

Johnathan Bush

College of Liberal Arts & Sciences

College of Liberal Arts and
Sciences

Home

Research

Publications

Courses

Curriculum Vitae

TDA seminar –
Spring 2024

Linear Algebra for
Data Science: MAS
4115, Spring 2024

Linear Algebra for Data Science: MAS 4115, Spring 2024

Instructor:

Johnathan Bush (bush.j@ufl.edu)

Time and location

Monday, Wednesday, Friday in Little Hall 121, period 6 (12:50-1:40 pm ET)

Office Hours

Tuesday, Wednesday in Little Hall 496, period 7 (1:55-2:45 pm ET), or by appointment.

From the course catalog:

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 large data-sets and machine learning. Topics include LU, QR, and singular-value decompositions; conditioning and stability; the DFT and filters; deep learning; fully connected and convolutional nets.

Prerequisites

(MAS 3114 or MAS 4105) and MAC 2313.

Specific to our section:

Programming prerequisites

Labs and certain homework assignments will use Python. You **are not** expected to have any previous experience programming with Python at the start of the course. On the other hand, you **are** expected to have enough experience with a computer to be able to pick up the very basics of Python quickly.

The first few weeks of the course are designed to help you pick up the very basics of Python from scratch. We will go over basic syntax and functions. We will use [Google Colab](#) and Jupyter notebooks, so you will not need to set up a

complicated programming environment.

You are not required to own a laptop, but **if you have a laptop, please bring it to class every day**. Most of our meetings will include time to work on demos and exercises, and you will be able to get the most out of our time together if you can actively participate.

Course goals and objectives

The core objective of this course is to develop an understanding of the *mathematics* behind concepts from linear algebra that are relevant to data science. A secondary objective is to learn how to *implement* these concepts in practice.

A student who successfully completes this course will be able to:

Map data analysis problems to concepts of linear algebra.

Articulate and clearly explain mathematical concepts in the context of data science.

Perform basic linear algebra computations by hand and in Python.

Understand and implement various standard matrix decompositions and use their numerical implementation for data analysis and solving linear problems.

Perform (and understand the theoretical foundations behind) basic machine learning tasks.

Tentative schedule:

Week 1-2: Review of basic linear algebra: core definitions including linear independence, basis, dimension, matrices, linear transformations, subspaces, etc. Intro to Python.

Week 3-4: Systems of equations, LU decomposition. More Python and libraries (focusing on NumPy).

Week 5-7: Inner products, orthogonality, norms. Clustering, nonlinear dimensionality reduction (e.g., t-SNE), gradient descent.

Week 8-12: Eigenvalues, eigenvectors, spectral theorem, orthogonalization, matrix decompositions focusing on singular value decomposition (SVD) and principal component analysis (PCA); Problems and techniques based on these concepts (linear dimensionality reduction, low-rank approximations, power method etc).

Week 13-14: Introduction to basic concepts from machine learning including simple linear classifiers, backpropagation, various loss functions, stochastic gradient descent etc. Intro to Keras library and neural networks.

Week 15: overflow, review.

This schedule is tentative and may change in response to the pace of our discussion.

Assignments and exams:

Homework. Homework assignments will be posted on Canvas. Some exercises

will require basic programming. In general, late submissions will not be accepted.

Written project. There will be one longer project assigned toward the end of the semester. Your goal will be to both *research* a topic of your choice and *implement* tools from class in the context of data science.

Exams: There will be two midterm exams and a final exam. These will be traditional paper-and-pencil exams taken without the use of technology. The purpose of the exams is to check your understanding of the theoretical foundations for our methods.

Activity and participation: I will encourage discussion and participation during class. Most class meetings will include time to work on demos and exercises, so there will be ample opportunity to participate.

Final grades:

Grades will be assigned based on homework (40%), the written project (10%), exams (40%), and activity and participation (10%). Final letter grades will be assigned on a scale no stricter than the following: 93-100 A, 90-92 A-, 87-89 B+, 83-86 B, 80-82 B-, 77-79 C+, 73-76 C, 70-72 C-, 67-69 D+, 63-66 D, 60-62 D-, 0-59 E.

Collaboration:

You can, of course, use the internet to look up definitions and Python documentation. You are also encouraged to work with other students on assignments, but **you must write and submit your own solutions/code**. Additionally, you must follow the guidelines below regarding authorized aid and let me know if you have questions.

In this course, authorized aid on assignments consists of talking to me, talking to other students, reading the documentation for your computational platform, and looking at the notes/demos/text-based resources for this course. Therefore, unauthorized aid includes, for example, reading solutions to problems online (or elsewhere), submitting work that is not your own, etc. See the “academic honesty” section below for more detail.

Text-based resources:

This course will be largely based on notes and demos. We will not follow a textbook throughout the semester, but here are some resources that may be useful:

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 <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/>

Additional information:

Attendance:

Attendance is required. We will adhere to the university attendance policies that can be found here: <https://catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/>.

Grade points and grading policies:

We will adhere to the university grading policies that can be found here: <https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/>.

Diversity statement:

The University of Florida and the Department of Mathematics are committed to diversity and inclusion of all students. We recognize the diversity of backgrounds and learning needs of our students and strive to create a more inclusive and welcoming environment for everyone. We strongly believe that an inclusive learning environment promotes higher academic achievements.

Disabilities statement:

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/>. 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.

Course evaluation:

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/>.

Academic honesty:

University of Florida 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 (<https://sccr.dso.ufl.edu/process/student-honor-code/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. If you have any questions or concerns, please consult with the instructor.

Campus resources:

Health and wellness:

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

Counseling and Wellness Center: Visit the Counseling and Wellness Center website at <https://counseling.ufl.edu/> or call 352-392-1575 for information on crisis services as well as non-crisis services.

Student Health Care Center: Call 352-392-1161 for 24/7 information to help you find the care you need, or visit the Student Health Care Center website at <https://shcc.ufl.edu/>.

University Police Department: Visit the UF Police Department website at <https://police.ufl.edu/> or call 352-392-1111 (or 9-1-1 for emergencies).

UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; Visit the UF Health Emergency Room and Trauma Center website at <https://ufhealth.org/emergency-room-trauma-center/>.

GatorWell Health Promotion Services: For prevention services focused on optimal wellbeing, including Wellness Coaching for Academic Success, visit the GatorWell website at <https://gatorwell.ufsa.ufl.edu/> or call 352-273-4450.

Academic resources:

Career Resource Center: Reitz Union, 352-392-1601, <https://career.ufl.edu/>. Career assistance and counseling.

Library Support: ask@ufl.libanswers.com. Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center: Broward Hall, 352-392-2010, 352-392-6420, or <http://teachingcenter.ufl.edu/>. General study skills and tutoring.

Writing Studio: 302 Tigert Hall, 352-846-1138, <http://writing.ufl.edu/writing-studio/>. Help brainstorming, formatting, and writing papers.