A Spline functions

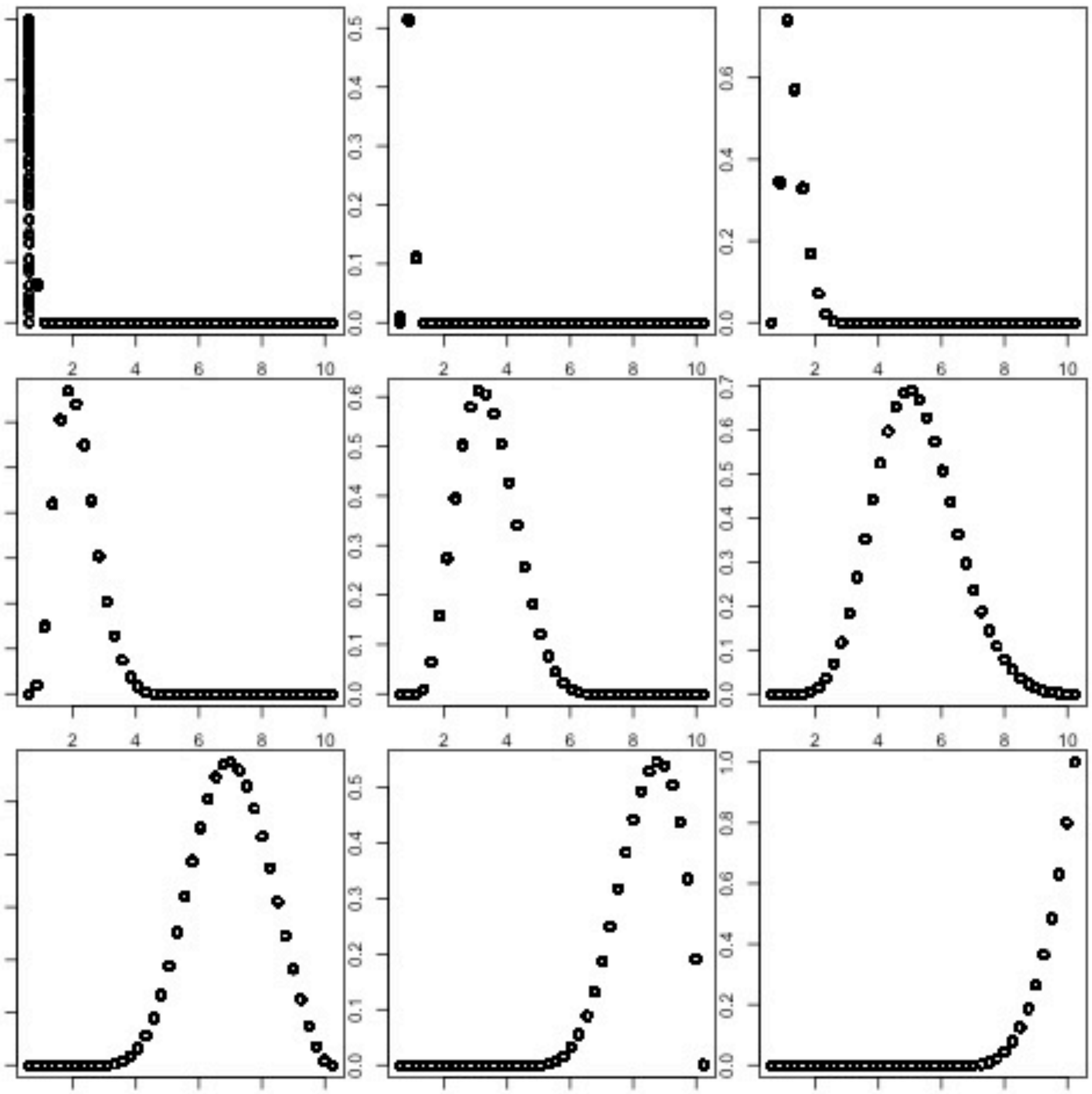

Figure A.1: Spline functions of time used as covariates in $\boldsymbol{D}^*_{ij}$ with profile-specific coefficients. The x-axis represents time in years.

B Posterior computation

Supplemental materials for the posterior computation under marginal and joint profiling.

B.1 Marginal profiling

The prior distributions are specified as: $\beta \sim N(0, \mathbf{\Sigma}_\beta)$, $\alpha_l \sim N(0, \mathbf{\Sigma}_\alpha)$, $\sigma_l^2 \sim$ Inverse-Gamma: IG(a, b), $\mathbf{\Sigma}_r \sim$ Inverse-Wishart: IW$(\nu_b, \mathbf{\Sigma}_b)$ and $\eta_l \sim N(b_l, B_l)$, where we set the hyper-parameters equal to 1 and assume independence in the covariance matrix specification. The posterior updating follows the steps:

1. Update C_i: for $i = 1, \ldots, n$, draw C_i from a multinomial distribution with probability

$$\Pr(C_i = l) = \frac{\pi_l(\boldsymbol{X}_{i0}, \eta_l) \prod_j f(Y_{ij}|\beta, \alpha_l, b_i, \sigma_l^2, D_i, D_i^*, D_i^{**})}{\sum_l \pi_l(\boldsymbol{X}_{i0}, \eta_l) \prod_j f(Y_{ij}|\beta, \alpha_l, b_i, \sigma_l^2, D_i, D_i^*, D_i^{**})}$$

2. Update α_l: for $l = 2, \ldots, L$, where $\alpha_1 = 0$ for identification purpose, $\pi(\alpha_l|-) = N(\mu_\alpha, V_\alpha)$, where $V_{\alpha_l} = (\Sigma_\alpha^{-1} + \sigma_l^{-2} D_l^{*\prime} D_l^*)^{-1}$, $\mu_\alpha = V_{\alpha_l} \sigma_l^{-2} D_l^{*\prime} (X_l - D_l\beta - D_l^{**} b_l)$. Here $D_l, D_l^*, D_l^{**}, X_l, b_l$ are subsets of the design matrix and random effects for all i, j such that $C_i = l$

3. Update η_l: for $l = 2, \ldots, L$, let $c_{il} = I(C_i = l)$, $\pi(\eta_l) \sim N(m_l^*, S_l^*)$, where $S_l^* = (V^T \Omega_l V + B_l^{-1})^{-1}$, V is the design matrix with each row $\boldsymbol{X}_{i0}^T$, $\Omega_l = diag(w_{1l}, \ldots, w_{nl})$, and $m_l^* = S_l^*(B_l^{-1} b_l + V^T m_l)$, where m_l is a vector in R^n, and the ith component is $m_l^{(i)} = c_{il} - 1/2 + w_{il}\{\log[\sum_{k \neq l} \exp(\boldsymbol{X}_{i0}^T \eta_k)]\}$.

4. Update $\beta \sim N(\mu_\beta, V_\beta)$,where $\boldsymbol{X}_{i0}^* = D_i^{**} \mathbf{\Sigma}_r D_i^{**\prime} + \sigma_{C_i}^2 I_{T_i}$, $V_\beta = (\mathbf{\Sigma}_\beta^{-1} + \sum_i D_i' \boldsymbol{X}_{i0}^{*\,-1} D_i)^{-1}$, and $\mu_\beta = V_\beta \sum_i D_i' \boldsymbol{X}_{i0}^{*\,-1}(Y_i - D_i^* \alpha_{C_i})$.

5. Update σ_l^2, for $l = 1, \ldots, L$ from $IG(\sum_{i:C_i=l} T_i/2 + a, (Y_l - D_l\beta - D_l^*\alpha - D_l^{**} b_l)'(Y_l - D_l\beta - D_l^*\alpha - D_l^{**} b_l)/2 + b)$.

6. Update $\mathbf{\Sigma}_r$ from $IW(n + \nu_b, b'b + \mathbf{\Sigma}_b)$.

7. Update b_i from $N(\mu_b, V_b)$, here $V_b = (\mathbf{\Sigma}_r^{-1} + \sigma_{C_i}^{-2} D_i^{**\prime} D_i^{**})^{-1}$, and $\mu_b = V_b \sigma_{C_i}^{-2} D_i^{**\prime}(Y_i - D_i\beta - D_i^* \alpha_{C_i})$

After convergence, we impute missing data from the assumed model for Y since this step does not need to be included into the iterations.

B.2 Joint profiling

The prior distributions are specified as $\beta \sim N(0, \mathbf{\Sigma}_\beta)$, $\alpha_l \sim N(0, \mathbf{\Sigma}_\alpha)$, $\sigma_l^2 \sim$ IG(a, b), $\mathbf{\Sigma}_r \sim$ IW$(\nu_b, \mathbf{\Sigma}_b)$, $\eta_l \sim N(b_l, B_l)$, $\nu \sim N(0, \mathbf{\Sigma}_\nu)$, $\gamma_l \sim N(0, \mathbf{\Sigma}_\gamma)$ and $E \sim IW(\nu_e, \mathbf{\Sigma}_e)$, where we set the hyper-parameters equal to 1 and assume independence in the covariance matrix specification. The posterior computation is the following.

1. Update C_i: for $i = 1, \ldots, n$, draw C_i from a multinomial distribution with probability

$$\Pr(C_i = l) = \frac{\pi_l(\boldsymbol{X}_{i0}, \eta_l) \prod_j p_{ij|C_i=l}^{R_{ij}} (1 - p_{ij|C_i=l})^{1-R_{ij}} f(Y_{ij}|\beta, \alpha_l, b_i, \sigma_l^2, D_i, D_i^*, D_i^{**})}{\sum_l \pi_l(\boldsymbol{X}_{i0}, \eta_l) \prod_j p_{ij|C_i=l}^{R_{ij}} (1 - p_{ij|C_i=l})^{1-R_{ij}} f(Y_{ij}|\beta, \alpha_l, b_i, \sigma_l^2, D_i, D_i^*, D_i^{**})}.$$

2. Update α_l: for $l = 2, \ldots, L$, where $\alpha_1 = 0$, $\pi(\alpha_l|-) = N(\mu_\alpha, V_\alpha)$, where $V_{\alpha_l} = (\Sigma_\alpha^{-1} + \sigma_l^{-2} D_l^{*'} D_l^*)^{-1}$, $\mu_\alpha = V_{\alpha_l} \sigma_l^{-2} D_l^{*'} (Y_l - D_l\beta - D_l^{**} b_l)$.

3. Update η_l: for $l = 2, \ldots, L$, $\pi(\eta_l) \sim N(m_l^*, S_l^*)$, where $S_l^* = (V^T \Omega_l V + B_l^{-1})^{-1}$, V is the design matrix with each row $\boldsymbol{X}_{i0}^T$, $\Omega_l = diag(w_{1l}, \ldots, w_{nl})$, and $m_l^* = S_l^*(B_l^{-1} b_l + V^T m_l)$, where m_l is a vector in R^n, and the ith component is

$$m_l^{(i)} = c_{il} - 1/2 + w_{il}\{\log[\sum_{k \neq l} \exp(\boldsymbol{X}_{i0}^T \eta_k)]\}$$

4. Update β: $\pi(\beta|-) = N(\mu_\beta, V_\beta)$, $\boldsymbol{X}_{i0}^* = D_i^{**} \boldsymbol{\Sigma}_r D_i^{**'} + \sigma_{C_i}^2 I_{T_i}$, $V_\beta = (\boldsymbol{\Sigma}_\beta^{-1} + \sum_i D_i' \boldsymbol{X}_{i0}^{*\,-1} D_i)^{-1}$, and $\mu_\beta = V_\beta \sum_i D_i' \boldsymbol{X}_{i0}^{*\,-1} (Y_i - D_i^* \alpha_{C_i})$.

5. Update σ_l^2, for $l = 1, \ldots, L$, from $IG(\sum_{i:C_i=l} T_i/2 + 1, (Y_l - D_l\beta - D_l^*\alpha - D_l^{**}b)'(Y_l - D_l\beta - D_l^*\alpha - D_l^{**}b)/2 + 1)$.

6. Update $\boldsymbol{\Sigma}_r$ from $IW(n + \nu_b, b'b + \boldsymbol{\Sigma}_b)$.

7. Update b_i from $\pi(b_i|-) = N(\mu_b, V_b)$, here $V_b = (\boldsymbol{\Sigma}_r^{-1} + \sigma_{C_i}^{-2} D_i^{**'} D_i^{**})^{-1}$, and $\mu_b = V_b \sigma_{C_i}^{-2} D_i^{**'} (Y_i - D_i\beta - D_i^* \alpha_{C_i})$.

8. Update ν: $[v|-] \sim N(m_v^*, S_v^*)$, where $m_v^* = S_v^*(B^\top k_v)$, $B = \left[\begin{array}{c} B_{11}^\top, ..., B_{1T_1}^\top, B_{21}^\top, ..., B_{nT_n}^\top \end{array}\right]^\top$ and k_v is a vector with length $\sum_{i=1}^n T_i$:

$$k_v = \left[\begin{array}{c} k_{11}^* - w_{11}^*((B_{11}^*)^\top \gamma_{C_1} + (B_{11}^{**})^\top e_1)), ..., k_{nT_n}^* - w_{nT_n}^*((B_{nT_n}^*)^\top \gamma_{C_n} + (B_{nT_n}^{**})^\top e_n)) \end{array}\right]^\top$$

and $S_v^* = (\boldsymbol{\Sigma}_v^{-1} + B^\top \Omega_v B)^{-1}$, where $\Omega_v = diag(w_{11}^*, ..., w_{1T_1}^*, w_{21}^*, ..., w_{2T_2}^*, ..., w_{n1}^*, ..., w_{nT_n}^*)$.

9. Update γ_l, for $l = 2, \ldots, k$, from $[\gamma_l|-] \sim N(m_{\gamma_l}^*, S_{\gamma_l}^*)$, where $m_{\gamma_l}^* = S_{\gamma_l}^*((B^*)^\top k_{\gamma_l})$, $B_l^* = \left[\begin{array}{c} (B_{11}^*)^\top I(C_1 = l), ..., (B_{1T_1}^*)^\top I(C_1 = l), (B_{21}^*)^\top I(C_2 = l), ..., B_{nT_n}^*)^\top I(C_n = l) \end{array}\right]^\top$ and k_{γ_l} is a vector with length $\sum_{i:C_i=l} T_i$:

$$k_{\gamma_l} = \left[\begin{array}{c} \{k_{11}^* - w_{11}^*(B_{11}^\top \nu + (B_{11}^{**})^\top e_1)\} I(C_1 = l), ..., \{k_{nT_n}^* - w_{nT_n}^*(B_{nT_n}^\top \nu + (B_{nT_n}^{**})^\top e_n)\} I(C_n = l) \end{array}\right]$$

and $S_{\gamma_l}^* = (\boldsymbol{\Sigma}_{\gamma_l}^{-1} + (B^*)^\top \Omega_{\gamma_l} B^*)^{-1}$, where $\Omega_{\gamma_l} = diag(w_{11}^* I(C_1 = l), ..., w_{1T_1}^* I(C_1 = l), w_{21}^* I(C_2 = l), ..., w_{2T_2}^* I(C_2 = l), ..., w_{n1}^* I(C_n = l), ..., w_{nT_n}^* I(C_n = l))$.

10. Update e_i from $[e_i|-] \sim N(m_{e_i}^*, S_{e_i}^*)$, where $m_{e_i}^* = S_{e_i}^*((B_i^{**})^\top k_{e_i})$, $B_i^{**} = \left[\begin{array}{c} (B_{i1}^*)^\top, ..., (B_{iT_i}^*)^\top \end{array}\right]^\top$ and k_{e_i} is a vector of R^{T_i} as the following:

$$k_{e_i} = \left[\begin{array}{c} k_{i1}^* - w_{i1}^*(B_{i1}^\top \nu + (B_{i1}^*)^\top \gamma_{C_1}), ..., k_{iT_i}^* - w_{iT_i}^*(B_{iT_i}^\top \nu + (B_{iT_i}^*)^\top \gamma_{C_i}) \end{array}\right]^\top$$

and $S_{e_i}^* = (E^{-1} + ((B_i^{**})^\top \Omega_{e_i} B_i^{**})^{-1}$, where $\Omega_{e_i} = diag(w_{i1}^*, ..., w_{iT_i}^*)$.

11. Update E from $IW(n + \nu_e, \sum_{i=1}^n e_i e_i^\intercal + \boldsymbol{\Sigma}_e)$.

12. Impute missing data: draw Y_{ij}^* from $f(Y_{ij}|\beta, \alpha_{C_i}, b_i, \sigma_{C_i}^2, D_i, D_{ij}^*, D_i^{**})$.