

Spike-and-Slab Generalized Additive Models and Fast Algorithms for High-Dimensional Data

Boyi Guo and Nengjun Yi

Department of Biostatistics
University of Alabama at Birmingham

August 8th, 2021

High Dimensional Generalized Additive Model

- ▶ Models that can handle grouped predictors
 - ▶ Penalized model with grouped penalty: Lasso, SCAD
 - ▶ Bayesian regularized models with grouped prior: spike-and-slab prior
- ▶ Overly shrinking due to sparse penalty
- ▶ Reduced interpretation due to the “all-in-all-out” functional selection
- ▶ Lack of uncertainty measure for penalized models
- ▶ Scaling problem for Bayesian regularized models

Bayesian Hierarchical Additive Model

- ▶ Re-parameterization of smoothing function design matrix
 - ▶ Eigen-decompose the smoothing matrix of a smoothing function
 - ▶ Deploy smoothing penalty on the smoothing function
 - ▶ Isolate the linear and polynomial spaces of a smoothing function
- ▶ Two-part spike-and-slab lasso prior
 - ▶ Independent SS-Lasso priors of linear and polynomial spaces for sparse solution
 - ▶ Two latent indicator variables for bi-level selection
 - ▶ Hyperprior of latent indicator variables for local adaptation

Optimization Algorithms

- ▶ EM algorithms
- ▶ Combine with coordinate descent algorithm for fast computing and sparse solution
- ▶ Combine with iterative weighted least square for uncertainty measures
 - ▶ Accomodate other priors: Bayesian ridge, lasso priors, spike-and-slab priors

Software Support

- ▶ R package: BHAM
 - ▶ Ancillary functions for setting up high-dimensional spline formulation
 - ▶ Model summary and variable selection
 - ▶ Curve plotting
 - ▶ Covariate adjustment without penalty
 - ▶ Website via *boyiguo1.github.io/BHAM*