1 Introduction

We consider a problem were we measure p continuous variables $(X^{(1)}, \cdots, X^{(p)})$ on n samples. We further assume that the n sample are collected under two Condition 1 and Condition 2 let n_1 and n_2 be the number of samples collected under Condition 1, respectively Condition 2. Eventually we assume that we measure a continuous response y on this n samples and that this response is center.

We want to study the model :

$$y_{j,k} = \mu + \alpha_j + \sum_{i=1}^{p} (\beta_{i,j} + \gamma_i) X_{j,k}^{(i)} + E_{j,k}, \tag{1}$$

where $y_{j,k}$ is the value of the response y for the k^{th} sample of Condition j.

We can re-write this problem as :

$$y_{j,k} = \delta_j + \sum_{i=1}^{p} (d_{i,j}) X_{j,k}^{(i)} + E_{j,k}, \tag{2}$$

where $\delta_j = \mu + \alpha_j$ and $d_{i,j} = \beta_{i,j} + \gamma_i$. As y as a mean of 0, $\mu = 0$.

Goal : Here we want to find which regressors can explain the best the responses and whether or not each regressors explain indifferently the response between the two conditions. In other terms we want to assess which $\beta_{i,j} = 0$ and which $\gamma_i = 0$ in (1) or which $b_{i,j} = 0$ and which $b_{i,1} = b_{i,2}$ in (2). To answer this problem we propose to use the fused lasso methodology for Model 2 and we will show that we can re-write it as a classical lasso problem in Model 1.

Let's call X the ANCOVA design matrix, for model 2 link to our data, in other terms X has $2(p+1)$ column and looks like :

$$X = \begin{pmatrix} 1 & 0 & X_1^{(1)} & 0 & X_1^{(2)} & 0 & \cdots \\ 1 & 0 & X_2^{(1)} & 0 & X_2^{(2)} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 1 & 0 & X_{n_1}^{(1)} & 0 & X_{n_1}^{(2)} & 0 & \cdots \\ 0 & 1 & 0 & X_{n_1+1}^{(1)} & 0 & X_{n_1+1}^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 1 & 0 & X_{n_1+n_2}^{(1)} & 0 & X_{n_1+n_2}^{(2)} & \cdots \end{pmatrix}. \tag{3}$$

Model 2 can then be re-written as:

$$y = XD + E, \tag{4}$$

where D is a sparse vector of length $2(p+1)$. For any $2(p+1)$ vector D lets defined $D_i^{(1)}$ the coefficients for the regressor i for the 1st modalities and $D_i^{(2)}$ the coefficients for the regressor i for the 2nd modalities it correspond respectively to $d_{i,1}$ and $d_{i,2}$ from Model 2. Lets also call $D_0^{(1)}$ and $D_0^{(2)}$ the effect of Condition 1, respectively 2 on y, it corresponds to δ_1 respectively δ_2 of 2. In order to estimate D we want to resolve

$$\widetilde{D}_{\lambda_1,\lambda_2} = \text{Argmin}_D \left\{ \|y - XD\|_2^2 + \lambda_1 \|D\|_1 + \lambda_2 \sum_{i=1}^{p} |D_i^{(1)} - D_i^{(2)}| \right\}. \tag{5}$$

Let us now re-write Model 1. Let X_2 be the matrix X in which we add all the regressors at the end. X_2 has then $2(p+1) + p$ column and looks like :

$$X_2 = \begin{pmatrix} 1 & 0 & X_1^{(1)} & 0 & X_1^{(2)} & 0 & \cdots & X_1^{(1)} & X_1^{(2)} & \cdots \\ 1 & 0 & X_2^{(1)} & 0 & X_2^{(2)} & 0 & \cdots & X_2^{(1)} & X_2^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots \\ 1 & 0 & X_{n_1}^{(1)} & 0 & X_{n_1}^{(2)} & 0 & \cdots & X_{n_1}^{(1)} & X_{n_1}^{(2)} & \cdots \\ 0 & 1 & 0 & X_{n_1+1}^{(1)} & 0 & X_{n_1+1}^{(2)} & \cdots & X_{n_1+1}^{(1)} & X_{n_1+1}^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots \\ 0 & 1 & 0 & X_{n_1+n_2}^{(1)} & 0 & X_{n_1+n_2}^{(2)} & \cdots & X_{n_1+n_2}^{(1)} & X_{n_1+n_2}^{(2)} & \cdots \end{pmatrix}. \tag{6}$$

Then Model 1 can be re-written as :

$$y = X_2 B + E, \tag{7}$$

where B is a sparse vector of length $2(p+1)+p$. For any $2(p+1)+p$ vector B let us call $B_i^{(12)}$, $B_i^{(2)}$, $B_i^{(2)}$ the coefficients of B associate to the regressor i, the regressor i for the first condition and the one associate with the regressor i for the second condition respectively. Using the notation of 1 we get $B_i^{(12)} = \gamma_i$, $B_i^{(2)} = \beta_{i,1}$ and $B_i^{(2)} = \beta_{i,2}$ Lets also call $B_0^{(1)}$ and $B_0^{(2)}$ the effect of Condition 1, respectively 2 on y, it corresponds to α_1 respectively $\alpha 2$ of 1.

The LASSO criterion applied to X_2 is then :

$$B_{\lambda,a,b} = \text{Argmin}_B \left\{ \|y - X_2 B\|_2^2 + \lambda \sum_{i=1}^{p} (|B_i^{(1)}| + |B_i^{(2)}| + |B_i^{(12)}|) + \lambda(|B_0^{(1)}|, |B_0^{(2)}|) \right\}.$$

We don't want to penalised all the coefficient equally. In order to do that we will put different weight according if the coefficients is associate to one whole regressor, to one condition of the regressor or to the effect of a condition alone.

$$\widetilde{B}_{\lambda,a,b} = \text{Argmin}_B \left\{ \|y - X_2 B\|_2^2 + \lambda \sum_{i=1}^{p} (b|B_i^{(1)}| + b|B_i^{(2)}| + a|B_i^{(12)}|) + c\lambda(|B_0^{(1)}|, |B_0^{(2)}|) \right\}. \tag{8}$$

2 Theorem

Here we want to find which regressors can explain the best the response and whether or not each regressors explain indifferently the response between the two conditions. One way to do this could be to resolve In other terms we want to assess which $\beta_{i,j} = 0$ and which $\gamma_i = 0$ in (1) or which $b_{i,j} = 0$ and which $b_{i,1} = b_{i,2}$ in (2). To answer this problem we propose to use the fused lasso methodology for Model 2 and we will show that we can re-write it as a classical lasso problem in Model 1.

Theorem 1. *Let's define :*

1. $a = \frac{2\lambda_1}{\lambda}$
2. $\lambda = \frac{(\lambda_1 + \lambda_2)}{b}$
3. $c = \frac{\lambda_1}{\lambda}$

If $2b > a > b > 0$ *(* *ie* $\lambda_2 > 0$ *and* $\lambda_1 > \lambda_2$ *) then:*

$$\widetilde{D}_0^{(1)} = \widetilde{B}_0^{(1)},\ \widetilde{D}_0^{(2)} = \widetilde{B}_0^{(2)},\ \widetilde{D}_i^{(1)} = \widetilde{B}_i^{(1)} + \widetilde{B}_i^{(12)} \text{ and } \widetilde{D}_i^{(2)} = \widetilde{B}_i^{(2)} + \widetilde{B}_i^{(12)}, \forall i \in \{1, \ldots, p\}, \tag{9}$$

where $\widetilde{D} = \widetilde{D}_{\lambda_1,\lambda_2}$ *and* $\widetilde{B} = \widetilde{B}_{\lambda,a,b}$ *are define in 5 and 8 respectively.*

3 In practice

In practice : to use the package glmnet we must have : $pa + 2pb + 2c = 3p + 2 \Rightarrow pa + 2pb + b = 3p + 2 \Rightarrow b = \frac{3p+2-ap}{2p+1}$ where the second implication come from the fact that $c = b/2$. **What are the value of a, b,λ using λ_1 and λ_2?**

$$\begin{aligned}
a &= \frac{2\lambda_1 b}{\lambda_1+\lambda_2} \\
&= \frac{2\lambda_1}{\lambda_1+\lambda_2}\frac{3p+2-ap}{2p+1} \\
&= \frac{2\lambda_1}{\lambda_1+\lambda_2}\frac{3p+2}{2p+1} - a\frac{2\lambda_1}{\lambda_1+\lambda_2}\frac{p}{2p+1} \\
&\Rightarrow a\frac{2\lambda_1 p + (\lambda_1+\lambda_2)(2p+1)}{(\lambda_1+\lambda_2)(2p+1)} = \frac{2\lambda_1}{\lambda_1+\lambda_2}\frac{3p+2}{2p+1} \\
&\Rightarrow a = \frac{2\lambda_1(3p+2)}{2\lambda_1 p + (\lambda_1+\lambda_2)(2p+1)}
\end{aligned}$$

We just have to use the value of a to get the value of b and c and then the value of λ.
Where the value of a can be ?
we must have $2b > a > b > 0$

$$\begin{aligned}
2b > a &\Rightarrow \frac{3p+2-ap}{p+1/2} > a \\
&\Rightarrow \frac{3p+2}{p+1/2} > a(\frac{1/2+2p}{p+1/2}) \\
&\Rightarrow \frac{3p+2}{1/2+2p} > a
\end{aligned}$$

$$\begin{aligned}
a > b &\Rightarrow \frac{3p+2-ap}{2p+1} < a \\
&\Rightarrow \frac{3p+2}{2p+1} < a(\frac{3p+1}{2p+1}) \\
&\Rightarrow \frac{3p+2}{3p+1} < a
\end{aligned}$$

4 Proof

Of Theorem 1. Let's call $\widetilde{B} = \widetilde{B}_{\lambda,a,b}$ and $\widetilde{D} = \widetilde{D}_{\lambda_1,\lambda_2}$ with $\widetilde{B}_{\lambda,a,b}$ and $\widetilde{D}_{\lambda_1,\lambda_2}$ define in 8 and 5 respectively. We will prove the Theorem using a "reductio ad absurdum". If $\exists i \in (1,\dots,p)$ such that, $\widetilde{D}_i^{(1)} \neq \widetilde{B}_i^{(1)} + \widetilde{B}_i^{(12)}$ or if $C_0^{(1)} \neq \widetilde{B}_0^{(1)}$ or $C_0^{(2)} \neq \widetilde{B}_0^{(2)}$.

Let C be $2(p+1)$ vector such that $\forall i \in (1,\dots,p)$ $C_i^{(1)} = \widetilde{B}_i^{(1)} + \widetilde{B}_i^{(12)}$, $C_i^{(2)} = \widetilde{B}_i^{(2)} + \widetilde{B}_i^{(12)}$, $C_0^{(1)} = \widetilde{B}_0^{(1)}$ and $C_0^{(2)} = \widetilde{B}_0^{(2)}$. Then using Remark 1 and Lemma 2 we get

$$\|y-XC\|_2^2+\lambda_1\|C\|_1+\lambda_2\sum_{i=1}^{p}|C_i^{(1)}-C_i^{(2)}| = \|y-X_2\widetilde{B}\|_2^2+\lambda\sum_{i=1}^{p}(b|\widetilde{B}_i^{(1)}|+b|\widetilde{B}_i^{(2)}|+a|\widetilde{B}_i^{(12)}|+c\lambda(|\widetilde{B}_0^{(1)}|,|\widetilde{B}_0^{(2)}|). \tag{10}$$

By definition of $\widetilde{D}$ and uniqueness of the LASSO solution we have :

$$\|y-XC\|_2^2+\lambda_1\|C\|_1+\lambda_2\sum_{i=1}^{p}|C_i^{(1)}-C_i^{(2)}| > \|y-X\widetilde{D}\|_2^2+\lambda_1\|\widetilde{D}\|_1+\lambda_2\sum_{i=1}^{p}|\widetilde{D}_i^{(1)}-\widetilde{D}_i^{(2)}|. \tag{11}$$

Using 10 we get that :

$$\|y-X_2\widetilde{B}\|_2^2+\lambda\sum_{i=1}^{p}(b|\widetilde{B}_i^{(1)}|+b|\widetilde{B}_i^{(2)}|+a|\widetilde{B}_i^{(12)}|+c\lambda(|\widetilde{B}_0^{(1)}|,|\widetilde{B}_0^{(2)}|) > |y-X\widetilde{D}\|_2^2+\lambda_1\|\widetilde{D}\|_1+\lambda_2\sum_{i=1}^{p}|\widetilde{D}_i^{(1)}-\widetilde{D}_i^{(2)}|. \tag{12}$$

Let's define E such as $\forall i$ $\widetilde{D}_i^{(1)} = E_i^{(1)} + E_i^{(12)}$, $\widetilde{D}_i^{(2)} = E_i^{(2)} + E_i^{(12)}$, $\widetilde{D}_0^{(1)} = E_0^{(1)}$, $\widetilde{D}_0^{(2)} = E_0^{(2)}$ and that respect Assumptions 7 and6 of 2.

Then using Remark 1 and Lemma 2 we get that

$$\|y-X\widetilde{D}\|_2^2+\lambda_1\|\widetilde{D}\|_1+\lambda_2\sum_{i=1}^{p}|\widetilde{D}_i^{(1)}-\widetilde{D}_i^{(2)}| = \|y-X_2E\|_2^2+\lambda\sum_{i=1}^{p}(b|E_i^{(1)}|+b|E_i^{(2)}|+a|E_i^{(12)}|). \tag{13}$$

Then by definition of $\widetilde{B}$

$$\|y-X_2E\|_2^2+\lambda\sum_{i=1}^{p}(b|E_i^{(1)}|+b|E_i^{(2)}|+a|E_i^{(12)}|)+c\lambda(|E_0^{(1)}|,|E_0^{(2)}|) > \|y-X_2\widetilde{B}\|_2^2+\lambda\sum_{i=1}^{p}(b|\widetilde{B}_i^{(1)}|+b|\widetilde{B}_i^{(2)}|+a|\widetilde{B}_i^{(12)}|+c\lambda(|\widetilde{B}_0^{(1)}|, \tag{14}$$

using 13 we get that :

$$\|y-X\widetilde{D}\|_2^2+\lambda_1\|\widetilde{D}\|_1+\lambda_2\sum_{i=1}^{p}|\widetilde{D}_i^{(1)}-\widetilde{D}_i^{(2)}| > \|y-X_2\widetilde{B}\|_2^2+\lambda\sum_{i=1}^{p}(b|\widetilde{B}_i^{(1)}|+b|\widetilde{B}_i^{(2)}|+a|\widetilde{B}_i^{(12)}|+c\lambda(|\widetilde{B}_0^{(1)}|,|\widetilde{B}_0^{(2)}|) \tag{15}$$

which is contradictory with 12, and it conclude the proof.

□

5 Lemmas and remarks

Remark 1. *Let us notice that for any $2(p+1)$ vector C and $3p+2$ vector B such that $C_0^{(1)} = B_0^{(1)}$, $C_0^{(2)} = B_0^{(2)}$, $C_i^{(1)} = B_i^{(1)} + B_i^{(12)}$ and $C_i^{(2)} = B_i^{(2)} + B_i^{(12)}$ we have $\|y - XC\|_2^2 = \|y - X_2B\|_2^2$.*

Lemma 2. *Let C be a $2(p + 1)$ vector. Let us call $P_1 = \lambda_1\|C\|_1 + \lambda_2\sum_{i=1}^{p}|C_i^{(1)} - C_i^{(2)}|$ and $P_2 = \lambda\sum_{i=1}^{p}(b|B_i^{(1)}| + b|B_i^{(2)}| + a|B_i^{(12)}|) + c\lambda(|B_0^{(1)}|, |B_0^{(2)}|)$. Under the following assumptions :*

1. $2b > a > b > 0$
2. $a = \frac{2\lambda_1}{\lambda}$
3. $b = \frac{(\lambda_1+\lambda_2)}{\lambda}$
4. $c = \frac{\lambda_1}{\lambda}$
5. $C_i^{(1)} = B_i^{(1)} + B_i^{(12)}$, $C_i^{(2)} = B_i^{(2)} + B_i^{(12)}$, $|B_0^{(1)}| = C_0^{(1)}|$ *and* $|B_0^{(2)}| = C_0^{(2)}|$
6. $\mathrm{sign}(\mathrm{B_i^{(1)}}) \neq \mathrm{sign}(\mathrm{B_i^{(2)}})$ *or both zero,*
7. *if $B_i^{(12)} \neq 0$ then $B_i^{(1)} = 0$ or $B_i^{(2)} = 0$ and the sign of the non zero one is equal to the sign of $B_i^{(12)}$ (or it is also zero),*

$P_1 = P_2$.

Proof of Lemma 2. Let's defined I_1 the set of regressors where $\forall i \in I_1, B_i^{(12)} \neq 0$ and $|B_i^{(2)}| < |B_i^{(1)}|$; I_2 the set of regressors where $\forall i \in I_2, B_i^{(12)} \neq 0$ and $|B_i^{(1)}| < |B_i^{(2)}|$and J such as $\forall i \in I, B_i^{(12)} = 0$. Then

$$\begin{aligned}
P_1 &= \lambda_1(|C_0^{(1)}| + |C_0^{(2)}|) + \sum_{i\in I_1}(\lambda_1(|C_i^{(1)}| + |C_i^{(2)}|) + \lambda_2|C_i^{(1)} - C_i^{(2)}|) \\
&\quad + \sum_{i\in I_2}(\lambda_1(|C_i^{(1)}| + |C_i^{(2)}|) + \lambda_2|C_i^{(1)} - C_i^{(2)}|) \\
&\quad + \sum_{i\in J}(\lambda_1(|C_i^{(1)}| + |C_i^{(2)}|) + \lambda_2|C_i^{(1)} - C_i^{(2)}|) \\
&= \lambda_1(|C_0^{(1)}| + |C_0^{(2)}|) + P_1^{I_1} + P_1^{I_2} + P_1^{J},
\end{aligned}$$

$$\begin{aligned}
P_2 &= c\lambda(|B_0^{(1)}| + |B_0^{(2)}|) + \lambda\sum_{i\in I_1}(b|B_i^{(1)}| + b|B_i^{(2)}| + a|B_i^{(12)}|) + \sum_{i\in I_2}(b|B_i^{(1)}| + b|B_i^{(2)}| + a|B_i^{(12)}|) \\
&\quad + \lambda\sum_{i\in J}(b|B_i^{(1)}| + b|B_i^{(2)}| + a|B_i^{(12)}|) \\
&= c\lambda(|B_0^{(1)}| + |B_0^{(2)}|) + P_2^{I_1} + P_2^{I_2} + P_2^{J}.
\end{aligned}$$

Proof that $c\lambda(|B_0^{(1)}| + |B_0^{(2)}|) = \lambda_1(|C_0^{(1)}| + |C_0^{(2)}|)$ It is straightforward using 4 and 5.

Proof that $P_1^{I_1} = P_2^{I_1}$
Using Assumptions 5 of Lemma 2 and 6 of Lemma 3 we get that

$$\begin{aligned}
P_1^{I_1} &= \sum_{i\in I_1}\lambda_1(|B_i^{(1)} + B_i^{(12)}| + |B_i^{(12)}|) + \lambda_2|B_i^{(1)}| \\
&= \sum_{i\in I_1}\lambda_1(|B_i^{(1)}| + |B_i^{(12)}| + |B_i^{(12)}|) + \lambda_2|B_i^{(1)}| \\
&= \sum_{i\in I_1}|B_i^{(1)}|(\lambda_1 + \lambda_2) + |B_i^{(12)}|2\lambda_1,
\end{aligned}$$

where the second equality come from $\forall i \in I_1 \mathrm{sign}(\mathrm{B_i^{(1)}}) = \mathrm{sign}(\mathrm{B_i^{(12)}})$ (Assumptions 7).

In this case $P_2^{I_1} = \sum_{i\in I_1}(\lambda b|B_i^{(1)}| + a\lambda|B_i^{(12)}|)$.

Thus, it came from assumptions 2 and3 of Lemma 2 that $P_1^{I_1} = P_2^{I_1}$.
Proof that $P_1^{I_2} = P_2^{I_2}$
By replacing $B_i^{(1)}$ by $B_i^{(2)}$ in the precedent calculus we get that $P_1^{I_2} = P_2^{I_2}$.
Proof that $P_1^{J} = P_2^{J}$

$$\begin{aligned}
P_1^{J} &= \sum_{i\in J}(\lambda_1(|C_i^{(1)}| + |C_i^{(2)}|) + \lambda_2|C_i^{(1)} - C_i^{(2)}|) \\
&= \sum_{i\in J}\lambda_1(|B_i^{(1)}| + |B_i^{(2)}|) + \lambda_2|B_i^{(1)} - B_i^{(2)}| \\
&= \sum_{i\in J}\lambda_1(|B_i^{(1)}| + |B_i^{(2)}|) + \lambda_2(|B_i^{(1)}| + |B_i^{(2)}|),
\end{aligned}$$

where the last equality come from the fact that $\mathrm{sign}(\mathrm{B_i^{(1)}}) \neq \mathrm{sign}(\mathrm{B_i^{(2)}})$.

In this case $P_2^{J} = \sum_{i\in J}\lambda b(|B_i^{(1)}| + |B_i^{(2)}|)$. Using 3 of Lemma 2 the equality is proven. □

Lemma 3. *For all* $i \in \{1, \ldots, p+1\}$ *if* $2b > a > b > 0$

1. $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{i}}^{(1)}) \neq \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{i}}^{(2)})$ *or both zero.*

2. If $\widetilde{B}_i^{(12)} \neq 0$ then $\widetilde{B}_i^{(1)} = 0$ or $\widetilde{B}_i^{(2)} = 0$ and the sign of the non zero one is equal to the sign of $\widetilde{B}_i^{(12)}$ (or it is also zero).

In the above statements $\widetilde{B} = \widetilde{B}_{\lambda,a,b}$ define in (8).

Proof of Lemma 3. Let A be a $2(p+1)+p$ vector such that $A = \widetilde{B}_{\lambda,a,b}$ except that $\exists j$ such that $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(1)}) = \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(2)}) \neq 0$, lets consider here that $|\widetilde{B}_j^{(1)}| \geq |\widetilde{B}_j^{(2)}|$ and $A_j^{(1)} = \widetilde{B}_j^{(1)} - \widetilde{B}_j^{(2)}$, $A_j^{(2)} = 0$ and $A_j^{(12)} = \widetilde{B}_j^{(12)} + \widetilde{B}_j^{(2)}$ then $\|y - X_2\widetilde{B}\|_2 = \|y - X_2 A\|_2$.

if $\widetilde{B}_j^{(2)} \neq 0$ by uniqueness of the Lasso solution:

$$\begin{aligned}
b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}| &< b|A_j^{(1)}| + b|A_j^{(2)}| + a|A_j^{(12)}| \\
&= b|\widetilde{B}_j^{(1)} - \widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)} + \widetilde{B}_j^{(2)}| \\
&\leq b(|\widetilde{B}_j^{(1)}| - |\widetilde{B}_j^{(2)}|) + a(|\widetilde{B}_j^{(12)}| + |\widetilde{B}_j^{(2)}|) \\
&= b(|\widetilde{B}_j^{(1)}|) + a(|\widetilde{B}_j^{(12)}|) + (b-a)|\widetilde{B}_j^{(2)}| \\
&< b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}|,
\end{aligned}$$

where the last inequality come from $2b > a$ and the second inequality from the fact that $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(1)}) = \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(2)})$ and $|\widetilde{B}_j^{(1)}| \geq |\widetilde{B}_j^{(2)}|$. As the first line is contradictory with the last line of the previous calculus there is no such j and so $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(1)}) \neq \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(2)})$ or both zero.

Let us now consider that $\exists j$ such that $\widetilde{B}_j^{(12)} \neq 0$ and both $\widetilde{B}_j^{(1)} \neq 0$ and $\widetilde{B}_j^{(2)} \neq 0$, using the first part of the proof we know that $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(1)}) \neq \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(2)})$ lets consider here that $\mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(1)}) = \mathrm{sign}(\widetilde{\mathrm{B}}_{\mathrm{j}}^{(12)})$ and

First scenario $|\widetilde{B}_j^{(2)}| \leq |\widetilde{B}_j^{(12)}|$

$A_j^{(1)} = \widetilde{B}_j^{(1)} - \widetilde{B}_j^{(2)}$, $A_j^{(2)} = 0$ and $A_j^{(12)} = \widetilde{B}_j^{(12)} + \widetilde{B}_j^{(2)}$ then $\|y - X_2\widetilde{B}\|_2 = \|y - X_2 A\|_2$. Let us first notice that $\mathrm{sign}(\mathrm{A}_{\mathrm{j}}^{(1)}) = \mathrm{sign}(\mathrm{A}_{\mathrm{j}}^{(12)})$. As $\widetilde{B}_j^{(2)} \neq 0$ by uniqueness of the Lasso solution

$$\begin{aligned}
b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}| &< b|A_j^{(1)}| + b|A_j^{(2)}| + a|A_j^{(12)}| \\
&= b|\widetilde{B}_j^{(1)} - \widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)} + \widetilde{B}_j^{(2)}| \\
&\leq b(|\widetilde{B}_j^{(1)}| + |\widetilde{B}_j^{(2)}|) + a(|\widetilde{B}_j^{(12)}| - |\widetilde{B}_j^{(2)}|) \\
&= b(|\widetilde{B}_j^{(1)}|) + a(|\widetilde{B}_j^{(12)}|) + (b-a)|\widetilde{B}_j^{(2)}| \\
&< b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}|,
\end{aligned}$$

where the second inequality come from $\mathrm{sign}(\mathrm{B}_{\mathrm{j}}^{(2)}) \neq \mathrm{sign}(\mathrm{B}_{\mathrm{j}}^{(12)})$ and $|\widetilde{B}_j^{(2)}| \leq |\widetilde{B}_j^{(12)}|$. As the first line is contradictory with the last line of the previous calculus there is no such j.

Second scenario $|\widetilde{B}_j^{(2)}| > |\widetilde{B}_j^{(12)}|$

$A_j^{(1)} = \widetilde{B}_j^{(1)} + \widetilde{B}_j^{(12)}$, $A_j^{(2)} = \widetilde{B}_j^{(2)} + \widetilde{B}_j^{(12)}$ and $A_j^{(12)} = 0$ then $\|y - X_2\widetilde{B}\|_2 = \|y - X_2 A\|_2$.

As $\widetilde{B}_j^{(12)} \neq 0$ by uniqueness of the Lasso solution :

$$\begin{aligned}
b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}| &< b|A_j^{(1)}| + b|A_j^{(2)}| + a|A_j^{(12)}| \\
&= b|\widetilde{B}_j^{(1)} + \widetilde{B}_j^{(12)}| + a|\widetilde{B}_j^{(12)} + \widetilde{B}_j^{(2)}| \\
&\leq b(|\widetilde{B}_j^{(1)}| + |\widetilde{B}_j^{(12)}|) + b(|\widetilde{B}_j^{(2)}| - |\widetilde{B}_j^{(12)}|) \\
&= b(|\widetilde{B}_j^{(1)}|) + b|\widetilde{B}_j^{(2)}| \\
&< b|\widetilde{B}_j^{(1)}| + b|\widetilde{B}_j^{(2)}| + a|\widetilde{B}_j^{(12)}|.
\end{aligned}$$

As the first line is contradictory with the last line of the previous calculus there is no such j.

More over using the fact that either $|\widetilde{B}_j^{(2)}| > |\widetilde{B}_j^{(12)}|$ or $|\widetilde{B}_j^{(2)}| \leq |\widetilde{B}_j^{(12)}|$ and the fact that we can do exactly the same calculus if $\mathrm{sign}(\widetilde{\mathrm{B}}_\mathrm{j}^{(2)}) = \mathrm{sign}(\widetilde{\mathrm{B}}_\mathrm{j}^{(12)})$ we prove that there is no j such as $\widetilde{B}_i^{(12)} \neq 0$ and both $\widetilde{B}_i^{(1)} \neq 0$ and $\widetilde{B}_i^{(2)} \neq 0$ and that If $\widetilde{B}_i^{(12)} \neq 0$ then $\widetilde{B}_i^{(1)} = 0$ or $\widetilde{B}_i^{(2)} = 0$ and the sign of the non zero one is equal to the sign of $\widetilde{B}_i^{(12)}$ (or it is also zero).

$\square$