Home

MAS 4115, Linear Algebra for Data Science, S22 (OLD)

MAS 4115, Linear Algebra for Data Science, F22

MAS 4115, Linear Algebra for Data Science, F22

Note: The class format and conduct may change based on University Guidelines and recommendations.

Course Number: MAS 4115

Instructor: Hubert Wagner, hwagner[…]ufl.edu

Time and Location:

Section 1205 | M,W,F | Period 5 (11:45 AM – 12:35 PM), Location: LIT 0233

Section 9199 | M,W,F | Period 6 (12:50 PM – 1:40 PM), Location: LIT 0233

Office Hours: M,W,F | Period 7 (1:55 PM – 2:45 PM), Location: LIT 428 or zoom (please let me know at least one day before, with brief information what it is about).

Course Description: A second course in linear algebra, focusing on topics that are the most essential for data science. Introduces theory and numerical methods required for linear problems associated with large data-sets and machine learning.

Course Focus: We will highlight linear algebra concepts using practical examples in data analysis, many of which coming from my experience in industry. We will also put much emphasis on developing intuitions and communication skills necessary in such jobs.

Course Goals and Objectives: A student who successfully completes this course will be able to:

- Map data analysis problems to concepts of linear algebra.
- Articulately discuss and clearly explain mathematical concepts in the context of data science.
- Use Python libraries to solve concrete data analysis problems. In particular, construct and tune feedforward neural networks using a high-level library such

as Keras.

- Implement basic low-level routines avoiding common pitfalls related to conditioning, stability and computational complexity.
- Continue learning more advanced techniques on their own or by taking more specialized courses.

Schedule overview:

- Week 1-2: Review of (very) basic Linear Algebra: vectors, Euclidean norm and distance. Data as a high-dimensional vector space. Basic data analysis techniques and problems based on these concepts (including similarity search, k-means). Intro to Python + numpy.
- Week 3-4: Basic concepts related to optimization, gradient descent; nonlinear dimensionality reduction (focusing on tSNE). More Python and libraries.
- Week 5-6: More Linear Algebra: dot product, orthogonality; techniques based on hyperplanes (SVMs, kd-trees).
- Week 7-10: More advanced Linear Algebra: matrices, linear transformations, various matrix decompositions focusing on SVD; Problems and techniques based on these concepts (including linear regression, (linear) dimensionality reduction, low-rank approximations).
- Week 11-14: Various neural networks, backpropagation, various loss functions, elements of information theory. Intro to Keras library.
- Week 15: Summary.

Logistics:

Prerequisites: A course in linear algebra (MAS 3114, MAS4105 or equivalent course) is required.

Programming Prerequisite: Class demos, examples and homework assignments will **use Python**.

However, you **are not expected** to be proficient in Python at the start of the course. You **are expected** to have enough programming experience in another language to pick up the basics of Python quickly.

The first couple of weeks of classes are meant to help you pick up Python and its basic libraries (numpy, matplotlib, scipy, later sklearn and Keras).

We will use google colab(https://colab.research.google.com/) and similar online programming environments, so no complicated setup of a programming environment is necessary.

Participation: This is a synchronous, face-to-face class.

Work and grading.

The grade will depend on:

- homework: 50%
- exams: 30%
- written project: 10%
- participation and activity : 10%

Homework. Homework will be posted on Canvas and you will upload your solution in the Assignments section before the stated deadline. Usually you will have one week to complete the homework, which will be graded within a week (from a deadline). **Most homework assignments will be mini-projects and will require programming.**

Written project. There will be one project with longer time duration. The goal will be to research a topic of your choice and write down your observations.

Exams: There will be three exams (including the final exam). The format will be announced for each exam. The exam tasks will check your understanding of details and of the theoretical foundations of the used methods.

Activity and participation: I will encourage discussion during classes. Good questions and answers will be rewarded.

Grading. The grade ranges for the total scores will be no tougher than: 93-100% A, 90-92% A-, 88-89% B+,83-87% B, 80-82% B-, 78-79% C+,73-77% C, 70-72% C-, 60-69% D, <60% F.

The outlined arrangements may change based on University guidelines and student needs. We will discuss and finalize them during the first week of classes.

Additional information

Resources: The course will be based on lecture notes. The following resources may be useful as additional references:

- *Linear Algebra and Learning from Data,* by Gilbert Strang, Wellesley-Cambridge Press; First edition (2019).
- Mathematical Foundations for Data Analysis, by Jeff M. Philips, online version: Springer, 2021 (online version, 95% complete: https://mathfordata.github.io/versions/M4D-v0.6.pdf
- *Neural Networks and Deep Learning* by Michael Nielsen http://neuralnetworksanddeeplearning.com/index.html
- *Deep Learning* by Ian Goodfellow and Yoshua Bengio and Aaron Courville http://www.deeplearningbook.org/

Honor Code and Collaboration: In this course authorized aid on projects and hw consists of talking to me, other students, reading the documentation for your computational platform, and looking at the notes for this course. This means that you are not allowed to look on-line, in other books specifically for solutions to the hw or projects, or at the written solutions of other students. Looking up general stuff like

definitions, usage of Python libraries is of course fine. You can collaborate with fellow students but must write up and code individually.

Excused Absences: In certain circumstances a student will be able to make up a missed exam. These circumstances could include medical situations, family emergencies, travel for University activities (eg. band, debating club, etc), and religious observances. In these cases the student must inform me before or within one week after the missed work and **provide written documentation**.

Additional Information:

Grades: Grading will be in accord with the UF policy stated at https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx.

Honor Code: "UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class."

Class Attendance: "Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx."

Grading Disputes: Any issues or questions about the grading of homework or exams must be brought to my attention within one week after the exams or homework are returned to the class.

Diversity Statement: I am committed to diversity and inclusion of all students in this course. I acknowledge, respect, and value the diverse nature, background and perspective of students and believe that it furthers academic achievements. It is my intent to present materials and activities that are respectful of diversity: race, color, creed, gender, gender identity, sexual orientation, age, religious status, national origin, ethnicity, disability, socioeconomic status, and any other distinguishing qualities.

Accommodations for Students with Disabilities: "Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting https://disability.ufl.edu/students/get-started/. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester."

Online Evaluations: "Students are expected to provide professional and respectful

feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/."

Contact information for the Counseling and Wellness Center: https://counseling.ufl.edu/, 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

U Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit umatter.ufl.edu/ to refer or report a concern and a team member will reach out to the student in distress.