Mathematical Methods for Structural Biology

Julie C. Mitchell and Shravan Sukumar

2:25PM MWF, 1116 Biochemistry

Description

The course gives a broad survey of molecular modeling along with numerical and computational tools for practical applications. Given the diverse nature of the students taking the course, and the range of topics from elementary to advanced, most will find both known and new topics among the course materials. The ultimate goal of the course is to expose students to practical working knowledge in the field of computational structural biology, with which they will undertake readings in the past and current literature and complete a project of their choosing.

Prerequisites

This course is intended for advanced undergraduate students or beginning graduate students from mathematics, biochemistry and biophysics. The minimal prerequisite is mathematical background through vector calculus, though additional math background is helpful. Those enrolling in the math version of the course should have some CS background, whereas this is waived for those enrolling in the biochem version. I will do my best to ensure that everyone finds the material both accessible and challenging. Note that the university is moving toward requiring different assessment of undergraduate and graduate student in 600-level courses. For this reason, graduate students must give an oral presentation in Biochem 906 during the last few weeks of class.

Grading

Homework	30%	About 6-7 assignments
Quizzes	30%	About one quiz/chapter
Final Project	40%	Class project writeup
Presentation		Mandatory for grad students

Textbooks

Julie C Mitchell, *Mathematical Methods for Structural Biology*. These free course notes will be distributed via the Learn@UW website. I wrote the notes by copying things off Wikipedia during my sabbatical. Corrections are appreciated.

Course Project

The course project will consist of a computational study on which a short report will be written. Graduate students need to give an oral report on their project during the last three weeks of the semester. Students enrolled for Math 606 or BMI 606 should do a project in which some programming work is performed. Students enrolled in Bioc 606 or BMC 606 can do a project in which they apply existing software toward a modeling project. Students are encouraged to choose projects that are related to their thesis work or future career goals. Students who have neither thesis work nor career goals are welcome to consult with me for ideas.

Homework and Quizzes

There will be several quizzes on the lecture and reading material, which will be online. There will also be 5-6 homework assignments consisting of modeling exercises using either Python or popular molecular modeling software. You will need a computer and install Python, Numerical Python, and Scientific Python on it. The easiest way to do this is to install the Enthought Canopy Python distribution, which is free for students. You can download it here:

https://store.enthought.com/#canopy-academic

If you need to turn in homework late, you can have a one week extension, but it costs you 50% credit on the assignment. Note that early homework is always gladly accepted and is generally the better way to deal with coinciding deadlines.

In Class Policy

I enjoy teaching as a function of student interest, so paying attention and asking questions are appreciated. Your attention in class, including not talking, eating loud snack foods, or banging on keyboards, is not only appreciated but expected. You will be asked nicely not to do this the first time, asked not so nicely the second time, the third time you will be asked to leave class.

Contact Information

Below is contact information for the professor and TA for the course. Our place is located on the southwest corner of Henry Mall, and if you come up to the top of the stairs and go left down the hall, you'll find us.

Julie Mitchell (`jcmitchell@wisc.edu`)
2104D Biochemistry
Office Hours: TBD

Shravan Sukumar (`sukumar@wisc.edu`)
2104 Biochemistry
Office Hours: TBD

Course Topics

The course covers a large range of topics from many STEM fields in a fairly accessible manner. Occasionally, a student has commented that they wanted greater depth in some subject, but the problem is that this is different for each student, and thus I'd get a different comment if I changed the format saying I didn't cover the topics that student was interested in. Moreover, I am not taking equations out of the course no matter how many times students write this on the evaluations, because you should buck up and learn hard things if you want to be well-educated in science.

Protein/DNA/RNA Biochemistry/Biophysics

- Structure (primary/secondary/tertiary structure; atoms, bonds, backbones, torsion angles)
- Amino and Nucleic Acids (hydrophobicity, polarity, charge, aromatic rings)
- Force Fields (charge, radii, electrostatics, van der Waals, stacking, hydrogen bonds)
- Structure Determination (Crystallography, NMR)
- Structure Modeling (Folding, docking, dynamics, mutagenesis effects)

Applied Math

- Linear algebra, Vector Geometry, Rotation Groups
- Complex Numbers, Fourier Analysis
- Optimization and Sampling
- Ordinary and Partial Differential Equations
- Numerical Methods

Computing

- Python
- Machine Learning
- High Performance and GPU Computing

Modeling Tools and Resources

- Protein Databases: BLAST/NCBI, Protein Data Bank, SCOP, OrthoMCL
- Structure: PyMol, ConSurf, Modeller, TMalign, DALI, DSSP
- Macromolecular Docking/Interfaces: ZDOCK, PatchDock, KFC2
- Molecular Dynamics: OpenMM (CHARMM, AMBER, GROMACS FFs)
- Electrostatics: PDB2PQR, APBS

Week 1. Protein Structure Basics
Vectors and Matrices
Computing: PDB, PyMol, BLAST

Week 2. Bond Lengths, Bond Angles, Protein Backbones, Ramachandran Plots
Atomic charges, radii, bond lengths/angles/torsions
Computing: PDB2PQR, Molprobity

Week 3. Clustering and Machine Learning

Week 4. Complex numbers, Gaussian functions, Fourier series and transforms
Crystallography, X-ray diffraction, electron density and structure refinement
Nuclear Magnetic Resonance and spectra

Week 5. Protein Folding, Docking and Mutagenesis
Computing: SWISS-MODEL, Modeller, TMalign, SCOP, DALI, ZDOCK, HADDOCK, KFC, Consurf

Week 6. Linear and Quadratic Approximation, Jacobians, Hessians
Quadratic Approximation, Normal Modes and Protein Motion
Computing: elNemo

Week 7. Numerical basics, norms, error, convergence, least squares, Gauss-Newton, Levenberg-Marquardt
Gradient descent, Conjugate Gradient, BFGS, Lagrange multipliers

Week 8. Monte Carlo, Simulated Annealing
Optimization-based Approaches to Protein Folding and Docking

Week 9. Ordinary differential equations
Linear Multistep Methods

Week 10. Molecular Dynamics Simulation
Computing: NAMD, Amber, CHARMM, OpenMM

Week 11. Numerical PDE's, finite differences, finite elements, multi grid

Week 12. The Poisson-Boltzmann Equation and Protein Electrostatics
Computing: APBS, CHARMM-GUI

Week 13. Designing High Performance Algorithms, Parallel and GPU Computing

Week 14. Students will give brief presentations on their course projects

Week 15. Students will give brief presentations on their course projects