Mathematical Foundations for Computational Social Science
ECON GA 4002

Sargent

Fall Semester 2021

3 credits

E-mail: cc7768@gmail.com — Web: **http://www.chasegcoleman.com**

E-mail: spencerlyon2@gmail.com — Web: **http://www.spencerlyon.com**

E-mail: ts43@nyu.edu — Web: **http://www.tomsargent.com/**

Office Hours: TBD — Class Hours: TBD

Course Description

This course introduces core concepts underlying computational social science, including, optimization, probability, and statistical modeling. We learn how to apply these concepts to a variety of social science questions.

Pace

Being a graduate course, the level and amount of material and the rate at which it is to be presented and digested will be about twice those encountered in a challenging undergraduate class.

Course Materials

- **Linear Algebra Done Right, 3rd edition** by Sheldon Axler
- **All of Statistics** by Larry Wasserman
- **Convex Optimization** by Stephen Boyd
- **Probability and Information: An Integrated Approach** by David Applebaum
- **QuantEcon** Various lectures from QuantEcon available at **https://quantecon.org**
- **Software Carpentry** We will use materials from the Software Carpentry organization available at **https://software-carpentry.org/**
- **Statistical Rethinking** by Richard McElreath

Prerequisites

Students should have taken at least a course in calculus and a course in linear algebra. Prior courses in probability and statistics will also be helpful.

Highly motivated graduate students lacking some prerequisites can still succeed in this course but will have to put in more time and effort. Such students should meet with us well in advance of the course so that we can recommend readings that will help close the gaps in their preparations. Also, a willingness to consult Wikepedia for math and statistics concepts will help all students.

Course Objectives

This course teaches powerful mathematical tools and topics that will

- Demonstrate how computational tools can be used to formulate and to solve problems in the social sciences that may not have otherwise had good answers
- Allow students themselves to answer questions in the social sciences using computational tools
- Give students a concise "language" (i.e., mathematics) that will allow them to learn other new tools that are not covered in this course

Course Structure

Class Structure

This course will meet once a week for 3 hours.

Class will be treated as a mixture of lecture time and lab time. Students should bring, and expect to use, their laptops every time the class meets.

Assessments

A mixture of homework assignments, in-class quizzes, exams, and a final project will be used to evaluate students' learning achievements.

Homework: At the beginning of the course, homework will be assigned each week. Later, there will be less frequent homework assignments because then you will be spending substantial time on your class project. Your two lowest homework grades will be dropped.
In-class quizzes: There will be 4 in-class quizzes during the semester. These quizzes will be open book. Only your three highest scores will contribute to your grade (your lowest score will be dropped).
Exams: There will be 2 take-home exams.
Project: The class project will be structured to give you practice creatively applying the tools that you have learned to a real-world problem.

Except for quizzes and exams, we highly encourage students to work together. We have found that groups of 3-4 seem to work best. We believe that collaborative work is the best way to learn the type of material that we cover. We advise students not to rely on others to do work that you do not understand.

Grading Policy

Assignments described above will be the main inputs to the grade for the course. Assignments will be weighted evenly within groups and overall according to the following decision rule:

- Homework assignments: 25%
- In-class quizzes: 15%
- Tests: 20%
- Project: 40%

This weighting reflects our opinion that the most important skills to be acquired in this class are communicated by one's ability successfully to apply the tools that you learn to an interesting question in the social sciences.

Grading policy: A =100-90%; A- = 90-85%; B+ = 84-80%; B = 79 - 70%; C = 69 - 50%; D = 49 and lower

Schedule and weekly learning goals

The schedule is tentative and subject to change. Several of the modules below will occupy more than one week. The learning goals target key concepts to be mastered after each module. Successive modules build on early modules.

Week 01: The calculus

Sources and tools:

- `sympy package:` https://www.sympy.org/en/index.html
- https://python.quantecon.org/complex_and_trig.html
- https://python.quantecon.org/complex_and_trig.html

Topics to be mastered:

- Derivatives
- Integrals
- Complex valued functions and their derivatives
- Fourier transforms and the convolution theorem
- The fast Fourier transform and its inverse
- First and second order conditions for extremization
- Euler's method for solving ordinary differential equations
- Runge-Kutta numerical intergration
- Gaussian quadrature integration
- Introduction to genetic algorithms and other procedures for extremizing over rugged landscapes

Week 02: Applications of Linear algebra

Sources and tools:

- https://python.quantecon.org/linear_algebra.html
- `scipy.linalg`

Topics to be mastered:

- Vectors and matrices
- Elementwise operations, matrix multiplication, determinants
- Linear independence and dependence

- Linear transformations
- Matrix inverses
- Determinants and volumes
- Eigenvalues and eigenvectors
- Courant-Fischer minimax theorem and some of its applications in machine learning
- Inner product spaces
- Hilbert spaces
- Orthogonal projection
- Gram-Schmidt orthogonalization
- The Cholesky decomposition

Week 03: Convex optimization, I

Sources and tools:

- Chapters 1-3 of Stephen Boyd's text.
- `CVXpy:` https://www.cvxpy.org
- `scipy.optimize.linprog`

Topics to be mastered:

- Linear programming problems
- Primal and dual problems
- Relationship of a "welfare theorem" and prices
- Pathologies
- Simplex algorithm
- Diet problem
- Least absolute deviations regressions

Week 04: Convex optimization, II

Sources and tools:

- Chapters 4-5 of Stephen Boyd's text.
- `CVXpy`
- `scipy.optimize`

Topics to be mastered:

- Quadratic programming
- Nonlinear programming
- Nonlinear regression

Week 05: Probability, I

Sources and tools:

- **Probability and Information**
- **Statistical Rethinking**
- **All of Statistics**

Topics to be mastered:

- Laws of probability
- Continuous and discrete probability distributions
- Conditional and marginal probabilities
- Stochastic processes
- Stationarity
- Ergodicity
- Concepts of statistical convergence
- Strong laws of large numbers
- Central limit theorems
- A *statistical model* as a probability distribution over a sequence indexed by a vector of parameters

Week 06: Probability, II

Sources and tools:

- https://python.quantecon.org/finite_markov.html
- https://python.quantecon.org/ar1_processes.html
- https://python.quantecon.org/stationary_densities.html

Topics to be mastered:

- The *direct problem* given parameters
  - Computation
  - Random simulations of paths
- The *inverse problem* given data
  - Likelihood function
  - Prior over parameters
  - Maximum likelihood and Bayesian estimation
- Time series models
- Discrete state Markov chains
- Continuous state Markov processes
- First order autoregressions
- Moving average processes
- Mixed autoregressive, moving average processes
- Nonlinear models
- Illustrations of stationarity and ergodicity

Week 07: Statistics, I

Sources and tools:

- **Statistical Rethinking**
- https://python.quantecon.org/heavy_tails.html
- https://python.quantecon.org/arma.html

Topics to be mastered:

- Linear regression

- Population and sample regressions
- Nonlinear population and sample regressions
- Mathematical expectations versus linear least squares projections
- Gram-Schmidt process as recursive projection
- Link of Gram-Schmidt to efficient computation and interpretation

Week 08: Statistics, II

Sources and tools:

- **Statistical Rethinking**
- https://python.quantecon.org/mle.html
- https://python.quantecon.org/exchangeable.html
- quantecon Jupyter notebook on estimating a first-order a.r.

Topics to be mastered:

- Bayesian parameter estimation
- Bayesian model evaluation
- Frequentist inference
- Frequentist model evaluation
- Learning versus hypothesis testing
- Exchangeability and de Finetti theorem
- David Kreps's famous "chapter 11" story about Totrep also known as "Totally Rational Economic Person"

Week 09: Statistics, III

Sources and tools:

- **Statistical Rethinking**
- https://python.quantecon.org/wald_friedman.html
- https://python.quantecon.org/odu.html

Topics to be mastered:

- Applications of Bayesian updating to decision making
- Job search with learning
- The classic problem that stumped Milton Friedman
- How Abraham Wald solved Milton Friedman's problem

Week 10: Statistics, IV

Sources and tools:

- **Statistical Rethinking**
- **All of Statistics**
- **Linear Models of Dynamic Stochastic Economies**

Topics to be mastered:

- Linear rational expectations models
- Simulation
- Inference
- Connections between simulation and inference (two sides of one coin)

Week 11: Dynamic Equilibrium Models, I

Sources and tools:

- https://python.quantecon.org/cattle_cycles.html
- https://python.quantecon.org/rosen_schooling_model.html

Topics to be mastered:

- Cattle Cycles
- The Rosen schooling model

Week 12: Case studies, I

Sources and tools:

- https://python.quantecon.org/black_litterman.html
- Cedric Villani on optimal transport https://www.youtube.com/watch?v=zo46TEp6FB8l (start at minute 28)

Topics to be mastered:

- Portfolio theory and its challenges
- The Black Litterman model and its relationship to Bayesian statistics and robust control theory
- Monge-Kantorovich transport problem and applications in social sciences

Week 13: Case studies, II

Sources and tools:

- `pgmpy`: https://www.youtube.com/watch?v=DEHqIxX1Kq4

Topics to be mastered:

- Using Bayes to infer who wrote the Federalist papers?
- Probabilistic graphic models with the Python package `pgmpy`

Week 14: Case studies, III

Sources and tools:

- https://link.springer.com/chapter/10.1007/978-94-011-2410-2_19

Topics to be mastered:

- From political science: the Law of Cubic Proportions
- Student presentations of class projects

Week 15: Wrapping up and looking forward

Sources and tools:

- Lectures notes on duality between filtering and control

Topics to be mastered:

- Useful math and going forward
  - Why "forecasting" and "decision" are related mathematically
  - How this can be exploited
- Student presentations of class projects

Course Policies

Professional Behavior

Attend class. They say "eighty percent of success is just showing up." We have found that those who show up perform systematically better.

Arrive to class on time and stay until the end of class. Chronically arriving late or leaving class early is unprofessional and disruptive to the rest of the class.

We understand that the electronic recording of notes will be important for class and so computers will be allowed in class. Please refrain from using computers for anything but activities related to the class. Phones are prohibited as they are rarely useful for anything in the course. Eating and drinking are allowed in class but please refrain from it affecting the course. Try not to eat your lunch in class as the classes are typically active.