Matrix Methods in Machine Learning

University of Wisconsin–Madison

ECE/CS/ME 532 – Fall 2017

Outline: This course is an introduction to machine learning that focuses on matrix methods and features real-world applications ranging from classification and clustering to denoising and data analysis. Mathematical topics covered include: linear equations, regression, regularization, the singular value decomposition, and iterative algorithms. Machine learning topics include: the lasso, support vector machines, kernel methods, clustering, dictionary learning, neural networks, and deep learning. Students are expected to have taken a course in calculus and have exposure to numerical computing (e.g. Matlab, Python, Julia, R). Appropriate for graduate students or advanced undergraduates.

Class place and time: 1610 Engineering Hall, 11-12:15pm Mondays and Wednesdays

Instructor: Rebecca Willett (AKA Rebecca Lu; NOT Rebekah Willett in School of Information)
E-mail: `mailto:rmwillett@wisc.edu`
Course webpage: `http://willett.ece.wisc.edu/teaching/fall-2017-532/`
Office: 3537 Engineering Hall
Office hours: 12:15-1pm Mondays and 12-1pm Tuesdays when classes are in session

Prerequisites: (MATH 222 and (ECE 203 or CS 200, 300, 302)) or (graduate or professional standing).

Textbook: *Matrix Methods in Data Mining and Pattern Recognition* by Lars Elden. Textbook is freely available for anybody on the UW–Madison network:

`http://epubs.siam.org/doi/book/10.1137/1.9780898718867`

The textbook will be supplemented with additional notes and readings.

Learning Outcomes: This applies to both graduate and undergraduate students enrolled in the class. Upon successful completion of this course, students will:

- Understand machine learning methods and algorithms through matrix-vector methods and optimization theory.
- Formulate a wide variety of machine learning problems as optimization models and solve them numerically. Understand practical implications of norm choice, regularization, and convexity.
- Investigate an applied machine topic not explicitly covered in class and produce a research project that explains, analyzes, and discusses the topic.

Evaluation: Graduate and undergraduate students will be expected to perform at the graduate level and will be evaluated equally. All students will be evaluated by regular homework assignments, exams, and a final project. The final grade will be allocated to the different components as follows:

- *Homework*: 20%. There are roughly weekly homework assignments (about 10 total). Homework problems include both mathematical derivations and proofs as well as more applied problems that involve writing code and working with real or synthetic data sets.
- *Exams*: 40%. Two midterm exams (20% each), to conclude Parts I and II. No final exam.
- *Final project*: 40%. Students will work in groups (up to 3 students per group) to investigate a machine learning problem or technique using tools learned in class.

Letter grades will be assigned using the following hard cutoffs:

A: 93% or higher
AB: 87% or higher
B: 80% or higher
BC: 70% or higher
C: 60% or higher
D: 50% or higher
F: less than 50%

We reserve the right to curve the grades, but only in a fashion that would improve the grade earned by the stated rubric.

Tentative lecture calendar:

Week 1. Vectors and Matrices in Machine Learning Models
Week 2. Vectors and Matrices in Machine Learning
Week 3. Vectors and Matrices (Bioinformatics examples)
Week 4. Linear Systems and Least Squares (Face recognition)
Week 5. Linear Independence and Orthogonality (Classification)
Week 6. Singular Value Decomposition (Principal Component Analysis)
Week 7. Applications of Principle Component Analysis
Week 8. Singular Value Decomposition (Dimensional Reduction Methods)
Week 9. Ridge Regression and Kernel Methods
Week 10. The Lasso and Proximal Point Algorithms
Week 11. Beyond Least Squares: Alternate Loss Functions
Week 12. Support Vector Machines and Hinge Loss
Week 13. Stochastic Gradient Descent Algorithms
Week 14. Unsupervised Learning and Clustering
Week 15. Neural Networks and Deep Learning