COLLOQUIUM SERIES: FALL 2015

SCALABLE BAYES VIA FAST COMPUTATION OF BARYCENTER OF SUBSET POSTERIORS

Abstract:

*Joint work with David Dunson and Cheng Li, both from Duke University. For further details, see the manuscript: http://arxiv.org/abs/1508.05880v1

The promise of Bayesian methods for big data sets has not fully been realized due to the lack of scalable computational algorithms. For massive data, it is necessary to store and process subsets on different machines in a distributed manner. We propose a simple, general, and highly efficient approach, which first runs a posterior sampling algorithm in parallel on different machines for subsets of a large data set. To combine these subset posteriors, we calculate the Wasserstein barycenter via a highly efficient linear program. The resulting estimate for the Wasserstein posterior (WASP) has an atomic form, facilitating straightforward estimation of posterior summaries of functionals of interest. The WASP approach allows posterior sampling algorithms for smaller data sets to be trivially scaled to huge data. We provide theoretical justification in terms of posterior consistency and algorithm efficiency. Examples are provided in complex settings including Gaussian process regression and nonparametric Bayes mixture models.

WWW. STAT.UIOWA.EDU

DEPARTMENT OF STATISTICS & ACTUARIAL SCIENCE
UNIVERSITY OF IOWA
319-335-0712

PRESENTER

Sanvesh Srivastava
Assistant Professor
Department of Statistics and Actuarial Science
The University of Iowa

PROFESSIONAL INTERESTS:

DEVELOPING FLEXIBLE BAYESIAN METHODS AND EFFICIENT COMPUTATIONAL ALGORITHMS FOR BIG DATA SETS, TAILORED FOR BOTH THEIR COMPLEXITY AND SIZE

WHEN

September 17, 2015
3:30 p.m.

WHERE

61 Schaeffer Hall

RECEPTION

241 Schaeffer Hall
3:00 p.m.