STAT 7630 Nonparametric Function Estimation

Term: Spring 2023

Lectures: M 10:05–11:55 AM (2 credit hours) in Baker Systems Engineering 184

Instructor: Yoonkyung Lee
Office: 440H Cockins Hall
Office Hours: M 4:10–5:10 PM and F 10:30–11:30 AM (Virtually via CarmenZoom)
or by appointment
Virtual Office: Zoom link available on the Carmen website
Email: yklee@stat.osu.edu or lee.2272@osu.edu

Grader: Rui Zhang
Office Hours: by appointment only
Email: zhang.9473@osu.edu

Course Website: `https://carmen.osu.edu`

Course Description:
Statistics 7630 aims to introduce a nonparametric function estimation method with roughness penalties. Starting from smoothing splines for univariate data, a unified framework for penalized likelihood approach will be developed for flexible model building with splines covering multivariate data with both Gaussian and non-Gaussian responses. Mathematical formulation of smoothing splines, reproducing kernel Hilbert space methods, selection of a smoothing parameter, computation, and their applications will be treated in detail. In addition, connection between spline models and kernel methods in machine learning (especially support vector machines) will be discussed.

Prerequisites: Mathematical maturity in analysis and linear algebra, and a good knowledge of basic statistical inference (6802) and regression (6450/6950) are expected. Some knowledge of functional analysis (familiarity with Hilbert spaces), multivariate analysis (6560), and generalized linear models (7430) would be helpful, but not required.

Expected Learning Outcomes:
By the end of this course, students should successfully be able to

- Comprehend the mathematical formulation of smoothing splines and penalized likelihood approach in general
- Understand the role of kernels in defining modeling procedures
- Understand the relation between parametric and nonparametric modeling procedures
- Understand the bias-variance tradeoff with a tunable parameter in nonparametric modeling procedures
- Build statistical models for data using smoothing splines and kernel methods and interpret the results in the context of the data problem.

Textbooks:

- Required:
  Smoothing Spline ANOVA Models by Chong Gu (2013), 2nd edition, Springer.
  (An electronic version is available online through the OSU library website)

- Recommended:
  1. *Spline Models for Observational Data* by Grace Wahba.
  (An electronic version is available online through the OSU library website)
  2. *Nonparametric Regression and Generalized Linear Models* by Peter Green and Bernard Silverman.
  3. *Learning with kernels* by Bernhard Schölkopf and Alexander Smola.

Coursework:

- **Homework Assignments:** Homework will be assigned regularly (about every two weeks) throughout the semester using the Assignments link on Carmen. Typically, homework assignments will involve analytical exercises, computing and data analysis. You should submit your own individual work. In addition, computer code must be separately submitted as an appendix to each assignment. Your code should include comment statements that indicate what sections of the code correspond to the specific homework questions so that, if needed, the grader can read and check your code for its accuracy. Homework solutions will be posted on the course webpage. Due dates for homework assignments will be announced on Carmen and in class. Late submissions will NOT be accepted (unless an alternative arrangement has been made with the instructor prior to the deadline for valid reasons).

- **Final Project:**

  The final project will consist of selecting a research paper on topics related to the main theme of the course, presenting the paper through a recorded video (by April 28), and submitting a final report (by May 1) using the Assignments page. A list of suggested papers will be provided, but students can also choose other papers in consultation with the instructor. Depending on the enrollment size, students may complete the project in groups of two members or individual projects may also be allowed. More details will be provided on the course website later.

- **Participation:**

  In addition to regular class participation, there will be several activities requiring your participation for building connections with other students or formulating potential projects (e.g., posting introduction video, proposing papers for project). These activities will be announced in class and on Carmen.

Grading: There will be no in-class written exam. Course grades will be assigned on the basis of performance on homework assignments (50%), final project (40%) and participation (10%).

Tentative Course Schedule:
Homework due dates and project due dates are tentatively as follows. Please refer to in-class announcements (also on Carmen) for official dates.

Week	Dates	Topics, Assignments, Deadlines
1	1/9–1/13	Introduction to smoothing splines (CG 1.1, GW Foreword, GS Ch 1)
2	1/16–1/20	1/16(M): Martin Luther King Jr Day
3	1/23–1/29	Splines for interpolation and smoothing (GS 2.1–2.3), Natural splines, B-splines Homework assignment 1 due
4	1/30–2/3	Functional analytic approach to smoothing splines (CG 2.3, GW 1.2–1.3)
5	2/6–2/10	Characterizing the solution to the smoothing problem, Representer theorem (CG 3.1, GW 1.3) Homework assignment 2 due
6	2/11–2/17	Influence of the tuning parameter on smoother matrix (CG 3.1, GW 1.3)
7	2/20–2/24	Smoothing parameter selection, cross-validation (CG 3.2, GW 4.2–4.3, GS 3.1–3.4) Homework assignment 3 due
8	2/27–3/3	Smoothing splines as Bayes estimates (CG 2.5, GW 1.5, GS 3.8)
9	3/6–3/10	Confidence intervals (CG 3.3, GW Ch 5), Numerical examples (CG 3.6) Homework assignment 4 due
	3/13–3/17	Spring Break
10	3/20–3/24	Introduction to Reproducing Kernel Hilbert Spaces (CG 2.1, GW 1.1, SS 2.1-2.2)
11	3/27–3/31	Properties of reproducing kernels (CG 2.1, 2.4) Project proposal
12	4/3–4/7	Smoothing Spline ANOVA models (CG 3.1, GW Ch 10) Homework assignment 5 due
13	4/10–4/14	Generalized spline models with non-Gaussian responses (CG 5.1–5.3, GS 5.1–5.3)
14	4/17–4/21	Support vector machines (SS 7.1–7.3)
15	4/24–4/28	4/24 (M): Last day of class Nonlinear support vector machines, Constrained optimization (SS 7.4–7.5) Project presentation video due by 4/28 (F)
16	5/1–5/5	Project report due by 5/1 (M)

Disclaimer

This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular we reserve the right to change due dates or the methods of grading and/or assessment if necessary. Any changes will be communicated to you through official course announcements.

Academic Integrity Policy

Although students are encouraged to work together on assignments, each student is expected to write and submit individual solutions to homework problems.

Academic misconduct will not be tolerated and will be dealt with procedurally in accordance with university policy. It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct at `https://studentconduct.osu.edu/`.

Health and Safety Policy

Guidelines and requirements for campus safety can be found at `https://safeandhealthy.osu.edu`. All students, faculty and staff are required to comply with and stay up to date on all university safety and health guidelines.

Accessibility accommodations for students with disabilities

The university strives to make all learning experiences as accessible as possible. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. If you anticipate or experience academic barriers based on your disability (including mental health, chronic, or temporary medical conditions), please let the instructor know immediately so that we can privately discuss options. To establish reasonable accommodations, we may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: slds@osu.edu; 614-292-3307; `http://slds.osu.edu`; 098 Baker Hall, 113 W. 12th Avenue.