Bayesian variable selection approach in varying coefficient model: application in functional mapping

Marie Denis

in collaboration with B. Heuclin, F. Mortier, JM. Gion, S. Tisné

October 6, 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements N° 840383 and N° 773383

Outline

1. Introduction
2. Bayesian variable selection
3. Varying coefficient model
   - Estimation of $\beta_j(t)$
   - Selection of relevant variables X_j
   - VCGSS package
4. Bibliography

Objectives

Objectives during this lecture are:

1. to introduce Bayesian variable selection methods,
2. to present a Bayesian variable selection method for selecting variables with effects evolving over time,
3. to apply Bayesian variable selection methods in the genetic context.

Outline

1. Introduction
2. Bayesian variable selection
3. Varying coefficient model
   - Estimation of $\beta_j(t)$
   - Selection of relevant variables X_j
   - VCGSS package
4. Bibliography

Why use variable selection approaches ?

In various domains **high-dimensional data**:

- Genetic/Genomic: high-throughput sequencing,
- Ecophysiology/Ecology: high-throughput phenotyping,
- Economic: more and more variables,
- ....

resulting in **a high number** of variables p measured on **a limited number** of individuals $n \Rightarrow$ p$>$ n

For example in genetic:

- To identify the genomic regions involved in the variability of a phenotype Y

$\hookrightarrow$ $n = 100$ individuals genotyped with $p = 5000$ SNPs (variables, X)

How to select the relevant SNPs? How to select the best subset of variables?

Variable selection in linear model

A linear model is a statistical model assuming that the response variable Y may be written as a linear combination of variables X:

$$Y = \mu + X_1\beta_1 + \cdots + X_p\beta_p + \varepsilon, \quad \varepsilon \sim \mathcal{N}_n(0, \sigma^2 Id_n)$$

$$\Leftrightarrow Y = \mu + X\beta + \varepsilon, \quad \varepsilon \sim \mathcal{N}_n(0, \sigma^2 Id_n)$$

$$\Leftrightarrow Y \sim \mathcal{N}_n(\mu + X\beta, \sigma^2 Id_n)$$

with n the number of observations and p the number of variables.

Which variables are relevant ? $\Leftrightarrow$ *Which* $\beta_j \neq 0$ *?*

For example in genetic:

$$Y = \mu + SNP_1\beta_1 + \cdots + SNP_p\beta_p + \varepsilon$$

Which SNP is relevant $\Leftrightarrow$ *Which* $\beta_j \neq 0$ *?* $\Leftrightarrow$ *Does the position* j *affect the variability of phenotype* Y

Variable selection in linear model

Classical approaches:

- Student test with multiple testing and adjustment for multiplicity,
- Comparison of all models by using criteria (R^2, AIC, BIC, cross-validation, Fisher test),
- Backward stepwise selection, Forward stepwise selection

Problems: approaches not feasible or not optimal when $p > n$: need to use statistical methods allowing to **regularize** models

- ↬ Frequentist context: likelihood penalized approaches (Lasso, Ridge, Elastic-Net,...) (Hoerl and Kennard, 1970; Tibshirani, 1996)
- ↬ Bayesian context: shrinkage priors (spike-and-slab prior, Bayesian Lasso prior,...)

In the following we will focus on Bayesian approaches

Introduction to Bayesian linear model

In the Bayesian context **prior** distributions are placed on the parameters (here: μ, β, σ^2)

Bayesian linear model

$$
\begin{aligned}
Y|\mu, \beta, \sigma^2 &\sim \mathcal{N}_n(\mu + X\beta, \sigma^2 I_n) \\
\beta_j &\sim p_\beta(\beta_j), \quad j = 1, \ldots, p \\
\mu &\sim p_\mu(\mu) \\
\sigma^2 &\sim p_{\sigma^2}(\sigma^2)
\end{aligned}
$$

Usual prior distributions:

- σ^2: Inverse Gamma distribution
- β: Normal distribution (no selection)
- μ: Uniform distribution

Bayesian variable selection

In the Bayesian context regularization is done by specifying shrinkage priors on the regression coefficients β.

Two classes of shrinkage priors:

- **Spike-and-slab priors**: Discrete mixture of two distributions (Mitchell and Beauchamp, 1988; George and McCulloch, 1997)
- **Continuous shrinkage priors**: Unimodal continuous distributions (Bayesian Lasso prior, Horseshoe prior, Elastic-Net prior, ...) (Kyung et al., 2010; Carvalho et al., 2008)

$\rightarrowtail$ To shrink towards zero small effects while allowing large signals to escape from the overall shrinkage

Spike-and-slab prior

- Introduction of γ:

$$\gamma_j = \begin{cases} 1 \text{ if variable } j \text{ is selected} \\ 0 \quad \text{otherwise} \end{cases}$$

$$\beta_j|(\gamma_j = 1) \sim p_{Slab}(\beta_j), \qquad \beta_j|(\gamma_j = 0) \sim p_{Spike}(\beta_j)$$

- Bayesian hierarchical model:

$$Y|\mu, \beta, \sigma^2 \sim \mathcal{N}_n(\mu + X_1\beta_1 + \cdots + X_p\beta_p, \sigma^2 I_n)$$

$$\beta_j|\gamma \sim \gamma_j \mathcal{N}(0, \sigma_\beta^2) + (1 - \gamma_j)\delta_0, \quad j = 1, \ldots, p$$

- The estimation of $\mathbb{P}(\gamma_j = 1|Y)$ gives access to the a posteriori probability of variable selection

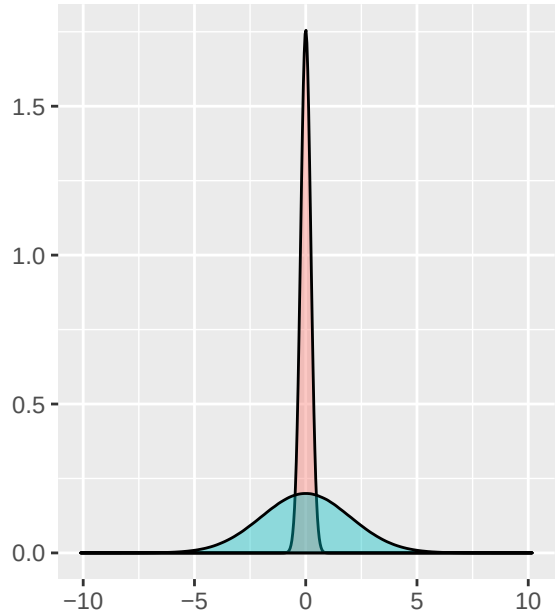


Figure 1: Spike-and-Slab prior distribution. Slab part in blue and spike part in red

Continuous shrinkage prior

- Penalized likelihood approaches:

$$\hat{\beta} = \text{argmin}_{\beta}\{\| Y - X\beta \|^2 + \nu \sum_{j=1}^{p} \phi(\beta_j^2)\}$$

with ν the penalty parameter and ϕ the penalization function.

- Bayesian version of penalized likelihood approaches:

$$\beta_j | \tau^2, \omega_j^2 \sim \mathcal{N}(0, \tau^2 \omega_j^2) \;\; j = 1, \ldots, p$$

$$\tau^2, \omega_j^2 \sim \mathcal{F}(\tau^2; \omega_j^2)$$

where

- $\mathcal{F}$ is a distribution to specify,
- τ^2 controls the global shrinkage and ω_j^2 controls the individual shrinkage

↬ Horseshoe prior, Bayesian Lasso prior (Laplace prior), ...

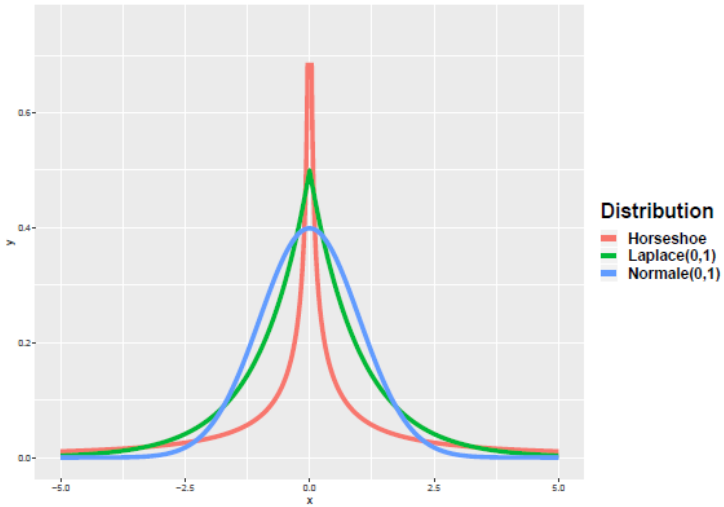


Figure 2: Continuous shrinkage prior distributions

R packages for implementing Bayesian variable selection

Spike-and-slab prior

- `BGLR` (Pérez and de Los Campos, 2014)
- `BoomSpikeSlab` (Scott et al., 2021)

Continuous shrinkage priors

- `bayesreg` (Makalic and Schmidt, 2016)
- `BGLR` (Pérez and de Los Campos, 2014)
- `horseshoe` (van der Pas et al., 2016)
- `fastHorseshoe` (Hahn et al., 2016)
- `dlbayes` (Zhang and Li, 2018)

Monte Carlo Markov Chain (MCMC) algorithms are used to infer parameters

Application in genetic

Objective: to select the genetic markers involved in the variation of the compactness of *Arabidopsis thaliana* (data publicly available at phenotypes (Loudet, 2018) and genotypes)

- Individuals: $n = 357$ under well-watered environmental condition
- Markers: SNP, $q = 532$
- Measurement frequency: daily for $T = 21$
- Phenotypic trait: compactness

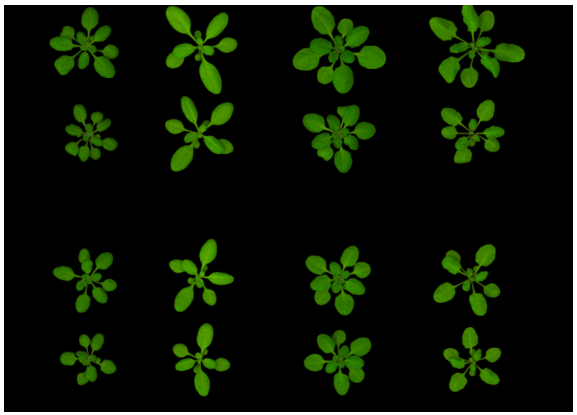
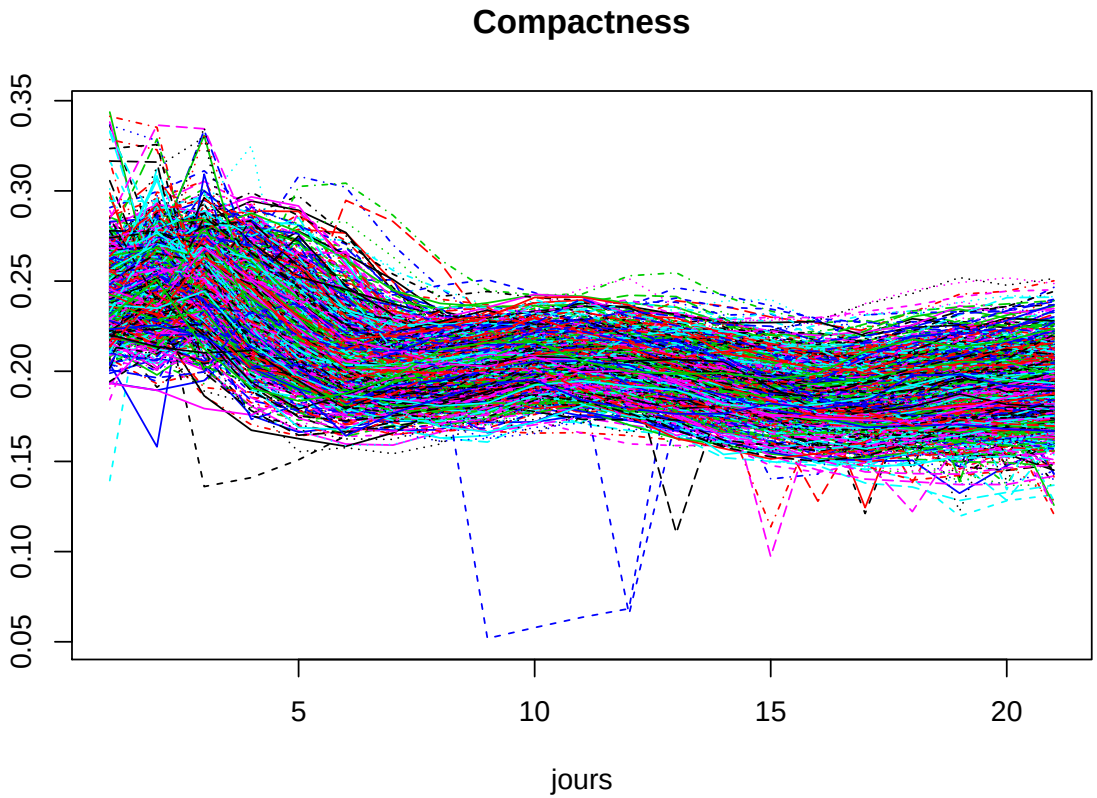


Figure 3: Arabidopsis thaliana (Marchadier et al., 2018)

Application in genetic

Evolution of the compactness over time for each individual:



In this first part we focus on the last time point $T = 21$

Bayesian variable selection with R

We will explore two R packages for implementing Bayesian variable selection with the spike-and-slab prior and the horseshoe prior:

- `bayesreg` allows to fit Bayesian regression models with continuous shrinkage priors,
- `BoomSpikeSlab` allows to fit Bayesian regression models with a spike-and-slab prior.

Arabidopsis Thaliana data

Analysis of the compactness at T=21 according to q = 50 markers:

To download the data

```
> load(file = "DataPrac.RData")
> dim(Y)
[1] 357  21
> dim(X.50)
[1] 357  50
> y.21 <- Y[,21] # to analyze the compactness at T=21
> y.21 <- y.21-mean(y.21)
```

Number of iterations of MCMC algorithms:

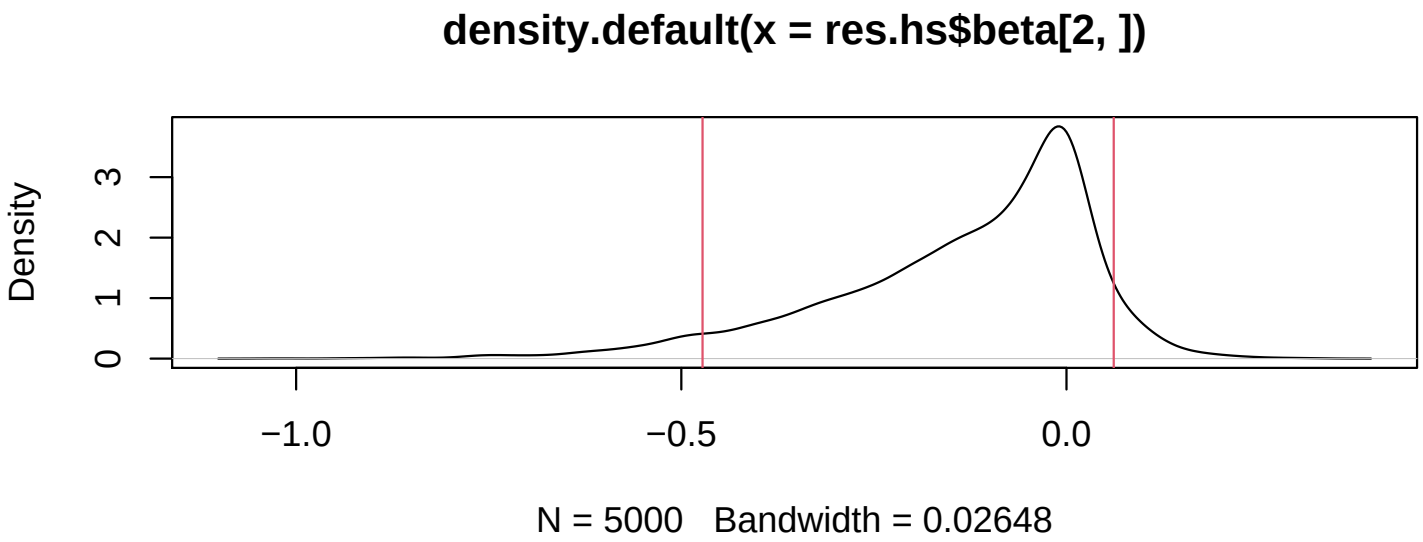
```
> niter <- 5000
```

bayesreg

To run the main function bayesreg

```
> library(bayesreg)
> res.hs <- bayesreg(y.21 ~ X.50, data = data.frame(y.21,X.50),
+                    model = "normal", prior = "hs", n.samples = niter)
> # 90% credible intervals
> ci <- apply(res.hs$beta, 1 , quantile, c(0.05,0.95))
> # selected markers
> which(apply(ci, 2, function(y) findInterval(0,y)) != 1)
 [1]  4 17 19 21 23 31 33 38 46 50
```

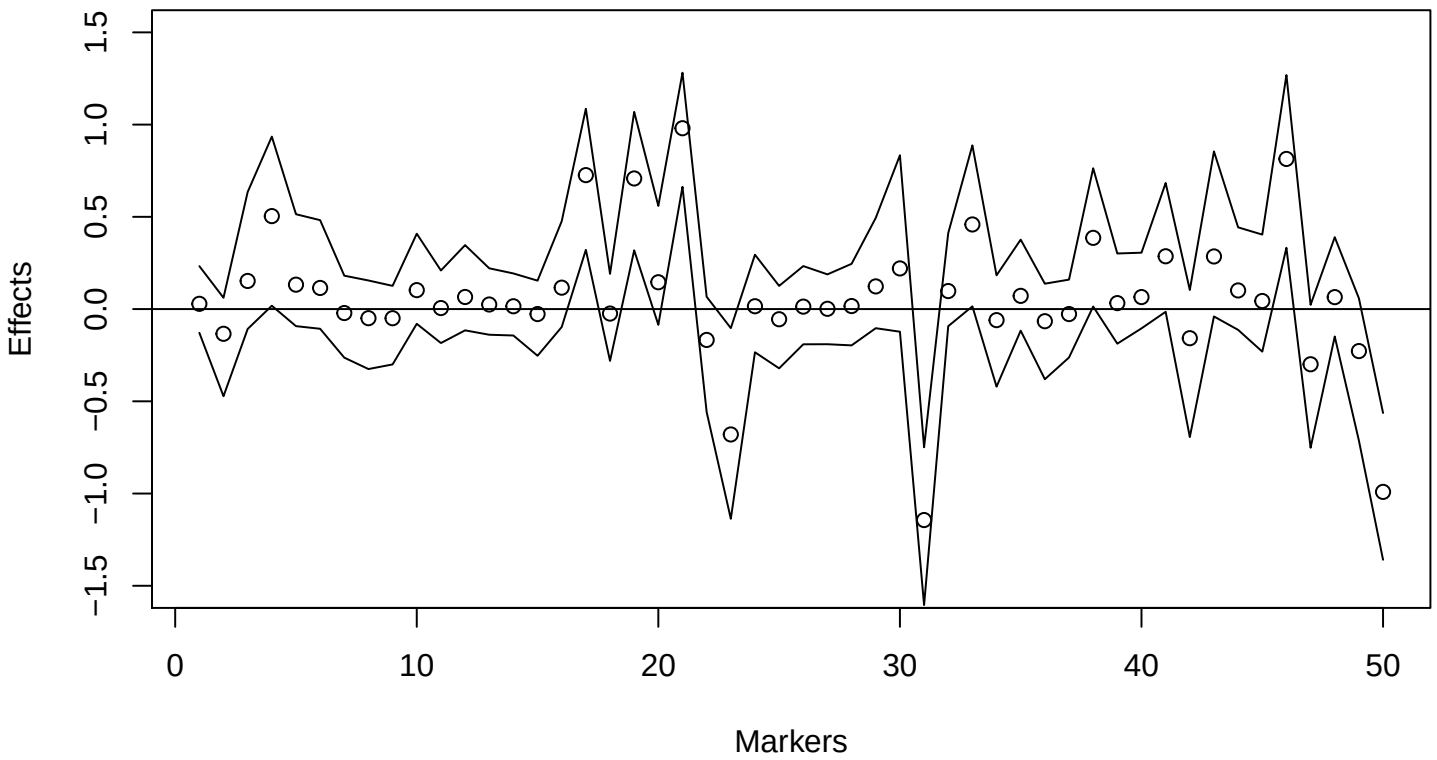
Posterior density of the marker 2 along with its 90% credible interval in red:



bayesreg

To plot the estimated effects:

```
> plot(rowMeans(res.hs$beta), ylab = "Effects", xlab = "Markers",
+      ylim = c(-1.5, 1.5))
> points(ci[1,], t = "l"); points(ci[2,], t = "l")
> abline(h = 0)
```



bayesreg

This package may be used to implement others shrinkage priors:

- `prior`: ridge, lasso, horseshoe, horseshoe+

and to analyze data with non normal distribution:

- `model`: normal, poisson, geometric, logistic, binomial, Laplace, student

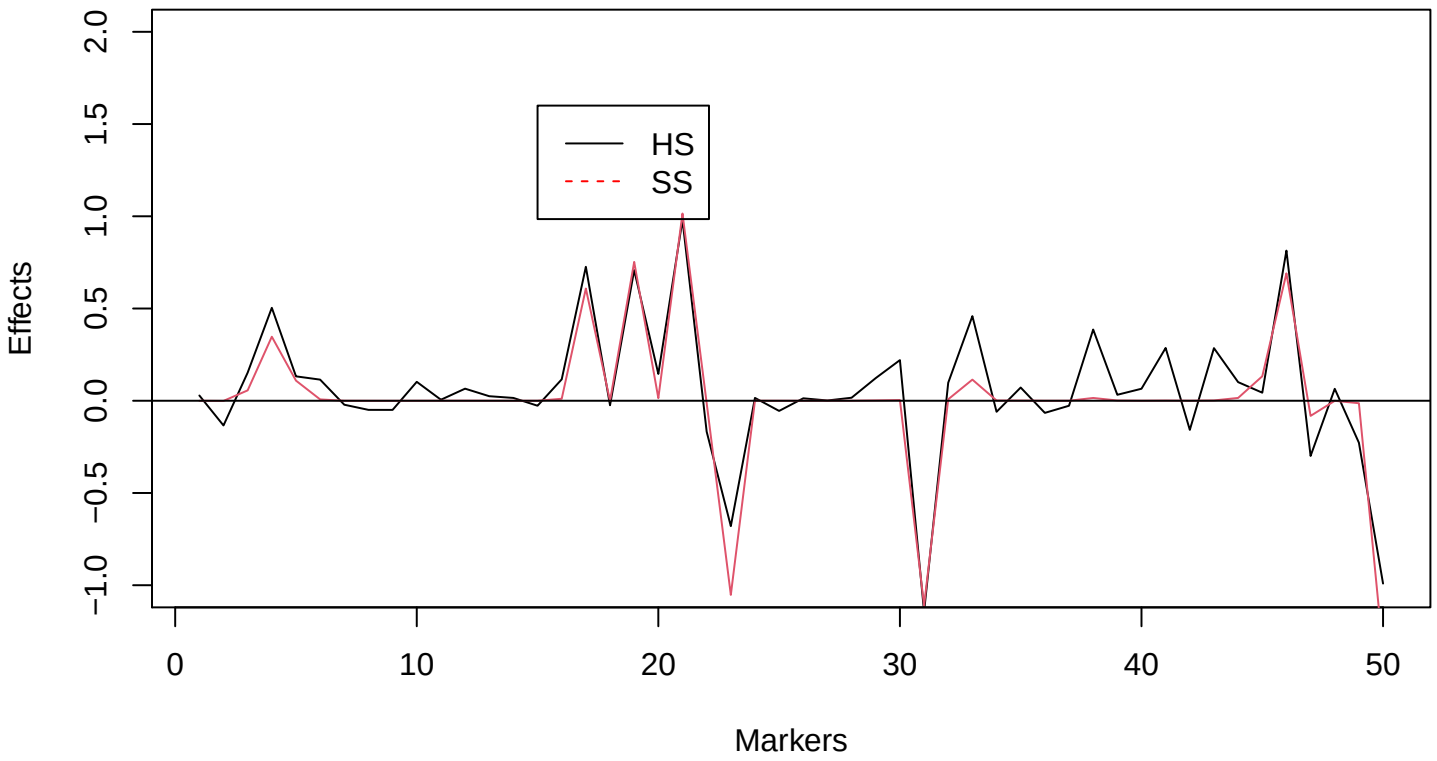
BoomSpikeSlab

To run the main function lm.spike

```
> library(BoomSpikeSlab)
> # to specify the prior inclusion probabilities equal to 0.1
> prior <- SpikeSlabPrior(X.50, y.21, prior.inclusion.probabilities = rep(0.1,ncol(X.50)))
> res.ss <- lm.spike(y.21 ~ X.50-1, niter = niter, prior = prior)
> # posterior inclusion probabilities
> pp <- apply(res.ss$beta, 2, function(c){sum(c != 0) })/niter
> #selected markers
> which(pp > 0.5)
  4  17  19  21  23  31  46  50
> beta.ss <- colMeans(res.ss$beta)
```

To plot the estimated effects

```
> plot(rowMeans(res.hs$beta), t="l", ylab = "Effects", xlab = "Markers", ylim = c(-1,2))
> lines(beta.ss, col =2)
```



How does the genetic architecture evolve over time?

We have identified the genetic markers involved in the variability of the compactness measured at $T = 21$

- Which are the genetic markers involved in the variability of the compactness measured at different time points ? Are they the same over time ?
- How are their effects? Are they the same over time ?

Need to analyze all the measures simultaneously, to select the relevant markers and to estimate their effects over time

$\Rightarrow$ Varying coefficient model

Outline

1 Introduction

2 Bayesian variable selection

3 Varying coefficient model
- Estimation of $\beta_j(t)$
- Selection of relevant variables X_j
- VCGSS package

4 Bibliography

From linear model to varying coefficient model

We assume that the response variables is followed over T times such that for the individual i we have: $y_i = (y_i^{t_1}, \ldots, y_i^{t^T})'$.

Linear model:

$$y_i^{t_1} = \mu^{t_1} + (\beta_1^{t_1}, \ldots, \beta_q^{t_1}) \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix} + \varepsilon_i^{t_1}, \quad \varepsilon_i^{t_1} \sim N(0, \sigma^2)$$

From linear model to varying coefficient model

We assume that the response variables is followed over T times such that for the individual i we have: $y_i = (y_i^{t_1}, \ldots, y_i^{t^T})'$.

Linear model:

$$\begin{array}{lllllll} y_i^{t_1} & = & \mu^{t_1} & + & (\beta_1^{t_1}, \ldots, \beta_q^{t_1}) & + & \varepsilon_i^{t_1}, \quad \varepsilon_i^{t_1} \sim N(0, \sigma^2) \\ y_i^{t_2} & = & \mu^{t_2} & + & (\beta_1^{t_2}, \ldots, \beta_q^{t_2}) \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix} & + & \varepsilon_i^{t_2}, \quad \varepsilon_i^{t_2} \sim N(0, \sigma^2) \end{array}$$

From linear model to varying coefficient model

We assume that the response variables is followed over T times such that for the individual i we have: $y_i = (y_i^{t_1}, \ldots, y_i^{t^T})'$.

Linear model:

$$\begin{array}{rcccccccl}
y_i^{t_1} & = & \mu^{t_1} & + & (\beta_1^{t_1}, & \ldots, & \beta_q^{t_1}) & + \varepsilon_i^{t_1}, & \varepsilon_i^{t_1} \sim N(0, \sigma^2) \\
y_i^{t_2} & = & \mu^{t_2} & + & (\beta_1^{t_2}, & \ldots, & \beta_q^{t_2}) & + \varepsilon_i^{t_2}, & \varepsilon_i^{t_2} \sim N(0, \sigma^2) \\
\vdots & & \vdots & & \vdots & & \vdots & \vdots & \\
y_i^{t_T} & = & \mu^{t_T} & + & (\beta_1^{t_T}, & \ldots, & \beta_q^{t_T}) & + \varepsilon_i^{t_T}, & \varepsilon_i^{t_T} \sim N(0, \sigma^2)
\end{array} \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix}$$

- Simple analysis at each time point does not take into account the correlations over the time
  - ↪ Can lead to false positive detection and loss of statistical power

Varying coefficient model

Varying coefficient model (Hastie and Tibshirani, 1993)

$$\begin{pmatrix} y_i^{t_1} \\ \vdots \\ y_i^{t_T} \end{pmatrix} = \begin{pmatrix} \mu^{t_1} \\ \vdots \\ \mu^{t_T} \end{pmatrix} + \begin{pmatrix} \beta_1^{t_1} & \dots & \beta_q^{t_1} \\ \vdots & & \vdots \\ \beta_1^{t_T} & \dots & \beta_q^{t_T} \end{pmatrix} \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix} + \begin{pmatrix} \varepsilon_i^{t_1} \\ \vdots \\ \varepsilon_i^{t_T} \end{pmatrix}, \quad \begin{array}{l} \varepsilon_i \sim N_T(0, \sigma^2 \Gamma) \\ \Gamma_{i,j} = \rho^{|i-j|} \\ -1 < \rho < 1 \end{array}$$

Varying coefficient model

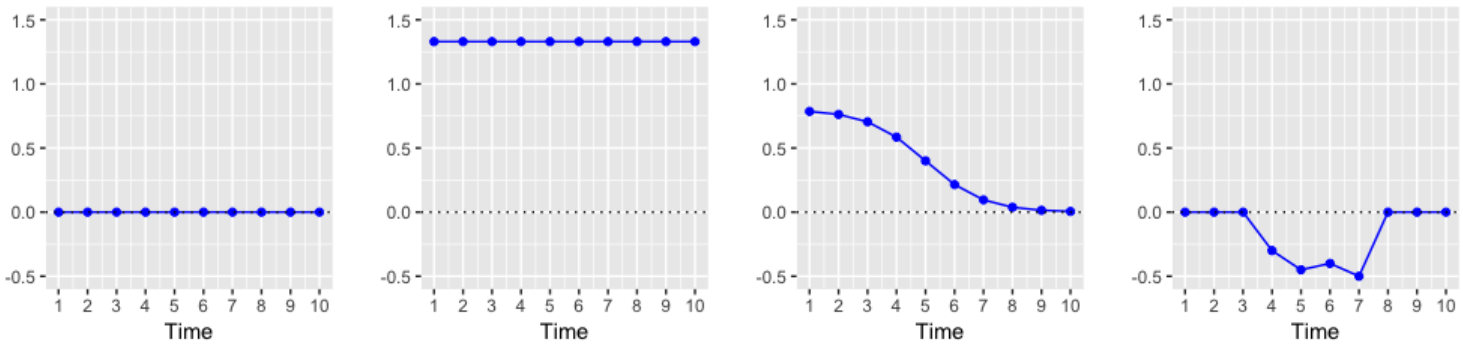
Varying coefficient model (Hastie and Tibshirani, 1993)

$$\begin{pmatrix} y_i^{t_1} \\ \vdots \\ y_i^{t_T} \end{pmatrix} = \begin{pmatrix} \mu^{t_1} \\ \vdots \\ \mu^{t_T} \end{pmatrix} + \begin{pmatrix} \beta_1^{t_1} & \dots & \beta_q^{t_1} \\ \vdots & & \vdots \\ \beta_1^{t_T} & \dots & \beta_q^{t_T} \end{pmatrix} \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix} + \begin{pmatrix} \varepsilon_i^{t_1} \\ \vdots \\ \varepsilon_i^{t_T} \end{pmatrix}, \quad \begin{array}{l} \varepsilon_i \sim N_T(0, \sigma^2 \Gamma) \\ \Gamma_{i,j} = \rho^{|i-j|} \\ -1 < \rho < 1 \end{array}$$

Varying coefficient model

Varying coefficient model (Hastie and Tibshirani, 1993)

$$\begin{pmatrix} y_i^{t_1} \\ \vdots \\ y_i^{t_T} \end{pmatrix} = \begin{pmatrix} \mu^{t_1} \\ \vdots \\ \mu^{t_T} \end{pmatrix} + \begin{pmatrix} \beta_1^{t_1} & \dots & \beta_q^{t_1} \\ \vdots & & \vdots \\ \beta_1^{t_T} & \dots & \beta_q^{t_T} \end{pmatrix} \begin{pmatrix} X_{i,1} \\ \vdots \\ X_{i,q} \end{pmatrix} + \begin{pmatrix} \varepsilon_i^{t_1} \\ \vdots \\ \varepsilon_i^{t_T} \end{pmatrix}, \quad \begin{array}{l} \varepsilon_i \sim N_T(0, \sigma^2 \Gamma) \\ \Gamma_{i,j} = \rho^{|i-j|} \\ -1 < \rho < 1 \end{array}$$



$(\beta_j^{t_1}, \dots, \beta_j^{t_T})'$ are assumed to be a realization of a function $\beta_j(t)$

- ↪ Estimation of $\beta_j(t)$ with functional or non functional methods
- ↪ Selection of significant variables X_j such that $(\beta_j^{t_1}, \dots, \beta_j^{t_T})' = (0, \dots, 0)'$ with a spike-and-slab prior

Outline

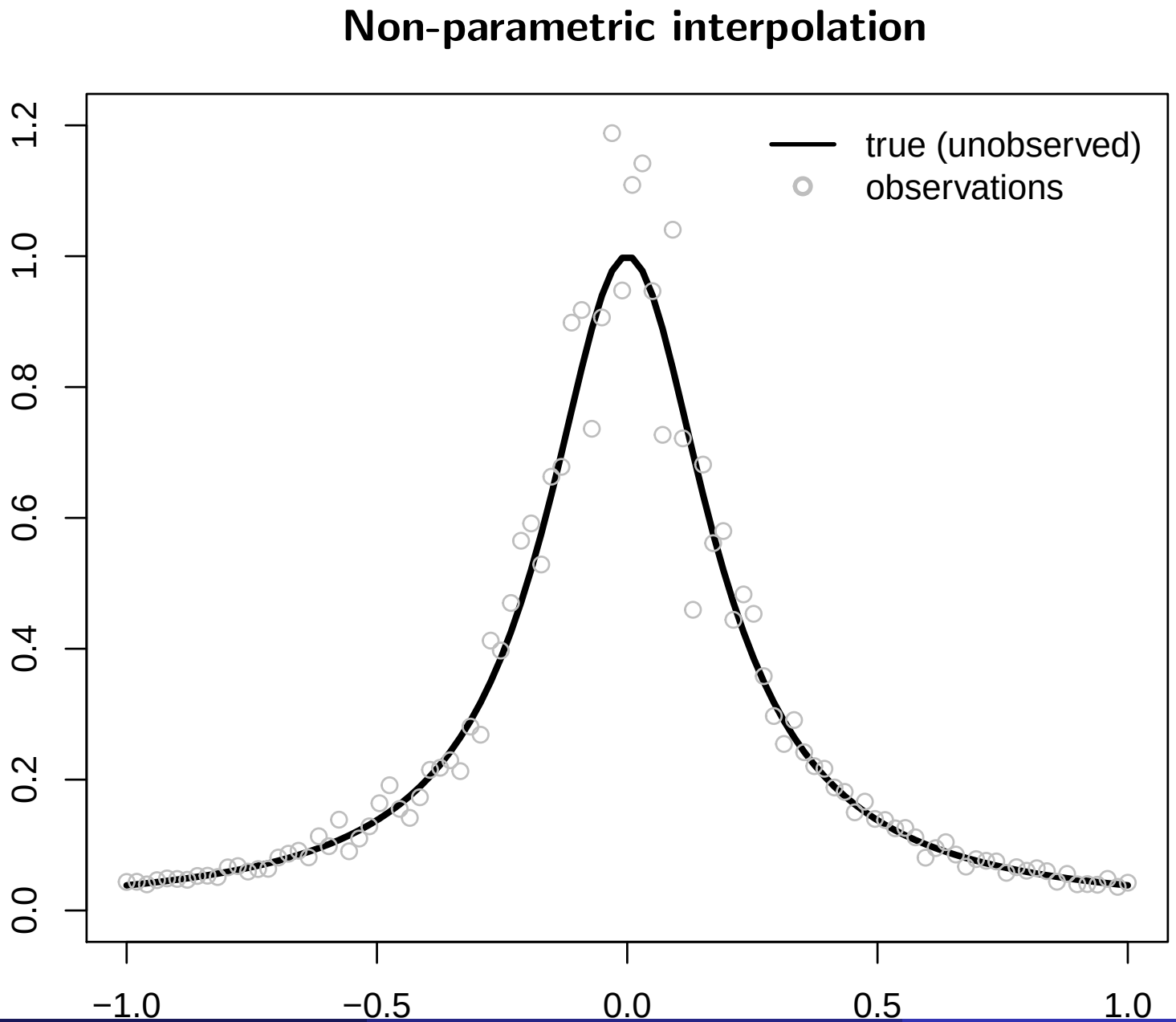
1 Introduction

2 Bayesian variable selection

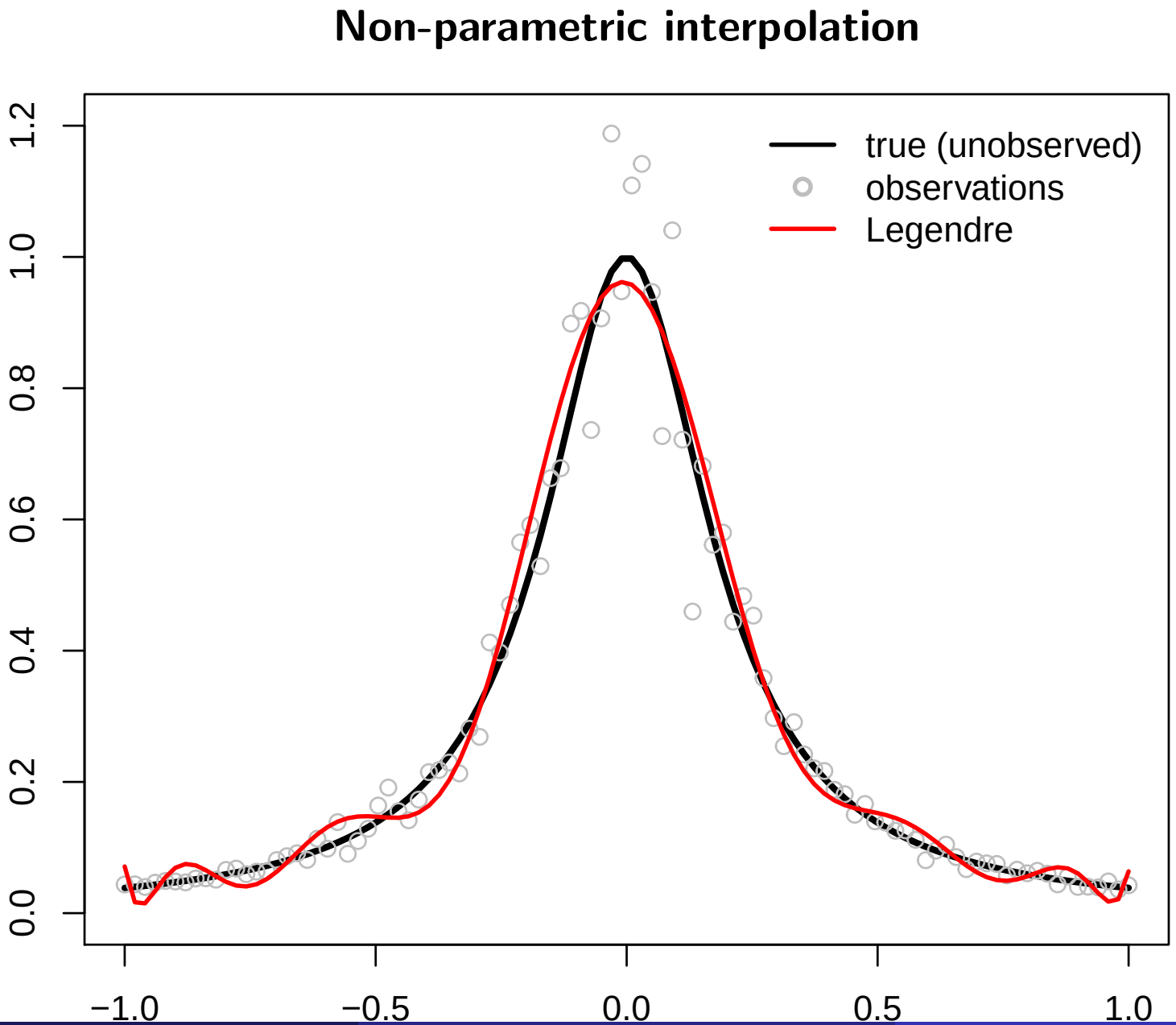
3 Varying coefficient model
- Estimation of $\beta_j(t)$
- Selection of relevant variables X_j
- VCGSS package

4 Bibliography

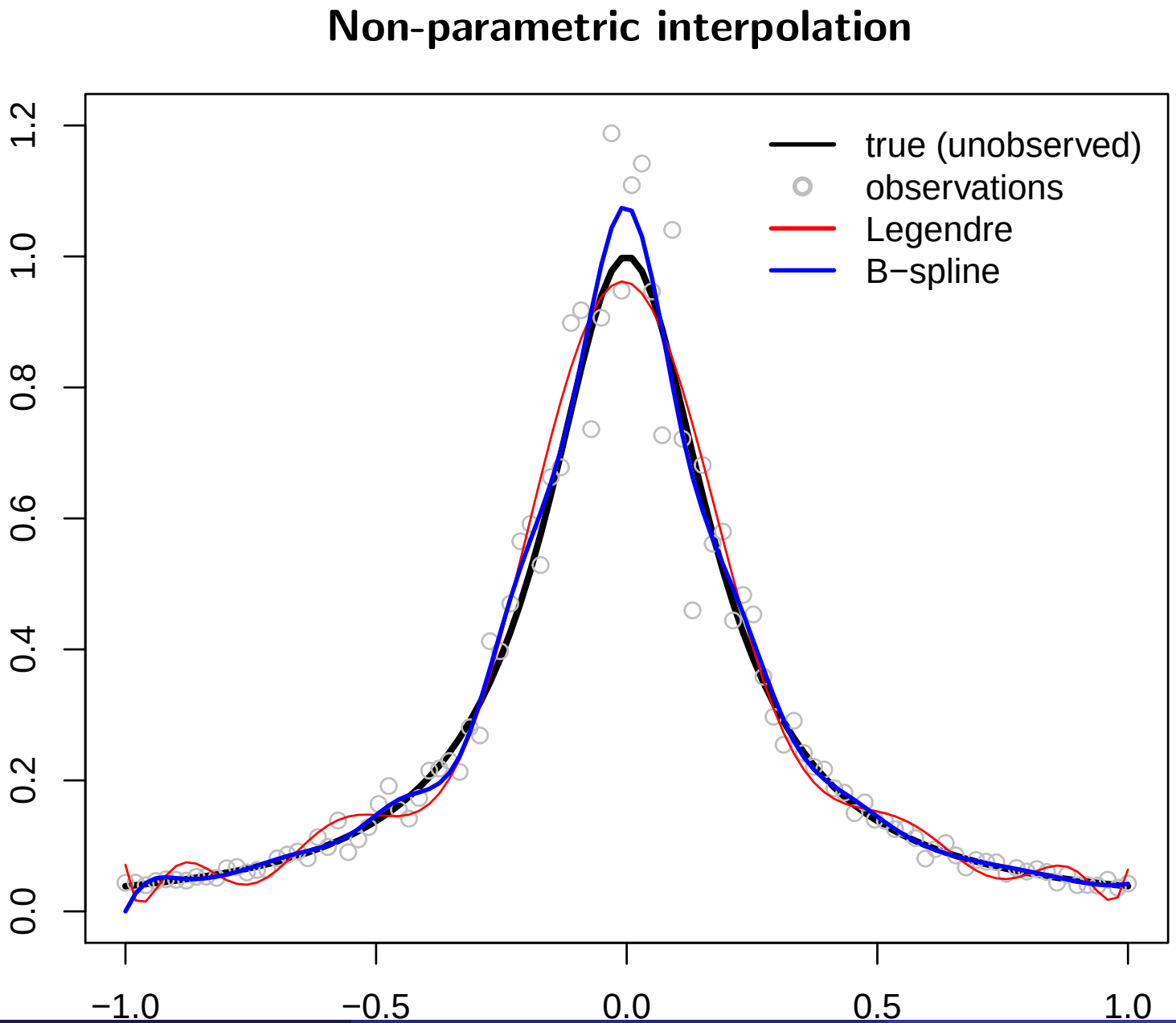
Functional method



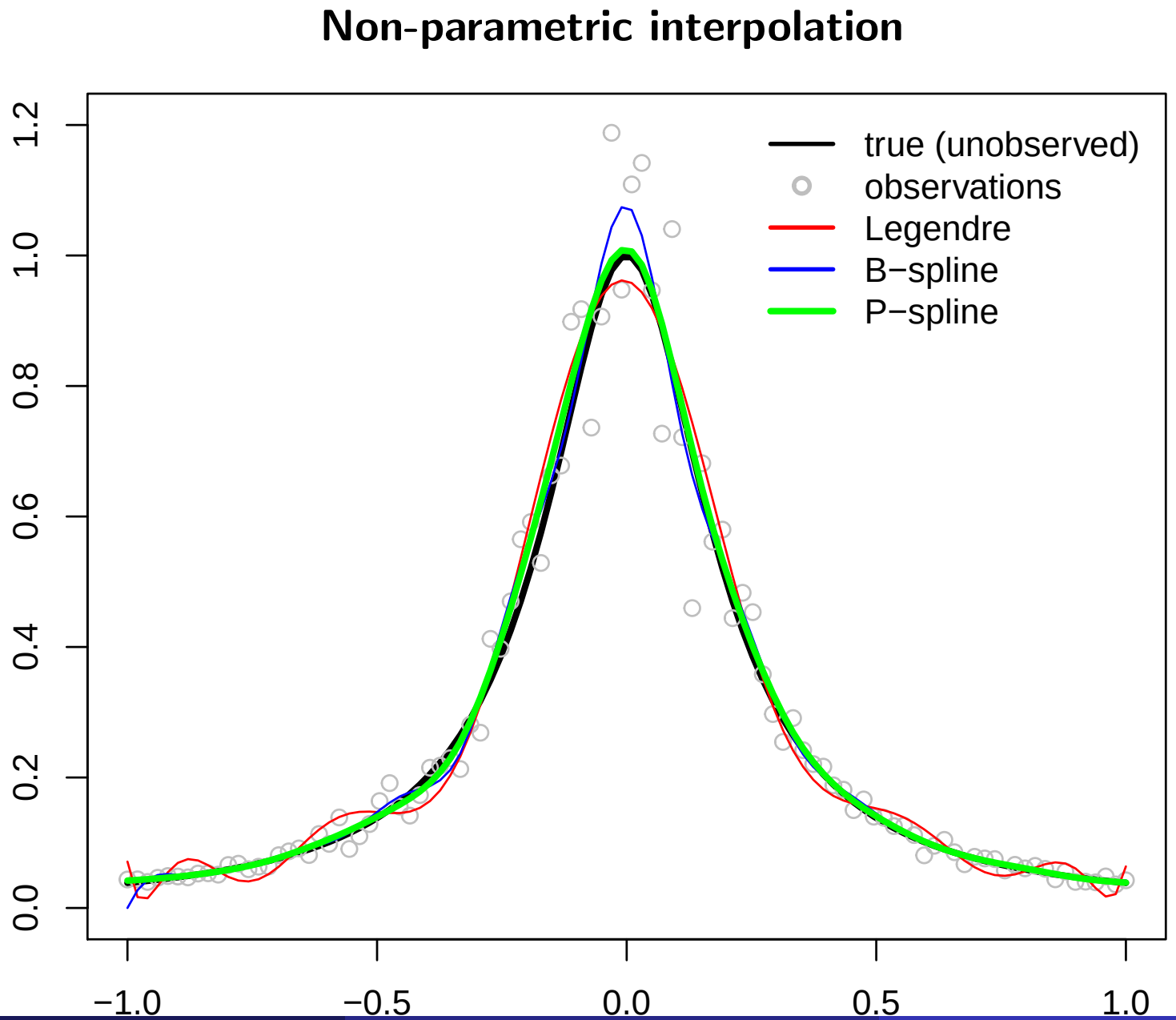
Functional method



Functional method



Functional method



Functional method

What is P-spline ?

P-spline = B-spline + Penalisation

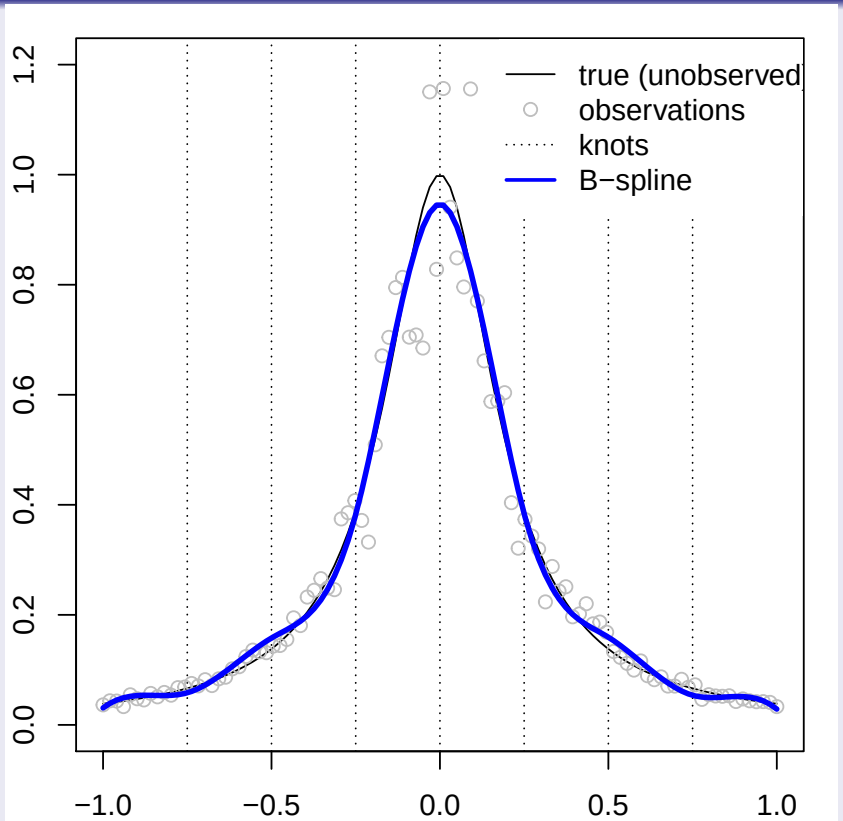
Functional method

B-spline (Eubank, 1999)

- define knots
- interpolate third degree polynomials on each piece
- conditions: C^0, C^1, C^2
- can be formulated as linear combination
  - ↪ define a new basis: B

$$\begin{pmatrix} \beta_j^{t_1} \\ \vdots \\ \beta_j^{t_T} \end{pmatrix} = \sum_{k=1}^{v} B_k b_{k,j} = B\ b_j,$$

- Disadvantage: sensitive to the choice of knots



Functional method

P-spline: penalized log-likelihood by the first order differences of adjacent B-spline coefficients (Eilers and Marx, 1996):

$$L = \sum_{i=1}^{n} l(Y_i, m, b_1, \ldots, b_q, \rho, \sigma^2)$$
$$-\lambda_0 \sum_{k=2}^{v} (m_k - m_{k-1})^2 - \lambda_1 \sum_{k=2}^{v} (b_{k,1} - b_{k-1,1})^2 - \cdots - \lambda_q \sum_{k=2}^{v} (b_{k,q} - b_{k-1,q})^2$$

- smooth curves
- not sensitive to the knot positions
- disadvantage: choice of $\lambda_0, \ldots, \lambda_q$ via cross-validation is computationally intensive

↬ **Bayesian formulation is convenient**

Functional method

Bayesian P-splines (Lang and Brezger, 2004):

Replace the penalties by their stochastic analogues

$\hookrightarrow$ first-order random walk such that $b_{k,j} \sim N(b_{k-1,j}, \lambda_j^{-1})$

$\Leftrightarrow$ $b_j|\lambda_j \sim N_v(0, (\lambda_j K)^{-1})$, K the known appropriate penalty matrix defined by

$$K = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}.$$

Functional versus non functional methods

Functional method: P-spline interpolation (Eilers and Marx, 1996)

$$\begin{pmatrix} \beta_j^{t_1} \\ \vdots \\ \beta_j^{t_T} \end{pmatrix} = \sum_{k=1}^{v} B_k b_{k,j} = B\ b_j,$$

+ penalization on $(b_{k,j} - b_{k-1,j})$ of order 1 or 2
$\hookrightarrow$ PS_1 and PS_2

Non functional method: direct estimation of time coefficient functions) Li and Sillanpää (2013)

$$\begin{pmatrix} \beta_j^{t_1} \\ \vdots \\ \beta_j^{t_T} \end{pmatrix} = b_j$$

+ penalization on $(b_{t_T,j} - b_{t_{T-1},j})$ of order 1 or 2
$\hookrightarrow$ RW_1 and RW_2

In both methods we assume that:

$$b_j \sim \mathcal{N}(0, (\lambda_j K)^{-1})$$

with K a structured matrix.

Outline

1 Introduction

2 Bayesian variable selection

3 Varying coefficient model
- Estimation of $\beta_j(t)$
- Selection of relevant variables X_j
- VCGSS package

4 Bibliography

Selection in varying coefficient model

Selection of relevant variables X_j, $j = 1, \ldots, p$:

$$(b_{t_1,j}, \ldots, b_{t_T,j})' = (0, \ldots, 0)'?$$

↬ Group spike-and-slab prior on b_j (Ghosh and Ghattas, 2015; Yang and Narisetty, 2020):

$$b_j|\gamma_j, \lambda_j \sim \gamma_j N_v(0, (\lambda_j K)^{-1}) + (1 - \gamma_j)\delta_v(0)$$
$$\lambda_j \sim Gamma(s, r),$$
$$\gamma_j \sim Ber(\pi),$$

Bayesian Varying Coefficient model using Group Spike-and-Slab prior Heuclin et al. (2021)

Bayesian hierarchical model

$$Y_i|m, b, \rho, \sigma^2 \sim N_T(Bm + BbX_i, \sigma^2\Gamma)$$

$$m|\lambda_0 \sim N_v(0, (\lambda_0 K)^{-1})$$

$$b_j|\gamma_j, \lambda_j \sim \gamma_j N_v(0, (\lambda_j^2 K)^{-1}) + (1-\gamma_j)\delta_v(0), \quad j = 1, \ldots, q$$

$$\lambda_j \sim Gamma(s, r), \quad j = 0, \ldots, q$$

$$\gamma_j \sim Ber(\pi), \quad j = 1, \ldots, q$$

$$\rho \sim U_{[-1,1]}$$

$$\sigma^2 \sim I - Gamma(s_{\sigma^2}, r_{\sigma^2})$$

To infer the distribution of $m, b_j, \lambda_j, \gamma_j, \rho, \sigma^2|Y$**:**

↪ Gibbs algorithm (Markov Chain Monte Carlo algorithm)

Outline

1 Introduction

2 Bayesian variable selection

3 Varying coefficient model
- Estimation of $\beta_j(t)$
- Selection of relevant variables X_j
- VCGSS package

4 Bibliography

VCGSS package

VCGSS: an R package for implementing the sparse Bayesian **V**arying **C**oefficient model using **G**roup **S**pike-and-**S**lab prior

↬ Available on https://github.com/Heuclin/VCGSS

We will explore the two main functions:

`VCM_fct()`

To run the Bayesian Varying Coefficient model using Group Spike-and-Slab prior.

- allows to implement functional and non functional methods with penalty of order 1 or 2,
- calls an MCMC sampler implementation in C++,
- allows to run many repetitions in parallel,
- applies convergence diagnostics

`plot_functional_effects()`:

To visualize the dynamic effects

Application in functional mapping

Objectives: to select the genetic markers involved on the variation of the compactness of *Arabidopsis thanliana* over time and to estimate their effects over time

↬ Same data than in the first part but here we analyze jointly all the measures

To download and install VCGSS package

```
> library(devtools)
> devtools::install_github("Heuclin/VCGSS")
```

If it does not work:

```
> source("~/Documents/Bourse/doc/data/NIH/Weight analysis/VCM-function.R")
> Rcpp::sourceCpp("~/Documents/Bourse/doc/data/NIH/Weight analysis/VCM.cpp"
```

To run the model

To run the model with a P-spline interpolation using a second order difference penalty

```
> library(VCGSS)
> load("DataPrac.RData")
> fit <- VCM_fct(Y, X.50, ENV = NULL, selection = TRUE,
+                interpolation = "P-spline",
+                order_diff = 2,
+                save = FALSE, core = -1,
+                rep = 2, niter = 1000, burnin = 500)
```

Arguments:

- `selection`: TRUE/FALSE
- `interpolation`: functional method = "P-spline", "B-spline", "Legendre", non-functional method = "RW",
- `order_diff`: order of the difference penalty,
- `niter, burnin, thin`: MCMC parameters,
- ...

The outputs: Convergence diagnostics

- Gelman-Rubin's potential scale reduction factor for all parameters except β

```
> fit$gelman.diag
```

- Gelman-Rubin's potential scale reduction factor for β

```
>   fit$gelman.diag.b.psrf.median
```

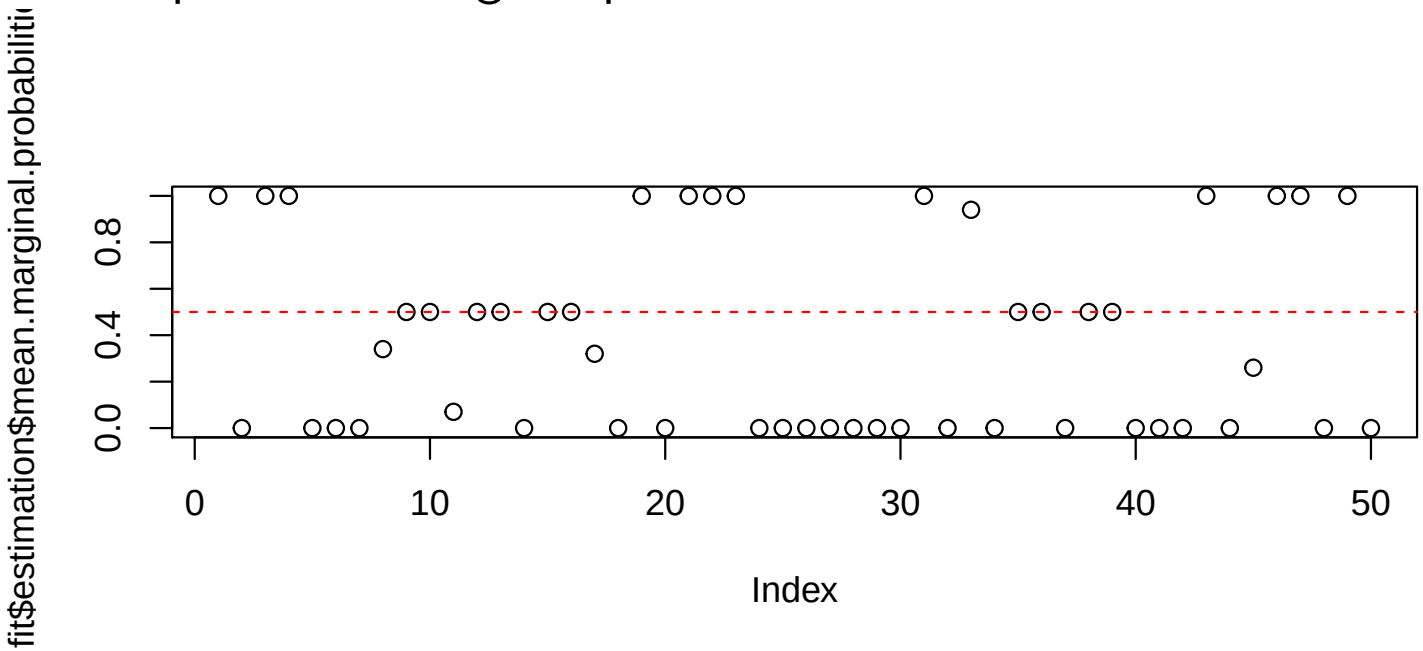
- Trace plot with posterior densities

```
>   plot(fit$mcmc_list)
```

To plot the posterior marginal probabilities

```
>   plot(fit$estimation$mean.marginal.probabilities)
> abline(0.5, 0, lty =2, col = "red")
> # markers selected
> which(fit$estimation$mean.marginal.probabilities> 0.5)
```

Plot of posterior marginal probabilities of inclusion of markers:

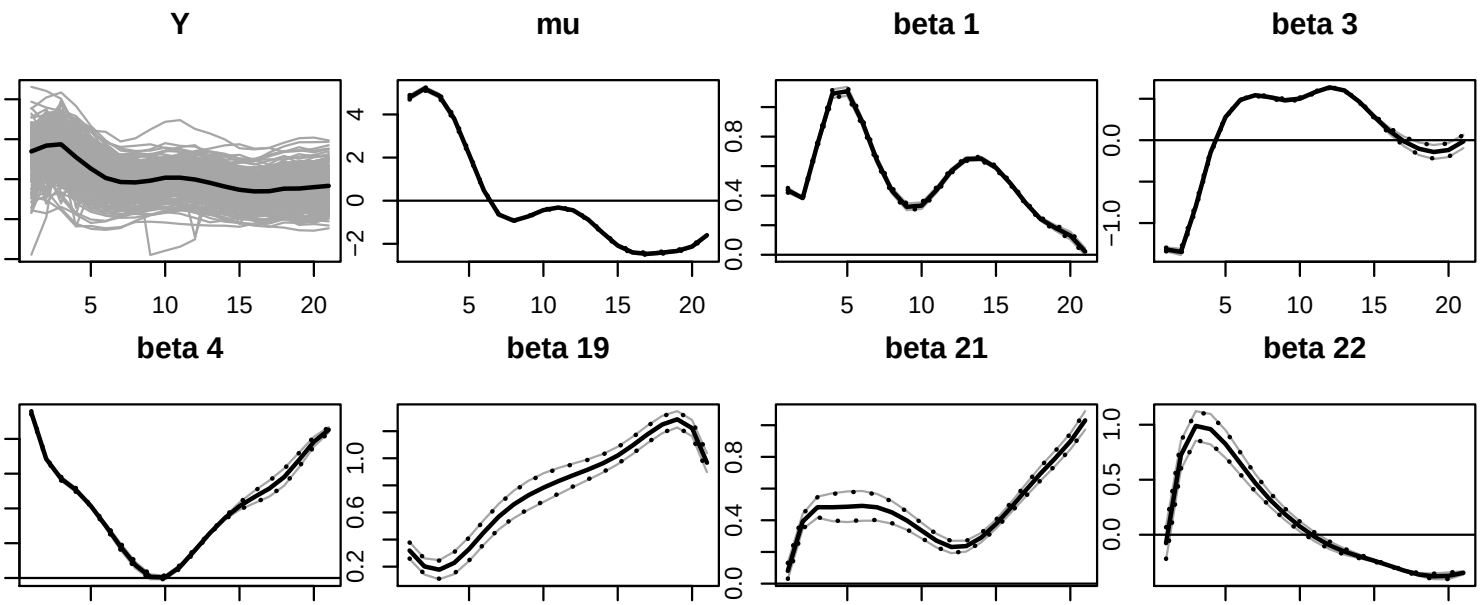


Markers with posterior inclusion probabilities greater than 0.5:

```
 [1]  1  3  4 19 21 22 23 31 33 43 46 47 49
```

To visualize the dynamic effects

```
> plot_functional_effects(fit, plot=c("Y", "mu", "beta"), mfrow = c(2, 4),
+                         add = c("matplot", "quantile"))
```



Estimated effects of the intercept (μ) and selected markers. Gray lines are the estimation for each repetition, black lines are the mean of the estimation over the repetitions and dotted lines are the credible intervals.

To run the model with a non functional method: RW1

To set the arguments `interpolation` and `order_diff` to "RW" and 1, respectively.

```
>   fit.rw <- VCM_fct(Y, X.50, ENV = NULL, selection = TRUE,
+                     interpolation = "RW",
+                     order_diff = 1,
+                     save = FALSE, core = -1,
+                     rep = 2, niter = 1000, burnin = 500)
```

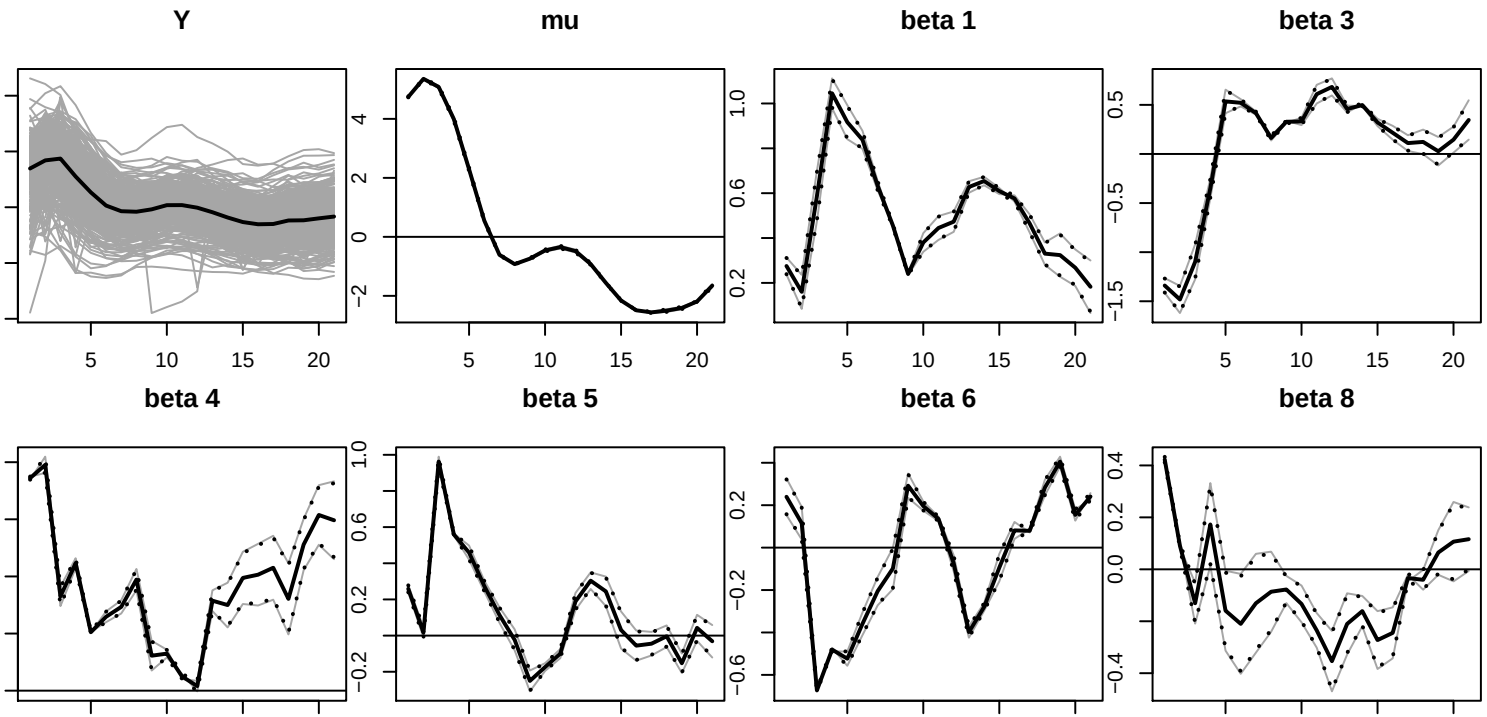
The selected markers are:

```
>   which(fit.rw$estimation$mean.marginal.probabilities > 0.5)
 [1]  1  3  4  5  6  8  9 10 11 12 15 16 18 19 21 22 23 24 27 28 29 30 31 33 34 35 36 42 43 44 4
```

To run the model with a non functional method: RW1

The estimated dynamic effects:

```
>   plot_functional_effects(fit.rw, plot=c("Y", "mu", "beta"), mfrow = c(2, 4),
+                           add = c("matplot", "quantile"))
```



Outline

1. Introduction
2. Bayesian variable selection
3. Varying coefficient model
   - Estimation of $\beta_j(t)$
   - Selection of relevant variables X_j
   - VCGSS package
4. Bibliography

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2008). The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B -splines and penalties. *Statistical Science*, 11(2):89–121.

Eubank, R. L. (1999). *Nonparametric Regression and Spline Smoothing*. CRC press.

George, E. I. and McCulloch, R. E. (1997). Approaches for bayesian variable selection. *Statistica sinica*, pages 339–373.

Ghosh, J. and Ghattas, A. (2015). Bayesian Variable Selection Under Collinearity. *The American Statistician*, 69(3):165–173.

Hahn, P. R., He, J., and Lopes, H. (2016). Elliptical slice sampling for bayesian shrinkage regression with applications to causal inference. *U RL http://faculty. chicagobooth. edu/richard. hahn/research. html*.

Hastie, T. and Tibshirani, R. (1993). Varying-Coefficient Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4):757–796.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55.

Kyung, M., Gill, J., Ghosh, M., Casella, G., et al. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5(2):369–411.

Lang, S. and Brezger, A. (2004). Bayesian P-Splines. *Journal of Computational and Graphical Statistics*, 13(1):183–212.

Li, Z. and Sillanpää, M. J. (2013). A Bayesian Nonparametric Approach for Mapping Dynamic Quantitative Traits. *Genetics*, 194(4):997–1016.

Loudet, O. (2018). Raw phenotypic data obtained on the arabidopsis rils with the phenoscope robots (marchadier, hanemian, tisné et al., 2018).

Makalic, E. and Schmidt, D. F. (2016). High-dimensional bayesian regularised regression with the bayesreg package. *arXiv preprint arXiv:1611.06649*.

Marchadier, E., Hanemian, M., Tisne, S., Bach, L., Bazakos, C., Gilbault, E., Haddadi, P., Virlouvet, L., and Loudet, O. (2018). The complex genetic architecture of shoot growth natural variation in Arabidopsis thaliana.

Mitchell, T. J. and Beauchamp, J. J. (1988). Bayesian variable selection in linear regression. *Journal of the american statistical association*, 83(404):1023–1032.

Pérez, P. and de Los Campos, G. (2014). Genome-wide regression and prediction with the bglr statistical package. *Genetics*, 198(2):483–495.

Scott, S. L., Scott, M. S. L., and Boom, D. (2021). Package 'boomspikeslab'.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

van der Pas, S., Scott, J., Chakraborty, A., and Bhattacharya, A. (2016). horseshoe: Implementation of the horseshoe prior. *R package version 0.1. 0*.

Yang, X. and Narisetty, N. N. (2020). Consistent group selection with bayesian high dimensional modeling. *Bayesian Analysis*.

Zhang, S. and Li, M. (2018). "dlbayes" available at cran, r package for implementing the dirichlet-laplace shrinkage prior in bayesian linear regression and variable selection.