Correction to “Variable selection in semiparametric linear regression with censored data,” JRSS-B, **70**(2) 351–370.

Address for correspondence:

Brent A. Johnson

Department of Biostatistics

Rollins School of Public Health

Emory University

1518 Clifton Rd., NE

Atlanta, GA 30322

U. S. A.

Email: bajohn3@emory.edu

I thank Professor Howard Bondell, North Carolina State University, for alerting me to mistakes in the proof provided in Appendix B. The chief concern is whether the penalised estimating function yields a sparse solution; that is, can the solution $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_d)^{\mathrm{T}}$ to $0 = \mathbf{U}^P(\boldsymbol{\beta})$ ever possess an element $\widehat{\beta}_j = 0$? The answer is found using a more general and careful definition of "solution" to the estimating equation.

First, define the true active set $\mathcal{A} = \{j : \beta_{0j} \neq 0\}$ and the sample active set $\mathcal{A}_n = \{j : \widehat{\beta}_j \neq 0\}$. Note, the original article assumes that, without loss of generality, the first s covariables are active, i.e. $\mathcal{A} = \{1, \ldots, s\}$. Second, partition the estimate $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_{\mathcal{A}}^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$, where $\widehat{\boldsymbol{\beta}}_{\mathcal{A}}$ pertains to the s coefficient estimates on the active set and $\mathbf{0}$ is a $(d-s)$-vector of zeros; similarly, partition the vector of true coefficients $\boldsymbol{\beta}_0 = (\boldsymbol{\beta}_{\mathcal{A}}^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$. The goal is to show that $0 \approx \mathbf{U}^P(\widehat{\boldsymbol{\beta}})$ in the sense that $\widehat{\boldsymbol{\beta}}$ is a zero-crossing of the estimating equations. To define zero-crossing, adopt the short-hand notation,

$$U_j^P(\widehat{\boldsymbol{\beta}}+) \cdot U_j^P(\widehat{\boldsymbol{\beta}}-) = \lim_{\tau \to 0+} U_j^P(\widehat{\boldsymbol{\beta}} + \tau \mathbf{u}_j) \cdot U_j^P(\widehat{\boldsymbol{\beta}} - \tau \mathbf{u}_j),$$

where $\mathbf{u}_j$ is the j-th canonical unit vector and $\mathbf{U}^P = (U_1^P, \ldots, U_d^P)^{\mathrm{T}}$. Then, a zero-crossing $\widehat{\boldsymbol{\beta}}$ of the penalised estimating equations is given through the element-wise product, $U_j^P(\widehat{\boldsymbol{\beta}}+) \cdot U_j^P(\widehat{\boldsymbol{\beta}}-) \leq 0$ for $j = 1, \ldots, d$. When the estimating function $\mathbf{U}^P$ pertains to the d-dimensional gradient of a penalised loss function, then the new definition of solution agrees with the Karush-Kuhn-Tucker conditions. Namely, $U_j^P(\widehat{\boldsymbol{\beta}}+) \cdot U_j^P(\widehat{\boldsymbol{\beta}}-) = 0$ for $j \in \mathcal{A}$ and $U_j^P(\widehat{\boldsymbol{\beta}}+) \cdot U_j^P(\widehat{\boldsymbol{\beta}}-) < 0$ for $j \notin \mathcal{A}$; note, the latter implies there is a sign change at zero on the inactive set. Thus, the coefficient estimate $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_{\mathcal{A}}^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$ satisfies $U_j^P(\widehat{\boldsymbol{\beta}}) = 0$ for all $j \in \mathcal{A}$ and $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + O_p(n^{-1/2})$.

After adopting the partitioned form of coefficient estimate $\widehat{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_{\mathcal{A}}^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}}$, the remaining portions of the proof are corrected by restricting one's attention to asymptotic behaviour on the active set. That is, replace $\widehat{\boldsymbol{\beta}}$, $\boldsymbol{\beta}_0$, and $\mathbf{A}^{\mathrm{T}}\mathbf{U}^P(\boldsymbol{\beta})$ with $\widehat{\boldsymbol{\beta}}_{\mathcal{A}}$, $\boldsymbol{\beta}_{\mathcal{A}}$, and $\mathbf{A}_{\mathcal{A}}^{\mathrm{T}}\mathbf{U}_{\mathcal{A}}^P(\boldsymbol{\beta})$, respectively, where $\mathbf{A}$ pertains to the d-dimensional asymptotic slope matrix of $\mathbf{U}$ and $\mathbf{A}_{\mathcal{A}}$ is the s-dimensional active subset of $\mathbf{A}$. The rest of the proof follows.

Acknowledgements. I gratefully acknowledge Professor Wood for his assistance with this correction as well as the original article. I also acknowledge the intellectual contributions of Donglin Zeng who, as I say in the original article, contributed significantly to the Theorem 1 and its proof. Finally, I thank Danyu Lin, one of three postdoc mentors, for generously providing general insight in censored data theory and applications.