

# Grouped variable selection via nested spike and slab priors

Tso-Jung Yen<sup>1\*</sup> (顏佐榕) and Yu-Min Yen<sup>2</sup>

<sup>1</sup>Institute of Statistical Science, Academia Sinica

<sup>2</sup>Institute of Economics, Academia Sinica

## Abstract

In this paper we study grouped variable selection problems. We propose a specified prior, called the nested spike and slab prior, to model collective behavior of regression coefficients. At the group level, the nested spike and slab prior puts positive mass on the probability that the  $\ell_2$ -norm of the grouped coefficients is equal to zero. At the individual level, it assigns each coefficient a spike and slab prior. To estimate the model, we adopt a maximum a posteriori approach. The estimation problem involves an approximate objective function modified with the majorization-minimization technique. Simulation studies show that the proposed estimator performs relatively well when both the true and redundant covariates are included in the same group. Asymptotic analysis further shows that the proposed estimator can achieve a smaller  $\ell_2$  estimation error if groups that include the true covariates do not include too many redundant covariates. In addition, given some regular conditions hold, the proposed estimator can achieve model selection consistency regardless of the irrepresentable-type conditions.

Keywords: group sparsity, irrepresentable conditions, log-sum approximation, majorization-minimization algorithms, subgradients