CAS MA 586 / CDS DS 522

Stochastic Methods for Algorithms

Table of Contents

Overview

Instructor	Dr. Jonathan Huggins (huggins@bu.edu)
Office	CDS 427
Drop-in Hours	Wednesday 12:30–2:00pm and Thursday 10:30–11:30am (CDS 427)
TF	Andrew Roberts (arober@bu.edu)
Drop-in Hours	Tuesday 4:00–5:00pm (CDS 1326)
Meetings	Tuesday & Thursday 2–3:15pm (MCS B37) Friday 9:05–9:55am (CAS 222) or 10:10–11:00am (IEC B12)
Textbook	*Stochastic Methods for Algorithms* (available on Perusall)
Course Content	Blackboard and Perusall
Communication	Perusall
Other software	Python, LaTeX editor (e.g., Overleaf, TeXShop, TeXstudio, or LyX)
Hub Units	Creativity/Innovation (CRI) and Writing-Intensive Course (WIN)
Prerequisites	See below

Welcome / About This Course

“There is nothing more practical than a good theory.” – *Kurt Lewin*

This course concerns the use of stochastic processes for designing and analyzing algorithms, with a focus on applications in statistics and machine learning. You will learn about core concepts and results about stochastic processes, then use this machinery to (1) develop algorithms and (2) characterize their statistical and numerical properties. The two recurring applications will concern (large-scale) optimization using stochastic gradients and sampling from complex distributions such as Bayesian posterior distributions and energy-based models using Markov chain Monte Carlo algorithms. In addition to statistics and machine learning, the optimization and sampling algorithms derived and analyzed have diverse applications including to problems in computer science, physics, chemistry, ecology, biology, and operations research. As such, a strong emphasis is placed on the ability to describe and investigate the practical implications of the results.

Considering the application-oriented motivations, the stochastic processes material focuses on intuitive understanding of definitions and theoretical properties rather than their rigorous development. This approach will allow us to almost immediately begin to investigate algorithms and their properties. Toward the same ends, we will mostly forgo using advanced tools probability theory and functional analysis. Thus, we will tend not to dwell on regularity conditions and other (still important!) technicalities. However, throughout the course I will provide pointers to sources where rigorous treatments can be found.

When describing the practical implications of algorithm analyses and empirical results aimed at validating such analyses, clear communication is a must. Thus, you will do a significant amount of expository writing in this course. These writing opportunities will enable you to learn how to synthesize complex mathematics and sophisticated experiments in a clear way that can be understood by someone who is familiar with the underlying algorithms but not the mathematical theory.

Learning Outcomes

After successful completion of this course, you will be able to…

1. Explain how to use the theory of Markov chains and stochastic differential equations to analyze Markov chain Monte Carlo and stochastic optimization algorithms.
2. Evaluate the implications of theoretical analyses of Markov chain Monte Carlo and stochastic optimization algorithms, including law of large numbers, geometric ergodicity, central limit theorems, error analyses, and scaling limits.
3. Interpret results and arguments in the modern statistics and machine learning literature about the design and analysis of Markov chain Monte Carlo and stochastic optimization algorithms.

Prerequisites

The formal prerequisites are (recommended versions marked with a *):

- **Writing** *[undergraduate only]***:** First-Year Writing Seminar (e.g., WR 120)
- **Programming:** *CAS CS111, *CDS DS110, *ENG EK125, or equivalent.
- **Vector calculus:** *CAS MA225, *CAS CS235, *CDS DS122, or equivalent
- **Linear algebra:** *CAS MA242, *CAS CS132, *CDS DS121, or equivalent
- **Probability theory:** *CAS MA581, CAS CS237, ENG EK381, *ENG EK500, or equivalent.

However, there are some additional "soft" requirements/recommendations that do not correspond to a specific course:

- You must **have some experience writing scientific code**, preferably in Python (which will be used for all course assignments). For example, you may have taken CAS MA415, CAS MA615, CDS DS210, or a more advanced CS programming course.
- You must **have experience writing rigorous mathematical proofs**, with proofs by induction being particularly important. For example, you may have taken one or more 500-level proof-based math courses, or a course on analysis algorithms, or a computer science theory course.

- Having some **previous exposure to stochastic process theory** will be invaluable but is not strictly required. For example, you may have taken CAS MA583, CAS MA783, or ENG EC505.
- Having some **previous exposure to statistics and/or machine learning** – particularly regression, statistical models, Markov chain Monte Carlo, and/or (stochastic) optimization – is very helpful but not required.

Communication

Course content will primarily be on Perusall. Blackboard will be used for homework submission and recording grades.

All communication should be done via Perusall. I will typically respond within 24 hours during the week. However, I have three kids, so my weekends are very busy with family commitments. Therefore, <u>you should not expect me to be responsive between approximately 5pm Friday and 9am Monday.</u>

I encourage you to come to my drop-in hours sometime in the first month of class, as I would very much like to get to know you all. As the semester gets going, you can come to drop-in hours to discuss the class material – or math, statistics, and machine learning in general. If you can't make it during my regularly scheduled hours or if you need to talk to me privately, please email me and I will do my best to schedule a time for us to meet.

Textbook

The primary textbook will be the provided course notes. However, the following may be useful supplements with more detailed treatments of certain topics:

- Probability theory roughly at the same level as this course ("just short of measure theory"): *Probability and Random Processes* by Grimmett & Stirzaker

- Background on (probabilistic) models and algorithms: *Machine Learning: A Probabilistic Approach* by Murphy[1]
- Mathematical background: *Mathematics for Machine Learning* by Deisenroth, Faisal & Ong[2]
- Detailed but accessible treatment of stochastic differential equations: *Applied Stochastic Differential Equations* by Särkkä & Solin[3]

Class Meetings

BEFORE class: You'll complete **pre-class assignments** in which you'll read parts of the text and either answer some comprehension-focused questions or ask a question you have about the reading.

DURING class: We will discuss the reading in depth, focusing on the big picture takeaways and the most challenging technical materials such as proofs. **There are no timed tests or quizzes during class meetings**.

[1] Available at https://probml.github.io/pml-book/book0.html

[2] Available at https://mml-book.github.io

[3] Available at https://users.aalto.fi/~asolin/sde-book/sde-book.pdf

Equity/Inclusion

As your professor, I pledge to work to create an equitable learning environment where all students belong. Statistics can seem like an "objective" subject, but like all education, mathematical education is a cultural activity, and many aspects of who we are affect how we experience statistics classes. There is considerable research on how students' "sense of belonging" in their classes impacts their learning. Feeling like you don't belong in a class can impact your cognitive load and diminish your ability to focus on the mathematical and statistical content; so, for example, you might use your brain to worry whether the professor or your classmates will think you're stupid if you make a mistake; to deal with the impact of a racist, sexist, classist, xenophobic, homophobic, transphobic, or ableist comments; or to wonder whether the other students are better than you.

Although frustration and struggle are also part of the educational enterprise, ultimately, I see my work as setting up situations where all students can experience the joy of mathematics and statistics, to feel a sense of belonging, and to use our brains for learning.

Assignments

There will be four types of assignments:

1. **Pre-class Assignments.** In addition to readings, there will often be exercises and questions to be submitted prior to class. These will help you learn the basics, reflect on new material, and prepare you for lectures and discussions during class time. There are 20 of these planned.
2. **Exit Tickets.** At the end of each week, you will complete an "exit ticket" survey through Blackboard that will ask you questions about how the past week went (What did you learn? What are you confused about?). These exit tickets facilitate meta-cognition (e.g., ask yourself questions like "Which topics/concepts do I understand?" and "Which ones am I having more

difficulty with?”). They also help me to understand which topics or concepts I haven’t done a good enough job teaching you. There are 12 of these planned.

3. **Problem Sets.** More challenging problems with a mix of derivations, proofs, and coding problems. These can be done in collaboration with other students. There are 7 of these planned.
4. **Mini Project.** This will be more cohesive and open-ended than problem sets. It will feature a mix of scientific programming, mathematical derivations, and some synthesis that provides a chance to develop writing skills. There will be one of these.
5. **Article Evaluations.** These will involve reading a published paper, then evaluating the writing and scientific content following a provided rubric. There are two of these planned.
6. **Final Project.** The final project will involve further investigation of a topic closely related to what is covered in the course. After critiquing a paper in your second Article Evaluation, you will make a novel contribution that is motivated by your critique. The final project is similar in scope to the Mini Project.

There is no final exam in this course.

Grades

Philosophy

Your grade in the course is earned by **demonstrating evidence of skill on the main concepts in the course** and by **showing appropriate engagement with the course**. This is done by completing the assignments outlined above, at a reasonably high level of quality. The class should be a learning community, where students support each other to increase everyone’s learning. You should not be competing with other students for grades. You should have some choice in your trajectory through the course.

Therefore, in this course, **there are no points or percentages** on any items. Instead, the work you turn in will be evaluated against **quality standards** that will be made clear on each assignment. If your work meets the standard, then you will receive full credit for it. Otherwise, you will get helpful feedback and, on most items, the chance to reflect on the feedback, revise your work, and then resubmit it for regrading.

This feedback loop represents and supports the way that people learn: By trying things, making mistakes, reflecting on those mistakes, and then trying again. **You can make mistakes without penalty** if you *eventually* demonstrate evidence of skill.

Determining Your Course Grade

The individual kinds of assignments are marked as follows:

Assignment	How it's marked
Pre-class Assignments	**Pass** or **No Pass**
Exit Tickets	**Pass** or **No Pass**
Problem Sets	**Pass** or **Not Yet**
Mini Project	**Pass**, **Almost Pass**, **Partial Pass**, or **Not Yet**
Article Evaluations	**Pass**, **Almost Pass**, **Partial Pass**, or **Not Yet**
Project	**Pass**, **Almost Pass**, **No Pass**

The criteria for each mark are explained below in the "Grading Criteria" section below.

Your final grade in the course is determined by the following table. Each grade has a *requirement* specified in its column in the table. **To earn a grade, you will need to meet *all* the requirements in the column for that grade.** Put differently, your grade is the **highest** grade level for which **all** the requirements in a column of the table have been met or exceeded.

	A	B	C	D
Pre-class Assignments + Exit Tickets passed (32)	28	22	16	8
Problem Sets passed (7)	6	5	4	2
Mini Project mark	Pass	Almost Pass	Partial Pass	Partial Pass
Article Evaluation minimal marks (2)	2 Passes	1 Pass, 1 Almost Pass	1 Pass	1 Almost Pass
Final Project mark	Pass	Almost Pass	N/A	N/A

A grade of **F** is given if all the requirements for a **D** are not met.

Plus/minus grades: Plus/minus grades will be assigned at my discretion based on how close you are to the next higher grade level.

Grading Criteria

- **Pre-class Assignments** and **Exit Tickets:** A **Pass** mark is given if it is turned in before its deadline and if each item has a response that represents a good faith effort to be right. Mistakes are not penalized. A **No Pass** is given if an item is left blank (even accidentally), has an answer but it shows insufficient effort (including responses like "I don't know"), or if it is turned in late.
- **Problem Sets**, **Mini Projects**, **Article Evaluations**, and **Final Project**: Each type has its own standards which will be included with the assignment.

Revising and Resubmitting Work

Instead of earning partial credit, on most assignments you will have the opportunity to revise and resubmit your work based on feedback that I or the TF provide, if the work doesn't meet its standard for acceptability. **Mistakes, and work that does not meet the standard for**

acceptability, are typically not penalized. Instead, if your work has enough room for improvement that it would benefit from redoing parts of it or the whole thing, you'll get the chance to do so. This again is because **human beings learn from making mistakes and fixing them with feedback and reflection.**

- **Pre-class Assignments** and **Exit Tickets** may not be revised or resubmitted. They are graded on completeness and effort only, and therefore can only be done once.
- Each **Problem Set** may be revised **twice**. One problem set revision may be submitted each week, where the week begins at midnight on Monday and ends at 11:59pm on Sunday. To revise, simply reflect on the feedback that's given, make corrections or rewrites to the original, and upload the work again to Blackboard, then submit a regrade request using this form.
- The **Mini Project** and each **Article Evaluation** may be revised **once**. One such revision may be submitted each week, where the week begins at midnight on Monday and ends at 11:59pm on Sunday. To revise, simply reflect on the feedback that's given, make corrections or rewrites to the original, upload the work again to Blackboard, then submit a regrade request using this form.
- **Final Project:** The project may not be revised, but you will have the chance to get feedback on a draft and some partial credit.

Tokens

We all have things – good and bad – that come up in our lives that affect our ability to get work done on time. So, at the beginning of the semester, you will receive 5 “tokens,” which you can use to break course rules in prescribed ways:

- You may spend **1 token** to convert a **No Pass** on a Pre-class Assignment or Exit Ticket to a **Pass**.
- You may spend **1 token** to submit an additional Problem Set for regrading in a week.
- You may spend **1 token** to receive a 5-day extension on an assignment with a flexible deadline. Assignments with flexible deadlines will be explicitly labeled as such.
- You may spend **2 tokens** to submit a Mini Project or Article Evaluation for a second regrade. However, note that you can still only submit one regrade per week.

In general, I will offer a 24-hour grace period for late submissions (excluding pre-class assignments and others explicitly labeled as not having a flexible deadline).

Use of Generative AI

I will follow the Faculty of Computing & Data Science's **Generative AI Assistance (GAIA) Policy**, which can be found in full at https://www.bu.edu/cds-faculty/culture-community/gaia-policy/. You must read the GAIA policy if you plan to use Generative AI tools. The intent of the policy is as follows:

> *Students should learn how to use AI text generators and other AI-based assistive resources (collectively, AI tools) to enhance rather than damage their developing abilities as writers, coders, communicators, and thinkers. Instructors should ensure fair grading for both those who do and do not use AI tools. The GAIA policy stresses transparency, fairness, and honoring relevant stakeholders such as students eager to learn and build careers, families who send students to the university, professors who are charged with teaching vital skills, the university that has a responsibility to attest to student competency with diplomas, future employers who invest in student because of their abilities and character, and colleagues who lack privileged access to valuable resources. To that end, the GAIA policy adopts a few commonsense limitations on an otherwise embracing approach to AI tools.*

If you have questions or aren't sure if something is allowed, ask me!

Collaboration and Academic Honesty

I strongly encourage you to collaborate with your classmates whenever it is allowed. However, realize that collaboration is not always allowed, and, in all cases, there are limitations on how you can collaborate. In particular:

- On **Exit Tickets**, **Article Evaluations**, and the **Final Project**, your work must represent your own understanding in your own words using your own code. You may not use solutions, directly or indirectly, from any sources not explicitly allowed – including other students, past students, online sources, or other textbooks.
- On **all other assignments**, you may collaborate with others, but you must contribute significantly to the assignment, and your work must represent your own understanding in your own words and using your own code.

You are responsible for understanding this policy and [Boston University's Academic Conduct Code]. Violations will result, at minimum, in a mark of **No Pass / Not Yet** on the assignment with no chance to resubmit. Serious or repeat violations of this policy will result in increasingly unfortunate consequences, including being barred from further submissions of the assignment, or even receiving an **F** in the course.

Hub Learning Outcomes

Creativity/Innovation

Students will earn a Creativity/Innovation credit by satisfy the learning outcomes as follows:

1. *Students will demonstrate understanding of creativity as a learnable, iterative process of imagining new possibilities that involves risk-taking, use of multiple strategies, and*

reconceiving in response to feedback, and will be able to identify individual and institutional factors that promote and inhibit creativity.

2. *Students will be able to exercise their own potential for engaging in creative activity by conceiving and executing original work either alone or as part of a team.*
   Students will demonstrate creativity in structuring and creating original mathematical proofs and implementing numerical experiments to validate their results. Students will improve their solutions and results on assignments through an iterative process of peer and instructor feedback. Students will be asked to show increasing degrees of creativity in the progression from In-class Exercises to Problem Sets, Article Evaluations, and Mini Projects to the Final Project. Overall, students will exercise their own potential for creativity in all assignment types, which will include both individual and small-group work.

Writing-Intensive Course

Students will earn a Writing-Intensive Course credit by satisfy the learning outcomes as follows:

1. *Students will be able to craft responsible, considered, and well-structured written arguments, using media and modes of expression appropriate to the situation.*
   Students will write formal mathematical proofs and expository explanations of mathematical and experimental results through In-class Exercises, Problem Sets (4–8 pages total), the Mini Project (10–12 pages total), Article Evaluations (8–10 pages total), and the Final Project (10–15 pages). They will improve the clarity and style of their writing through peer and instructor feedback, which they will incorporate into revisions that they will resubmit. The course grade will be almost completely determined by student performance on these assignment types.

2. *Students will be able to read with understanding, engagement, appreciation, and critical judgment.*

   By completing structured Problem Sets and homework readings from the primary literature, students will learn how to *(i)* read, understand, and deconstruct complex mathematical proofs and *(ii)* analyze and critically judge numerical experiments designed to support mathematical theory of algorithms. Readings will also be discussed during class in small groups and with the whole class. Students will also learn how to read and respond to each other's writing in ways that are beneficial and useful. As a class and in smaller groups, including instructors and students, students will learn how to give and respond to feedback on assignments.

3. *Students will be able to write clearly and coherently in a range of genres and styles, integrating graphic and multimedia elements as appropriate.*

   N/A

Tentative Schedule

The course will consist of 7 modules:

0. **Preliminaries (Weeks 1–2).** Course overview, background material review, high-level overview of stochastic optimization and Markov chain Monte Carlo.
1. **Probability theory and Markov chains (Weeks 2–4).** Review of introductory probability theory and extension to a level just short of measure theory. Basics of Markov chains, including manipulating probability kernels and finding stationary distributions.
2. **Markov chains for stochastic optimization (Weeks 5–6).** Convex analysis, Taylor series error, and the application of Markov chain theory to analyzing the error and convergence of SGD.

3. **Markov chains for sampling (Weeks 7–8).** Application of Markov chain theory to designing Markov chain Monte Carlo algorithms and analyzing their convergence properties.
4. **Stochastic differential equations (Weeks 8–10).** Basics of SDEs including Itô calculus and generators. Application of Langevin diffusions and scaling limits to design and analyze SGD and MCMC algorithms.
5. **More! (Weeks 10–11).** You will have four options for topics (chosen to align with your project):
    a. Markov chain LLN and geometric ergodicity
    b. Geometric ergodicity and MCMC central limit theorems
    c. Polyak–Ruppert averaging and Markov chain CLTs
    d. Analysis of the unadjusted Langevin algorithm and scaling limits
6. **Final project (Weeks 12–13).** Paper critiquing and expository scientific writing. Peer feedback

I reserve right to make changes to the assessment system and to other aspects of the syllabus to better meet the needs of students in the class. If appropriate, students might have input into these changes. Any changes will be clearly documented with sufficient notice for students to adapt.

Community of Learning: Additional Class and University Policies

Accommodations for Students with Documented Disabilities

If you are a student with a disability or believe you might have a disability that requires accommodations, please contact the Office for Disability Services (ODS) at (617) 353-3658 or access@bu.edu to coordinate any reasonable accommodation requests. ODS is located at 19 Buick Street.

Attendance & Absences

Attendance is critical to your success in this course, and if you want to do well, you need to be present and prepared. In this class you'll be part of a learning community, and we will miss you when you aren't here. However, there are many reasons why people need to miss labs and discussions, such as illness, religious holidays, family emergencies and milestones, civic responsibilities (jury duty, citizenship ceremonies, etc.), dangerous commutes during bad weather, etc. If you know in advance that you're going to miss a lab or discussion section, please let the lab instructor or TF know. If you do miss a lab or discussion section, it's your responsibility to find out what you missed.

If there's something in your life that's interfering with your ability to engage in the course, please come talk to me about it (you only need to share the details you want to share). Note that you don't need to bring me doctor's notes; I don't read them. Also note that I have considerable experience with chronic illness and how it impacts academic life.

Incompletes

If you have health issues, an emergency, or find yourself in other difficult circumstances that affect your performance in the course, you may be eligible for an incomplete, where you would finish the work after the semester ends. Please feel free to talk to me about this possibility.

Acknowledgments

Parts of this syllabus are borrowed from / based on the syllabi of Debra Borkovitz (BU, CAS MA 293) and Robert Talbert (GVSU, MTH 350).