Annals of the Institute of Statistical Mathematics (2022) 74:1109–1142
https://doi.org/10.1007/s10463-022-00824-8

A sequential feature selection procedure for high-dimensional Cox proportional hazards model

Ke Yu[1] · **Shan Luo**[1]

Received: 8 December 2021 / Revised: 16 March 2022 / Accepted: 17 March 2022 /
Published online: 7 May 2022

Abstract
Feature selection for the high-dimensional Cox proportional hazards model (Cox model) is very important in many microarray genetic studies. In this paper, we propose a sequential feature selection procedure for this model. We define a novel partial profile score to assess the impact of unselected features conditional on the current model, significant features are thereby added into the model sequentially, and the Extended Bayesian Information Criteria (EBIC) is adopted as a stopping rule. Under mild conditions, we show that this procedure is selection consistent. Extensive simulation studies and two real data applications are conducted to demonstrate the advantage of our proposed procedure over several representative approaches.

Keywords Sequential feature selection · Selection consistency · Cox proportional hazards model · High-dimensional · Extended Bayesian information criteria

✉ Shan Luo
sluomath@sjtu.edu.cn

[1] School of Mathematical Sciences, Shanghai Jiao Tong University, 800 Dongchuan RD, Minhang District, Shanghai 200240, China