Technical Report on

"Identifiability Conditions For The Linear Regression Model Under Right Censoring".

Dong, J.Y. and Yu, Q.Q.

We give an example of obtaining SMLE when $p = 2$.

Example 5. Let $X, Z \sim$ bin(1,0.5), $W \sim$ U(0,1), $\beta_1 = \beta_2 = 1$, $Y = X + Z + W$, and $C \equiv 1$. Here $\tau_C = \tau_o$=1, $\mathbf{x}_0 = (0,0)$, $\mathbf{x}_1 = (0,1)$ with $w_1 = 0$ and $\mathbf{x}_2 = (1,0)$ with $w_2 = 0$. By Theorem 1, since $\mathscr{B}_{x_0} \neq \emptyset$ and $\mu(\mathscr{B}_{x_0}) = 0$, the SMLE of $\boldsymbol{\beta} = (\beta_1, \beta_2)$ is consistent.

Proof of Example 5. Assume there are N_0 exact observations $(X, Z, M, \delta) = (0, 0, M, 1)$, N_1 RC observations $(X, Z, M, \delta) = (1, 0, 1, 0)$, N_2 RC observations $(X, Z, M, \delta) = (0, 1, 1, 0)$, N_3 = RC observations $(X, Z, M, \delta) = (1, 1, 1, 0)$ and $N_0 + N_1 + N_2 + N_3 = n$. WLOG, assume $M_1 < M_2 < \cdots < M_{N_1} < 1 = M_{N_1+1} = \cdots = M_n$. The exact observations divide the real line into $N_0 + 1$ intervals, $I_i = (M_i, M_{i+1}]$, $i = 0, 1, ..., N_0$, $M_0 = -\infty$ and $M_{N_0+1} = \infty$. Let $T_i = T_i(b_1, b_2) = M_i - b_1 X_i - b_2 Z_i$, $i = 1, 2, ..., n$, then $T_1 = M_1$, ..., $T_{N_1} = M_{N_1}$, $T_{N_1+1} = 1 - b_1 = \cdots = T_{N_1+N_2}$, $T_{N_1+N_2+1} = 1 - b_2 = \cdots = T_{n-N_3}$, $T_{n-N_3+1} = 1 - b_1 - b_2 = \cdots = T_n$. Assume $M_i \le 1 - b_1 < M_{i+1}$, $M_j \le 1 - b_2 < M_{j+1}$, and $M_k \le 1 - b_1 - b_2 < M_{k+1}$.

Case 1. If $0 \le i \le j \le k \le N_0$, then we can order all $T_i's$ as $M_1 < M_2 < ... < M_i \le 1 - b_1 < M_{i+1} < ... < M_j \le 1 - b_2 < M_{j+1} < ... < M_k \le (1 - b_1 - b_2) < M_{k+1} < ... < M_{N_0}$. Then

$$
\begin{aligned}
&\mathscr{L}(\mathbf{b}, \hat{S}_{\mathbf{b}}) \\
&= [\prod_{r=1}^{i} \hat{f}(M_r)]\ [(\hat{S}(1-b_1))^{N_1}]\ [\prod_{r=i+1}^{j} \hat{f}(M_r)]\ [(\hat{S}(1-b_2))^{N_2}]\ [\prod_{r=j+1}^{k} \hat{f}(M_r)]\ [(\hat{S}(1-b_1-b_2))^{N_3}]\ [\prod_{r=k+1}^{N_0} \hat{f}(M_r)] \\
&= (\frac{1}{n})^i\ (\frac{n-i}{n})^{N_1}\ (\frac{n-i}{n}\frac{1}{N_2+N_3+N_0-i})^{j-i}\ (\frac{n-i}{n}\frac{N_2+N_3+N_0-j}{N_2+N_3+N_0-i})^{N_2} \\
&\quad (\frac{n-i}{n}\frac{N_2+N_3+N_0-j}{N_2+N_3+N_0-i}\frac{1}{N_2+N_0-j})^{k-j}\ (\frac{n-i}{n}\frac{N_2+N_2+N_0-j}{N_2+N_3+N_0-i}\frac{N_3+N_0-k}{N_3+N_0-j})^{N_3} \\
&\quad (\frac{n-i}{n}\frac{N_2+N_3+N_0-j}{N_2+N_3+N_0-i}\frac{N_3+N_0-k}{N_3+N_0-j}\frac{1}{N_0-k})^{N_0-k} \\
&= \frac{1}{n^n}\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}\frac{(N_0+N_2+N_3-j)^{N_0+N_2+N_3-j}}{(N_0+N_3-j)^{N_0+N_3-j}}\frac{(N_0+N_3-k)^{N_0+N_3-k}}{(N_0-k)^{N_0-k}}
\end{aligned}
$$

By Lemma 3, since $n - i = N_1 + N_2 + N_3 + N_0 - i > N_0 + N_2 + N_3 - i$, $N_0 + N_2 + N_3 - j > N_0 + N_3 - j$, $N_0 + N_3 - k > N_0 - k$, $\frac{(N_0+N_3-k)^{N_0+N_3-k}}{(N_0-k)^{N_0-k}}$ is a decreasing function and is maximized when $k = j$, $\frac{(N_0+N_2+N_3-j)^{N_0+N_2+N_3-j}}{(N_0+N_3-j)^{N_0+N_3-j}}$ is a decreasing function and is maximized when $j = i$, and $\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}$ is a decreasing function and is maximized when $i = 0$. So $\mathscr{L}(\mathbf{b}, \hat{S}_{\mathbf{b}})$ is maximized when $i = j = k = 0$.

Case 2. If $0 \le i \le k \le j \le N_0$, then we can order all $T_i's$ as $M_1 < M_2 < ... < M_i \le 1 - b_1 < M_{i+1} <$

$... < M_k \leq 1-b_1-b_2 < M_{k+1} < ... < M_j \leq 1-b_2 < M_{j+1} < ... < M_{N_0}$. Then

$$\begin{aligned}
&\mathscr{L}(\mathbf{b},\hat{S}_{\mathbf{b}})\\
&= [\prod_{r=1}^{i}\hat{f}(M_r)]\ [(\hat{S}(1-b_1))^{N_1}]\ [\prod_{r=i+1}^{k}\hat{f}(M_r)]\ [(\hat{S}(1-b_1-b_2))^{N_3}]\ [\prod_{r=k+1}^{j}\hat{f}(M_r)]\ [(\hat{S}(1-b_2))^{N_2}]\ [\prod_{r=k+1}^{N_0}\hat{f}(M_r)]\\
&= (\frac{1}{n})^i\ (\frac{n-i}{n})^{N_1}\ (\frac{n-i}{n}\frac{1}{N_2+N_3+N_0-i})^{k-i}\ (\frac{n-i}{n}\frac{N_2+N_3+N_0-k}{N_2+N_3+N_0-i})^{N_2}\\
&\quad (\frac{n-i}{n}\frac{N_2+N_3+N_0-k}{N_2+N_3+N_0-i}\frac{1}{N_2+N_0-k})^{j-k}\ (\frac{n-i}{n}\frac{N_2+N_3+N_0-k}{N_2+N_3+N_0-i}\frac{N_2+N_0-j}{N_2+N_0-k})^{N_2}\\
&\quad (\frac{n-i}{n}\frac{N_2+N_3+N_0-k}{N_2+N_3+N_0-i}\frac{N_2+N_0-j}{N_2+N_0-k}\frac{1}{N_0-j})^{N_0-j}\\
&= \frac{1}{n^n}\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}\frac{(N_0+N_2+N_3-k)^{N_1+N_2+N_3-k}}{(N_0+N_2-k)^{N_0+N_2-k}}\frac{(N_0+N_2-j)^{N_0+N_2-j}}{(N_0-j)^{N_0-j}}
\end{aligned}$$

By Lemma 3, since $n-i = N_0+N_1+N_2+N_3-i > N_0+N_2+N_3-i$, $N_0+N_2+N_3-j > N_0+N_2-j$, $N_0+N_2-k = N_0-k$, $\frac{(N_0+N_2-j)^{N_0+N_2-j}}{(N_0-j)^{N_0-j}}$ is a decreasing function and is maximized when $j=i$, $\frac{(N_0+N_2+N_3-k)^{N_1+N_2+N_3-k}}{(N_0+N_2-k)^{N_0+N_2-k}}$ is a decreasing function and is maximized when $k=j$, and $\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}$ is a decreasing function and is maximized when $i=0$. So $\mathscr{L}(\mathbf{b},\hat{S}_{\mathbf{b}})$ is maximized when $i=j=k=0$.

Other cases. By the similar argument as in Case 1 and 2, one can obtain

$$\mathscr{L}(\mathbf{b},\hat{S}_{\mathbf{b}}) =$$
$$\begin{cases}
\frac{1}{n^n}\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}\frac{(N_0+N_2+N_3-j)^{N_0+N_2+N_3-j}}{(N_0+N_3-j)^{N_0+N_3-j}}\frac{(N_0+N_3-k)^{N_0+N_3-k}}{(N_0-k)^{N_0-k}} & \text{if } 0\leq i\leq j\leq k\leq N_0\\
\frac{1}{n^n}\frac{(n-i)^{n-i}}{(N_0+N_2+N_3-i)^{N_0+N_2+N_3-i}}\frac{(N_0+N_2+N_3-k)^{N_0+N_2+N_3-k}}{(N_0+N_2-k)^{N_0+N_2-k}}\frac{(N_0+N_2-j)^{N_0+N_2-j}}{(N_0-j)^{N_0-j}}, & \text{if } 0\leq i\leq k\leq j\leq N_0\\
\frac{1}{n^n}\frac{(n-j)^{n-j}}{(N_0+N_1+N_3-j)^{N_0+N_1+N_3-j}}\frac{(N_0+N_1+N_3-i)^{N_0+N_1+N_3-i}}{(N_0+N_3-i)^{N_0+N_1-i}}\frac{(N_0+N_3-k)^{N_0+N_3-k}}{(N_0-k)^{N_0-k}}, & \text{if } 0\leq j\leq i\leq k\leq N_0\\
\frac{1}{n^n}\frac{(n-j)^{n-j}}{(N_0+N_3+N_1-j)^{N_0+N_3+N_1-j}}\frac{(N_0+N_3+N_1-k)^{N_0+N_3+N_1-k}}{(N_0+N_1-k)^{N_0+N_1-k}}\frac{(N_0+N_1-i)^{N_0+N_1-i}}{(N_0-i)^{N_0-i}}, & \text{if } 0\leq j\leq k\leq i\leq N_0\\
\frac{1}{n^n}\frac{(n-k)^{n-k}}{(N_0+N_1+N_2-k)^{N_0+N_1+N_2-k}}\frac{(N_0+N_1+N_2-i)^{N_0+N_1+N_2-i}}{(N_0+N_2-i)^{N_0+N_2-i}}\frac{(N_0+N_2-j)^{N_0+N_2-j}}{(N_0-j)^{N_0-j}}, & \text{if } 0\leq k\leq i\leq j\leq N_0\\
\frac{1}{n^n}\frac{(n-k)^{n-k}}{(N_0+N_2+N_1-k)^{N_0+N_2+N_1-k}}\frac{(N_0+N_2+N_1-j)^{N_0+N_2+N_1-j}}{(N_0+N_1-j)^{N_0+N_1-j}}\frac{(N_0+N_1-i)^{N_0+N_1-i}}{(N_0-i)^{N_0-i}}, & \text{if } 0\leq k\leq j\leq i\leq N_0
\end{cases}$$

For each case, by Lemma 3, since the last three terms are decreasing functions of i, j and k respectively, each case is always maximized when $i=j=k=0$. The order of i, j, or k doesn't matter while maximizing $\mathscr{L}(\mathbf{b},\hat{S}_{\mathbf{b}})$ and the maximum is always obtained when $i=j=k=0$.

In other words, $\mathscr{L}$ is maximized when $1-b_1, 1-b_2, 1-b_1-b_2 < M_1$, that is, when $b_1, b_2, b_1+b_2 > 1-M_1$. To obtain three inequalities simultaneously, we need to make $b_1 > 1-M_1$ and $b_2 > 1-M_1$. Similar to $p=1$ case, we can let $b_1 = 1-M_1+\eta$, $b_2 = 1-M_1+\eta$, e.g., $\eta = \frac{1}{n}$. When sample size is large enough, $M_1 = \min(M_i) \to 0$ and thus $b_1 \to 0$ and $b_2 \to 0$. They are both consistent.

Note. If we let $b_1 = 1 - M_1 + \eta$ and $b_1 + b_2 = 1 - M_1 + \eta$, then we cannot guarantee $b_2 > 1 - M_1$. So the only way to make all inequalities hold is to set $b_1 = 1 - M_1 + \eta$, $b_2 = 1 - M_1 + \eta$.

Remark 4. Example 5 shows that the order of i, j, or k doesn't matter while maximizing $\mathscr{L}(\mathbf{b}, \hat{S}_{\mathbf{b}})$ and the maximum is always obtained when $i = j = k = 0$. So, WLOG, we can assume $i \le j \le k$.