(1)

(a) $\bar{y} = \mu + \bar{x}_1\beta_1 + \text{———} \bar{x}_p\beta_p$

$\Leftrightarrow \mu = \bar{y} - \bar{x}_1\beta_1 - \dots - \bar{x}_p\beta_p$

(b) $y_i = (\bar{y} - \bar{x}_1\beta_1 - \dots - \bar{x}_p\beta_p) + x_{i1}\beta_1 + \dots + x_{1p}\beta_p + \varepsilon_i$

$\Leftrightarrow (y_i - \bar{y}) = (x_{i1} - \bar{x})\beta_1 + \dots + (x_{ip} - \bar{x}_p) + \varepsilon_i$

$\Leftrightarrow \tilde{y} = (x_{1i} - \bar{x})\beta_1 + \dots + (x_{ip} - \bar{x}_p) + \varepsilon_i$

(c) $\pi(\beta,\sigma^2) = \pi(\beta|\sigma^2)\pi(\sigma^2) = \pi(\sigma^2)\left(\frac{\lambda}{2\sigma}\right)^p e^{-\frac{\lambda}{\sigma}\sum_{i=1}^p|\beta_i|} = \pi(\sigma^2)\left(\frac{\lambda}{2\sigma}\right)^p e^{-\frac{\lambda}{\sigma}\|\beta\|_1}$

(d) $L(\tilde{y};\beta,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma^{n-1}} \exp\left(-\frac{1}{2}{}^t(\tilde{y}-X\beta)(\sigma^2 I_n)^{-1}(\tilde{y}-X\beta)\right)$

y a un -1 ici mais je ne trouve pas pourquoi

$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma^{n-1}} \exp\left(-\frac{1}{2\sigma^2}\|\tilde{y}-X\beta\|_2^2\right)$

(e) $\pi(\beta,\sigma^2;\tilde{y}) \propto L(\tilde{y};\beta,\sigma^2)\pi(\beta,\sigma^2)$

$\propto \left[\frac{1}{\sigma^n}\exp\left(-\frac{1}{2\sigma^2}\|\tilde{y}-X\beta\|_2^2\right)\right]\left[\pi(\sigma^2)\cancel{\left(\frac{\lambda}{2}\right)^p}\left(\frac{1}{\sigma}\right)^p e^{-\frac{\lambda}{\sigma}\|\beta\|_1}\right]$

$\propto \pi(\sigma^2)\frac{1}{\sigma^{n+p-1}}\exp\left(-\frac{1}{2\sigma^2}\|\tilde{y}-X\beta\|_2^2 - \frac{\lambda}{\sigma}\|\beta\|_1\right)$

?

(f) $\ln(\pi(\beta,\sigma^2;\tilde{y})) \propto \ln(\pi(\sigma^2)) - \frac{n+p-1}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\|\tilde{y}-X\beta\|_2^2 - \frac{\lambda}{\sqrt{\sigma^2}}\|\beta\|_1$

$\propto \ln(\pi(\sigma^2)) - \frac{n+p-1}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\left(\|\tilde{y}-X\beta\|_2^2 - 2\sqrt{\sigma^2}\lambda\|\beta\|_1\right)$

CRS ENSAE

$\hat{\beta}_{MAP} \in \underset{\beta}{\text{argmax}}\ \pi(\beta,\sigma^2;\tilde{y})$

$\in \text{argmax} \ln \pi(\beta,\sigma^2;\tilde{y})$ i.e $\in \text{argmax}(\ln L(\tilde{y};\beta,\sigma^2) + \ln\pi(\beta,\sigma^2))$

$\in$ argmax (log vraissemblance + penalisation)

La loi à priori peut être vu comme une pénalisation
Le bayesien est ainsi un moyen de donner de la régularisation
via des contraintes 'douces'

NB avec $\mathcal{L}$ fct de coût uniforme, l'estimateur bayesien assoc à $\mathcal{L}$ i.e qui minimise $E^\pi(\mathcal{L}(\beta,d)|\tilde{y})$ est le MAP

(9) $$\pi(\beta|\sigma^2) = \int_{\mathbb{R}_+^p} \pi(\beta|\sigma^2, s_1 \ldots s_p)\, dP_{(\tau_1^2 \ldots \tau_p^2)}(s_1 \ldots s_p)$$

$$= \int_{\mathbb{R}_+^p} \left[\left(\frac{1}{\sqrt{2\pi}}\right)^p \frac{1}{\sqrt{\det(\sigma^2 D_\tau)}} \exp\left(-\frac{1}{2}\, {}^t\beta (\sigma^2 D_\tau)^{-1} \beta\right) \times \prod_{j=1}^p \frac{\lambda^2}{2} \exp\left(-\frac{\lambda^2}{2}\tau_j^2\right)\right] d\tau_j^2$$

$$= \int_{\mathbb{R}_+^p} \left(\frac{1}{\sqrt{2\pi}}\right)^p \left(\frac{1}{\sqrt{\sigma^2}}\right)^p \frac{1}{\prod_{j=1}^p \sqrt{\tau_j^2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^p \beta_j^2 \frac{1}{\tau_j^2}\right) \times \prod_{j=1}^p \frac{\lambda^2}{2} \exp\left(-\frac{\lambda^2}{2}\tau_j^2\right) d\tau_j^2$$

$$= \prod_{j=1}^p \left(\int_{\mathbb{R}_+} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}\, e^{-\frac{1}{2\sigma^2}\beta_j^2 \frac{1}{\tau_j^2}}\, \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2} d\tau_j^2 \right)$$

En passant par la f° génératrice des moments,

$$M_{\beta_j|\sigma^2}(t) = \int_{\mathbb{R}} e^{t\beta_j}\, \pi(\beta_j|\sigma^2)\, d\beta_j$$

$$= \int_{\mathbb{R}} e^{t\beta_j} \int_{\mathbb{R}_+} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}\, e^{-\left(\frac{1}{2\sigma^2}\beta_j^2\frac{1}{\tau_j^2}\right)} \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2}\, d\tau_j^2\, d\beta_j$$

$$= \int_{\mathbb{R}_+} \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2} \int_{\mathbb{R}} e^{t\beta_j} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}\, e^{-\frac{1}{2\sigma^2\tau_j^2}\beta_j^2}\, d\beta_j\, d\tau_j^2$$

$$= \int_{\mathbb{R}_+} \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2} \left(\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}}\, e^{-\frac{1}{2\sigma^2\tau_j^2}(\beta_j - \sigma^2\tau_j^2 t)^2} d\beta_j \right) e^{\frac{\sigma^2\tau_j^2 t^2}{2}} d\tau_j^2$$

densité loi N intégrée sur IR

$$= \int_{\mathbb{R}_+} \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2}\, e^{\frac{\sigma^2\tau_j^2 t^2}{2}} d\tau_j^2 = \int_{\mathbb{R}_+} \frac{\lambda^2}{2} e^{\left(-\frac{\lambda^2}{2} + \frac{\sigma^2 t^2}{2}\right)\tau_j^2} d\tau_j^2$$

$$= \left[\frac{\lambda^2}{2} \left(\frac{1}{-\frac{\lambda^2}{2} + \frac{\sigma^2 t^2}{2}} \right) e^{\left(-\frac{\lambda^2}{2} + \frac{\sigma^2 t^2}{2}\right)\tau_j^2} \right]_0^\infty$$

$$= -\frac{\lambda^2}{2}\left(\frac{1}{-\frac{\lambda^2}{2} + \frac{\sigma^2 t^2}{2}}\right) = \frac{1}{1 - \frac{\sigma^2}{\lambda^2} t^2}$$

(3)

On reconnait la fn generatrice des mmts de la loi laplace $\mathcal{L}(0, \frac{\sqrt{\sigma^2}}{\lambda})$.

Nous avons ainsi montré que,

$$(h) \quad \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_j|/\sqrt{\sigma^2}} = \int_{\mathbb{R}_+} \frac{1}{\sqrt{2\pi\sigma^2\tau_j^2}} e^{-\left(\frac{1}{2\sigma^2}\beta_j^2 \frac{1}{\tau_j^2}\right)} \frac{\lambda^2}{2} e^{-\frac{\lambda^2}{2}\tau_j^2} d\tau_j^2$$

i.e la distribution de laplace peut s'exprimer en tant que melange de gaussienne (avec une densite de melange exponentielle)

et que, $\pi(\beta|\sigma^2)$ est de la forme (3) désirée.

(i)

$$\pi(\beta, \sigma^2, \tau_1^2 \_\_ \tau_p^2 \mid \tilde{y}) \propto L(\tilde{y} \mid \beta, \sigma^2, \tau_1^2 \_\_ \tau_p^2) \times \pi(\beta \mid \sigma^2, \tau_1^2 \_\_ \tau_p^2) \times \pi(\sigma^2, \tau_1^2 \_\_ \tau_p^2)$$

$$\propto \frac{1}{\sigma^{n-1}} e^{-\frac{1}{2\sigma^2}(\tilde{y}-X\beta)^T(\tilde{y}-X\beta)} \times \frac{1}{\sigma^p \sqrt{\det D_\tau}} e^{-\frac{1}{2}\beta^T(\sigma^2 D_\tau)^{-1}\beta} \times \pi(\sigma^2) \prod_{j=1}^{p} \cancel{\frac{\lambda^2}{2}} \times e^{-\frac{\lambda^2}{2}\tau_j^2}$$

$$\propto \frac{1}{\sigma^{n+p-1}} \frac{1}{\sqrt{\det D_\tau}} \pi(\sigma^2) \left(\cancel{\frac{\lambda^2}{2}}\right)^p e^{-\frac{\lambda^2}{2}\|\tau_j^2\|_2^2} \exp\left[-\frac{1}{2\sigma^2}\left((\tilde{y}-X\beta)^T(\tilde{y}-X\beta) - \beta^T D_\tau^{-1}\beta\right)\right]$$

- En intégrant par rapport à β, on obtient,

$$\pi(\sigma^2, \tau_1^2 \_\_ \tau_p^2 \mid \tilde{y}) = \int_{\mathbb{R}^p} \pi(\beta, \sigma^2, \tau_1^2 \_\_ \tau_p^2 \mid \tilde{y})\, d\beta$$

$$\propto \frac{1}{\sigma^{n+p-1}} \frac{1}{\sqrt{\det D_\tau}} \pi(\sigma^2) e^{-\frac{\lambda^2}{2}\|\tau_j^2\|_2^2}$$

$$\times \int_{\mathbb{R}^p} \exp\Big(-\frac{1}{2\sigma^2}\big(\underbrace{\tilde{y}^T\tilde{y} - \tilde{y}^T X\beta - (X\beta)^T\tilde{y} + (X\beta)^T X\beta - \beta^T D_\tau^{-1}\beta}_{\tilde{y}^T\tilde{y} - 2\tilde{y}^T(X\beta) + \underbrace{\beta^T X^T X\beta + \beta^T D_\tau^{-1}\beta}_{\beta^T(X^TX + D_\tau^{-1})\beta}}\big)$$

On a $X \sim \mathcal{N}(\mu, \Sigma)$ alors $f_X(x) \propto \exp\left((X-\mu)^T \Sigma^{-1}(X-\mu)\right)$

$$f_X(x) \propto \exp\left(X^T\Sigma^{-1}X + X^T\Sigma^{-1}\mu - \mu^T\Sigma^{-1}X + \mu^T\Sigma^{-1}\mu\right)$$

$$\propto \exp\left(X^T\Sigma^{-1}X - 2\mu^T\Sigma^{-1}X + \mu^T\Sigma^{-1}\mu\right)$$

D'où par identification on reconnait la densité d'une loi normale $\mathcal{N}(A^{-1}X^T\tilde{y}, \sigma^2 A^{-1})$ avec $A = X^TX + D_\tau^{-1}$

Ainsi $\pi(\sigma^2, \tau_1^2 \_\_ \tau_p^2 \mid \tilde{y}) \propto \dfrac{1}{\sigma^{n+p-1}} \dfrac{1}{\sqrt{\det D_\tau}} \pi(\sigma^2) e^{-\frac{\lambda^2}{2}\|\tau_j^2\|_2^2}$

→ On a alors $\pi(\beta \mid \sigma^2, \tau_1^2 \_\_\_ \tau_p^2, \tilde{y}) = \dfrac{\pi(\beta, \sigma^2, \tau_1^2 \_\_\_ \tau_p^2 \mid \tilde{y})}{\pi(\sigma^2, \tau_1^2 \_\_\_ \tau_p^2 \mid \tilde{y})}$

$$\propto f_{\mathcal{N}}(A^{-1}X^T\tilde{y}, \sigma^2 A^{-1})$$

• En intégrant par rapport à σ^2, on obtient

$$\pi(\beta, \tau_1^2 \_\_\_ \tau_p^2 \mid \tilde{y}) \propto \frac{1}{\sqrt{\det D_\tau}} e^{-\frac{\lambda^2}{2}\|\tau_j^2\|_2^2}$$

$$\times \int_{\mathbb{R}} \frac{1}{\sigma^{n+p-1}} \pi(\sigma^2) \exp\left(-\left[\frac{(\tilde{y}-X\beta)^T(\tilde{y}-X\beta) - \beta^T D_\tau^{-1}\beta}{2}\right]\frac{1}{\sigma^2}\right) d\sigma^2$$

$$\propto \frac{1}{\sqrt{\det D_\tau}} e^{-\frac{\lambda^2}{2}\|\tau_j^2\|_2^2} \times 1$$

en reconnaissant la densité d'une loi $IG\left(\frac{n+p-1}{2}, \frac{(\tilde{y}-X\beta)^T(\tilde{y}-X\beta)}{2} + \frac{\beta^T D_\tau^{-1}\beta}{2}\right)$

→ d'où $\pi(\sigma^2 \mid \beta, \tau_1^2 \_\_\_ \tau_p^2, \tilde{y}) = \dfrac{\pi(\beta, \sigma^2, \tau_1^2 \_\_\_ \tau_p^2 \mid \tilde{y})}{\pi(\beta, \tau_1^2 \_\_\_ \tau_p^2 \mid \tilde{y})}$

$$\propto f_{IG_{amma}}(\cdot, \cdot)$$

• Enfin, en intégrant par rapport à $1/\tau_j^2$, on obtient

[_ _ _] compléter

(j) $\beta | \sigma^2, \tau_1^2 - \tau_p^2, \tilde{y} \sim \mathcal{N}(A^{-1}X^T\tilde{y}, \sigma^2 A^{-1})$ avec $A = X^TX + D_\tau^{-1}$

$\sigma^2 | \beta, \tau_1^2 - \tau_p^2, \tilde{y} \sim \underset{\text{gamma}}{IG}\left(\frac{n+p-1}{2}, \frac{(\tilde{y}-X\beta)^T(\tilde{y}-X\beta)}{2} + \frac{\beta^T D_\tau^{-1}\beta}{2}\right)$

$\frac{1}{\tau_j^2} | \beta, \sigma^2, \tilde{y} \sim \text{IGaussian}(\mu', \lambda')$ avec $\begin{cases} \mu' = \sqrt{\frac{\lambda^2\sigma^2}{\beta_j^2}} \\ \lambda' = \lambda^2 \end{cases}$

de densité $f(x) = \sqrt{\frac{\lambda'}{2\pi}}\, x^{-3/2} \exp\left(-\frac{\lambda'(x-\mu')^2}{2(\mu')^2 x}\right), x > 0$

(k) sachant (β_t, σ_t^2), générer

~~① $\tau_1^2 - \tau_p^2$ iid de loi $\frac{1}{\tau_i^2} \sim \text{IGaussian}(\mu_t', \lambda_t')$~~

①
$\forall j \in \{1 - p\}$, $\frac{1}{\tau_{j,t+1}^2} | \beta_t, \sigma_t^2, \tilde{y} \sim \text{IGaussian}\left(\sqrt{\frac{\lambda^2\sigma_t^2}{\beta_{jt}^2}}, \lambda^2\right)$

② $\sigma_{t+1}^2 | \beta_t, \tau_{1,t+1}^2, \dots, \tau_{p,t+1}^2, \tilde{y} \sim \text{IGamma}\left(\frac{n+p-1}{2}, \frac{(\tilde{y}-X\beta_t)^T(\tilde{y}-X\beta_t)}{2} + \frac{\beta_t^T D_{\tau,t+1}^{-1}\beta_t}{2}\right)$

③ $\beta_{t+1} | \sigma_{t+1}^2, \tau_{1,t+1}^2, \dots, \tau_{p,t+1}^2, \tilde{y} \sim \mathcal{N}(A_{t+1}^{-1} X^T\tilde{y}, \sigma_{t+1}^2 A_{t+1}^{-1})$

À l'étape 2

(l) • la log vraissemblance des données complétées pour λ est

$$\ln\left[\pi(\beta, \sigma^2, \tau_1^2 \_\_\_ \tau_p^2, \lambda \mid \tilde{y})\right]$$

$$\propto \ln\left[L(\tilde{y} \mid \beta, \sigma^2, \tau_1^2 \_\_\_ \tau_p^2)\right] + \ln\left[\pi(\beta \mid \sigma^2, \tau_1^2 \_\_\_ \tau_p^2)\right]$$

$$\ln\left[\pi(\sigma^2, \tau_1^2 \_\_\_ \tau_p^2, \lambda)\right]$$

$$\propto \ln\left[\pi(\sigma^2, \tau_1^2 \_\_\_ \tau_p^2)\right]$$

en négligeant les termes additifs ne dependant pas de λ (en effet $L(\tilde{y} \mid \beta, \sigma^2, \tau_1^2 \_\_\_ \tau_p^2)$ et $\pi(\beta \mid \sigma^2, \tau_1^2 \_\_\_ \tau_p^2)$ ne dépendent pas de λ)

$$\propto p\ln(\lambda^2) - \frac{\lambda^2}{2}\sum_{j=1}^{p}\tau_j^2$$

• l'esperance de la log vraissemblance des données complétées pour λ, conditT à $\lambda^{(k)}$ notée $\varphi(\lambda, \lambda^{(k)})$ est alors,

$$\varphi(\lambda, \lambda^{(k)}) \propto E_{\lambda^{(k)}}\left(p\ln(\lambda^2) - \frac{\lambda^2}{2}\sum_{j=1}^{p}\tau_j^2\right)$$

$$\propto p\ln(\lambda^2) - \frac{\lambda^2}{2}\sum_{j=1}^{p} E_{\lambda^{(k)}}\left[\tau_j^2 \mid \tilde{y}\right]$$

À l'étape 3

On obtient alors, $\lambda^{(k+1)} = \underset{\lambda}{\operatorname{argmax}}\left(\varphi(\lambda, \lambda^{(k)})\right)$

$$= \underset{\lambda}{\operatorname{argmax}}\left(2p\ln(\lambda) - \frac{\lambda^2}{2}\sum_{j=1}^{p} E_{\lambda^{(k)}}\left[\tau_j^2 \mid \tilde{y}\right]\right)$$

(CPO) $\quad 2p\frac{1}{\lambda} - \lambda\sum_{j=1}^{p} E_{\lambda^{(k)}}(\tau_j^2 \mid \tilde{y}) = 0$

$$\Leftrightarrow 2p\frac{1}{\lambda} = \lambda\sum_{j=1}^{p} E_{\lambda^{(k)}}(\tau_j^2 \mid \tilde{y})$$

$$\Rightarrow \quad \lambda = \sqrt{\frac{2p}{\sum_{j=1}^{p} E_{\lambda^{(k)}}(\tau_j^2 | \tilde{y})}}$$

On montre aisément que $\varphi(\lambda, \lambda^{(k)})$ est concave

D'où $\lambda^{(k+1)} = \sqrt{\frac{2p}{\sum_{j=1}^{p} E_{\lambda^{(k)}}(\tau_j^2 | \tilde{y})}}$

Ainsi, l'étape 2 revient à estimer $E_{\lambda^{(k)}}(\tau_j^2 | \tilde{y})$
avec l'échantillon du Gibbs sampling
et l'étape 3 revient à calculer $\lambda^{(k+1)} = \sqrt{\frac{2p}{\sum_{j=1}^{p} E_{\lambda^{(k)}}(\tau_j^2 | \tilde{y})}}$

(m) la loi à postériori devient

$$\pi(\beta, \sigma^2, \tau_1^2 \_\_ \tau_p^2, \lambda^2 | \tilde{y})$$

$$\propto \pi(\sigma^2, \tau_1^2 \_\_ \tau_p^2, \lambda^2)$$ (en négligeant les termes ne dépendant pas de λ)

$$\propto \pi(\sigma^2, \tau_1^2 \_\_ \tau_p^2 | \lambda^2)\, \pi(\sigma^2)$$

$$\propto \prod_{j=1}^{p} \frac{\lambda^2}{2} e^{-\frac{\lambda^2 \tau_j^2}{2}} \times \frac{\delta^r}{\Gamma(r)} (\lambda^2)^{r-1} e^{-\delta \lambda^2}$$

$$\propto (\lambda^2)^{p+r-1} \exp\left(-\lambda^2 \left(\tfrac{1}{2} \sum_{j=1}^{p} \tau_j^2 + \delta\right)\right)$$

On en déduit, par des calculs similaires à ceux effectués pour obtenir les lois conditionnelles (10)

que $\lambda^2 | \beta, \tau_1^2 \_\_ \tau_p^2, \sigma^2, \tilde{y} \sim \text{Gamma}\left(p+r, \sum_{i=1}^{p} \frac{\tau_i^2}{2} + \delta\right)$