6. $X\beta(\lambda) = X_{A_2}(X_{A_2}^T X_{A_2})^{-1}(X_{A_2}^T Y - n\lambda t_{A_2})$

$= X_{A_2}(X_{A_2}^T X_{A_2})^{-1} X_{A_2}^T Y - n\lambda \underbrace{X_{A_2}(X_{A_2}^T X_{A_2})^{-1} t_{A_2}}_{u_{A_2}}$

$X_{A_2} = [X^{j_1}\ X^{j_2}]$

$\Rightarrow (X_{A_2}^T X_{A_2})^{-1} X_{A_2}^T Y = \hat{\beta}_{A_2}$

7. $X_{A_2}^T u_{A_2} = X_{A_2}^T X_{A_2}(X_{A_2}^T X_{A_2})^{-1} t_{A_2} = t_{A_2} \leftarrow$ vecteur de signe (donc de $+/- 1$)

8. $\|Y - X\beta(\lambda)\|^2 = \|Y - X_{A_2}\hat{\beta}_{A_2}\|^2 + \|n\lambda u_{A_2}\|^2 - 2\langle Y - X_{A_2}\hat{\beta}_{A_2}, n\lambda u_{A_2}\rangle$

mais $Y - X_{A_2}\hat{\beta}_{A_2} \in \mathrm{Vect}(X_{A_2})^{\perp}$ par définition de $\hat{\beta}_{A_2}$.

et $u_{A_2} = X_{A_2}(X_{A_2}^T X_{A_2})^{-1} t_{A_2}$ est une CL des colonnes de $X_{A_2} \Rightarrow \in \mathrm{Vect}(X_{A_2})$

On a $X_{A_2}^T(Y - X\beta(\lambda)) = X_{A_2}^T(Y - X_{A_2}(X_{A_2}^T X_{A_2})^{-1} X_{A_2}^T Y) + X_{A_2}^T n\lambda u_{A_2}$

$= 0 + n\lambda X_{A_2}^T u_{A_2}$

même chose que 8)

$= +n\lambda t_{A_2}.$

9. En λ_2 $\quad \beta_{j_1}(\lambda_2) =$

$\beta_{j_2}(\lambda_2) = 0$

En λ_2^+ $\quad \beta^{(\lambda_2)}_{[j_1, j_2]} = (X_{A_2}^T X_{A_2})^{-1}(X_{A_2}^T Y - n\lambda_2 t_{A_2})$

si $t_{A_2} = \begin{pmatrix}1\\1\end{pmatrix}$

$t_{A_1} = 1$

cf questions 5

$= (X_{A_2}^T X_{A_2})^{-1}\left(X_{A_2}^T Y - n t_{A_2} \frac{1}{n} \frac{X^{j_2 T}(Y - \frac{1}{n}X^{j_1}X^{j_1 T}Y)}{1 - \frac{1}{n}t_{A_1} X^{j_2 T}X^{j_1}}\right)$

$= (X_{A_2}^T X_{A_2})^{-1} \frac{1}{1 - \frac{1}{n}t_{A_1} X^{j_2 T} X^{j_1}}\left(\begin{pmatrix}X^{j_1 T}Y\\ X^{j_2 T}Y\end{pmatrix} - \frac{1}{n}t_{A_1}\begin{pmatrix}X^{j_2 T}X^{j_1}X^{j_1 T}Y\\ X^{j_2 T}X^{j_1}X^{j_1 T}Y\end{pmatrix} - \begin{pmatrix}\cdots\\ X^{j_2 T}Y - \frac{1}{n}X^{j_2 T}X^{j_1}X^{j_1 T}Y\end{pmatrix}\right)$

$= (X_{A_2}^T X_{A_2})^{-1} \frac{1}{-} \begin{pmatrix} \\ 0 \end{pmatrix}$

c'est bien continu en λ_2

Examen 2016-2017

Exercice 2:

1) $$\widehat{\mathrm{corr}}(Y, X^j) = \frac{\frac{1}{n}\sum (Y_i - \bar{Y})(X_i^j - \bar{X}^j)}{\sqrt{\frac{1}{n}\sum (Y_i - \bar{Y})^2 \frac{1}{n}\sum (X_i^j - \bar{X}^j)^2}} \underset{\text{variables centrées}}{=} \frac{\frac{1}{n}\sum (Y_i X_i^j)}{\sqrt{\frac{1}{n}\sum Y_i^2 \frac{1}{n}\sum X_i^{j\,2}}}$$

par définition

$$= \frac{\frac{1}{n} Y^T X^j}{\sqrt{\frac{1}{n}\sum Y_i^2}} \propto \frac{1}{n}|Y^T X^j|$$

2) Si $X^T X = Id$, on sait cf cours, que la solution du problème du lasso est donnée par $\hat{\beta}(\lambda) = S_{\lambda n}(Y^T X)$

soit $\hat{\beta}^j(\lambda) = \mathrm{sign}(Y^T X^j)\left(|Y^T X^j| - \lambda n\right)_+$

$$= n\,\mathrm{sign}(Y^T X^j)\left(\frac{1}{n}|Y^T X^j| - \lambda\right)_+$$

Pour tout $\lambda \geq \lambda_1 = \frac{1}{n}\max_j |Y^T X^j|$ $\quad \hat{\beta}^j(\lambda) = 0$.

3) $\lambda \in]\lambda_1, \lambda_2] \to [\beta(\lambda)]_{j_1} = (X_{A_1}^T X_{A_1})^{-1}(X_{A_1}^T Y - n\lambda t_{A_1})$ linéaire donc continue en λ.

Il reste à vérifier que la fonction est bien continue en $\lambda = \lambda_1$. Elle vaut 0 en λ_1 (cf question précédente)

En λ_1^+, elle vaut $(X_{A_1}^T X_{A_1})^{-1}(X_{A_1}^T Y - n\lambda_1^+ \,\mathrm{signe}(X_{A_1}^T Y))$

$$= (X^{j_1 T} X^{j_1})^{-1}\left(X^{j_1 T} Y - n\frac{1}{n}|Y^T X^{j_1}|\,\mathrm{signe}(X^{j T} Y)\right) = 0$$

par définition de λ_1

Pour $k \neq j_1$, $\lambda \to [\beta(\lambda)]_k$ est constante et nulle

4) $$X^{j_1 T}(Y - X\beta(\lambda)) = X^{j_1 T}(Y - X_{j_1}(\beta(\lambda))_{j_1}) \quad \text{car } \left((\beta(\lambda))_k = 0 \;\; k \neq j_1\right)$$

$$= X^{j_1 T}\left(Y - X_{j_1}(X_{j_1}^T X_{j_1})^{-1}(X^{j_1 T} Y - n\lambda t_{A_1})\right)$$

$$= X^{j_1 T} Y - X^{j_1 T} X_{j_1}(X_{j_1}^T X_{j_1})^{-1} X^{j_1 T} Y + n\lambda t_{A_1} = n\lambda t_{A_1}.$$

5) $X^{j_1 T}(Y - X\beta(\lambda_2)) = n\lambda_2 t_{A_1} \geq 0 \quad t_{A_1} \geq 0$

$$X^{j_2 T}(Y - X\beta(\lambda_2)) = X^{j_2 T} Y - X^{j_2 T} X^{j_1}(X^{j_1 T} X^{j_1})^{-1} X^{j_1 T} Y + \frac{1}{n} X^{j_2 T} X^{j_1}\, n\lambda_2 t_{A_1} = n\lambda_2$$

$$\to \lambda_2 = \frac{1}{n}\,\frac{X^{j_2 T} Y - X^{j_2 T} X^{j_1}(X^{j_1 T} X^{j_1})^{-1} X^{j_1 T} Y}{(1 \mp \frac{1}{n} t_{A_1} X^{j_2 T} X^{j_1})}$$

$$= \frac{1}{n}\,\frac{X^{j_2 T}(Y - X^{j_1}(X^{j_1 T} X^{j_1})^{-1} X^{j_1 T} Y)}{1 \mp X^{j_2 T} X^{j_1} \frac{t_{A_1}}{n}} = \frac{1}{n}\,\frac{X^{j_2 T}(Y - \frac{1}{n} X^{j_1} X^{j_1 T} Y)}{1 - \frac{1}{n} t_{A_1} X^{j_2 T} X^{j_1}} = \frac{\frac{1}{n} X^{j_2 T}(Y - \frac{1}{n} X^{j_1} X^{j_1 T} Y)}{1 - \frac{1}{n} X^{j_2 T} X^{j_1}}$$

$t_{A_1} \geq 0$.