8-1-2014

A Study of Joinpoint Models for Longitudinal Data

Libo Zhou
University of Nevada, Las Vegas

A STUDY OF JOINPOINT MODELS FOR LONGITUDINAL DATA

by

Libo Zhou

July 21, 2014

This dissertation was prepared by using the class file unlvmathesis.cls according to the requirement of dissertation format of the Graduate College of the University of Nevada at Las Vegas, USA.

unlvmathesis.cls: A LaTeX Class for UNLV–MATH thesis and dissertation.

Version: 2.0

Authors: Anthony D. Holmes and Hongtao Yang

Dates: March 13, 2010 & April 29, 2014

A STUDY OF JOINPOINT MODELS FOR LONGITUDINAL DATA

by

Libo Zhou

Bachelor of Arts in Actuarial Mathematics
Concordia University, Canada
2002

Master of Science in Mathematics
University of Nevada, Las Vegas, USA
2006

A dissertation submitted in partial fulfillment of
the requirements for the

Doctor of Philosophy - Mathematical Sciences

Department of Mathematical Sciences
College of Sciences
The Graduate College

University of Nevada, Las Vegas
August 2014

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

THE GRADUATE COLLEGE

We recommend the dissertation prepared under our supervision by

Libo Zhou

entitled

A Study of Joinpoint Models for Longitudinal Data

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Mathematical Sciences

Department of Mathematical Sciences

Kaushik Ghosh, Ph.D., Committee Chair

Chih-Hsiang Ho, Ph.D., Committee Member

Hokwon Cho, Ph.D., Committee Member

Malwane Ananda, Ph.D., Committee Member

Guogen Shen, Ph.D., Graduate College Representative

Kathryn Hausbeck Korgan, Ph.D., Interim Dean of the Graduate College

August 2014

ABSTRACT

A STUDY OF JOINPOINT MODELS FOR LONGITUDINAL DATA

by

Libo Zhou

Dr. Kaushik Ghosh, Examination Committee Chair
Associate Professor, Biostatistics
University of Nevada, Las Vegas, USA

In many medical studies, data are collected simultaneously on multiple biomarkers from each individual. Levels of these biomarkers are measured periodically over certain time duration, giving rise to longitudinal trajectories. The subjects under study may also be subject to dropout due to several competing causes, the likelihood of which may be affected by the levels of these biomarkers.

In this dissertation, we investigate flexible Bayesian modeling of such data, taking into account any available covariate information as well as possible censoring of the drop-out times. We propose joint models for multiple biomarkers with multiple causes of dropout. Our proposed models allow the trajectories to have multiple joinpoints, the locations of which are estimated from the data. We explore two ways of modeling longitudinal data incorporating the dropout information. Dirichlet process priors are used to make the models robust to misspecification. The Dirichlet process also leads to a natural clustering of subjects with similar trajectories, which can be of importance in efficiently estimating the joinpoints.

Efficient Markov chain Monte Carlo algorithms are developed for fitting the proposed models. The performance of all the methods is investigated through simulation

studies. One of the proposed models is seen to give rise to improved estimates of individual trajectories. Data from ACTG 398 study is used to illustrate the applicability of that model.

ACKNOWLEDGEMENTS

First, I give my most sincere thanks to my advisor, Dr. Kaushik Ghosh. For more than three years, he has patiently guided and inspired me to explore my dissertation work along the way. His support and encouragement was indispensable to the completion of this dissertation. At the same time, his attitudes towards research and collaboration provided me with a living example, which I will aspire to follow in my future career.

I also give my sincere thanks to Dr. Anton Westveld. His guidance, teaching of statistical knowledge and training of R program are essential for me to complete this dissertation.

I would like to thank Dr. Malwane Ananda, Dr. Chih-Hsiang Ho, Dr. Hokwon Cho and Dr. Guogen Shan for serving as my dissertation committee. Their advice and comments are so valuable. I am also very grateful to Dr. Malwane Ananda, Dr. Chih-Hsiang Ho and Dr. Hokwon Cho for all the classes they have taught me and their mentoring for all these years.

I would also like to thank Dr. Amei Amei for her valuable suggestions and advice regarding my work. Her attitudes towards research and teaching is also a living example to me.

Finally I would like to thank my wife, Xiaolin Wang, for her dedication and sacrifice, her love when I needed it most, her encouragement when I was down, and her blind trust in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In many medical studies, both longitudinal biomarker and survival data are collected on each subject. These observed biomarker series are important health indicators that represent the progression of a disease. Such data will typically have additional features and complications associated with them, including the presence of treatment group indicators and baseline covariates, measurement error in the biomarkers, and right censoring of the event time with the possibility of dependent censoring. The goals for studies with data of these types can be quite variable. The goal might be assessing how the biomarker changes with time and how this is influenced by the baseline covariates; it might be determining how the risk of the event is influenced by the biomarker and the covariates; it could be determining whether the biomarker can be used as a surrogate endpoint or as an auxiliary variable in a clinical trial, or whether it could be used to make individual predictions of future event times for patients who are censored.

Joint models are frequently used in survival analysis to assess the relationship between time-to-event data and some time-independent and time-dependent covariates that are measured longitudinally but often with error. A common framework consists of using the Cox regression model for the survival time data and a linear mixed-effects

model for the longitudinal observations. This dissertation investigates joint models with multiple biomarkers and multiple causes of dropout. Our proposed model will allow the trajectories to have multiple joinpoints or changepoints, the locations of which will be estimated from the data. Dirichlet process priors will be used to model the distribution of the individual random effects. That will lead to a natural clustering of subjects with similar trajectories, which can be of importance in efficiently estimating the changepoints.

1.2 Past work

Methods for jointly modeling longitudinal and survival data are recently becoming more popular. Tsiatis et al. (1995) used the Cox proportional hazards regression model to study the relationship between CD4 counts as a time-dependent covariate and survival. A two-stage procedure by plugging the estimates for modeling the longitudinal data was proposed. Although their model reduces bias compared to using the raw covariate data directly in a Cox model, a likelihood approach based on specification of a joint likelihood may make more efficient use of data. This joint likelihood is constructed by assuming conditional independence of the longitudinal and survival data, given the longitudinal trajectory. The trajectory function represents the true latent longitudinal measures derived. They also considered methods to account for missing data patterns. Wulfsohn and Tsiatis (1997) implemented an EM algorithm to fit a proportional hazards model for survival, conditional on the latent trajectory function. Faucett and Thomas (1996) adopted a Bayesian Markov chain Monte Carlo

(MCMC) technique to do the estimation for this model. Kiuchi et al. (1995) presented a changepoint model to estimate the distribution of the time before AIDS when rapid decline begins. They proposed both empirical and hierarchical Bayes changepoint models using the EM algorithm and Markov chain Monte Carlo technique to estimate the parameters. Henderson et al. (2000) introduced a stationary Gaussian process allowing the trajectory to vary with time. They developed a flexible methodology for handling combined longitudinal and event history data, incorporating the most commonly used first-choice assumptions from both subject areas. A latent bivariate Gaussian process $W(t) = \{W_1(t), W_2(t)\}$ was postulated assuming that the measurement and event process are conditionally independent given $W(t)$ and covariates. Linear random effects models and EM estimations were used to estimate the parameters. Wang and Taylor (2001) used a mixed effects model but proposed an integrated Ornstein-Uhlenbeck (IOU) process for longitudinal CD4 count data in a joint model. Using random intercepts and fixed slope, they used IOU process to allow the path of an individual's biomarker, also known as the trajectory, to fluctuate around a straight line. It is an improvement in the fit of the longitudinal markers, but we need models that can provide more flexibility than this model.

Tsiatis and Davidian (2001) specified a nonparametric distribution for the random effects from a frequentist perspective. They assumed that the survival is related to covariate through a proportional hazards relationship with the underlying random effects. They developed a simple method for inference that does not put any restrictions on the distribution of random effects by exploring the conditional score approach of Stefanski & Carroll (1987). Brown and Ibrahim (2003) specified a nonparametric

distribution for the random effects from a Bayesian perspective. A new semiparametric Bayesian hierarchical model for the joint modeling for longitudinal and survival data was proposed. A Dirichlet process prior on the parameters defining the longitudinal model was used to relax the distributional assumptions for the longitudinal model. This makes the posterior distribution of the longitudinal parameters free of parametric constraints which results in more robust estimates. Lin et al. (2002) developed a latent class model that allowed the polynomial trajectory to depend on class membership. These approaches do allow for more flexibility for modeling the longitudinal data. However, they still impose parametric assumptions on the path of an individual's longitudinal marker.

Xu and Zeger (2001), Song et al. (2002) and Ibrahim et al. (2004) all extended the longitudinal model to the multivariate case. Sinha et al. (2001) provide a detailed discussion of joint modeling. Xu and Zeger (2001) proposed a joint model for a time to clinical event and for repeated measures over time on multiple biomarkers that are potential surrogates. Two complementary measures of the relative benefit of multiple surrogates were proposed as opposed to a single one. Song et al. (2002) proposed a semiparametric likelihood approach for a joint model for survival and longitudinal data in which parametric assumptions on the distribution of random effects may be relaxed to that of a smooth density. An important feature of the procedure is that it makes possible the study of robustness to parametric assumptions on the random effects in joint models. Ibrahim et al. (2004) developed a Bayesian joint model for multivariate longitudinal and survival data. A model assessment tool called the multivariate L measure was presented which allowed them to formally compare different

models.

Brown et al. (2005) proposed a joint longitudinal and survival model that has a nonparametric model for the longitudinal biomarkers. Cubic B-splines were used to specify the longitudinal model and a proportional hazards model to link the longitudinal measure to the hazards. They also used the Conditional Predictive Ordinate (CPO) and the Deviance Information Criterion (DIC) to select the number of knots for the cubic B-spline model. The method was applied to examine the link between viral load, CD4 counts, and time to event in data from an AIDS clinical trial. The longitudinal cubic B-spline model for a single longitudinal outcome was extended to the Bayesian settings. The univariate Bayesian B-spline model was generalized to accommodate a multivariate outcome. They also showed how to incorporate this multivariate B-spline model in a joint longitudinal and survival.

Let Y_{ij} be the ith subject's set of observed biomarkers at time t_{ij}, the longitudinal cubic B-spline model was extended to the multivariate case as follows:

$$Y_{ij} = \psi_{\alpha,\beta}(t_{ij}) + \epsilon_{ij},$$

where

$$\psi_{\alpha,\beta}(t_{ij}) = \sum_{k=1}^{q} \boldsymbol{\beta}_{ik} B_k(t_{ij}) + \boldsymbol{x}_i^{'} \boldsymbol{\alpha}.$$

Furthermore,

$$\boldsymbol{\beta}_{ik} \sim N_p(\boldsymbol{b}_{0k}, \boldsymbol{V}_{0k})$$

where $\boldsymbol{\beta}_{ik} = (\beta_{ik1}, \ldots, \beta_{ikp})^{'}, \boldsymbol{b}_{0k} = (b_{0k1}, \ldots, b_{0kp})^{'}$, $\boldsymbol{V}_{0k}$ is a $p \times p$ covariate matrix, $\boldsymbol{\epsilon}_{ij} \sim N_p(0, \Sigma)$, and $N_p(\boldsymbol{a}, \boldsymbol{b})$ is the p-dimensional multivariate normal distribution

with mean vector $\boldsymbol{a}$ and covariance matrix $\boldsymbol{b}$. $B_k(t_{ij})$ denotes the value of the kth basis function at time t_{ij}, $\boldsymbol{\alpha}$ is a vector of parameters linking the vector of baseline covariates $\boldsymbol{x}_i$ to the longitudinal outcome.

With p being the number of longitudinal outcomes measured at a time point, the observation, trajectory, and error vectors for subject i are then,

$$\begin{aligned}
\boldsymbol{Y}_{ij} &= \begin{pmatrix} Y_{ij1} \\ \vdots \\ Y_{ijp} \end{pmatrix}, \\
\boldsymbol{\psi}_{\alpha,\beta}(t_{ij}) &= \begin{pmatrix} \psi_{\alpha,\beta,1}(t_{ij}) \\ \vdots \\ \psi_{\alpha,\beta,p}(t_{ij}) \end{pmatrix} \\
&= \begin{pmatrix} \sum_{k=1}^{q} \beta_{ik1} B_k(t_{ij}) + x_i^{'}\alpha_1 \\ \vdots \\ \sum_{k=1}^{q} \beta_{ikp} B_k(t_{ij}) + x_i^{'}\alpha_p \end{pmatrix}, \\
\text{and} \qquad \boldsymbol{\epsilon}_{ij} &= \begin{pmatrix} \epsilon_{ij1} \\ \vdots \\ \epsilon_{ijp} \end{pmatrix}.
\end{aligned}$$

The hazard function was expressed as

$$h(t|Y) = \lambda(t)\exp(\boldsymbol{\gamma}^{'}\psi_\beta(t) + \boldsymbol{z}^{'}\boldsymbol{\zeta})$$

where $\psi_\beta(t) = \psi_{\alpha=0,\beta}(t)$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)^{'}$ is a vector of parameters linking the trajectory to the hazard function, $\lambda(t)$ is the baseline hazard, and $\boldsymbol{\zeta}$ is the parameter vector linking a vector $\boldsymbol{z}$ of baseline covariates to the failure time. The contribution to likelihood for an individual's time to event, s_i is given by

$$\begin{aligned}
f(s_i, \nu_i|\boldsymbol{Y}_i) =& \lambda(s_i)^{\nu_i}\exp\{\nu_i(\boldsymbol{\gamma}^{'}\psi_\beta(s_i) + \boldsymbol{z}_i^{'}\boldsymbol{\zeta})\} \\
& \times \exp\left\{ -\int_0^{s_i} \lambda(u) e^{\boldsymbol{\gamma}^{'}\psi_\beta(u) + \boldsymbol{z}_i^{'}\boldsymbol{\zeta}} \mathrm{d}u \right\},
\end{aligned}$$

where ν_i is the censoring indicator for subject i.

Assuming a piecewise constant baseline hazard function

$$\lambda(u) = \lambda_j, \quad u_{j-1} < u < u_j, \quad j = 1, \ldots, J,$$

we get

$$\begin{aligned} f(s_i, \nu_i | \boldsymbol{Y}_i) =& \lambda(s_i)^{\nu_i} \exp\{\nu_i(\boldsymbol{\gamma}' \psi_\beta(s_i) + \boldsymbol{z}_i' \boldsymbol{\zeta})\} \\ & \times \exp\left\{ - e^{\boldsymbol{z}_i' \boldsymbol{\zeta}} \sum_{j=1}^{J} H_{ij}(\beta, \boldsymbol{\gamma}, \lambda) \right\}, \end{aligned}$$

where

$$H_{ij}(\beta, \boldsymbol{\gamma}, \lambda) = I_{\{s_i \geq u_{j-1}\}} \int_{u_{j-1}}^{u_j \wedge s_i} e^{\boldsymbol{\gamma}' \psi_\beta(u) + \boldsymbol{z}_i' \boldsymbol{\zeta}} \mathrm{d}u$$

Chi and Ibrahim (2006) proposed a likelihood-based approach to extend both longitudinal and survival components. A multivariate mixed effects model was presented to explicitly capture two different sources of dependence among longitudinal measures over time as well as dependence between different variables. For the survival component of the joint model, they introduced a shared frailty, which is assumed to have a positive stable distribution, to induce correlation between failure times. The proposed marginal univariate survival model, which accommodates both zero and nonzero cure fractions for the time to event, was applied to each marginal survival function. The proposed multivariate survival model had a proportional hazards structure for the population hazard, conditionally as well as marginally, when the baseline covariates were entered biologically through the mean function of the Poisson process. The model was capable of dealing with survival functions with different cure rate structures and thus accommodated a mixture of proper and improper survival functions.

The model was computationally feasible with the incorporation of longitudinal data. They developed MCMC algorithms to sample from the joint posterior by introducing latent variables. With the use of the modified version of the collapsed Gibbs technique, the computational development facilitated an efficient Gibbs sampling scheme for the posterior distribution. A new bivariate survival model was discussed and the simulation study was conducted to examine the feasibility as well as properties of the proposed multivariate survival model.

Ghosh et al. (2009) developed a joint model of longitudinal data and informative dropout time in the presence of multiple changepoints. A multiple-joinpoint model was proposed for the longitudinal response and a Cox proportional hazards model was used to connect the longitudinal part to the dropout part, giving rise to unbiased estimates of the underlying parameters. The model has multiple known changepoints with univariate biomarker and univariate dropout cause. Dirichlet process (DP) priors are used to model the distribution of the individual random effects and error distribution. This allowed for clustering of trajectories, which can help in identifying patients with similar trajectories and thus pool information to get better estimates. The baseline hazard for the survival model was assumed to have a piecewise constant structure.

The joinpoint model was assumed to have the following form:

$$y_{ij} = \psi_i(t_{ij}) + e_{ij},$$

where the random errors were assumed to have independent zero-mean normal distributions as follows:

$$e_{ij}|(\sigma_e^2, \eta_i) \overset{iid}{\sim} N(0, \frac{\sigma_e^2}{\eta_i}), \quad i = 1, \ldots, N, \quad j = 1, \ldots, n_i$$

The trajectory function was assumed as the following form:

$$\psi_i(t) = x_{i1}\alpha_1 + \cdots + x_{ic}\alpha_c + \delta_{i0} + \delta_{i1}t + \sum_{l=1}^{L} \beta_{il}(t - \tau_l)_+,$$

where $\tau_1 < \cdots < \tau_L$ are completely known joinpoints and $\alpha_1, \ldots, \alpha_c$ are the fixed effects of the c covariates. The notation u_+ is used to define a spline:

$$u_+ = \begin{cases} u & \text{if } u > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The baseline hazard function was assumed piecewise constant as

$$\lambda_0(u) = \lambda_{0j}, \quad u_{j-1} < u < u_j, \quad j = 1, \ldots, J,$$

where the number of steps J and the endpoints u_j are pre-specified with $u_0 \equiv 0$ and $u_J \equiv \infty$.

The contribution to the likelihood from survival part of ith subject was then,

$$\begin{aligned} f(s_i, \nu_i|\psi_i) =& \lambda_0(s_i)^{\nu_i} \exp\{\nu_i \gamma \psi_i(s_i)\} \\ & \times \exp\Big\{ - \int_0^{s_i} \lambda_0(u) \exp\big(\gamma \psi_i(u)\big) \mathrm{d}u \Big\}, \end{aligned}$$

which was rewritten as

$$f(s_i, \nu_i|\psi_i) = \lambda_0(s_i)^{\nu_i} \exp\Big\{ \nu_i \gamma \psi_i(s_i) - \sum_{j=1}^{L} \lambda_{0j} H_{ij}(\gamma) \Big\},$$

where

$$H_{ij}(\gamma) = I_{\{s_i \geq u_{j-1}\}} \int_{u_{j-1}}^{u_j \wedge s_i} \exp\big(\gamma \psi_j(u)\big) \mathrm{d}u.$$

Fieuws and Verbeke (2006) and Fieuws et al. (2007) had proposed an approach for modeling multivariate longitudinal data whereby all possible pairs of longitudinal data were separately modeled and were then combined in a final step. Albert and Shih (2010) proposed a regression calibration approach for jointly modeling multiple longitudinal measurements and discrete time-to-event data. A regression calibration approach which appropriately accounts for informative dropout was proposed in that article. An approach was proposed for jointly modeling multivariate longitudinal and discrete time-to-event data which easily accommodates many longitudinal biomarkers. Complete data are then simulated based on estimates from these pairwise conditional models, and regression calibration is used to estimate the relationship between longitudinal data and time-to-event data using the complete data.

Martinez-Beneito et al. (2011) introduced a reparametrization of the usual joinpoint regression model, which made it convenient as a first step to assign prior distributions. Starting from the reparametrization proposed, they introduced a longitudinal modeling proposal with an unknown number of joinpoints. That proposal was carried out as a variable selection process, and prior distributions was discussed in detail in order to get reasonable results from the previous model selection problem.

The model proposed was the following:

$$Y_i \sim \text{Poisson}(\mu_i), \quad i = 1, \ldots, n,$$

with

$$\log(\mu_i) = \log(P_i) + \alpha + \beta_0(t_i - \bar{t}) + \sum_{j=1}^{J^*} \delta_j \beta_j B_{\tau_j}(t_i),$$

where

$$B_{\tau_j}(t) = \begin{cases} a_{0j} + b_{0j}t, & t \leq 0, \\ a_{1j} + b_{1j}t, & t > 0. \end{cases}$$

where Pi is the population under study during year i. The quantity of $B_{\tau_j}(t)$ was restricted using a number of conditions, which made it unambiguously determined. Rizopoulos (2012) provided an overview of the theory and application of joint models for longitudinal and survival data. The focus was on random effects joint model that uses latent variables to capture the association between the two outcomes and several extensions were also presented. All the analyses included in the book were implemented in the R software for statistical computing and graphics, using the freely available package JM written by the author.

1.3 Organization

The organization of this dissertation is as follows: In Chapter 2, we present a longitudinal model formulation and do the estimation of the parameters. In Chapter 3, we combine the longitudinal model with the dropout part and do the estimation of the parameters. In Chapter 4, we present an alternative reparametrization of the longitudinal model and do the estimation of the parameters. In Chapter 5, we combine the alternative parametrization model with the dropout part and do the estimation of the parameters. The performance of all the methods is investigated through simulation studies. In Chapter 6, we analyze part of the ACTG 398 data and the results of the analysis are presented. In Chapter 7, we present the conclusions of our studies.

CHAPTER 2

LONGITUDINAL MODEL FOR MULTIPLE BIOMARKERS

In this chapter, we present a longitudinal model for multiple biomarkers, each with multiple joinpoints. Each subject is modeled using random spline coefficients that are normally distributed around a common population mean. Dirichlet process priors are used to make the models robust to misspecification.

2.1 Longitudinal model with multiple joinpoints

Consider a scenario in which there are several subjects, with each subject being measured on multiple biomarkers repeatedly over time. On each subject, we also measure certain covariates, which we expect to influence the biomarker levels. Additionally, the biomarker trajectories are assumed to have several joinpoints.

Let y_{ijk} denote the kth biomarker response of subject i measured at time t_{ij}, ($i = 1, \ldots, N; \quad j = 1, \ldots, n_i; \quad k = 1, \ldots, K.$). We assume a model of the following form:

$$y_{ijk} = \psi_{ik}(t_{ij}) + e_{ijk},$$

where $\psi_{ik}(\cdot)$ is the trajectory function for the kth biomarker response of the ith individual and e_{ijk} is the associated random error.

Let there be C covariates measured on each subject, with those for subject i being denoted by $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iC})'$. Additionally, let there be L joinpoints for each

biomarker trajectory, with those for biomarker k being $\tau_{k1}, \ldots, \tau_{kL}$ and $\tau_{k1} < \ldots < \tau_{kL}$. We assume a trajectory function of the following form:

$$\psi_{ik}(t) = \boldsymbol{\alpha}_k^{'} \boldsymbol{x}_i + \delta_{i0k} + \delta_{i1k} t + \sum_{l=1}^{L} \beta_{ikl} (t - \tau_{kl})_+,$$

where the notation u_+ is used to define:

$$u_+ = \begin{cases} u & \text{if } u > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Here, $\boldsymbol{\alpha}_k = (\alpha_{k1}, \ldots, \alpha_{kC})^{'}$ are the effects of the C covariates on the kth biomarker. For subject i and biomarker k, δ_{i0k} can be interpreted as the random intercept, δ_{i1k} as the random slope and β_{ikl} as the random change in slope at the potential joinpoint τ_{kl}.

The previous model can be rewritten as:

$$\begin{pmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijK} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1^{'} \\ \boldsymbol{\alpha}_2^{'} \\ \vdots \\ \boldsymbol{\alpha}_K^{'} \end{pmatrix} \boldsymbol{x}_i + \begin{pmatrix} \delta_{i01} & \delta_{i11} \\ \delta_{i02} & \delta_{i12} \\ \vdots & \vdots \\ \delta_{i0K} & \delta_{i1K} \end{pmatrix} \begin{pmatrix} 1 \\ t_{ij} \end{pmatrix}$$
$$+ \begin{pmatrix} \sum_{l=1}^{L} \beta_{i1l} (t_{ij} - \tau_{1l})_+ \\ \sum_{l=1}^{L} \beta_{i2l} (t_{ij} - \tau_{2l})_+ \\ \vdots \\ \sum_{l=1}^{L} \beta_{iKl} (t_{ij} - \tau_{Kl})_+ \end{pmatrix} + \begin{pmatrix} e_{ij1} \\ e_{ij2} \\ \vdots \\ e_{ijK} \end{pmatrix},$$

where

$$\boldsymbol{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iC} \end{pmatrix}, \quad i = 1, \ldots, N \quad \text{and} \quad \boldsymbol{\alpha}_k = \begin{pmatrix} \alpha_{k1} \\ \alpha_{k2} \\ \vdots \\ \alpha_{kC} \end{pmatrix}, \quad k = 1, \ldots, K.$$

Note that if $\beta_{ikl} = 0$, there is no change in slope of the trajectory at the location τ_{kl}, implying that the corresponding location is not a joinpoint after all. Hence, the

above model allows for the possibility of having different number of joinpoints across subjects and biomarkers.

The random error $\boldsymbol{e}_{ij}$ is assumed to have independent zero-mean multivariate normal distributions as follows:

$$\boldsymbol{e}_{ij} \sim \boldsymbol{N}_K(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma}$ is a $K \times K$ covariance matrix.

Define $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}, \boldsymbol{Y}_{i2}, \ldots, \boldsymbol{Y}_{i,n_i})'$ to be the observations on the ith subject, with $\boldsymbol{Y}_{ij} = (y_{ij1}, \ldots, y_{ijK})'$. Also define $\boldsymbol{\phi}_i = (\boldsymbol{\phi}_{i1}, \ldots, \boldsymbol{\phi}_{in_i})'$, $\boldsymbol{\phi}_{ij} = (\psi_{i1}(t_{ij}), \ldots, \psi_{iK}(t_{ij}))'$. The contribution of the ith subject to the likelihood is then:

$$f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \propto \frac{1}{|\boldsymbol{\Sigma}|^{\frac{n_i}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \right\}.$$

2.2 Prior distributions

Joinpoint problems often give rise to irregular likelihoods, making them difficult to handle using the standard maximum-likelihood setup. A natural choice to handle complex models as above is to use a Bayesian approach, which is what we use throughout this dissertation. Specifications of prior distributions of the model parameters are provided in this section.

Let the parameters corresponding to the first biomarker of the ith subject be denoted by $\boldsymbol{\theta}_{i,1} = (\delta_{i01}, \delta_{i11}, \beta_{i11}, \beta_{i12}, \ldots, \beta_{i1L})'$. We will use a Dirichlet process prior for the common underlying distribution of the parameters $\boldsymbol{\theta}_{i,1}$. The natural clustering property for Dirichlet Process will be helpful in ascertaining clusters of subjects that have the same mean trajectory. Another advantage of Dirichlet Process prior choice

is the availability of a rich class of algorithms for drawing posterior samples. It also gives rise to flexible models robust to misspecification.

We model the prior distribution of $\boldsymbol{\theta}_{i,1}$ as follows:

$$\boldsymbol{\theta}_{i,1} \overset{iid}{\sim} G$$

$$G \sim DP(M_\theta, G_0)$$

$$G_0 \equiv \boldsymbol{N}_{L+2}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta),$$

where, $\boldsymbol{\mu}_\theta \sim \boldsymbol{N}(\boldsymbol{\mu}_{\theta 0}, \boldsymbol{\Sigma}_{\theta 0})$ and $\boldsymbol{\Sigma}_\theta \sim \boldsymbol{IW}(\nu_{\theta 0}, \boldsymbol{S}_{\theta 0})$.

Let the parameters corresponding to the rest of the biomarkers (other than the first) for the ith subject be denoted by

$$\begin{aligned}\boldsymbol{\theta}_{i,-1} &= (\delta_{i02}, \delta_{i12}, \beta_{i21}, \beta_{i22}, \ldots, \beta_{i2L}, \ldots, \delta_{i0K}, \delta_{i1K}, \beta_{iK1}, \beta_{iK2}, \ldots, \beta_{iKL})' \\ &\equiv (\boldsymbol{\theta}_{i,2}, \ldots, \boldsymbol{\theta}_{i,K})' \quad \text{(say)}.\end{aligned}$$

We assume

$$\boldsymbol{\theta}_{i,2}, \ldots, \boldsymbol{\theta}_{i,K} \overset{iid}{\sim} \boldsymbol{N}_{(L+2)}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta),$$

where, as before

$$\boldsymbol{\mu}_\theta \sim \boldsymbol{N}(\boldsymbol{\mu}_{\theta 0}, \boldsymbol{\Sigma}_{\theta 0}),$$

$$\text{and} \quad \boldsymbol{\Sigma}_\theta \sim \boldsymbol{IW}(\nu_{\theta 0}, \boldsymbol{S}_{\theta 0}).$$

Additionally we assume that the error covariance $\boldsymbol{\Sigma}$ has the prior $\boldsymbol{\Sigma} \sim \boldsymbol{IW}(\nu_{\Sigma 0}, \boldsymbol{S}_{\Sigma 0})$ and the covariate effects $\boldsymbol{\alpha}_k$ have the prior $\boldsymbol{\alpha}_k \overset{iid}{\sim} \boldsymbol{N}(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha) \quad k = 1, \ldots, K$.

Let the joinpoints of the trajectory for the kth biomarker be $\tau_{k1}, \tau_{k2}, \ldots, \tau_{kL}$, $k =$

$1, 2, \ldots, K$ with $0 < \tau_{k1} < \tau_{k2} < \cdots < \tau_{kL} < T^*$, where T^* is specified . Let the time intervals between the joinpoints be $D_{k1}, D_{k2}, \ldots, D_{kL}, D_{k,L+1}$. That is

$$D_{k1} = \tau_{k1} - 0,$$

$$D_{k2} = \tau_{k2} - \tau_{k1},$$

$$\vdots$$

$$D_{kL} = \tau_{kL} - \tau_{k,L-1},$$

$$\text{and} \quad D_{k,L+1} = T^* - \tau_{kL}.$$

Note that

$$D_{kl} > 0 \quad \text{and} \quad \sum_{l=1}^{L+1} D_{kl} = T^*.$$

Let

$$\boldsymbol{D}_1 = (D_{11}, D_{12}, \ldots, D_{1,L+1})'$$

$$\boldsymbol{D}_2 = (D_{21}, D_{22}, \ldots, D_{2,L+1})'$$

$$\vdots$$

$$\text{and} \quad \boldsymbol{D}_K = (D_{K1}, D_{K2}, \ldots, D_{K,L+1})'$$

Potentially for each biomarker if we put $\boldsymbol{\tau}_k$s with some distribution we can not estimate them. The clustering of $\boldsymbol{D}_k^*$s will translate of clustering of $\boldsymbol{\tau}_k$s. Clustering of $\boldsymbol{\tau}_k$ with $\boldsymbol{\tau}_{k'}$ means the corresponding biomarkers have the same changepoints. Let $\boldsymbol{D}_k^* = \boldsymbol{D}_k / T^*, \quad k = 1, \ldots, K$. We assume

$$\boldsymbol{D}_1^*, \boldsymbol{D}_2^*, \ldots, \boldsymbol{D}_K^* \overset{iid}{\sim} G,$$

where

$$G \sim DP(\alpha_0, G_0),$$

$$\text{and} \quad G_0 \equiv \text{Dirichlet}(\alpha_1, \alpha_2, \ldots, \alpha_{L+1}).$$

2.3 The full conditional of $\boldsymbol{\theta}_{i,1}|\cdot$

The above semiparametric Bayesian model does not give rise to a closed-form expression of the posterior distribution. We will use Markov chain Monte Carlo approach to approximate samples from the posterior. Fortunately, the conditional distributions are easy to sample from and we implement the Gibbs sampler. Since we have a conjugate baseline prior, we can use the Algorithm 2 from Neal (2000) to sample $\boldsymbol{\theta}_{i,1}$ from its posterior. Details of Algorithm 2 are in Appendix where the following conditional is obtained:

$$\boldsymbol{\theta}_{i,1}|\boldsymbol{\theta}_{-i,1}, \boldsymbol{Y}_i \sim \sum_{j' \neq i} q_{i,j'} \delta_{\boldsymbol{\theta}_{j'}} + r_i \boldsymbol{H}_i.$$

Here $\boldsymbol{H}_i$ is the posterior distribution for $\boldsymbol{\theta}_{i,1}$ based on the prior G_0 and observation $\boldsymbol{Y}_i$, with likelihood $L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i)$. The values of the $q_{i,j'}$ and r_i are defined by

$$q_{i,j'} = bL(\boldsymbol{\theta}_{j',1}|\boldsymbol{Y}_i),$$

$$r_i = bM_\theta \int L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i) \mathrm{d}G_0(\boldsymbol{\theta}_{i,1}),$$

where b is such that $\sum_{j' \neq i} q_{i,j'} + r_i = 1$.

Note that

$$\boldsymbol{\phi}_{ij}(t_{ij}) = \begin{pmatrix} \psi_{i1}(t_{ij}) \\ \psi_{i2}(t_{ij}) \\ \vdots \\ \psi_{iK}(t_{ij}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1^{'} \\ \boldsymbol{\alpha}_2^{'} \\ \vdots \\ \boldsymbol{\alpha}_K^{'} \end{pmatrix} \boldsymbol{x}_i + \begin{pmatrix} \boldsymbol{Z}_{ij1}^{'} & \cdots \\ \vdots & \ddots & \vdots \\ \cdots & \boldsymbol{Z}_{ijK}^{'} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta}_{i,1} \\ \boldsymbol{\theta}_{i,2} \\ \vdots \\ \boldsymbol{\theta}_{i,K} \end{pmatrix},$$

where

$$\boldsymbol{\theta}_{i,k} = \begin{pmatrix} \delta_{i0k} \\ \delta_{i1k} \\ \beta_{ik1} \\ \beta_{ik2} \\ \cdots \\ \beta_{ikL} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{Z}_{ijk} = \begin{pmatrix} 1 \\ t_{ij} \\ (t_{ij} - \tau_{k1})_+ \\ (t_{ij} - \tau_{k2})_+ \\ \vdots \\ (t_{ij} - \tau_{kL})_+ \end{pmatrix}.$$

Let $\sigma^{(k,k')}$ be the $(k,k')th$ element of $\boldsymbol{\Sigma}^{-1}$ and the likelihood of $\boldsymbol{\theta}_{i,1}$ be denoted by $L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i)$. Define $\psi_{ijk} = \psi_{ik}(t_{ij})$. Then

$$
\begin{aligned}
L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i) &= f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \\
&\propto \exp\left\{-\frac{1}{2}\sum_{j=1}^{n_i}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{j=1}^{n_i}(y_{ij1} - \psi_{ij1}, \ldots, y_{ijK} - \psi_{ijK})\Sigma^{-1}\begin{pmatrix} y_{ij1} - \psi_{ij1} \\ \vdots \\ y_{ijK} - \psi_{ijK} \end{pmatrix}\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{j=1}^{n_i}\sum_{k=1}^{K}\sum_{k'=1}^{K}(\psi_{ijk} - y_{ijk})\sigma^{(k,k')}(\psi_{ijk'} - y_{ijk'})\right\} \\
&= \exp\Bigg\{-\frac{1}{2}\sum_{j=1}^{n_i}\sum_{k=1}^{K}\sum_{k'=1}^{K}(\boldsymbol{\alpha}_k'\boldsymbol{x}_i + \boldsymbol{\theta}_{i,k}'\boldsymbol{Z}_{ijk} - y_{ijk})\sigma^{(k,k')} \\
&\qquad (\boldsymbol{\alpha}_{k'}'\boldsymbol{x}_i + \boldsymbol{\theta}_{i,k'}'\boldsymbol{Z}_{ijk'} - y_{ijk'})\Bigg\} \\
&= \exp\Bigg\{-\frac{1}{2}\sum_{j=1}^{n_i}\sum_{k'=1}^{K}(\boldsymbol{\theta}_{i,1}'\boldsymbol{Z}_{ij1} + \boldsymbol{\alpha}_1'\boldsymbol{x}_i - y_{ij1})\sigma^{(1,k')} \\
&\qquad (\boldsymbol{\theta}_{i,k'}'\boldsymbol{Z}_{ijk'} + \boldsymbol{\alpha}_{k'}'\boldsymbol{x}_i - y_{ijk'}) - \frac{1}{2}\sum_{j=1}^{n_i}\sum_{k=2}^{K}(\boldsymbol{\theta}_{i,k}'\boldsymbol{Z}_{ijk} + \boldsymbol{\alpha}_k'\boldsymbol{x}_i - y_{ijk}) \\
&\qquad \sigma^{(k,1)}(\boldsymbol{\theta}_{i,1}'\boldsymbol{Z}_{ij1} + \boldsymbol{\alpha}_1'\boldsymbol{x}_i - y_{ij1})\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&\propto \exp\Bigg\{ -\frac{1}{2}\sum_{j=1}^{n_i}\sum_{k'=1}^{K}\Big[\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})}\boldsymbol{Z}_{ijk'}^{'}\boldsymbol{\theta}_{i,k'}+\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})} \\
&\qquad\qquad (\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i-y_{ijk'})+\boldsymbol{\theta}_{i,k'}^{'}\boldsymbol{Z}_{ijk'}\sigma^{(1,k^{'})}(\boldsymbol{\alpha}_1^{'}\boldsymbol{x}_i-y_{ij1})\Big] \\
&\qquad -\frac{1}{2}\sum_{j=1}^{n_i}\sum_{k=2}^{K}\Big[\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}\boldsymbol{Z}_{ijk}^{'}\boldsymbol{\theta}_{i,k}+\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}(\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i-y_{ijk})\Big]\Bigg\} \\
&=\exp\Bigg\{ -\frac{1}{2}\Big[\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\boldsymbol{Z}_{ij1}\sigma^{(1,1)}\boldsymbol{Z}_{ij1}^{'}\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\sum_{k'=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})}\boldsymbol{Z}_{ijk'}^{'}\boldsymbol{\theta}_{i,k'} \\
&\qquad\qquad +\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\sum_{k'=1}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})}(\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i-y_{ijk'}) \\
&\qquad\qquad +\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\boldsymbol{Z}_{ij1}\sigma^{(1,1)}(\boldsymbol{\alpha}_1^{'}\boldsymbol{x}_i-y_{ij1}) \\
&\qquad\qquad +\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}\boldsymbol{Z}_{ijk}^{'}\boldsymbol{\theta}_{i,k} \\
&\qquad\qquad +\boldsymbol{\theta}_{i,1}^{'}\sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}(\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i-y_{ijk})\Big]\Bigg\} \\
&=\exp\Bigg\{ -\frac{1}{2}\boldsymbol{\theta}_{i,1}^{'}\Big[\sum_{j=1}^{n_i}\boldsymbol{Z}_{ij1}\sigma^{(1,1)}\boldsymbol{Z}_{ij1}^{'}\Big]\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}\left(-\frac{1}{2}\right)\Big[\sum_{j=1}^{n_i}\sum_{k'=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})}\boldsymbol{Z}_{ijk'}^{'}\boldsymbol{\theta}_{i,k'} \\
&\qquad\qquad +\sum_{j=1}^{n_i}\sum_{k'=1}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k^{'})}(\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i-y_{ijk'}) \\
&\qquad\qquad +\sum_{j=1}^{n_i}\boldsymbol{Z}_{ij1}\sigma^{(1,1)}(\boldsymbol{\alpha}_1^{'}\boldsymbol{x}_i-y_{ij1}) \\
&\qquad\qquad +\sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}\boldsymbol{Z}_{ijk}^{'}\boldsymbol{\theta}_{i,k} \\
&\qquad\qquad +\sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}(\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i-y_{ijk})\Big]\Bigg\} \\
&=\exp\left\{-\frac{1}{2}\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{B}_1\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{A}_1\right\},
\end{aligned}
$$

where

$$\boldsymbol{B}_1 = \sum_{j=1}^{n_i} \boldsymbol{Z}_{ij1}\sigma^{(1,1)}\boldsymbol{Z}_{ij1}^{'},$$

and

$$\begin{aligned}\boldsymbol{A}_1 = \left(-\frac{1}{2}\right)\Bigg[&\sum_{j=1}^{n_i}\sum_{k'=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k')}\boldsymbol{Z}_{ijk'}^{'}\boldsymbol{\theta}_{i,k'} + \sum_{j=1}^{n_i}\sum_{k'=1}^{K}\boldsymbol{Z}_{ij1}\sigma^{(1,k')}(\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i - y_{ijk'}) \\ &+ \sum_{j=1}^{n_i}\boldsymbol{Z}_{ij1}\sigma^{(1,1)}(\boldsymbol{\alpha}_1^{'}\boldsymbol{x}_i - y_{ij1}) + \sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}\boldsymbol{Z}_{ijk}^{'}\boldsymbol{\theta}_{i,k} \\ &+ \sum_{j=1}^{n_i}\sum_{k=2}^{K}\boldsymbol{Z}_{ij1}\sigma^{(k,1)}(\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i - y_{ijk})\Bigg].\end{aligned}$$

The baseline posterior of $\boldsymbol{\theta}_{i,1}$ is

$$H_i \propto L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i)\pi(\boldsymbol{\theta}_{i,1}),$$

which gives

$$H_i \sim \boldsymbol{N}(\boldsymbol{\mu}_{H1}, \boldsymbol{\Lambda}_{H1}),$$

where

$$\boldsymbol{\Lambda}_{H1} = (\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{B}_1)^{-1},$$

$$\text{and} \qquad \boldsymbol{\mu}_{H1} = (\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{B}_1)^{-1}(\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta + \boldsymbol{A}_1).$$

Moreover,

$$
\begin{aligned}
&\int L(\boldsymbol{\theta}_{i,1}|\boldsymbol{Y}_i)\mathrm{d}G_0(\boldsymbol{\theta}_{i,1}) \\
&\propto \frac{1}{(2\pi)^{\frac{L+2}{2}}|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\int \exp\left\{-\frac{1}{2}\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{B}_1\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{A}_1\right\} \\
&\qquad\qquad \times \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}_{i,1}-\boldsymbol{\mu}_\theta)^{'}\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{i,1}-\boldsymbol{\mu}_\theta)\right\}\mathrm{d}(\boldsymbol{\theta}_{i,1}) \\
&= \frac{1}{(2\pi)^{\frac{L+2}{2}}|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\int \exp\left\{-\frac{1}{2}\boldsymbol{\theta}_{i,1}^{'}(\boldsymbol{B}_1+\boldsymbol{\Sigma}_\theta^{-1})\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}(\boldsymbol{A}_1+\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta)\right\}\mathrm{d}(\boldsymbol{\theta}_{i,1}) \\
&\qquad\qquad \times \exp\left\{-\frac{1}{2}\boldsymbol{\mu}_\theta^{'}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta\right\} \\
&= \frac{1}{(2\pi)^{\frac{L+2}{2}}|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\int \exp\left\{-\frac{1}{2}\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{B}^*\boldsymbol{\theta}_{i,1}+\boldsymbol{\theta}_{i,1}^{'}\boldsymbol{A}^*\right\}\mathrm{d}(\boldsymbol{\theta}_{i,1})\exp\left\{-\frac{1}{2}\boldsymbol{\mu}_\theta^{'}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta\right\} \\
&= \frac{1}{(2\pi)^{\frac{L+2}{2}}|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\int \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}_{i,1}-\boldsymbol{\mu}_1)^{'}\boldsymbol{B}^*(\boldsymbol{\theta}_{i,1}-\boldsymbol{\mu}_1)\right\}\mathrm{d}(\boldsymbol{\theta}_{i,1}) \\
&\qquad\qquad \times \exp\left\{-\frac{1}{2}\boldsymbol{\mu}_\theta^{'}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta\right\}\exp\left\{\frac{1}{2}\boldsymbol{\mu}_1^{'}\boldsymbol{B}^*\boldsymbol{\mu}_1\right\} \\
&= \frac{1}{|\mathbf{B}^*|^{\frac{1}{2}}|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\exp\left\{\frac{1}{2}(\boldsymbol{\mu}_1^{'}\boldsymbol{B}^*\boldsymbol{\mu}_1-\boldsymbol{\mu}_\theta^{'}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta)\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{B}^* &= \boldsymbol{B}_1+\boldsymbol{\Sigma}_\theta^{-1}, \\
\boldsymbol{A}^* &= \boldsymbol{A}_1+\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta, \\
\text{and} \qquad \boldsymbol{\mu}_1 &= (\boldsymbol{B}^*)^{-1}\boldsymbol{A}^*.
\end{aligned}
$$

Hence, using

$$
r_i = b\cdot M_\theta\frac{1}{|\mathbf{B}^*|^{\frac{1}{2}}\cdot|\boldsymbol{\Sigma_\theta}|^{\frac{1}{2}}}\exp\left\{\frac{1}{2}(\boldsymbol{\mu}_1^{'}\boldsymbol{B}^*\boldsymbol{\mu}_1-\boldsymbol{\mu}_\theta^{'}\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta)\right\}
$$

and

$$
q_{i,j'} = b\cdot\exp\left\{-\frac{1}{2}\boldsymbol{\theta}_{j',1}^{'}\boldsymbol{B}_1\boldsymbol{\theta}_{j',1}+\boldsymbol{\theta}_{j',1}^{'}\boldsymbol{A}_1\right\},
$$

we can use Algorithm 2 to update $\boldsymbol{\theta}_{i,1}|\cdot$.

2.4 The full conditional of $\boldsymbol{\theta}_{i,-1}|\cdot$

The derivation of the conditionals of $\boldsymbol{\theta}_{i,k}(k = 2, \ldots, K)$ follows similar to the last section. The posterior of $\boldsymbol{\theta}_{i,k}$ is given by:

$$\boldsymbol{\theta}_{i,k}|\boldsymbol{Y}_i \sim \boldsymbol{N}(\boldsymbol{\mu}_{Hk}, \boldsymbol{\Lambda}_{Hk}),$$

where

$$\boldsymbol{\Lambda}_{Hk} = (\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{B}_k)^{-1},$$

$$\boldsymbol{\mu}_{Hk} = (\boldsymbol{\Sigma}_\theta^{-1} + \boldsymbol{B}_k)^{-1}(\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\mu}_\theta + \boldsymbol{A}_k),$$

with

$$\boldsymbol{B}_k = \sum_{j=1}^{n_i} \boldsymbol{Z}_{i,k}\sigma^{(k,k)}\boldsymbol{Z}_{i,k}^{'},$$

and

$$\begin{aligned}\boldsymbol{A}_k = \left(-\frac{1}{2}\right)\Bigg[&\sum_{j=1}^{n_i}\sum_{k'\neq k} \boldsymbol{Z}_{i,k}\sigma^{(k,k')}\boldsymbol{Z}_{i,k'}^{'}\boldsymbol{\theta}_{i,k'} + \sum_{j=1}^{n_i}\sum_{k'=1}^{K} \boldsymbol{Z}_{i,k}\sigma^{(k,k')}(\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i - y_{ijk'}) \\ &+ \sum_{j=1}^{n_i} \boldsymbol{Z}_{i,k}\sigma^{(k,k)}(\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i - y_{ijk}) + \sum_{j=1}^{n_i}\sum_{k'\neq k} \boldsymbol{Z}_{i,k}\sigma^{(k',k)}\boldsymbol{Z}_{i,k'}^{'}\boldsymbol{\theta}_{i,k'} \\ &+ \sum_{j=1}^{n_i}\sum_{k'\neq k} \boldsymbol{Z}_{i,k}\sigma^{(k',k)}(\boldsymbol{\alpha}_{k'}^{'}\boldsymbol{x}_i - y_{ijk'})\Bigg].\end{aligned}$$

2.5 The full conditionals of $\boldsymbol{\mu}_\theta|\cdot$ and $\boldsymbol{\Sigma}_\theta|\cdot$

Let the distinct values of $\boldsymbol{\theta}_{i,1}$ be $\boldsymbol{\theta}_{11}^*, \boldsymbol{\theta}_{12}^*, \ldots, \boldsymbol{\theta}_{1N_1^*}^*$ and the corresponding counts be $k_{11}, k_{12}, \ldots, k_{1N_1^*}$ respectively, with $k_{11} + k_{12} + \ldots + k_{1N_1^*} = N$. We assume

$$\boldsymbol{\theta}_{11}^*, \boldsymbol{\theta}_{12}^*, \ldots, \boldsymbol{\theta}_{1N_1^*}^* \overset{iid}{\sim} \boldsymbol{N}_{(L+2)}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$$

Then

$$\begin{aligned}\pi(\boldsymbol{\mu}_\theta|\cdot) &\propto L(\boldsymbol{\theta})\cdot\pi(\boldsymbol{\mu}_\theta)\\ &\propto \exp\left\{-\frac{1}{2}\sum_{l=1}^{N_1^*}\left[k_{1l}(\boldsymbol{\theta}_{1l}^*-\boldsymbol{\mu}_\theta)'\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{1l}^*-\boldsymbol{\mu}_\theta)\right]\right\}\\ &\times\exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\sum_{k=2}^{K}(\boldsymbol{\theta}_{i,k}-\boldsymbol{\mu}_\theta)'\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{i,k}-\boldsymbol{\mu}_\theta)\right\}\\ &\times\exp\left\{-\frac{1}{2}(\boldsymbol{\mu}_\theta-\boldsymbol{\mu}_{\theta0})'\boldsymbol{\Sigma}_{\theta0}^{-1}(\boldsymbol{\mu}_\theta-\boldsymbol{\mu}_{\theta0})\right\}\end{aligned}$$

After a little algebra, we get

$$\boldsymbol{\mu}_\theta|\cdot\sim\boldsymbol{N}_{(L+2)}(\boldsymbol{\mu}_{n\theta},\boldsymbol{\Lambda}_{n\theta}),$$

where

$$\boldsymbol{\Lambda}_{n\theta}=(\boldsymbol{\Sigma}_{\theta0}^{-1}+KN\cdot\boldsymbol{\Sigma}_\theta^{-1})^{-1},$$

$$\boldsymbol{\mu}_{n\theta}=(\boldsymbol{\Sigma}_{\theta0}^{-1}+KN\cdot\boldsymbol{\Sigma}_\theta^{-1})^{-1}(\boldsymbol{\Sigma}_{\theta0}^{-1}\boldsymbol{\mu}_{\theta0}+\boldsymbol{\Sigma}_\theta^{-1}\boldsymbol{\theta}),$$

with

$$\boldsymbol{\theta}=k_{11}\boldsymbol{\theta}_{11}^*+k_{12}\boldsymbol{\theta}_{12}^*+\cdots+k_{1N_1^*}\boldsymbol{\theta}_{1N_1^*}^*+\sum_{i=1}^{N}\boldsymbol{\theta}_{i,2}+\cdots+\sum_{i=1}^{N}\boldsymbol{\theta}_{i,K}.$$

Similarly, we can show that

$$\boldsymbol{\Sigma}_\theta|\cdot\sim\mathrm{IW}(\nu_{\theta0}+KN,\boldsymbol{S}_{\theta0}+\boldsymbol{S}_\theta),$$

where

$$\boldsymbol{S}_\theta=\sum_{l=1}^{N_1^*}\left[k_{1l}(\boldsymbol{\theta}_{1l}^*-\boldsymbol{\mu}_\theta)'\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{1l}^*-\boldsymbol{\mu}_\theta)\right]+\sum_{i=1}^{N}\sum_{k=2}^{K}(\boldsymbol{\theta}_{i,k}-\boldsymbol{\mu}_\theta)'\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{i,k}-\boldsymbol{\mu}_\theta)$$

2.6 The full conditional of $\boldsymbol{\Sigma}|\cdot$

The posterior of $\boldsymbol{\Sigma}|\cdot$ is given by

$$\pi(\boldsymbol{\Sigma}|\cdot) \propto \prod_{i=1}^{N} f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \cdot \pi(\boldsymbol{\Sigma}|\nu_{\Sigma 0}, \boldsymbol{S}_{\Sigma 0}),$$

from which we get

$$\boldsymbol{\Sigma}|\cdot \sim \boldsymbol{IW}(\nu_{\Sigma 0} + N, \boldsymbol{S}_{\Sigma 0} + \boldsymbol{S}_{\Sigma}),$$

where

$$\boldsymbol{S}_{\Sigma} = \sum_{i=1}^{N} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})^{'}.$$

2.7 The full conditional of $\boldsymbol{\alpha}_k|\cdot$

The full conditional of $\boldsymbol{\alpha}_k|\cdot$ is obtained using

$$\begin{aligned}
\pi(\boldsymbol{\alpha}_k|\cdot) &\propto \prod_{i=1}^{N} f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \cdot \pi(\boldsymbol{\alpha}_k|\boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_{\alpha}) \\
&\propto \exp\left\{ -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})^{'} \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \right\} \\
&\quad \times \exp\left\{ -\frac{1}{2} (\boldsymbol{\alpha}_k - \boldsymbol{\mu}_{\alpha})^{'} \boldsymbol{\Sigma}_{\alpha}^{-1} (\boldsymbol{\alpha}_k - \boldsymbol{\mu}_{\alpha}) \right\} \\
&\propto \exp\Bigg\{ -\frac{1}{2} \boldsymbol{\alpha}_k^{'} \Big(\sum_{i=1}^{N} \sum_{j=1}^{n_i} \boldsymbol{x}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{x}_i^{'} + \boldsymbol{\Sigma}_{\alpha}^{-1} \Big) \boldsymbol{\alpha}_k \\
&\qquad + \boldsymbol{\alpha}_k^{'} \Big(\sum_{i=1}^{N} \sum_{j=1}^{n_i} \boldsymbol{x}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{A}^{*} + \boldsymbol{\Sigma}_{\alpha}^{-1} \boldsymbol{\mu}_{\alpha} \Big) \Bigg\},
\end{aligned}$$

where

$$\boldsymbol{A}^{*} = \boldsymbol{Y}_{ij} - \delta_{i0k} - \delta_{i1k} t - \sum_{l=1}^{L} \beta_{ikl} (t - \tau_{kl})_{+}.$$

Hence

$$\boldsymbol{\alpha}_k|\cdot \sim \boldsymbol{N}(\boldsymbol{\mu}_{\alpha k}, \boldsymbol{\Lambda}_{\alpha k}),$$

where

$$\boldsymbol{\Lambda}_{\alpha k} = \Big(\sum_{i=1}^{N} \sum_{j=1}^{n_i} \boldsymbol{x}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{x}_i^{'} + \boldsymbol{\Sigma}_{\alpha}^{-1} \Big)^{-1},$$

$$\boldsymbol{\mu}_{\alpha k} = \Big(\sum_{i=1}^{N} \sum_{j=1}^{n_i} \boldsymbol{x}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{x}_i^{'} + \boldsymbol{\Sigma}_{\alpha}^{-1} \Big)^{-1} \Big(\sum_{i=1}^{N} \sum_{j=1}^{n_i} \boldsymbol{x}_i \boldsymbol{\Sigma}^{-1} \boldsymbol{A}^{*} + \boldsymbol{\Sigma}_{\alpha}^{-1} \boldsymbol{\mu}_{\alpha} \Big).$$

2.8 Estimating the joinpoints

The baseline prior distribution of $\boldsymbol{D}_k^*$ is:

$$\begin{aligned} \pi(\boldsymbol{D}_k^*) &= f(D_{k1}^*, D_{k2}^*, \ldots, D_{k,L+1}^*) \\ &= \frac{\Gamma\left(\sum_{l=1}^{L+1} \alpha_l \right)}{\prod_{l=1}^{L+1} \Gamma\left(\alpha_l\right)} \prod_{l=1}^{L+1} \left(D_{kl}^*\right)^{\alpha_l - 1}, \quad D_{kl}^* \geq 0 \quad \text{and} \quad \sum_{l=1}^{L+1} D_{kl}^* = 1. \end{aligned}$$

The corresponding baseline posterior of $\boldsymbol{D}_k^* | \cdot$ is given by

$$\pi(\boldsymbol{D}_k^* | \cdot) \propto \exp\left\{ -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})^{'} \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \right\} \cdot \prod_{l=1}^{L+1} \left(D_{kl}^*\right)^{\alpha_l - 1},$$

where

$$D_{kl}^* \geq 0 \quad \text{and} \quad \sum_{l=1}^{L+1} D_{kl}^* = 1.$$

We set $\alpha_1 = \alpha_2 = \cdots = \alpha_{L+1} = 1$ to denote a non-informative baseline prior. Since the prior is not conjugate we use Algorithm 8 from Neal (2000) to sample $\boldsymbol{D}_k^* | \cdot$. Algorithm 8 is outlined in Appendix. Once we have the D_{kl}^* values, the joinpoints are

obtained using

$$\begin{aligned}
\tau_{k1} &= (D^*_{k1}) \times T^* \\
\tau_{k2} &= (D^*_{k2} + D^*_{k1}) \times T^* \\
&\vdots \\
\tau_{kL} &= \left(D^*_{kL} + D^*_{k,L-1} + \cdots + D^*_{k1}\right) \times T^*
\end{aligned}$$

2.9 Simulation study

Without loses of generality, we set $\boldsymbol{\alpha}$ to be zero to simplify our simulation study. We assume that the number of biomarkers $K = 2$. The mean trajectory function for one biomarker is then

$$\psi(t) = \delta_0 + \delta_1 t + \beta_1 (t - \tau_1)_+ + \beta_2 (t - \tau_2)_+.$$

We set $\boldsymbol{\theta} = (\delta_0, \delta_1, \beta_1, \beta_2)'$ and generate two groups of data using two $\boldsymbol{\theta}$ values. To make the joinpoints more prominent, we use the following $\boldsymbol{\theta}$ values:

$$\boldsymbol{\theta}_1 = (0.8, -20.6, 50.2, -34.6)' \quad \text{and} \quad \boldsymbol{\theta}_2 = (2.1, 10.6, -40.2, 32.6)'.$$

The joinpoints are set as $\boldsymbol{\tau} = (5, 15)'$ for both biomarkers.

Using $e_{ij} \sim (0, 1)$ we generated data for the two groups, with 50 subjects each. The generated data is shown in Figure 2.1. We ran 50,000 MCMC scans with 20,000 burn-in and thinning of 10.

To get the cluster graph, we calculated an $N \times N$ matrix whose (i, j)th element gives us proportion times that $\boldsymbol{\theta}_{i,1}$ and $\boldsymbol{\theta}_{j,1}$ are equal (i.e. ith and jth individuals cluster

together.) The elements of this matrix are plotted in Figure 2.2. Higher propensity to cluster together is indicated by deeper red color. We have a faint cluster of $\boldsymbol{\theta}_{i,1}$ in Figure 2.2. Hence the true cluster structure is not recovered. Nevertheless, based on the true cluster structure, we decided to examine trace plots of the parameters corresponding to first and fifty first individual. They are in Figures 2.3 and 2.4. Figures 2.5 and 2.6 are the trace plots of joinpoints for the first and the second biomarkers. We see that the estimated values are nowhere close to the true values.

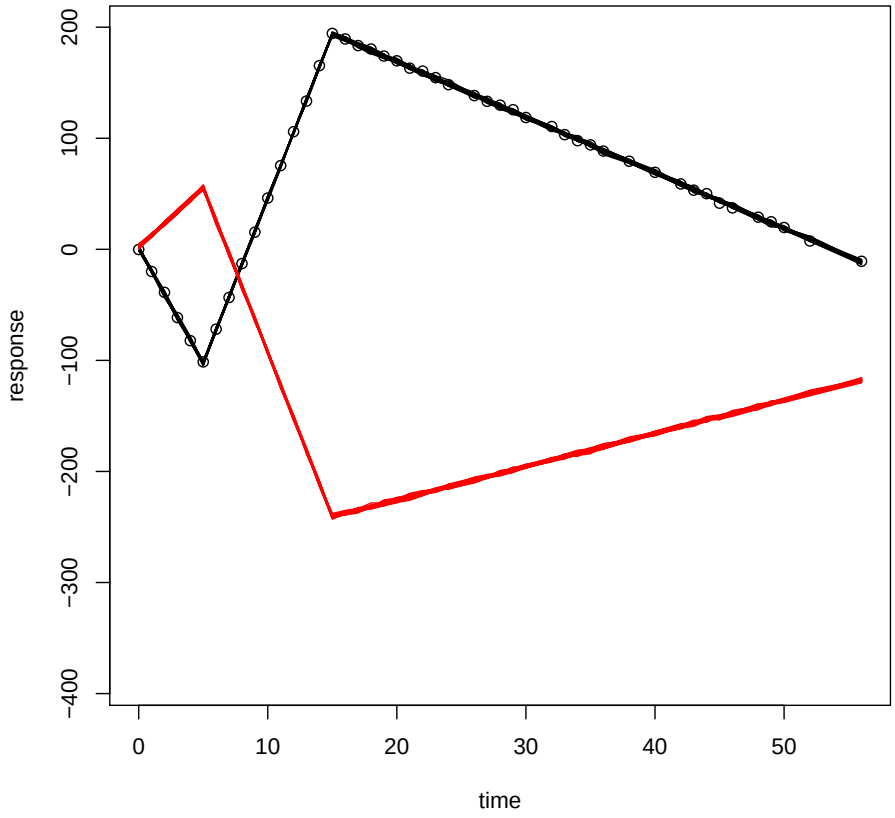

Figure 2.1. Longitudinal trajectories for the generated data.

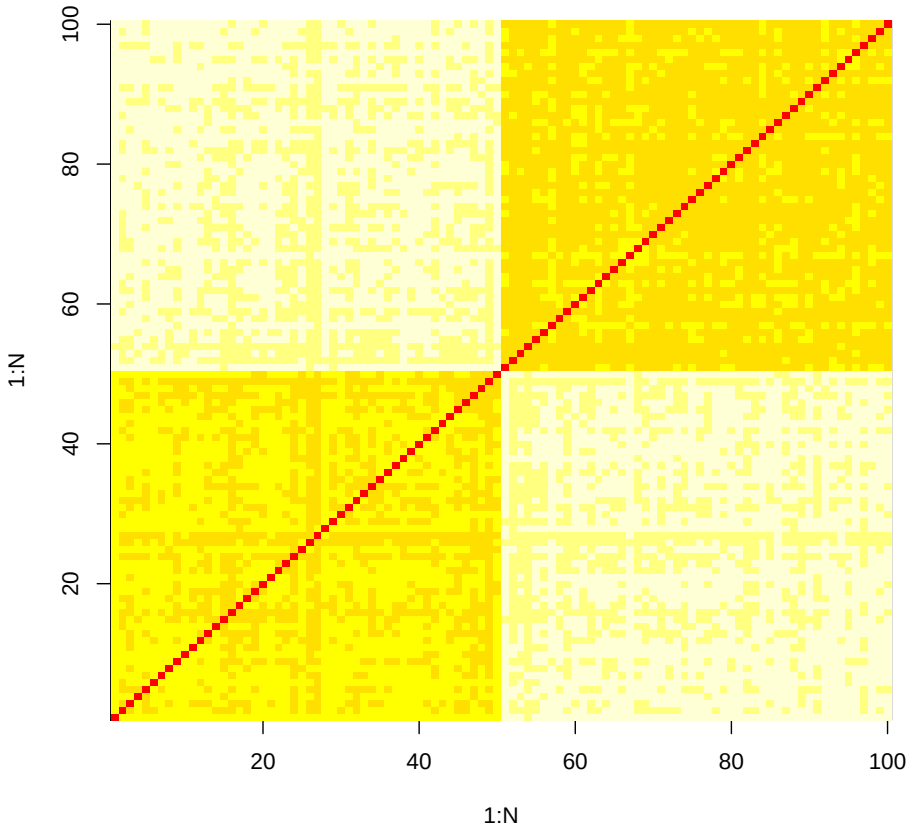

Figure 2.2. Plot of the heatmap of the clustering probabilities of individuals based on $\boldsymbol{\theta}_{i,1}$ values. Higher propensity to cluster together is indicated by deeper red color.

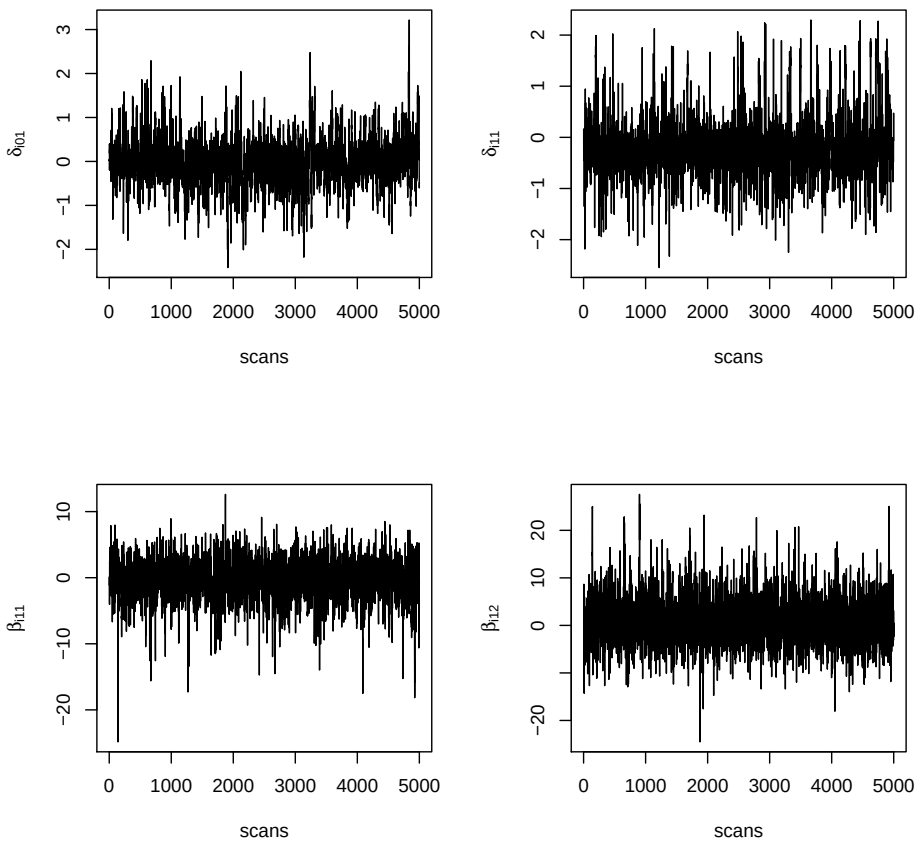

Figure 2.3. Trace plot of the parameters associated with the first individual.

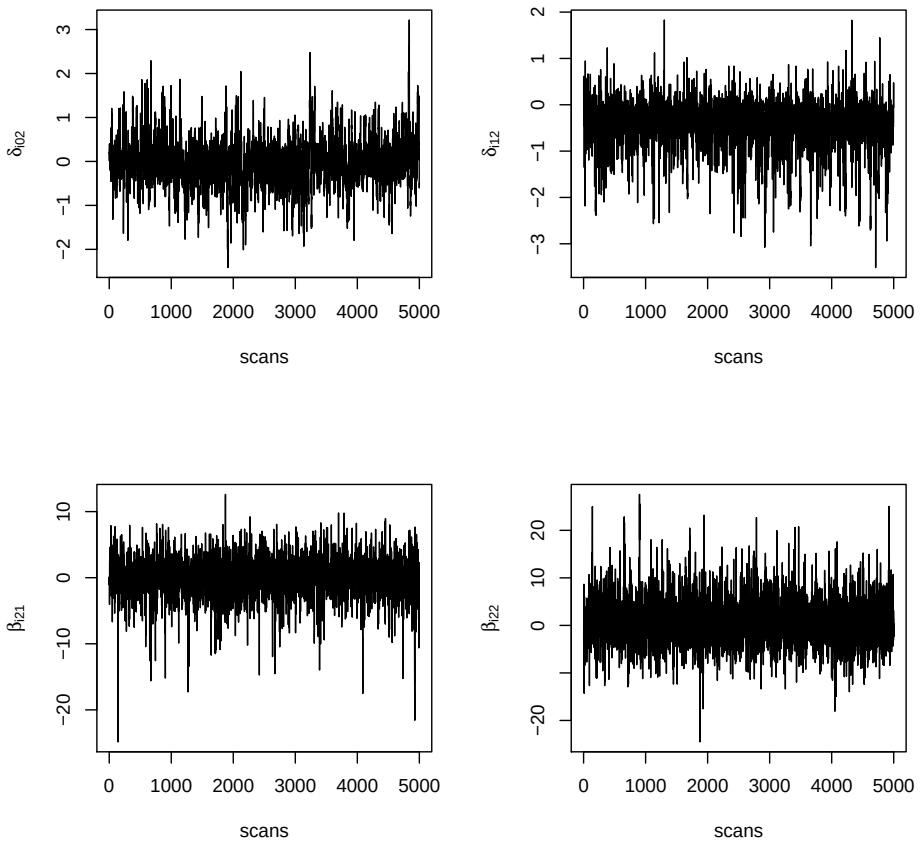

Figure 2.4. Trace plot of the parameters associated with the fifty first individual.

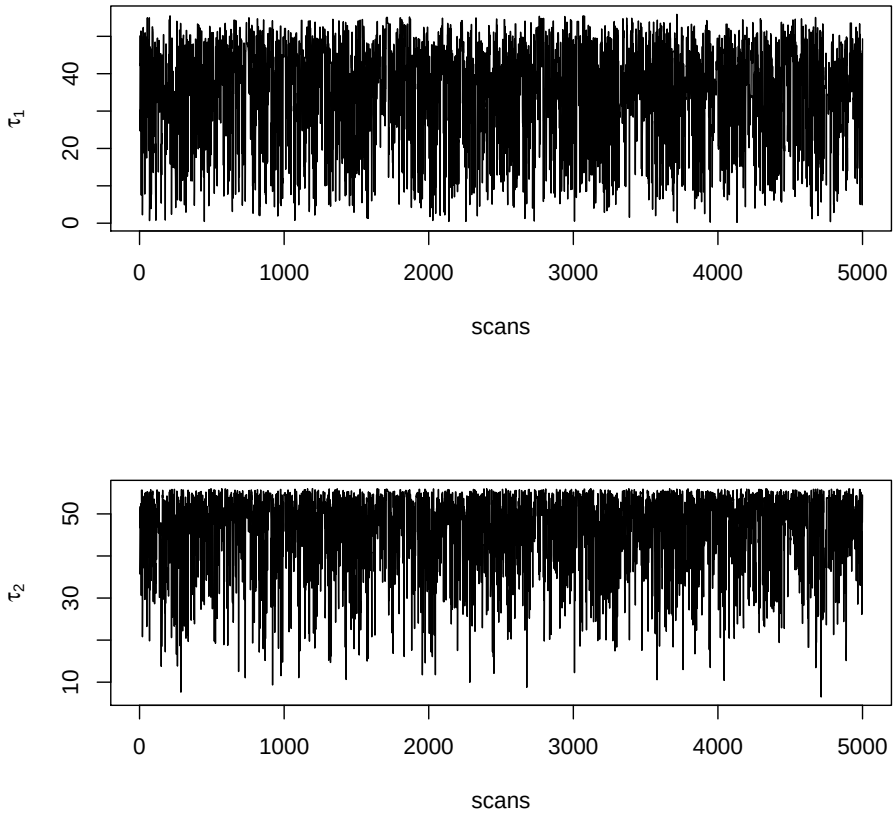

Figure 2.5. Trace plot of joinpoints for the first biomarker.

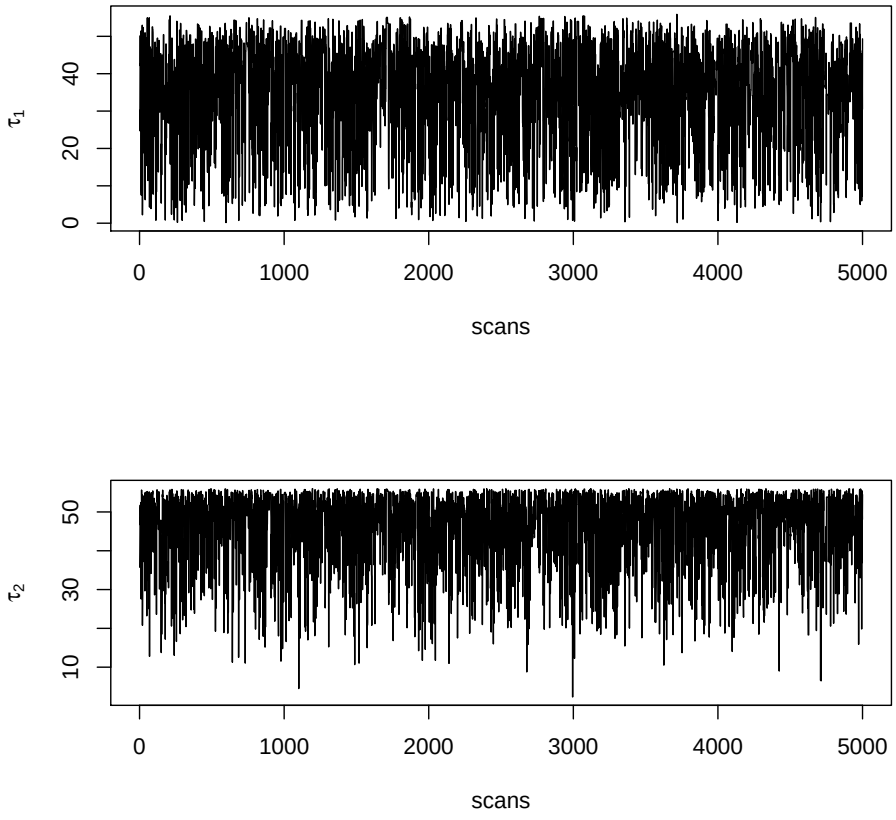

Figure 2.6. Trace plot of joinpoints for the second biomarker.

Table 2.1 gives the results of $\boldsymbol{\theta}$s and $\hat{\boldsymbol{\theta}}$s, where $\boldsymbol{\theta}$s are the parameters used to generate data and $\hat{\boldsymbol{\theta}}$s are the estimated values.

Table 2.1. Results of parameter estimation

$\boldsymbol{\theta}_1$	$\hat{\boldsymbol{\theta}}_1$	$\boldsymbol{\theta}_2$	$\hat{\boldsymbol{\theta}}_2$
0.8	-0.01	2.1	-0.01
-20.6	-0.27	10.6	-0.32
50.2	-0.13	-40.2	0.14
-34.6	0.62	32.6	0.53

Table 2.2 gives the results of $\boldsymbol{\tau}$s and $\hat{\boldsymbol{\tau}}$s, where $\boldsymbol{\tau}$s are the parameters used to generate data and $\hat{\boldsymbol{\tau}}$s are the estimated values.

Table 2.2. Results of joinpoint estimation

	τ_1	$\hat{\tau}_1$	τ_2	$\hat{\tau}_2$
First biomarker	5	38.9	15	49.4
Second biomarker	5	38.7	15	49.3

The proposed model was then failed to give the true cluster structure and the close estimates of the parameters for the generated data.

CHAPTER 3

JOINT MODEL FOR MULTIPLE BIOMARKERS AND MULTIPLE DROPOUT

In addition to longitudinal trajectories, information is available on dropout. Dropout occurs when a subject drop out of study because of some causes. For example, in ACTG 398 study, if toxicity level is too high the individuals drop out of study. To correctly account for the possibly informative censoring in the data due to dropout from the treatment or patients lost to follow up, in this chapter we jointly model the longitudinal trajectory and a dropout process. We construct the joint likelihood as the product of the time-to-event (dropout) likelihood conditional on the longitudinal marker, multipled by the likelihood of the longitudinal trajectory.

3.1 Longitudinal model with multiple joinpoints

We model the longitudinal part in a similar manner as in Chapter 2. The contribution of the ith subject to the likelihood is:

$$f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \propto \frac{1}{|\boldsymbol{\Sigma}|^{\frac{n_i}{2}}} \exp\left\{-\frac{1}{2}\sum_{j=1}^{n_i}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})\right\}$$

3.2 Modeling dropout

We assume there are M causes of dropout and the hazard of dropout due to cause m is a function of trajectory levels of the K biomarkers at time t using a Cox

proportional model of the form:

$$\lambda_m(t|\psi) = \lambda_{0m}(t)\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_k(t)\right\}, \quad m = 1, 2, ...M.$$

Here $\psi_k(t)$ is the trajectory for the kth biomarker of a generic subject, and for a particular subject i it will be just $\psi_{ik}(t)$. Note that $\gamma_{mk} = 0$ implies that the kth biomarker has no effect on the dropout due to cause m. When $\gamma_{mk} > 0$ implies that higher dropout rates will be associated with higher trajectory levels of the kth biomarker and vice-versa.

For each m, the baseline hazard function $\lambda_{0m}(\cdot)$ is assumed to be piecewise constant, given by

$$\lambda_{0m}(u) = \lambda_{0mj}, \quad u_{j-1} \leq u < u_j \qquad j = 1, \ldots, J,$$

where the number of steps J and the endpoints u_j are pre-specified with $u_0 \equiv 0$, and $u_J \equiv \infty$.

Let $\boldsymbol{\delta}_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{iM})$ denote the vector of indicators of the M dropout causes for subject i, where

$$\delta_{im} = \begin{cases} 1 & \text{if dropout is due to cause } m \text{ ,} \\ 0 & \text{otherwise.} \end{cases}$$

For a subject i experiencing dropout, let π_{im} be the probability that dropout is due to cause m. We assume that this probability is a function of the covariates in the following manner:

$$\pi_{im} = \pi_{im}(\boldsymbol{x}_i) = \frac{\exp\{\alpha_{0m} + \boldsymbol{\beta}_{0m}^{'}\boldsymbol{x}_i\}}{1 + \sum_{m=1}^{M-1}\exp\{\alpha_{0m} + \boldsymbol{\beta'}_{0m}\boldsymbol{x}_i\}}, \quad m = 1, \ldots, M-1,$$

where α_{0m} and $\boldsymbol{\beta}_{0m}$ are unknown coefficients. The contribution to the likelihood from the survival part of the ith subject is then

$$\prod_{m=1}^{M} \left[\pi_{im}\lambda_m(s_i)S(s_i)\right]^{\delta_{im}} S(s_i)^{1-\sum_{m=1}^{M}\delta_{im}} = \prod_{m=1}^{M} \left[\pi_{im}\lambda_m(s_i)\right]^{\delta_{im}} S(s_i).$$

An individual is exposed to M mutually exclusive causes of failure and when a dropout occurs, we observe the time T of dropout and the cause L of dropout. The cause specific hazard function for dropout due to cause m is given by

$$\lambda_m(t) = \lim_{\triangle t \to 0} \frac{Pr\{t \le T \le t + \triangle t, L = m | T \ge t\}}{\triangle t},$$

The overall hazard function for dropout at time t is then

$$\tilde{\lambda}(t) = \sum_{m=1}^{M} \left[\pi_{im}\lambda_m(t)\right].$$

Then the corresponding cumulative hazard is

$$H(t) = \int_0^t \tilde{\lambda}(u)\mathrm{d}u = \int_0^t \sum_{m=1}^{M} \left[\pi_{im}\lambda_m(u)\right] \mathrm{d}u = \sum_{m=1}^{M} \int_0^t \left[\pi_{im}\lambda_m(u)\right] \mathrm{d}u$$

The survival function is then

$$\begin{aligned} S(t) &= \exp\left\{-H(t)\right\} \\ &= \exp\left\{-\sum_{m=1}^{M} \int_0^t \left[\pi_{im}\lambda_m(u)\right] \mathrm{d}u\right\} \end{aligned}$$

Hence,

$$S(s_i) = \exp\left\{-\sum_{m=1}^{M} \int_0^{s_i} \pi_{im}\lambda_{0m}(u)\exp\left[\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u)\right] \mathrm{d}u\right\}$$

The contribution to the likelihood from the survival part of ith subject is then

$$\begin{aligned} f(s_i, \boldsymbol{\delta}_i | \boldsymbol{\psi}_i) = &\prod_{m=1}^{M} \left\{\pi_{im}\lambda_{0m}(s_i)\exp\left[\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(s_i)\right]\right\}^{\delta_{im}} \\ &\times \exp\left\{-\sum_{m=1}^{M} \int_0^{s_i} \pi_{im}\lambda_{0m}(u)\exp\left[\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u)\right] \mathrm{d}u\right\}, \end{aligned}$$

which can be written as

$$
\begin{aligned}
f(s_i, \boldsymbol{\delta}_i|\boldsymbol{\psi}_i) &= \prod_{m=1}^{M} [\pi_{im}\lambda_{0m}(s_i)]^{\delta_{im}} \exp\left\{ \sum_{m=1}^{M} \left[\delta_{im} \sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(s_i) \right] \right\} \\
&\quad \times \exp\left\{ -\sum_{m=1}^{M} \int_0^{s_i} \pi_{im}\lambda_{0m}(u) \exp\left[\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u) \right] \mathrm{d}u \right\} \\
&= \prod_{m=1}^{M} [\pi_{im}\lambda_{0m}(s_i)]^{\delta_{im}} \exp\Bigg\{ \sum_{m=1}^{M} \left[\delta_{im} \sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(s_i) \right] \\
&\quad - \sum_{m=1}^{M} \int_0^{s_i} \pi_{im}\lambda_{0m}(u) \exp\left[\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u) \right] \mathrm{d}u \Bigg\} \\
&= \prod_{m=1}^{M} [\pi_{im}\lambda_{0m}(s_i)]^{\delta_{im}} \exp\Bigg\{ \sum_{m=1}^{M} \left[\delta_{im} \sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(s_i) \right] \\
&\quad - \sum_{m=1}^{M} \sum_{j=1}^{J} \pi_{im}\lambda_{0mj} H_{ijm} \Bigg\},
\end{aligned}
$$

where

$$
H_{ijm} = I(s_i \geq u_{j-1}) \int_{u_{j-1}}^{u_j \wedge s_i} \exp\left\{ \sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u) \right\} \mathrm{d}u.
$$

The full likelihood is thus proportional to:

$$
\begin{aligned}
&\prod_{m=1}^{M} \prod_{i=1}^{N} [\pi_{im}]^{\delta_{im}} \prod_{m=1}^{M} \prod_{j=1}^{J} \lambda_{0mj}^{\sum_{i=1}^{N} \delta_{im} I(u_{j-1} \leq s_i \leq u_j)} \frac{1}{|\boldsymbol{\Sigma}|^{\sum_{i=1}^{N} \frac{n_i}{2}}} \\
&\times \exp\Bigg\{ -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \\
&\qquad + \sum_{i=1}^{N} \sum_{m=1}^{M} \delta_{im} \sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(s_i) - \sum_{i=1}^{N} \sum_{m=1}^{M} \sum_{j=1}^{J} \pi_{im}\lambda_{0mj} H_{ijm} \Bigg\}
\end{aligned}
$$

3.3 Prior distributions

The parameters $\boldsymbol{\theta}_i, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}, \boldsymbol{\alpha}_k, D_k^*$ are assigned to the same priors as in Section 2.2.

We assume the following priors for the other parameters:

$$\lambda_{0mj} \overset{iid}{\sim} G(a_{mj}, b_{mj}), \quad m = 1, \ldots, M, \quad j = 1, \ldots, J,$$

$$\gamma_{mk} \overset{iid}{\sim} N(\mu_\gamma, \sigma^2_\gamma), \quad m = 1, \ldots, M, \quad k = 1, \ldots, K,$$

$$\alpha_{0m} \sim N(0, \sigma^2_{\alpha m}), \quad m = 1, \ldots, M,$$

$$\boldsymbol{\beta}_{0m} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\beta m}), \quad m = 1, \ldots, M.$$

where $a_{mj}, b_{mj}, \mu_\gamma, \sigma^2_\gamma, \sigma^2_{\alpha m}, \boldsymbol{\Sigma}_{\beta m}$ are known values.

3.4 The full conditional of $\boldsymbol{\theta}^*_{1,p}|\cdot$

As before, let $\boldsymbol{\theta}^*_{1,p} = (\delta^*_{01p}, \delta^*_{11p}, \beta^*_{11p}, \beta^*_{12p}, \ldots, \beta^*_{1Lp})'$ be the pth distinct value of the random effect vector for the first biomarker ($p = 1, \ldots, k_\theta$.) With the dropout added, the baseline prior is not conjugate to the sampling distribution. The configuration structure $\boldsymbol{c}^\theta = (c^\theta_1, c^\theta_2, \ldots, c^\theta_N)$ of the random effects is updated using Algorithm 8 in Neal (2000).

We get

$$\begin{aligned}
f(\boldsymbol{\theta}^*_{1,p}|\cdot) &\propto L(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta}^*_{1,p}|\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}^*_{1,p} - \boldsymbol{\mu}_\theta)' \boldsymbol{\Sigma}^{-1}_\theta (\boldsymbol{\theta}^*_{1,p} - \boldsymbol{\mu}_\theta)\right\} \\
&\times \exp\Bigg\{ -\frac{1}{2} \sum_{i:c^\theta_i = p} \sum_{j=1}^{n_i} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \\
&\qquad + \sum_{i:c^\theta_i = p} \sum_{m=1}^{M} \delta_{im} \sum_{k=1}^{K} \gamma_{mk} \psi_{ik}(s_i) - \sum_{i:c^\theta_i = p} \sum_{m=1}^{M} \sum_{j=1}^{J} \pi_{im} \lambda_{0mj} H_{ijm} \Bigg\},
\end{aligned}$$

which can be updated by using Metropolis-Hastings algorithm.

3.5 The full conditional of $\boldsymbol{\theta}_{i,-1}|\cdot$

For the remaining biomarkers (except the first one) of the ith subject, let

$$\begin{aligned}\boldsymbol{\theta}_{i,-1} &= (\delta_{i02}, \delta_{i12}, \beta_{i21}, \beta_{i22}, \ldots, \beta_{i2L}, \ldots, \delta_{i0K}, \delta_{i1K}, \beta_{iK1}, \beta_{iK2}, \ldots, \beta_{iKL})' \\ &= (\boldsymbol{\theta}_{i,2}, \ldots, \boldsymbol{\theta}_{i,K})'.\end{aligned}$$

As before, we assume

$$\boldsymbol{\theta}_{i,2}, \ldots, \boldsymbol{\theta}_{i,K} \overset{iid}{\sim} \boldsymbol{N}_{(L+2)}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta).$$

Let $\boldsymbol{\theta}_{i,k}$ be the random effect vector for the kth biomarker of the ith subject. Then the posterior conditional of $\boldsymbol{\theta}_{i,k}$ is given by

$$\begin{aligned}f(\boldsymbol{\theta}_{i,k}|\cdot) &\propto L(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) \\ &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}_{i,k} - \boldsymbol{\mu}_\theta)'\boldsymbol{\Sigma}_\theta^{-1}(\boldsymbol{\theta}_{i,k} - \boldsymbol{\mu}_\theta)\right\} \\ &\times \exp\Bigg\{-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \\ &\qquad + \sum_{i=1}^{N}\sum_{m=1}^{M}\delta_{im}\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(s_i) - \sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{j=1}^{J}\pi_{im}\lambda_{0mj}H_{ijm}\Bigg\},\end{aligned}$$

which can be updated by using Metropolis-Hastings algorithm.

3.6 The full conditionals of $\boldsymbol{\mu}_\theta|\cdot$, $\boldsymbol{\Sigma}_\theta|\cdot$ and $\boldsymbol{\Sigma}|\cdot$

The full conditionals of $\boldsymbol{\mu}_\theta|\cdot$, $\boldsymbol{\Sigma}_\theta|\cdot$ and $\boldsymbol{\Sigma}|\cdot$ are the same as in Chapter 2.

3.7 The full conditional of $\boldsymbol{\alpha}_k|\cdot$

The full conditional of $\boldsymbol{\alpha}_k|\cdot$ is obtained as

$$\begin{aligned} f(\boldsymbol{\alpha}_k|\cdot) &\propto \prod_{i=1}^{N} f(\boldsymbol{Y}_i|\boldsymbol{\phi}_i, \boldsymbol{\Sigma}) \cdot \pi(\boldsymbol{\alpha}_k|\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha) \\ &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\alpha}_k - \boldsymbol{\mu}_\alpha)' \boldsymbol{\Sigma}_\alpha^{-1}(\boldsymbol{\alpha}_k - \boldsymbol{\mu}_\alpha)\right\} \\ &\quad \cdot \exp\Bigg\{ -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij})' \boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_{ij} - \boldsymbol{\phi}_{ij}) \\ &\qquad + \sum_{i=1}^{N}\sum_{m=1}^{M} \delta_{im}\gamma_{mk}\psi_{ik}(s_i) - \sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{j=1}^{J} \pi_{im}\lambda_{0mj}H_{ijm} \Bigg\} \end{aligned}$$

3.8 The full conditional of $\lambda_{0mj}|\cdot$

The full conditional of $\lambda_{0mj}|\cdot$ is obtained using

$$\begin{aligned} \pi(\lambda_{0mj}|\cdot) &\propto {\lambda_{0mj}}^{\sum_{i=1}^{N} \delta_{im} I(u_{j-1} \le s_i \le u_j)} \exp\left\{ -\sum_{i=1}^{N} \pi_{im}\lambda_{0mj}H_{ijm} \right\} \\ &\quad \times {\lambda_{0mj}}^{a_{mj}-1} \exp\left\{-b_{mj}\lambda_{0mj}\right\}, \quad \lambda_{0mj} > 0, \end{aligned}$$

which results in

$$\lambda_{0mj}|\cdot \sim G\left(a_{mj} + \sum_{A_j} \delta_{im}, \quad b_{mj} + \sum_{i=1}^{N} \pi_{im}H_{ijm} \right),$$

where

$$A_j = \{i : s_i \in [u_{j-1}, u_j]\}.$$

3.9 The full conditional of $\boldsymbol{\gamma}_{mk}|\cdot$

The full conditional of $\boldsymbol{\gamma}_{mk}|\cdot$ is given by

$$\log f(\boldsymbol{\gamma}|\cdot) = const. - \sum_{m=1}^{M}\sum_{k=1}^{K}\frac{\gamma_{mk}^2 - 2\mu_\gamma\gamma_{mk}}{2\sigma_\gamma^2} + \sum_{m=1}^{M}\sum_{k=1}^{K}\sum_{i=1}^{N}\delta_{im}\gamma_{mk}\psi_{ik}(s_i) - \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\pi_{im}\lambda_{0mj}H_{ijm},$$

where

$$\boldsymbol{\gamma} = \begin{pmatrix} \gamma_{11}\dots\gamma_{1K} \\ \vdots \quad \vdots \\ \gamma_{M1}\dots\gamma_{MK} \end{pmatrix}.$$

3.10 The full conditional of $\boldsymbol{\alpha}_0|\cdot$

Define

$$\boldsymbol{\alpha}_0 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_M \end{pmatrix},$$

The full conditional of $\boldsymbol{\alpha}_0|\cdot$ is

$$\log f(\boldsymbol{\alpha_0}|\cdot) = const. + \log\left[\prod_{m=1}^{M}\prod_{i=1}^{N}(\pi_{im})^{\delta_{im}}\right] - \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\pi_{im}\lambda_{0mj}H_{ijm} - \frac{1}{2\sigma_{\alpha m}^2}\sum_{m=1}^{M}\alpha_{0m}^2.$$

3.11 The full conditional of $\boldsymbol{\beta}_0|\cdot$

Define

$$\boldsymbol{\beta}_{0m} = \begin{pmatrix} \beta_{m1} \\ \vdots \\ \beta_{mC} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\beta_0} = \begin{pmatrix} \beta_{11}\dots\beta_{1C} \\ \vdots \quad \vdots \\ \beta_{M1}\dots\beta_{MC} \end{pmatrix},$$

The full conditional of $\boldsymbol{\beta}_0|\cdot$ is

$$\log f(\boldsymbol{\beta_0}|\cdot) = const. + \log\left[\prod_{m=1}^{M}\prod_{i=1}^{N}(\pi_{im})^{\delta_{im}}\right] - \sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\pi_{im}\lambda_{0mj}H_{ijm} - \frac{1}{2}\sum_{m=1}^{M}\boldsymbol{\beta}_{0m}'\boldsymbol{\Sigma}_{\beta m}^{-1}\boldsymbol{\beta}_{0m}.$$

3.12 Estimating the changepoints

Similarly to Section 2.8, with the dropout part added, the posterior conditional of $\boldsymbol{D}_k^*|\cdot$ will be

$$\pi(\boldsymbol{D}_k^*|\cdot) = \exp\Bigg\{ -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}(\boldsymbol{Y}_{ij}-\boldsymbol{\phi}_{ij})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_{ij}-\boldsymbol{\phi}_{ij}) + \sum_{i=1}^{N}\sum_{m=1}^{M}\delta_{im}\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(s_i) - \sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{j=1}^{J}\pi_{im}\lambda_{0mj}H_{ijm}\Bigg\} \cdot \prod_{l=1}^{L+1}(D^*)_{kl}^{\alpha_l-1}, \quad D_{kl}^* \geq 0 \quad \text{and} \quad \sum_{l=1}^{L+1}D_{kl}^* = 1,$$

which can be updated by using Metropolis-Hastings algorithm.

3.13 Calculation of H_{ijm}

Below, we present calculation of

$$H_{ijm} = I(s_i \geq u_{j-1})\int_{u_{j-1}}^{u_j \wedge s_i}\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}\mathrm{d}u.$$

Let $\tau_{(1)} < \tau_{(2)} < \ldots < \tau_{(KL)}$ be the ordered values of τ_{kl}, $k = 1, \ldots, K$ and $l = 1, \ldots, L$ (note that $\tau_{(j)} = \tau_{kl}$ for some $k, l.$) If $\tau_{(j)} = \tau_{kl}$, denote $\beta_{ij}^* = \beta_{ikl}$ and $\gamma_{mj}^* = \gamma_{mk}$.

Hence

$$\begin{aligned}
\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u) &= \sum_{k=1}^{K}\gamma_{mk}\left\{\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i+\delta_{i0k}+\delta_{i1k}u+\sum_{l=1}^{L}\beta_{ikl}(u-\tau_{kl})_+\right\}\\
&= \sum_{k=1}^{K}\gamma_{mk}\left\{\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i+\delta_{i0k}\right\}+\sum_{k=1}^{K}\gamma_{mk}\delta_{i1k}u+\sum_{k=1}^{K}\sum_{l=1}^{L}\gamma_{mk}\beta_{ikl}(u-\tau_{kl})_+\\
&= \begin{cases}
a_{i0}+a_{i1}u \quad \text{if} \quad u\le\tau_{(1)}\\
a_{i0}+a_{i1}u+\gamma^*_{m1}\beta^*_{i1}(u-\tau_{(1)}) \quad \text{if} \quad \tau_{(1)}<u\le\tau_{(2)}\\
a_{i0}+a_{i1}u+\gamma^*_{m1}\beta^*_{i1}(u-\tau_{(1)})+\gamma^*_{m2}\beta^*_{i2}(u-\tau_{(2)})\\
\qquad\qquad\qquad\qquad \text{if} \quad \tau_{(2)}<u\le\tau_{(3)}\\
\quad\vdots\\
a_{i0}+a_{i1}u+\gamma^*_{m1}\beta^*_{i1}(u-\tau_{(1)})+\gamma^*_{m2}\beta^*_{i2}(u-\tau_{(2)})\\
\qquad\qquad +\cdots\gamma^*_{mK}\beta^*_{i,KL}(u-\tau_{(KL)}) \quad \text{if} \quad u>\tau_{(KL)}
\end{cases}\\
&= \sum_{l=0}^{KL}(a_{il}+\rho_{il}u)\cdot 1(u)_{\{\tau_{(l)},\tau_{(l+1)}\}},
\end{aligned}$$

where

$$\begin{aligned}
a_{i0} &= \sum_{k=1}^{K}\gamma_{mk}\left\{\boldsymbol{\alpha}_k^{'}\boldsymbol{x}_i+\delta_{i0k}\right\},\\
a_{il} &= a_{i,l-1}-\gamma^*_{mj}\beta^*_{il}\tau_{(l)}, \quad l=1,2,\ldots,KL, \quad \tau_{(j)}=\tau_{kl},\\
\rho_{i0} &= a_{i1}=\sum_{k=1}^{K}\gamma_{mk}\delta_{i1k},\\
\rho_{il} &= \rho_{i,l-1}+\gamma^*_{mj}\beta^*_{il}, \quad l=1,2,\ldots,KL, \quad \tau_{(j)}=\tau_{kl},\\
\tau_{(0)} &= 0 \quad \text{and} \quad \tau_{(KL+1)}=T^*.
\end{aligned}$$

Then

$$\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}=\sum_{l=0}^{KL}\exp\Big\{(a_{il}+\rho_{il}u)\Big\}\cdot 1(u)_{(\tau_{(l)},\tau_{(l+1)})}.$$

Hence

$$
\begin{aligned}
&\int_0^x \exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}\mathrm{d}u \\
&=\sum_{l=0}^{KL}\int_{\tau_{(l)}\wedge x}^{\tau_{(l+1)}\wedge x}\exp\Big\{(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u \\
&=\begin{cases} x \quad \text{if} \quad \gamma_{mk}=0 \quad \text{for all } m,\, k \\
\displaystyle\sum_{l=0}^{KL}\exp\{a_{il}\}\left[\Big(\big(\exp\{\rho_{il}(\tau_{(l+1)}\wedge x)\}-\exp\{\rho_{il}(\tau_{(l)}\wedge x)\}\big)/\rho_{il}\Big)\right. \\
\left.\cdot 1_{\{\rho_{il}\neq 0\}}+\Big(\tau_{(l+1)}\wedge x-\tau_{(l)}\wedge x\Big)\cdot 1_{(\rho_{il}=0)}\right] \quad \text{if} \quad \gamma_{mk}\neq 0, \quad \text{for some } m,\, k.
\end{cases}
\end{aligned}
$$

Then

$$
\begin{aligned}
&\int_{x_1}^{x_2}\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}\mathrm{d}u \\
&=\int_0^{x_2}\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}\mathrm{d}u-\int_0^{x_1}\exp\left\{\sum_{k=1}^{K}\gamma_{mk}\psi_{ik}(u)\right\}\mathrm{d}u \\
&=\sum_{l=0}^{KL}\int_{\tau_{(l)}\wedge x_2}^{\tau_{(l+1)}\wedge x_2}\exp\Big\{(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u-\sum_{l=0}^{KL}\int_{\tau_{(l)}\wedge x_1}^{\tau_{(l+1)}\wedge x_1}\exp\Big\{(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u \\
&=\begin{cases} x_2-x_1 \quad \text{if} \quad \gamma_{mk}=0 \quad \text{for all } m,\, k \\
\displaystyle\sum_{l=0}^{KL}\exp\{a_{il}\}\left[\left(\Big(\big(\exp\{\rho_{il}(\tau_{(l+1)}\wedge x_2)\}-\exp\{\rho_{il}(\tau_{(l)}\wedge x_2)\}\big)\right.\right. \\
\left.-\big(\exp\{\rho_{il}(\tau_{(l+1)}\wedge x_1)\}-\exp\{\rho_{il}(\tau_{(l)}\wedge x_1)\}\big)\Big)/\rho_{il}\right)\cdot 1_{(\rho_{il}\neq 0)} \\
\left.+\Big(\big(\tau_{(l+1)}\wedge x_2-\tau_{(l)}\wedge x_2\big)-\big(\tau_{(l+1)}\wedge x_1-\tau_{(l)}\wedge x_1\big)\Big)\right. \\
\left.\qquad\cdot 1_{(\rho_{il}=0)}\right] \quad \text{if} \quad \gamma_{mk}\neq 0, \quad \text{for some } m,\, k.
\end{cases}
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
H_{ijm} &= I(s_i \geq u_{j-1}) \int_{u_{j-1}}^{u_j \wedge s_i} \exp\left\{\sum_{k=1}^{K} \gamma_{mk}\psi_{ik}(u)\right\} \mathrm{d}u \\
&= \begin{cases}
I(s_i \geq u_{j-1})(u_j \wedge s_i - u_{j-1}) \quad \text{if} \quad \gamma_{mk} = 0 \quad \text{for all } m,\, k \\
\sum_{l=0}^{KL} \exp\{a_{il}\} \Bigg[\Bigg(\Big(\big(\exp\{\rho_{il}(\tau_{(l+1)} \wedge u_j \wedge s_i)\} \\
-\exp\{\rho_{il}(\tau_{(l)} \wedge u_j \wedge s_i)\}\big) - \big(\exp\{\rho_{il}(\tau_{(l+1)} \wedge u_{j-1})\} \\
-\exp\{\rho_{il}(\tau_{(l)} \wedge u_{j-1})\}\big)\Big)/\rho_{il}\Bigg) \cdot 1_{(\rho_{il}\neq 0)} \\
+\Big(\big(\tau_{(l+1)} \wedge u_j \wedge s_i - \tau_{(l)} \wedge u_j \wedge s_i\big) \\
-\big(\tau_{(l+1)} \wedge u_{j-1} - \tau_{(l)} \wedge u_{j-1}\big)\Big) \cdot 1_{(\rho_{il}=0)}\Bigg] \quad \text{if} \quad \gamma_{mk} \neq 0, \quad \text{for some } m,\, k.
\end{cases}
\end{aligned}
$$

3.14 Simulation study

The generation of the data for the longitudinal part is the same as that in Chapter 2. We generate the dropout data according to the structure of data of plasma HIV RNA from the AIDS Clinical Trials Group (ACTG) 398 study. ACTG 398 was a randomized double-blinded placebo-controlled study comparing a single protease inhibitor (PI) with double-PI antiretroviral regimens in treating HIV-infected patients. The primary objective of the study was to compare the proportion of subjects who had virologic failure after 24 weeks on study between the double-PI arms and the single-PI arm. Four hundred and eighty-one subjects were followed for varying durations (some, as long as 72 weeks) and their viral loads were tracked.Detailed description of the data and results of primary analysis are available in Hammer et al. (2002).

In ACTG 398, individuals are assumed to drop out of study due to two competing causes-namely toxicity and other causes. We denote the time to drop out due to toxicity by TTOX and the time to drop out due to other causes by TOTHER. A third variable TOT is defined as TOT=min(TTOX,TOTHER,100)-it is the overall time to drop out of study. Here TOT=100 indicates that the individual never drop out.

Since the survival function takes values in the interval $(0, 1)$, we first generate a value from uniform distribution and set it equal to the survival function, which is as the following:

$$Y \sim U(0, 1),$$

$$Y = e^{-\int_0^t H(s)\mathrm{d}s}.$$

where $H(s)$ is the cumulative hazard function in Section 3.2. For a fixed set of parameters, we can easily solve for t values and we use them as the dropout times in our generated data. We first assign generated t values randomly to these time for these 2 causes. We generate the same portion of 100's as in ACTG 398 data and randomly generate the same number of positions and then assign the 100's to these positions in our generated data. We also generate the off-study time (TOS) randomly according to the proportion of these times in the ACTG 398 data. Then we can easily get the individual time to event, si, according to si=min(TOT,TOS). We assign either cause as cause 1 and the other as cause 2. Then the vector of indicators of the 2 dropout causes is easy to obtained according to $\delta_{im} = I(si = si_m^d)$, where si_m^d is the dropout time.

We used two groups of generated data with 10 subjects each. We ran 50,000 scans

of MCMC with 20,000 burn-in and thinning of 100. We have a faint cluster of $\boldsymbol{\theta}_{i,1}$ in Figure 3.1. Hence the true cluster structure is not recovered. Nevertheless, based on the true cluster structure, we decided to examine trace plots of the parameters corresponding to 1st and 11th individual. They are in Figures 3.2 and 3.3. We see that the estimated values are nowhere close to the true values.

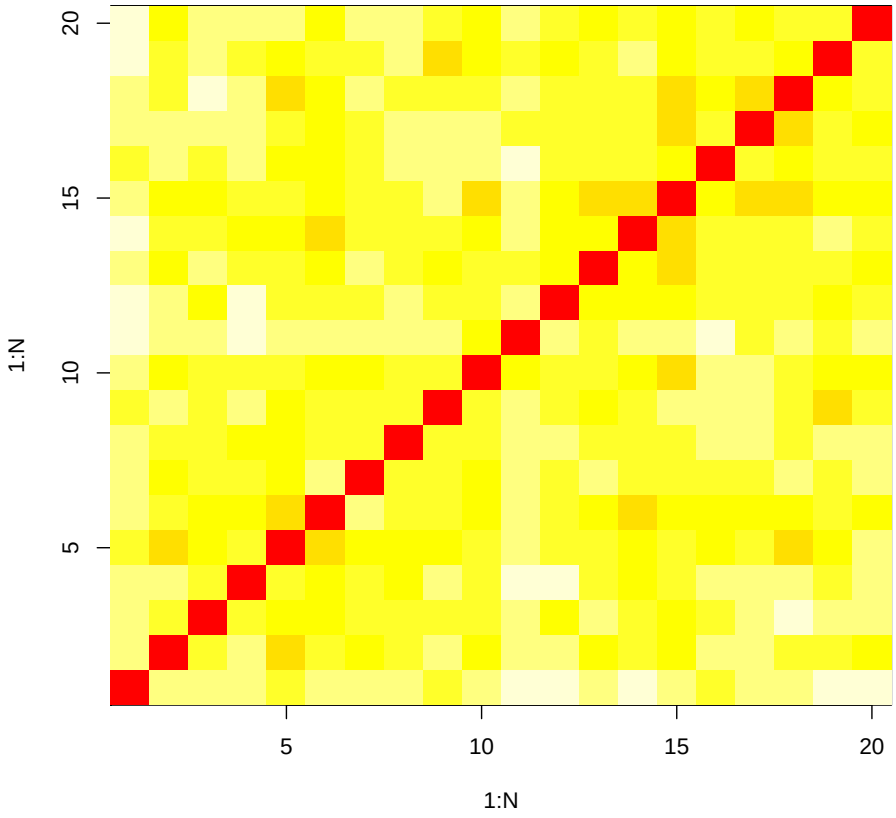

Figure 3.1. Plot of the heatmap of the clustering probabilities of individuals based on $\boldsymbol{\theta}_{i,1}$ values. Higher propensity to cluster together is indicated by deeper color.

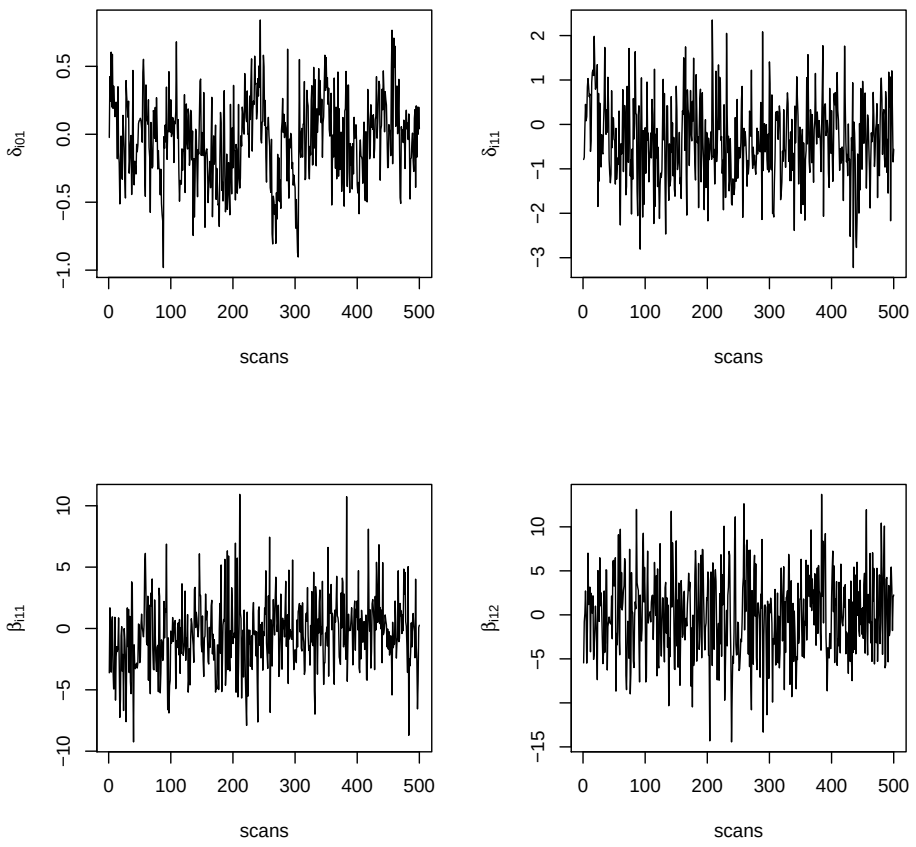

Figure 3.2. Trace plot of the parameters associated with the first individual.

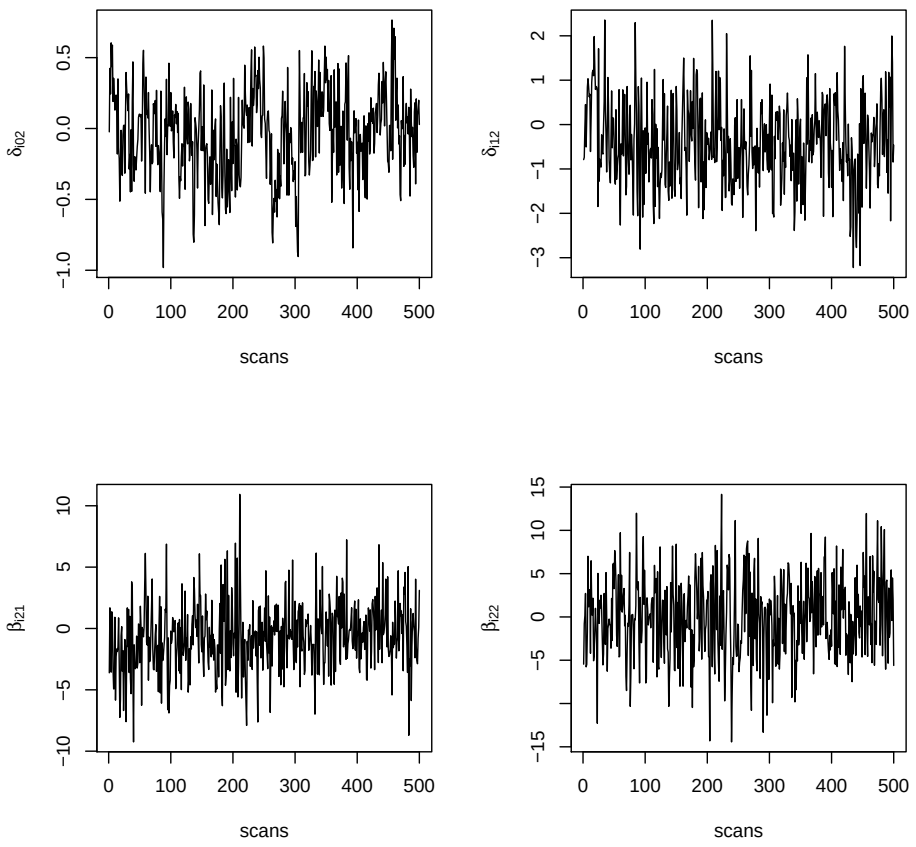

Figure 3.3. Trace plot of the parameters associated with the eleventh individual.

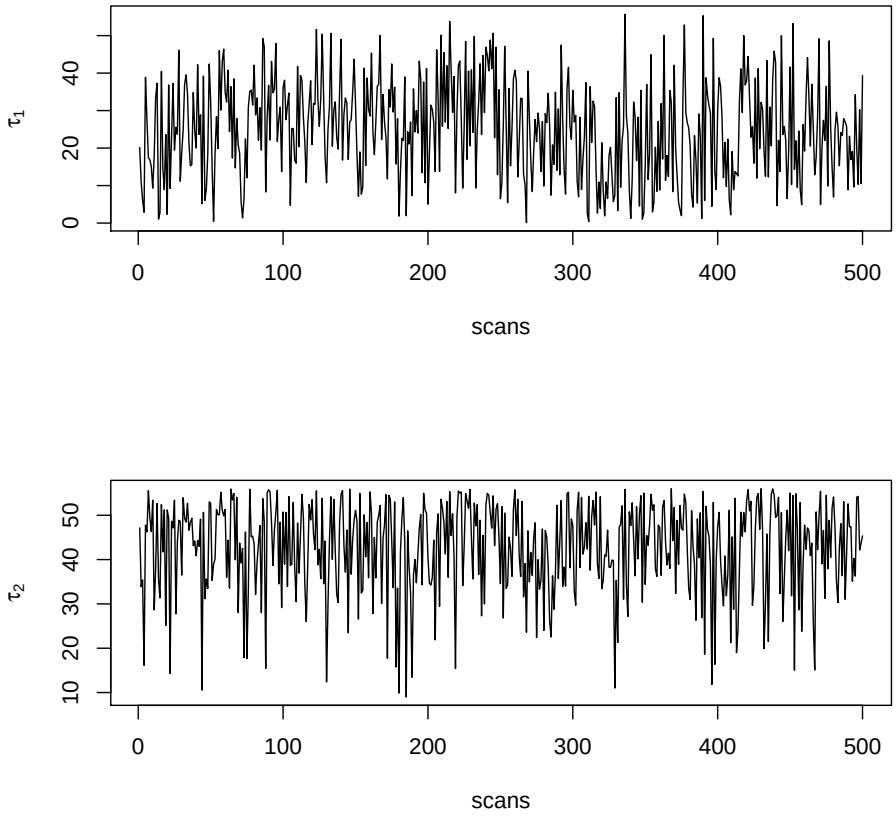

Figure 3.4. Trace plot of joinpoints for the first biomarker.

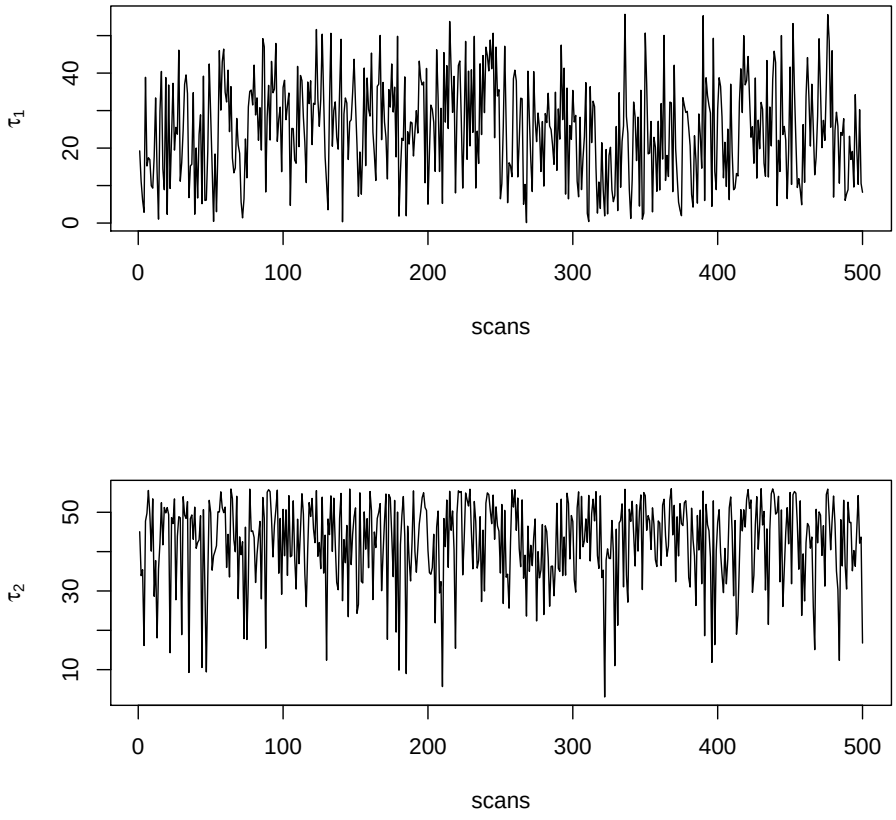

Figure 3.5. Trace plot of joinpoints for the second biomarker.

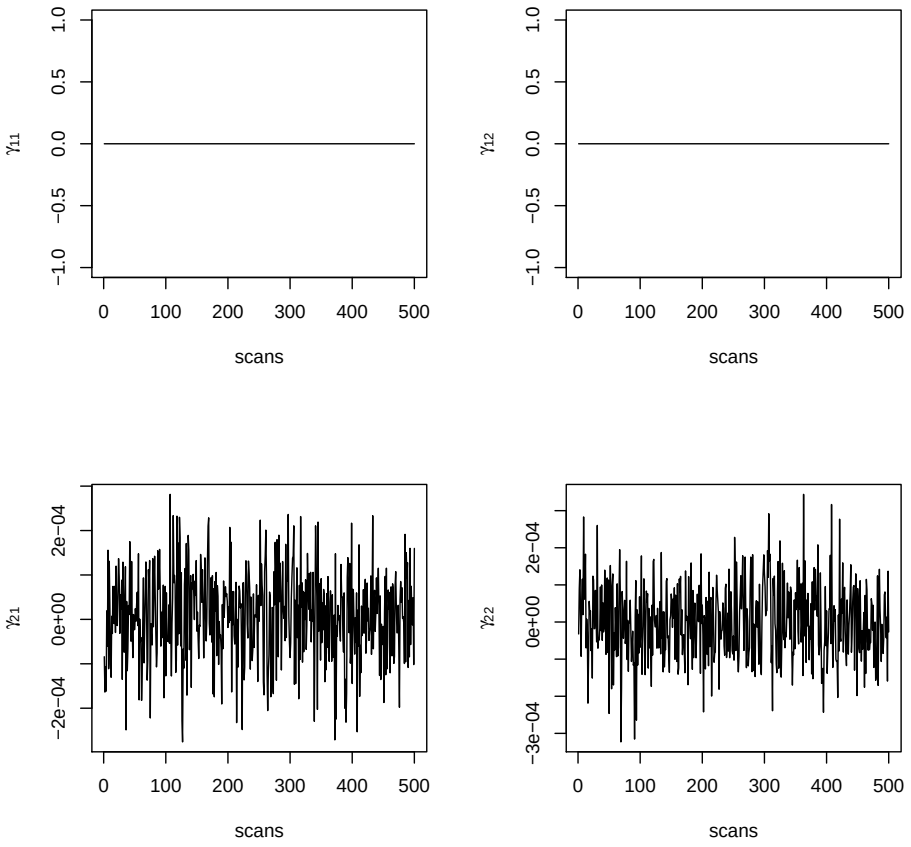

Figure 3.6. Trace plot of γ_{mk}.

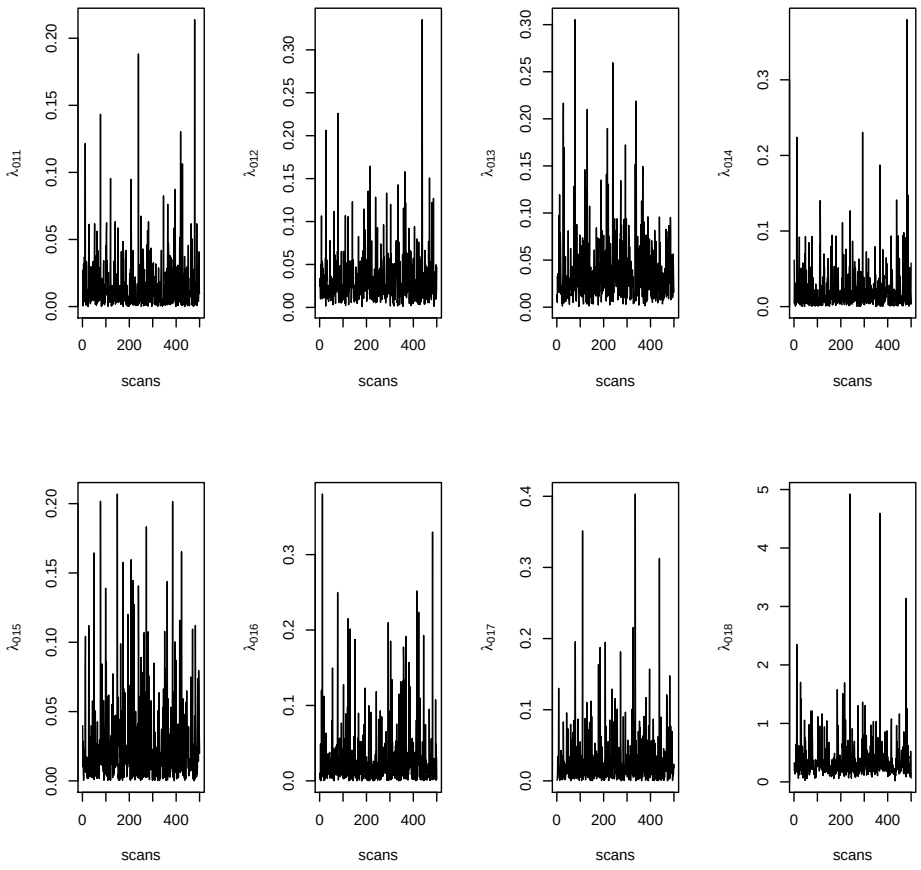

Figure 3.7. Trace plot of λ_{0mj} for cause 1 ($m = 1$).

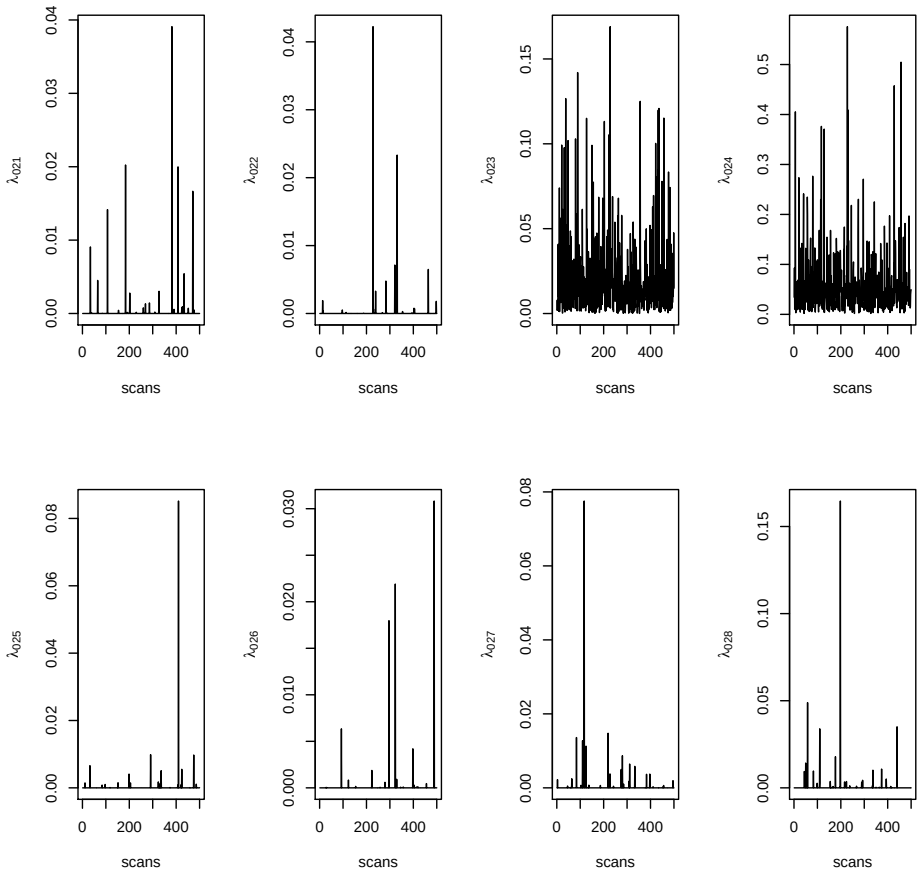

Figure 3.8. Trace plot of λ_{0mj} for cause 2 ($m = 2$).

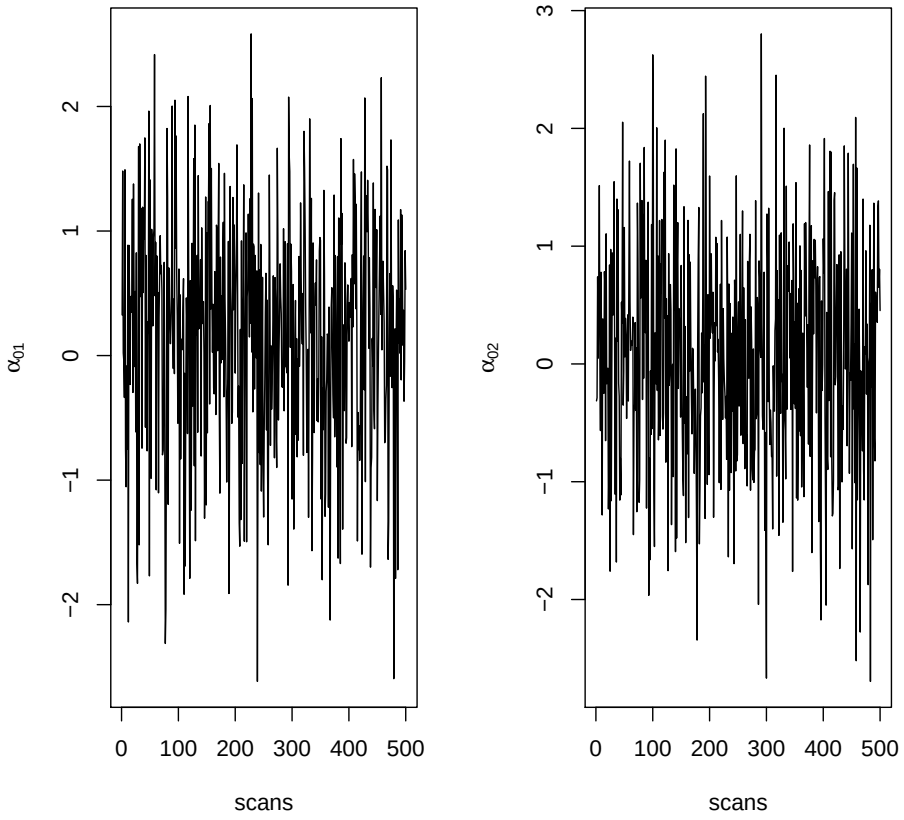

Figure 3.9. Trace plot of α_{0m}.

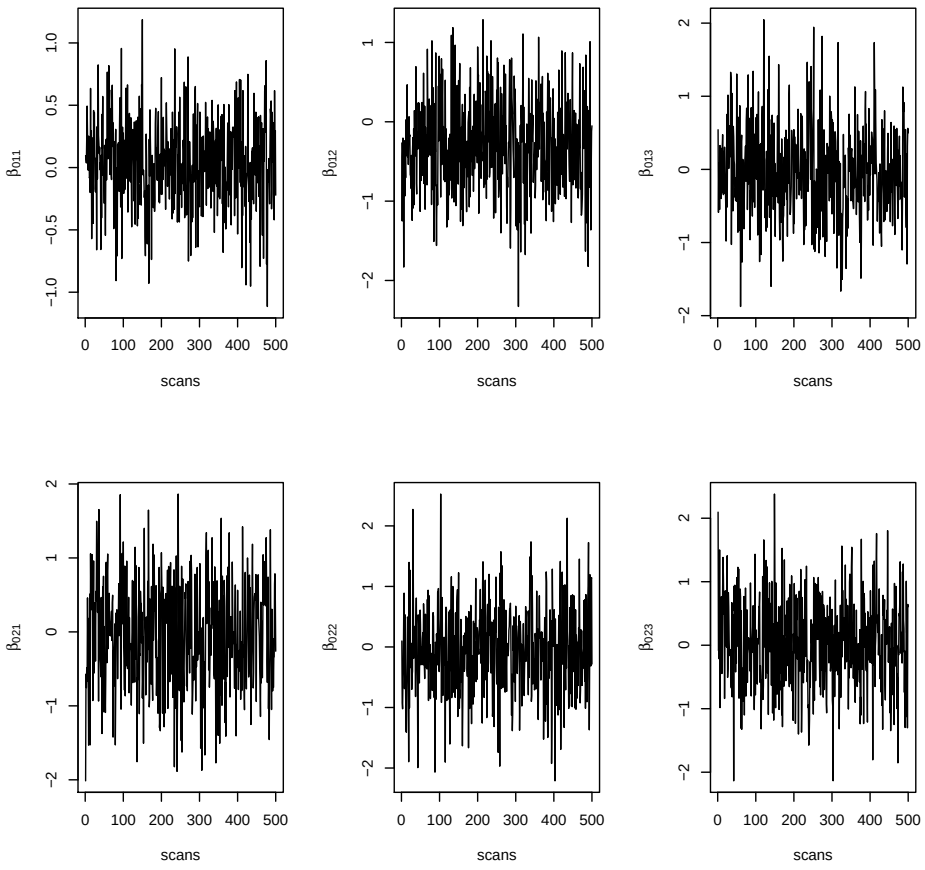

Figure 3.10. Trace plot of $\boldsymbol{\beta}_{0m}$.

Table 3.1 gives the results of $\boldsymbol{\theta}$s and $\hat{\boldsymbol{\theta}}$s, where $\boldsymbol{\theta}$s are the parameters used to generate the data and $\hat{\boldsymbol{\theta}}$s are the corresponding estimated values.

Table 3.1. Results of parameter estimation

$\boldsymbol{\theta}_1$	$\hat{\boldsymbol{\theta}}_1$	$\boldsymbol{\theta}_2$	$\hat{\boldsymbol{\theta}}_2$
0.8	-0.02	2.1	-0.01
-20.6	-0.40	10.6	-0.43
50.2	-0.46	-40.2	-0.71
-34.6	0.18	32.6	0.17

Table 3.2 gives the results of estimation of $\boldsymbol{\tau}$s, with the corresponding trace plots are presented in Figures 3.4 and 3.5.

Table 3.2. Results of joinpint estimation

	τ_1	$\hat{\tau}_1$	τ_2	$\hat{\tau}_2$
First biomarker	5	25	15	42.7
Second biomarker	5	24.9	15	42.4

Since γ_{mk} is non-identifiable, we tried to estimate the identifiable differences of $\gamma_{21} - \gamma_{11}$ and $\gamma_{22} - \gamma_{12}$. We set γ_{11} and γ_{12} zero thus the positive values of γ_{21} and γ_{22} showed in Figure 3.6 implies that the first biomarker has higher effect in the dropout probability due to cause 2 than cause 1. Figures 3.7 showed that the piecewise constant baseline hazard for the dropout, λ_{0mj}, is high in the last interval, which means individuals tend to drop out in this interval with high probability. Figures 3.9

and 3.10 showed that the quantities α_{0m} and $\boldsymbol{\beta}_{0m}$ are close to zero, which implies that the probability of dropout is the same for all causes.

CHAPTER 4

AN ALTERNATIVE PARAMETRIZATION OF THE JOINPOINTS

Since there were difficulties in estimating the joinpoints by the model used in Chapter 2 and Chapter 3, we will try a new model with a convenient parametrization of the joinpoints. This idea is inspired by Martinez-Beneito et al. (2011). For simplicity, we use only one biomarker.

4.1 The longitudinal model of the joinpoints

As before, let y_{ij} denote the response of the ith subject measured at time t_{ij}, $(i = 1, \ldots, N; j = 1, \ldots, n_i)$. Our longitudinal model assumes the following form:

$$y_{ij} = \varphi_i(t_{ij}) + e_{ij},$$

where $\varphi_i(\cdot)$ is the trajectory function of the ith individual, e_{ij} is the random error associated with the jth measurement on the ith individual with

$$e_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

We assume the following form of the trajectory function for the ith individual:

$$\varphi_i(t) = \boldsymbol{\alpha}^{'} \boldsymbol{x}_i + \delta_{i0} + \beta_{i0}(t - \bar{t}_i) + \sum_{l=1}^{L} \delta_{il}\beta_{il} B_{\tau_{il}}(t),$$

where L is the maximum number of joinpoints and $B_{\tau_{il}}(t)$ is defined as the following piecewise linear function:

$$B_{\tau_{il}}(t) = \begin{cases} a_{i0l} + b_{i0l}t, & t \leq \tau_{il}, \\ a_{i1l} + b_{i1l}t, & t > \tau_{il}. \end{cases}$$

The functions $B_{\tau_{il}}(t)$ are restricted to a number of conditions given below:

1. $B_{\tau_{il}}(t)$ is continuous at the joinpoints, which is expressed as $\lim_{t\to\tau_{il}^-} B_{\tau_{il}}(t) = \lim_{t\to\tau_{il}^+} B_{\tau_{il}}(t), \quad l = 1, \ldots, L$. This guarantee that $\varphi_i(t)$ is continuous all around, independently of the number of joinpoints.

2. At all the observed points, the sum of the elements of the joinpoint evaluated must be zero. This is expressed as $\sum_{l=1}^{n_i} B_{\tau_{il}}(t_{ij}) = 0, \quad l = 1, \ldots, L$. This way, the addition of any joinpoints in the model would not alter the mean value of the regression function and it would not change the meaning and the estimation of parameter $\boldsymbol{\alpha}$ across the model.

3. To guarantee the addition of any joinpoints would not change the slope of the regression function and the meaning and the estimate of parameters β_0 would not be changed either, the slope of the break-points along the whole period of study should be zero. This is expressed as $\sum_{l=1}^{n_i} B_{\tau_{il}}(t_{ij}) \cdot t_{ij} = 0, \quad l = 1, \ldots, L$.

4. To make the parameter β_j take the role of measuring the magnitude of the break-point in the location where the change takes place, we need to have $B_{\tau_{il}}(\tau_{il}) = 1, \quad l = 1, \ldots, L$. This makes the value of β_j identifiable.

The above conditions can be translated to the following 8 equations (for the case of

$L = 2$):

$$\begin{cases} a_{i01} + b_{i01} \cdot \tau_{i1} = a_{i11} + b_{i11} \cdot \tau_{i1} \\ a_{i02} + b_{i02} \cdot \tau_{i2} = a_{i12} + b_{i12} \cdot \tau_{i2} \\ n_{i1} a_{i01} + b_{i01} \sum\limits_{t_{ij} < \tau_{i1}} t_{ij} + a_{i11}(n_i - n_{i1}) + b_{i11} \sum\limits_{t_{ij} > \tau_{i1}} t_{ij} = 0 \\ n_{i2} a_{i02} + b_{i02} \sum\limits_{t_{ij} < \tau_{i2}} t_{ij} + a_{i12}(n_i - n_{i2}) + b_{i12} \sum\limits_{t_{ij} > \tau_{i2}} t_{ij} = 0 \\ a_{i01} \sum\limits_{t_{ij} < \tau_{i1}} t_{ij} + b_{i01} \sum\limits_{t_{ij} < \tau_{i1}} t_{ij}^2 + a_{i11} \sum\limits_{t_{ij} > \tau_{i1}} t_{ij} + b_{i11} \sum\limits_{t_{ij} > \tau_{i1}} t_{ij}^2 = 0 \\ a_{i02} \sum\limits_{t_{ij} < \tau_{i2}} t_{ij} + b_{i02} \sum\limits_{t_{ij} < \tau_{i2}} t_{ij}^2 + a_{i12} \sum\limits_{t_{ij} > \tau_{i2}} t_{ij} + b_{i11} \sum\limits_{t_{ij} > \tau_{i2}} t_{ij}^2 = 0 \\ a_{i01} + b_{i01} \cdot \tau_{i1} = 1 \\ a_{i02} + b_{i02} \cdot \tau_{i2} = 1 \end{cases}$$

where n_i is the number of time points for the ith subject. n_{i1} is the number of time points smaller or equal to τ_{i1} and n_{i2} is the number of time points smaller or equal to τ_{i2}. Hence, for given value of τ_{i1} and τ_{i2} there is a unique set of values $a_{i01}, a_{i02}, a_{i11}, a_{i12}, b_{i01}, b_{i02}, b_{i11}, b_{i12}$ that can be solved from the above equations. The contribution of the ith subject to the likelihood is then:

$$f(\boldsymbol{Y}_i | \varphi_i, \sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^{n_i}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_i(t_{ij}) \Big)^2 \right\}$$

4.2 Prior distributions

For the ith subject, let $\boldsymbol{\theta}_i = (\delta_{i1}, \ldots, \delta_{iL}, \delta_{i0}, \beta_{i0}, \beta_{i1}, \ldots, \beta_{iL}, \tau_{i1}, \ldots, \tau_{iL})'$ denote the associated parameters. We model the prior distribution of $\boldsymbol{\theta}_i$ as follows:

$$\boldsymbol{\theta}_i \overset{iid}{\sim} G,$$

where

$$G \sim DP(M_\theta, G_0).$$

This prior choice of $\boldsymbol{\theta}_i$ make it robust to missepecification, as well as allows for clustering of subject trajectories.

The baseline prior G_0 for $\boldsymbol{\delta}_i, \boldsymbol{\eta}_i, \boldsymbol{\beta}_i$ is specified below:

Define

$$\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{iL})^{'}$$

$$\boldsymbol{\eta}_i = (\delta_{i0}, \beta_{i0})^{'}$$

$$\text{and} \qquad \boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{iL})^{'}$$

We assume

$$\pi(\boldsymbol{\delta}_i) = (L)^{-L}(L-1)^{L-\sum_{l=1}^{L}\delta_{il}}, \quad \text{where} \quad \delta_{il} = 0, 1, \quad l = 1, \ldots, L$$

In particular, for $L = 2$, this becomes

$$\pi(\delta_{i1} = u, \delta_{i2} = v) = \frac{1}{4}, \quad \text{for} \quad u = 0, 1, \quad v = 0, 1.$$

We also assume

$$\boldsymbol{\eta}_i \sim \boldsymbol{N}_2(\boldsymbol{\mu}_{\eta 0}, \boldsymbol{\Sigma}_{\eta 0}),$$

$$\text{and} \qquad \boldsymbol{\beta}_i \sim \boldsymbol{N}_L(\mathbf{0}, \gamma\boldsymbol{\Sigma}),$$

where

$$\gamma \sim \boldsymbol{IG}(0.5, 0.5).$$

We assume $\sigma^2 \sim IG(a_\sigma, b_\sigma)$.

In our problem, the block (corresponding to $\boldsymbol{\beta}_i$) of the Fisher information matrix evaluated at $\boldsymbol{\beta}_i = \mathbf{0}$ is $L = \boldsymbol{\Delta}\boldsymbol{B}^{'}\boldsymbol{B}\boldsymbol{\Delta}$, where $\boldsymbol{B} = \{B_{\tau_{il}}(t_{ij})$ (the matrix of covariates),

and $\boldsymbol{\Delta} = \text{diag}(\boldsymbol{\delta}_i)$. Since L is not a positive-definite matrix for every $\boldsymbol{\delta}_i$, it can not be used directly to define $\boldsymbol{\Sigma}$ above. Instead, we propose

$$\boldsymbol{\Sigma} = n_i(\boldsymbol{\Delta}\boldsymbol{B}^{'}\boldsymbol{B}\boldsymbol{\Delta} + \text{diag}(\boldsymbol{B}^{'}\boldsymbol{B} - \boldsymbol{\Delta}\boldsymbol{B}^{'}\boldsymbol{B}\boldsymbol{\Delta}))^{-1}$$

which is a positive-definite matrix for every $\boldsymbol{\delta}_i$ and is fixed for each subject.

To avoid identifiability problems, we impose a number of restrictions on the locations of the joinpoints. In previous chapters we found that when τ values are very close, it created calculation problems. Moreover in real life we don't expect joinpoints to occur very closely. Hence we assume a minimum separation d between the corresponding joinpoints. In particular, we assume that the parameter space for $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_L)$, which we call Ω, is

$$\Omega = \{(\tau_1, \tau_2, \ldots, \tau_L) : t_1 + d < \tau_1, \tau_1 + d < \tau_2, \tau_2 + d < \tau_3, \ldots, \tau_L + d < T^*\}$$

We put minimum separation $d = 2$ in our applications. We specify $t_1 = 0, T^* = 56$ as in the previous chapters.

We assume a uniform prior for $\tau_1 < \tau_2 < \ldots < \tau_L$ on the set Ω. Note that when $L = 2$, the marginal distribution of τ_1 becomes

$$f_{\tau_1}(\tau_1) = \begin{cases} \frac{1}{1250}(52 - \tau_1), & 2 < \tau_1 < 52 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\tau_2 | \tau_1 \sim \text{Uniform}(\tau_1 + 2, 54).$$

Consequently, the cumulative distribution of τ_1 becomes

$$F_{\tau_1}(s) = \int_2^s \frac{1}{1250}(52 - \tau_1) = \frac{1}{1250}(-\frac{1}{2}s^2 + 52s - 102).$$

Hence the steps to sample τ_1 and τ_2 from the prior will be:

Step1: Generate $u \sim \text{Uniform}(0, 1)$,

Step2: Generate $\tau_1 = 52 - 50\sqrt{1-u}$,

Step3: Generate $\tau_2|\tau_1 \sim \text{Uniform}(\tau_1 + 2, 54)$.

4.3 Posterior Calculations

The likelihood of $\boldsymbol{\theta}_i$ is

$$L(\boldsymbol{\theta}_i) = \frac{1}{(\sqrt{2\pi\sigma^2})^{n_i}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_{ij}(t_{ij}) \Big)^2 \right\}$$

The baseline posterior of $\boldsymbol{\delta}_i|\cdot$ is then:

$$\begin{aligned} \pi^*(\delta_{i1} = u, \delta_{i2} = v|\cdot) \propto L(\boldsymbol{\theta}_i|\delta_{i1} = u, \delta_{i2} = v, \delta_{i0}, \beta_{i0}, \beta_{i1}, \beta_{i2}, \tau_{i1}, \tau_{i2}) \\ \times \pi(\delta_{i1} = u, \delta_{i2} = v) \quad \text{for} \quad u, v \in \{0, 1\}. \end{aligned}$$

Let $\boldsymbol{Q} = (1, t_{ij} - \bar{t}_i)^{'}$ and $\boldsymbol{B} = (\delta_{i1}\beta_{i1}B_{\tau_{i1}}(t_{ij}), \delta_{i2}\beta_{i2}B_{\tau_{i2}}(t_{ij}))^{'}$, We first write the likelihood in terms of $\boldsymbol{\eta}_i = (\delta_{i0}, \beta_{i0})^{'}$ as:

$$
\begin{aligned}
L(\boldsymbol{\theta}_i) &\propto \exp\left\{ -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij} - \varphi_{ij}(t_{ij})\Big)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) - \sum_{l=1}^{2}\delta_{il}\beta_{il}B_{\tau_{il}}(t_{ij})\Big)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) \right. \\
&\qquad\qquad \left. - \delta_{i1}\beta_{i1}B_{\tau_{i1}}(t_{ij}) - \delta_{i2}\beta_{i2}B_{\tau_{i2}}(t_{ij})\Big)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \mathbf{1}^{'}\boldsymbol{B} - \boldsymbol{\eta}_i^{'}\boldsymbol{Q}\Big)^2 \right\} \\
&\propto \exp\left\{ -\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(\boldsymbol{\eta}_i^{'}\boldsymbol{Q}\boldsymbol{Q}^{'}\boldsymbol{\eta}_i - 2(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \mathbf{1}^{'}\boldsymbol{B})\boldsymbol{\eta}_i^{'}\boldsymbol{Q}\Big) \right\} \\
&= \exp\left\{ -\frac{1}{2}\boldsymbol{\eta}_i^{'}\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\boldsymbol{Q}\boldsymbol{Q}^{'}\boldsymbol{\eta}_i + \boldsymbol{\eta}_i^{'}\frac{1}{\sigma^2}\sum_{j=1}^{n_i}(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \mathbf{1}^{'}\boldsymbol{B})\boldsymbol{Q} \right\}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
f(\boldsymbol{\eta}_i|\cdot) &\propto L(\boldsymbol{\theta}_i)\cdot\exp\left\{ -\frac{1}{2}(\boldsymbol{\eta}_i^{'} - \boldsymbol{\mu}_{\eta 0})^{'}(\boldsymbol{\Sigma}_{\eta 0})^{-1}(\boldsymbol{\eta}_i^{'} - \boldsymbol{\mu}_{\eta 0}) \right\} \\
&\propto L(\boldsymbol{\theta}_i)\cdot\exp\left\{ -\frac{1}{2}\boldsymbol{\eta}_i^{'}\boldsymbol{\Sigma}_{\eta 0}^{-1}\boldsymbol{\eta}_i + \boldsymbol{\eta}_i^{'}\boldsymbol{\Sigma}_{\eta 0}^{-1}\boldsymbol{\mu}_{\eta 0} \right\} \\
&\propto \exp\left\{ -\frac{1}{2}\boldsymbol{\eta}_i^{'}\Big(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\boldsymbol{Q}\boldsymbol{Q}^{'} + \boldsymbol{\Sigma}_{\eta 0}^{-1}\Big)\boldsymbol{\eta}_i \right. \\
&\qquad \left. + \boldsymbol{\eta}_i^{'}\Big(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}(y_{ij} - \boldsymbol{\alpha}^{'}\boldsymbol{x}_i - \mathbf{1}^{'}\boldsymbol{B})\boldsymbol{Q} + \boldsymbol{\Sigma}_{\eta 0}^{-1}\boldsymbol{\mu}_{\eta 0}\Big) \right\}.
\end{aligned}
$$

Hence we have

$$
\boldsymbol{\eta}_i|\cdot \sim \boldsymbol{N}(\boldsymbol{\mu}_{\eta n}, \boldsymbol{\Sigma}_{\eta n}),
$$

where

$$\boldsymbol{\Sigma}_{\eta n} = \left(\frac{1}{\sigma^2} \sum_{j=1}^{n_i} \boldsymbol{Q}\boldsymbol{Q}^{'} + \boldsymbol{\Sigma}_{\eta 0}^{-1} \right)^{-1},$$

and

$$\boldsymbol{\mu}_{\eta n} = \left(\frac{1}{\sigma^2} \sum_{j=1}^{n_i} \boldsymbol{Q}\boldsymbol{Q}^{'} + \boldsymbol{\Sigma}_{\eta 0}^{-1} \right)^{-1} \left(\frac{1}{\sigma^2} \sum_{j=1}^{n_i} (y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \mathbf{1}^{'} \boldsymbol{B}) \boldsymbol{Q} + \boldsymbol{\Sigma}_{\eta 0}^{-1} \boldsymbol{\mu}_{\eta 0} \right).$$

To get the posterior of $\boldsymbol{\beta}_i$, we set $\boldsymbol{B}^* = \left(\delta_{i1} B_{\tau_{i1}}(t_{ij}), \delta_{i2} B_{\tau_{i2}}(t_{ij}) \right)^{'}$ and we write the likelihood in terms of $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2})^{'}$ as:

$$\begin{aligned}
L(\boldsymbol{\theta}_i) &\propto \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \left(y_{ij} - \varphi_{ij}(t_{ij}) \right)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \left(y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) - \sum_{l=1}^{2} \delta_{il} \beta_{il} B_{\tau_{il}}(t_{ij}) \right)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \Big(y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) \right. \\
&\qquad \left. - \delta_{i1} \beta_{i1} B_{\tau_{i1}}(t_{ij}) - \delta_{i2} \beta_{i2} B_{\tau_{i2}}(t_{ij}) \Big)^2 \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \left(y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) - \boldsymbol{\beta}_i^{'} \boldsymbol{B}^* \right)^2 \right\} \\
&\propto \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \left(\boldsymbol{\beta_i}^{'} \boldsymbol{B}^* \boldsymbol{B}^{*'} \boldsymbol{\beta}_i - 2 \Big(y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) \Big) \boldsymbol{\beta}_i^{'} \boldsymbol{B}^* \right) \right\} \\
&= \exp\left\{ -\frac{1}{2} \boldsymbol{\beta}_i^{'} \frac{1}{\sigma^2} \sum_{j=1}^{n_i} \boldsymbol{B}^* \boldsymbol{B}^{*'} \boldsymbol{\beta}_i + \boldsymbol{\beta}_i^{'} \frac{1}{\sigma^2} \sum_{j=1}^{n_i} \Big(y_{ij} - \boldsymbol{\alpha}^{'} \boldsymbol{x}_i - \delta_{i0} - \beta_{i0}(t_{ij} - \bar{t}_i) \Big) \boldsymbol{B}^* \right\}.
\end{aligned}$$

Thus we have

$$
\begin{aligned}
f(\boldsymbol{\beta}_i|\cdot) &\propto L(\boldsymbol{\theta}_i)\cdot\pi(\boldsymbol{\beta}_i) \\
&\propto L(\boldsymbol{\theta}_i)\cdot\pi(\boldsymbol{\beta}_i|\mathbf{0},\gamma\boldsymbol{\Sigma})\cdot \\
&\propto L(\boldsymbol{\theta}_i)\cdot\frac{1}{|\gamma\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\Big\{-\frac{1}{2}\boldsymbol{\beta}_i^{'}(\gamma\boldsymbol{\Sigma})^{-1}\boldsymbol{\beta}_i\Big\} \\
&\propto \exp\Bigg\{-\frac{1}{2}\boldsymbol{\beta}_i^{'}\bigg(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\boldsymbol{B}^{*}\boldsymbol{B}^{*'}+(\gamma\boldsymbol{\Sigma})^{-1}\bigg)\boldsymbol{\beta}_i \\
&\qquad +\boldsymbol{\beta}_i^{'}\bigg(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}(y_{ij}-\boldsymbol{\alpha}_i^{'}\boldsymbol{x}_i-\delta_{i0}-\beta_{i0}(t_{ij}-\bar{t})\bigg)\boldsymbol{B}^{*}\Bigg\}.
\end{aligned}
$$

Hence we have

$$\boldsymbol{\beta}_i|\cdot\sim\boldsymbol{N}(\boldsymbol{\mu}_{\beta n},\boldsymbol{\Sigma}_{\beta n}),$$

where

$$\boldsymbol{\Sigma}_{\beta n}=\bigg(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\boldsymbol{B}^{*}\boldsymbol{B}^{*'}+(\gamma\boldsymbol{\Sigma})^{-1}\bigg)^{-1},$$

and

$$\boldsymbol{\mu}_{\beta n}=\bigg(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\boldsymbol{B}^{*}\boldsymbol{B}^{*'}+(\gamma\boldsymbol{\Sigma})^{-1}\bigg)^{-1}\bigg(\frac{1}{\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij}-\boldsymbol{\alpha}_i^{'}\boldsymbol{x}_i-\delta_{i0}-\beta_{i0}(t_{ij}-\bar{t})\Big)\boldsymbol{B}^{*}\bigg).$$

Furthermore

$$f(\gamma|\cdot)\propto\prod_{i=1}^{N}\frac{1}{|\gamma\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\Big\{-\frac{1}{2}\boldsymbol{\beta}_i^{'}(\gamma\boldsymbol{\Sigma})^{-1}\boldsymbol{\beta}_i\Big\}\cdot\gamma^{-0.5-1}\exp\Big\{-\frac{0.5}{\gamma}\Big\},$$

which implies

$$\gamma|\cdot\sim IG\left(0.5+\frac{N}{2},\quad 0.5+0.5\sum_{i=1}^{N}\boldsymbol{\beta}_i^{'}\boldsymbol{\Sigma}\boldsymbol{\beta}_i\right).$$

4.4 The full conditional of $\boldsymbol{\alpha}|\cdot$

The conditional of $\boldsymbol{\alpha}|\cdot$ is is obtained using

$$\begin{aligned}
f(\boldsymbol{\alpha}|\cdot) &\propto \prod_{i=1}^{N} f(\boldsymbol{Y}_i|\varphi_i)\cdot\pi(\boldsymbol{\alpha}|\boldsymbol{\mu}_\alpha,\boldsymbol{\Sigma}_\alpha)\\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\mu}_\alpha)'\boldsymbol{\Sigma}_\alpha^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu}_\alpha)\right\}\\
&\quad\times\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_i(t_{ij})\Big)^2\right\}\\
&\propto \exp\Bigg\{-\frac{1}{2}\boldsymbol{\alpha}'\Big(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma^2}\boldsymbol{x}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_i'+\boldsymbol{\Sigma}_\alpha^{-1}\Big)\boldsymbol{\alpha}\\
&\qquad+\boldsymbol{\alpha}'\Big(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma^2}\boldsymbol{x}_i\boldsymbol{\Sigma}^{-1}A^*+\boldsymbol{\Sigma}_\alpha^{-1}\boldsymbol{\mu}_\alpha\Big)\Bigg\},
\end{aligned}$$

where

$$A^* = y_{ij}-\delta_{i0}-\beta_{i0}(t-\bar{t}_i)-\sum_{l=1}^{L}\delta_{il}\beta_{il}B_{\tau_{il}}(t).$$

Hence

$$\boldsymbol{\alpha}|\cdot\sim\boldsymbol{N}(\boldsymbol{\mu}_{\alpha n},\boldsymbol{\Lambda}_{\alpha n}),$$

where

$$\begin{aligned}
\boldsymbol{\Lambda}_{\alpha n} &= \Big(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma^2}\boldsymbol{x}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_i'+\boldsymbol{\Sigma}_\alpha^{-1}\Big)^{-1},\\
\boldsymbol{\mu}_{\alpha n} &= \Big(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma^2}\boldsymbol{x}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_i'+\boldsymbol{\Sigma}_\alpha^{-1}\Big)^{-1}\Big(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{1}{\sigma^2}\boldsymbol{x}_i\boldsymbol{\Sigma}^{-1}A^*+\boldsymbol{\Sigma}_\alpha^{-1}\boldsymbol{\mu}_\alpha\Big).
\end{aligned}$$

4.5 The full conditional of $\sigma^2|\cdot$

The conditional of $\sigma^2|\cdot$ is

$$\begin{aligned} f(\sigma^2|\cdot) &\propto \prod_{i=1}^{N} f(\boldsymbol{Y}_i|\varphi_i)\cdot\pi(\sigma^2|a_\sigma,b_\sigma) \\ &\propto \prod_{i=1}^{N}\frac{1}{(\sigma^2)^{\frac{n_i}{2}}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_i(t_{ij})\Big)^2\right\}\cdot(\sigma^2)^{-a_\sigma-1}\exp\left\{\frac{-b_\sigma}{\sigma^2}\right\} \\ &\propto \exp\left\{-\frac{1}{\sigma^2}\left(\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_i(t_{ij})\Big)^2+b_\sigma\right)\right\}\cdot(\sigma^2)^{-a_\sigma-\frac{\sum_{i=1}^{N}n_i}{2}-1} \end{aligned}$$

Then we have

$$\sigma^2|\cdot\sim IG\left(a_\sigma+\frac{\sum_{i=1}^{N}n_i}{2},\quad \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_i(t_{ij})\Big)^2+b_\sigma\right)$$

4.6 Simulation study

We use the same mean trajectory functions as in Chapter 2 to generate data for our simulation studies. Note that in Chapter 2, our mean trajectory was of the form:

$$\psi(t)=\delta_0+\delta_1 t+\beta_1(t-\tau_1)_+ +\beta_2(t-\tau_2)_+.$$

Denoting $\boldsymbol{\theta}=(\delta_0,\delta_1,\beta_1,\beta_2)'$, we used two sets of $\boldsymbol{\theta}$ values in Chapter 2:

$$\boldsymbol{\theta}_1=(0.8,-20.6,50.2,-34.6)' \quad \text{and} \quad \boldsymbol{\theta}_2=(2.1,10.6,-40.2,32.6)'.$$

The mean trajectory function in this chapter is of the form:

$$\varphi(t)=A+B(t-\bar{t})+CB_{\tau_1}(t)+DB_{\tau_2}(t).$$

Let us denote $\tilde{\boldsymbol{\theta}} = (A, B, C, D)'$. If $\psi(t)$ and $\varphi(t)$ are alternative representations of the same trajectory, the relation between $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{\theta}}$ is given by

$$\begin{aligned}
\delta_0 + \delta_1 t &= A + B(t - \bar{t}_i) + C(a_{01} + b_{01}t) \\
&\quad + D(a_{02} + b_{02}t), \quad t \leq \tau_1 \\
\delta_0 + \delta_1 t + \beta_1(t - \tau_1) &= A + B(t - \bar{t}_i) + C(a_{11} + b_{11}t) \\
&\quad + D(a_{02} + b_{02}t), \quad \tau_1 < t \leq \tau_2 \\
\delta_0 + \delta_1 t + \beta_1(t - \tau_1) + \beta_2(t - \tau_2) &= A + B(t - \bar{t}_i) + C(a_{11} + b_{11}t) \\
&\quad + D(a_{12} + b_{12}t), \quad t > \tau_2,
\end{aligned}$$

which leads to

$$\begin{aligned}
C &= \frac{\beta_1}{b_{11} - b_{01}}, \\
D &= \frac{\beta_2}{b_{12} - b_{02}}, \\
B &= \delta_1 - b_{01}C - b_{02}D, \\
\text{and} \quad A &= \delta_0 + \bar{t}_i B - a_{01}C - a_{02}D.
\end{aligned}$$

Thus, the two sets of $\boldsymbol{\theta}$ values in Chapter 2 correspond to the the following two sets of $\tilde{\boldsymbol{\theta}}$ values:

$$\tilde{\boldsymbol{\theta}}_1 = (72.76, 0.66, -45.97, 158.24)' \quad \text{and} \quad \tilde{\boldsymbol{\theta}}_2 = (-141.26, -2.54, 36.82, -149.09)'$$

For each group, we generate data on 10 subjects. Since we only used one biomarker the computation speed is much faster than that for the model in Chapter 2. We ran 500,000 scans with 200,000 burn-in and thinning as 100.

Figure 4.2 and Figure 4.3 are the trace plots of selected parameters when σ^2 is updat-

ed. Since the data was generated with $\sigma^2 = 1$, we thus ran with $\sigma^2 = 1$ fixed. Figure 4.4 and Figure 4.5 are the trace plots of selected parameters when σ^2 is fixed at 1.

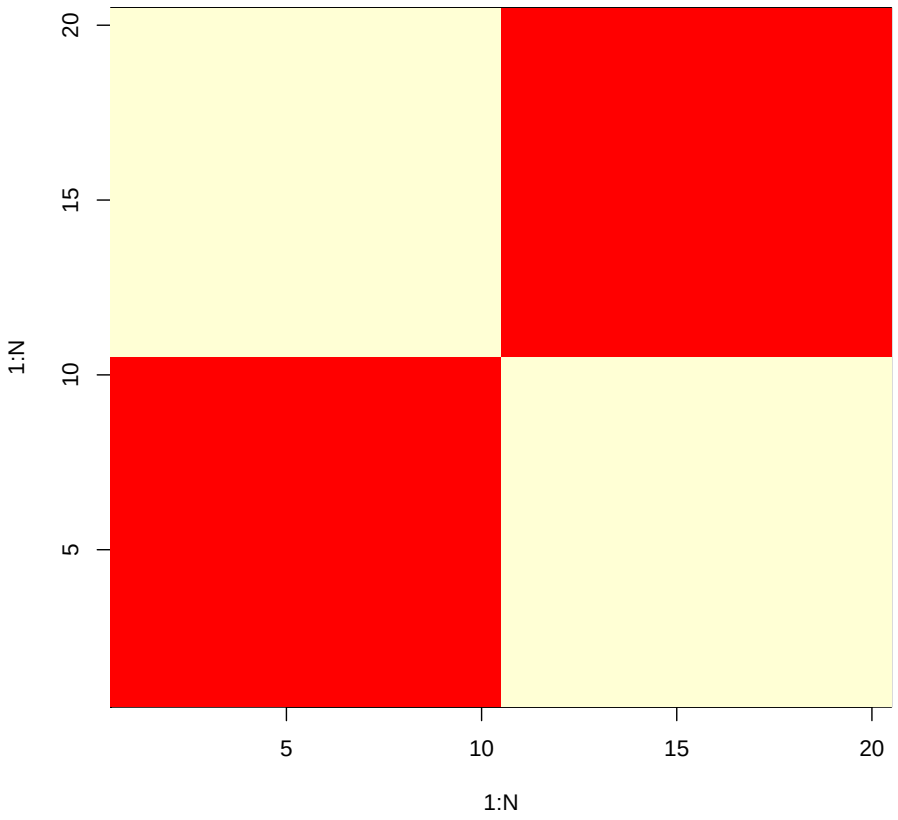

Figure 4.1. Plot of the heatmap of the clustering probabilities with σ^2 fixed at 1. Higher propensity to cluster together is indicated by deeper red color.

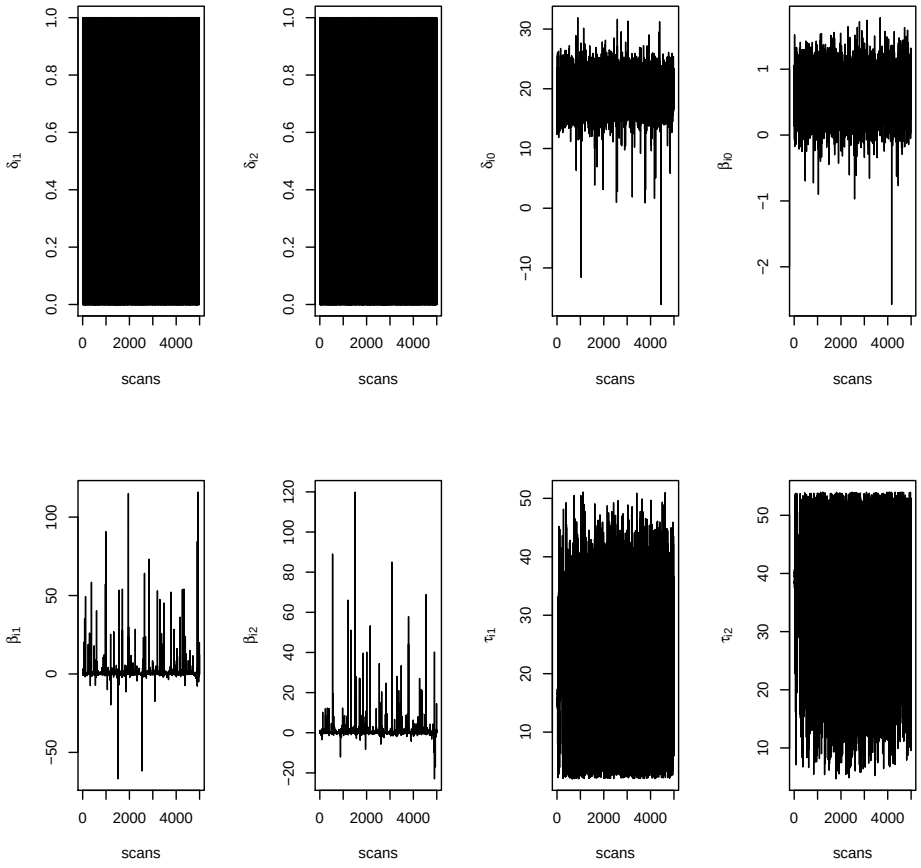

Figure 4.2. Trace plot of the parameters for the first individual with σ^2 updated.

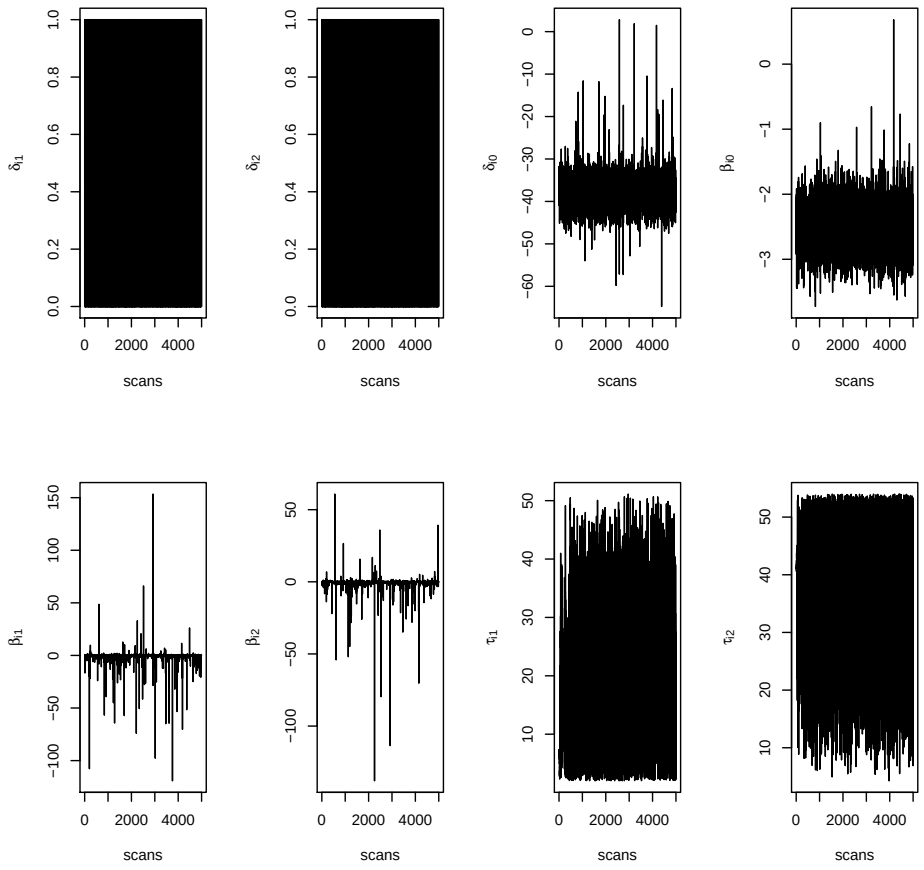

Figure 4.3. Trace plot of the parameters for the eleventh individual with σ^2 updated.

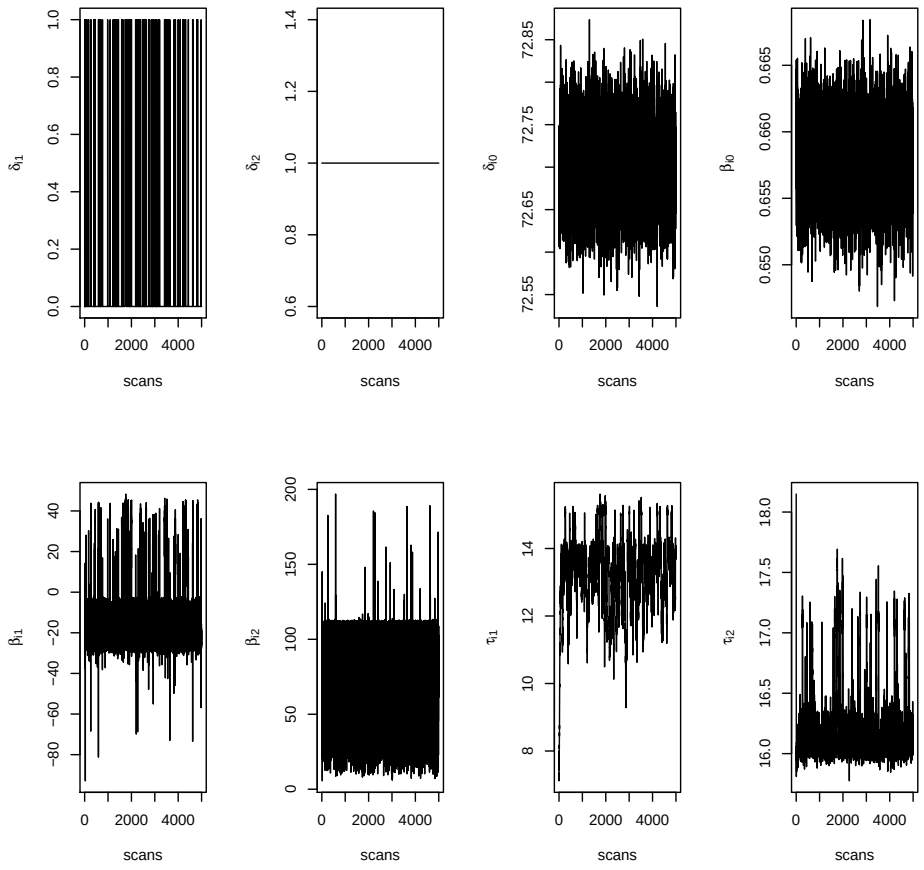

Figure 4.4. Trace plot of the parameters for the first individual with σ^2 fixed at 1.

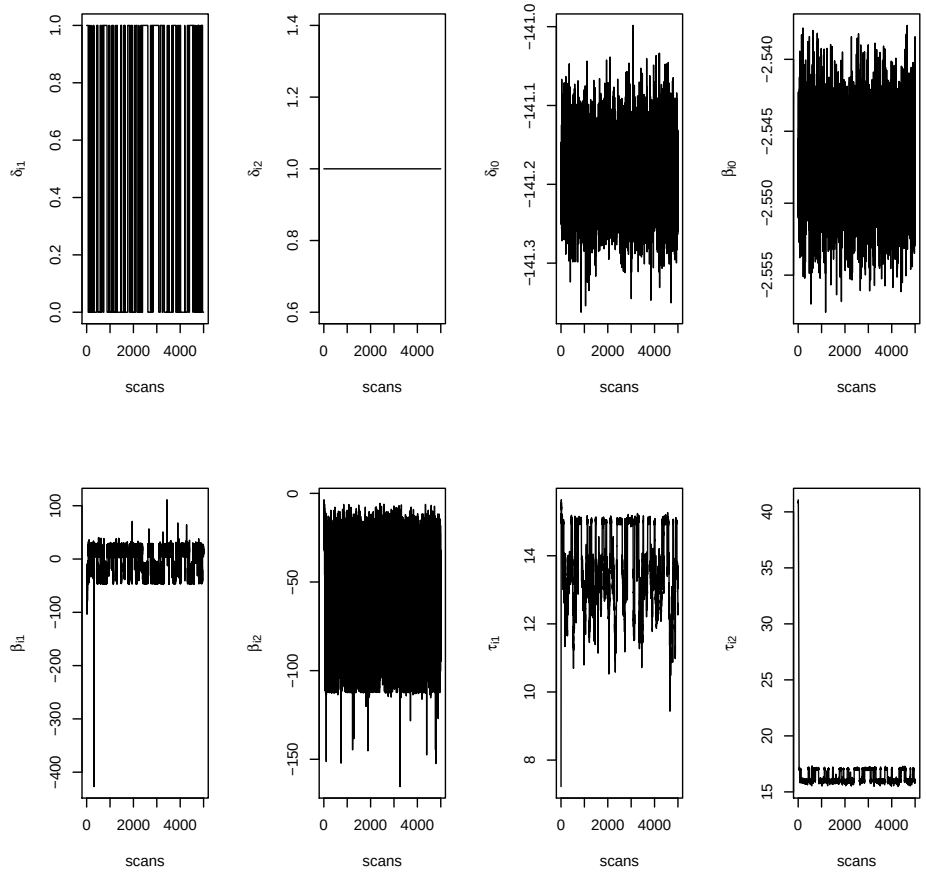

Figure 4.5. Trace plot of the parameters for the eleventh individual with σ^2 fixed at 1.

The following tables are the results of $\tilde{\boldsymbol{\theta}}$s and $\hat{\tilde{\boldsymbol{\theta}}}$s, where $\tilde{\boldsymbol{\theta}}$s are the parameter used to generate data and $\hat{\tilde{\boldsymbol{\theta}}}$s are the estimations. Table 4.1 gives the results with σ^2 updated and Table 4.2 gives the results with σ^2 fixed at 1. Note that although the estimated values of β_1 for the first individual has a different sign from the true value, since δ_1 is estimated to be close to zero, the effect is not significant.

Table 4.1. Results of parameter estimation with σ^2 updated

$\tilde{\boldsymbol{\theta}}_1$	$\hat{\tilde{\boldsymbol{\theta}}}_1$	$\tilde{\boldsymbol{\theta}}_2$	$\hat{\tilde{\boldsymbol{\theta}}}_2$
72.76	19.5	-141.26	-37.9
0.66	0.65	-2.54	-2.5
-45.97	0.01	36.82	-0.01
158.24	0.01	-149.09	0

Table 4.2. Results of parameter estimation with σ^2 fixed at 1

$\tilde{\boldsymbol{\theta}}_1$	$\hat{\tilde{\boldsymbol{\theta}}}_1$	$\tilde{\boldsymbol{\theta}}_2$	$\hat{\tilde{\boldsymbol{\theta}}}_2$
72.76	72.7	-141.26	-141.2
0.66	0.66	-2.54	-2.55
-45.97	17.5	36.82	10.1
158.24	69.7	-149.09	-55.4

The following tables are the results of $\boldsymbol{\tau}$s and $\hat{\boldsymbol{\tau}}$s, where $\boldsymbol{\tau}$s are the joinpoints used to generate data and $\hat{\boldsymbol{\tau}}$s are the estimates. Table 4.3 gives the results with σ^2 updated and Table 4.4 gives the result with σ^2 fixed at 1.

Table 4.3. Results of joinpoint estimation with σ^2 updated

	τ_1	$\hat{\tau}_1$	τ_2	$\hat{\tau}_2$
First individual	5	16.2	15	38.9
Eleventh individual	5	16.3	15	39

Table 4.4. Results of joinpoint estimation with σ^2 fixed at 1

	τ_1	$\hat{\tau}_1$	τ_2	$\hat{\tau}_2$
First individual	5	13.2	15	16.2
Eleventh individual	5	13.7	15	16.5

We see that the proposed method is able to correctly cluster the data. If we update σ^2, the estimated values of $\boldsymbol{\theta}$ and $\boldsymbol{\tau}$ are not close to the true values. If we fix σ^2 at 1, part of the estimated values of $\boldsymbol{\theta}$ are close to the true values and the estimated values of τ_2 are close to the true values.

CHAPTER 5

AN ALTERNATIVE PARAMETRIZATION WITH DROPOUT

In this chapter, we we extend the model in the previous chapter to incorporate the dropout information. As in Section 4.1, the contribution of the longitudinal observations of the ith subject to the likelihood is:

$$f(\boldsymbol{Y}_i|\varphi_i,\sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^{n_i}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_i(t_{ij})\Big)^2\right\}$$

5.1 Modeling dropout

Let s_i^d be the dropout time for the ith subject and c_i be the corresponding censoring time. The observed survival data then consists of the pair (s_i, δ_i) where $s_i = min(s_i^d, c_i)$ and $\delta_i = I(s_i = s_i^d)$. We assume that the hazard of dropout at time t is a function of trajectory at time t using a Cox proportional model of the form:

$$\lambda(t|\varphi) = \lambda_0(t)\exp\{\gamma\varphi(t)\}$$

Note that $\gamma = 0$ implies that the survival part is not affected by the longitudinal trajectory whereas $\gamma > 0$ implies that higher dropout rates will be associated with higher trajectory levels and vice-versa.

The baseline hazard function $\lambda_0(\cdot)$ is assumed to be piecewise constant, given by

$$\lambda_0(u) = \lambda_{0j}, \quad u_{j-1} \le u < u_j, \quad j = 1, \ldots, J,$$

where the number of steps J and the endpoints u_j are pre-specified with $u_0 \equiv 0$, and $u_J \equiv \infty$.

The contribution to the likelihood from the survival part of ith subject is then

$$f(s_i, \delta_i|\varphi_i) = \lambda_0(s_i)^{\delta_i} \exp\left\{\delta_i \gamma \varphi_i(s_i)\right\} \times \exp\left\{-\int_0^{s_i} \lambda_0(u) \exp\left[\gamma \varphi_i(u)\right] \mathrm{d}u\right\},$$

which can be written as

$$f(s_i, \delta_i|\varphi_i) = \lambda_0(s_i)^{\delta_i} \exp\Big\{\gamma \delta_i \varphi_i(s_i) - \sum_{j=1}^{J} \lambda_{0j} H_{ij}(\gamma)\Big\},$$

where

$$H_{ij} = I(s_i \geq u_{j-1}) \int_{u_{j-1}}^{u_j \wedge s_i} \exp\Big\{\gamma \varphi_i(u)\Big\} \mathrm{d}u.$$

The full likelihood is proportional to:

$$\prod_{j=1}^{J} \lambda_{0j}^{\sum_{i=1}^{N} \delta_i I(u_{j-1} \leq s_i \leq u_j)} \exp\Bigg\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_i(t_{ij})\Big)^2 + \gamma \sum_{i=1}^{N} \delta_i \varphi_i(s_i) - \sum_{i=1}^{N} \sum_{j=1}^{J} \lambda_{0j} H_{ij}(\gamma) \Bigg\}.$$

5.2 Prior distributions

Parameters $\boldsymbol{\delta}_i, \boldsymbol{\eta}_i, \boldsymbol{\beta}_i, \gamma, \sigma^2$ are assigned the same priors as in Section 4.1. Priors for the other parameters are taken to be

$$\lambda_{0j} \overset{iid}{\sim} G(a_j, b_j), \quad j = 1, \ldots, J,$$

$$\gamma \sim N(\mu_\gamma, \sigma_\gamma^2),$$

where $a_j, b_j, \mu_\gamma, \sigma_\gamma^2$ are known.

5.3 Posteriors of $\boldsymbol{\theta}_p^*$

Let $\boldsymbol{\theta}_p^* = (\delta_1^*, \delta_2^*, \delta_0^*, \beta_0^*, \beta_1^*, \beta_2^*, \tau_1^*, \tau_2^*)$ be the pth distinct value of the random effect vector with ($p = 1, \ldots, k_\theta$.) With the dropout added, the baseline prior is not conjugate to the sampling distribution. The configuration structure $\boldsymbol{c}^\theta = (c_1^\theta, c_2^\theta, \ldots, c_N^\theta)$ of the random effects is updated using Algorithm 8 in Neal (2000).

Let $\boldsymbol{\delta}_p^* = (\delta_1^*, \delta_2^*)'$, the posterior of $\boldsymbol{\delta}_p^*|\cdot$ is

$$
\begin{aligned}
f(\boldsymbol{\delta}_p^*|\cdot) &\propto L(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\delta}) \\
&\propto \exp\Bigg\{ -\frac{1}{2\sigma^2} \sum_{i:c_i^\theta=p} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_{ij}(t_{ij})\Big)^2 \\
&\qquad + \gamma \sum_{i:c_i^\theta=p} \delta_i \varphi_i(s_i) - \sum_{i:c_i^\theta=p} \sum_{j=1}^{J} \lambda_{0j} H_{ij}(\gamma) \Bigg\} \cdot \pi(\boldsymbol{\delta}).
\end{aligned}
$$

Since $\pi(\boldsymbol{\delta}) = \frac{1}{4}$, we then have

$$
\begin{aligned}
\log f(\boldsymbol{\delta}_p^*|\cdot) = const. &- \frac{1}{2\sigma^2} \sum_{i:c_i^\theta=p} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_{ij}(t_{ij})\Big)^2 \\
&+ \gamma \sum_{i:c_i^\theta=p} \delta_i \varphi_i(s_i) - \sum_{i:c_i^\theta=p} \sum_{j=1}^{J} \lambda_{0j} H_{ij}(\gamma).
\end{aligned}
$$

Let $\boldsymbol{\eta}_p^* = (\delta_0^*, \beta_0^*)'$, the posterior of $\boldsymbol{\eta}_p^*|\cdot$ is:

$$
\begin{aligned}
f(\boldsymbol{\eta}_p^*|\cdot) &\propto L(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\eta}) \\
&\propto \exp\Bigg\{ -\frac{1}{2\sigma^2} \sum_{i:c_i^\theta=p} \sum_{j=1}^{n_i} \Big(y_{ij} - \varphi_{ij}(t_{ij})\Big)^2 \\
&\qquad + \gamma \sum_{i:c_i^\theta=p} \delta_i \varphi_i(s_i) - \sum_{i:c_i^\theta=p} \sum_{j=1}^{J} \lambda_{0j} H_{ij}(\gamma) \Bigg\} \\
&\quad \times \exp\Big\{ -\frac{1}{2} (\boldsymbol{\eta}_p^* - \boldsymbol{\mu}_{\eta 0})' \boldsymbol{\Sigma}_{\eta 0}^{-1} (\boldsymbol{\eta}_p^* - \boldsymbol{\mu}_{\eta 0}) \Big\}.
\end{aligned}
$$

Thus we have

$$\begin{aligned}\log f(\boldsymbol{\eta}_p^*|\cdot) = const. &- \frac{1}{2\sigma^2}\sum_{i:c_i^\theta=p}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_{ij}(t_{ij})\Big)^2 \\ &+ \gamma\sum_{i:c_i^\theta=p}\delta_i\varphi_i(s_i) - \sum_{i:c_i^\theta=p}\sum_{j=1}^{J}\lambda_{0j}H_{ij}(\gamma) \\ &- \frac{1}{2}(\boldsymbol{\eta}_p^*-\boldsymbol{\mu}_{\eta 0})'\boldsymbol{\Sigma}_{\eta 0}^{-1}(\boldsymbol{\eta}_p^*-\boldsymbol{\mu}_{\eta 0}).\end{aligned}$$

Let $\boldsymbol{\beta}_p^* = (\beta_1^*, \beta_2^*)'$, we first need to find the prior marginal density of $\boldsymbol{\beta}$, which is given by

$$\begin{aligned}\pi(\boldsymbol{\beta}) &= \int_0^\infty p(\boldsymbol{\beta}|\boldsymbol{\gamma})\pi(\gamma)\mathrm{d}\gamma \\ &= \int_0^\infty \frac{e^{-\frac{1}{2\gamma}(\boldsymbol{\beta}'\Sigma^{-1}\boldsymbol{\beta})}}{(2\pi)^{\frac{J^*}{2}}|\gamma\Sigma|^{\frac{1}{2}}}\cdot\frac{(0.5)^{0.5}}{\Gamma(0.5)}\cdot\gamma^{-0.5-1}e^{-\frac{0.5}{\gamma}}\mathrm{d}\gamma \\ &= \frac{(0.5)^{0.5}}{(2\pi)^{\frac{J^*}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(0.5)}\int_0^\infty e^{-\frac{1}{\gamma}(\frac{\boldsymbol{\beta}'\Sigma^{-1}\boldsymbol{\beta}+1}{2})}\cdot\gamma^{-0.5-1-\frac{J^*}{2}}\mathrm{d}\gamma \\ &= \frac{(0.5)^{0.5}}{(2\pi)^{\frac{J^*}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(0.5)}\cdot\frac{\Gamma(0.5+\frac{J^*}{2})}{(\frac{\boldsymbol{\beta}'\Sigma^{-1}\boldsymbol{\beta}+1}{2})^{0.5+\frac{J^*}{2}}}.\end{aligned}$$

Then the posterior of $\boldsymbol{\beta}_p^*|\cdot$ is

$$\begin{aligned}f(\boldsymbol{\beta}_p^*|\cdot) &\propto L(\boldsymbol{\theta})\cdot\pi(\boldsymbol{\beta}) \\ &\propto \exp\Bigg\{-\frac{1}{2\sigma^2}\sum_{i:c_i^\theta=p}\sum_{j=1}^{n_i}\Big(y_{ij}-\varphi_{ij}(t_{ij})\Big)^2 \\ &\quad + \gamma\sum_{i:c_i^\theta=p}\delta_i\varphi_i(s_i) - \sum_{i:c_i^\theta=p}\sum_{j=1}^{J}\lambda_{0j}H_{ij}(\gamma)\Bigg\} \\ &\quad\times\frac{1}{(\frac{\boldsymbol{\beta}^{*\prime}\Sigma^{-1}\boldsymbol{\beta}^*+1}{2})^{0.5+\frac{J^*}{2}}},\end{aligned}$$

resulting in

$$\log f(\boldsymbol{\beta}_p^*|\cdot) = const. - \frac{1}{2\sigma^2}\sum_{i:c_i^\theta=p}\sum_{j=1}^{n_i}\Big(y_{ij} - \varphi_{ij}(t_{ij})\Big)^2 + \gamma\sum_{i:c_i^\theta=p}\delta_i\varphi_i(s_i) - \sum_{i:c_i^\theta=p}\sum_{j=1}^{J}\lambda_{0j}H_{ij}(\gamma) + \log\frac{1}{(\frac{\boldsymbol{\beta}^{*\prime}\Sigma^{-1}\boldsymbol{\beta}^*+1}{2})^{0.5+\frac{J^*}{2}}}.$$

5.4 The full conditional of $\boldsymbol{\alpha}|\cdot$

The posterior of $\boldsymbol{\alpha}|\cdot$ is

$$f(\boldsymbol{\alpha}|\cdot) \propto L(\boldsymbol{\theta}) \cdot \pi(\boldsymbol{\alpha}|\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha),$$

which results in

$$\log f(\boldsymbol{\alpha}|\cdot) = const. - \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\mu}_\alpha)'\boldsymbol{\Sigma}_\alpha^{-1}(\boldsymbol{\alpha} - \boldsymbol{\mu}_\alpha) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\Big(y_{ij} - \varphi_i(t_{ij})\Big)^2 + \gamma\sum_{i=1}^{N}\delta_i\varphi_i(s_i) - \sum_{i=1}^{N}\sum_{j=1}^{J}\lambda_{0j}H_{ij}(\gamma).$$

5.5 The full conditional of $\sigma^2|\cdot$

The full conditional of $\sigma^2|\cdot$ is the sam as in Section 4.5.

5.6 The full conditional of $\lambda_{0j}|\cdot$

The posterior of $\lambda_{0j}|\cdot$ is

$$f(\lambda_{0j}|\cdot) \propto \lambda_{0j}^{\sum_{i=1}^{N}\delta_i I(u_{j-1}\leq s_i\leq u_j)}\exp\left\{-\sum_{i=1}^{N}\lambda_{0j}H_{ij}(\gamma)\right\}$$
$$\cdot\, {\lambda_{0j}}^{a_j-1}\exp\left\{-b_j\lambda_{0j}\right\}.$$

Then

$$\lambda_{0j}|\cdot \sim G\left(a_j+\sum_{A_j}\delta_i,\quad b_j+\sum_{i=1}^{N}H_{ij}(\gamma)\right),$$

where

$$A_j=\{i: s_i\in[u_{j-1},u_j]\}.$$

5.7 The full conditional of $\gamma|\cdot$

The posterior of $\gamma|\cdot$ is

$$f(\gamma|\cdot)\propto L(\boldsymbol{\theta})\cdot\pi(\gamma|\mu_\gamma,\sigma^2_\gamma),$$

which results in

$$\log f(\gamma|\cdot)=const.-\frac{\gamma^2-2\mu_\gamma\gamma}{2\sigma^2_\gamma}+\gamma\sum_{i=1}^{N}\delta_i\varphi_i(s_i)-\sum_{i=1}^{N}\sum_{j=1}^{J}\lambda_{0j}H_{ij}(\gamma).$$

5.8 Calculation of H_{ij}

To calculate

$$H_{ij}=I(s_i\geq u_{j-1})\int_{u_{j-1}}^{u_j\wedge s_i}\exp\Big\{\gamma\varphi_i(u)\Big\}\mathrm{d}u,$$

we write

$$\begin{aligned}\varphi_i(u) &= \boldsymbol{\alpha}^{'}\boldsymbol{x}_i + \delta_{i0} + \beta_{i0}(u - \bar{t}_i) + \sum_{l=1}^{L} \delta_{il}\beta_{il}B_{\tau_{il}}(u) \\ &= \boldsymbol{\alpha}^{'}\boldsymbol{x}_i + \delta_{i0} - \beta_{i0}\bar{t}_i + \beta_{i0}u + \delta_{i1}\beta_{i1}B_{\tau_{i1}}(u) + \cdots + \delta_{iL}\beta_{iL}B_{\tau_{iL}}(u).\end{aligned}$$

Note that

$$\begin{aligned}&\delta_{i1}\beta_{i1}B_{\tau_1}(u) + \cdots \delta_{iL}\beta_{iL}B_{\tau_{iL}}(u) \\ &= \begin{cases} \delta_{i1}\beta_{i1}(a_{i01} + b_{i01}u) + \delta_{i2}\beta_{i2}(a_{i02} + b_{i02}u) & \\ \quad + \cdots + \delta_{iL}\beta_{iL}(a_{i0L} + b_{i0L}u) \quad \text{if} \quad u \le \tau_{i1} \\ \delta_{i1}\beta_{i1}(a_{i11} + b_{i11}u) + \delta_{i2}\beta_{i2}(a_{i02} + b_{i02}u) & \\ \quad + \cdots + \delta_{iL}\beta_{iL}(a_{i0L} + b_{i0L}u) \quad \text{if} \quad \tau_{i1} < u \le \tau_{i2} \\ \delta_{i1}\beta_{i1}(a_{i11} + b_{i11}u) + \delta_{i2}\beta_{i2}(a_{i12} + b_{i12}u) + \delta_{i3}\beta_{i3}(a_{i03} + b_{i03}u) & \\ \quad + \cdots + \delta_{iL}\beta_{iL}(a_{i0L} + b_{i0L}u) \quad \text{if} \quad \tau_{i2} < u \le \tau_{i3} \\ \qquad \vdots & \\ \delta_{i1}\beta_{i1}(a_{i11} + b_{i11}u) + \delta_{i2}\beta_{i2}(a_{i12} + b_{i12}u) + \delta_{i3}\beta_{i3}(a_{i13} + b_{i13}u) & \\ \quad + \cdots + \delta_{iL}\beta_{iL}(a_{i1L} + b_{i1L}u) \quad \text{if} \quad u > \tau_{iL}. \end{cases}\end{aligned}$$

Hence

$$\varphi_i(u) = \sum_{l=0}^{L}(a_{il} + \rho_{il}u) \cdot 1(u)_{(\tau_l, \tau_{l+1})},$$

where

$$a_{il} = \boldsymbol{\alpha}^{'}\boldsymbol{x}_i + \delta_{i0} - \beta_{i0}\bar{t}_i + \sum_{\substack{t \le l \\ t \ge 1}} \delta_{it}\beta_{it}a_{i1t} + \sum_{\substack{t > l \\ t \ge 1}} \delta_{it}\beta_{it}a_{i0t}, \quad l = 0, 1, 2, \ldots, L,$$

and

$$\rho_{il} = \beta_{i0} + \sum_{\substack{t \le l \\ t \ge 1}} \delta_{it}\beta_{it}b_{i1t} + \sum_{\substack{t > l \\ t \ge 1}} \delta_{it}\beta_{it}b_{i0t}, \quad l = 0, 1, 2, \ldots, L.$$

Then

$$\exp\{\gamma\varphi_i(u)\} = \exp\left\{\gamma\sum_{l=0}^{L}(a_{il}+\rho_{il}u)\right\}\cdot 1(u)_{(\tau_l,\tau_{l+1})}$$
$$= \sum_{l=0}^{L}\exp\{\gamma(a_{il}+\rho_{il}u)\}\cdot 1(u)_{(\tau_l,\tau_{l+1})}.$$

Hence

$$\int_0^x \exp\{\gamma\varphi_i(u)\}\,\mathrm{d}u$$
$$= \sum_{l=0}^{L}\int_{\tau_l\wedge x}^{\tau_{l+1}\wedge x}\exp\Big\{\gamma(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u$$
$$= \begin{cases} x \quad \text{if} \quad \gamma = 0, \\ \displaystyle\sum_{l=0}^{L}\exp\{\gamma a_{il}\}\left[\Big(\big(\exp\{\gamma\rho_{il}(\tau_{l+1}\wedge x)\}-\exp\{\gamma\rho_{il}(\tau_l\wedge x)\}\big)/\gamma\rho_{il}\Big)\right. \\ \left.\cdot 1_{\{\rho_{il}\neq 0\}} + \Big(\tau_{l+1}\wedge x - \tau_l\wedge x\Big)\cdot 1_{\{\rho_{il}=0\}}\right] \quad \text{if} \quad \gamma\neq 0. \end{cases}$$

Then

$$\int_{x_1}^{x_2}\exp\left\{\sum_{l=0}^{L}\gamma\varphi_i(u)\right\}\mathrm{d}u$$
$$= \int_0^{x_2}\exp\left\{\sum_{l=0}^{L}\gamma\varphi_i(u)\right\}\mathrm{d}u - \int_0^{x_1}\exp\left\{\sum_{l=0}^{L}\gamma\varphi_i(u)\right\}\mathrm{d}u$$
$$= \sum_{l=0}^{L}\int_{\tau_l\wedge x_2}^{\tau_{l+1}\wedge x_2}\exp\Big\{\gamma(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u - \sum_{l=0}^{L}\int_{\tau_l\wedge x_1}^{\tau_{l+1}\wedge x_1}\exp\Big\{\gamma(a_{il}+\rho_{il}u)\Big\}\mathrm{d}u$$
$$= \begin{cases} x_2 - x_1 \quad \text{if} \quad \gamma = 0, \\ \displaystyle\sum_{l=0}^{L}\exp\{\gamma a_{il}\}\left[\left(\Big(\big(\exp\{\gamma\rho_{il}(\tau_{l+1}\wedge x_2)\}-\exp\{\gamma\rho_{il}(\tau_l\wedge x_2)\}\big)\right.\right. \\ \left.\left.-\big(\exp\{\gamma\rho_{il}(\tau_{l+1}\wedge x_1)\}-\exp\{\gamma\rho_{il}(\tau_l\wedge x_1)\}\big)\Big)/\gamma\rho_{il}\right)\cdot 1_{\{\rho_{il}\neq 0\}} \\ \left.+\Big(\big(\tau_{l+1}\wedge x_2-\tau_l\wedge x_2\big)-\big(\tau_{l+1}\wedge x_1-\tau_l\wedge x_1\big)\Big)\cdot 1_{\{\rho_{il}=0\}}\right] \quad \text{if} \quad \gamma\neq 0. \end{cases}$$

Thus we have

$$
\begin{aligned}
H_{ij} &= I(s_i \geq u_{j-1}) \int_{u_{j-1}}^{u_j \wedge s_i} \exp\{\gamma \varphi_i(u)\}\, \mathrm{d}u \\
&= \begin{cases}
I(s_i \geq u_{j-1})(u_j \wedge s_i - u_{j-1}) \quad \text{if} \quad \gamma = 0, \\
\sum_{l=0}^{L} \exp\{\gamma a_{il}\} \Bigg[\Bigg(\Big(\big(\exp\{\gamma \rho_{il}(\tau_{l+1} \wedge u_j \wedge s_i)\} - \exp\{\gamma \rho_{il}(\tau_l \wedge u_j \wedge s_i)\} \big) \\
- \big(\exp\{\gamma \rho_{il}(\tau_{l+1} \wedge u_{j-1})\} - \exp\{\gamma \rho_{il}(\tau_l \wedge u_{j-1})\} \big) \Big) / \gamma \rho_{il} \Bigg) \times 1_{\{\rho_{il} \neq 0\}} \\
+ \Big(\big(\tau_{l+1} \wedge u_j \wedge s_i - \tau_l \wedge u_j \wedge s_i \big) - \big(\tau_{l+1} \wedge u_{j-1} - \tau_l \wedge u_{j-1} \big) \Big) \\
\times 1_{\{\rho_{il} = 0\}} \Bigg] \quad \text{if} \quad \gamma \neq 0.
\end{cases}
\end{aligned}
$$

5.9 Simulation study

We used the same set of generated data for the longitudinal part as in Chapter 4 and the same set of generated data for the dropout part as in Chapter 3. However, in this chapter, the dropout has only one cause. We used two groups of generated data each with 50 subjects. We ran 50,000 scans with 20,000 burn-in and thinning of 10. Since in the previous chapter we saw that when σ^2 is fixed at 1, we get better estimates than when σ^2 is updated, we do our simulation in this chapter only with σ^2 fixed at 1.

Figure 5.1 is the trace plot of $\boldsymbol{\theta}$ associated with the first individual and Figure 5.2 is the trace plot of $\boldsymbol{\theta}$ associated with the fifty first individual. From Figure 5.3 we see that he estimation of $\boldsymbol{\alpha}$ is close to the true values which are all zeros. From Figure 5.4 we see that γ is convergent to 0.0006, which implies that the implies that higher dropout rates will be associated with higher trajectory levels and vice-versa. From

Figure 5.5 we see that the piecewise constant baseline hazard for the dropout, λ_{0j}, are all fairly small except in the last interval, which means individuals tend to drop out in this interval with high probability. From Figure 5.6 we see that the proposed method is able to correctly cluster the data.

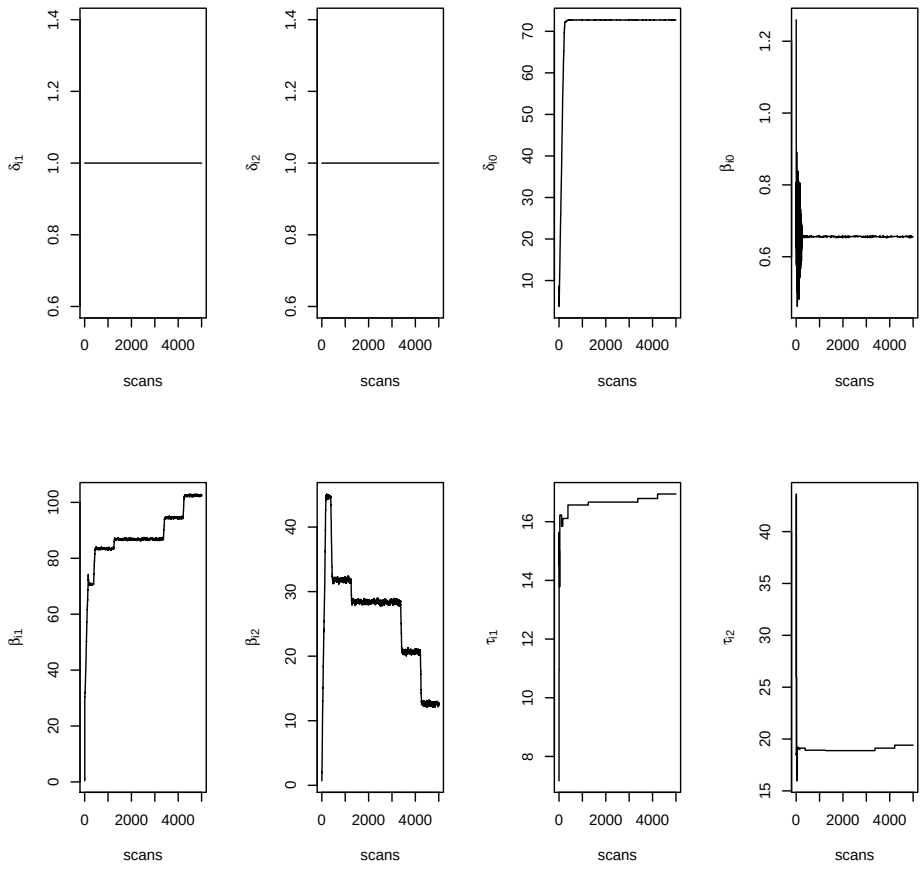

Figure 5.1. Trace plot of the parameters associated with the first individual.

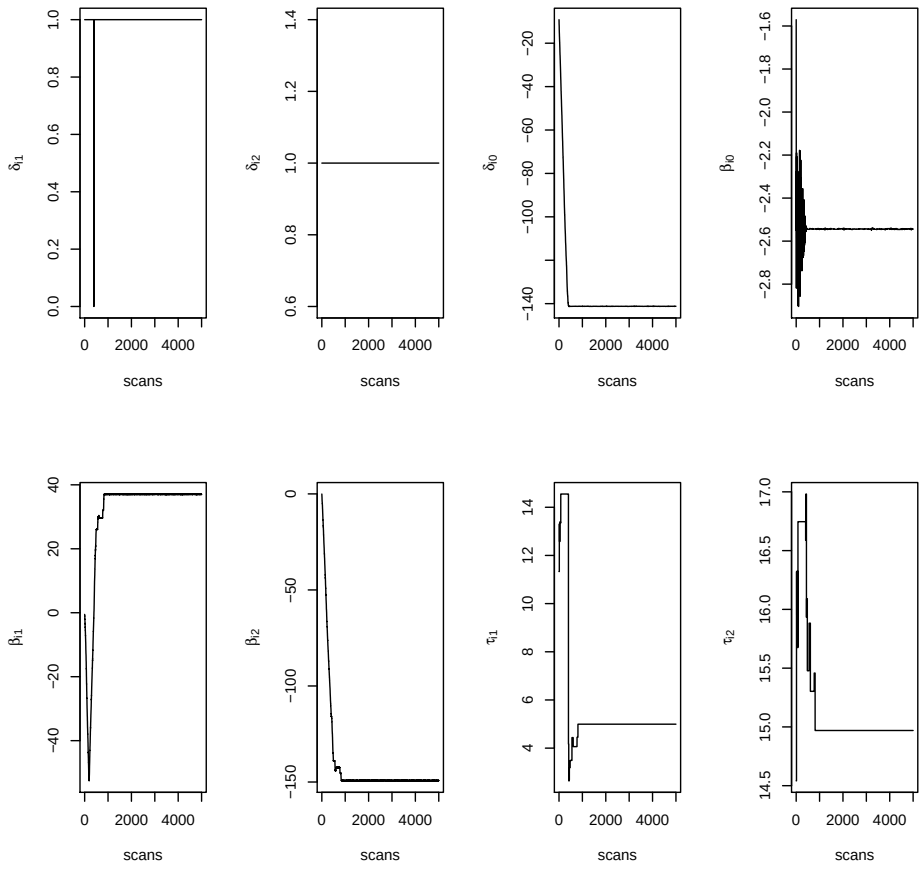

Figure 5.2. Trace plot of the parameters associated with the fifty first individual.

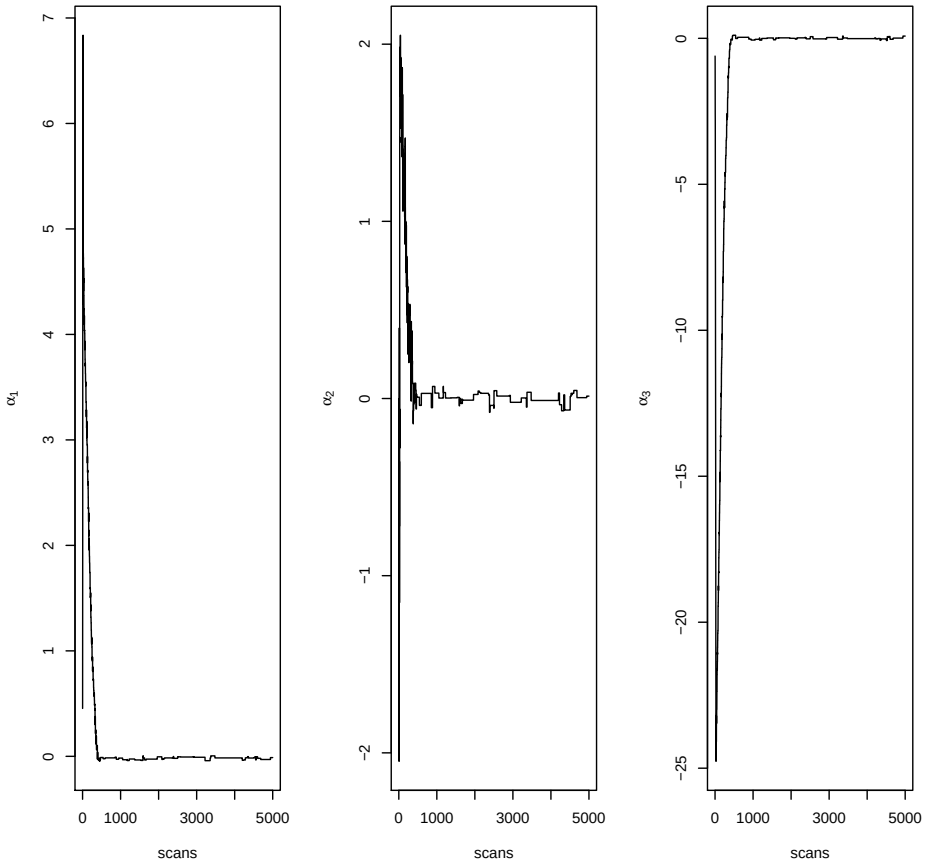

Figure 5.3. Trace plot of $\boldsymbol{\alpha}$.

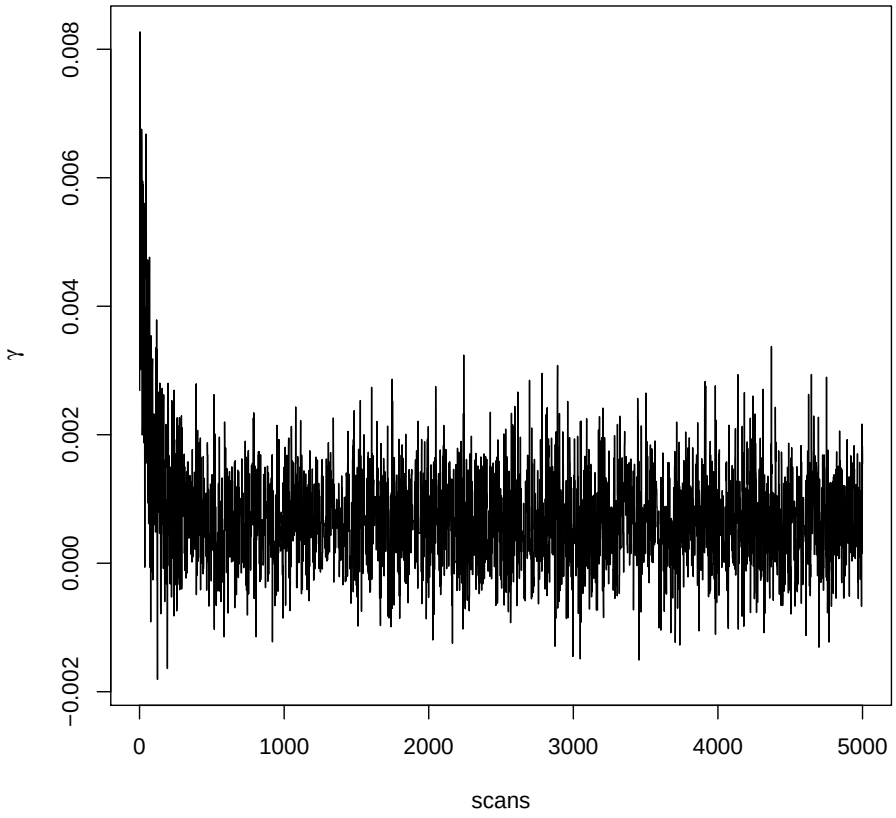

Figure 5.4. Trace plot of γ.

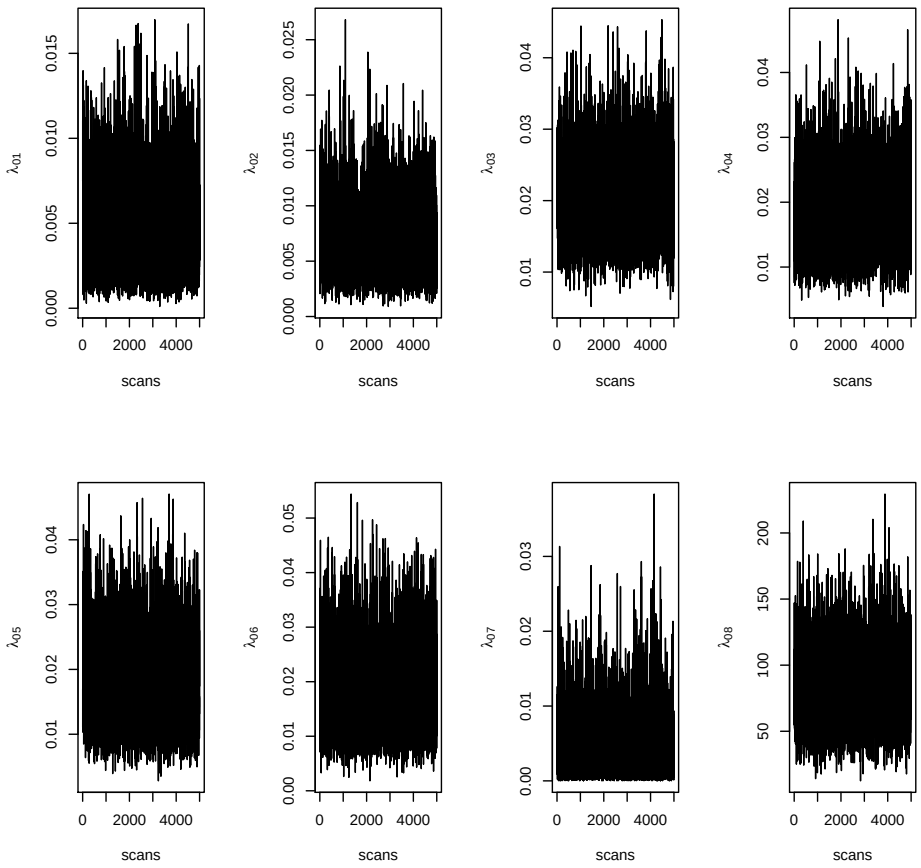

Figure 5.5. Trace plot of λ_{0j}.

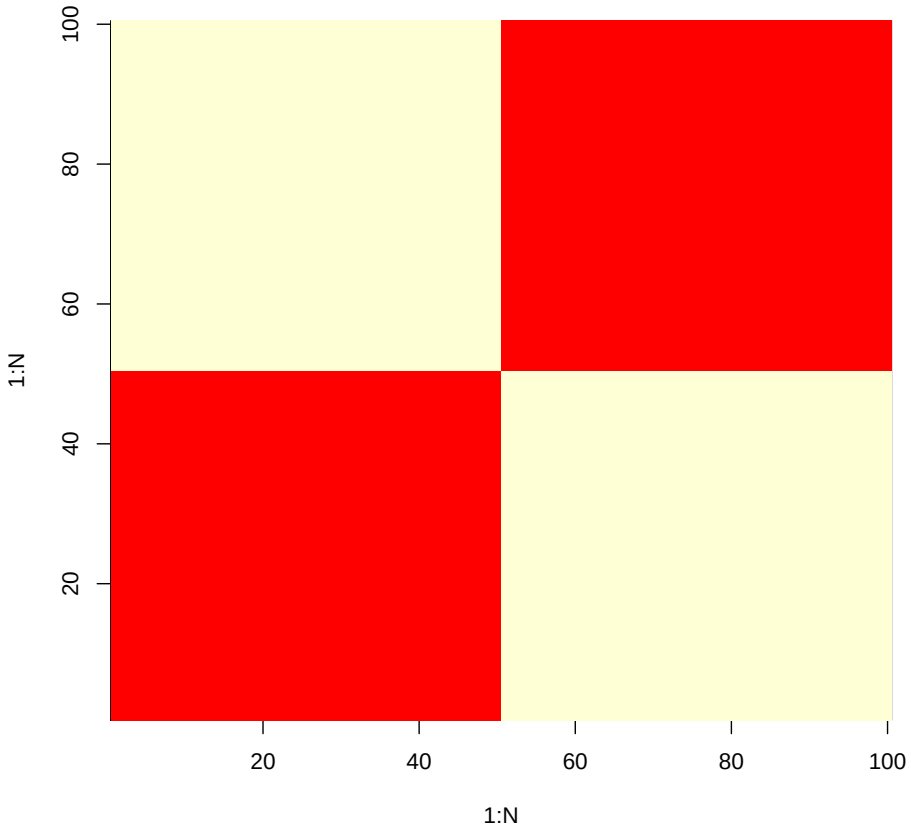

Figure 5.6. Plot of the heatmap of the clustering probabilities. Higher propensity to cluster together is indicated by deeper red color.

Table 5.1 gives the results of estimations for $\tilde{\boldsymbol{\theta}}$s, where $\tilde{\boldsymbol{\theta}}$s are the parameters used to generate data and $\hat{\tilde{\boldsymbol{\theta}}}$s are the estimated values.

Table 5.1. Results of parameter estimation

$\tilde{\boldsymbol{\theta}}_1$	$\hat{\tilde{\boldsymbol{\theta}}}_1$	$\tilde{\boldsymbol{\theta}}_2$	$\hat{\tilde{\boldsymbol{\theta}}}_2$
72.76	72.73	-141.26	-141.2
0.66	0.66	-2.54	-2.54
-45.97	86.9	36.82	37.01
158.24	28.3	-149.09	-149.2

Table 5.2 gives the results of estimations for $\boldsymbol{\tau}$s, where $\boldsymbol{\tau}$s are the parameters used to generate data and $\hat{\tau}$s are the estimations.

Table 5.2. Results of joinpint estimation

	τ_1	$\hat{\tau}_1$	τ_2	$\hat{\tau}_2$
First individual	5	16.7	15	18.9
Fifty first individual	5	5.7	15	15.2

From Table 5.1, we see that the first selected individual has the first and the second number close to the true values and the fifty first selected individual has all the numbers close to the true values. From Table 5.2, we see that for the first individual the estimations of joinpoints are not close to the true values but for the eleventh individual the estimations of the joinpoints are close to the true values.

Our model is able to correctly cluster the individuals and correctly estimate the

intercept and slope of the trajectory function. However, it can not get the accurate estimates for the change of slopes and joinpoints.

CHAPTER 6

AN APPLICATION OF JOINT MODELING

We motivate the need for a flexible joint model using data of plasma HIV RNA from the AIDS Clinical Trials Group (ACTG) 398 study. Detailed description of the data and results of primary analysis are available in Hammer et al. (2002). ACTG 398 was a randomized double-blinded placebo-controlled study comparing a single protease inhibitor (PI) with double-PI antiretroviral regimens in treating HIV-infected patients. The primary objective of the study was to compare the proportion of subjects who had virologic failure after 24 weeks on study between the double-PI arms and the single-PI arm. Four hundred and eighty-one subjects were followed for varying durations (some, as long as 72 weeks) and their viral loads were tracked. We are interested in modeling these multiple phases of change in the viral load trajectory. As noted by Liang (2007), apart from better modeling of the trajectory, identifying these changes can have important clinical and biological ramifications, such as helping us to identify the time to change drugs and thus avoid the problem of drug resistance. It will also help us to investigate the extent to which viral rebound can be predicted by initial response to treatment, which can improve overall prognosis.

Patients with HIV who show more rapid increase in viral RNA counts are more likely to terminate the study due to sickness or death, or may withdraw from the trial altogether to seek alternate treatment options. One aspect of the ACTG 398 trial is the high toxicity rate, due to the high drug burden and the advanced stage of

infection in the study population. Approximately 49 per cent of the subjects went off-study-treatment (stopped at least one drug) due to toxicity by week 72. When subjects go off-study-treatment, their viral load trajectories are immediately affected. To account for this phenomenon, the viral RNA values are censored at the time of subject going off-study-treatment. This is potentially informative dropout mechanism in the sense that the tendency to drop out at any point is related to the level of the (longitudinal) marker variable. Without correctly accounting for this informative dropout, parameter estimation in a longitudinal model becomes biased. Thus, for modeling the longitudinal data with informative dropout, a joint statistical model is needed for the dropout process, in addition to the model for the longitudinal viral RNA marker.

Due to complexity of a joint model, in this chapter we only use the longitudinal model from Chapter 4 to explore the estimating of the parameters in the trajectory function. The full data set makes the computing speed too slow thus we use just the first 100 subjects. We ran 50,000 scans with 20,000 burn-in and thinning as 100. We selected one individual each from the first two cluster groups and gave the results of the parameter estimations. The trace plot of the parameters associated with them are in Figure 6.1 and Figure 6.2. Figure 6.3 shows that the covariates effects are not zeros, which implies that it does have some effects to the longitudinal outcome.

The way of identifying the cluster-structure in the data is through the use of heatmap routine in R. The routine uses the D matrix as a form of distance matrix, and a hierarchical clustering algorithm is then run using the distance matrix specified. Individuals more likely to cluster together in the MCMC iterations are thus identified,

with darker colors indicating higher propensity to cluster together. The heatmap of the clustering likelihood of the individual patients for our case is given in Figure 6.4 and Figure 6.5 where Figure 6.5 is the grouped graph of Figure 6.4. Figure 6.5 shows that our non-parametric procedure classified the 100 patients into 8 clusters.

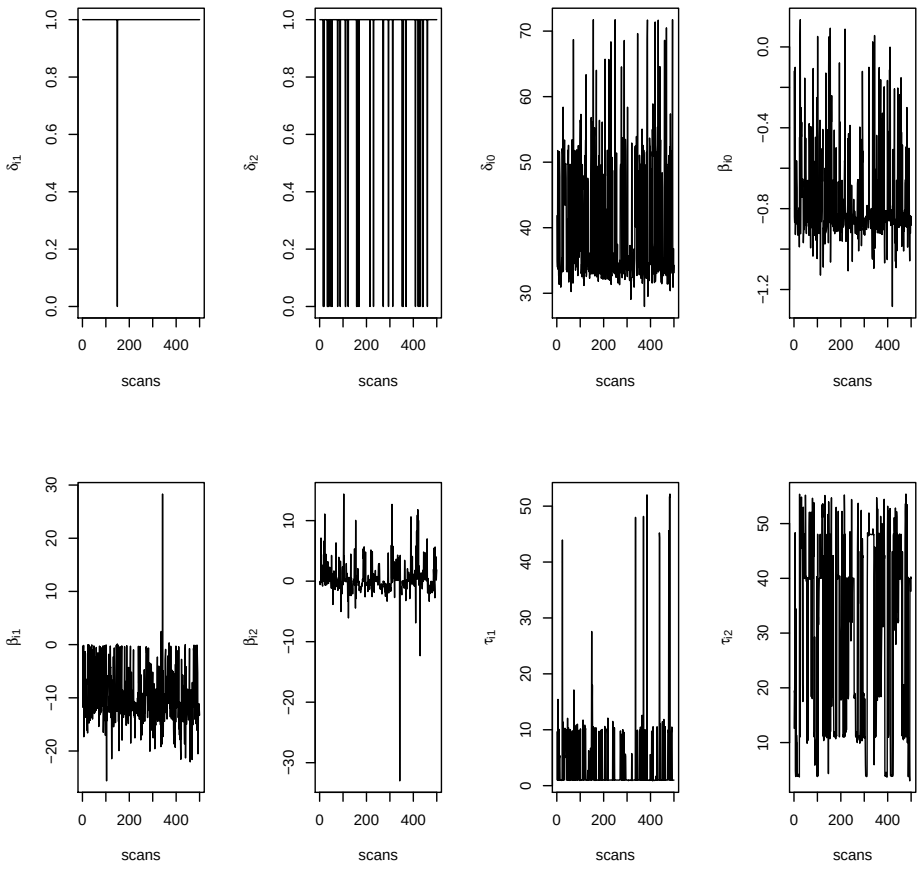

Figure 6.1. Trace plot of the parameters associated with the fifty first individual.

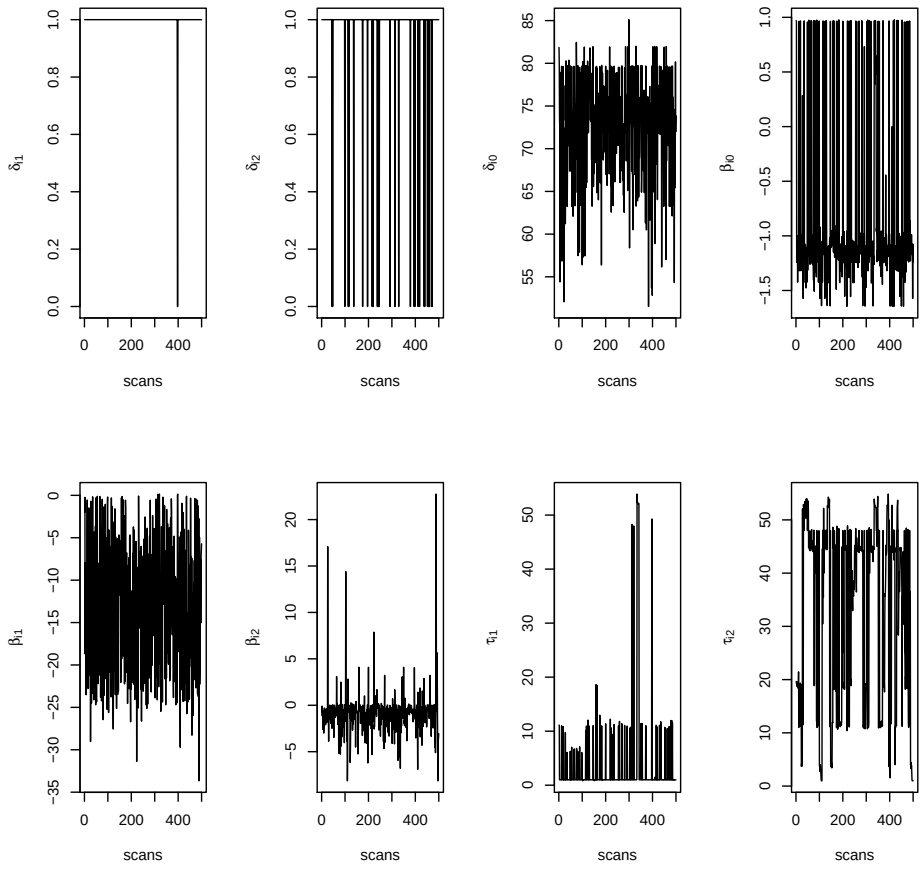

Figure 6.2. Trace plot of the parameters associated with the tenth individual.

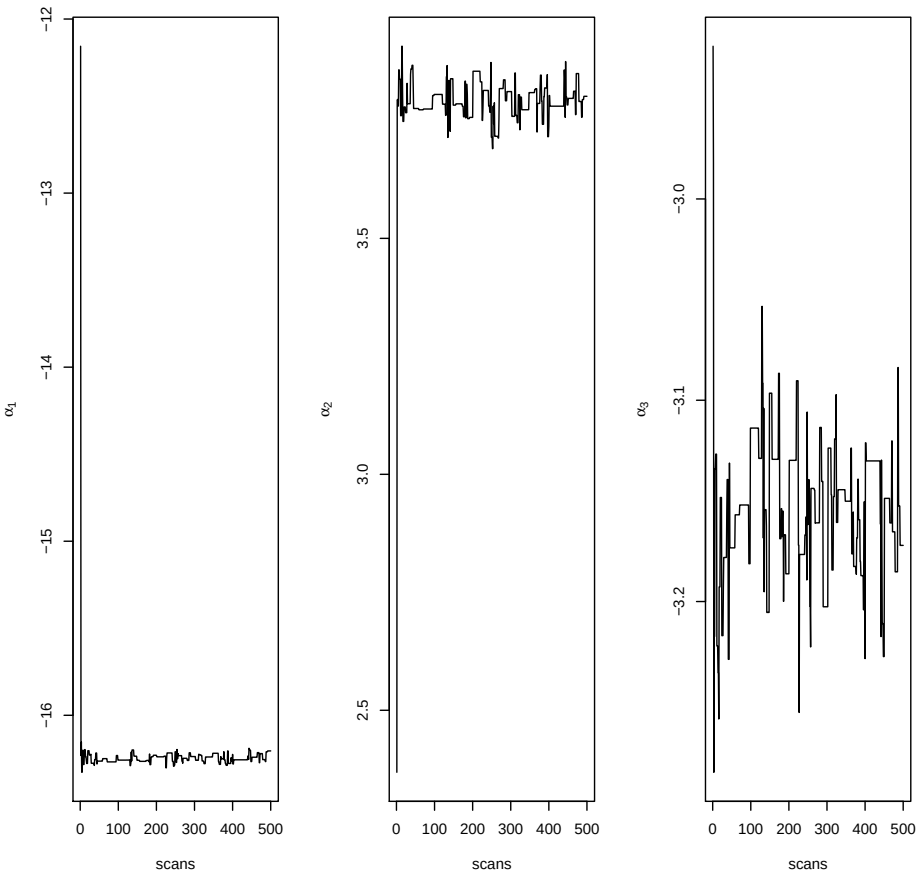

Figure 6.3. Trace plot of $\boldsymbol{\alpha}$.

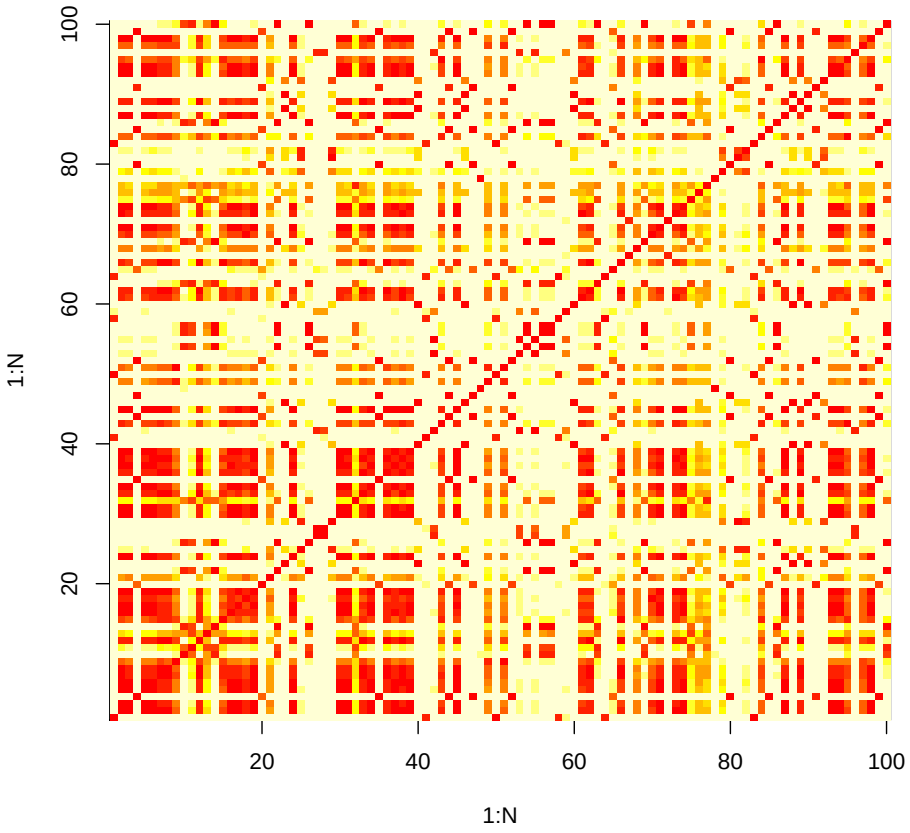

Figure 6.4. Plot of the heatmap of the clustering probabilities. Higher propensity to cluster together is indicated by deeper red color.

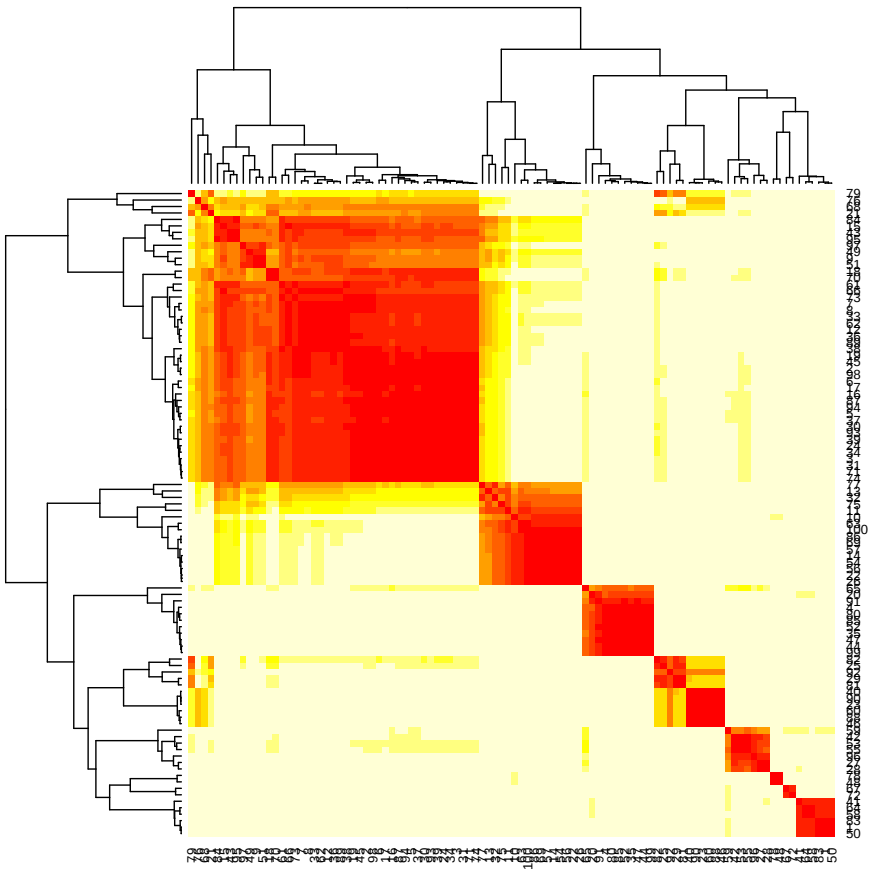

Figure 6.5. Plot of the heatmap of the clustering probabilities. Higher propensity to cluster together is indicated by deeper red color.

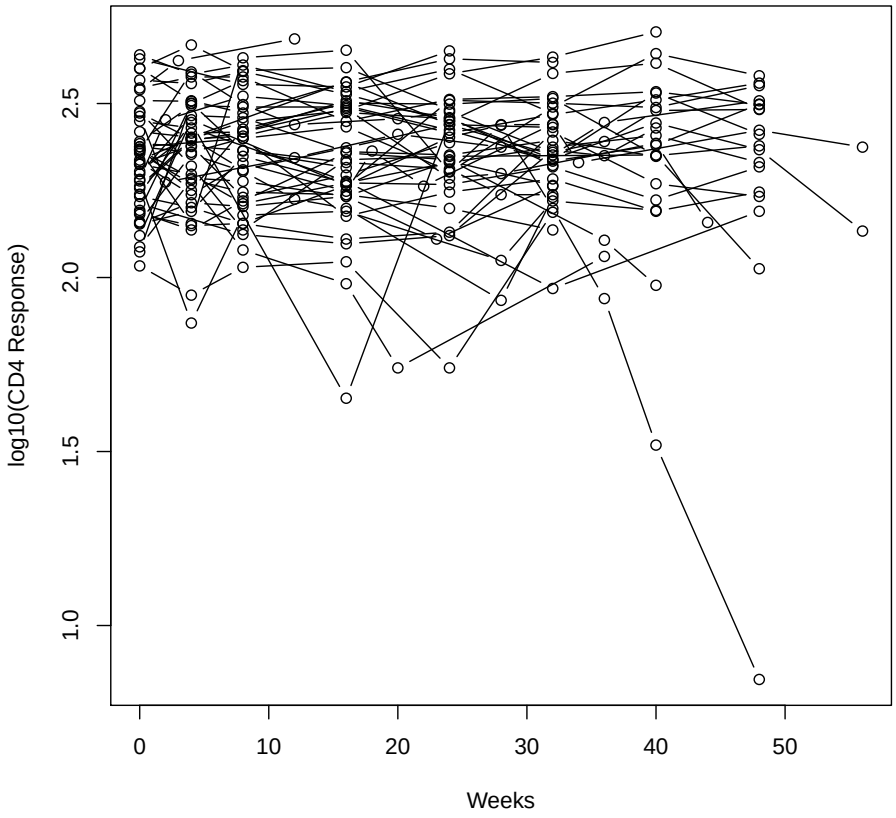

Figure 6.6. Trajectory plot of the first cluster.

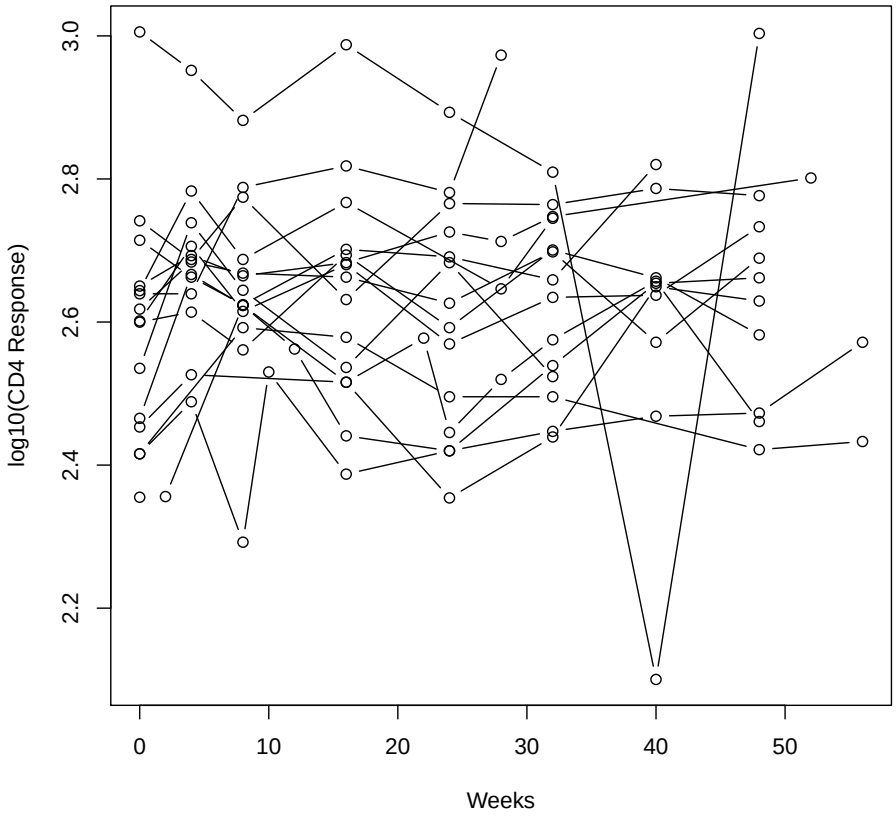

Figure 6.7. Trajectory plot of the second cluster.

Table 6.1 gives the results of $\hat{\boldsymbol{\theta}}$s for the two selected individuals from the first and the second cluster groups.

Table 6.1. Results of parameter estimation

$\hat{\boldsymbol{\theta}}_1$	$\hat{\boldsymbol{\theta}}_2$
41.14	72.72
-0.75	-0.88
-9.26	-13.25
0.43	-0.89
3.99	3.40
29.46	31.61

From figure 6.1 and 6.2 we can see that the model can detect the first joinpoint stronger than the second. Figure 6.6 and Figure 6.7 are the trajectory plots for the first and second cluster group. We can see that our model is able to correctly cluster the individuals with the same parameters and detect the most two obvious joinpoints.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

In this dissertation, we have first proposed a longitudinal model for multiple biomarkers, each with multiple joinpoints. Dirichlet process (DP) priors are used to model the distribution of the individual random effects. A simulation study of this model showed that the underlying trace cluster structure is barely revealed and the estimated parameters are not close to the true values. However, when the jointpoins are fixed but not estimated, the close estimation of parameters can be obtained.

Next, we combined the longitudinal part with a dropout part with multiple dropout causes. A simulation study of this new model showed that the parameter estimation and cluster structure are inaccurate.

Since there were difficulties in estimating the joinpoints by the above models, we have tried a reparametrized model for the longitudinal trajectory. The model has been explored with only one biomarker. Dirichlet process (DP) priors are also used to model the distribution of the individual random effects. In a simulation study, this model was able to correctly cluster the individuals. Howerver, we can not get accurate estimates of all the parameters. If the jointpoins are fixed but not estimated, the remaining parameters can be accurately estimated.

Finally, we combined the longitudinal part with the dropout part for this model. We can also correctly cluster the individuals but can't get accurate estimates for all the parameters. We can get the intercept and slope of the trajectory correctly estimated.

However we can not get the changes of the slopes and joinpoints accurately estimated. If we fix the joinpoints we can get close estimations for the longitudinal models we have proposed. The estimation of joinpoints is the main difficulty in this dissertation and none of the models we tried was able to successfully address this issue. New methods for the parametrization of joint model for the longitudinal trajectory could be helpful in the future. Due to complexity of the computation of the models, we were not able to run the simulation code for very long time. Improved algorithms will also be helpful.

APPENDIX

A BRIEF REVIEW OF DIRICHLET PROCESS MIXTURE MODELS

The basic model applies to data $y_1, \ldots, y_n$ which can be regarded as part of an infinite exchangeable sequence, or equivalently, as being independently drawn from some unknown distribution. The y_i maybe multivariate, with components that maybe real-valued or categorical. The distribution from which the y_i are drawn is modeled as a mixture of distributions of the form $F(\theta)$, with the mixing distribution over θ being G. The prior for this mixing distribution is assumed to be a Dirichlet process, with concentration parameter α and base distribution G_0. This gives the following model:

$$y_i|\theta_i \sim F(\theta_i),$$

$$\theta_i|G \sim G,$$

$$G \sim \mathrm{DP}(G_0, \alpha).$$

Since realizations of the Dirichlet process are discrete with probability one, these models can be viewed as countably infinite mixtures. This is also apparent when we integrate over G to obtain a representation of the prior distribution of the θ_i in terms of successive conditional distributions of the following form:

$$\theta_i|\theta_1, \ldots, \theta_{i-1} \sim \frac{1}{i-1+\alpha} \sum_{j=1}^{i-1} \delta_{\theta_j} + \frac{\alpha}{i-1+\alpha} G_0.$$

Here, δ_θ is the distribution concentrated at the single point θ. Note that the notation of the form $pR+(1-p)S$, where R and S are distributions, represents the distribution

that is the mixture of R and S, with proportions p and $1-p$, respectively.

Equivalent models can also be obtained by taking the limit as K goes to infinity of finite mixture models with K components having the following form:

$$y_i|c_i, \boldsymbol{\phi} \sim F(\phi_{c_i}),$$

$$c_i|\boldsymbol{p} \sim \text{Discrete}(p_1, \ldots, p_K),$$

$$\phi_c \sim G_0,$$

$$\boldsymbol{p} \sim \text{Dirichlet}(\alpha/K, \ldots, \alpha/K).$$

Here, c_i indicates which latent class is associated with observation y_i, with the numbering of the c_i being of no significance. For each class c, the parameter ϕ_c determines the distribution of observations from that class; the collection of all such ϕ_c is denoted by $\boldsymbol{\phi}$. The mixing proportions for the classes $\boldsymbol{p} = (p_1, \ldots, p_K)$, are given a symmetric Dirichlet prior, with concentration parameter written as α/K, so that it approaches zero as K goes to infinity.

By integrating over the mixing proportions $\boldsymbol{p}$, we can write the prior for the c_i as the product of conditional probabilities of the following form:

$$P(c_i = c|c_1, \ldots, c_{i-1}) = \frac{n_{i,c} + \alpha/K}{i - 1 + \alpha},$$

where $n_{i,c}$ is the number of c_j for $j < i$ that are equal to c.

If we let K go to infinity, the conditional probabilities in the above equation reach the following limits:

$$P(c_i = c|c_1, \ldots, c_{i-1}) \to \frac{n_{i,c}}{i - 1 + \alpha},$$

$$P(c_i \neq c_j \text{ for all } j < i|c_1, \ldots, c_{i-1}) \to \frac{\alpha}{i - 1 + \alpha}.$$

The following conditional distribution can be used for Gibbs sampling:

$$\theta_i|\theta_{-i}, y_i \sim \sum_{j \neq i} q_{i,j}\delta(\theta_j) + r_i H_i$$

Here, H_i is the posterior distribution for θ based on the prior G_0 and the single observation y_i, with likelihood $F(y_i, \theta)$. The values of the $q_{i,j}$ and of r_i are defined by

$$q_{i,j} = bF(y_i, \theta_j),$$

$$r_i = b\alpha \int F(y_i, \theta)\mathrm{d}G_0(\theta),$$

where b is such that $\sum_{j \neq i} q_{i,j} + r_i = 1$. For this Gibbs sampling method to be feasible, computing the integral defining r_i and sampling from H_i must be feasible operations. This will generally be so when G_0 is the conjugate prior for the likelihood given by F.

Neal (2000) present 2 algorithms for Markov chain Monte Karlo sampling from the posterior distribution when one has a DPM set up as before. Usually Algorithm 2 is used for conjugate cases and Algorithm 8 is used for conjugate or non-conjugate cases.

Algorithm 2 Let the state of the Markov chain consist of $\boldsymbol{c} = (c_1, \ldots, c_n)$ and $\boldsymbol{\phi} = (\phi_c : c \in \{c_1, \ldots, c_n\})$. Repeatedly sample as follows:

- For $i = 1, \ldots, n$: If the present value of c_i is associated with no other observation ($i.e., n_{-i,c_i} = 0$), remove ϕ_{c_i} from the state. Draw a new value for c_i from $c_i|c_{-i}, y_i, \boldsymbol{\phi}$ as defined by the equations for c_i. If the new c_i is not associated with any other observation, draw a value for ϕ_{c_i} from H_i and add it to the state.

- For all $c \in \{c_1, \ldots, c_n\}$: Draw a new value from $\phi_c | \text{all } y_i$ for which $c_i = c$, that is from the posterior distribution based on the prior G_0 and all the data points currently associated with latent class c.

Gibbs sampling for the c_i is based on the following conditional probabilities (with $\boldsymbol{\phi}$ here being the set of ϕ_c currently associated with at least one observation):

$$\text{If } c = c_j \text{ for some } j \neq i : P(c_i = c | c_{-i}, y_i, \boldsymbol{\phi}) = b \frac{n_{-i,c}}{n - 1 + \alpha} F(y_i, \phi_c)$$

$$P(c_i \neq c_j \text{ for all } j \neq i | c_{-i}, y_i, \boldsymbol{\phi}) = b \frac{\alpha}{n - 1 + \alpha} \int F(y_i, \phi) \mathrm{d}G_0(\phi)$$

Algorithm 8 Let the state of the Markov chain consist of $\boldsymbol{c} = (c_1, \ldots, c_n)$ and $\boldsymbol{\phi} = (\phi_c : c \in \{c_1, \ldots, c_n\})$. Repeatedly sample as follows:

- For $i = 1, \ldots, n$: Let k^- be the number of distinct c_j for $j \neq i$, and let $h = k^- + m$. Label these c_j with values in $\{1, \ldots, k^-\}$. If $c_i = c_j$ for some $j \neq i$, draw values independently from G_0 for those ϕ_c for which $k^- < c \leq h$. If $c_i \neq c_j$ for all $j \neq i$, let c_i have the label $k^- + 1$, and draw values independently from G_0 for those ϕ_c for which $k^- + 1 < c \leq h$. Draw a new value for c_i from $\{1, \ldots, h\}$ using the following probabilities:

$$P(c_i = c | c_{-i}, y_i, \phi_1, \ldots, \phi_h) = \begin{cases} b \frac{n_{-i,c}}{n-1+\alpha} F(y_i, \phi_c) & \text{for } 1 \leq c \leq k^- \\ b \frac{\alpha/m}{n-1+\alpha} F(y_i, \phi_c) & \text{for } k^- < c \leq h \end{cases}$$

  where $n_{-i,c}$ is the number of c_j for $j \neq i$ that are equal to c, and b is the appropriate normalizing constant. Change the state to contain only those ϕ_c that are now associated with one or more observations.

- For all $c \in \{c_1, \ldots, c_n\}$: Draw a new value from $\phi_c | y_i$ such that $c_i = c$, or perform some other update to ϕ_c that leaves this distribution invariant.

BIBLIOGRAPHY

Albert, P. S. and Shih, J. H. (2010). An approach for joint modeling multivariate longitudinal measurements and discrete time-to-event data. *The Annals of Applied Statistics*, 4(3):1517–1532.

Brown, E. R. and Ibrahim, J. G. (2003). A Bayesian semiparametric joint hierarchical model for longitudinal and survial data. *Biometrics*, 59:221–228.

Brown, E. R., Ibrahim, J. G., and DeGruttola, V. (2005). A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics*, 61:64–73.

Chi, Y.-Y. and Ibrahim, J. G. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics*, 62:432–445.

Faucett, C. L. and Thomas, D. C. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine*, 15:1663–1685.

Fieuws, S. and Verbeke, G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics*, 62:424–431.

Fieuws, S., Verbeke, G., and Molenberghs, G. (2007). Random-effects models for multivariate repeated measures. *Statistical Methods in Medical Research*, 16:387–397.

Ghosh, P., Ghosh, K., and Tiwari, R. C. (2009). Joint modeling of longitudinal data and informative dropout time in the presence of multiple changepoints. *Statistics in Medicine*, 30:611–626.

Hammer, S. M., Vaida, F., Bennett, K. K., Holohan, M. K., Sheiner, L., Eron, J. J., Wheat, L. J., Mitsuyasu, R. T., Gulick, R. M., Valentine, F. T., Aberg, J. A., Rogers, M. D., Karol, C. N., Saah, A. J., Lewis, R. H., Bessen, L. J., Brosgart, C., DeGruttola, V., and Mellors, J. W. (2002). Dual vs single protease inhibitor therapy following antiretroviral treatment failure: a randomized trial. *Journal of the American Medical Association*, 288(2):169–180.

Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics*, 1(4):465–480.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2004). Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine trials. *Statistica Sinica*, 14:863–883.

Kiuchi, A. S., Hartigan, J. A., Holford, T. R., Rubinstein, P., and Stevens, C. E. (1995). Change points in the series of T4 counts prior to AIDS. *Biometrics*, 51:236–248.

Liang, H. (2007). Segmental modeling of changing viral load to assess drug resistance in hiv infection. *Statistical Methods in Medical Research*, 17:365–373.

Lin, H., Turnbull, B. W., McCulloch, C. E., and Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association*, 97(457):53–65.

Martinez-Beneito, M. A., Garcia-Donato, G., and Salmeron, D. (2011). A Bayesian joinpoint regression model with an unknown number of break-points. *The Annals of Applied Statistics*, 5(3):2150–2168.

Neal, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9(2):249–265.

Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-Event Data With Applications in R.* Chapman and Hall/CRC Biostatistics Series.

Sinha, D., Chen, M.-H., and Ibrahim, J. G. (2001). *Bayesian Survival Analysis.* John Wiley and Sons.

Song, X., Davidian, M., and Tsiatis, A. A. (2002). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics*, 58:742–753.

Tsiatis, A. A. and Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika*, 88(2):447–458.

Tsiatis, A. A., DeGruttola, V., and Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error - applications to survival and CD4 counts in patients with aids. *Journal of the American Statistical Association*, 90(429):27–37.

Wang, Y. and Taylor, J. M. G. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association*, 96(455):895–905.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics*, 53(1):330–339.

Xu, J. and Zeger, S. L. (2001). The evaluation of multiple surrogate endpoints. *Biometrics*, 57:81–87.

VITA

Graduate College
University of Nevada, Las Vegas, USA

Libo Zhou

Degrees:

Bachelor of Arts in Actuarial Mathematics, 2002
Concordia University, Canada

Master of Science in Mathematics, 2006
University of Nevada, Las Vegas, USA

Thesis Title:

A Study of Joinpoint Models for Longitudinal Data

Thesis Examination Committee:

Chairperson, Kaushik Ghosh, Ph.D.
Committee Member, Malwane Ananda, Ph.D.
Committee Member, Chih-Hsiang Ho, Ph.D.
Committee Member, Hokwon Cho, Ph.D.
Graduate Faculty Representative, Guogen Shan, Ph.D.