STA 790 Advanced Regression

Lectures

- WF 1:45 -- 3:00 PM.
- Zoom link under Zoom Meetings on the left.
- Password: "advancereg".

Topics

Our main focus is going to be Bayesian data smoothing within a regression framework. We will discuss why this is necessary -- with motivations coming from causal inference in observational studies and in analyzing extreme observations. We will study smooth Gaussian process regression, starting from its original motivation in the smoothing spline literature of Grace Wahba. We will review the theoretical foundation of Gaussian processes and associated reproducing kernel Hilbert space theory and more modern advances in understanding how a Gaussian process measure spreads its mass across different function classes. We will study basic mean regression and binary classification based on the basic theory and then move to more advanced density estimation and quantile regression.

Theme

We will see a mix of theory, method, computation and application -- giving us the perspectives of how to judge whether a Bayesian model is useful, expressive, computable and well behaved. A brief overview will be given of asymptotic properties of the posterior, touching upon the very general notion of frequentist properties of Bayesian methods. Many of our explorations will require computational exercises.

From Lectures

Smoothing: 3/24 (code), 3/31

Gaussian Processes: 3/31, 4/02, 4/07 4/09 (Rmd file, Slides, main R/C code)

Frequentist Theory: 4/16, 4/21

Quantile Regression: 4/21, 4/23

Suggested Exercise Problems: Set 1

Additional Reading

I will list here some interesting reading material with some brief annotations. This will expand as we proceed with the course.

Foundation

- The preface of the book Bayesian Nonparametric (2010) edited by N Hjort, C Holmes, P Muller and S Walker. Read Sections 1.1 and 1.2 to get a very general introduction to *Bayesian Nonparametrics* as a subarea of statistics.
- Bayarri and Berger (Stat Sinica 2004) on *The Interplay of Bayesian and Frequentist Analysis*. Read Sections 1 through 3. These discussions are very pertinent to the philosophical foundation of Bayesian nonparametrics. But the whole article is an excellent piece to read for statistics PhD students.

Splines etc.

- Teaching slides by *Thierry Denoeux* on Splines and generalized additive models (with worked out examples in R). You may also check out his other slides on Machine Learning.
- Hall (2005) on the *asymptotic theory* of penalized regression (i.e., smoothing splines). The technical details may be a little too dense and dry (hallmarks of Hall's writing), but the pay attention to how the theory questions are set up, the assumptions made, and the results stated. We will see equivalent or better results for Bayesian regression smoothing with Gaussian processes later in the course.

Gaussian Processes

- Gaussian Porcesses for Machine Learning, book by Williams and Rasmussen is a good way to get introduced to GP regression smoothin (and classification) from a machine learning perspective. The book came out in 2006, when the ML community was quite interested in GPs for their connection to artifical neural networks (NNs). A GP could be seen as a limiting version of some NN with infinite complexity. And GP was much easier to fit than NNs. Of course this all changed after the rise of deep learning -- which could also get increase the complexity of ordinary NNs, but are amenable to faster computing (but still with a fair amount of tuning). But perhaps the biggest advantage of deep NNs over GPs is that they assume and exploit structures, while GP can be an agnostic smoother. A good example is image processing, where the success of convolutional NN (CNN) comes mostly from the idea of how CNN uses local summation to create features that are meaningful in the context of images (and could be spectacularly bad, e.g., in genetics). So, GPs have stopped being the go-to tool in ML, particularly in AI applications. But they still remain widely used in less structured problems.
- This lecture note by Steve Lalley on a more probabilistic/analytic introduction to Gaussian processes. You will find several overlapping themes/details with my slides. But it's more technical than what I need for this course.
- These lecture slides by Arthur Gretton on RKHS. This is a good supplement to what we will discuss here and connects with the GPML book mentioned above in discussing various different GP covariance kernels. In contrast we will use mostly one type, the squared exponential kernel.
- Kanagawa et al (2018) on *Gaussian Processes and Kernel Methods: A Review on Connections and Equivalences.* An excellent treatment of the theoretical basis of Gaussian processes with an emphasis of comparing and constrasting Bayesian vs. kernel ridge regression approaches toward regression, quadrature and nonparametric test of independence.

Quantile Regression

- Tokdar and Kadane (2012) on our first attempt at the problem and a detailed analysis of the hurricane intensity study.
- Yang and Tokdar (2017) which delivered the full solution. The associated R package is qrjoint which you can install from CRAN. This vignette primarily by Erika could be useful in understanding the method and the software use at a deeper level.
- Chen and Tokdar (2021+, under reivew) for application to spatial data.
- Tokdar and Cunningham (2021+, under review) for the related concept of semiparametric Bayesin density estimation in the context of heavy tailed data and tail index estimation.