

POLITECNICO DI MILANO
MATHEMATICAL ENGINEERING - STATISTICAL LEARNING

A.Y. 2021/2022



Bayesian registration of functional data

Project of: Matteo CACCAVALE
Federico CARRARA
Chiara ESPOSITO
Flavio LA MANNA
Giorgio MASATTI
Chiara SCRIMIERI

Supervisors: Prof. Raffaele ARGIENTO
Prof. Alessia PINI

Contents

1	Introduction	3
2	Dataset	4
3	Model	5
3.1	Hierarchical model	5
3.2	Assumptions	5
3.2.1	Assumptions on \mathbf{Y}_{ij}	5
3.2.2	Assumptions on a_{ij} and c_{ij}	5
3.2.3	Assumptions on ϵ_{ij}	6
3.2.4	Assumptions on β_j	6
3.2.5	Assumptions on Φ_{ij}	6
4	Gibbs-within-Metropolis algorithm	8
4.1	Full conditional distributions	8
4.1.1	Full conditional of a_{ij} and c_{ij}	9
4.1.2	Full conditional of a_0 and c_0	9
4.1.3	Full conditional of σ_c^2 and σ_a^2	11
4.1.4	Full conditional of σ_ϵ^2	12
4.1.5	Full conditional of β_j	12
4.1.6	Full conditional of μ_β	13
4.1.7	Full conditional of λ	15
4.1.8	Full conditional of S_l	16
4.1.9	Full conditional of Φ_{ij}	16
4.1.10	Full conditional of σ_Φ^2	16
4.2	Metropolis-Hastings step for S_l	17
4.2.1	Random walk	17
4.2.2	Introduction of the parameters $(P_l)_{ij}$	17
4.2.3	Adaptive Proposal method	18
4.3	Metropolis-Hastings step for Φ_{ij}	18
5	Fake data simulation	20
6	Results	21
6.1	Tecnical results	21
6.2	Medical results	21
7	Conclusions	23
8	Other plots	24
8.1	Results of the algorithm with the use of S_l	24
8.2	Traceplots	25
9	Bibliography	27
10	Github	27

1 Introduction

The functional data have become a key part of data analysis because of the growth of complexity of data.

Every functional data is observed on a discrete grid, since we don't have direct access to the function, but just to some points of it which have been sampled. The main idea is to take a very fine grid, to have continuous point, which indeed represents a function.

The project is based on a dataset containing the measurements of the movement angle generated by a knee during one hop over time performed by 100 patients.

Since that not all the measurements are taken at the same time the goal of our analysis is to find a warping function that aligns all the curves, which represent the measurements of the patients, modeling the amplitude of the features and estimating the velocity function.

2 Dataset

The dataset is characterized by three different groups of patients: healthy people, people who had physiotherapy and those who have surgery. Those who were treated all suffered a common type of injury. The purpose of the analysis is to compare the medical performance of the treatments. The number of the features is way bigger than that of the observations so then functional data are needed.

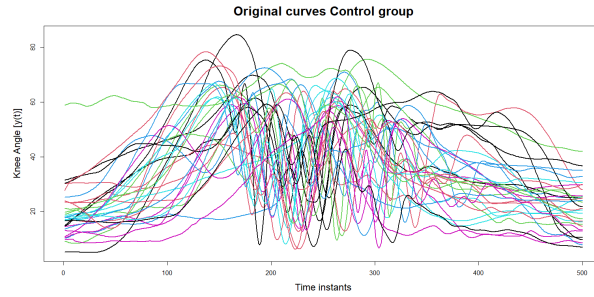


Figure 1: Healthy people

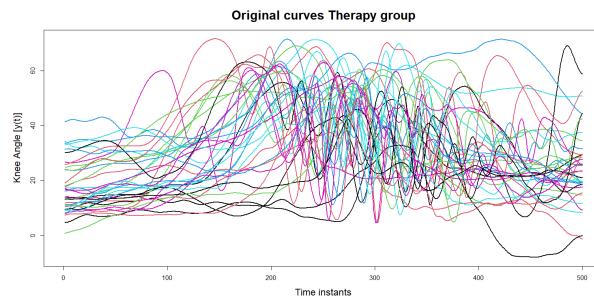


Figure 2: People who had physiotherapy

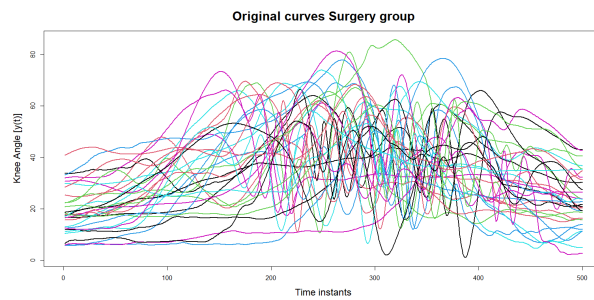


Figure 3: People who underwent surgery

3 Model

3.1 Hierarchical model

Let $y_{ij}(t)$ denote the observation of the ij -th curve related to the i -th patient and j -th group, j is equal to 1, 2 or 3 and i goes from 1 to n_j (where n_j is the number of patients in group j).

The shape function generating the individual curves is

$$m_{ij}(t; \theta_{ij}) = c_{ij} + a_{ij} \mathcal{B}'_m(t) \beta_j$$

And curve-specific time transformation function

$$\tau_{ij}(t) = \mathcal{B}'_\mu(t) \Phi_{ij}$$

And the composite function

$$m_{ij}(\tau_{ij}(t); \theta_{ij}) = c_{ij} + a_{ij} \mathcal{B}'_m(\tau_{ij}(t)) \beta_j$$

The observed value for each curve at time t is modeled as:

$$\begin{aligned} y_{ij}(t) &= m_{ij}(\tau_{ij}(t); \theta_{ij}) + \epsilon_{ij} \\ &= c_{ij} + a_{ij} \mathcal{B}'_m(\tau_{ij}(t)) \beta_j + \epsilon_{ij} \end{aligned}$$

Where m_{ij} and τ_{ij} , are modeled as linear combinations of a set of basis functions (\mathcal{B}'_m and \mathcal{B}'_μ) and a set of basis coefficients (β_j and Φ_{ij}).

3.2 Assumptions

3.2.1 Assumptions on Y_{ij}

As first thing we have assumed that our data are distributed as a normal as follow

$$Y_{ij} | a_{ij}, c_{ij}, \beta_j, \sigma_\epsilon^2, \Phi_{ij} \stackrel{ind}{\sim} \mathcal{N}(c_{ij} \mathbb{1}_{N_{ij}} + a_{ij} \mathcal{B}'_m(\tau_{ij}(t)) \beta_j, \sigma_\epsilon^2 \mathbf{I}_{N_{ij}})$$

Where $\mathbb{1}_{N_{ij}}$ is the N_{ij} -dimensional unitary vector and $\mathbf{I}_{N_{ij}}$ is the N_{ij} -dimensional identity matrix.

3.2.2 Assumptions on a_{ij} and c_{ij}

The parameter c_{ij} is the intercept of the common shape function (not constant but related to the i -th curve and to the j -th group) and a_{ij} is a parameter belonging to the model that characterizes the individual shape function (still i -th and j -th specific and therefore not constant).

$$a_{ij} \stackrel{iid}{\sim} \mathcal{N}(a_0, \sigma_a^2) \quad \text{and} \quad c_{ij} \stackrel{iid}{\sim} \mathcal{N}(c_0, \sigma_c^2)$$

Where we have assumed that

$$a_0 \sim \mathcal{N}(m_{a_0}, \sigma_{a_0}^2) \quad \text{and} \quad c_0 \sim \mathcal{N}(m_{c_0}, \sigma_{c_0}^2)$$

$$\sigma_a^2 \sim \text{Inv} - \text{Gamma}(a_a, b_a) \quad \text{and} \quad \sigma_c^2 \sim \text{Inv} - \text{Gamma}(a_c, b_c)$$

3.2.3 Assumptions on ϵ_{ij}

The parameter ϵ_{ij} are the error terms and are distributed as

$$\epsilon_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$$

Where

$$\sigma_\epsilon^2 \sim \text{Inv} - \text{Gamma}(a_\epsilon, b_\epsilon)$$

3.2.4 Assumptions on β_j

It is well known that, when using B-splines, the main issue is the choice of the positions and number of interior knots. We refer to an alternative approach that relies on penalized regression splines (Eilers and Marx, 1996).

We considered a first order random walk shrinkage prior on the shape coefficients β_j which depends on a smooth parameter λ .

$$(\beta_j)_k = (\beta_j)_{k-1} + \epsilon_k \quad \epsilon_k \sim \mathcal{N}(0, \lambda)$$

Where β_j is such that

$$\beta_j | \Sigma_\beta, \mu_\beta \stackrel{ind}{\sim} \mathcal{N}(\mu_\beta, \Sigma_\beta) \quad (j = 1, 2, 3)$$

$$\mu_\beta \sim \mathcal{N}(m_0, \Sigma_0)$$

$$\Sigma_\beta^{-1} = \frac{\Omega}{\lambda}$$

The matrix Ω is defined as

$$\Omega = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

And $\lambda \sim \text{Inv} - \text{Gamma}(a_\lambda, b_\lambda)$.

3.2.5 Assumptions on Φ_{ij}

Individual curves may exhibit features occurring at different times. Thus we use curve-specific random time-transformation functions $\tau_{ij}(t)$, parameterized with curve-specific parameters Φ_{ij} , which are Q -dimensional vectors.

The parameters Φ_{ij} must satisfy monotonicity constraint

$$0 = \Phi_1^{ij} < \Phi_2^{ij} < \cdots < \Phi_{Q-1}^{ij} < \Phi_Q^{ij} = 1$$

We decided to use two different specification of Φ_{ij}

1. Therefore, we assume

$$\Phi_l^{ij} = \frac{\sum_{k=1}^l S_k^{ij}}{T_{ij}} \quad \text{where} \quad T_{ij} = \sum_{l=1}^Q S_l^{ij}$$

$$S_1^{ij} = 0$$

$$S_l^{ij} \stackrel{iid}{\sim} \text{gamma}(\gamma_l, 1) \quad \text{where} \quad l = 2, \dots, Q$$

We decided to use Q equal to 5.

2. As model two we assume that the time-transformation function coefficients have multivariate normal distribution (Telesca et al., 2008)

$$\Phi_{ij} \stackrel{iid}{\sim} \mathcal{N}(\Upsilon, \Sigma_\Phi)$$

The matrix Σ_Φ is equal to $\sigma_\Phi^2 \mathbf{P}^{-1}$

- $\sigma_\Phi^2 \sim \text{Inv} - \text{Gamma}(a_\Phi, b_\Phi)$ is the smoothing parameter associated with the transformation functions $\tau_{ij}(t; \Phi_{ij})$;
- \mathbf{P} is the precision penalization matrix constructed as the matrix Ω .

Υ is the vector $(v_1, \dots, v_{h+r})'$ where its components are the identity coefficients if $f(t) = \mathcal{B}'(t)\Upsilon = t$. The coefficients Υ are found setting the derivative of $\mathcal{B}'\beta$ with respect to t to 1 for all $t \in [t_1, t_n]$. Given the order r and the interior knots $\{\kappa\}_h$, Υ can be obtained using the recursion

$$v_{q+1} = (v_{q+r} - v_{q+1})/(r+1) + v_q \quad q = 1, \dots, h+r-1$$

$$v_1 = t_1$$

The dimension of Φ_{ij} is equal to 5 also in this case.

A more detailed discussion about performances and feature of the two approach are in section 6.

4 Gibbs-within-Metropolis algorithm

4.1 Full conditional distributions

The parameters vector of the presented model is

$$\boldsymbol{\theta} = (\boldsymbol{c}', \boldsymbol{a}', \boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \boldsymbol{\beta}'_3, c_0, a_0, \sigma_\epsilon^2, \sigma_c^2, \sigma_a^2, \mu_\beta, \lambda, \boldsymbol{\Phi}')$$

Where \boldsymbol{c} and \boldsymbol{a} are vectors defined as

$$\boldsymbol{c} = \begin{pmatrix} c_{11} \\ \vdots \\ c_{1n_1} \\ c_{21} \\ \vdots \\ c_{2n_2} \\ c_{31} \\ \vdots \\ c_{3n_3} \end{pmatrix} \quad \text{and} \quad \boldsymbol{a} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n_1} \\ a_{21} \\ \vdots \\ a_{2n_2} \\ a_{31} \\ \vdots \\ a_{3n_3} \end{pmatrix}$$

We also define $\boldsymbol{\Phi}$ as

$$\boldsymbol{\Phi} = \begin{pmatrix} \Phi_{11} \\ \vdots \\ \Phi_{1n_1} \\ \Phi_{21} \\ \vdots \\ \Phi_{2n_2} \\ \Phi_{31} \\ \vdots \\ \Phi_{3n_3} \end{pmatrix}$$

The vector containing all the observations is \boldsymbol{Y} .

The posterior distribution of $\boldsymbol{\theta}$ is

$$\begin{aligned} \pi(\boldsymbol{\theta}|\boldsymbol{Y}) &= \pi(\boldsymbol{Y}|\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, c_0, a_0, \sigma_\epsilon^2, \sigma_c^2, \sigma_a^2, \mu_\beta, \lambda, \boldsymbol{\Phi}) \\ &\propto \pi(\boldsymbol{Y}|\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\Phi}, \sigma_\epsilon^2) \cdot \pi(\boldsymbol{c}, \boldsymbol{a}|c_0, a_0, \sigma_c^2, \sigma_a^2) \cdot \pi(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3|\mu_\beta, \lambda) \\ &\quad \cdot \pi(\boldsymbol{S}|\boldsymbol{\gamma}_l) \cdot \pi(\sigma_c^2|a_c, b_c) \cdot \pi(\sigma_a^2|a_a, b_a) \cdot \pi(a_0|\sigma_{a_0}^2) \cdot \pi(c_0|\sigma_{c_0}^2) \\ &\quad \cdot \pi(\sigma_\epsilon^2|a_\epsilon, b_\epsilon) \cdot \pi(\mu_\beta|m_0, \Sigma_0) \cdot \pi(\lambda|a_\lambda, b_\lambda) \end{aligned}$$

Where the vector \mathbf{S} is defined as

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} \\ \vdots \\ \mathbf{S}_{1n_1} \\ \mathbf{S}_{21} \\ \vdots \\ \mathbf{S}_{2n_2} \\ \mathbf{S}_{31} \\ \vdots \\ \mathbf{S}_{3n_3} \end{pmatrix}$$

The problem of calculating the posterior distribution has a high complexity since the number of parameters is high so we decided to use MCMC simulation algorithms to make posterior inference. In particular, we used the Gibbs-within-Metropolis algorithm. To be able to implement it we first calculated all the full conditionals.

4.1.1 Full conditional of a_{ij} and c_{ij}

$$\pi\left(\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} \middle| \boldsymbol{\theta}_{-(c_{ij}, a_{ij})}, \mathbf{Y}\right) \propto \pi\left(\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} \middle| a_0, c_0, \sigma_a^2, \sigma_c^2\right) \cdot \pi(\mathbf{Y} | \mathbf{c}, \mathbf{a}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\Phi}, \sigma_\epsilon^2)$$

Therefore

$$\begin{aligned} \pi\left(\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} \middle| \boldsymbol{\theta}_{-(c_{ij}, a_{ij})}, \mathbf{Y}\right) &\propto \exp\left\{-\frac{1}{2}\left[\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} - \begin{pmatrix} c_0 \\ a_0 \end{pmatrix}\right]' \Sigma_{c,a}^{-1} \left[\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} - \begin{pmatrix} c_0 \\ a_0 \end{pmatrix}\right]\right\} \\ &\quad \cdot \exp\left\{-\frac{1}{2}(\mathbf{Y}_{ij} - m_{ij})' \frac{1}{\sigma_\epsilon^2 \mathbf{I}_{N_{ij}}} (\mathbf{Y}_{ij} - m_{ij})\right\} \end{aligned}$$

So we have

$$\begin{aligned} \begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} \middle| \boldsymbol{\theta}_{-(c_{ij}, a_{ij})}, \mathbf{Y} &\sim \mathcal{N}(\mathbf{m}_l, \boldsymbol{\Sigma}_l) \\ (\mathbf{m}_l)_{ij} &= \boldsymbol{\Sigma}_l \left[\Sigma_{c,a}^{-1} \begin{pmatrix} c_0 \\ a_0 \end{pmatrix} + \frac{1}{\sigma_\epsilon^2} \mathbf{W}' \mathbf{Y}_{ij} \right] \\ (\boldsymbol{\Sigma}_l^{-1})_{ij} &= \Sigma_{c,a}^{-1} + \frac{1}{\sigma_\epsilon^2} \mathbf{W}' \mathbf{W} \\ \mathbf{W}_{ij} &= (\mathbb{1}_{N_{ij}} \quad \mathbf{B}'_m(\tau_{ij}(t)) \boldsymbol{\beta}_j) \end{aligned}$$

4.1.2 Full conditional of a_0 and c_0

In the following lines we have found the full conditional of a_0 and those of c_0 can be found in the same way.

$$\pi(a_0 | \mathbf{Y}, \boldsymbol{\theta}_{-a_0}) \propto \pi(\mathbf{c}_j, \mathbf{a}_j | c_0, a_0, \sigma_c^2, \sigma_a^2) \cdot \pi(a_0 | \sigma_{a_0}^2)$$

We decided to define a vector \mathbf{z} as follow

$$\mathbf{z} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \mathbf{c}_3 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} c_{11} \\ \vdots \\ c_{1n_1} \\ c_{21} \\ \vdots \\ c_{2n_2} \\ c_{31} \\ \vdots \\ c_{3n_3} \\ a_{11} \\ \vdots \\ a_{1n_1} \\ a_{21} \\ \vdots \\ a_{2n_2} \\ a_{31} \\ \vdots \\ a_{3n_3} \end{pmatrix} \sim \mathcal{N}(\mathbf{m}, \mathbf{V})$$

Where

$$\mathbf{m} = \begin{pmatrix} \mathbf{c}_0 \mathbb{1}_{n_1} \\ \mathbf{c}_0 \mathbb{1}_{n_2} \\ \mathbf{c}_0 \mathbb{1}_{n_3} \\ \mathbf{a}_0 \mathbb{1}_{n_1} \\ \mathbf{a}_0 \mathbb{1}_{n_2} \\ \mathbf{a}_0 \mathbb{1}_{n_3} \end{pmatrix} \quad \text{and} \quad \mathbf{V} = \begin{pmatrix} \sigma_c^2 I_{n_1} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_c^2 I_{n_2} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_c^2 I_{n_3} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \sigma_a^2 I_{n_1} & \cdots & 0 \\ 0 & 0 & 0 & 0 & \sigma_a^2 I_{n_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_a^2 I_{n_3} \end{pmatrix}$$

Where $\mathbb{1}_{n_j}$ are unitary n_j -dimensional vector.

Then

$$\pi(a_0 | \mathbf{Y}, \boldsymbol{\theta}_{-a_0}) \propto \exp \left\{ -\frac{1}{2} (\mathbf{z} - \mathbf{m})' \mathbf{V}^{-1} (\mathbf{z} - \mathbf{m}) \right\} \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(a_0 - m_{a_0})^2}{\sigma_{a_0}^2} \right\}$$

The terms that just rely on \mathbf{z} can be discarded thanks to the fact that \mathbf{z} does not depend on a_0 .

$$\begin{aligned} \pi(a_0 | \mathbf{Y}, \boldsymbol{\theta}_{-a_0}) &\propto \exp \left\{ (\mathbf{z}' \mathbf{V}^{-1} \mathbf{z}) - \frac{1}{2} (\mathbf{m}' \mathbf{V}^{-1} \mathbf{m}) \right\} \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(a_0 - m_{a_0})^2}{\sigma_{a_0}^2} \right\} \\ &\propto \exp \left\{ \mathbf{a}' (\sigma_a^2 \mathbf{I}_N)^{-1} \mathbf{m}_a - \frac{1}{2} \mathbf{m}_a' (\sigma_a^2 \mathbf{I}_N) \right\} \cdot \exp \left\{ -\frac{1}{2} \cdot \frac{(a_0 - m_{a_0})^2}{\sigma_{a_0}^2} \right\} \end{aligned}$$

Where

$$\mathbf{m}_a = a_0 \cdot \mathbb{1}_N$$

And

$$N = \sum_{j=1}^3 n_j$$

$$\begin{aligned}
\pi(a_0|\mathbf{Y}, \boldsymbol{\theta}_{-a_0}) &\propto \exp\left\{-\frac{1}{2}\underbrace{\left(\frac{1}{\sigma_a^2}N + \frac{1}{\sigma_{a_0}^2}\right)}_W a_0^2 - 2\underbrace{\left(\frac{1}{\sigma_a^2}\sum_{i=1}^N a_{ij} + \frac{m_{a_0}}{\sigma_{a_0}^2}\right)}_V a_0\right\} \\
&= \exp\left\{-\frac{1}{2}\left[W\left(a_0 - \frac{V}{2W}\right)^2\right]\right\}
\end{aligned}$$

Therefore

$$\begin{aligned}
a_0|\mathbf{Y}, \boldsymbol{\theta}_{-a_0} &\sim \mathcal{N}(m_{a_0}^*, \sigma_{a_0}^{*2}) \\
m_{a_0}^* &= \frac{V}{2W} = \frac{1}{2}\left(\frac{1}{\sigma_a^2}N + \frac{1}{\sigma_{a_0}^2}\right)^{-1} \cdot 2\left(\frac{1}{\sigma_a^2}\sum_{j=1}^3\sum_{i=1}^{n_j} a_{ij} + \frac{m_{a_0}}{\sigma_{a_0}^2}\right) \\
\sigma_{a_0}^{*2} &= W^{-1} = \left(\frac{1}{\sigma_a^2}N + \frac{1}{\sigma_{a_0}^2}\right)^{-1}
\end{aligned}$$

The full conditional distribution of c_0 can be derived in a similar way:

$$\begin{aligned}
c_0|\mathbf{Y}, \boldsymbol{\theta}_{-c_0} &\sim \mathcal{N}(m_{c_0}^*, \sigma_{c_0}^{*2}) \\
m_{c_0}^* &= \frac{V}{2W} = \frac{1}{2}\left(\frac{1}{\sigma_c^2}N + \frac{1}{\sigma_{c_0}^2}\right)^{-1} \cdot 2\left(\frac{1}{\sigma_c^2}\sum_{j=1}^3\sum_{i=1}^{n_j} c_{ij} + \frac{m_{c_0}}{\sigma_{c_0}^2}\right) \\
\sigma_{c_0}^{*2} &= W^{-1} = \left(\frac{1}{\sigma_c^2}N + \frac{1}{\sigma_{c_0}^2}\right)^{-1}
\end{aligned}$$

4.1.3 Full conditional of σ_c^2 and σ_a^2

$$\begin{aligned}
\pi(\sigma_c^2|\mathbf{Y}, \boldsymbol{\theta}_{-\sigma_c^2}) &\propto \pi(\mathbf{c}, \mathbf{a}|c_0, a_0, \sigma_c^2, \sigma_a^2) \cdot \pi\left(\frac{1}{\sigma_c^2} \middle| a_c, b_c\right) \\
&\propto \prod_{j=1}^3 \prod_{i=1}^{n_j} \frac{1}{\sqrt{2\pi\sigma_c^2}} \cdot \exp\left\{-\frac{1}{2}(c_{ij} - c_0)' \frac{1}{\sigma_c^2}(c_{ij} - c_0)\right\} \cdot \left(\frac{1}{\sigma_c^2}\right)^{a_c-1} \cdot \exp\left\{-\frac{b_c}{\sigma_c^2}\right\} \\
&\propto \prod_{j=1}^3 \prod_{i=1}^{n_j} \frac{1}{\sqrt{2\pi\sigma_c^2}} \cdot \exp\left\{-\frac{1}{2\sigma_c^2}(c_{ij} - c_0)^2\right\} \cdot \left(\frac{1}{\sigma_c^2}\right)^{a_c-1} \cdot \exp\left\{-\frac{b_c}{\sigma_c^2}\right\} \\
&= \exp\left\{-\frac{1}{\sigma_c^2}\left[b_c + \frac{1}{2}\sum_{j=1}^3\sum_{i=1}^{n_j}(c_{ij} - c_0)^2\right]\right\} \cdot \left(\frac{1}{\sigma_c^2}\right)^{a_c-1+\frac{1}{2}\sum_{j=1}^3 n_j}
\end{aligned}$$

So we obtain

$$\sigma_c^2|\boldsymbol{\theta}_{-\sigma_c^2}, \mathbf{Y} \sim \text{Inv-Gamma}(a_c^*, b_c^*)$$

Where

$$\begin{aligned}
a_c^* &= a_c + \frac{1}{2}\sum_{j=1}^3 n_j = a_c + \frac{N}{2} \\
b_c^* &= b_c + \frac{1}{2}\sum_{j=1}^3\sum_{i=1}^{n_j}(c_{ij} - c_0)^2
\end{aligned}$$

Similarly to the full conditional of σ_c^2 that of σ_a^2 is

$$\sigma_a^2 | \boldsymbol{\theta}_{-\sigma_a^2}, \mathbf{Y} \sim \text{Inv} - \text{Gamma}(a_a^*, b_a^*)$$

$$a_a^* = a_a + \frac{N}{2}$$

$$b_a^* = b_a + \frac{1}{2} \sum_{j=1}^3 \sum_{i=1}^{n_j} (a_{ij} - a_0)^2$$

4.1.4 Full conditional of σ_ϵ^2

$$\begin{aligned} \pi(\sigma_\epsilon^2 | \mathbf{Y}, \boldsymbol{\theta}_{-\sigma_\epsilon^2}) &\propto \pi(\mathbf{Y} | \boldsymbol{\theta}_{-\sigma_\epsilon^2}) \cdot \pi\left(\frac{1}{\sigma_\epsilon^2} | a_\epsilon, b_\epsilon\right) \\ &\propto \prod_{j=1}^3 \prod_{i=1}^{n_j} \frac{1}{\sqrt{\text{Det}(\sigma_\epsilon^2 I_{N_{ij}})^{-1}}} \cdot \exp\left\{-\frac{1}{2}(\mathbf{Y}_{ij} - m_{ij})'(\sigma_\epsilon^2 I_{N_{ij}})^{-1}(\mathbf{Y}_{ij} - m_{ij})\right\} \left(\frac{1}{\sigma_\epsilon^2}\right)^{a_\epsilon-1} \cdot \exp\left\{-\frac{b_\epsilon}{\sigma_\epsilon^2}\right\} \\ &\propto \left(\frac{1}{\sigma_\epsilon^2}\right)^{a_\epsilon + \frac{1}{2}(\sum_{j=1}^3 \sum_{i=1}^{n_j} N_{ij})-1} \cdot \exp\left\{-\frac{1}{\sigma_\epsilon^2}\left[b_\epsilon + \frac{1}{2} \sum_{j=1}^3 \sum_{i=1}^{n_j} (\mathbf{Y}_{ij} - m_{ij})'(\mathbf{Y}_{ij} - m_{ij})\right]\right\} \end{aligned}$$

So we obtain

$$\sigma_\epsilon^2 | \boldsymbol{\theta}_{\sigma_\epsilon^2}, \mathbf{Y} \sim \text{Inv} - \text{Gamma}(a_\epsilon^*, b_\epsilon^*)$$

Where

$$\begin{aligned} a_\epsilon^* &= a_\epsilon + \frac{\sum_{j=1}^3 \sum_{i=1}^{n_j} N_{ij}}{2} \\ b_\epsilon^* &= b_\epsilon + \frac{1}{2} \sum_{j=1}^3 \sum_{i=1}^{n_j} (\mathbf{Y}_{ij} - m_{ij})'(\mathbf{Y}_{ij} - m_{ij}) \end{aligned}$$

4.1.5 Full conditional of β_j

Recalling the assumptions on β_j and \mathbf{Y}

$$\beta_j \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$$

$$\begin{aligned} Y_{ij} | c_{ij}, a_{ij}, \beta_j, \Phi_{ij}, \sigma_\epsilon^2 &\sim \mathcal{N}\left(\begin{pmatrix} c_{ij} + a_{ij} \cdot \mathbf{B}'_m(\tau_{ij}(t_1)) \cdot \beta_j \\ c_{ij} + a_{ij} \cdot \mathbf{B}'_m(\tau_{ij}(t_2)) \cdot \beta_j \\ \vdots \end{pmatrix}, \sigma_\epsilon^2 \cdot I_{N_{ij}}\right) \\ &\sim \mathcal{N}(c_{ij} \mathbb{1}_{N_{ij}} + a_{ij} \mathbf{B}'_m(\tau_{ij}(\mathbf{t})) \cdot \beta_j, \sigma_\epsilon^2 I_{N_{ij}}) \end{aligned}$$

Let's introduce some notation that will help us in the calculations

$$\mathbf{C}_j = [(c_{1j} \mathbb{1}_{N_{1j}})', \dots, (c_{n_j j} \mathbb{1}_{N_{n_j j}})']'$$

$$\mathbf{M}_j = [(a_{1j} \mathbf{B}'_m(\tau_{1j}(\mathbf{t})))', \dots, (a_{n_j j} \mathbf{B}'_m(\tau_{n_j j}(\mathbf{t})))']'$$

These things imply

$$\mathbf{Y}_j | \mathbf{c}_j, \mathbf{a}_j, \beta_j, \Phi_j, \sigma_\epsilon^2 \sim \mathcal{N}(\mathbf{C}_j + \mathbf{M}_j \beta_j, \sigma_\epsilon^2 I_{\sum_{i=1}^{n_j} N_{ij}})$$

Therefore

$$\begin{aligned}
\pi(\beta_j | Y_j, \theta_{-\beta_j}) &\propto \exp \left\{ -\frac{1}{2} \left[Y_j - (C_j + M_j \beta_j) \right]' (\sigma_\epsilon^2 I_{\sum_{i=1}^{n_j} N_{ij}})^{-1} [Y_j - (C_j + M_j \beta_j)] \right\} \\
&\quad \cdot \exp \left\{ -\frac{1}{2} (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) \right\} \\
&\propto \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[Y_j - (C_j + M_j \beta_j) \right]' \left[Y_j - (C_j + M_j \beta_j) \right] \right\} \cdot \exp \left\{ -\frac{1}{2} (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) \right\} \\
&\propto \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[-Y_j' M_j \beta_j + C_j' M_j \beta_j - \beta_j' M_j' C_j + \beta_j' M_j' M_j \beta_j \right] - \frac{1}{2} \beta_j' \frac{\Omega}{\lambda} \beta_j + \mu_\beta' \frac{\Omega}{\lambda} \beta_j + \mu_\beta' \frac{\Omega}{\lambda} \mu_\beta \right\} \\
&= \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[-2Y_j' M_j \beta_j + 2C_j' M_j \beta_j + \beta_j' M_j' M_j \beta_j \right] + \frac{1}{2} \beta_j' \frac{\Omega}{\lambda} \beta_j + \mu_\beta' \frac{\Omega}{\lambda} \beta_j \right\} \\
&= \exp \left\{ -\frac{1}{2} \beta_j' \left(\frac{M_j' M_j}{\sigma_\epsilon^2} + \frac{\Omega}{\lambda} \right) \beta_j + \beta_j' \left[\frac{1}{\sigma_\epsilon^2} M_j' (Y_j - C_j) + \frac{\Omega}{\lambda} \mu_\beta \right] \right\}
\end{aligned}$$

Let

$$\begin{aligned}
V_{\beta_j} &= \left(\frac{M_j' M_j}{\sigma_\epsilon^2} + \frac{\Omega}{\lambda} \right)^{-1} \\
m_{\beta_j} &= \left[\frac{1}{\sigma_\epsilon^2} M_j' (Y_j - C_j) + \frac{\Omega}{\lambda} \mu_\beta \right]
\end{aligned}$$

So this implies

$$\pi(\beta_j | Y_j, \theta_{-\beta_j}) \propto \exp \left\{ -\frac{1}{2} \beta_j' V_{\beta_j}^{-1} \beta_j + \beta_j' m_{\beta_j} \right\}$$

With some calculations, like those done for μ_β , it can be showed that

$$\beta_j | Y_j, \theta_{-\beta_j} \sim \mathcal{N}(m_{\beta_j}, V_{\beta_j})$$

4.1.6 Full conditional of μ_β

Indicating k as the dimension of β_j and μ_β we proceeded as follow

$$\begin{aligned}
\pi(\mu_\beta | m_o, \Sigma_0, \lambda, \beta) &\propto \pi(\beta | \mu_\beta, \lambda) \cdot \pi(\mu_\beta | m_o, \Sigma_0) \\
&= \prod_{j=1}^3 \frac{1}{\sqrt{(2\pi)^k \text{Det}[(\frac{\Omega}{\lambda})^{-1}]}} \cdot \exp \left\{ -\frac{1}{2} \left[(\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) \right] \right\} \\
&\quad \cdot \frac{1}{\sqrt{(2\pi)^k \text{Det}[(\Sigma_0)^{-1}]}} \cdot \exp \left\{ -\frac{1}{2} (\mu_\beta - m_o)' \Sigma_0^{-1} (\mu_\beta - m_o) \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \sum_{j=1}^3 (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) \right\} \cdot \exp \left\{ -\frac{1}{2} (\mu_\beta - m_o)' \Sigma_0^{-1} (\mu_\beta - m_o) \right\} \\
&= \exp \left\{ -\frac{1}{2} \sum_{j=1}^3 (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) + (\mu_\beta - m_o)' \Sigma_0^{-1} (\mu_\beta - m_o) \right\}
\end{aligned}$$

- Analyze as first the term $-\frac{1}{2} \sum_{j=1}^3 (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta)$

$$\begin{aligned}
-\frac{1}{2} \sum_{j=1}^3 (\beta_j - \mu_\beta)' \frac{\Omega}{\lambda} (\beta_j - \mu_\beta) &= -\frac{1}{2} \sum_{j=1}^3 \beta_j' \frac{\Omega}{\lambda} \beta_j - \mu_\beta' \frac{\Omega}{\lambda} \beta_j - \beta_j' \frac{\Omega}{\lambda} \mu_\beta + \mu_\beta' \frac{\Omega}{\lambda} \mu_\beta \\
&= -\frac{1}{2} \underbrace{\sum_{j=1}^3 \beta_j' \frac{\Omega}{\lambda} \beta_j}_{\text{independent of } \mu_\beta} - \frac{1}{2} \sum_{j=1}^3 \left[-\mu_\beta' \frac{\Omega}{\lambda} \beta_j - \underbrace{(\mu_\beta' \frac{\Omega}{\lambda} \beta_j)'}_{\Omega = \Omega'} \right] - \frac{1}{2} \sum_{j=1}^3 \mu_\beta' \frac{\Omega}{\lambda} \mu_\beta \\
&= \sum_{j=1}^3 \mu_\beta' \frac{\Omega}{\lambda} \beta_j - \frac{1}{2} \cdot 3 \mu_\beta' \frac{\Omega}{\lambda} \mu_\beta
\end{aligned}$$

- Now analyze the term $-\frac{1}{2} (\mu_\beta - m_0)' \Sigma_0^{-1} (\mu_\beta - m_0)$

$$\begin{aligned}
-\frac{1}{2} (\mu_\beta - m_0)' \Sigma_0^{-1} (\mu_\beta - m_0) &= -\frac{1}{2} \mu_\beta' \Sigma_0^{-1} (\mu_\beta - m_0) + \frac{1}{2} \mu_\beta' \Sigma_0^{-1} m_0 + \underbrace{\frac{1}{2} m_0' \Sigma_0^{-1} \mu_\beta}_{\frac{1}{2} (\mu_\beta' \Sigma_0^{-1} m_0)'} - \frac{1}{2} m_0' \Sigma_0^{-1} m_0 \\
&= -\frac{1}{2} \mu_\beta' \Sigma_0^{-1} \mu_\beta + \mu_\beta' \Sigma_0^{-1} m_0
\end{aligned}$$

All this implies that

$$\begin{aligned}
\pi(\mu_\beta | m_0, \Sigma_0, \lambda, \beta) &\propto \exp \left\{ -\frac{1}{2} \cdot 3 \mu_\beta' \frac{\Omega}{\lambda} \mu_\beta + \sum_{j=1}^3 \mu_\beta' \frac{\Omega}{\lambda} \beta_j - \frac{1}{2} \mu_\beta' \Sigma_0^{-1} \mu_\beta + \mu_\beta' \Sigma_0^{-1} m_0 \right\} \\
&= \exp \left\{ -\frac{1}{2} \mu_\beta' \left(3 \frac{\Omega}{\lambda} + \Sigma_0^{-1} \right) \mu_\beta + \mu_\beta' \left(\frac{\Omega}{\lambda} \sum_{j=1}^3 \beta_j + \Sigma_0^{-1} m_0 \right) \right\}
\end{aligned}$$

Some notation

$$\begin{aligned}
\Sigma_n &= \left(3 \frac{\Omega}{\lambda} + \Sigma_0^{-1} \right)^{-1} \rightarrow \Sigma_n^{-1} = \left(3 \frac{\Omega}{\lambda} + \Sigma_0^{-1} \right) \\
m_n &= \frac{\Omega}{\lambda} \sum_{j=1}^3 \beta_j + \Sigma_0^{-1} m_0
\end{aligned}$$

This implies

$$\begin{aligned}
\pi(\boldsymbol{\mu}_\beta | m_0, \Sigma_0, \lambda, \boldsymbol{\beta}_j) &\propto \exp\left\{-\frac{1}{2}\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_\beta + \boldsymbol{\mu}'_\beta \mathbf{m}_n\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_\beta + \frac{1}{2}\boldsymbol{\mu}'_\beta \mathbf{m}_n + \frac{1}{2}\boldsymbol{\mu}'_\beta \mathbf{m}_n\right\} \\
&= \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_\beta) + \frac{1}{2}\boldsymbol{\mu}'_\beta \mathbf{m}_n + \frac{1}{2}\mathbf{m}'_n \boldsymbol{\mu}_\beta - \frac{1}{2}\mathbf{m}'_n \boldsymbol{\Sigma}_n \mathbf{m}_n + \underbrace{\frac{1}{2}\mathbf{m}'_n \boldsymbol{\Sigma}_n \mathbf{m}_n}_{\text{independent of } \boldsymbol{\mu}_\beta}\right\} \\
&\propto \exp\left\{-\frac{1}{3}\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_\beta + \frac{1}{2}\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\Sigma}_n \mathbf{m}_n + \frac{1}{2}\mathbf{m}'_n \boldsymbol{\mu}_\beta - \frac{1}{2}\mathbf{m}'_n \boldsymbol{\Sigma}_n \mathbf{m}_n\right\} \\
&= \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}'_\beta \boldsymbol{\Sigma}_n^{-1} - \mathbf{m}'_n)(\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n \mathbf{m}_n)\right\} \\
&= \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}_\beta - \mathbf{m}'_n \boldsymbol{\Sigma}_n) \boldsymbol{\Sigma}_n^{-1} (\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n \mathbf{m}_n)\right\} \\
&= \exp\left\{-\frac{1}{2}(\boldsymbol{\mu}'_\beta - \boldsymbol{\Sigma}_n \mathbf{m}_n) \boldsymbol{\Sigma}_n^{-1} (\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n \mathbf{m}_n)\right\}
\end{aligned}$$

Let's take $\boldsymbol{\mu}_n = \boldsymbol{\Sigma}_n \mathbf{m}_n$.

This implies

$$(\boldsymbol{\mu}_\beta | m_0, \Sigma_0, \lambda, \boldsymbol{\beta}_j) \sim \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$$

Where

$$\begin{aligned}
\boldsymbol{\mu}_n &= \left(3\frac{\boldsymbol{\Omega}}{\lambda} + \boldsymbol{\Sigma}_0^{-1}\right) \left(\frac{\boldsymbol{\Omega}}{\lambda} \sum_{j=1}^3 \boldsymbol{\beta}_j + \boldsymbol{\Sigma}_0^{-1} m_0\right) \\
\boldsymbol{\Sigma}_n &= \left(3\frac{\boldsymbol{\Omega}}{\lambda} + \boldsymbol{\Sigma}_0^{-1}\right)^{-1}
\end{aligned}$$

4.1.7 Full conditional of λ

As first thing we have supposed that the distribution of λ is an inverse gamma so

$$\frac{1}{\lambda} \sim \text{Gamma}(a_\lambda, b_\lambda)$$

So the computation of the full conditional is

$$\begin{aligned}
\pi\left(\frac{1}{\lambda} \middle| \boldsymbol{\theta}_{-\lambda}, \mathbf{Y}\right) &\propto \pi(\boldsymbol{\beta} | \lambda, \boldsymbol{\mu}_\beta) \cdot \pi\left(\frac{1}{\lambda} \middle| a_\lambda, b_\lambda\right) \\
\pi\left(\frac{1}{\lambda} \middle| \boldsymbol{\theta}_{-\lambda}, \mathbf{Y}\right) &\propto \prod_{j=1}^3 \frac{1}{\sqrt{\text{Det}\left(\frac{\boldsymbol{\Omega}}{\lambda}\right)^{-1}}} \cdot \exp\left\{\frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)' \frac{1}{\boldsymbol{\Omega}} \lambda (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\right\} \cdot \left(\frac{1}{\lambda}\right)^{a_\lambda-1} \cdot \exp\left\{-b_\lambda \frac{1}{\lambda}\right\} \\
&\propto \frac{1}{\lambda^{\frac{3k}{2}}} \left(\frac{1}{\lambda}\right)^{a_\lambda-1} \cdot \exp\left\{-\frac{1}{\lambda} \left[\frac{1}{2} \sum_{j=1}^3 ((\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)' \boldsymbol{\Omega} (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)) + b_\lambda\right]\right\} \\
&= \left(\frac{1}{\lambda}\right)^{a_\lambda + \frac{3k}{2} - 1} \cdot \exp\left\{-\frac{1}{\lambda} \left[\frac{1}{2} \sum_{j=1}^3 ((\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)' \boldsymbol{\Omega} (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)) + b_\lambda\right]\right\}
\end{aligned}$$

Therefore

$$\frac{1}{\lambda} \sim \text{Gamma}(a_\lambda^*, b_\lambda^*)$$

Where

$$a_\lambda^* = a_\lambda + \frac{3k}{2} \quad \text{and} \quad b_\lambda^* = b_\lambda + \frac{1}{2} \cdot \sum_{j=1}^3 (\beta_j - \mu_\beta)' \Omega (\beta_j - \mu_\beta)$$

4.1.8 Full conditional of S_l

$$\begin{aligned} \pi(S_l | \mathbf{Y}, \boldsymbol{\theta}_{-S_l}) &\propto \pi(\mathbf{Y} | \boldsymbol{\theta}_{-S_l}) \cdot \pi(S_l) \\ &\propto \left(\prod_{j=1}^3 \prod_{i=1}^{n_j} \frac{1}{\sqrt{\text{Det}(\sigma_\epsilon^2 I_{n_j})^{-1}}} \cdot \exp \left\{ -\frac{1}{2} (\mathbf{Y}_{ij} - \mathbf{m}_{ij})' (\sigma_\epsilon^2 I_{n_j})^{-1} (\mathbf{Y}_{ij} - \mathbf{m}_{ij}) \right\} \right) \\ &\quad \cdot \frac{1}{\Gamma(\gamma_l)} \cdot S_l^{\gamma_l - 1} \cdot \exp\{-S_l\} \end{aligned}$$

The full conditional of S_l is not in closed form so to we decided to implement a Metropolis-Hastings step to sample from it.

4.1.9 Full conditional of Φ_{ij}

$$\begin{aligned} \pi(\Phi_{ij} | \mathbf{Y}, \boldsymbol{\theta}_{-\Phi_{ij}}) &\propto \pi(\mathbf{Y} | \boldsymbol{\theta}_{-\Phi_{ij}}) \cdot \pi(\Phi_{ij}) \\ &\propto \left(\prod_{j=1}^3 \prod_{i=1}^{n_j} \frac{1}{\sqrt{\text{Det}(\sigma_\epsilon^2 I_{n_j})^{-1}}} \cdot \exp \left\{ -\frac{1}{2} (\mathbf{Y}_{ij} - \mathbf{m}_{ij})' (\sigma_\epsilon^2 I_{n_j})^{-1} (\mathbf{Y}_{ij} - \mathbf{m}_{ij}) \right\} \right) \\ &\quad \cdot \exp \left\{ -\frac{1}{2} (\Phi_{ij} - \boldsymbol{\Upsilon})' \frac{\mathbf{P}}{\sigma_\Phi^2} (\Phi_{ij} - \boldsymbol{\Upsilon}) \right\} \end{aligned}$$

The full conditional of Φ_{ij} is not in closed form so to we decided to implement a Metropolis-Hastings step to sample from it.

4.1.10 Full conditional of σ_Φ^2

$$\begin{aligned} \pi\left(\frac{1}{\sigma_\Phi^2} \middle| \mathbf{Y}, \boldsymbol{\theta}_{-\sigma_\Phi^2}\right) &\propto \pi\left(\Phi_{ij} \middle| \frac{1}{\sigma_\Phi^2}\right) \cdot \pi\left(\frac{1}{\sigma_\Phi^2}\right) \\ &\propto \prod_{j=1}^3 \prod_{i=1}^{n_j} \left(\frac{1}{\sqrt{\det(\mathbf{P}/\sigma_\Phi^2)^{-1}}} \right) \cdot \exp \left\{ -\frac{1}{2} (\Phi_{ij} - \boldsymbol{\Upsilon})' \frac{\mathbf{P}}{\sigma_\Phi^2} (\Phi_{ij} - \boldsymbol{\Upsilon}) \right\} \\ &\quad \cdot \left(\frac{1}{\sigma_\Phi^2} \right)^{a_\Phi - 1} \cdot \exp \left\{ -b_\Phi \frac{1}{\sigma_\Phi^2} \right\} \end{aligned}$$

With similar calculations like those for parameter λ we get

$$\sigma_\Phi^2 \sim \text{Inv-Gamma}(a_\Phi^*, b_\Phi^*)$$

Where

$$a_{\Phi}^* = a_{\Phi} + \frac{N \cdot Q}{2}$$

With $N = \sum_{j=1}^3 n_j$ is the total number of patients.

And

$$b_{\Phi}^* = b_{\Phi} + \frac{1}{2} \sum_{j=1}^3 \sum_{i=1}^{n_j} (\Phi_{ij} - \Upsilon)' P (\Phi_{ij} - \Upsilon)$$

4.2 Metropolis-Hastings step for S_l

4.2.1 Random walk

We decided to use a Random Walk Metropolis-Hastings (RWMH) to sample from the full conditionals of the parameters $(S_l)_{ij}$.

Random walk MH (Metropolis Hastings) is a particular case of MH in which the proposal density is $q(X, Y) = f(Y - X)$, where f is a density. For example when using the Gaussian distribution we have

$$Y = X + Z \quad Z \sim \mathcal{N}(0, \sigma^2)$$

A common choice for the distribution of the step is a symmetric one like in the example above. The acceptance rate in this case is computed as follow

$$\alpha(x, y) = \min\left(\frac{\pi(y)f(x-y)}{\pi(x)f(y-x)}, 1\right) = \min\left(\frac{\pi(y)}{\pi(x)}, 1\right)$$

Where π is the invariant distribution of the chain and the target.

4.2.2 Introduction of the parameters $(P_l)_{ij}$

The strategy has the following issues

- $(S_l)_{ij}$ are distributed as a *Gamma* that has a positive support while adopting a random walk strategy with a symmetric proposal could generate a negative step. This negative step added to the previous sample could lead to a negative sample.
- in the ratio $\frac{\pi(y)}{\pi(x)}$ the quantity $\pi(x)$ could be near to zero for some x and this may lead to numerical problems

A solution for the former problem is to introduce the parameters $(P_l)_{ij}$ that are the logarithm of $(S_l)_{ij}$ for $l = 2, \dots, 5$

$$(P_l)_{ij} = \log((S_l)_{ij})$$

Instead for $l = 1$ we don't need the update because $(S_1)_{ij}$ is always equal to zero.

So by applying the random variable transformation we can find the density function of $(P_l)_{ij}$ which is

$$f((P_l)_{ij}) = \frac{1}{\Gamma(\gamma_l)} \cdot \exp\left\{(P_l)_{ij} \cdot \gamma_l - \exp\{(P_l)_{ij}\}\right\}$$

When we first introduced our model we built the quantities m_{ij} which depends on the parameters $(S_l)_{ij}$. The values m_{ij} do not change with the introduction of the new parameters $(P_l)_{ij}$. So we can evaluate the likelihood of our data as done before.

Since that the RWMH is an iterative algorithm we calculate at each iteration h the densities ratio

$$r[h] = \frac{\prod_{j=1}^3 \left(\prod_{i=1}^{n_j} \frac{1}{\sqrt{\text{Det}(\sigma_\epsilon^2 I_{n_j})}} \exp\left\{-\frac{1}{2}(\mathbf{Y}_{ij} - \mathbf{m}_{ij})'(\sigma_\epsilon^2 I_{N_{ij}})^{-1}(\mathbf{Y}_{ij} - \mathbf{m}_{ij})\right\} \right) \cdot \prod_{l=2}^5 \exp\left\{(P_l)_{ij}^{(h+1)} \cdot \gamma_l - \exp\{(P_l)_{ij}^{(h+1)}\}\right\}}{\prod_{j=1}^3 \left(\prod_{i=1}^{n_j} \frac{1}{\sqrt{\text{Det}(\sigma_\epsilon^2 I_{n_j})}} \exp\left\{-\frac{1}{2}(\mathbf{Y}_{ij} - \mathbf{m}_{ij})'(\sigma_\epsilon^2 I_{N_{ij}})^{-1}(\mathbf{Y}_{ij} - \mathbf{m}_{ij})\right\} \right) \cdot \prod_{l=2}^5 \exp\left\{(P_l)_{ij}^{(h)} \cdot \gamma_l - \exp\{(P_l)_{ij}^{(h)}\}\right\}}$$

In order to solve the second problem we compute the logarithm for each $r[h]$. Then we calculated the acceptance probability for each density ratio defined as follow

$$\alpha[h] = \min\{\log(r[h]), 0\}$$

4.2.3 Adaptive Proposal method

The ratio between the number of iterations in which the new sample was accepted and the total of iterations resulting from the RWMH was too low therefore we considered it unsatisfactory for the analysis. We therefore decided to use the Adaptive Proposal method to try to get better results.

(Haario et al. 1999) The basic idea is to update the proposal distribution with the knowledge we have learned up to the current iteration about the target distribution.

Assume that the target distribution $\pi(\cdot)$ has dimension d . Suppose that at time h we have sampled at least H points $\{X_1, \dots, X_{h-H+1}, \dots, X_h\}$. The proposal distribution q_h for sampling proposal state Y is chosen as follows

$$q_h(\cdot | X_1, \dots, X_h) \sim \mathcal{N}(X_h, s_d^2 R_h)$$

Where R_h is the $d \times d$ - dimensional covariance matrix determined by the H points $\{X_{h-H+1}, \dots, X_h\}$ and s_d is the scaling factor which depends only on the dimension d .

The covariance R_h may be calculated by collecting the points X_{h-H+1}, \dots, X_h in a $H \times h$ matrix K , where each row represents one sampled point. Then

$$R_h = \frac{1}{H-1} \tilde{K}^T \tilde{K}$$

Where we denote \tilde{K} the centered matrix $\tilde{K} = K - \mathbb{E}[K]$.

The scaling parameter s_d has been chosen equal to $2.4/\sqrt{d}$.

We have not updated the covariance of the proposal distribution at each step of the process, but rather kept it fixed for U steps. Thus in this case the R_h is updated only at times that are multiples of U and kept fixed in between.

4.3 Metropolis-Hastings step for Φ_{ij}

Since that the results for the Metropolis-Hastings step for S_l , reported in section 8.1, were not satisfactory we decided to use the second assumption on Φ_{ij} (reported in paragraph 3.2.5).

Each component of Φ_{ij} is updated individually using uniform proposal on their constrained support with the range of the proposal calibrated at the burn-in of the MCMC runs. In particular we have tried different ranges for the uniform proposal to reach an acceptance rate between 35% and 75%.

$$(\text{proposed } (\Phi_l)_{ij})^{(h)} \sim \mathcal{U}(\text{lower bound}, \text{upper bound})$$

- *lower bound* $= (\Phi_l)_{ij}^{(h-1)} - \varepsilon \cdot \left[(\Phi_l)_{ij}^{(h-1)} - (\Phi_{l-1})_{ij}^{(h)} \right];$
- *upper bound* $= (\Phi_l)_{ij}^{(h-1)} + \varepsilon \cdot \left[(\Phi_{l+1})_{ij}^{(h-1)} - (\Phi_l)_{ij}^{(h-1)} \right];$
- h represents the current iteration;
- $l = 2, \dots, Q - 1$ (so in our case l goes from 2 to 4);
- $\varepsilon \in (0, 1)$ is the range-varying coefficient. After several attempts we decided to use ε equal to 0.05.

As in the case of the Metropolis-Hastings for S_l , also in this case we compute the acceptance rate and the densities ratio in a logarithmic scale.

5 Fake data simulation

Since the dataset has a very large size, we decided that in the preliminary phase of our project it was appropriate to fit the model on simulated data and then switch to the original ones once everything worked.

We generated three sets of smooth curves as transformations of common Beta density function. This helped in reducing of times steps needed and also served us to check if the algorithm worked well for the parameters that determine the shape of the functions, i.e. a_{ij} and c_{ij} , and β_j .

$$y_{ij}(t) = c_{ij} + a_{ij}\mathcal{B}'_m(\tau_{ij}(t))\beta_j + \epsilon_{ij}$$

Once the algorithm starts working we switched to the original data.

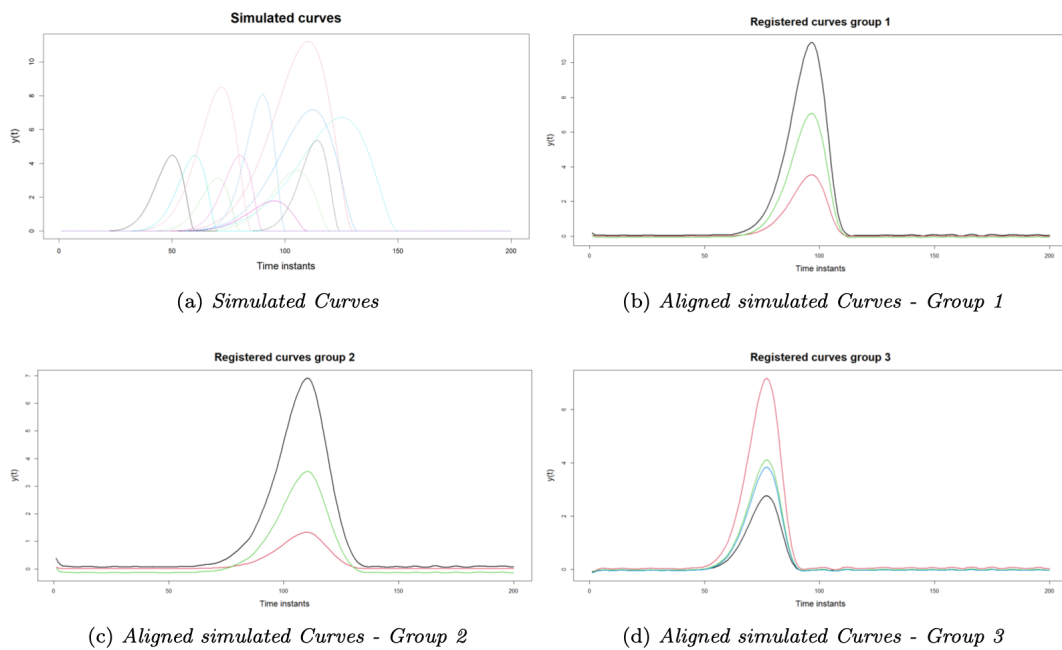


Figure 4: Fake simulated data

6 Results

For the MCMC simulations

- we decided to use a number of iterations between 30.000 and 50.000;
- burn-in equal to 20.000;
- we have thinned the samples by discarding one every ten.

6.1 Technical results

Given M draws from the posterior distribution of the model parameter θ , we can drive the posterior distribution of any parameter/function of interest. To register the observed curve, we use

$$y_{ij}^*(t) = y_{ij}(t) \circ \tau_{ij}^{-1}(t)$$

Given the posterior samples $\Phi_{ij}^{(h)}$

$$\tau_{ij}^{(h)}(t) = \tau(t; \Phi_{ij}^{(h)}) = \mathcal{B}'_{\tau}(t) \Phi_{ij}^{(h)}$$

Besides registering the observed curves, we are interested in the scaled shape functions, which we define as

$$(\mathbf{m}_s)_{ij}(t) = c_{ij} + a_{ij} \mathcal{B}'_m(t) \beta_j$$

that is, the shape function rescaled by the individual parameters a_{ij} and c_{ij} . Denoting as usual h has the current iteration

$$(\mathbf{m}_s)_{ij}^{(h)}(t) = c_{ij}^{(h)} + a_{ij}^{(h)} \mathcal{B}'_m(t) \beta_j^{(h)}$$

6.2 Medical results

The final considerations are made by comparing the results we got from the analysis of each group.

As we can see in the plots below, the group of healthy people is characterized by two main peaks and a very smooth shape. Instead having a lot of oscillations could indicate a weak knee that is unable to withstand the hop. Moreover we notice that people who had surgery reach the maximum angle of the knee later than healthy people. So we can suppose that people who had surgery flex the knee less quickly. Since the plots of patients who had physiotherapy oscillations more often than the others plots we can also suppose that they have a weaker knee probably because treatment with therapy is a longer healing process.

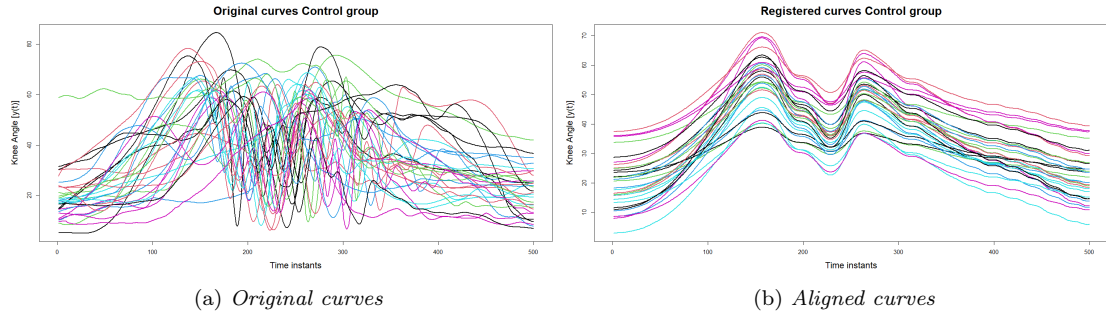


Figure 4: Healthy people

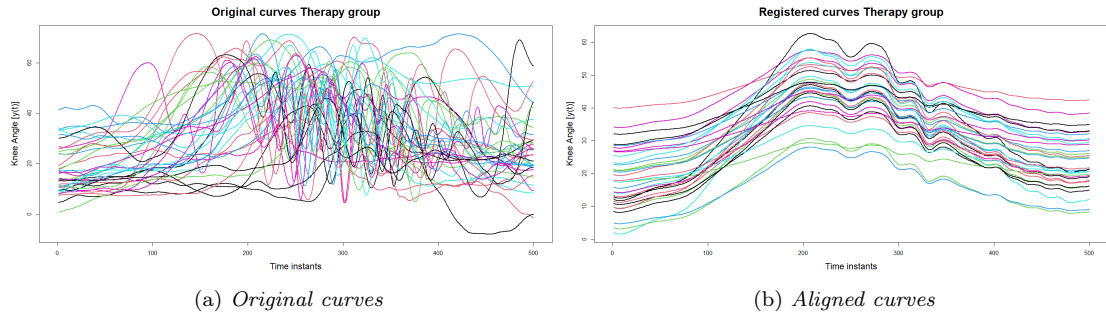


Figure 5: People who had physiotherapy

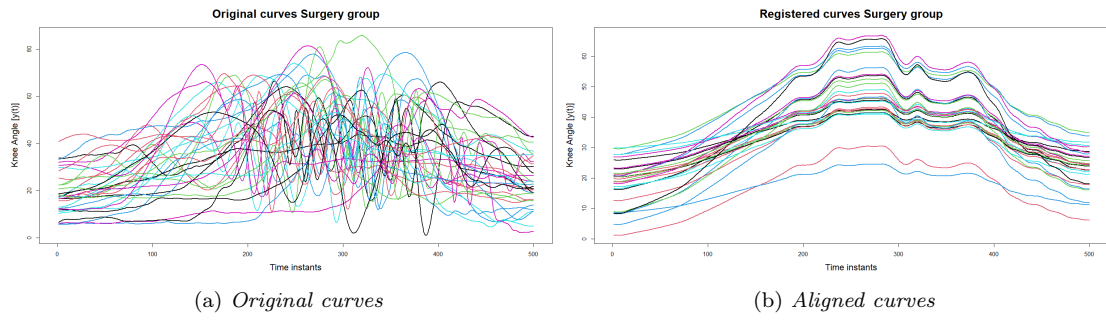


Figure 6: People who had surgery

7 Conclusions

Our goal was to build and test a hierarchical model that could align the curves of the three different groups introducing group-specific parameters. Regarding the shape of the functions in a group the algorithm works well however the smoothing can still be improved.

After the analyzes made, we can say that there is indeed a difference between the groups. In particular, the common shape function of the group of people who had physiotherapy shows more oscillations than that of the other groups analyzed. This difference could be linked to the fact that those who had physiotherapy have a weaker knee that cannot fully sustain the hop.

A possible future development could be to study the initial values assumed by the hyper-parameters of the algorithm. In particular, a grid-search approach could be adopted to identify an optimal combination.

Another possible improvement that could be done is to invest in the identifiability of the chain built for the algorithm. A model is said to be unidentifiable if there are two different combinations of parameters for which the model does not change. In our case the parameters that could cause the non-identifiability of the model are a_{ij} and β_j because they can be represented synthesized in a single parameter.

8 Other plots

8.1 Results of the algorithm with the use of S_t

The following plots are the results for the first model.

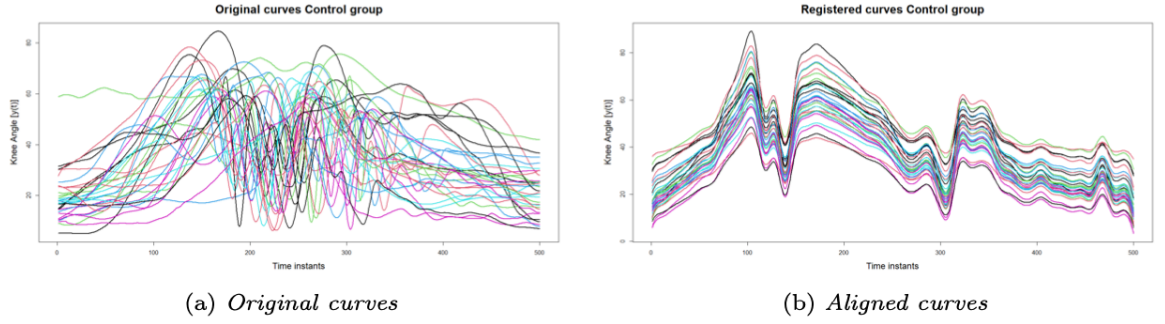


Figure 8: Healthy people

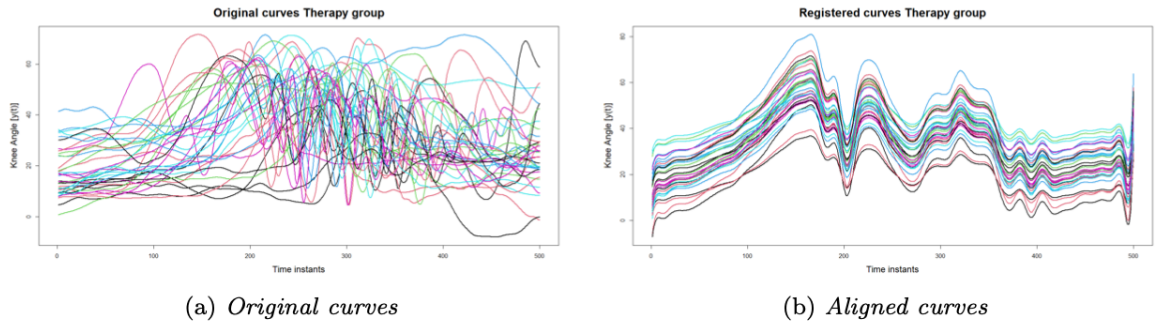


Figure 9: People who had physiotherapy

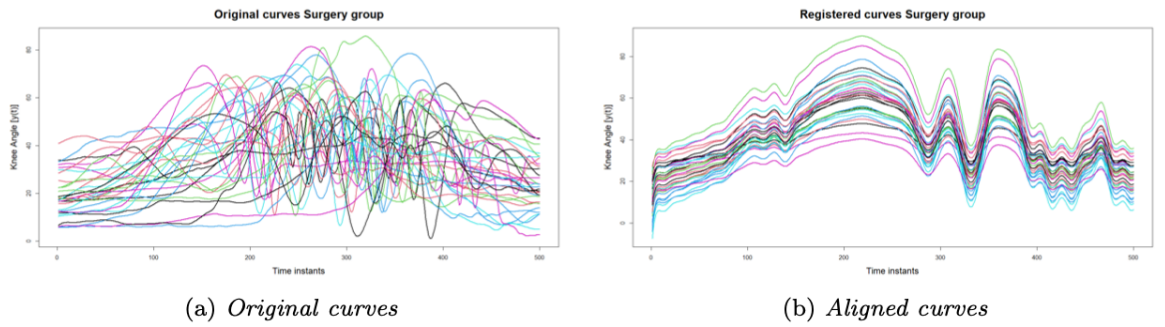
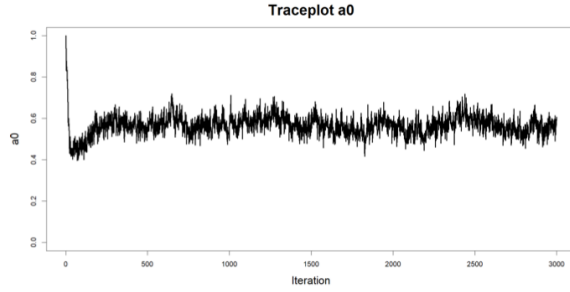
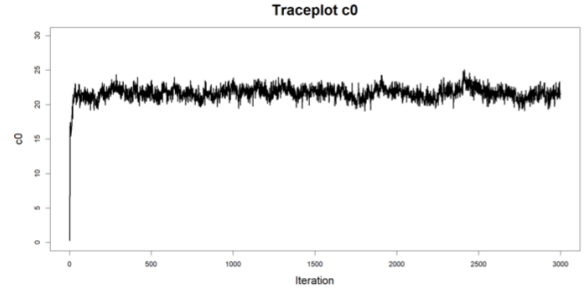


Figure 10: People who had surgery

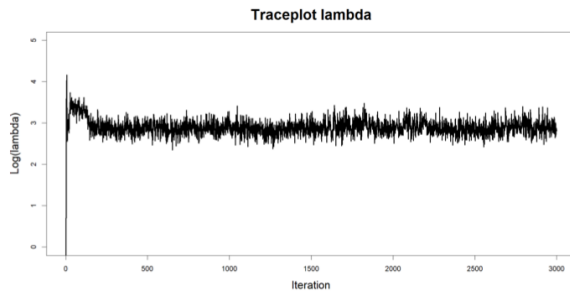
8.2 Traceplots



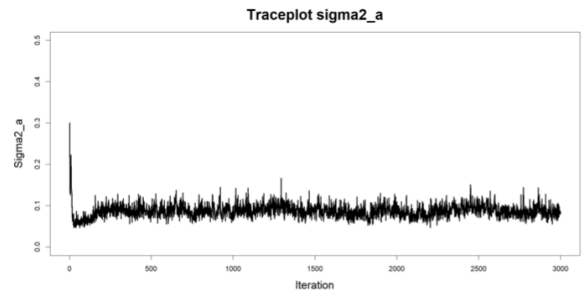
(a) Traceplots of a_0



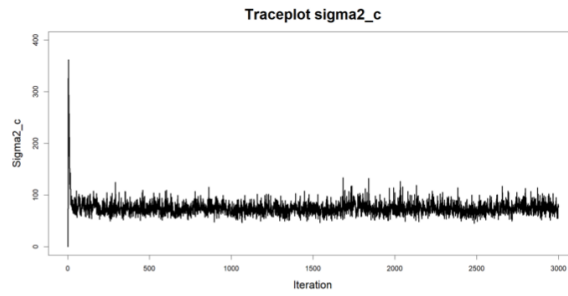
(b) Traceplots of c_0



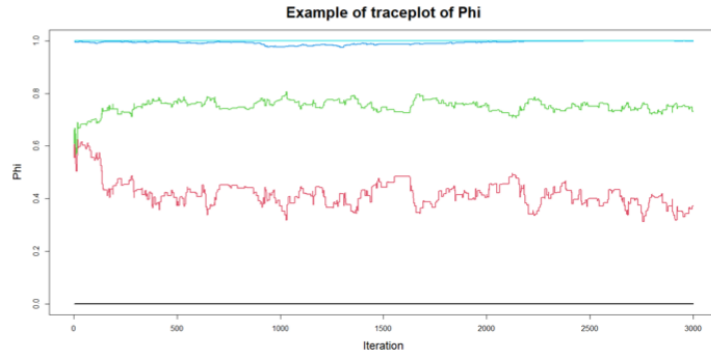
(c) Traceplots of λ



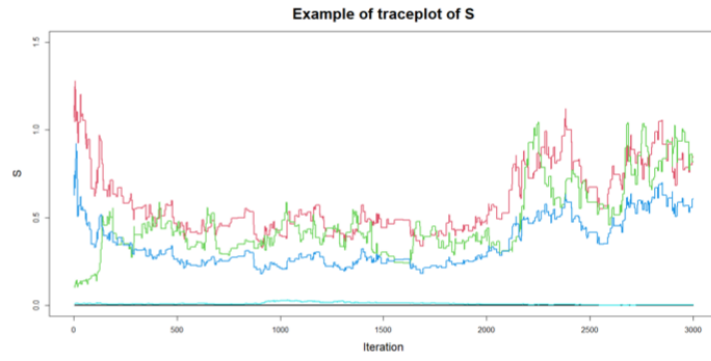
(d) Traceplots of σ_a^2



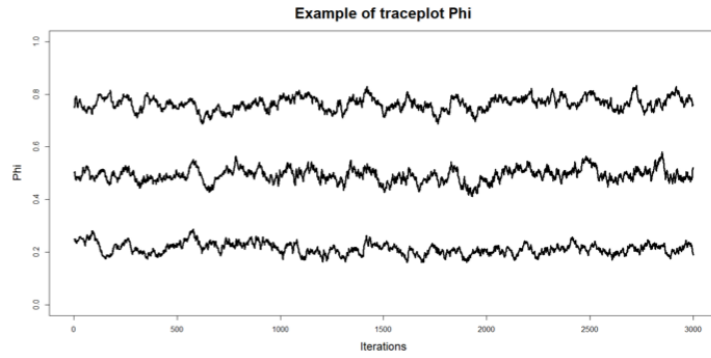
(e) Traceplots of σ_c^2



(a) Traceplots of Φ (first model)



(b) Traceplots of S_l



(c) Traceplots of Φ (second model)

9 Bibliography

Donatello Telesca, Lurdes Y. T. Inoue.

Bayesian Hierarchical Curve Registration.

Journal of the American Statistical Association (2008).

Heikki Haario, Eero Saksman, Johanna Tamminen.

An Adaptive Metropolis Algorithm.

Bernoulli Journal (2001).

Eilers, P., and Marx, B. (1996)

Flexible Smoothing Using B-Splines and Penalized Likelihood (with discussion).

Statistical Science, 11, 1200–1224.

10 Github

<https://github.com/Bayes-Proj2022/Bayesian-Statistics-Project>