

PhDay-EIO

Santiago de Compostela, 4 de outubro do 2019

Dealing with the LASSO adjustment problems as a variable selector

Laura Freijeiro González¹, Manuel Febrero Bande¹, Wenceslao González Manteiga¹

¹ Departamento de Estatística, Análise Matemática e Optimización. Universidade de Santiago de Compostela

RESUMO

In a Big Data context where the number of covariates p is large and greater than the number of samples n ($p \gg n$), it takes special importance the necessity of synthesizing this big amount of information in simple models. Furthermore, these models should be able to select the important covariates and avoid unnecessary noise. The penalized linear regression models face these problems and provide useful solutions, specially those of the LASSO family.

Nevertheless, this algorithm exhibits some drawbacks related to the correct selection of important covariates and the exclusion of redundant information. Thus, particular controlled scenarios are simulated with the purpose of analysing the drawbacks of this algorithm, giving an explanation of its behaviour and comparing with current approaches with the aim of achieving a better performance.