Computational Statistics and Data Analysis 84 (2015) 135–151

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A new MM algorithm for constrained estimation in the proportional hazards model

Jieli Ding [a], Guo-Liang Tian [b,*], Kam Chuen Yuen [b]

[a] *School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, PR China*

[b] *Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China*

ARTICLE INFO

Article history:
Received 3 April 2013
Received in revised form 7 November 2014
Accepted 9 November 2014
Available online 17 November 2014

Keywords:
Asymptotic properties
Bootstrap approach
Constrained estimation
Karush–Kuhn–Tucker conditions
MM algorithm
Proportional hazards model

ABSTRACT

The constrained estimation in Cox's model for the right-censored survival data is studied and the asymptotic properties of the constrained estimators are derived by using the Lagrangian method based on Karush–Kuhn–Tucker conditions. A novel *minorization–maximization* (MM) algorithm is developed for calculating the maximum likelihood estimates of the regression coefficients subject to box or linear inequality restrictions in the proportional hazards model. The first M-step of the proposed MM algorithm is to construct a surrogate function with a diagonal Hessian matrix, which can be reached by utilizing the convexity of the exponential function and the negative logarithm function. The second M-step is to maximize the surrogate function with a diagonal Hessian matrix subject to box constraints, which is equivalent to separately maximizing several one-dimensional concave functions with a lower bound and an upper bound constraint, resulting in an explicit solution via a median function. The ascent property of the proposed MM algorithm under constraints is theoretically justified. Standard error estimation is also presented via a nonparametric bootstrap approach. Simulation studies are performed to compare the estimations with and without constraints. Two real data sets are used to illustrate the proposed methods.

1. Introduction

Survival data arise in a number of fields such as reliability engineering, economics, sociology, public health, epidemiology and medicine (especially, clinical trials). Survival analysis is used to model the relationship between the time-to-event (e.g., death or disease) and a set of covariates or predictors. When the period of observation expires, or an individual is removed from or drops out the study prior to the event occurs, survival data are considered as right-censored. The proportional hazards model originally introduced by Cox (1972) may be the most widely used method for analyzing survival data with censoring. Since the publication of Cox (1972), numerous extensions and developments in various aspects have been proposed during the past 40 years by many authors including Cox (1975), Andersen and Gill (1982), Bickel et al. (1993), Lin and Ying (1993), Lin (1994), Huang (1996), Chen and Little (1999), and Chen and Lo (1999). A comprehensive review was given by Kalbfleisch and Prentice (2002).

In many practical problems, it may be available as prior information that restrictions on some model parameters would result in a more reasonable interpretation. Such restrictions cannot be ignored; otherwise the statistical inference may be

* Corresponding author.
E-mail address: gltian@hku.hk (G.-L. Tian).

http://dx.doi.org/10.1016/j.csda.2014.11.005

misled and an underestimate of the effect may be caused (Tan et al., 2005; Fang et al., 2006). Therefore, it is reasonable to expect that the analysis would perform better if parameter constraints are taken into account in the modeling process. However, the complication resulting from such restrictions raises statistical challenges. Statistical inferences on constrained problems have been studied by many authors (e.g. Liew, 1976; Nyquist, 1991; Silvapulle, 1997). For example, Wang (1996, 2000) studied asymptotic properties of constrained estimators in nonlinear regressions. Moore and Sadler (2006) and Moore et al. (2008) discussed the asymptotic theory for the constrained *maximum likelihood estimator* (MLE) and presented a constrained Cramér–Rao bound. However, to our knowledge, asymptotic properties of constrained estimators for the regression coefficients in Cox's model have never been studied.

The first objective of this paper is to derive two asymptotic properties of the constrained MLEs for the regression coefficients in the proportional hazards model with right-censored data. These asymptotic results are useful in statistical inferences for Cox's model with box and linear inequality constraints. We use the Karush–Kuhn–Tucker conditions, a well-known approach in optimization with inequality constraints, to overcome the difficulty caused by the constraint. Similar techniques were adopted by Wang (2000), Xu and Wang (2008) for constrained least-squares estimator, and Moore and Sadler (2006), Moore et al. (2008) for constrained Cramér–Rao bound in parametric models.

Böhning and Lindsay (1988) developed a *quadratic lower bound* (QLB) algorithm with monotone convergence like the EM algorithm for Cox's model without constraints. Since the construction of the quadratic surrogate function in the QLB algorithm is based on the second-order Taylor expansion of the partial log-likelihood function in the neighborhood of the maximum likelihood estimate, this QLB algorithm cannot be applied to Cox's model with box and/or linear inequality constraints. Note that the QLB algorithm is a special case of *minorization–maximization* (MM) algorithms (Becker et al., 1997; Hunter and Lange, 2004; Lange, 2004, 2010). In addition, the existing MM algorithms (Lange et al., 2000) such as De Pierro's algorithm (De Pierro, 1995) cannot be applied to Cox's model even for the case without constraints. Hunter and Lange (2002) proposed an MM algorithm for finding the MLEs of the regression coefficients in the semiparametric proportional odds model just for the case without constraints.

Thus, the second objective of this paper is to develop a novel MM algorithm for calculating the MLEs of the regression coefficients with box or linear inequality restrictions in the proportional hazards model. The key to the proposed MM algorithm is to construct a surrogate function $Q(\boldsymbol{\beta}|\boldsymbol{\beta}^{(m)})$ with a diagonal Hessian matrix, which can be reached by utilizing the convexity of the exponential function e^x and the negative logarithm function $-\log x$. Maximizing this surrogate function with a diagonal Hessian matrix subject to box constraints is equivalent to separately maximizing several one-dimensional concave functions with a lower bound and an upper bound constraint, which has an explicit solution via a median function.

The rest of the article is organized as follows. In Section 2, we formulate the proportional hazards model with constraints and derive two asymptotic properties for the constrained estimator. In Section 3, we develop a new MM algorithm for calculating the constrained estimation in Cox's model. The ascent property of the proposed MM algorithm under constraints is derived. Standard error estimation is also presented via a non-parametric bootstrap approach. We conduct several simulation studies in Section 4 to compare the estimations with and without constraints. In Section 5, two real data sets are used to illustrate the proposed methods. A discussion is presented in Section 6. Detailed proofs on asymptotic properties are put in the Appendix.

2. Constrained estimation in Cox's model

2.1. The formulation of constrained Cox's model

Consider Cox's proportional hazards model with constrained regression coefficients. Suppose that there are n subjects drawn randomly from the population of interest. For the ith subject ($i = 1, \ldots, n$), let $\widetilde{T}_i, C_i, T_i = \min(\widetilde{T}_i, C_i)$ and $\mathbf{Z}_i$ denote the failure time, the censoring time, the observed time and the covariate vector of p dimension, respectively. We assume that given the covariate vector $\mathbf{Z}_i$, the failure time $\widetilde{T}_i$ and the censoring time C_i are conditionally independent. Furthermore, let $\Delta_i = I(T_i \leqslant C_i)$, $Y_i(t) = I(T_i \geqslant t)$ and $N_i(t) = \Delta_i I(T_i \leqslant t)$ respectively denote the right censoring indicator, at-risk process and counting process for subject i, where $I(\cdot)$ is the indicator function.

The proportional hazards model specifies the hazard function of the failure time conditional on covariates taking the following form:

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)\exp\{\boldsymbol{\beta}^\top \mathbf{Z}(t)\}, \tag{2.1}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, $\mathbf{Z}(t)$ is a p-dimensional vector of time-varying covariates, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a p-dimensional vector of regression coefficients. We are interested in estimating the unknown parameter vector $\boldsymbol{\beta}$ subject to the following equality and inequality constraints:

$$\boldsymbol{\beta} \in \mathbb{S}(\boldsymbol{f}, \boldsymbol{g}) = \{\boldsymbol{\beta}: \boldsymbol{f}(\boldsymbol{\beta}) = \mathbf{0}_r,\ \boldsymbol{g}(\boldsymbol{\beta}) \leqslant \mathbf{0}_s\}, \tag{2.2}$$

where both $\boldsymbol{f}(\boldsymbol{\beta}) = (f_1(\boldsymbol{\beta}), \ldots, f_r(\boldsymbol{\beta}))^\top$ and $\boldsymbol{g}(\boldsymbol{\beta}) = (g_1(\boldsymbol{\beta}), \ldots, g_s(\boldsymbol{\beta}))^\top$ are assumed to have continuous second-order partial derivatives, and functional constraints are consistent; i.e. $\mathbb{S}(\boldsymbol{f}, \boldsymbol{g})$ is a non-empty convex set. Note that if $r = 0$, then there is no equality constraint; if $s = 0$, there is no inequality constraint.