

Sparse Modelling of Cancer Patients' Survival Based on Genomic Copy Number Alterations

Authors : Khaled M. Alqahtani

Abstract : Copy number alterations (CNA) are structural variations in the genome in which some regions exhibit more or less than the normal two chromosomal copies. This genomic CNA profile provides critical information in tumor progression and is, therefore, informative for patients' survival. It is currently a statistical challenge to model patients' survival using their genomic CNA profiles while at the same time identifying regions in the genome that are associated with patients' survival. Some methods have been proposed, including the Cox proportional hazard (PH) model with ridge, lasso, or elastic net penalties. However, these methods do not take the general dependencies between genomic regions into account and produce results that are difficult to interpret. In this paper, we extend the elastic net penalty by introducing an additional penalty that takes into account general dependencies between genomic regions. This model produces smooth parameter estimates while simultaneously performing variable selection via sparse solution. The results indicate that the proposed method shows a better prediction performance than other models in our simulation study while enabling us to investigate regions in the genome that are associated with the patient's survival with sensible interpretation. We illustrate the method using a real dataset from a lung cancer cohort and simulated data.

Keywords : copy number alterations, cox proportional hazard, lung cancer, regression, sparse solution

Conference Title : ICAEM 2024 : International Conference on Applied and Engineering Mathematics

Conference Location : Venice, Italy

Conference Dates : June 20-21, 2024