Computing extended maximum likelihood estimates for Cox proportional-hazards models

Douglas B. Clarkson, MathSoft Inc.
Robert I. Jennrich, University of California, Los Angeles

Abstract

Infinite parameter estimates can occur in Cox proportional–hazards models when a linear combination of the covariates monotonically increases or decreases with the failure times. Considered here are methods for computing *extended maximum likelihood* estimates for Cox models. An extended estimate is a pair of vectors: a direction vector for the infinite component, and a vector that optimizes the "stratified" log-likelihood of Bryson and Johnson (1981). A method of identifying problems that have extended estimates and an algorithm for finding them is given along with an example illustrating their use.

Key Words: Computer algorithms; Cox regression; Linear programming; Solutions at infinity.

1 Introduction

When the maximizer of a likelihood is at infinity, many standard computational algorithms fail because they are not designed to deal with such solutions. Haberman (1974, appendix B) defines and gives an example of such estimates for frequency data. He calls the estimates obtained *extended maximum likelihood* estimates. Because these estimates contain infinite values, they do not exist in the usual sense, and many authors (Silvapulle and Burridge, 1986, Albert and Anderson, 1984, Hamada and Tse, 1988) have felt that detecting the presence of infinite estimates is sufficient. Bryson and Johnson (1981) note that infinite estimates can be common in the Cox (1972) proportional-hazards model, at least in a simulation study of a model with sample size 20. They also note that infinite estimates occur when the failure times are monotone with a linear function of the covariates, and they give an algorithm in which parameter estimates are computed based upon a "stratified" likelihood, where the stratification is determined by inspection of the covariates. Baker, Clarke, and Lane (1985) discuss an algorithm for computing extended maximum likelihood estimates in sparse contingency