USC Viterbi School of Engineering

MATLAB: Advanced Data Processing and Applications

ITP 499 (3 Units)
Spring 2018

Catalogue Description	This class covers methods for handling medium to large amounts of digital information. Students with many professional aspirations, not just engineering, will find this course useful as it teaches MATLAB programming techniques in order to solve general and practical problems ranging from large-scale data processing and analysis, to text file manipulation and organization.
Objective	Upon successful completion of this course, students will be able to perform pre-processing, manipulation, organization, and analysis of large amounts of data, which can be in various forms (text, number, or a mixture of both). The goal is to teach students several fundamentals, and then guide them through approximately five projects, which will take, on average, 3 - 4 weeks each. Successful completion of these projects will give the students a small library from which they can draw on to solve many of the data handling problems they may face in the biology, chemistry, medical, economics, accounting, finance, entertainment, engineering, and business industries.
Recommended Preparation	ITP 168 or equivalent is highly recommended
Instructor	Tyler Davis, Ph.D. (tylerdav@usc.edu)
Office Hours	Listed on Blackboard under Contacts
Lab Assistants	Listed on Blackboard under Contacts
Course Hours	Tuesdays, 6:00 - 8:50 pm
Course Structure	The class meets for two hours and 50 minutes once a week. These sessions include lecture and laboratory time. Lectures will consist of instruction, and laboratory time will often be available immediately following a lecture, in which students are free to work on their projects and solicit help from the instructor and teaching assistant(s).
Required Textbook	None
Grading	The following percentage breakdown is used to determine the final grade:

Project 0	5%
Project 1	15%
Project 2	15%
Project 3	30%
Project 4	25%
Laboratory Performance	10%
TOTAL	**100**

Grading Scale The following scale is used to determine the letter grade:

93% and above	A
90 – 92%	A-
87 – 89%	B+

83 – 86%	B
80 – 82%	B-
77 – 79%	C+
73 – 76%	C
70 – 72%	C-
65 – 69%	D+
67 – 68%	D
66%	D-
65% and below	F

If you are taking the class with a grade of P/NP, you must earn a grade of 70% or higher in order to receive a P. Final grade percentages are calculated to two decimal places and rounded to the nearest hundredth. For example, 89.99 is a B+, while 89.995 is rounded to a 90, which is an A-.

Homework

Homework for this course will consist of projects. The projects will be posted on Blackboard under the "Assignments" section. Each project will include instructions, a due date, and a link for electronic submission. Assignments must be submitted using this link. All assignments will be submitted digitally through Blackboard except where specified. Do not email them to the instructor or lab assistant(s).

Students will also be evaluated on their Laboratory Performance. Lecture and Laboratory participation, thus attendance, will be a significant contribution to a student's overall grade. For larger, more complicated projects, extra time with the instructor has been scheduled into the lecture series to aid student success. Students are highly encouraged to attend these sessions (which are scheduled during the regular lecture time) and make use of this time with the instructor to improve their code.

It is the students' responsibility to submit their projects on or before the due date.

You will not be able to save your work on the ITP lab computers. ITP is not responsible for lost work.

Policies

ITP offers open lab use for all students enrolled in ITP classes. These open labs are held beginning the second week of classes through the last weeks of classes. Hours are listed at **http://itp.usc.edu/labs/**. The open labs will not have a lab assistant for this specific class. These lab times are available in case you do not have a computer or extra time is needed to complete an assignment.

Computer Programming

This course requires the use of MATLAB. MATLAB is available to USC students at no additional cost. Please visit USC ITS at **https://itservices.usc.edu/software/** to find out more information about downloading the software.

Incomplete and Missing Grades

The University Grading Handbook (see link below) contains details on incomplete and missing grades, as well as other grading concerns.
http://www.usc.edu/dept/APR/grades/gradinghandbook/index.html

A Missing Grade (MG) should only be assigned in unique or unusual situations such as for those cases in which a student does not complete work for the course before the semester ends. All missing grades must be resolved by the instructor through

	the Correction of Grade Process. One calendar year is allowed to resolve an MG. If an MG is not resolved within one year, the grade is changed to Unofficial Withdrawal (UW) and will be calculated into the grade point average as zero grade points. An Incomplete Grade (IN) is assigned when work is not completed because of a documented illness or other "emergency" occurring after the 12th week of the semester (or 12th week equivalency for any course scheduled for less than 15 weeks).
Viterbi Honor Code	Engineering enable and empowers our ambitions and is integral to our identities. In the Viterbi Community, accountability is reflected in all of our endeavors. Engineering + Integrity. Engineering + Responsibility. Engineering + Community. Think Good. Do better. Be great. These are the pillars we stand upon as we address the challenges of society and enrich lives.
Academic Integrity	USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by the instructor, and the obligations both to protect one's own academic work from the misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. SCampus is USC's Student Guide to Policies and Conduct Code and can be found at **http://scampus.usc.edu**. Section 11 contains the Behavior Violating University Standards and Appropriate Sanctions and can be found at **http://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions/**. Students will be referred to the Office of Student Judicial Affairs and Community Standards (SJACS) for further review, should there be any suspicion of academic dishonesty. The review process can be found at **http://www.usc.edu/student-affairs/SJACS/**. An academic integrity tutorial can be found at **http://www.usc.edu/libraries/about/reference/tutorials/academic_integrity/index.php** Examples of behavior violating University standards: • The submission of material authored by another person but represented as the student's own work, whether that material is paraphrased or copied in verbatim or near-verbatim form. • Obtaining for oneself or providing for another person a solution to homework, a project or other assignments, or a copy of an exam or exam key without the knowledge and expressed consent of the instructor. • Unauthorized collaboration on a project, homework, or other assignment. • Fabrication: Submitting material for lab assignments, class projects or other assignments which is wholly or partially falsified, invented or otherwise does not represent work accomplished or undertaken by the

	student. If the instructor, a grader, or a lab assistant suspects you of academic dishonesty, it has to be reported to SJACS. Do not share lab assignments with another student. Do not submit another student's work as your own. Do not look at other students' papers during exams. Do not leave the room during an exam without permission. **Do not cheat! As Trojans, we are faithful, scholarly, skillful, courageous, and ambitious.**
Support Systems	Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* at **http://equity.usc.edu/** or to the *Department of Public Safety* at **http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us**. This is important for the safety of the whole USC community. Another member of the university community - such as a friend, classmate, advisor, or faculty member - can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* (http://www.usc.edu/student-affairs/cwm/) provides 24/7 confidential support. The sexual assault resource center webpage at **sarc.usc.edu** describes reporting options and other resources.
Disability Services	*The Office of Disability Services and Programs* (**http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html**) Provides certification for students with disabilities and helps arrange the relevant accommodations. Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. **Please be sure the letter is delivered to our course instructor as early in the semester as possible.** If you need accommodations for an exam, the form needs to be given to the instructor at least two weeks before the exam, but preferably at the beginning of the semester.
Emergency Preparedness	If an officially declared emergency makes travel to campus infeasible, ISC Emergency Information (**http://emergency.usc.edu/)** will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology. Additional information about Campus Safety and Emergency Preparedness can be found at **http://preparedness.usc.edu**.

Course Outline*				
Week/Lect.	**Day**	**Topic**	**Description**	**Due**
1	1/9	MATLAB ENVIRONMENT	Course Outline, how to operate in MATLAB	
			Finish Fundamental MATLAB practices	
2	1/16		Scripts and Functions	1/16 Project 0
3	1/23	FUNDAMENTALS	Reading data from a single file	
			2D and 3D arrays, basic array calculations	
4	1/30		Filtering, Checks	
			Sorting	
5	2/6		Plotting	
			Class Time for Project 1	
6	2/13		Digitizing data from a reference	2/13 Project 1
			Handling Data too Large to Store in MATLAB (MapReduce)	
7	2/20		Regular Expressions	
8	2/27	LARGE-SCALE DATA PROCESSING & ANALYSIS	Directory Hierarchies Part I: Getting in (2 lectures)	2/27 Project 2
9	3/6		Directory Hierarchies Part II: Finding out what contents are inside	
			Directory Hierarchies Part III: Getting back out	
-- --	**3/13**		**SPRING RECESS**	
10	3/20		Preprocessing: changing file names, formats, locations	
			Data Analysis	
11	3/27		For loop plotting	
			Organizing & saving MATLAB-based analyses	
12	4/3		Large-scale data handling: putting it all together	
			Class Time for Project 3	4/10 Project 3
13	4/10		Status and checks, REVIEW	
		PARSING, MANIPULATING TEXT FILES	Reading in text-dominated data	
14	4/17		Changing portions of text file	
			Building larger text files	
15	4/24		Making use of that final product	
			Class Time for Project 4	5/1 Project 4
16	5/1		COURSE REVIEW	
5/2-5/9			**Finals Week**	

*This course outline is for planning purposes and is subject to change.

Project	Description**
0	A relatively simple set of calculations will be conducted using a script and a function file. This project is intended to get students familiar with how to operate in MATLAB, edit their scripts and functions, and debug their own code.
1	Project 1 will require students to load data into MATLAB, and then pick out several columns of data by searching for specific text to locate columns of interest. Once the correct columns of data are located, students will be required to organize the data in a given order. Once organized, students will produce multiple plots and conduct some basic analyses.
2	When writing a report or preparing a presentation, far too often do students simply "Print Screen" and/or use a visual editing tool to reproduce data from a reference. When this is done, students lose the ability to define any relationship that may exist for that data set, as well as lose any technical understanding of that information. Furthermore, the original data representation is often skewed, resulting in an unprofessional copy. For the second project, the students will be required to find two references specific to a research field of their interest. These references must contain at least one plot. The students will then digitize that plot, as well as gain a deeper understanding of the underlying data by determining whether the data can be fit and how best to fit them.
3	This project will be the largest of the semester. It is an extension of Project 1, but will also include pre-processing, which at a minimum consists of filtering data based on specified conditions. It will also require the student to traverse several layers of a directory hierarchy, access specific data files within the deepest sub-directory (there will also be other files to serve as noise within the directory), conduct some processing on those data files, exporting their analysis to a created directory, and then exit the directory hierarchy without altering the original files or original data directory structure.
4	Often times, a large task, such as organizing thousands of data files (e.g. audio or .mp3 or .mp4 files), will consist of many steps, where just a handful of details change from file to file. Students will be required to develop a text file which is approximately 100 lines long. This text file, when executed will accomplish a relatively simple task (e.g. moving a file into a folder). Students will then alter several specific lines or characters of their original text file, generating a second text file. The second text file represents the code required to manipulate or organize the second file of many. Students will iterate this process generating a larger text file, which is a concatenation of the manipulated text files stemming from the original. Upon execution of this larger text file, which will now be on the order of 10,000 lines (two orders of magnitude larger than the original), will be required to accomplish something useful on a larger scale, like alphabetizing several hundred .mp4 files. A significant amount of freedom will be given to the students in deciding the overall end goal this project will accomplish.

**Projects are subject to change.