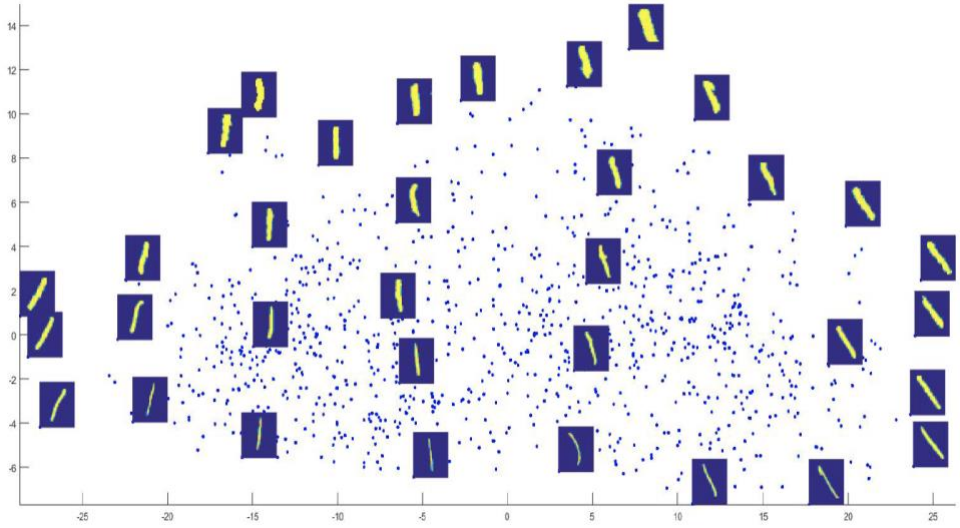
San José State University
Department of Mathematics & Statistics

Math 250 Mathematical Methods for Data Visualization, Spring 2021



Course and Contact Information

Instructor:	Dr. Guangliang Chen
Email:	guangliang.chen@sjsu.edu
Class Days/Time:	TR 9-10:15am
Meeting mode:	Online, synchronous (Zoom ID: 898 1836 7340, registration required)
Office Hours:	TWR 3-4pm (Zoom ID: 851 6038 0197), and by appointment
Prerequisites:	Math 32, Math 39 and Math 163 (each with a grade of B or better)

Faculty Web Page and MYSJSU Messaging

Course materials such as syllabus, lecture slides, and reading material can be found on the course page at http://www.sjsu.edu/faculty/guangliang.chen/Math250.html. Assignments and grades will be posted on Canvas Leaning Management System course login website at http://sjsu.instructure.com.

Piazza

The course will use Piazza at piazza.com/sjsu/spring2021/math250 as a venue for communication and discussions outside of the class meetings. Please post all course-related questions on piazza for fastest response and broadest benefit.

Course Description

Programming basics; data plotting and graphing in 3D or less; advanced linear algebra; dimensionality reduction; visualization of high dimensional data; and application to clustering. 3 units.

Course Learning Outcomes (CLO)

Upon successful completion of this course, students will be able to:

- Use software to carry out various linear algebra operations and statistical computing tasks
- Create publication-quality pictures and graphs
- Perform matrix singular value decomposition and other advanced linear algebra operations
- Apply various dimensionality reduction techniques to high dimensional data and visualize them in low dimensions
- Acquire first-hand experience with large data sets
- Develop a basic understanding of machine learning tasks such as clustering and classification

Required Texts/Readings

Required Textbook

Linear Algebra and Optimization with Applications to Machine Learning: Volume I: Linear Algebra for Computer Vision, Robotics, and Machine Learning, Jean Gallier and Jocelyn Quaintance. Publisher: WSPC (January 15, 2020). ISBN-13: 978-9811207716

Recommended Readings

Foundations of Data Science, Avrim Blum, John Hopcroft, and Ravindran Kannan. Cambridge University Press; 1st edition (January 1, 2020). An unofficial January 2018 version of the book is publicly available from the authors' website at https://www.cs.cornell.edu/jeh/book.pdf.

Technology and Equipment Requirements

The course will make intensive use of specialized software (MATLAB and/or Python) to perform various computing tasks on large data sets. Familiarity with either of them is very helpful but not required.

Students are required to have an electronic device (laptop or desktop) with a camera and built-in microphone in order to attend the Zoom-based classes and the Proctorio-based test. Additionally, they should have access to a scanner (physical or cell phone app) in order to scan and submit their work.

Students are responsible for ensuring that they have access to reliable Wi-Fi during tests. If students are unable to have reliable Wi-Fi, they must inform the instructor, as soon as possible or at the latest one week before the test date to determine an alternative. See the Learn Anywhere website at https://www.sjsu.edu/learnanywhere/equipment/index.php for current Wi-Fi options on campus.

Any student that needs accommodations or assistive technology due to a disability should work with the Accessible Education Center (AEC), and notify the instructor at the earliest possible time.

Zoom Classroom Etiquette

Please arrive at each Zoom meeting on time. If you have to miss a class occasionally, please find out from the instructor or your classmates regarding what's said/done in that class and act accordingly.

Students are encouraged to turn on their cameras (when without privacy concerns) during each Zoom meeting. If using a virtual background, it should be appropriate and professional and should NOT suggest or include content that is objectively offensive or demeaning.

Please raise your hand to get the instructor's permission before you speak up in class. Alternatively, you can type your question or answer (when responding to the instructor's question) in the chat window.

To help keep background noise to a minimum, make sure you mute your microphone when not speaking.

Recording Policy

All lectures will be recorded and shared with the whole class; however, you should still make every effort to attend all classes. The recordings will be deleted at the end of the semester. If you would prefer to remain anonymous during these recordings, then please speak with the instructor about possible alternatives.

Students are not allowed to record without instructor permission: Students are prohibited from recording class activities (including class lectures, office hours, advising sessions, etc.), distributing class recordings, or posting class recordings. Materials created by the instructor for the course (syllabi, lectures and lecture notes, presentations, etc.) are copyrighted by the instructor. The university policy (S12-7) is in place to protect the privacy of students in the course, as well as to maintain academic integrity through reducing the instances of cheating. Students who record, distribute, or post these materials will be referred to the Student Conduct and Ethical Development office. Unauthorized recording may violate university and state law.

Course Requirements and Assignments

Course requirements include weekly homework assignments, two midterm exams, and a final project.

Students are expected to attend all classes and actively participate in class discussions, as they often lead to a deeper understanding of the concepts and are also strongly associated with course grade.

The homework assignments will typically contain both theory and coding questions. For the theory questions, you must show steps to earn full credit, while for the programming questions, you need to present your results in an organized, meaningful way, interpret them carefully, and attach the code you used to obtain the results.

The midterms will be open-book and open notes, but communication with anybody in or outside of the class is not allowed during each exam.

Final Examination or Evaluation

This course ends with a data visualization project that aims to provide students with an opportunity to practice and apply the methods learned in class to large, high dimensional data sets from the internet.

The class will be divided into groups of size two to work on the projects. In special circumstances, a single-person group can be approved by the instructor.

The data sets used by different groups must be distinct. Each data set must have at least 5000 instances and 10 features, and requires advanced approval of the instructor. It is advised that you select a data set as early as possible, because data sets will be available on a first-propose, first-get basis and you also need enough time to complete your project.

You will be asked to report your results through a short oral presentation in class and meanwhile submit a report that contains all the details:

- Your presentation needs to present a high-level summary of your work but you should still give some necessary specifics, such as data information and parameter values for certain algorithms. It should be clear, organized, logical, and self-sustained. We will reserve the final exam day for your presentations.
- Your report must be written using your own language (copying from other places is strictly prohibited and will be given zero points). In addition, it needs to contain a clear structure with the following components: Title, Author, Abstract, Introduction, Experiments, Conclusions (or Discussions), and References. Your report will also be due on the scheduled final exam day.

Your presentation and report will be graded based on clarity, completeness, correctness and originality.

Grading Information

Students may collaborate on homework but must write their own codes and solutions independently. Copying and other forms of cheating will not be tolerated and will be reported to the SJSU Office of Student Conduct.

You must submit homework on time to receive full credit. Late submissions within 24 hours of the due time can still be accepted but will receive a penalty of 10% of the total number of points. Submissions that are late for more than one day (24 hours) will not be accepted for any reason.

No make-up exam will be given if you miss a midterm exam, unless you have a legitimate excuse (such as illness or other personal emergencies) and can provide documented proof.

For both homework and tests, it is your work (in terms of correctness, completeness, and clarity), not just your answer, that is graded. Thus, correct answers with no or poorly written supporting steps may receive very little credit.

The weights in determining the semester average are:

- Homework: 20%
- Midterm 1: 30%
- Midterm 2: 35%
- Final project: 15%

The following cutoffs will be used for assigning students' course grades (however, the instructor reserves the right to slightly adjust these percentages in order to better reflect the actual distribution of the class in the end):

A+: 98% to 100%	B+: 86% to 89%	C+: 73% to 75%	D+: 63% to 64%	F: 0% to 55%
A: 93% to 97%	B: 80% to 85%	C: 68% to 72%	D: 58% to 62%	
A-: 90% to 92%	B-: 76% to 79%	C-: 65% to 67%	D-: 56% to 57%	

University Policies

Per University Policy S16-9 *(http://www.sjsu.edu/senate/docs/S16-9.pdf)*, relevant information to all courses, such as academic integrity, accommodations, dropping and adding, consent for recording of class, etc. is available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/".

Disclaimer: *The instructor reserves the final right to interpret, and make changes to, all the policies that are stated in this course syllabus.*

Math 250 Tentative Course Schedule, Spring 2021

Week	Date		Topics
1	January	28	Course introduction and overview
2	February	2	Review of linear algebra
2		4	Review of multivariable calculus
3		9	Matrix computing in MATLAB
3		11	High quality data plotting in 1D
4		16	High quality data plotting in 2D
4		18	High quality data plotting in 3D
5		23	Rayleigh quotient
5		25	Singular value decomposition (SVD)
6	March	2	Singular value decomposition (SVD)
6		4	Generalized inverse and pseudoinverse
7		9	Generalized inverse and pseudoinverse
7		11	Matrix norm and low-rank approximation
8		16	Review
8		18	Midterm 1 (cumulative)
9		23	Principal component analysis (PCA)
9		25	Principal component analysis (PCA)
March 29 – April 2: Spring Recess			
10	April	6	Multidimensional scaling (MDS)
10		8	ISOmap
11		13	Linear discriminant analysis (LDA)
11		15	Linear discriminant analysis (LDA)
12		20	Laplacian Eigenmaps
12		22	Nonnegative Matrix Factorization
13		27	Review
13		29	Midterm 2 (comprehensive)
14	May	4	Introduction to data clustering
14		6	Introduction to data clustering
15		11	Introduction to data clustering
15		13	Introduction to data clustering
		20	Project presentations and reports due (Thursday, 7:15am - 9:30am)