Ann Inst Stat Math (2017) 69:997–1025
DOI 10.1007/s10463-016-0571-z

A doubly sparse approach for group variable selection

Sunghoon Kwon[1] · **Jeongyoun Ahn**[2] · **Woncheol Jang**[3] · **Sangin Lee**[4] · **Yongdai Kim**[3]

Received: 17 December 2014 / Published online: 28 June 2016

Abstract We propose a new penalty called the doubly sparse (DS) penalty for variable selection in high-dimensional linear regression models when the covariates are naturally grouped. An advantage of the DS penalty over other penalties is that it provides a clear way of controlling sparsity between and within groups, separately. We prove that there exists a unique global minimizer of the DS penalized sum of squares of residuals and show how the DS penalty selects groups and variables within selected groups, even when the number of groups exceeds the sample size. An efficient optimization algorithm is introduced also. Results from simulation studies and real data analysis show that the DS penalty outperforms other existing penalties with finite samples.

Keywords Doubly sparse penalty · Group selection · Group selection consistency · Variable selection

✉ Yongdai Kim
ydkim0903@gmail.com

[1] Department of Applied Statistics, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Korea

[2] Department of Statistics, University of Georgia, Athens, GA 30602, USA

[3] Department of Statistics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea

[4] Quantitative Biomedical Research Center, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA