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

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 in 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

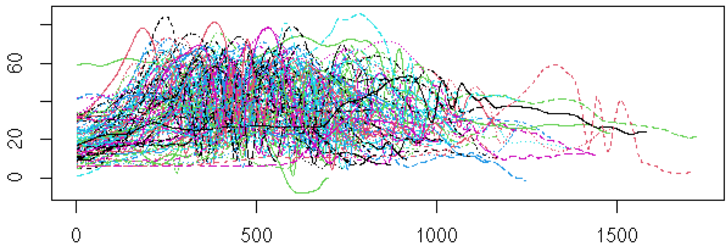

Figure 4: Merged dataset

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 common shape function generating the individual curves is

$$m_{ij}(t;\boldsymbol{\theta}_{ij}) = c_{ij} + a_{ij}\boldsymbol{\mathcal{B}}'_m(t)\boldsymbol{\beta}_j$$

And curve-specific time transformation function

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

And the composite function

$$m_{ij}(\tau_{ij}(t);\boldsymbol{\theta}_{ij}) = c_{ij} + a_{ij}\boldsymbol{\mathcal{B}}'_m(\tau_{ij}(t))\boldsymbol{\beta}_j$$

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

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

Where m_{ij} and τ_{ij}, are modeled as linear combinations of a set of basis fuctions ($\boldsymbol{\mathcal{B}}'_m$ and $\boldsymbol{\mathcal{B}}'_\mu$) and a set of basis coefficients (β_j and $\boldsymbol{\Phi}_{ij}$).

3.2 Assumptions

3.2.1 Assumptions on $\boldsymbol{Y}_{ij}$

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

$$\mathbf{Y}_{ij}|a_{ij}, c_{ij}, \boldsymbol{\beta}_j, \sigma^2_\epsilon, \boldsymbol{\Phi}_{ij} \overset{ind}{\sim} \mathcal{N}(c_{ij}\mathbb{1}_{N_{ij}} + a_{ij}\boldsymbol{\mathcal{B}}'_m(\tau_{ij}(t))\boldsymbol{\beta}_j, \sigma^2_\epsilon \boldsymbol{I}_{N_{ij}})$$

Where $\mathbb{1}_{N_{ij}}$ is the N_{ij}-dimensional unitary vector and $\boldsymbol{I}_{N_{ij}}$ 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 common shape function (still i-th and j-th specific and therefore not constant).

$$a_{ij} \overset{iid}{\sim} \mathcal{N}(a_0, \sigma^2_a) \quad and \quad c_{ij} \overset{iid}{\sim} \mathcal{N}(c_0, \sigma^2_c)$$

Where we have assumed that

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

$$\frac{1}{\sigma^2_a} \sim Gamma(a_a, b_a) \quad and \quad \frac{1}{\sigma^2_c} \sim Gamma(a_c, b_c)$$

3.2.3 Assumptions on ϵ_{ij}

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

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

3.2.4 Assumptions on $\boldsymbol{\beta}_j$

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 $\boldsymbol{\beta}_j$ which depends on a smooth parameter λ.

$$(\boldsymbol{\beta}_j)_k = (\boldsymbol{\beta}_j)_{k-1} + \epsilon_k \qquad \epsilon_k \sim \mathcal{N}(0, \lambda)$$

Where $\boldsymbol{\beta}_j$ is such that

$$\boldsymbol{\beta}_j | \boldsymbol{\Sigma}_\beta, \boldsymbol{\mu}_\beta \overset{ind}{\sim} \mathcal{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \quad (j = 1, 2, 3)$$
$$\boldsymbol{\mu}_\beta \sim \mathcal{N}(m_0, \Sigma_0)$$
$$\boldsymbol{\Sigma}_\beta^{-1} = \frac{\boldsymbol{\Omega}}{\lambda}$$

The matrix Ω is defined as

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

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

3.2.5 Assumptions on $\boldsymbol{\Phi}_{ij}$

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 $\boldsymbol{\Phi}_{ij}$, which are Q-dimensional vectors.
The parameters $\boldsymbol{\Phi}_{ij}$ 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 $\boldsymbol{\Phi}_{ij}$

1. Therefore, we assume

$$\Phi_l^{ij} = \frac{\sum_{k=1}^{l} S_k}{T} \quad where \quad T = \sum_{l=1}^{Q} S_l$$
$$S_1 = 0$$
$$S_l \sim gamma(\gamma_l, 1) \quad where \quad l = 2, \ldots, Q$$

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)

$$\boldsymbol{\Phi}_{ij} \overset{iid}{\sim} \mathcal{N}(\boldsymbol{\Upsilon}, \boldsymbol{\Sigma}_{\Phi})$$

The matrix $\boldsymbol{\Sigma}_{\Phi}$ is equal to $\sigma^2_{\Phi}\boldsymbol{P}$

- $\sigma^2_{\Phi} \sim Inv - Gamma(a_{\Phi}, b_{\Phi})$ is the smoothing parameter associated with the transformation functions $\tau_{ij}(t; \boldsymbol{\Phi}_{ij})$;
- $\boldsymbol{P}$ is the precision penalization matrix constructed as the matrix $\boldsymbol{\Omega}$.

$\boldsymbol{\Upsilon}$ is the vector $(\upsilon_1, \ldots, \upsilon_{h+r})'$ where its components are the identity coefficients if $f(t) = \boldsymbol{\mathcal{B}}'(t)\boldsymbol{\Upsilon} = t$. The coefficients $\boldsymbol{\Upsilon}$ are found setting the derivative of $\boldsymbol{\mathcal{B}}'\boldsymbol{\beta}$ with respect to t to 1 for all $t \in [t_1, t_n]$. Given the order r and the interior knots $\{\kappa\}_h$, $\boldsymbol{\Upsilon}$ can be obtained using the recursion

$$\upsilon_{q+1} = (\upsilon_{q+r} - \upsilon_{q+1})/(r+1) + \upsilon_q \qquad q = 1, \ldots, h+r-1$$

$$\upsilon_1 = t_1$$

The dimension of $\boldsymbol{\Phi}_{ij}$ is equal to 5 also in this case.

A more detailed discussion about performances and feature of the two approach are in paragraph INSERIRE NOME PARAGRAFO

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

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

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

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

The vector containing all the observations is $\boldsymbol{Y}$.
The posterior distribution od $\boldsymbol{\theta}$ is

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

Where the vector $\boldsymbol{S}$ is defined as

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

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}

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

Therefore

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

So we have

$$\begin{pmatrix} c_{ij} \\ a_{ij} \end{pmatrix} \middle| \boldsymbol{\theta}_{-(c_{ij},a_{ij})}, \boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{m}_l, \boldsymbol{\Sigma}_l)$$

$$(\boldsymbol{m}_l)_{ij} = \boldsymbol{\Sigma}_l \left[\Sigma_{c,a}^{-1} \begin{pmatrix} c_0 \\ a_0 \end{pmatrix} + \frac{1}{\sigma_\epsilon^2} \boldsymbol{W}' \boldsymbol{Y}_{ij} \right]$$

$$(\boldsymbol{\Sigma}_l^{-1})_{ij} = \Sigma_{c,a}^{-1} + \frac{1}{\sigma_\epsilon^2} \boldsymbol{W}' \boldsymbol{W}$$

$$\boldsymbol{W}_{ij} = (\mathbb{1}_{N_{ij}} \quad \boldsymbol{\mathcal{B}}'_m(\tau_{ij}(t))\boldsymbol{\beta}_j)$$

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.

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

We decide to define a vector $\boldsymbol{z}$ as follow

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

Where

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

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

Then

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

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

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

Where

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

And

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

$$\pi(a_0|\boldsymbol{Y},\boldsymbol{\theta}_{-a_0}) \propto exp\Bigg\{-\frac{1}{2}\underbrace{\left(\frac{1}{\sigma_a^2}N+\frac{1}{\sigma_{a_0}^2}\right)}_{W}a_0^2 - 2\underbrace{\left(\frac{1}{\sigma_a^2}\sum_{i=1}^{N}a_{ij}+\frac{m_{a_0}}{\sigma_{a_0}^2}\right)}_{V}a_0\Bigg\}$$

$$= exp\Bigg\{-\frac{1}{2}\Bigg[W\Big(a_0-\frac{V}{2W}\Big)^2\Bigg]\Bigg\}$$

Therefore

$$a_0|\boldsymbol{Y},\boldsymbol{\theta}_{-a_0} \sim \mathcal{N}(m^*_{a_0},\sigma^*_{a_0})$$

$$m^*_{a_0} = \frac{V}{2W} = \frac{1}{2}\Big(\frac{1}{\sigma_a^2}N+\frac{1}{\sigma_{a_0}^2}\Big)^{-1}\cdot 2\Big(\frac{1}{\sigma_a^2}\sum_{j=1}^{3}\sum_{i=1}^{n_j}a_{ij}+\frac{m_{a_0}}{\sigma_{a_0}^2}\Big)$$

$${\sigma^*_{a_0}}^2 = W^{-1} = \left(\frac{1}{\sigma_a^2}N+\frac{1}{\sigma_{a_0}^2}\right)^{-1}$$

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

$$c_0|\boldsymbol{Y},\boldsymbol{\theta}_{-c_0} \sim \mathcal{N}(m^*_{c_0},\sigma^*_{c_0})$$

$$m^*_{c_0} = \frac{V}{2W} = \frac{1}{2}\Big(\frac{1}{\sigma_c^2}N+\frac{1}{\sigma_{c_0}^2}\Big)^{-1}\cdot 2\Big(\frac{1}{\sigma_c^2}\sum_{j=1}^{3}\sum_{i=1}^{n_j}c_{ij}+\frac{m_{c_0}}{\sigma_{c_0}^2}\Big)$$

$${\sigma^*_{c_0}}^2 = W^{-1} = \left(\frac{1}{\sigma_c^2}N+\frac{1}{\sigma_{c_0}^2}\right)^{-1}$$

4.1.3 Full conditional of σ_c^2 and σ_a^2

$$\pi(\sigma_c^2|\boldsymbol{Y},\boldsymbol{\theta}_{-\sigma_c^2}) \propto \pi(\boldsymbol{c},\boldsymbol{a}|c_0,a_0,\sigma_c^2,\sigma_a^2)\cdot\pi\left(\frac{1}{\sigma_c^2}\middle|a_c,b_c\right)$$

$$\propto \prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\sqrt{2\pi\sigma_c^2}}\cdot exp\Bigg\{-\frac{1}{2}(c_{ij}-c_0)'\frac{1}{\sigma_c^2}(c_{ij}-c_0)\Bigg\}\cdot\left(\frac{1}{\sigma_c^2}\right)^{a_c-1}\cdot exp\Bigg\{-\frac{b_c}{\sigma_c^2}\Bigg\}$$

$$\propto \prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\sqrt{2\pi\sigma_c^2}}\cdot exp\Bigg\{-\frac{1}{2\sigma_c^2}(c_{ij}-c_0)^2\Bigg\}\cdot\left(\frac{1}{\sigma_c^2}\right)^{a_c-1}\cdot exp\Bigg\{-\frac{b_c}{\sigma_c^2}\Bigg\}$$

$$= exp\Bigg\{-\frac{1}{\sigma_c^2}\Bigg[b_c+\frac{1}{2}\sum_{j=1}^{3}\sum_{i=1}^{n_j}(c_{ij}-c_0)^2\Bigg]\Bigg\}\cdot\left(\frac{1}{\sigma_c^2}\right)^{a_c-1+\frac{1}{2}\sum_{j=1}^{3}n_j}$$

So we obtain

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

Where

$$a_c^* = a_c+\frac{1}{2}\sum_{j=1}^{3}n_j = a_c+\frac{N}{2}$$

$$b_c^* = b_c+\frac{1}{2}\sum_{j=1}^{3}\sum_{i=1}^{n_j}(c_{ij}-c_0)^2$$

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

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

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

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

4.1.4 Full conditional of σ_ϵ^2

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

So we obtain

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

Where

$$a_\epsilon^* = a_\epsilon + \frac{\sum_{j=1}^{3}\sum_{i=1}^{n_j} N_{ij}}{2}$$

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

4.1.5 Full conditional of $\boldsymbol{\beta}_j$

Recalling the assumptions on β_j and $\boldsymbol{Y}$

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

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

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

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

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

These things imply

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

Therefore

$$\begin{aligned}
\pi(\boldsymbol{\beta}_j|\boldsymbol{Y}_j, \theta_{-\beta_j}) &\propto exp\bigg\{-\frac{1}{2}\bigg[\boldsymbol{Y}_j - (\boldsymbol{C}_j + \boldsymbol{M}_j\boldsymbol{\beta}_j)\bigg]'(\sigma_\epsilon^2 I_{\sum_{i=1}^{n_j} N_{ij}})^{-1}[\boldsymbol{Y}_j - (\boldsymbol{C}_j + \boldsymbol{M}_j\boldsymbol{\beta}_j)\bigg]\bigg\} \\
&\cdot exp\{-\frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\} \\
&\propto exp\bigg\{-\frac{1}{2\sigma_\epsilon^2}\bigg[Y_j - (\boldsymbol{C}_j + \boldsymbol{M}_j\boldsymbol{\beta}_j)\bigg]'\bigg[Y_j - (\boldsymbol{C}_j + \boldsymbol{M}_j\boldsymbol{\beta}_j)\bigg]\bigg\} \cdot exp\bigg\{-\frac{1}{2}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\bigg\} \\
&\propto exp\bigg\{-\frac{1}{2\sigma_\epsilon^2}\bigg[-Y_j'\boldsymbol{M}_j\boldsymbol{\beta}_j + \boldsymbol{C}_j'\boldsymbol{M}_j\boldsymbol{\beta}_j - \boldsymbol{\beta}_j'\boldsymbol{M}_j'\boldsymbol{C}_j + \boldsymbol{\beta}_j'\boldsymbol{M}_j'\boldsymbol{M}_j\boldsymbol{\beta}_j\bigg] - \frac{1}{2}\boldsymbol{\beta}_j'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j + \boldsymbol{\mu}_\beta\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j + \boldsymbol{\mu}_\beta'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j\bigg\} \\
&= exp\bigg\{-\frac{1}{2\sigma_\epsilon^2}\bigg[-2Y_j'\boldsymbol{M}_j\boldsymbol{\beta}_j + 2\boldsymbol{C}_j'\boldsymbol{M}_j\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\boldsymbol{M}_j'\boldsymbol{M}_j\boldsymbol{\beta}_j\bigg] + \frac{1}{2}\boldsymbol{\beta}_j'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j + \boldsymbol{\mu}_\beta'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j\bigg\} \\
&= exp\bigg\{-\frac{1}{2}\boldsymbol{\beta}_j'\bigg(\frac{\boldsymbol{M}_j'\boldsymbol{M}_j}{\sigma_\epsilon^2} + \frac{\Omega}{\lambda}\bigg)\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\bigg[\frac{1}{\sigma_\epsilon^2}\boldsymbol{M}_j(\boldsymbol{Y}_j - \boldsymbol{C}_j) + \frac{\Omega}{\lambda}\boldsymbol{\mu}_\beta\bigg]\bigg\}
\end{aligned}$$

Let

$$\boldsymbol{V}_{\beta_j} = \bigg(\frac{\boldsymbol{M}_j'\boldsymbol{M}_j}{\sigma_\epsilon^2} + \frac{\boldsymbol{\Omega}}{\lambda}\bigg)^{-1}$$

$$\boldsymbol{m}_{\beta_j} = \bigg[\frac{1}{\sigma_\epsilon^2}\boldsymbol{M}_j'(\boldsymbol{Y}_j - \boldsymbol{C}_j) + \frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\mu}_\beta\bigg]$$

So this imply

$$\pi(\boldsymbol{\beta}_j|\boldsymbol{Y}_j, \boldsymbol{\theta}_{-\beta_j}) \propto exp\bigg\{-\frac{1}{2}\boldsymbol{\beta}_j'\boldsymbol{V}_{\beta_j}^{-1}\boldsymbol{\beta}_j + \boldsymbol{\beta}_j'\boldsymbol{m}_{\beta_j}\bigg\}$$

With some calculations, like those done for $\boldsymbol{\mu}_\beta$, it can be showed that

$$\boldsymbol{\beta}_j|\boldsymbol{Y}_j, \boldsymbol{\theta}_{-\beta_j} \sim \mathcal{N}(\boldsymbol{m}_{\boldsymbol{\beta}_j}, \boldsymbol{V}_{\beta_j})$$

4.1.6 Full conditional of $\boldsymbol{\mu}_\beta$

Indicating k as the dimension of $\boldsymbol{\beta}_j$ and $\boldsymbol{\mu}_\beta$ we proceeded as follow

$$\begin{aligned}
\pi(\boldsymbol{\mu}_\beta|m_o, \Sigma_0, \lambda, \boldsymbol{\beta}) &\propto \pi(\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \lambda) \cdot \pi(\boldsymbol{\mu}_\beta|m_0, \Sigma_0) \\
&= \prod_{j=1}^{3}\frac{1}{\sqrt{(2\pi)^k Det[(\frac{\boldsymbol{\Omega}}{\lambda})^{-1}]}} \cdot exp\bigg\{-\frac{1}{2}\bigg[(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\bigg]\bigg\} \cdot \\
&\cdot \frac{1}{\sqrt{(2\pi)^k Det[(\Sigma_0)^{-1}]}} \cdot exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta - m_0)'\Sigma_0^{-1}(\boldsymbol{\mu}_\beta - m_0)\bigg\} \\
&\propto exp\bigg\{-\frac{1}{2}\sum_{j=1}^{3}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)\bigg\} \cdot exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta - m_0)'\Sigma_0^{-1}(\boldsymbol{\mu}_\beta - m_0)\bigg\} \\
&= exp\bigg\{-\frac{1}{2}\sum_{j=1}^{3}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta) + (\boldsymbol{\mu}_\beta - m_0)'\Sigma_0^{-1}(\boldsymbol{\mu}_\beta - m_0)\bigg\}
\end{aligned}$$

- Analize as firt the term $-\frac{1}{2}\sum_{j=1}^{3}(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)'\frac{\boldsymbol{\Omega}}{\lambda}(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)$

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

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

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

All this implies that

$$
\begin{aligned}
\pi(\boldsymbol{\mu}_\beta|m_0,\Sigma_0,\lambda,\boldsymbol{\beta}) &\propto exp\Big\{-\frac{1}{2}\cdot 3\boldsymbol{\mu}_\beta'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\mu}_\beta+\sum_{j=1}^{3}\boldsymbol{\mu}_\beta'\frac{\boldsymbol{\Omega}}{\lambda}\boldsymbol{\beta}_j-\frac{1}{2}\boldsymbol{\mu}_\beta'\Sigma_0^{-1}\boldsymbol{\mu}_\beta+\boldsymbol{\mu}_\beta'\Sigma_0^{-1}m_0\Big\} \\
&= exp\Big\{-\frac{1}{2}\boldsymbol{\mu}_\beta'\Big(3\frac{\boldsymbol{\Omega}}{\lambda}+\Sigma_0^{-1}\Big)\boldsymbol{\mu}_\beta+\boldsymbol{\mu}_\beta'\Big(\frac{\boldsymbol{\Omega}}{\lambda}\sum_{j=1}^{3}\boldsymbol{\beta}_j+\Sigma_0^{-1}m_0\Big)\Big\}
\end{aligned}
$$

Some notation

$$
\boldsymbol{\Sigma}_n=(3\frac{\boldsymbol{\Omega}}{\lambda}+\Sigma_0^{-1})^{-1}\quad\rightarrow\quad\boldsymbol{\Sigma}_n^{-1}=\Big(3\frac{\boldsymbol{\Omega}}{\lambda}+\Sigma_0^{-1}\Big)
$$

$$
\boldsymbol{m}_n=\frac{\boldsymbol{\Omega}}{\lambda}\sum_{j=1}^{3}\boldsymbol{\beta}_j+\Sigma_0^{-1}m_0
$$

This implies

$$\pi(\boldsymbol{\mu}_\beta|m_0,\Sigma_0,\lambda,\boldsymbol{\beta}_j) \propto exp\bigg\{-\frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_\beta + \boldsymbol{\mu}_\beta'\boldsymbol{m}_n\bigg\}$$

$$= exp\bigg\{-\frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_\beta + \frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{m}_n + \frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{m}_n\bigg\}$$

$$= exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_\beta) + \frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{m}_n + \frac{1}{2}\boldsymbol{m}_n'\boldsymbol{\mu}_\beta - \frac{1}{2}\boldsymbol{m}_n'\boldsymbol{\Sigma}_n\boldsymbol{m}_n + \underbrace{\frac{1}{2}\boldsymbol{m}_n'\boldsymbol{\Sigma}_n\boldsymbol{m}_n}_{\text{independent of } \boldsymbol{\mu}_\beta}\bigg\}$$

$$\propto exp\bigg\{-\frac{1}{3}\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_\beta + \frac{1}{2}\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\Sigma}_n\boldsymbol{m}_n + \frac{1}{2}\boldsymbol{m}_n'\boldsymbol{\mu}_\beta - \frac{1}{2}\boldsymbol{m}_n'\boldsymbol{\Sigma}_n\boldsymbol{m}_n\bigg\}$$

$$= exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta'\boldsymbol{\Sigma}_n^{-1} - \boldsymbol{m}_n')(\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n\boldsymbol{m}_n)\bigg\}$$

$$= exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta - \boldsymbol{m}_n'\boldsymbol{\Sigma}_n)\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n\boldsymbol{m}_n)\bigg\}$$

$$= exp\bigg\{-\frac{1}{2}(\boldsymbol{\mu}_\beta' - \boldsymbol{\Sigma}_n\boldsymbol{m}_n)\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\mu}_\beta - \boldsymbol{\Sigma}_n\boldsymbol{m}_n)\bigg\}$$

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

This implies

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

Where

$$\boldsymbol{\mu}_n = \left(3\frac{\boldsymbol{\Omega}}{\lambda} + \Sigma_0^{-1}\right)\left(\frac{\boldsymbol{\Omega}}{\lambda}\sum_{j=1}^{3}\boldsymbol{\beta}_j + \Sigma_0^{-1}m_0\right)$$

$$\boldsymbol{\Sigma}_n = \left(3\frac{\boldsymbol{\Omega}}{\lambda} + \Sigma_0^{-1}\right)^{-1}$$

4.1.7 Full conditional of λ

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

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

So the computation of the full conditional is

$$\pi\left(\frac{1}{\lambda}\middle|\boldsymbol{\theta}_{-\lambda},\boldsymbol{Y}\right) \propto \pi(\boldsymbol{\beta}|\lambda,\mu_\beta)\cdot\pi\left(\frac{1}{\lambda}\middle|a_\lambda,b_\lambda\right)$$

$$\pi\left(\frac{1}{\lambda}|\boldsymbol{\theta}_{-\lambda},\boldsymbol{Y}\right) \propto \prod_{j=1}^{3}\frac{1}{\sqrt{Det(\frac{\boldsymbol{\Omega}}{\lambda})^{-1}}}\cdot exp\bigg\{\frac{1}{2}(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)'\frac{}{\Omega}\lambda(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)\bigg\}\cdot\left(\frac{1}{\lambda}\right)^{a_\lambda-1}\cdot exp\bigg\{-b_\lambda\frac{1}{\lambda}\bigg\}$$

$$\propto \frac{1}{\lambda^{\frac{3k}{2}}}\left(\frac{1}{\lambda}\right)^{a_\lambda-1}\cdot exp\bigg\{-\frac{1}{\lambda}\bigg[\frac{1}{2}\sum_{j=1}^{3}((\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)'\boldsymbol{\Omega}(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta))+b_\lambda\bigg]\bigg\}$$

$$= \left(\frac{1}{\lambda}\right)^{a_\lambda+\frac{3k}{2}-1}\cdot exp\bigg\{-\frac{1}{\lambda}\bigg[\frac{1}{2}\sum_{j=1}^{3}((\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta)'\boldsymbol{\Omega}(\boldsymbol{\beta}_j-\boldsymbol{\mu}_\beta))+b_\lambda\bigg]\bigg\}$$

Therefore

$$\frac{1}{\lambda} \sim Gamma(a^*_\lambda, b^*_\lambda)$$

Where

$$a^*_\lambda = a_\lambda + \frac{3k}{2} \quad and \quad b^*_\lambda = b_\lambda + \frac{1}{2} \cdot \sum_{j=1}^{3} (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)' \boldsymbol{\Omega} (\boldsymbol{\beta}_j - \boldsymbol{\mu}_\beta)$$

4.1.8 Full conditional of S_l

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

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 $\boldsymbol{\Phi}_{ij}$

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

The full conditional of $\boldsymbol{\Phi}_{ij}$ 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^2_Φ

SCRIVERE DA FOTO MATTEO

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

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

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 introduce 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

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

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 $\mathbf{\Phi}_{ij}$

Since that the results for the Metropolis-Hastigs step for S_l, reported in section METTERE NUMERO SEZIONE RISULTATI , were not satisfactory we decided to use the second assumption on $\mathbf{\Phi}_{ij}$ (reported in paragraph 3.2.5).

Each component of $\mathbf{\Phi}_{ij}$ 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%.

$$(proposed\ (\Phi_{ij})_l)^{(h)} \sim \mathcal{U}(lower\ bound, upper\ bound)$$

- $lower\ bound = (\Phi_{ij})_l^{(h-1)} - \varepsilon \cdot [(\Phi_{ij})_l^{(h-1)} - (\Phi_{ij})_{l-1}^{(h)}]$;
- $upper\ bound = (\Phi_{ij})_l^{(h-1)} + \varepsilon \cdot [(\Phi_{ij})_{l+1}^{(h-1)} - (\Phi_{ij})_l^{(h-1)}]$;
- h represent 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 tthe 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 $\boldsymbol{\beta}_j$.

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

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

(a) *Simulated Curves - Group 1*

(b) *Aligned simulated Curves - Group 1*

(c) *Simulated Curves - Group 2*

(d) *Aligned simulated Curves - Group 2*

(e) *Simulated Curves - Group 3*

(f) *Aligned simulated Curves - Group 3*

Figure 5: 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;
- we have thinned the samples by discarding one every ten.

6.1 Tecnical results

Given M draws from the posterior distribution of the model parameter $\boldsymbol{\theta}$, we can drive the posterior distribution of any parameter/function of interest. To register the observed curve, we use

$$y^*_{ij}(t) = y_{ij}(t) \circ \tau^{-1}_{ij}(t)$$

Given the posterior samples $\boldsymbol{\Phi}^{(h)}_{ij}$

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

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

$$(\boldsymbol{m}_s)_{ij}(t) = c_{ij} + a_{ij}\boldsymbol{\mathcal{B}}'_m(t)\boldsymbol{\beta}_j$$

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

$$(\boldsymbol{m}_s)^{(h)}_{ij}(t) = c^{(h)}_{ij} + a^{(h)}_{ij}\boldsymbol{\mathcal{B}}'_m(t)\boldsymbol{\beta}^{(h)}_j$$

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 swings 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 swings 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.

(a) *Original curves* (b) *Aligned curves*

Figure 6: Healthy people

(a) *Original curves* (b) *Aligned curves*

Figure 7: People who had physiotherapy

(a) *Original curves* (b) *Aligned curves*

Figure 8: People who had surgery

7 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.

8 Github

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