibshirani, Jerome Friedman (Much more advanced on statistical learning for future study)
- other textbooks: see my recommended reading list:

  https://docs.google.com/spreadsheets/d/1j8LcyTxBTZ5Nh0-P1rb6NT2x_2iXFp1j8ptpvHYQWGA/edit?usp=sharing

Prerequisites: Required prerequisite for registration: **Math 2331**-Linear algebra. (Math 2341 Differential equations and linear algebra is an alternative but need instructor's permission.)

In addition to the required prerequisite, the following are highly recommended: **Math 3081**-Probability and statistics (basic probability concepts needed, e.g., conditional probability and Bayes theorem, probability distribution function, maximum likelihood) and **Math 2321**-Calculus 3 (partial derivatives, Lagrangian multiplier). Some concepts will be reviewed in class.

Prior knowledge of data analysis or machine learning is helpful but not required.

Software: Python will be used throughout the course. Students should be prepared to use *Python* for lab assignments and for the final project. Familiar with any one computer language is helpful, e.g., Python or Mathlab or R.

If you don't have any one of the above prerequisites, you can still be succeed in class. But you need to prepare to contribute double or triple time to study in this course.

Overview: Concepts and methods in linear algebra are key for understanding and creating machine learning and deep learning algorithms. This course provides a rigorous treatment of the concepts and computational techniques of linear algebra, including matrix factorization, positive definite matrices, inner product spaces. Applications to probability and statistics and optimization in high dimensional spaces include matrix calculus, gradient descents, and Newton's methods and principal components analysis. An explicit explanation of the mathematics behind the machine learning and neural network. This course also offers students opportunities to learn and practice Python skills in labs and projects.

List of main topics: (Order of lectures will be different.)

1. Overview of methods in machine learning and relations with linear algebra. Matrix multiplication and factoring matrices; eigenvalues and eigenvectors, symmetric matrices, spectral theorem, positive definite matrices, singular value decomposition, principal components analysis and PCA methods in unsupervised learning.

2. Linear spaces, basis, change of coordinates, Inner product spaces; norms of vectors and functions and matrices, Least squares problems; linear regression, linear methods in classification. Application to support vector machine(SVM) with kernel method.

3. Matrix calculus, optimization, gradient descent and stochastic gradient descent, Newton's method.

4. Neural network for deep learning, convolutional neural network.

Possible computer lab topics:

0. Start with Python: data cleaning, graphing, feature selections.
1. Linear Regressions (with regularization).
2. Logistic/softmax Regressions.
3. Neural network
4. CNN
5. SVM with kernel method.
6. PCA

Grade breakdown:

1. Homework (25%): There will be roughly 5-6 homework assignments which will focus on the theory. Homework will be assigned each week or two. You have one week to finish the homework. In addition to being mathematically correct, your write-ups are expected to show a high level of clarity and completeness.

2. Computer Labs (20%): There will be roughly 4-5 labs will focus on the applications to real world problems using Python. Computer labs are designed to help you learn the concepts and techniques of linear algebra by using them in an computer environment and master the computational tools for important applications of linear algebra. Labs will be graded based on completeness, correctness and clarity in written.

3. Attendance (5%)
4. Two Midterms (30%).
5. Final Projects (20%). The final project will consist of a proposal (1 page), middle stage progress report(2-3 pages), project report (5+ pages) and presentation (15 minutes with a poster or slides). Project groups should contain 4-7 people.

A: 93-100	A-: 90-92.9	B+: 87-89.9	B: 83-86.9	B-: 80-82.9	C+: 77-79.9
C: 73-76.9	C-: 70-72.9	D+: 67-69.9	D: 63-66.9	D-: 60-62.9	F: 0-59.9

To receive full credit, homework/lab assignments must normally be handed in when due. Assignments may be turned in late with a valid excuse; this should be discussed with me and approved in advance if possible or otherwise at the earliest opportunity.

Exam/Tests makeup policy: Makeup exams will be allowed only in the event of a documented medical or other unforeseen emergency. You are responsible for avoiding foreseeable conflicts with the exams.

Late submission policy. Late submission within a week of any assignment without permission will receive at most 90% of the grade. Late submission within a week but after the posting of the solution will receive at most 70% of the grade. Other late submissions will depend on instructor's discretion.

Collaboration: You are welcome, even encouraged, to collaborate on the homework and lab assignments, though we urge you to first attempt working out all of the problems by yourself. However, you are expected to **write answers yourself** and understand everything that you hand in. **Copy** results from any sources will be considered as violating Academic Integrity. Collaboration is not allowed on the quizzes and exams.

University Academic Integrity Policy: The university's academic integrity policy at OSCCR (`http://www.northeastern.edu/osccr/academic-integrity-policy`) discusses actions regarded as violations and their consequences for students.

Title IX: The University strictly prohibits sex or gender discrimination in all university programs and activities. Information on how to report an incident of such discrimination (which includes sexual harassment and sexual assault) is located at `http://www.northeastern.edu/titleix`.

Students with Disabilities: Students who have disabilities who wish to receive academic services and accommodations should follow the standard Disabilities Resource Center (DRC) procedures, `http://www.northeastern.edu/drc/getting-started-with-the-drc`.

College of Science Policies: The current College of Science Academic Course Policies are available at

`https://cos.northeastern.edu/wp-content/uploads/2012/10/COS-teaching-policies-April-2017.pdf`

TRACE: Every student is expected to complete the online TRACE survey at the end of the semester.

A tentative schedule of topics (Spring)

◊ **Week 1:**

Overview of matrix methods in data analysis and machine learning.

Groups, rings, fields

◊ **Week 2:**

Matrix algebra and factoring matrices.

Linear spaces over fields.

◊ **Week 3:**

Basis and dimension.

Transformation and matrices

◊ **Week 4:**

Inner product spaces.

Norm and metric.

◊ **Week 5:**

Least squares problems;

Linear regression

Ridge and LASSO regressions

◊ **Week 6:**

Matrix calculus, optimization,

Gradient descent, Newton's method.

Midterm 1

◊ **Week 7:**

Probability Review

Linear methods in classification.

◊ **Week 8:**

Logistic regression

Softmax regression

LDA/QDA (depending on time)

◊ **Week 9: Spring Break**

◊ **Week 10:**

Neural network for deep learning

CNN

◊ **Week 11:**

Support vector machine

Kernel method

◊ **Week 12:**

Change of coordinates.

Jordan decomposition (depending on time)

Midterm 2

◊ **Week 13:**

Dynamical system

Perron-Frobenius Theorem, Malkov chains

◊ **Week 14:**

Spectral theorem

Singular value decomposition and PCA

PCA methods in unsupervised learning.

◊ **Week 15:** Project Presentations