ON PROFILE MM ALGORITHMS FOR GAMMA FRAILTY SURVIVAL MODELS

Xifen Huang[1], Jinfeng Xu[1,*] and Guoliang Tian[2]

[1] *The University of Hong Kong and* [2] *Southern University of Science and Technology*

Supplementary Material

S1 Proof of Theorem 2

First, it is easy to see that the MM2 algorithm is an MM algorithm. By its construction, the minorizing function $Q_1(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ for $\ell_1(\theta, \boldsymbol{\beta}, \Lambda_0|Y_{\rm obs})$ satisfies that

$$\begin{aligned}
\ell_1(\theta, \boldsymbol{\beta}, \Lambda_0|Y_{\rm obs}) &\geqslant Q_1(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}), \quad \forall\, \theta, \boldsymbol{\beta}, \Lambda_0 \quad \text{and} \\
\ell_1(\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}|Y_{\rm obs}) &= Q_1(\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}).
\end{aligned}$$

Recall that $Q_1(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) = Q_{11}(\theta|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) + Q_{12}(\boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ and $\max_{\Lambda_0} Q_{12}(\boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) = Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. To maximize $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, the MM method is used again. The minorizing function for $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ is $Q_{15}(\beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, satisfying

$$\begin{aligned}
Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) &\geqslant Q_{15}(\beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \quad \forall\, \boldsymbol{\beta} \quad \text{and} \\
Q_{13}(\boldsymbol{\beta}^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) &= Q_{15}(\beta_1^{(k)}, \ldots, \beta_q^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}).
\end{aligned}$$

This follows from the fact that

$$
\begin{aligned}
& Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \\
= & \sum_{i=1}^{B}\sum_{j=1}^{M_i}\left\{I_{ij}\mathbf{X}_{ij}^{\top}\boldsymbol{\beta} - I_{ij}\log\left[\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta})\right]\right\}, \\
\geqslant & \sum_{i=1}^{B}\sum_{j=1}^{M_i}\left\{I_{ij}\mathbf{X}_{ij}^{\top}\boldsymbol{\beta} - I_{ij}\log\left[\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})\right]\right. \\
- & \left.\frac{I_{ij}\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta})}{\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})} + I_{ij}\right\}, \\
= & \sum_{i=1}^{B}\sum_{j=1}^{M_i}\left\{I_{ij}\mathbf{X}_{ij}^{\top}\boldsymbol{\beta} - I_{ij}\log\left[\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})\right] + I_{ij}\right. \\
- & \left.\frac{I_{ij}\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp\left(\sum_{p=1}^{q}\delta_{prs}[\delta_{prs}^{-1}X_{prs}(\beta_p-\beta_p^{(k)}) + \mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)}]\right)}{\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})}\right\}, \\
\geqslant & \sum_{i=1}^{B}\sum_{j=1}^{M_i}\left\{I_{ij}\mathbf{X}_{ij}^{\top}\boldsymbol{\beta} - I_{ij}\log\left[\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})\right] + I_{ij}\right. \\
- & \left.\frac{I_{ij}\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\sum_{p=1}^{q}\delta_{prs}\exp[\delta_{prs}^{-1}X_{prs}(\beta_p-\beta_p^{(k)}) + \mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)}]}{\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})}\right\}, \\
= & \sum_{p=1}^{q}\sum_{i=1}^{B}\sum_{j=1}^{M_i}\left[I_{ij}\beta_p X_{pij} - \frac{I_{ij}\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\delta_{prs}\exp[\delta_{prs}^{-1}(\beta_p-\beta_p^{(k)})X_{prs} + \mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)}]}{\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs}\geqslant t_{ij})\exp(\mathbf{X}_{rs}^{\top}\boldsymbol{\beta}^{(k)})}\right], \\
\hat{=} & \, Q_{15}(\beta_1,\ldots,\beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}).
\end{aligned}
$$

To prove the convergence of the MM2 algorithm, we first need to verify the convergence conditions for the inner loop MM algorithm constructed for maximizing $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. It is easy to check that conditions **C1, C2, C4** hold. The concavity of $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ as a function of $\boldsymbol{\beta}$ shows that condition **C5** holds. By (3.8) and (3.9), we can see that condition **C6** is satisfied. It remains to verify condition **C3**. It is to prove that the set $\Omega_c = \left\{\boldsymbol{\beta} \in \mathbb{R}^q : Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \geqslant c\right\}$ is compact. By the continuity of $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, the set Ω_c is closed. We now use proof by contradiction to show its boundedness. Assume that Ω_c is unbounded and there exist $\boldsymbol{\beta}_{0m} \in \Omega_c, m = 1, 2, ...$ s.t. $||\boldsymbol{\beta}_{0m}|| \to \infty$ as $m \to \infty$. Without loss of generality, let $O = \{r : \lim_{m\to\infty} \boldsymbol{\beta}_{0mr} \to \infty\}$ and $O^c = \{s : \lim_{m\to\infty} \boldsymbol{\beta}_{0ms} \to -\infty\}$. Note that

$$
\begin{aligned}
&\exp[Q_{13}(\boldsymbol{\beta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})] \\
=&\prod_{i=1}^{B}\prod_{j=1}^{M_i}\left\{\sum_{r=1}^{B}\frac{A_r^{(k)}}{\Pi_r^{(k)}}\sum_{s=1}^{M_r} I(t_{rs} \geqslant t_{ij}) \exp\left[(\mathbf{X}_{rs}^{\top} - \mathbf{X}_{ij}^{\top})\boldsymbol{\beta}_0\right]\right\}^{-I_{ij}}.
\end{aligned}
$$

By Condition A (ii), there exist the pairs (i, j), (i_1, j_1), and (i_2, j_2) such that $I_{ij} = 1$, $t_{i_1j_1} \geqslant t_{ij}$, $t_{i_2j_2} \geqslant t_{ij}$ and for any $r \in O$ and $s \in O^c = O_0 - O$, $X_{i_1j_1r} - X_{ijr} > 0$ and $X_{i_2j_2s} - X_{ijs} < 0$. It follows that as $m \to \infty$, $Q_{13}(\boldsymbol{\beta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \to -\infty$. Since $\boldsymbol{\beta}_{0m} \in \Omega_c$, we have $Q_{13}(\boldsymbol{\beta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)},$

$\Lambda_0^{(k)}) \geqslant c$. This yields contradiction and hence Ω_c is bounded. It follows that condition **C3** holds. By Lemma 1 and the unimodality of $Q_{13}(\boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, the limiting point of the inner-loop sequence $\{\boldsymbol{\beta}^{(k)}\}_k$ is its unique maximizer. Consequently, the limiting point, denoted by $\boldsymbol{\beta}^*$ together with $\hat{\Lambda}_0(t_{ij})$ calculated by (3.9) is the unique maximizer of $Q_{12}(\boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. Similarly as in Theorem 1, under Condition A, we can further show that the overall MM2 algorithm is convergent and the details are omitted here.

S2 Proof of Theorem 3.

In the MM3 algorithm, the minorizing function $Q^*(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ for $\ell_2(\theta, \boldsymbol{\beta}, \Lambda_0|Y_{\text{obs}})$ satisfies the two conditions

$$\ell_2(\theta, \boldsymbol{\beta}, \Lambda_0|Y_{\text{obs}}) \geqslant Q^*(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}), \quad \forall\, \theta, \boldsymbol{\beta}, \Lambda_0,$$

$$\ell_2(\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}|Y_{\text{obs}}) = Q^*(\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}).$$

Note that $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) = \max_{\Lambda_0} Q^*(\theta, \boldsymbol{\beta}, \Lambda_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ and an inner-loop MM algorithm to maximize $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ with respect to $\boldsymbol{\beta}$ and θ. The minorizing function for $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ is $Q(\theta, \beta_1, \ldots,$

$\beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. It satisfies the two conditions

$$\begin{aligned} Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) &\geqslant Q_2^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \\ &\geqslant Q_3^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \\ &\geqslant Q(\theta, \beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}), \end{aligned}$$

and

$$Q_1^*(\theta^{(k)}, \boldsymbol{\beta}^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) = Q(\theta^{(k)}, \beta_1^{(k)}, \ldots, \beta_q^{(k)}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}).$$

To prove the convergence of MM3 algorithm, we first show the convergence of the inner-loop MM sequence for maximizing $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. From the expressions of $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ and $Q(\theta, \beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, it is easy to verify that conditions **C1, C2, C4** hold. In addition, the parameters are separated in $Q(\theta, \beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$. It is easy to see that there exists a unique global maximum of $Q(\theta, \beta_1, \ldots, \beta_q|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, ensuring that **C6** holds. Next we verify condition **C3**. By the continuity of $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$, the set $\Omega_c = \big\{\boldsymbol{\eta} = (\theta, \boldsymbol{\beta}) : Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \geqslant c\big\}$ is closed. Similarly as in the proofs of Theorem 1 and Theorem 2, we use a proof by contradiction to prove its boundedness. If the set Ω_c is unbounded, then there exists a sequence $\boldsymbol{\eta}_{0m} = (\theta_{0m}, \boldsymbol{\beta}_{0m})$ s.t. $||\boldsymbol{\eta}_{0m}|| \to \infty$ as $m \to \infty$. This indicates that $\theta_{0m} \to \infty$ or $||\boldsymbol{\beta}_{0m}|| \to \infty$ as $m \to \infty$. When $||\boldsymbol{\beta}_{0m}|| \to \infty$, without loss of generality, let us assume that $O = \{r :$

$\lim_{m\to\infty} \boldsymbol{\beta}_{0mr} \to \infty\}$ and $O^c = \{s : \lim_{m\to\infty} \boldsymbol{\beta}_{0ms} \to -\infty\}$. Notice that

$$\begin{aligned}
&\exp[Q_1^*(\boldsymbol{\eta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})] \\
&= \exp\Bigg\{\sum_{i=1}^{B}\Bigg[\log\Gamma(D_i+\frac{1}{\theta}) - \log\Gamma(\frac{1}{\theta}) - \frac{\log(\theta)}{\theta} + \frac{1}{\theta}\Bigg(1 - \log(\Pi_i^{(k)}) - \frac{D_i}{\Pi_i^{(k)}}\Bigg) \\
&- \frac{1}{\Pi_i^{(k)}\theta^2}\Bigg]\Bigg\} \times \prod_{i=1}^{B}\prod_{j=1}^{M_i}\Bigg\{\sum_{r=1}^{B}\sum_{s=1}^{M_r} I(t_{rs} \geqslant t_{ij})\Bigg(\frac{D_r}{\Pi_r^{(k)}} + \frac{1}{\Pi_r^{(k)}\theta}\Bigg)\exp\big[(\mathbf{X}_{rs}^\top - \mathbf{X}_{ij}^\top)\boldsymbol{\beta}\big]\Bigg\}^{-I_{ij}}.
\end{aligned}$$

When $\theta_{0m} \to \infty$, we have

$$\begin{aligned}
&\exp\Bigg\{\sum_{i=1}^{B}\Bigg[\log\Gamma(D_i+\frac{1}{\theta}) - \log\Gamma(\frac{1}{\theta}) - \frac{\log(\theta)}{\theta} \\
&+ \frac{1}{\theta}\Bigg(1 - \log(\Pi_i^{(k)}) - \frac{D_i}{\Pi_i^{(k)}}\Bigg) - \frac{1}{\Pi_i^{(k)}\theta^2}\Bigg]\Bigg\} \to 0,
\end{aligned}$$

by Condition A (i). When $||\boldsymbol{\beta}_{0m}|| \to \infty$, by Condition A (ii), and there exist the pairs (i,j), (i_1,j_1), and (i_2,j_2) such that $I_{ij} = 1$, $t_{i_1j_1} \geqslant t_{ij}$, $t_{i_2j_2} \geqslant t_{ij}$ and for any $r \in O$ and $s \in O^c = O_0 - O$, $X_{i_1j_1r} - X_{ijr} > 0$ and $X_{i_2j_2s} - X_{ijs} < 0$. Consequently, we have

$$\sum_{r=1}^{B}\sum_{s=1}^{M_r} I(t_{rs} \geqslant t_{ij})\Bigg(\frac{D_r}{\Pi_r^{(k)}} + \frac{1}{\Pi_r^{(k)}\theta}\Bigg)\exp\big[(\mathbf{X}_{rs}^\top - \mathbf{X}_{ij}^\top)\boldsymbol{\beta}\big] \to \infty.$$

In either case, $\exp[Q_1^*(\boldsymbol{\eta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})] \to 0$ and $Q_1^*(\boldsymbol{\eta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \to -\infty$, yielding contradiction with the fact that $Q_1^*(\boldsymbol{\eta}_0|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)}) \geqslant c$. Hence the set Ω_c is bounded. By Lemma 1, we conclude that the inner loop sequence of the MM3 algorithm which maximizes $Q_1^*(\theta, \boldsymbol{\beta}|\theta^{(k)}, \boldsymbol{\beta}^{(k)}, \Lambda_0^{(k)})$ is convergent. Finally, the convergence of the overall MM3 algorithm can be proved similarly as in the MM1 algorithm. We omit the details here.