Minorization-Maximization-based Boosting for Large-scale Survival Analysis with Time-Varying Effects

Zhi (Kevin) He

University of Michigan
E-mail: kevinhe@umich.edu

Abstract: National disease registries have produced a vast amount of data. Many existing statistical methods that perform well for moderate sample sizes and small-dimensional data do not scale to such large-scale data, leading to a demand for statistical techniques that enable full utilization of these rich sources of information. For example, the time-varying effects model is a flexible and powerful tool for modeling the dynamic changes of covariate effects. However, in survival analysis, its computational burden increases quickly as the number of sample sizes or predictors grows. Traditional methods that perform well for moderate sample sizes and low-dimensional data do not scale to massive data. Analysis of national kidney transplant data with a massive sample size and large number of predictors defy any existing statistical methods and software. In view of these difficulties, we propose a Minorization-Maximization-based boosting procedure for estimating the time-varying effects. Leveraging the block structure formed by the basis expansions, the proposed procedure iteratively updates the optimal block-wise direction along which the approximate increase in the log-partial likelihood is maximized. The resulting estimates ensure the ascent property and serve as refinements of the previous step. The performance of the proposed method is examined by simulations and applications to the analysis of national kidney transplant data.