Baylor University
COLLEGE OF ARTS & SCIENCES
Statistical Science

Douglas Nychka

Department of Applied Mathematics and Statistics
Colorado School of Mines

Hybrid smoothing

Thursday, March 6, 2025 2:00 p.m.

Marrs McLean Science Building GL51

Abstract: Spline smoothing, and more generally Gaussian process smoothing, has become a successful methodology for estimating a smooth trend or surface from noisy data. Similarly, the LASSO and related L1 penalties have become important tools for variable selection and also admit a Bayesian version based on the Laplace distribution. This project combines these two approaches as a method to detect discontinuous behavior in an otherwise smooth signal.

Every day, the Foothills Facility of Denver Water filters more than 250 million gallons of water for the metropolitan area. This process runs continuously and is monitored across an array of filters, each the size of a small swimming pool, at five-minute intervals. Anomalies in the filter functioning often take the form of discontinuities or appear as step changes in the smooth filtering cycle. This application motivates a mixed smoothing approach, where normal operation is captured by a smoothing spline and the anomalies by basis function coefficients determined using an L1 penalty—or as a Laplace prior on the step changes.

A more general problem arises in Earth system science, where climate variables such as surface temperature can follow a smooth surface but exhibit discontinuities at land/ocean boundaries or over changes in terrain. In this context, we have found that more general priors are appropriate for hybrid surfaces observed in climate data.

This talk will examine how to fit these models to data and leverage the correspondences between frequentist estimates with penalties and Bayesian hierarchical models, along with important details to make the MCMC computations practical. The similarities between these frequentist and Bayesian models rely on the correspondence between splines and Gaussian processes. Some background on this historical connection will be provided as part of developing the Bayesian model.

This is joint work with Matthew Hofkes and Amanda Herring, see https://arxiv.org/pdf/2402.03459, https://arxiv.org/abs/2410.21469 and the R Package Darkroast

Bio: Dr. Nychka is a statistician who works in applications for the environment. He is a statistician and data scientist whose areas of research include the theory, computation and application of curve and surface fitting with a focus on geophysical and environmental applications. Currently he is a Professor in the Department of Applied Mathematics and Statistics at the Colorado School of Mines and Senior Scientist Emeritus at the National Center for Atmospheric Research (NCAR), Boulder, Colorado. Before moving to Mines he directed the Institute for Mathematics Applied to Geosciences at NCAR. His current focus in research is the computation of spatial statistics methods for large data sets and the migration of these algorithms into easy to use R packages.

He has coauthored more than 100 research articles and with an h-index of 50. He is a Fellow of the American Statistical Association, Fellow of the Institute for Mathematical Statistics and a recipient of the Jerry Sacks Award for interdisciplinary research.