AE740: Statistical Inference, Estimation, and Learning

Lectures: Tuesdays and Thursdays 1:30 PM – 3:00 PM
Lecture Room: 1008 FXB
Instructor: Alex Gorodetsky
Instructor Office: 3053 FXB
Office Hours: 10-11AM Tuesday/Thursday

Course Description

This course covers theory and algorithms for synthesizing models and data for general applications across science and engineering. Topics will include algorithms for maximum likelihood estimation, Bayesian inference, and regression for static inference problems and for estimation in dynamical systems. The first part of this course will focus on parameter inference and include recent advances in Monte Carlo sampling, importance sampling, Markov chain Monte Carlo (MCMC), and variational inference. The second part of the course will cover state estimation and include particle filters and variants of Kalman filters such as the extended, ensemble, unscented, and Gauss-Hermite filters. Theoretical foundations of the algorithms will be described and practical experience will be gained through projects that focus on implementation. PhD students will be encouraged to work on inference and estimation problems relevant to their ongoing research.

Topics Covered

The following topics will be discussed:

1. Introduction:
   - Probability, random variables, Gaussians
   - Concentration of measure inequalities (Markov, Chebyshev, ...)
   - Central Limit Theorem, Law of Large numbers
2. Sampling methods
   - Monte Carlo methods
   - Variance reduction methods: importance sampling, multilevel Monte Carlo, control variates
3. Inference and learning methodologies
   - Bayesian, frequentist, maximum entropy
   - Maximum likelihood estimation and regularization
   - Applications to linear-Gaussian and infinite dimensional models
   - Markov Chain Monte Carlo
   - Variational Inference methods and introduction to information theory concepts
4. Dynamical systems estimation
   - Particle filtering
   - Kalman filtering and smoothing (Extended and Ensemble Kalman filters)
   - Gaussian filters (Unscented, Cubature, Gauss-Hermite, etc.)
   - Joint parameter-state estimation

Learning Objectives

Following this course, students should be able to

1. Understand the methodology of modern techniques for fusing data and models together, and
2. Implement and design inference algorithms and identify their strengths and weakness.

Measurable Outcomes

Students successfully completing this subject will be able to:

1. Apply Monte Carlo methods and explain their convergence
2. Describe and apply standard approaches for variance reduction
3. Describe the Bayesian approach to inference and its relationship to regularization
4. Understand and assess the roles of the prior distribution in Bayesian computation
5. Applying Markov chain Monte Carlo for parameter inference both linear and nonlinear models
6. Understand the data assimilation problem as a problem of Bayesian inference
7. Explain and implement (Extended, Ensemble, Unscented) Kalman filter and particle filter algorithms
8. Design new algorithms for Gaussian filtering by leveraging ideas from numerical integration

Assignments and grading

Grading for this class will be primarily based on four large project-style assignments that are weighted equally. These projects will roughly be on the following topics

1. Monte Carlo simulation and variance reduction.
2. Bayesian inference with MCMC and regularization approaches.
3. State estimation and filtering.
4. Student's choice.

Proposals for the final project must be approved by the instructor; a list of papers and ideas can be provided upon request. While the project must investigate methodology, PhD students are encouraged to apply methods to problems within their own research. An oral presentation to the class will be required (depending on enrollment).

In addition to these projects students will be required to *scribe* at least one lecture. Two scribes will be assigned per lecture, so the total number of lectures a single student scribes will depend on enrollment. A template (in latex) will be provided. Scribing will be due **one week** after the lecture. Furthermore, a number of small homework assignments may be given throughout the term.

Table 1: Grade distribution

Component	Percentage
Projects	90
Scribing	10

Example projects from the past

In this course, I highly encourage applying the learned algorithms to problems that are of specific interest to the student because they are quite general and broadly applicable to problems throughout science and engineering. Sample projects done by students in the past have been

1. Learning provably safe paths through obstacles
2. Analyzing (social) network structures
3. Adaptive estimation and control for aircraft dynamics
4. Bayesian inference in neural networks
5. Particle filters for robot localization
6. and more. . . .

Prerequisites

This class will build upon the fundamentals of linear algebra, probability theory, and statistics. The course will consist both of theoretical (proofs) and applied components. Basic programming skills are essential, and any language can be used to complete the assignments.

References and resources

There is no single textbook for this subject. Lectures will draw from a variety of sources, including journal papers; specific references will be provided with the notes for each lecture. Below are some resources that you may find to be of general utility. Some of the books below have free copies available online, those that do not have free copies available have been placed on reserve at the library (Duderstadt).

1. Probability Theory and Inference
   - E T Jaynes, *Probability Theory: The Logic of Science* , Cambridge University Press, 2003 (Full book is available here).
   - 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, 2003, (Full book is available here).
   - J. S. Rosenthal, *A First Look at Rigorous Probability Theory*, World Scientific, 2006.
2. Methods and Algorithms
   - C. Robert, and G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004. (E-book available from UM library online)
   - T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2009, (Updated PDF of the book is free at the link).
   - D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford University Press, 2006. (E-book available from UM library online)
   - C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, MIT Press 2006 (Full book is available here).
   - A. Owen, *Monte Carlo theory, methods and examples*, 2013. Online book available here.
3. Estimation in Dynamical Systems
   - K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Springer, 2015, (A few free chapters here, E-book available form UM library online).
   - S. Särkkä, *Bayesian Filtering and Smoothing*, Cambrdige University Press, 2013, (Full book available here).

The books *Monte Carlo Statistical Methods* and *Data Assimilation: A Mathematical Introduction* are available to UM students for $25 as print-on-demand titles from Springer.