Statistics and Application 统计学与应用, 2021, 10(5), 864-874
Published Online October 2021 in Hans. http://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2021.105090

Hans汉斯

基于Cox回归模型的约束估计问题的研究

张露露，黄希芬*

云南师范大学数学学院，云南 昆明

收稿日期：2021年9月17日；录用日期：2021年10月2日；发布日期：2021年10月19日

摘 要

具有约束条件的右删失Cox回归模型的研究经常是基于偏似然函数，而该似然函数忽略了未知的基准风险率函数。文章致力于解决两个问题：基于完全似然函数得到回归参数和累积风险函数的极大似然估计；把有限制条件的优化问题转化为无限制条件的优化问题。ADMM算法能把限制条件引入到目标函数中，从而把有条件的优化问题转化为无条件的优化问题，MM算法在解决优化问题方面具有可分离参数的优点。因此，文章首先利用ADMM算法把有限制条件的优化问题转化为无限制条件的优化问题，然后将MM算法应用于极大化新的目标函数，实现了参数和非参数的分离，进而解决了回归参数和累积风险函数的非参数估计问题。同时，利用不等式放缩把高维优化问题转化为一维优化问题，避开了矩阵求逆的困难。

关键词

比例风险模型，完全似然函数，约束条件，MM算法，ADMM算法

The Constraint Estimation Based on Cox Regression Model

Lulu Zhang, Xifen Huang*

School of Mathematics, Yunnan Normal University, Kunming Yunnan

Received: Sep. 17th, 2021; accepted: Oct. 2nd, 2021; published: Oct. 19th, 2021

Abstract

The research on the right-censored Cox regression model with constraints is often based on the partial likelihood function, whereas the likelihood function ignores the unknown baseline hazard rate function. We devote to solving two problems. One is to obtain the maximum likelihood estimations of regression parameters and cumulative hazard function based on the complete likelih-

*通讯作者。

文章引用: 张露露, 黄希芬. 基于 Cox 回归模型的约束估计问题的研究[J]. 统计学与应用, 2021, 10(5): 864-874.
DOI: 10.12677/sa.2021.105090

ood function. Another is to transform the optimization problem with restrictions into the unrestricted optimization problem. The ADMM algorithm can introduce constraints into the objective function so as to transform the conditional optimization problem into an unconditional optimization problem. The MM algorithm has the advantage of separating parameters in solving optimization problems. Therefore, we first use the ADMM algorithm to transform the optimization problem with restrictions into an unconditional optimization problem. Then we apply the MM algorithm to maximize the new objective function in order to separate the parameters and the non-parameters, which is helpful to solve the problems of estimating the regression parameters and nonparametric cumulative hazard function. Meanwhile, the use of inequality which can transform the high-dimensional function into a sum of one-dimensional functions avoids the difficulty of matrix inversion.

Keywords

Proportional Hazard Model, Complete Likelihood Function, Constraint Conditions, MM Algorithm, ADMM Algorithm

1. 引言

Cox 回归模型自 1972 年被 Cox [1]首次提出后，由于该模型具有优良的理论性质，以及直观和解释性强的优点，已被广泛应用于生物医学、可靠性分析和经济领域等。Cox 模型是一种半参数生存分析模型，是分析生存数据最为广泛的统计模型之一。但是，在实际的生产实践中，由于实验期限已到、实验对象被移除或者提前退出实验等种种原因，导致观察收集到的生存数据往往又以右删失为主。因此，涌现了大量的科研工作者基于 Cox 回归模型对右删失数据进行分析研究。比如，王佳等[2]在其文章中基于 Cox 回归模型分析了三组医学类数据：乳腺癌数据、骨髓移植数据和原发性胆汁肝胆硬化数据。洪东跑等[3]证实了 Cox 回归模型可用于分析产品可靠性关于环境因素的动态变化特征。马超群和何文[4]基于 Cox 回归模型就上市公司财务困境问题进行判别能力和稳定性分析，等等。

在实际问题中，若我们提前知道有关 Cox 模型参数的一些信息，加以利用定会得到更加可靠的统计推断结果。然而，除了 Ding 等[5]，目前有关这方面的研究却比较少。但是，该文章是基于 Cox 模型的偏似然函数进行分析，而忽略了基准风险率函数的非参数估计问题，这显然具有一定的不合理性。因此，本文将基于 Cox 模型的完全似然函数，对具有限制条件的模型参数进行极大化估计。

显然，本文的难点之一在于极大化完全似然函数时，要考虑到限制条件对估计值的影响。幸而，ADMM 算法给出了一个较好的解决办法，即把有条件的优化问题转化为无条件的优化问题，进而避免了在求解过程中要考虑约束条件的限制。事实上，ADMM 算法已被广泛应用于解决各种带有约束条件的优化问题。例如，图像提取时的稀疏 lasso 和凸聚类等问题[6]，高维稀疏惩罚分位数回归模型[7]，基于凸函数的子组分析模型[8]，等等。这些文章都是运用 ADMM 算法把有限制条件的优化问题转化为无限制条件的优化问题，该算法提高了模型求解的速度。而且，ADMM 算法已在理论上被证明了其可靠性[9]。

再者，转化后的无条件优化问题没有解析解，故需要应用 Newton-Raphson 算法(以下简称“NR 算法”)进行迭代求解。但是，在面向高维数据集时，NR 算法需要在矩阵可逆的条件下才具有可行性。由于

minorization-maximization 算法(以下简称“MM 算法”)具有可分离参数的优点[10]，通过将目标函数进行参数分离，从而绕开了矩阵求逆问题。因此，本文将用两种形式的 MM 算法来解决基于 Cox 模型的完全似然函数的无条件优化问题，其主要区别在于处理基准风险率函数的非参数估计问题上的不同[11]。一种方法运用 profile 估计方法(Johansen [12])来估计基准风险率函数，另一种方法则是绕开了 profile 估计方法，直接应用 MM 算法把累积风险率从回归参数中分离出来。

本文剩余章节安排如下：第 2 节介绍受约束的 Cox 回归模型及其完全似然函数；第 3 节将 ADMM 算法和 MM 算法相结合，从而得到回归参数和非参数的累积风险函数的极大似然估计；第 4 节进行数值模拟，用于评估两种算法相结合的估计效果；第 5 节对一组癌症数据进行分析研究；第 6 节给出本文结论。

2. 受约束 Cox 回归模型

假设从某个总体中随机地抽取 n 个样本，并用 T_i， C_i 和 $X_i=\left(x_{i1},\cdots,x_{iq}\right)'$ 分别表示第 i 个样本的生存时间，删失时间以及 q 维解释变量值。进一步，假定在 X_i 给定的条件下，删失时间 C_i 独立于生存时间 T_i。从而，数据集 $Y_{obs}=\left\{\left(Y_i=\min\left(T_i,C_i\right),I_i,X_i\right),i=1,\cdots,n\right\}$，其中 Y_i， $I_i=I\left(T_i\le C_i\right)$ 和 $I\left(\cdot\right)$ 分别表示观测时间，右删失变量以及示性函数。因此，给定 X_i 的风险率函数为：

$$\lambda\left(t\mid X_i\right)=\lambda_0\left(t\right)\exp\left(X_i'\beta\right)$$

其中 $\lambda_0\left(\cdot\right)$ 是基准风险率函数， $\beta=\left(\beta_1,\cdots,\beta_q\right)'$ 是回归参数。

根据 Kim 等[13]，可得其完全似然函数为

$$L\left(\beta,\Lambda\mid Y_{obs}\right)=\prod_{i=1}^{n}\left(\lambda_0\left(t_i\right)\exp\left(X_i'\beta\right)\right)^{I_i}\exp\left(-\Lambda\left(t_i\right)\exp\left(X_i'\beta\right)\right)$$

其中 $\Lambda\left(\cdot\right)$ 为累积风险函数。因此，其对数似然函数为

$$l\left(\beta,\Lambda\mid Y_{obs}\right)=\sum_{i=1}^{n}\left(I_i\ln\left(\lambda_0\left(t_i\right)\right)+I_iX_i'\beta-\Lambda\left(t_i\right)\exp\left(X_i'\beta\right)\right)\tag{1}$$

我们感兴趣的是对受到某些条件限制的回归参数 β 的估计，比如边界限制条件或线性不等式的条件限制。在本文中，我们只考虑 $\beta\ge0$ 的约束条件，故本文最终考虑如下的基于完全似然函数的受约束 Cox 模型，即：

$$\max_{\beta\ge0}\quad l\left(\beta,\Lambda\mid Y_{obs}\right)\tag{2}$$

3. ADMM + MM 算法

根据 Boyd 等[9]在其文章的第 5 章中提出的理论依据，求解上述的受约束 Cox 模型等价于求解如下的模型：

$$\begin{aligned}\max\quad & l\left(\beta,\Lambda\mid Y_{obs}\right)+g\left(z\right)\\ & \beta-z=0\end{aligned}\tag{3}$$

其中 $g\left(z\right)=I\left(z\ge0\right)$。进而转化为极大化如下的无约束的目标函数：

$$l_\rho\left(\alpha\mid Y_{obs}\right)=l\left(\beta,\Lambda\mid Y_{obs}\right)+g\left(z\right)-\frac{\rho}{2}\left\|\beta-z+u\right\|_2^2\tag{4}$$

其中 $\alpha=\left(\beta,\Lambda,z,u\right)$，$\left\|\cdot\right\|_2$ 为向量的二范数。再根据 Boyd 等[12]的第 5 章的内容，可得极大化(4)式的 ADMM 形式为：

$$\left(\beta^{(k+1)},\Lambda^{(k+1)}\right)=\arg\max_{(\beta,\Lambda)}\left(l\left(\beta,\Lambda\mid Y_{obs}\right)-\frac{\rho}{2}\left\|\beta-z^{(k)}+u^{(k)}\right\|_2^2\right), \tag{5}$$

$$z^{(k+1)}=\left(\beta^{(k+1)}+u^{(k)}\right)_+, \tag{6}$$

$$u^{(k+1)}=u^{(k)}+\beta^{(k+1)}-z^{(k+1)}. \tag{7}$$

观察上述的迭代过程可知，难点在于求解(β,Λ)的极大似然估计。若直接利用 NR 算法，在高维情况下会涉及矩阵求逆问题，且非参数 Λ 的估计也是个棘手问题。幸而，近年发展迅速的 MM 算法为上述问题提供了解决办法。该算法的实现主要依赖于寻找到合适的不等式，进而通过一系列不等式放缩构造目标优化函数$l_\rho\left(\alpha\mid Y_{obs}\right)$的最小化函数(也叫替代函数) $Q\left(\alpha\mid\alpha^{(k)}\right)$，使其满足

$$\begin{cases} l_\rho\left(\alpha\mid Y_{obs}\right)\geq Q\left(\alpha\mid\alpha^{(k)}\right),\ \forall\alpha,\alpha^{(k)}\in\Theta \\ l_\rho\left(\alpha\mid Y_{obs}\right)=Q\left(\alpha^{(k)}\mid\alpha^{(k)}\right) \end{cases} \tag{8}$$

其中 $\alpha^{(k)}$ 为第 k 次对 $\hat{\alpha}$ 的近似值。然后，我们最大化替代函数，可得到如下迭代公式

$$\alpha^{(k+1)}=\arg\max_{\alpha\in\Theta}Q\left(\alpha\mid\alpha^{(k)}\right) \tag{9}$$

当第 $k+1$ 次逼近 $\hat{\alpha}$ 时，有

$$l_\rho\left(\alpha^{(k+1)}\mid Y_{obs}\right)\geq Q\left(\alpha^{(k+1)}\mid\alpha^{(k)}\right)\geq Q\left(\alpha^{(k)}\mid\alpha^{(k)}\right)=l_\rho\left(\alpha^{(k)}\mid Y_{obs}\right) \tag{10}$$

其中迭代过程(9)呈现上升趋势，使得目标函数不断增大，从而实现优化转移，即极大化 $Q\left(\cdot\mid\alpha^{(k)}\right)$等价于极大化$l_\rho\left(\cdot\mid Y_{obs}\right)$。

使用 MM 算法时，关键且困难的一步是寻找合适的替代函数。在实际过程中，尤其遇到高维的情况，构建一个参数可分离的替代函数尤为重要，因为这样可以避免矩阵求逆的困难。而本文从 Huang 等[11]处得到启发，利用 MM 算法把参数 β 和非参数 Λ 进行参数分离，且进一步把 β 的各个分量进行分离，从而避免了矩阵求逆问题。

3.1. ADMM + profile MM 算法

如上所述，MM 算法的关键是构造目标函数$l\left(\beta,\Lambda\mid Y_{obs}\right)-\frac{\rho}{2}\left\|\beta-z^{(k)}+u^{(k)}\right\|_2^2$的替代函数。

首先，观察 $l^A\left(\beta,\Lambda\mid\alpha^{(k)}\right)=l\left(\beta,\Lambda\mid Y_{obs}\right)-\frac{\rho}{2}\left\|\beta-z^{(k)}+u^{(k)}\right\|_2^2$ 可得，Johansen [12]在其文章中用到的 profile 方法可直接用来计算非参数 Λ，即令 $\partial l^A\left(\beta,\Lambda\mid\alpha^{(k)}\right)\Big/\partial\lambda_0\left(t_i\right)=0$ 可得

$$\lambda_0\left(t_i\right)=\frac{I_i}{\sum_{r=1}^{n}I\left(t_r\geq t_i\right)\exp\left(X_r'\beta\right)},i=1,\cdots,n \tag{11}$$

然后把(11)式带入 $l^A\left(\beta,\Lambda\mid\alpha^{(k)}\right)$，并把与参数 β 无关的部分忽略不计，进而整理得到函数

$$Q_1\left(\beta\mid\alpha^{(k)}\right)=\sum_{i=1}^{n}\left(I_iX_i'\beta-I_i\ln\sum_{r=1}^{n}I\left(t_r\geq t_i\right)\exp\left(X_r'\beta\right)\right)-\frac{\rho}{2}\left\|\beta-z^{(k)}+u^{(k)}\right\|_2^2$$

接下来，把 $Q_1\left(\beta\mid\alpha^{(k)}\right)$看作新的目标函数，对 β 的各个分量进行参数分离。根据支撑超平面不等式

$$-\ln x\geq-\ln x_0-\frac{x-x_0}{x_0}$$

可得

$$-\ln\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta\right)$$

$$\geq -\ln\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta^{(k)}\right) - \frac{\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta\right) - \sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta^{(k)}\right)}{\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta^{(k)}\right)}$$

从而可得 $Q_1\left(\beta \mid \alpha^{(k)}\right)$ 的替代函数

$$Q_{12}\left(\beta \mid \alpha^{(k)}\right) = \sum_{i=1}^{n}\sum_{p=1}^{q}\left(I_i x_{ip}\beta_p - \frac{I_i\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta\right)}{\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta^{(k)}\right)}\right) - \frac{\rho}{2}\left\|\beta - z^{(k)} + u^{(k)}\right\|_2^2$$

然后，基于 $X_i^{\mathrm{T}}\beta$ 构造权重 $\omega_{ip} = \left|x_{ip}\right| \Big/ \sum_{p=1}^{q}\left|x_{ip}\right|$，再把 $X_i^{\mathrm{T}}\beta$ 改写为

$$X_i^{\mathrm{T}}\beta = \sum_{p=1}^{q}\omega_{ip}\left[\omega_{ip}^{-1}x_{ip}\left(\beta_p - \beta_p^{(k)}\right) + X_i^{\mathrm{T}}\beta^{(k)}\right]$$

如果 $x_{ip} = 0$，则令 $1/\omega_{ip} = 0$。然后根据离散型 Jensen 不等式

$$\varphi\left(\sum_{i=1}^{n} a_i x_i\right) \geq \sum_{i=1}^{n} a_i\varphi\left(x_i\right)，\text{其中 } a_i \geq 0，\text{且 } \sum_{i=1}^{n} a_i = 1$$

得到 $Q_{12}\left(\beta \mid \alpha^{(k)}\right)$ 的替代函数

$$Q_{13}\left(\beta \mid \alpha^{(k)}\right) \triangleq \sum_{p=1}^{q} Q_{13p}\left(\beta_p \mid \alpha^{(k)}\right) \tag{12}$$

其中

$$Q_{13p}\left(\beta_p \mid \alpha^{(k)}\right) = \sum_{i=1}^{n}\left(I_i x_{ip}\beta_p - \frac{I_i\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\omega_{rp}\exp\left(\omega_{rp}^{-1}x_{rp}\left(\beta_p - \beta_p^{(k)}\right) + X_r'\beta^{(k)}\right)}{\sum_{r=1}^{n} I\left(t_r \geq t_i\right)\exp\left(X_r'\beta^{(k)}\right)}\right) - \frac{\rho}{2}\left(\beta_p - z_p^{(k)} + u_p^{(k)}\right)^2 \tag{13}$$

从(12)和(13)式可以看出，目标函数 $l^A\left(\beta, \Lambda \mid \alpha^{(k)}\right)$ 已经被分解成 q 个单变量函数之和，实现了把高维优化问题转化为低维优化问题。所得的 MM 算法在其最大化步骤中仅涉及 q 个单独的单变量优化问题，因此不需要矩阵求逆。令 $\mathrm{d}Q_{13p}\left(\beta_p \mid \alpha^{(k)}\right) \Big/ \mathrm{d}\beta_p = 0$，再利用 NR 算法进行迭代求解。

3.2. ADMM + non-profile MM 算法

本节提出一种绕过 profile 方法的 non-profile MM 算法。

首先，根据算术几何不等式

$$-\prod_{i=1}^{n} x_i^{a_i} \geq -\sum_{i=1}^{n}\frac{a_i}{\|a\|_1}x_i^{\|a\|_1}$$

其中 $\|\cdot\|_1$ 为向量的一范数。可得 $l^A\left(\beta, \Lambda \mid \alpha^{(k)}\right)$ 的替代函数

$$Q_2\left(\alpha \mid \alpha^{(k)}\right) \triangleq Q_{21}\left(\Lambda \mid \alpha^{(k)}\right) + Q_{22}\left(\beta \mid \alpha^{(k)}\right)$$

其中

$$Q_{21}\left(\Lambda\mid\alpha^{(k)}\right)=\sum_{i=1}^{n}\left(I_i\ln\left(\lambda_0\left(t_i\right)\right)-\frac{\exp\left(X_i'\beta^{(k)}\right)}{2\Lambda^{(k)}\left(t_i\right)}\Lambda^2\left(t_i\right)\right)$$

$$Q_{22}\left(\beta\mid\alpha^{(k)}\right)=\sum_{i=1}^{n}\left(I_iX_i'\beta-\frac{\Lambda^{(k)}\left(t_i\right)}{2\exp\left(X_i'\beta^{(k)}\right)}\exp\left(2X_i'\beta\right)\right)-\frac{\rho}{2}\left\|\beta-z^{(k)}+u^{(k)}\right\|_2^2$$

为了得到非参数 Λ 的估计，可极大化 $Q_{21}\left(\Lambda\mid\alpha^{(k)}\right)$，即令 $\mathrm{d}Q_{21}\left(\Lambda\mid\alpha^{(k)}\right)\Big/\mathrm{d}\lambda_0\left(t_i\right)=0$。并进一步整理可得 $\lambda_0\left(t_i\right)$ 的估计

$$\lambda_0\left(t_i\right)=\frac{I_i}{\sum_{r=1}^{n}I\left(t_r\geq t_i\right)\exp\left(X_r'\beta^{(k)}\right)}\tag{14}$$

类似于上一节的方法技巧，对 β 的各个分量进行参数分离。首先把 $2X_i^{\mathrm{T}}\beta$ 改写为 $2X_i^{\mathrm{T}}\beta=\sum_{p=1}^{q}\omega_{ip}\left[2\omega_{ip}^{-1}x_{ip}\left(\beta_p-\beta_p^{(k)}\right)+2X_i^{\mathrm{T}}\beta^{(k)}\right]$，然后把离散型 Jensen 不等式运用到函数 $-\exp(\cdot)$，从而得 $Q_{22}\left(\beta\mid\alpha^{(k)}\right)$ 的替代函数为

$$Q_{23}\left(\beta\mid\alpha^{(k)}\right)\triangleq\sum_{p=1}^{q}Q_{23p}\left(\beta_p\mid\alpha^{(k)}\right)\tag{15}$$

其中

$$Q_{23p}\left(\beta_p\mid\alpha^{(k)}\right)=\sum_{i=1}^{n}\left(I_iX_{ip}\beta_p-\frac{\Lambda^{(k)}\left(t_i\right)\omega_{ip}}{2\exp\left(X_i'\beta^{(k)}\right)}\exp\left(2\omega_{ip}^{-1}x_{ip}\left(\beta_p-\beta_p^{(k)}\right)+2X_i'\beta^{(k)}\right)\right)-\frac{\rho}{2}\left(\beta_p-z_p^{(k)}+u_p^{(k)}\right)^2\tag{16}$$

观察(15)和(16)式可以看出，我们得到了与(12)和(13)式类似的结果，即目标函数已经实现参数分离，在接下来的最大化步骤中绕开了矩阵求逆。同理，令 $\mathrm{d}Q_{23p}\left(\beta_p\mid\alpha^{(k)}\right)\Big/\mathrm{d}\beta_p=0$，再利用 NR 算法进行迭代求解即可。

综上所述，我们得到如下的迭代算法：

第一步：给定初始值 $\left(\beta^{(0)},\Lambda^{(0)},z^{(0)},u^{(0)}\right)$；

第二步：基于式(11)或式(14)计算 $\Lambda^{(k+1)},k=0,1,2,\cdots$；

第三步：对应于第二步，利用 NR 算法对式(13)或式(16)进行迭代求解 $\beta_p^{(k+1)}$，其中 $p=1,\cdots,q;k=0,1,2,\cdots$；

第四步：基于式(6)计算 $z^{(k+1)},k=0,1,2,\cdots$；

第五步：基于式(7)计算 $u^{(k+1)},k=0,1,2,\cdots$；

第六步：重复第二到五步，直到算法收敛。

4. 数值模拟

模拟研究设定为：参数向量 $\beta=(1,2)'$，则 $q=2$，且 q 个解释变量独立同分布，其数值由正态分布 $N(-1,1)$ 随机产生。基准风险率函数设置为 $\lambda_0\left(t\right)=2$，则对应的累积风险函数 $\Lambda\left(t\right)=2t$。在整个模拟实验中，我们通过控制样本量 n 和删失率 r 来评估 Cox 回归模型分别在无约束和有约束的情况下，对回归参数和累积风险率 $\left(\Lambda\left(0.1\right),\Lambda\left(0.5\right)\right)=\left(0.2,1\right)$ 的估计效果。其中，样本量 n 分别取 50 和 100；删失率 r 分别达到 20%，50%以及 70%。

整个模拟由 R 软件实现，所有实验的收敛条件均为

$$\frac{\left|l\left(\beta^{(k+1)}\mid Y_{obs}\right)-l\left(\beta^{(k)}\mid Y_{obs}\right)\right|}{\left|l\left(\beta^{(k)}\mid Y_{obs}\right)\right|+1}<10^{-8}$$

每种情况都进行 $R=500$ 次的模拟研究。具体模拟结果见表 1 和表 2。

Table 1. The simulations of the unconstrained and constrained Cox model under the profile MM algorithm
表 1. profile MM 算法下的无约束条件和有约束条件的 Cox 模型的模拟结果

			无约束条件			有约束条件		
n	r		Bias	S.D.	MSE	Bias	S.D.	MSE
50	0.20	β_1	0.0315	0.2361	0.0566	0.0254	0.2270	0.0521
		β_2	0.0869	0.3437	0.1255	0.0446	0.3150	0.1010
		$\Lambda(0.1)$	0.0101	0.1435	0.0206	0.0050	0.1239	0.0153
		$\Lambda(0.5)$	0.0881	0.5038	0.2611	0.0377	0.4485	0.2022
	0.50	β_1	0.0520	0.2988	0.0918	0.0504	0.2787	0.0801
		β_2	0.1362	0.4220	0.1963	0.1025	0.3960	0.1670
		$\Lambda(0.1)$	0.0015	0.1440	0.0207	0.0009	0.1300	0.0169
		$\Lambda(0.5)$	0.0922	0.5647	0.3268	0.0574	0.4833	0.2364
	0.70	β_1	0.1361	0.4400	0.2117	0.0614	0.3829	0.1501
		β_2	0.2476	0.6531	0.4870	0.1660	0.5684	0.3499
		$\Lambda(0.1)$	0.0079	0.1413	0.0200	0.0073	0.1361	0.0186
		$\Lambda(0.5)$	0.1486	0.8221	0.6966	0.0277	0.4888	0.2392
100	0.20	β_1	0.0147	0.1521	0.0233	0.0106	0.1563	0.0245
		β_2	0.0155	0.2188	0.0480	0.0174	0.2122	0.0453
		$\Lambda(0.1)$	0.0054	0.0919	0.0085	0.0049	0.0911	0.0083
		$\Lambda(0.5)$	0.0501	0.3311	0.1119	0.0158	0.2719	0.0740
	0.50	β_1	0.0417	0.2018	0.0424	0.0296	0.1886	0.0364
		β_2	0.0828	0.2582	0.0734	0.0715	0.2541	0.0695
		$\Lambda(0.1)$	0.0002	0.0969	0.0094	−0.0012	0.0903	0.0081
		$\Lambda(0.5)$	0.0591	0.3438	0.1215	0.0434	0.3289	0.1099
	0.70	β_1	0.0535	0.2643	0.0726	0.0285	0.2367	0.0568
		β_2	0.0874	0.3189	0.1092	0.0553	0.3081	0.0978
		$\Lambda(0.1)$	−0.0037	0.0966	0.0093	0.0034	0.0843	0.0071
		$\Lambda(0.5)$	0.0506	0.3670	0.1370	0.0431	0.3499	0.1240

表 1 是基于 profile MM 算法进行 500 次模拟研究得到的结果，其中有约束条件的 Cox 模型的极大化估计则结合了 ADMM 算法，具体过程见 3.1。表格中的数据来源于对 500 次模拟得到的估计值与真值的

偏差 BIAS 和均方误差 MSE 求平均，以及估计值的样本差 S.D.。从表 1 可以看出，无论样本量和删失率达到多少，对比 BIAS 数值可以看出，考虑受约束条件 $\beta \geq 0$ 的模拟结果比无约束条件的模拟结果总体上更接近真值。而在约束条件下得到的 S.D.和 MSE 数均比无约束条件下的小，故提前知道回归参数的部分信息有助于得到更可靠的估计值。其次，MM 算法和 ADMM + MM 算法在估计回归参数 β 和非参数 Λ 方面都表现出良好的估计效果。再者，通过对比观察可知，固定删失率，增加样本量，两种情况下模拟得到的估计值与真值之间的偏差越来越小，RMSE 和 S.D.的值也越来越小，即估计效果和模拟的稳定性越来越好。固定样本量，当删失率越小，评价指标 BIAS、RMSE 和 S.D.的值也越来越小，这显然符合实际情况。

表 2 是基于 non-profile MM 算法的模拟结果，其中有约束条件的 Cox 模型的极大化估计则结合了 ADMM 算法，具体过程见 3.2。其数据来源于对 500 次模拟得到的估计值与真值的偏差 BIAS 和 MSE 求平均，以及估计值的样本差 S.D.。对比表 2 中的 BIAS 数值，有约束条件下的估计值比无约束条件下的估计值更接近真值，且 S.D.和 MSE 数值也相对更小。这表示极大化有约束条件的 Cox 模型的完全似然函数更能得到可靠的结果。类似地可以看出，MM 算法和 ADMM + MM 算法在估计回归参数 β 和非参数 Λ 方面都表现出良好的估计效果。以及减小删失率或增加样本量，都使得估计值与真值之间的差异性越来越小，模拟效果越来越好。

Table 2. The simulations of the unconstrained and constrained Cox model under the non-profile MM algorithm
表 2. non-profile MM 算法下的无约束条件和有约束条件的 Cox 模型的模拟结果

			无约束条件			有约束条件		
n	r		Bias	S.D.	MSE	Bias	S.D.	MSE
50	0.20	β_1	0.0422	0.2416	0.0600	0.0395	0.2391	0.0586
		β_2	0.0844	0.3534	0.1318	0.0801	0.3221	0.1099
		$\Lambda(0.1)$	0.0016	0.1446	0.0209	0.0002	0.1384	0.0191
		$\Lambda(0.5)$	0.0652	0.5157	0.2696	0.0451	0.5109	0.2626
	0.50	β_1	0.0870	0.3176	0.1083	0.0414	0.2850	0.0828
		β_2	0.1525	0.4884	0.2613	0.0824	0.4008	0.1671
		$\Lambda(0.1)$	0.0054	0.1681	0.0282	−0.0006	0.1395	0.0194
		$\Lambda(0.5)$	0.0935	0.5781	0.3423	0.0364	0.4941	0.2450
	0.70	β_1	0.0954	0.5411	0.3013	0.0910	0.3982	0.1665
		β_2	0.2340	0.9259	0.9104	0.2008	0.6006	0.4003
		$\Lambda(0.1)$	−0.0033	0.1602	0.0256	−0.0016	0.1488	0.0221
		$\Lambda(0.5)$	0.0885	0.6466	0.4251	0.0841	0.6061	0.3736
100	0.20	β_1	0.0166	0.1501	0.0228	0.0120	0.1485	0.0222
		β_2	0.0306	0.2160	0.0475	0.0111	0.2066	0.0427
		$\Lambda(0.1)$	−0.0015	0.0933	0.0087	−0.0049	0.0811	0.0066
		$\Lambda(0.5)$	0.0173	0.2986	0.0893	0.0026	0.2701	0.0728

Continued

100	0.50	β_1	0.0250	0.1877	0.0358	0.0269	0.1839	0.0345
		β_2	0.0479	0.2720	0.0761	0.0487	0.2564	0.0680
		$\Lambda(0.1)$	0.0018	0.0954	0.0091	0.0026	0.0920	0.0085
		$\Lambda(0.5)$	0.0206	0.3121	0.0977	0.0047	0.3095	0.0956
	0.70	β_1	0.0473	0.2329	0.0564	0.0346	0.2300	0.0540
		β_2	0.1094	0.3300	0.1207	0.0744	0.3281	0.1129
		$\Lambda(0.1)$	0.0025	0.0944	0.0089	−0.0056	0.0912	0.0083
		$\Lambda(0.5)$	0.0515	0.3478	0.1233	0.0211	0.3348	0.1123

综上所述，无论样本量 n 取 50 还是 100，删失率 r 为 20%，50%还是 70%，极大化有约束条件的 Cox 回归模型的完全似然函数时，都表现出更好的估计效果。特别地，观察本实验中最不理想的组合 $(n,r)=(50,70\%)$，其模拟结果依然是令人满意的。由此可见，ADMM 算法和 MM 算法相结合的效果良好。进一步，该模拟结果告诉我们，可以通过增加样本量、减少删失率来改善估计效果，使其更加接近实际。

5. 实例分析

安装 R 语言中的 survival 程序包，调用内置的 cancer 数据集。该数据集共记录了 228 个来自中北部癌症治疗组的癌症晚期患者的相关数据[14]，其中包括机构代码、生存天数、患者生存状态(死亡或者实验数据删失)、年龄、性别、由医生评定的 GCOG 表现评分、医生评定的 Karnofsky 表现评分、由患者自己评定的 Karnofsky 表现评分、进食时消耗的卡路里以及最近 6 个月的体重下降数。

在本例中，我们只研究患者最近 6 个月的体重下降数与死亡风险率之间的关系，即回归变量 X 只取一维。从而，建立如下的 Cox 回归模型：

$$\lambda(t\mid X_i)=\lambda_0(t)\exp(X_i\beta)$$

然后根据对以上模型进行参数估计的结果可知(见表 3 中无约束条件下的估计结果)，本文可对回归参数做如下限制：

$$\beta\geq 0\text{。}$$

从而，针对有限制条件的 Cox 模型的估计，可用本文提出的 ADMM + MM (profile 和 non-profile)算法进行回归参数估计和累积风险函数的估计。由于该模型具有大样本性质，故可用 bootstrap 方法[5]求得样本估计值的标准差和 95%置信区间。最终得到如表 3 所示的数值分析结果。从表 3 中可以看出，两种 ADMM + MM 算法的估计结果在精确到小数点三位时是一致的，该结果表明最近 6 个月患者的体重下降越多，越会增加患者的死亡率，且增加到 $e^{0.015}\approx 1.02$ 倍，但在有约束条件下的 95%置信区间的区间长度更小，这表明该估计结果更可靠。因此，我们只画出有约束条件的模型的估计累积风险函数，如图 1 所示，两种算法下的估计累积风险函数的数值相差不大，走势一致，这表明两种 ADMM + MM 算法表现相当。

Table 3. The estimations by two proposed ADMM + MM algorithms and the confidence intervals based on bootstrap method
表 3. 两种 ADMM + MM 算法的估计结果以及 bootstrap 方法的置信区间

	Profile MM 算法	ADMM + profile MM 算法
	无约束条件	有约束条件

Continued

变量	估计值	标准差	95%置信区间	估计值	标准差	95%置信区间
免疫过氧化物酶	0.0148	0.0876	(−0.1569, 0.1865)	0.0148	0.0600	(−0.1029, 0.1325)
	Non-profile MM 算法			ADMM + non-profile MM 算法		
	无约束条件			有约束条件		
变量	估计值	标准差	95%置信区间	估计值	标准差	95%置信区间
免疫过氧化物酶	0.0146	0.0884	(−0.1586, 0.1879)	0.0146	0.0601	(−0.1031, 0.1324)

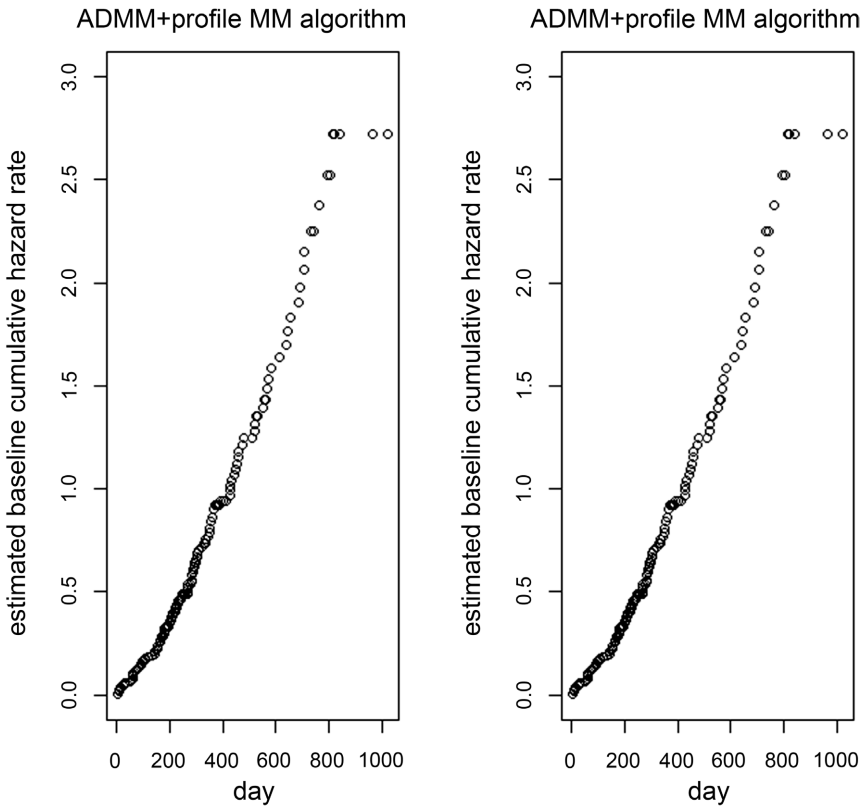

Figure 1. The estimated cumulative hazard function for patients
图 1. 患者的累积死亡风险估计函数

6. 结论

正如引言所言，若提前知道部分有关未知参数的信息，加以利用必会得到更可靠的估计结果。而针对有约束条件的 Cox 回归模型的研究很少，而 Ding 等的研究却是基于该模型的偏似然函数，忽略非参数部分的估计，则无法预知个体即将面临的死亡等风险率，这显然是不合理的。再者，基于该模型的完全似然函数的有约束条件的研究却不曾有过，因此，本文基于右删失型的 Cox 比例风险模型，在约束条件为 $\beta \geq 0$ 的情况下极大化该模型的完全似然函数，从而得到回归参数估计和非参数估计。

针对有约束条件的 Cox 模型，由于 ADMM 算法在有约束条件优化问题方面具有优良的理论性质而得到快速发展，本文应用 ADMM 算法把基于右删失 Cox 模型的有约束条件优化问题转化为无条件优化问题，从而减低了极大化其完全似然函数的难度。再者，极大化该完全似然函数时没有解析解，且涉及到非参数估计，因此我们把具有参数分离特点的 MM 算法应用到极大化过程中，该算法把非参数部分从

回归参数中分离出来，从而逐一击破。此外，在处理高维数据集时，MM 算法可以把回归参数彼此分离，实现把高维优化问题转化为低维优化问题，从而在 Newton-Raphson 迭代过程中避免了矩阵求逆。模拟实验的估计结果表明，有约束条件的模型估计结果更可靠，且进一步表明 ADMM 和 MM 算法相结合具有良好的收敛性，且通过增加样本量、减少删失率，可以改善估计结果。而对真实数据集，即乳腺癌数据的分析，则进一步表明考虑约束条件对结果的估计是更可靠的。

参考文献

[1] 王佳, 丁洁丽, 沈静. 比例风险模型下参数估计的两种 MM 算法的一些应用[J]. 数理统计与管理, 2016, 35(4): 649-661.

[2] 洪东跑, 赵宇, 马小兵. 基于比例风险模型的可靠性灵敏度分析[J].宇航学报, 2011, 32(8): 1865-1870.

[3] 马超群, 何文. 基于 Cox 的财务困境时点预测模型研究[J]. 统计与决策, 2010(21): 38-42.

[4] Ding, J.L., Tian, G.L. and Yuen, K.C. (2015) A New MM Algorithm for Constrained Estimation in the Proportional Hazards Model. *Computational Statistics and Data Analysis*, **84**, 135-151. https://doi.org/10.1016/j.csda.2014.11.005

[5] Zhu, Y.Z. (2017) An Augmented ADMM Algorithm with Application to the Generalized Lasso Problem. *Journal of Computational and Graphical Statistics*, **26**, 195-204. https://doi.org/10.1080/10618600.2015.1114491

[6] Gu, Y.W., Fan, J., Kong, L.C., Ma, S.Q. and Zou, H. (2018) ADMM for High-Dimensional Sparse Penalized Quantile Regression. *Technometrics*, **60**, 319-331. https://doi.org/10.1080/00401706.2017.1345703

[7] Ma, S.J. and Huang, J. (2017) A Concave Pairwise Fusion Approach to Subgroup Analysis. *Journal of the American Statistical Association*, **112**, 410-423. https://doi.org/10.1080/01621459.2016.1148039

[8] Boyd, S., Parikh, N. and Chu, E. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Now Publishers Inc., Massachusetts, 33-37. https://doi.org/10.1561/9781601984616

[9] Lange, K., Hunter, D.R. and Yang, I. (2000) Optimization Transfer Using Surrogate Objective Functions. *Journal of Computational and Graphical Statistics*, **9**, 1-20. https://doi.org/10.1080/10618600.2000.10474858

[10] Huang, X.F, Xu, J.F. and Tian, G.L. (2019) On Profile MM Algorithms for Gamma Frailty Survival Models. *Statistica Sinica*, **29**, 895-916. https://doi.org/10.5705/ss.202016.0516

[11] Johansen, S. (1983) An Extension of Cox's Regression Model. *International Statistical Review*, **51**, 165-174. https://doi.org/10.2307/1402746

[12] Kim, Y., Kim, B. and Jang, W. (2010) Asymptotic Properties of The Maximum Likelihood Estimator for the Proportional Hazards Model with Doubly Censored Data. *Journal of Multivariate Analysis*, **101**, 1339-1351. https://doi.org/10.1016/j.jmva.2010.01.010

[13] Loprinzi, C.L., Laurie, J.A., Wieand, H.S., Krook, J.E., Novotny, P.J., *et.al.* (1994) Prospective Evaluation of Prognostic Variables from Patient-Completed Questionnaires. North Central Cancer Treatment Group. *Journal of Clinical Oncology*, **12**, 601-607. https://doi.org/10.1200/JCO.1994.12.3.601

[14] Malik, A.S., Boyko, O., Atkar, N. and Young, W.F. (201) A Comparative Study of MR Imaging Profile of Titanium Pedicle Screws. *Acta Radiologica*, **42**, 291-293. https://doi.org/10.1080/028418501127346846