



THE UNIVERSITY OF CHICAGO

Department of Statistics

SCIENTIFIC AND STATISTICAL COMPUTING SEMINAR

GUANG CHENG

Department of Statistics
Purdue University

Bayesian Aggregation for Extraordinarily Large Dataset

THURSDAY, February 18, 2016 at 4:30 PM
133 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

In this talk, a set of scalable Bayesian inference procedures is developed for a general class of nonparametric regression models. Specifically, we first perform independent nonparametric Bayesian inference on each subset split from a massive dataset, and then aggregate those local results into global counterparts. This aggregation step is explicit without involving any additional computation cost. By a careful partition, we show that our aggregated inference results obtain the oracle rule in the sense that they are equivalent to those obtained directly from the entire data (which are computationally prohibitive). For example, an aggregated credible ball achieves desirable credibility level and also frequentist coverage while possessing the same radius as the oracle ball. This oracle matching phenomenon occurs due to a delicate geometric structure of the infinite-dimensional parameter space in consideration.

Organizers:

Lek-Heng Lim, Department of Statistics, lekheng@galton.uchicago.edu, Ridgway Scott, Departments of Computer Science and Mathematics, ridg@cs.uchicago.edu, Jonathan Weare, Department of Statistics and The James Franck Institute, weare@uchicago.edu. SSC Seminar URL: http://www.stat.uchicago.edu/seminars/SSC_seminars.shtml.

If you wish to subscribe to our email list, please visit the following website:
<https://lists.uchicago.edu/web/arc/statseminars>.