Variables selection in high dimension in a joint model of survival times and longitudinal outcomes with random effects.

Antoine Caillebotte*[1], Estelle Kuhn[1], and Sarah Lemler[2]

[1]Mathématiques et Informatique Appliquées du Génome à l'Environnement [Jouy-En-Josas] – Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France

[2]Mathématiques et Informatique pour la Complexité et les Systèmes – CentraleSupélec, Université Paris-Saclay – France

Résumé

A current challenge in many fields is understanding the interactions between dependent dynamic phenomena. For example, in plant science, the dynamics of plant development and the spread of an epidemic disease or pests in this field. Mathematical modeling is a powerful tool for capturing interactions between phenomena. Indeed, the joint modeling of several phenomena has shown its efficiency in several areas, notably in medicine, pharmacology, and biology. In this work, we consider a joint model of survival time and repeated measurements, longitudinal data in our setting, of a population of individuals. More precisely, we model the evolution of a dynamic variable of interest over time using a nonlinear mixed-effects model. This model is combined with a Cox model in which we introduce classical covariates characterizing the individual considered and also a link function based on the mixed-effects model function of the related dynamic. Moreover, we consider a high-dimensional setting where the number of covariates in the Cox model is larger than the sample size. It is necessary to adapt the statistical and numerical approaches for inference to deal with this high-dimensional setting since classical tools failed. Therefore we propose a new method for parameter inference and variable selection. We consider the penalized maximum likelihood estimate using a LASSO penalty to account for the high dimensionality of covariates. We implement a preconditioned proximal stochastic gradient algorithm to compute this estimate, dealing both with unobserved random effects of the mixed-effects model and the penalty term. We fixed the penalty coefficient according to the BIC criterion. Finally, we highlight the performance of our method through an intensive simulation study.

*Intervenant

sciencesconf.org:jps-2023:493931