Supplementary materials for accelerated fitting of joint models with cumulative variations

Yan Gao, Rodney A. Sparapani and Sergey Tarima

*Corresponding author(s). E-mail(s): yagao@mcw.edu;

1 Data notations and visualization

The mathematical symbols associated with data and statistical modeling are presented in Table 1. The illustration of the survival and longitudinal data are displayed in Tables 2 and 3, respectively.

Notation	Description
Data-associated symbols	
i	subject index
j	time point index
t_i	observed survival time
z_{ij}	observed longitudinal outcome for the ith subject at jth time point
n_i	number of time points for the ith subject
n	number of subjects
$N = \sum_{i=1}^{n} n_i$	number of observations
Model-associated symbols	
k	spline-function index
K	number of spline functions
d	GK node index
D	number of GK nodes
r	grid index for a partition
R	number of grids for a partition per subject
$\dot{R} = n \times R$	number of grids for all subjects
$\dot{K} = n \times R \times K$	number of all grids and spline functions for all subjects
$\ddot{R} = R - 1$	reduced number of grids
$\widetilde{R} = \dot{R} \times D$	number of grids and GK nodes per subject

Table 1 List of notations.

i	t_i	δ_i	x_{i1}	x_{i2}	$\ldots$	x_{iP}
1	t_1	δ_1	x_{11}	x_{12}	$\ldots$	x_{1P}
2	t_2	δ_2	x_{21}	x_{22}	$\ldots$	x_{2P}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	t_n	δ_n	x_{n1}	x_{n2}	$\ldots$	x_{nP}
n	t_n	δ_n	x_{n1}	x_{n2}	$\ldots$	x_{nP}

Table 2 Illustration of survival data.

i	z_{ij}	s_{ij}
1	z_{11}	s_{11}
1	z_{12}	s_{12}
1	z_{1n_1}	s_{1n_1}
2	z_{21}	s_{21}
2	z_{22}	s_{22}
2	z_{32}	s_{32}
2	z_{2n_2}	s_{2n_2}
$\vdots$	$\vdots$	$\vdots$
n	z_{n1}	s_{n1}
n	z_{n2}	s_{n2}
n	z_{n3}	s_{n3}
n	z_{n4}	s_{n4}
n	z_{nn_n}	s_{nn_n}

Table 3 Illustration of longitudinal data.

2 Matrix algebra for joint models

Appendix 2 elaborates matrix algebra for joint models. The detailed operations in the longitudinal submodel will be introduced in the subsections 2.1 and 2.2, followed by the details of the survival submodel in subsections **??** and **??**. Note that i' and j' in the matrix notation refer to the row and column indices of a matrix entry, respectively. And k' can index a vector or a matrix entry upon the context.

2.1 CSR structure in longitudinal submodel

In order to store the design matrix $\mathbf{B} = (B_{i'j'})$ and take advantages of the sparse matrices, B^s, we adopt the popular CSR (Compressed Sparse Row) data structure. Instead of the common term "array" under many contexts (Saad, 2003), we use the matrix term "vector" to align with our representations, because they are essentially equivalent in computation. Three vectors are constructed in the CSR structure $\mathbb{S}^B(\upsilon, \tau, \omega)$ to support efficient matrix-vector multiplication : 1) $\upsilon = \{\upsilon \in \mathbb{R} : \upsilon$ is an nonzero entry of $\mathbf{B}\}$ with $dim(\upsilon) = a$ (the number of nonzeros, NNZ) ; 2) $\tau = \{\tau \in \mathbb{N} : \tau$ is the column index of the nonzero entry of $\mathbf{B}\}$ with $dim(\tau) = a$; 3) $\omega = \{\omega \in \mathbb{N} : \omega$ is the row extent of nonzero entry of $\mathbf{B}\}$ with $dim(\omega) = N + 1$ and the last element, a. That is, if $\upsilon_{k'}$ stores $B_{i'j'}$, then $\tau_{k'} = j'$ and $\omega_{i'} \leq k' < \omega_{i'+1}$. The memory footprint and the complexity for the CSR format are of the order $\mathcal{O}(n)$ instead of $\mathcal{O}(n^2)$ as in a dense matrix (Hoffman, 2021).

2.2 Cholesky factorization in longitudinal submodel

To further accelerate the computation, the covariance matrix in (**??**) is decomposed into the product of the diagonal matrix diag(c) and the correlation matrix $\Lambda = (\Lambda_{i'j'})$: $\Sigma = (\Sigma_{i'j'}) = \mathrm{diag}(c) \times \Lambda \times \mathrm{diag}(c)$, where diag$(c)$ consists of the vector of the standard deviations $c = \{c_1, \ldots, c_K\}$. This mapping can be reversed by the following formula: $c_{k'} = \sqrt{\Sigma_{k'k'}}$; and $\Lambda_{i'j'} = \frac{\Sigma_{i'j'}}{c_{i'} c_{j'}}$. Given that the correlation matrix is a symmetric positive-definite matrix, it has a Cholesky factorization: $\Lambda = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L} = (l_{i'j'})$ denotes a real lower triangular matrix whose diagonal elements are positive. Each component can be derived based on the following formulas (Hoffman, 2021)

$$l_{i'i'} = \pm\sqrt{\Lambda_{i'j'} - \sum_{k'=1}^{i'-1} l_{i'k'}^2}$$

$$l_{i'j'} = \frac{1}{l_{i'j'}}\left(\Lambda_{i'j'} - \sum_{k'=1}^{j'-1} l_{i'k'} l_{j'k'}\right), \qquad i' > j'$$

which follows from

$$\begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \dots & \Lambda_{1K} \\ \Lambda_{21} & \Lambda_{22} & \dots & \Lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{K1} & \Lambda_{K2} & \dots & \Lambda_{KK} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{K1} & l_{K2} & \dots & l_{KK} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \dots & l_{K1} \\ 0 & l_{22} & \dots & l_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & l_{KK} \end{bmatrix}$$

$$= \begin{bmatrix} l_{11}^2 & l_{21}l_{11} & \dots & l_{K1}l_{11} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & \dots & l_{K1}l_{21} + l_{K2}l_{22} \\ \vdots & \vdots & \ddots & \vdots \\ l_{K1}l_{11} & l_{K1}l_{21} + l_{K2}l_{22} & \dots & l_{K1}^2 + \ldots + l_{KK}^2 \end{bmatrix}$$

References

Hoffman J (2021) Methods in Computational Science. Society for Industrial and Applied Mathematics, Philadelphia, PA

Saad Y (2003) Iterative Methods for Sparse Linear Systems, 2nd edn. Society for Industrial and Applied Mathematics