

## Group variable selection via convex Log-Exp-Sum penalty

Zhigeng Geng<sup>\*</sup>  
zgend@stat.wisc.edu

Sijian Wang<sup>\*†</sup>  
swang@biostat.wisc.edu

Grace Wahba<sup>\*</sup>  
wahba@stat.wisc.edu

---

In many scientific and engineering applications, predictors are naturally grouped, for example, in biological applications where assayed genes or proteins can be grouped by biological roles or biological pathways. Common statistical analysis methods such as ANOVA, factor analysis, and functional modeling with basis sets also exhibit natural variable groupings. When the group structures are available among predictors, people are usually interested in identifying both important groups and important variables within the selected groups. Among existing successful group variable selection methods, some methods fail to conduct the within group selection. Some methods are able to conduct both group and within group selection, but the criterion is non-convex which may require extra numerical effort. In this paper, we propose a novel Log-Exp-Sum penalty for group variable selection. The LES penalty is convex, which may yield numerical stability. It can identify important groups as well as select important variables within the group. We develop an efficient group-level coordinate descent algorithm for solving the corresponding optimization problem, and we also derive non-asymptotic theoretical bounds for the new method in the high-dimensional setting where the number of predictors can be much larger than the sample size. Numerical results indicate that the proposed method works well in terms of both variable selection and prediction accuracy. We also applied the proposed method to American Cancer Society breast cancer dataset.

*Key words* : Coordinate descent ; Group LASSO ; Group variable selection ; LASSO ; Log-Exp-Sum Penalty ; Oracle inequality

---

<sup>\*</sup>Department of Statistics, University of Wisconsin Madison, 1300 University Avenue, Madison, WI 53706, USA.

<sup>†</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin Madison, 600 Highland Avenue, Madison, WI 53792-4675, USA.