Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
6.438 Algorithms for Inference
Fall 2020

Information Sheet

Lecturers

Devavrat Shah	Caroline Uhler
Office: 32-D670	Office: 32-D634
Tel: x3-4670	Tel: x3-4181
E-mail: devavrat@mit.edu	E-mail: cuhler@mit.edu

Administrative Assistant
Kimberly Strampel
Office: E17-462B
Tel: x4-4934
E-mail: strampel@mit.edu

Teaching Assistants

Eshaan Nichani	E-mail: eshnich@mit.edu
Abhin Shah	E-mail: abhin@mit.edu
Cindy Yang	E-mail: cxy99@mit.edu
Chulhee 'Charlie' Yun	E-mail: chulheey@mit.edu

Lectures:	Tuesday and Thursday, 9:30am–11am, Zoom + Recorded
Recitations:	Friday, 10–11am or 11am–noon, Zoom + Recorded
Office Hours:	Days, times on stellar site, Zoom

Welcome to 6.438 !

This is a graduate-level introduction to the principles of statistical inference with probabilistic models defined using graphical representations. As such, it is a core graduate subject for students in the relevant subfields both within and beyond Course 6 (EECS). The material in this course constitutes a common foundation for work in, for example, machine learning and data science, artificial intelligence, signal processing, computer vision and computational imaging, communication, bioinformatics and neuroscience, finance and data analytics, and networks, robotics, and control. Moreover, it is the companion course to 6.437 (Inference and Information) which is offered each Spring; 6.437 and 6.438 may be taken in either order.

It is worth stressing that 6.438 is an *introductory* graduate subject: it is not an advanced graduate subject for students who already have a mastery of statistical inference algorithms, yet want to understand such material at an even more sophisticated level. That said, the structure of this subject will be somewhat different than other such introductions to the topic.

Ultimately, the subject is about teaching you contemporary approaches to—and perspectives on—problems of statistical inference and learning. The development of the material that forms the basis for this subject has historically been very much driven by applications. However, our focus in the course will not be on these applications—which form the basis for entire courses of their own—but rather on the common problem solving frameworks that they share. Nevertheless, we will cite various relevant applications as we develop the material and sometimes extract simplified examples from these contexts.

Note that the course has both lectures and recitations, which are designed to complement each other. Recitations begin the first week of classes. There are two possible recitation times to choose from, as indicated above. Select to attend whichever suits your schedule best. In addition, there are office hours scheduled throughout the week, as indicated on the class web site. You are welcome and encouraged to come to any and all of them you think might be helpful to you in clarifying your understanding of the material.

Prerequisites

The official prerequisites are 6.008, 6.041, 6.436, or 18.600, and 18.06, or their equivalents. Ultimately, what we require is *fluency* with both basic quantitative probabilistic analysis and linear algebra, together with some subsequent solid exposure to the engineering application of both. When in doubt, students whose undergraduate degrees are not from MIT should consult the staff to determine if they have had subjects that are effectively equivalent to the official prerequisites.

Reading

We will provide lecture notes whose content is under active development. These notes will be the primary reading material. You will find some sections of these notes more mature than others as we develop and refine the treatment of different topics, and we will count on you to help us identify bugs!

Also, some of the notes (or sections thereof) will specifically be indicated as "optional" reading. These notes develop aspects of the material that we will not be covering this semester, but which you may enjoy reading about, depending on your interests and as your time permits. During the semester we will not assume you have read any of the optional notes or sections.

You may also find sections of the following books to be useful and more in-depth auxiliary references for parts of the term.

M. J. Wainwright and M. I. Jordan, *Graphical Models, Exponential Families, and Variational Inference*, Foundations and Trends in Machine Learning, 2008.

C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

M. I. Jordan, *Introduction to Probabilistic Graphical Models*, Lecture notes (on class web site).

S. Lauritzen, *Graphical Models*, Clarendon Press, 1996.

D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques,* MIT Press, 2009.

D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms,* Cambridge University Press, UK, 2003. (on class web site; also available online at
`http://www.inference.phy.cam.ac.uk/mackay/itila/book.html`)

M. Mezard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press, 2009.

We suggest you initially hold off purchasing any books while you are gauging their usefulness to you. These books are available in the MIT Barker Library and in online form. If you find it difficult to acquire copy of any of the text, please contact teaching staff.

If you are interested in further reading, either to strengthen your background, reinforce some of the concepts from lecture, or to probe some topics in more detail, take a look at the additional references on the course web site. In particular, you'll find several papers containing a variety of useful insights, which are worth the effort to work through.

Finally, we kindly ask that you restrict course materials, including supplementary reading, for your personal use only, as we have specifically been asked to not share or otherwise distribute any portions of them to others outside the subject.

Problem Sets

There will be 4 problem sets (homework). They will be due at the time of lectures as indicted in the detailed class schedule. Solutions will be made available by the end of the day. While you should do all the assigned problems, only a randomly chosen subset will actually be graded. Don't be misled by the relatively few points assigned to homework grades in the final grade calculation! While the grade you get on your homework is only a minor component of your final grade, working through (and, yes, often struggling with at length!) the homework is a crucial part of the learning process and will invariably have a major impact on your understanding of the material.

In undertaking the problem sets, moderate collaboration in the form of joint problem solving with one or two classmates is permitted provided your writeup is your own and reflects your own understanding of the problem.

Computational Problems

Some of the questions in the problem sets / homeworks will involve a computational component, to help you explore additional aspects of the material. Officially, we will support PYTHON, but you're welcome to use e.g. MATLAB as long as you don't require assistance from the staff in converting any PYTHON-based code or data that may be provided. These problems will be graded separately from the rest of the problem set that they appear on. As with the other problems on the homework, moderate collaboration with one or two classmates is permitted provided the code and writeup you produce is your own and refelcts your own understanding of the problem.

Exams

There will be two quizzes in the subject. The first quiz will be held Thursday, October 15. The second quiz will be held on Thursday, December 3.

Course Grade

The final grade in the course is based upon our best assessment of your understanding of the material during the semester. Roughly, the weights used in grade assignment will be:

Quiz 1	40%
Quiz 2	40%
Problem Sets & Computational Problems	20%

Course Web Site and Email

We will make announcements via email, and we will post various information and handouts on the course web site.

You should first make sure that you have an active Athena account (by visiting `http://ist.mit.edu/support/accounts` if necessary) as well as a personal certificate (by visiting `https://ca.mit.edu/ca/` if necessary). If you have problems or if you are not a regular MIT student, please contact teaching staff for assistance.

The course web site is

`http://web.mit.edu/6.438`

You will need to have a valid certificate *and* be on the official course list to access the web site. If you have pre-registered for 6.438, this should already be set up;

just double-check that you can access the web site (try to download a handout, for example). Otherwise, contact one of the TAs and they will add you to the list. If you can access content on the web site, you should also be receiving all of the course announcements.

Additionally, we will use PIAZZA to post class announcements and as an on-line discussion forum. Through piazza you can post questions or comments (anonymously or otherwise) about any aspects of the material, which your peers and the staff can respond to. It has become a popular resource in the course for clarifying potential points of confusion in lecture notes, problem sets, recitation handouts, project descriptions, etc. You can access the piazza forum through the link provided on the 6.438 web site. Once you enroll yourself, you can fully participate in the discussions.

Finally, if you have any questions during the term, you can reach us by sending email to

`6.438-staff@mit.edu`

Tentative Syllabus and Schedule

HW	Inst'r		Lectures		
		Day	Date	#	Lecture Material
	Uhler	T	9/01	1	Introduction and overview
out: 1	Uhler	R	9/03	2	Directed graphical models
	Uhler	T	9/08	3	Undirected graphical models
	Uhler	R	9/10	4	Other graphical models and their relation
	Uhler	T	9/15	5	Gaussian graphical models
	Uhler	R	9/17	6	Learning: Parameter estimation
due: 1, out: 2	Uhler	T	9/22	7	Exponential families
	Uhler	R	9/24	8	Exponential families, parameter estimation
	Uhler	T	9/29	9	MTP2 exp. families (Markov prop., parameter est.)
	Shah	R	10/1	10	Marginalization and elimination algorithm
		F	10/2		ADD DATE
	Shah	T	10/6	11	Elimination, MAP and do-operation
due: 2, out: 3	Shah	R	10/08	12	Trees and Belief Propagation
		T	10/13		NO CLASS - Monday schedule
		R	10/15	Q	Quiz I (through L11, HW2)
	Shah	T	10/20	13	Gaussian BP and Kalman filter
	Shah	R	10/22	14	Junction tree and Belief Propagation
	Shah	T	10/27	15	Variational inference, mean field
	Shah	R	10/29	16	Variational inference, tree re-weighted
due: 3, out: 4	Shah	T	11/3	17	Variational inference, graph re-weighted
	Shah	R	11/05	18	Parameter estimation, EM
	Shah	T	11/10	19	Relationship between structure learning and inference
	Shah	R	11/12	20	Structure learning as regression
	Uhler	T	11/17	21	Causal structure discovery I
		W	11/18		DROP DATE
due: 4	Uhler	R	11/19	22	Causal structure discovery II
		T	11/24		NO CLASS (Thanksgiving Vacation)
		R	11/26		NO CLASS (Thanksgiving Vacation)
	Shah	T	12/01	23	Sampling and Inference I
			12/03	Q	Quiz 2 (through L21, HW4)
	Shah	T	12/08	24	Sampling and Inference II