

AIC for the non-concave penalized likelihood method

Yuta Umezu¹ · Yusuke Shimizu² ·
Hiroki Masuda³ · Yoshiyuki Ninomiya⁴

Received: 14 February 2017 / Revised: 23 December 2017 / Published online: 27 February 2018
© The Institute of Statistical Mathematics, Tokyo 2018

Abstract Non-concave penalized maximum likelihood methods are widely used because they are more efficient than the Lasso. They include a tuning parameter which controls a penalty level, and several information criteria have been developed for selecting it. While these criteria assure the model selection consistency, they have a problem in that there are no appropriate rules for choosing one from the class of information criteria satisfying such a preferred asymptotic property. In this paper, we derive an information criterion based on the original definition of the AIC by considering minimization of the prediction error rather than model selection consistency. Concretely speaking, we derive a function of the score statistic that is asymptotically equivalent to the non-concave penalized maximum likelihood estimator and then provide an estimator of the Kullback–Leibler divergence between the true distribution and the estimated distribution based on the function, whose bias converges in mean to zero.

The work of Y. Ninomiya (corresponding author) was partially supported by a Grant-in-Aid for Scientific Research (16K00050) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The works of H. Masuda and Y. Shimizu were partially supported by JST CREST Grant Number JPMJCR14D7, Japan.

✉ Yoshiyuki Ninomiya
nino@imi.kyushu-u.ac.jp

- ¹ Department of Computer Science, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
- ² Department of Mathematics, Josai University, 1-1 Keyakidai, Sakado, Saitama 350-0295, Japan
- ³ Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan
- ⁴ Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan

Keywords Information criterion · Kullback–Leibler divergence · ℓ_q regularization · Statistical asymptotic theory · Tuning parameter · Variable selection