Supplementary materials for Analysis of Cohort Studies with Multivariate, Partially Observed, Disease Classification Data

By Nilanjan Chatterjee
Division of Cancer Epidemiology and Genetics,
National Cancer Institute, NIH, DHHS. Rockville, MD 20852, USA.
chattern@mail.nih.gov

Samiran Sinha
Texas A&M University, College Station, TX 77843, USA.
sinha@stat.tamu.edu

W. Ryan Diver and Heather Spencer Feigelson
Department of Epidemiology and Surveillance Research,
American Cancer Society, Atlanta, GA 30303, USA.

Derivation of derivatives of the estimating function

Define

$$
\begin{aligned}
S^{(2)}_{\theta\theta}(V_i, y^{o_r}) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{y_i^{o_r}}(\mathcal{B}_Y\mathcal{B}_Y^\top|X_j)\otimes X_jX_j^\top\omega_{y_i^{o_r}}(X_j),\\
S^{(2)}_{\xi\xi}(V_i, u) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{u}(\mathcal{A}_Y\mathcal{A}_Y^\top|X_j)\omega_{u}(X_j),\\
S^{(2)}_{\theta\xi}(V_i, y^{o_r}) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{y_i^{o_r}}(\mathcal{B}_Y\mathcal{A}_Y^\top|X_j)\otimes X_j\omega_{y_i^{o_r}}(X_j),\\
S^{(2)}_{\xi\theta}(V_i, u) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{u}(\mathcal{A}_Y\mathcal{B}_Y^\top|X_j)\otimes X_j\omega_{u}(X_j),\\
S^{(1)}_{\xi}(V_i, y^{o_r}) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{y_i^{o_r}}(\mathcal{A}_Y|X_j)\omega_{y_i^{o_r}}(X_j),\\
S^{(1)}_{\theta}(V_i, u) &= \frac{1}{n}\sum_{j=1}^{n} I(V_j \geq V_i)\mathcal{E}_{u}(\mathcal{B}_Y|X_j)\otimes X_j\omega_{u}(X_j),
\end{aligned}
$$

and let $s^{(2)}_{\theta\theta}(V_i, y^{o_r})$, $s^{(2)}_{\xi\xi}(V_i, u)$,$s^{(2)}_{\theta\xi}(V_i, y^{o_r})$,$s^{(2)}_{\xi\theta}(V_i, u)$, $s^{(1)}_{\xi}(V_i, y^{o_r})$ and $s^{(1)}_{\theta}(V_i, u)$ denote the corresponding population expectations. Now we can write

$$
\frac{\partial S_\theta}{\partial\theta^T} = \sum_r \sum_{\Delta_i=1, R_i=r}\left[\mathcal{V}_{y_i^{o_r}}(\mathcal{B}_Y|X_i)\otimes X_iX_i^\top - \frac{S^{(2)}_{\theta\theta}(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})} + \frac{S^{(1)}_{\theta}(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})}\left\{\frac{S^{(1)}_{\theta}(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})}\right\}^\top\right],
$$

$$
\begin{aligned}
\frac{\partial S_\theta}{\partial \xi^T} &= \sum_r \sum_{\Delta_i=1,R_i=r} \left[\mathcal{C}_{y_i^{o_r}}(\mathcal{B}_Y, \mathcal{A}_Y|X_i) \otimes X_i - \frac{S^{(2)}_{\theta\xi}(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})} + \frac{S^{(1)}_\theta(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})} \left\{ \frac{S^{(1)}_\xi(V_i, y^{o_r})}{S^{(0)}(V_i, y^{o_r})} \right\}^\top \right], \\
\frac{\partial S_\xi}{\partial \xi^T} &= \sum_r \sum_{\Delta_i=1,R_i=r} \left[\mathcal{V}_{y_i^{o_r}}(\mathcal{A}_Y|X_i) - \frac{S^{(2)}_{\xi\xi}(V_i, u)}{S^{(0)}(V_i, u)} + \frac{S^{(1)}_\xi(V_i, u)}{S^{(0)}(V_i, u)} \left\{ \frac{S^{(1)}_\xi(V_i, u)}{S^{(0)}_\xi(V_i, u)} \right\}^\top \right], \\
\frac{\partial S_\xi}{\partial \theta^T} &= \sum_r \sum_{\Delta_i=1,R_i=r} \left[\mathcal{C}_{y_i^{o_r}}(\mathcal{A}_Y, \mathcal{B}_Y|X_i) \otimes X_i - \frac{S^{(2)}_{\theta\xi}(V_i, u)}{S^{(0)}(V_i, u)} + \frac{S^{(1)}_\xi(V_i, u)}{S^{(0)}(V_i, u)} \left\{ \frac{S^{(1)}_\theta(V_i, u)}{S^{(0)}(V_i, u)} \right\}^\top \right],
\end{aligned}
$$

where $\mathcal{V}_{y_i^{o_r}}$ and $\mathcal{C}_{y_i^{o_r}}$ denote variances and covariances with respect to the conditional distribution $Q^{y^r_m}_{y^r_o}$. Thus, the components of the matrix $\mathcal{I} = \lim_{n\to\infty}(1/n)\partial T_n/\partial\eta$ can now be obtained by replacing S by s throughout the above equations and then taking the expectations of the corresponding i.i.d. sums.

Regularity Conditions

In the following, let $|\theta|_p$ denote the sum of the absolute values of the θ-parameters associated with the vector of covariates X_p.

(A1) $\pi^{(r)}(T, S) > 0$ almost surely in (T, S) for $r = (1, 1, \ldots, 1)$.

(A2) The elements of the second-stage design matrices $\mathcal{B}$ and $\mathcal{A}$ remain uniformly bounded in absolute value by constants, say c_B and c_A, respectively.

(A3) Assume the function $X^{\otimes 3} \exp\left(c_B \sum_{p=1}^P |\theta|_p X_p\right)$ can be bounded by an integrable function of X uniformly in a open neighborhood of θ_0.

(A4) The functions $E_{V,X} I(V \geq v) \exp\left(-c_A - c_B \sum_{p=1}^P |\theta|_p X_p\right)$ are bounded away from zero uniformly in v and $\eta = (\theta^T, \xi^T)^T$ in a open neighborhood of η_0.

(A5) The matrix $\mathcal{I}$ is positive definite.