

Accepted Manuscript

A bootstrapping soft shrinkage approach for variable selection in chemical modeling

Bai-Chuan Deng, Yong-Huan Yun, Dong-Sheng Cao, Yu-Long Yin, Wei-Ting Wang,
Hong-Mei Lu, Qian-Yi Luo, Yi-Zeng Liang



PII: S0003-2670(16)30015-0

DOI: [10.1016/j.aca.2016.01.001](https://doi.org/10.1016/j.aca.2016.01.001)

Reference: ACA 234330

To appear in: *Analytica Chimica Acta*

Received Date: 31 August 2015

Revised Date: 14 December 2015

Accepted Date: 4 January 2016

Please cite this article as: B.-C. Deng, Y.-H. Yun, D.-S. Cao, Y.-L. Yin, W.-T. Wang, H.-M. Lu, Q.-Y. Luo, Y.-Z. Liang, A bootstrapping soft shrinkage approach for variable selection in chemical modeling, *Analytica Chimica Acta* (2016), doi: 10.1016/j.aca.2016.01.001.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A bootstrapping soft shrinkage approach for variable selection in chemical modeling

Bai-Chuan Deng^{a,b,d}, Yong-Huan Yun^b, Dong-Sheng Cao^{c,**}, Yu-Long Yin^{a,d},

Wei-Ting Wang^b, Hong-Mei Lu^b, Qian-Yi Luo^b, Yi-Zeng Liang^{b,*}

^aCollege of Animal Science, South China Agricultural University, Guangzhou 510642, P.R. China

^bSchool of Chemistry and Chemical Engineering, Central South University, Changsha 410083, P.R. China

^cSchool of Pharmaceutical Sciences, Central South University, Changsha 410083, P.R. China

^dInstitute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, P.R. China

Abstract

In this study, a new variable selection method called bootstrapping soft shrinkage (BOSS) method is developed. It is derived from the idea of weighted bootstrap sampling (WBS) and model population analysis (MPA). The weights of variables are determined based on the absolute values of regression coefficients. WBS is applied according to the weights to generate sub-models and MPA is used to analyze the sub-models to update weights for variables. The optimization procedure follows the rule of soft shrinkage, in which less important variables are not eliminated directly but are assigned smaller weights. The algorithm runs iteratively and terminates until the number of variables reaches one. The optimal variable set with the lowest root mean squared error of cross-validation (RMSECV) is selected. The method was tested on three groups of near infrared (NIR) spectroscopic datasets, i.e. corn datasets, diesel

* Corresponding author. Tel.: +86 731 8830824; fax: +86 731 8830831.

E-mail addresses: yizeng_liang@263.net

** Corresponding author. E-mail address: oriental-cds@163.com

Download English Version:

<https://daneshyari.com/en/article/7554913>

Download Persian Version:

<https://daneshyari.com/article/7554913>

[Daneshyari.com](https://daneshyari.com)