Department of Statistics

DISSERTATION PROPOSAL

JASON WILLWERSCHEID

Department of Statistics
The University of Chicago

Empirical Bayes Matrix Factorization: Methods and Applications

FRIDAY, January 31, 2020, at 1:30 PM
Jones 303, 5747 S. Ellis Avenue

ABSTRACT

The primary objective of the thesis is to develop methods and theory for fitting empirical Bayes matrix factorization (EBMF) models to large, unruly datasets.

EBMF, a model-based method for sparse matrix factorization proposed by Wang and Stephens, has been shown to yield highly interpretable factors in a variety of settings, and regularly fares better than competing methods on tasks such as data imputation.

The thesis builds on the work of Wang and Stephens. Improvements to the fitting algorithm have led to the successful application of the method to much larger datasets. Often, the difference of scale has turned out to be a difference of kind: the new applications have motivated a re-examination of many of the earlier theoretical and methodological assumptions.

I will begin the proposal presentation with a brief exposition of Wang and Stephen's model. Next, I will outline my own contributions to EBMF methods and theory. These innovations are largely responses to the particular questions posed by the applications, which I will then describe in some detail. Finally, I will give an overview of the software being developed concurrently with the project.

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: https://lists.uchicago.edu/web/subscribe/statseminars.