A Bayesian Procedure for Association Study of Copy Number Variation in Next Generation Sequencing

Yu-Chung Wei
Department of Statistics, Feng Chia University

Abstract

Model selection for sparse high-dimensional Cox models has broad applications to contemporary biostatistics, in particular, to extracting relevant biomarkers from high-dimensional survival data. In this talk, we propose using a greedy-type algorithm, Chebyshev Greedy Algorithm (CGA), to iteratively include covariates in the aforementioned models, and show that with probability tending to one, all relevant covariates can be included in a moderate number of iterations. We also devise a high-dimensional information criterion (HDIC) to remove the redundant covariates chosen by CGA, thereby leading to selection consistency. Finally, the proposed method is illustrated using simulated data and a diffuse large B-cell lymphoma (DLBCL) dataset.