Math 198 – Math for Machine Learning

Spring 2020 Syllabus

1 Course Description

This DeCal is intended for any student interested in the mathematics behind machine learning models, and particularly those who meet the math prerequisites for CS 189 but want to become more comfortable with the material and its applications to machine learning. The course assumes general knowledge in linear algebra, vector calculus, optimization problems, and statistics; however, students do not need to feel like experts in any of these fields to take this course. We will give general descriptions of things like vector spaces and random variables, but will not cover all of the important results from these fields. Rather, we will focus on the results which have the most direct applications to machine learning, and will go over some of these applications so they are more familiar when encountered in CS 189.

Importantly, this course is not meant to replace CS 189, Math 54/110, Math 53, or any of the many statistics and optimization courses available. We will be building on results from these classes which may be presented without proof, and familiarity with the basics of linear algebra, vector calculus, and statistics are essential. This course is meant to bridge the gap between these courses and CS 189, so students feel more comfortable taking 189 in a later semester.

2 Course Location and Instructor Contacts

Course Website: ocf.io/seanvernon/decal

Course Location and Time: MW 4-5 PM, Haviland 12

Facilitators:
Sean Vernon, CS '20, seanvernon@berkeley.edu
Nick Park, Math '20, npark@berkeley.edu

Faculty Sponsor: Prof. Per-Olof Persson, persson@berkeley.edu

3 Topics

Week 1: Syllabus/Introduction to Topics
Application: Perceptrons
Week 2: Review of Vector Spaces, Linear Maps, Matrices
Application: Orthogonal Projections
Week 3: Properties of Matrices, Spectral Theorem
No Application Lecture – President's Day Holiday
Week 4: Special Classes of Matrices
Application: Ordinary Least Squares
Week 5: Singular Value Decomposition
Application: PCA and Four Fundamental Subspaces
Week 6: *Further Applications: Ridge Regression, Total Least Squares, Feature Augmentation, Kernel Trick*
Week 7: Vector and Matrix Calculus
Application: Gradient Descent
Week 8: Convexity
Application: Proofs using Convexity
Week 9: Optimization of Non-Convex Functions
Application: Newton's Method, Gauss-Newton, etc.
Week 10: Probability Basics
Application: Chebyshev's Inequality and LLN
Week 11: Joint Distributions, Covariance, Correlation
Application: Bayes Nets
Week 12: Gaussian Distribution and Estimations
Application: Maximum A Priori

4 Grading and Structure

This course will meet twice per week, and we intend to alternate between math and applications, following the schedule above. Because this material is difficult to digest without practice, homework will be assigned with each math-based lecture and due two lectures later. However, homework will be graded on completion, not correctness. Homework may be submitted in-person or electronically (by emailing a PDF to both facilitators before lecture). Homework should not be completed in a group, nor should homework answers be shared.

There is no textbook for this course – notes will be posted on the course website, and are intended to present results and proofs from lecture in a more formal format. Nonetheless, lecture attendance is required. Students may miss up to two lectures unexcused, and must request excuses in advance from the facilitators for any further missed lectures.

To earn a "P" in this course, students must attend lectures in accordance with the requirements and complete at least 75% of the assigned homework problems. The lowest two homework scores will be dropped. There are no quizzes or tests. If at any point you would like to know your official grade in the course, feel free to email one or both facilitators. Compliments can be sent to Sean, and complaints to Nick.