MODEL SELECTION VIA PENALIZATION IN THE ADDITIVE COX MODEL

Marta Avalos[1], Yves Grandvalet[2], Christophe Ambroise[2]

[1] INSERM U738, Dpt of Epidemiology, Biostatistics and Clinical Research, CHU Bichat-Claude Bernard, 75018 Paris, FRANCE.

[2] Heudiasyc laboratory, UMR CNRS 6599, Compiègne University of Technology, BP 20529, 60205 Compiègne, FRANCE

ABSTRACT

The Cox proportional hazards model is the most popular model for the analysis of survival data. It allows estimating the relationship between covariates and a possibly censored failure time. The corresponding partial likelihood estimators are used for the estimation and prediction of relative risk of failure. However, if the explanatory variables are highly correlated or if the number of failures is not much greater than the number of covariates of interest, then partial likelihood estimators are unstable and have large variance.

Penalization is extensively used to address these difficulties. It decreases the predictor variability to improve the accuracy of prediction. Ridge regression (l2-penalization) is one of the main penalization procedures. It has been generalized to the nonparametric setting to reduce the possibility of overfitting with high dimensional models. Thus, smoothing splines are used to estimate flexibly covariate effects in the additive Cox model. Tibshirani's lasso (l1-penalization) has also been applied to the Cox model, providing an alternative to quadratic penalization. An attractive feature of the l1-penalization is that it shrinks coefficients and sets some of them to zero, performing parameter estimation and variable selection simultaneously.

We propose a new algorithm for variable selection and function estimation in the additive Cox model. The method is based on a generalization of the lasso to the nonparametric setting. Our proposal maximizes a penalized partial likelihood that includes a double penalty: on the l1-norm of linear components and on the (generalized) l1-norm of nonlinear components of spline coefficients. Because of their nature, these penalties shrink linear and nonlinear compounds, some of them reducing exactly to zero. Hence they give parsimonious models, select significant variables, and reveal nonlinearities in the effects of predictors. Our approach is compared to standard methods by simulations and an example. Different techniques to choose the penalty parameters are also tested.

Key words: penalization, variable selection, complexity tuning, penalized partial likelihood, proportional hazards model, survival data, smoothing splines, nonparametric regression, function estimation algorithm, constrained optimization.