EECS 281B / STAT 241B
Advanced Topics in Statistical Learning Theory
Spring 2009

Practical information

- Lectures: Mon/Wed from 13:00–14:30 Location: Evans Hall 330
- Course web page: `http://www.eecs.berkeley.edu/~wainwrig/stat241b/`
  All announcements and homeworks will be posted at this site; please check it regularly.

- Instructor:
  - *Name:* Martin Wainwright
  - *Email:* wainwrig AT SYMBOL {stat,eecs} DOT berkeley DOT edu
  - *Offices:* 263 Cory Hall or 421 Evans Hall
  - *Office hours:* Monday 14:30–15:30, 421 Evans Hall
    Thursday, 12:00–13:00, 421 Evans Hall
- Graduate student instructor: To be announced

Course outline

This course is a 3-unit course that covers a selection of more advanced topics in statistical machine learning, focusing on methods and theory of classification, density estimation and function estimation. We will also cover uniform convergence and concentration theorems with statistical applications. The emphasis will be on core methodology and theory in machine learning and non-parametric statistics. *Note:* If you are looking for an applications-focused course in machine learning, this course will not be suitable.

More specifically, a selection of topics to be covered include:

- Classification and kernel methods
  - Decision-theoretic formulation and basic algorithms
  - Constrained optimization and duality
  - Surrogate loss functions: support vector machines, logistic regression, boosting
  - Reproducing kernel Hilbert spaces and uses
- Non-parametric density estimation and regression
  - Methods (splines, orthogonal expansions, kernel methods, penalized MLEs)
  - Function classes and covering numbers
  - Minimax lower bounds

- Concentration and risk bounds
  - Concentration inequalities and applications
  - Covering numbers and metric entropy
  - Statistical consequences

Reading materials

There is no single reader for the course. In addition to lecture notes, some text books that you may find helpful include:

- *Probabilistic Theory of Pattern Recognition*, by L. Devroye, L. Gyorfi, G. Lugosi, Springer, New York, 1996.
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, by Trevor Hastie, Robert Tibshirani and Jerome Friedman, Springer, 2001.
- *Learning with Kernels*, by B. Schölkopf and A. Smola.

Links to other useful materials (e.g., books, tutorials, survey papers) will be posted on the website as the semester proceeds.

Prerequisites

The prerequisites are a strong background in linear algebra, multivariate calculus, as well as advanced training (introductory graduate level) in probability and statistics. An appropriate background would include at least one of the courses STAT 241A/EECS 281A, STAT 205A or STAT 210A, preferably with at least an A grade. Coursework in optimization theory (EE 227A) would also be helpful, but is not required. Please note that this course is more mathematically demanding than its precursor STAT 241A/EECS 281A.

Evaluation

Evaluation will be based on a combination of regular homework assignments (20%), an in-class midterm exam (40%) and a final project (40%).
Homework: Although it is acceptable for students to discuss the homework assignments with one another, each student *must* write up his/her homework on an individual basis. Each student must indicate with whom (if anyone) they discussed the homework problems. Late homeworks will not be accepted.
Final project: The final project can be in any area related to the topics of the course. Possibilities include implementing a survey of some sub-area (3–5 papers), implementing some algorithms, running experiments using an algorithm for a particular application, trying to extend an existing method or theoretical result, or doing some combination of these. You will need to submit a brief written report and give a presentation in class in the last week of semester (a poster presentation or a talk, depending on the class size).

Academic policy: Please see the EECS department policy on academic dishonesty at: `http://www.eecs.berkeley.edu/Policies/acad.dis.shtml`.