Web-based supporting materials for "Landmark Linear Transformation Model for Dynamic Prediction with Application to a Longitudinal Cohort Study of Chronic Disease" by

Yayuan Zhu, Liang Li, and Xuelin Huang

Appendix A. Algorithm for Solving Estimating Equations (8) and (9) in Section 2.2.

Given a landmark time s, for t_{ij} in the neighbourhood of s , we let $X_{ij} = X_i - t_{ij}$ be the observed residual failure times and $0 < u_1 < \cdots < u_{K_s}$ be the K_s uncensored residual failure times among X_{ij}'s. Then the resulting estimate of $H(u; s)$ is a step function with jumps only at $u_1, \dots, u_{K_s}$. The iterative algorithm is summarised as follows.

Step 1: Suppose we have an initial value of $\boldsymbol{\theta}(s)$ as $\boldsymbol{\theta}_0(s)$ and denote $\Lambda_\epsilon\{\boldsymbol{\theta}_0'(s)\widetilde{\boldsymbol{Z}}_i^p(t_{ij}-s) + H(u_k; s)\}$ by $\Lambda_\epsilon\{H(u_k; s)|\boldsymbol{\theta}_0(s)\}$ and let $w_{ij}(s) = K_h(t_{ij} - s)$. Note that $\Lambda_\epsilon(t) = \int_{-\infty}^t \lambda_\epsilon(u)du$ is the cumulative hazard function of ϵ_s in model (5), e.g., $\Lambda_\epsilon(t) = \exp(t)$ for the Cox models and $\Lambda_\epsilon(t) = \log\{1+\exp(t)\}$ for PO models, i.e., $\Lambda_\epsilon(t)$ is a known function given by the link function $g(\cdot)$ in (4). Then, $\widehat{H}(u_k; s)$, $k = 1, \dots, K_s$, can be obtained by numerically solving the following estimating equations one by one

$$\begin{aligned}
&\textstyle\sum_{i=1}^n \sum_{j=1}^{m_i} w_{ij}(s)[dN_i(u_1; t_{ij}) - Y_i(u_1; t_{ij})\Lambda_\epsilon\{H(u_1; s)|\boldsymbol{\theta}_0(s)\}] = 0\\
&\textstyle\sum_{i=1}^n \sum_{j=1}^{m_i} w_{ij}(s)[dN_i(u_2; t_{ij}) - Y_i(u_2; t_{ij})(\Lambda_\epsilon\{H(u_2; s)|\boldsymbol{\theta}_0(s)\} - \Lambda_\epsilon\{H(u_1; s)|\boldsymbol{\theta}_0(s)\})] = 0\\
&\vdots\\
&\textstyle\sum_{i=1}^n \sum_{j=1}^{m_i} w_{ij}(s)[dN_i(u_{K_s}; t_{ij}) - Y_i(u_{K_s}; t_{ij})(\Lambda_\epsilon\{H(u_{K_s}; s)|\boldsymbol{\theta}_0(s)\} - \Lambda_\epsilon\{H(u_{K_s-1}; s)|\boldsymbol{\theta}_0(s)\})] = 0\,.
\end{aligned} \tag{A1}$$

Step 2: As we obtain $\widehat{H}(u_k; s)$, $k = 1, \dots, K_s$, from Step 1, we can estimate $\widehat{\boldsymbol{\theta}}(s)$ by solving

$$\sum_{i=1}^n \sum_{j=1}^{m_i} \sum_{k=1}^{K_s} w_{ij}(s)\widetilde{\boldsymbol{Z}}_i^p(t_{ij}-s)[dN_i(u_k; t_{ij}) - Y_i(u_k; t_{ij})(\Lambda_\epsilon\{\widehat{H}(u_k; s)|\boldsymbol{\theta}(s)\} - \Lambda_\epsilon\{\widehat{H}(u_{k-1}; s)|\boldsymbol{\theta}(s)\})] = \mathbf{0}\,, \tag{A2}$$

where we assume $\Lambda_\epsilon\{\widehat{H}(u_0;s)|\boldsymbol{\theta}(s)\}=0$.

Then, we substitute $\boldsymbol{\theta}_0(s)$ with the new estimate of $\widehat{\boldsymbol{\theta}}(s)$ obtained in Step 2 and repeat Steps 1 & 2 until the specified convergence criteria are met.

For computational simplicity and efficiency, this algorithm can be simplified by the following linear approximation to avoid numerically solving the equations in (A1)

$$\Lambda_\epsilon\{H(u_k;s)|\boldsymbol{\theta}_0(s)\}-\Lambda_\epsilon\{H(u_{k-1};s)|\boldsymbol{\theta}_0(s)\}\approx\lambda_\epsilon\{H(u_{k-1};s)|\boldsymbol{\theta}_0(s)\}\Delta H(u_k;s)\,,$$

where $\Delta H(u_k;s)=H(u_k;s)-H(u_{k-1};s)$. Details are referred to Chen et al. (2002). In addition, if a landmark PH or PO model is desired, standard software for the Cox model or proportional odds model can be employed to solve the pooled score functions at t_{ij} weighted by kernel weights $w_{ij}(s)=K_h(t_{ij}-s)$.

Appendix B. General Data Generation Method for Simulations in Section 3.

Let $\psi_i(s)=\boldsymbol{Z}_i'(s)\boldsymbol{\beta}(s)$ and write the landmark linear transformation model as

$$P(T_{is}>u|\psi_i(s))=g^{-1}\{H(u;s)+\psi_i(s)\}\,. \tag{A3}$$

First, we generate failure time T_i at baseline $s=0$. Given the marginal distribution of $\psi_i(s)$ for $s=0$ and the function $H(t;0)$, we can generate T_i from the marginal distribution of the failure times, which we obtain by taking the expectation with respect to $\psi_i(0)$, i.e.,

$$P(T_i>t)=E_{\psi(0)}[g^{-1}\{H(t;0)+\psi_i(0)\}]\,. \tag{A4}$$

Second, it is known that for an individual who is at risk at $s>0$, $P(T_i>s+u)/P(T_i>s)=P(T_i>s+u|T_i>s)=P(T_{is}>u)$ for all $u\geq 0$. Then, by (A3) and (A4), we can solve $H(u;s)$ from

$$\frac{E_{\psi(0)}\{g^{-1}[H(s+u;0)+\psi_i(0)]\}}{E_{\psi(0)}\{g^{-1}[H(s;0)+\psi_i(0)]\}}=E_{\psi(s)}[g^{-1}\{H(u;s)+\psi_i(s)\}]\,, \tag{A5}$$

provided that all the expectations involved in (A5) are finite.

Third, as we obtain $H(u;s)$ and $T_{is}=T_i-s$, we define a variable transformation as $W_i(s)=H(T_{is};s)$. It can be shown that the distribution of $W_i(s)$ conditional on $\psi_i(s)$ as $f\{W_i(s)|\psi_i(s)\}=-\frac{\partial g^{-1}\{\psi_i(s)+W_i(s)\}}{\partial W_i(s)}$, based on the distribution $f\{T_{is}|\psi_i(s)\}$ given in model (A3). Then, the conditional distribution of $\psi_i(s)$ given $W_i(s)$ is derived by Bayes' formula

$$f\{\psi_i(s)|W_i(s)\}=\frac{f\{W_i(s)|\psi_i(s)\}f\{\psi_i(s)\}}{\int_{-\infty}^{\infty}f\{W_i(s)|\psi_i(s)\}f\{\psi_i(s)\}d\psi_i(s)}\,. \tag{A6}$$

Next, we generate $\psi_i(s)$ from the distribution given in (A6), which depends on T_{is}. This data-generating procedure is applicable to any transformation model. Specifically, we describe below about proportional hazards (PH) models and proportional odds models (PO) as examples.

For A Proportional Hazards Model in Sections 3.1.

The data generation algorithm described above is numerical in general, except for landmark PH models. For a landmark PH model, parametric distribution of $f\{\psi_i(s)|W_i(s)\}$ in (A6) can be found by using a Gamma conjugate prior. We consider a landmark Cox PH model for failure times conditional on one baseline variable and three time-dependent variables, i.e., $g(x) = \log\{-\log(x)\}$ and $\boldsymbol{Z}_i(s) = (Z_1(s), Z_2, Z_3(s), Z_4(s))$. Covariate coefficients are set to be $\beta_1(s) = 0.1 + 0.2\sqrt{s}$, $\beta_2(s) = -0.6 + 0.005s$, $\beta_3(s) = 0.4 + 0.01s$, and $\beta_4(s) = -0.3$.

We let $\psi_i(s) = \exp\{\boldsymbol{Z}_i'(s)\boldsymbol{\beta}(s)\}$ and assume that the marginal distribution of $\psi_i(s)$ is a Gamma distribution with shape parameter of $\alpha(s) = 1 - 0.05\sqrt{s}$ and rate parameter of $\eta(s) = 1.5 + 0.05s$ for any given s. In addition, we assume that the baseline cumulative hazard function $\Lambda_0(u; 0) = (0.2u)^3$ (i.e., Weibull distribution) for failure times at $s = 0$. As a conjugate distribution, it can be shown that the distribution of $\psi_i(s)$ conditional on T_{is} is also a Gamma distribution with parameters $\alpha(s) + 1$ and $\eta(s) + \Lambda_0\{T_i(s); s\}$. The idea is to find the marginal distribution of T_i and generate failure times first, and then generate $\psi_i(s)$ by a conditional distribution given T_{is}. To incorporate intra-subject correlation among the longitudinal measurements, the values of $\psi_i(s)$ at different landmark times s are generated from a Gaussian copula with an exchangeable correlation structure, where the correlation coefficient is given by $\rho = 0.3$. Additionally, we let $Z_{i2} \sim \text{Bernoulli}(p = 0.5)$ and $Z_{i3}(s) \sim N(2Z_{i2} - 1, 1)$ with independent serial correlation. Finally, we assume that $Z_{i4}(s) \sim N(0, 1)$ from a Gaussian copula with exchangeable serial correlation and set $\rho = 0.3$ as well. Then, Z_{i1} is obtained by solving the equation $\psi_i(s) = \exp\{\boldsymbol{Z}_i'(s)\boldsymbol{\beta}(s)\}$.

For A Proportional Odds Model in Sections 3.2.

For a proportional odds model, i.e. $g(x) = \log\{(1 - x)/x\}$. We let $\psi_i(s) = \boldsymbol{Z}_i'(s)\boldsymbol{\beta}(s)$ and assume that $\psi_i(s) \sim \text{Uniform}(-3, 3)$ for any s. We also define $H(u; 0) = \log\{(1 - S_0(u))/S_0(u)\}$, where $S_0(u) = \exp\{-(0.12u)^3\}$ is the baseline survival function of a Weibull distribution. Failure time T_i is generated first and $\psi_i(s)$ is simulated from its distribution conditional on T_{is}. Subsequently, we generate the trajectories, $Z_{i2} \sim \text{Bernoulli}(0.5)$, $Z_{i3}(s) \sim \text{Uniform}(-3 + Z_{i2}, 3 - Z_{i2})$, and $Z_{i4}(s) \sim \text{Uniform}(-3, 3)$, and $Z_{i1}(s)$ can be solved from

Table A1: Summary of the estimates of regression coefficients in the simulation for a landmark PH model. True value, bias, bias percentage (bias.perc), and mean squared error (MSE) are presented. Sample size of the training dataset is $n = 200$; number of simulation replicates is 100.

n=200			proposed method			LVCF method		
coef	s	True	bias	bias.perc (%)	MSE	bias	bias.perc (%)	MSE
β_1	2	0.383	0.019	4.908	0.003	-0.095	24.792	0.011
β_2	2	-0.590	-0.036	6.012	0.115	0.031	5.284	0.066
β_3	2	0.420	0.029	7.009	0.018	-0.030	7.091	0.013
β_4	2	-0.300	0.017	5.523	0.013	0.040	13.454	0.010
β_1	4	0.500	0.032	6.330	0.006	-0.071	14.230	0.008
β_2	4	-0.580	-0.002	0.295	0.153	0.057	9.902	0.084
β_3	4	0.440	0.032	7.327	0.017	-0.030	6.891	0.014
β_4	4	-0.300	-0.013	4.197	0.017	0.042	13.905	0.012
β_1	6	0.590	0.046	7.748	0.013	-0.063	10.596	0.014
β_2	6	-0.570	0.024	4.266	0.298	0.103	18.140	0.198
β_3	6	0.460	-0.003	0.602	0.031	-0.045	9.839	0.037
β_4	6	-0.300	-0.021	6.831	0.041	0.011	3.569	0.023

$\psi_i(s) = \boldsymbol{Z}_i'(s)\boldsymbol{\beta}(s)$, where $\boldsymbol{\beta}(s)$ is specified the same as before. We employ a Gaussian copula to incorporate serial correlations of $\psi_i(s)$ and $X_{i4}(s)$ with an exchangeable correlation structure and set the coefficient at $\rho = 0.3$.

As a result, trajectories of the time-varying biomarkers for 20 randomly selected individuals and the estimated Kaplan-Meier marginal survival curve based on a landmark PO model are plotted in Figure A1.

Table A2: Summary of the estimates of regression coefficients in the simulation for a landmark PH model. True value, bias, bias percentage (bias.perc), and mean squared error (MSE) are presented. Sample size of the training dataset is $n = 500$; number of simulation replicates is 100.

n=500			proposed method			LVCF method		
coef	s	True	bias	bias.perc (%)	MSE	bias	bias.perc (%)	MSE
β_1	2	0.383	0.015	3.826	0.002	-0.111	28.947	0.013
β_2	2	-0.590	-0.041	6.997	0.083	0.069	11.730	0.035
β_3	2	0.420	0.018	4.156	0.008	-0.060	14.369	0.007
β_4	2	-0.300	-0.008	2.625	0.012	0.029	9.485	0.005
β_1	4	0.500	0.017	3.448	0.003	-0.082	16.315	0.008
β_2	4	-0.580	-0.020	3.457	0.076	0.036	6.172	0.025
β_3	4	0.440	0.017	3.901	0.011	-0.036	8.148	0.006
β_4	4	-0.300	0.004	1.444	0.009	0.027	8.907	0.006
β_1	6	0.590	0.026	4.400	0.007	-0.078	13.222	0.010
β_2	6	-0.570	0.029	5.007	0.104	0.067	11.755	0.074
β_3	6	0.460	0.020	4.246	0.014	-0.039	8.536	0.010
β_4	6	-0.300	-0.020	6.793	0.016	0.026	8.569	0.009

Table A3: Summary of the estimates of regression coefficients in the simulation for a landmark PO model. True value, bias, bias percentage, and mean squared error (MSE) are presented. Sample size of the training dataset is $n = 200$; number of simulation replicates is 100.

n=200			proposed method			LVCF method		
coef	s	True	bias	bias.perc (%)	MSE	bias	bias.perc (%)	MSE
β_1	2	0.383	0.020	5.164	0.004	-0.123	32.204	0.017
β_2	2	-0.590	0.010	1.682	0.217	0.110	18.696	0.140
β_3	2	0.420	0.025	6.005	0.027	-0.073	17.346	0.024
β_4	2	-0.300	-0.001	0.223	0.014	0.049	16.262	0.011
β_1	4	0.500	0.011	2.121	0.007	-0.123	24.625	0.020
β_2	4	-0.580	-0.004	0.751	0.262	0.087	14.935	0.136
β_3	4	0.440	0.007	1.678	0.022	-0.079	17.982	0.023
β_4	4	-0.300	-0.010	3.273	0.018	0.051	16.984	0.014
β_1	6	0.590	0.025	4.281	0.014	-0.113	19.094	0.019
β_2	6	-0.570	-0.082	14.329	0.276	0.055	9.580	0.158
β_3	6	0.460	0.004	0.935	0.034	-0.072	15.652	0.030
β_4	6	-0.300	-0.005	1.558	0.018	0.033	10.940	0.019

Table A4: Summary of the estimates of regression coefficients in the simulation for a landmark PO model. True value, bias, bias percentage, and mean squared error (MSE) are presented. Sample size of the training dataset is $n = 500$; number of simulation replicates is 100.

n=500			proposed method			LVCF method		
coef	s	True	bias	bias.perc (%)	MSE	bias	bias.perc (%)	MSE
β_1	2	0.383	0.010	2.480	0.003	-0.133	34.834	0.019
β_2	2	-0.590	-0.040	6.721	0.116	0.099	16.853	0.048
β_3	2	0.420	0.018	4.216	0.013	-0.087	20.751	0.014
β_4	2	-0.300	-0.013	4.443	0.010	0.055	18.313	0.007
β_1	4	0.500	0.009	1.741	0.003	-0.119	23.844	0.017
β_2	4	-0.580	-0.022	3.824	0.093	0.081	13.970	0.050
β_3	4	0.440	-0.013	2.991	0.011	-0.070	15.875	0.011
β_4	4	-0.300	-0.010	3.342	0.012	0.051	17.102	0.007
β_1	6	0.590	0.006	0.971	0.007	-0.097	16.516	0.013
β_2	6	-0.570	-0.023	4.104	0.116	0.053	9.214	0.070
β_3	6	0.460	0.005	1.078	0.014	-0.038	8.260	0.012
β_4	6	-0.300	0.008	2.490	0.010	0.046	15.356	0.010

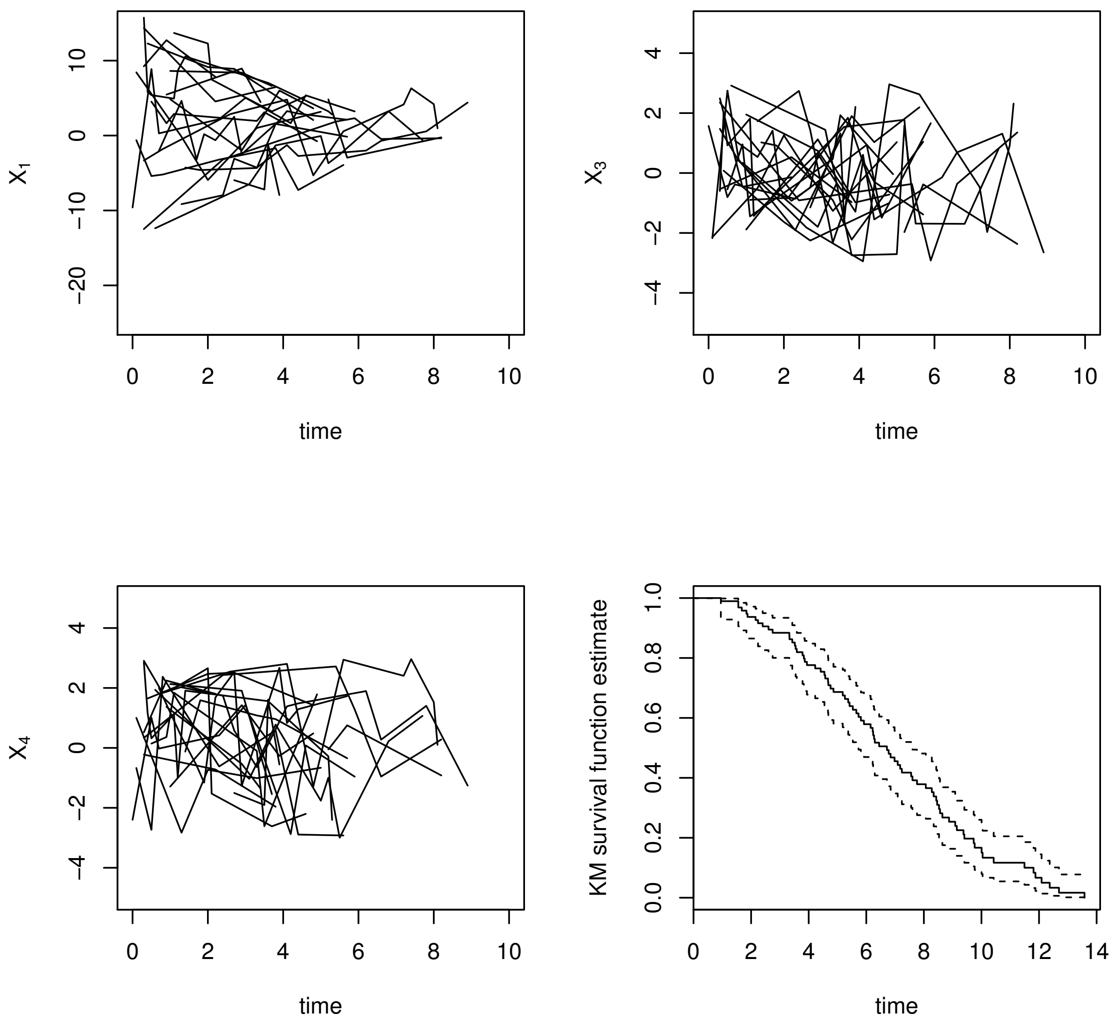

Figure A1: Trajectories of $X_1(t)$, $X_3(t)$, and $X_4(t)$ for 20 randomly selected subjects and the Kaplan-Meier (KM) estimate of the marginal survival curve for the whole sample of training data based on a landmark PO model from one simulation.

Appendix C. Additional Simulation Results in Sections 3.1 & 3.2.

References

Chen, K., Jin, Z. and Ying, Z. (2002) Semiparametric analysis of transformation models with censored data. *Biometrika*, **89**, 659–668.