Statistiques bayésiennes

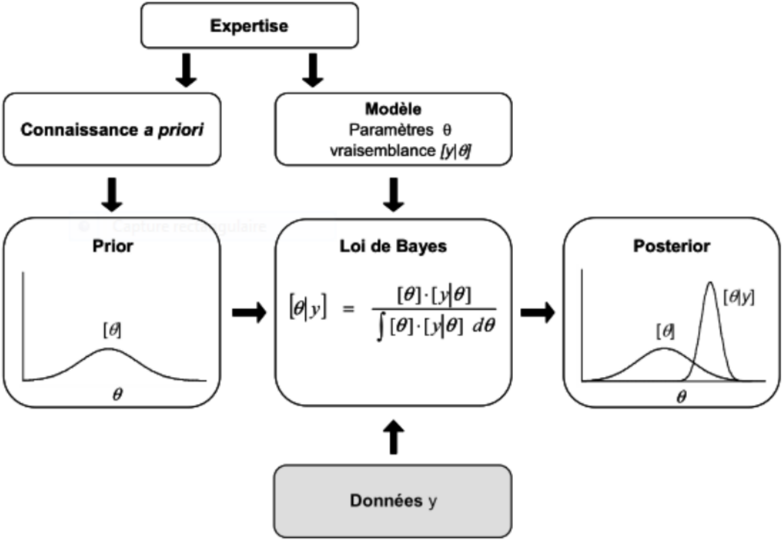
Othmane Ettaqi et Henry Watiotienne
Encadré par Mme Ophélie Guin

Département de Mathématiques
Université Lille

30 mai 2023

Plan de la Présentation

Principe général



Modèle gaussien à moyenne et variance inconnus

On considère le modèle bayésien sera le suivant :

$x_i|\mu,\tau \ \sim \mathcal{N}(\mu,\tau)$ *iid.* $i = 1, ..., n$

$\mu|\tau \sim \mathcal{N}(\mu_0, n_0\tau)$

$\tau \sim \mathcal{G}(\alpha,\beta)$

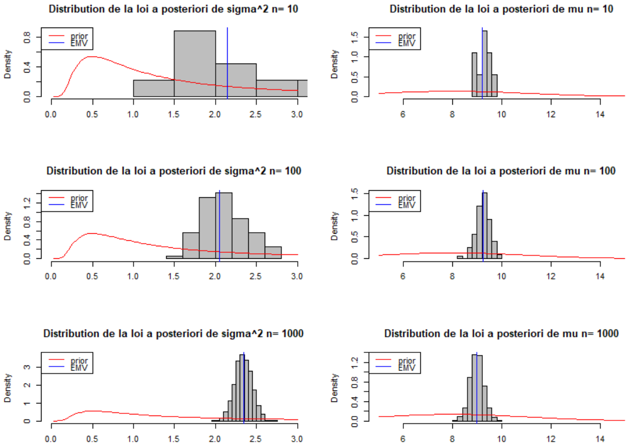
On trouve par le calcul les lois a posteriori suivantes :

$$\mu|x_1,..,x_n,\tau \sim \mathcal{N}\left(\frac{n}{n+n_0}\overline{x} + \frac{n_0}{n+n_0}\mu_0, \frac{1}{(n+n_0)\tau}\right)$$

$$\tau|x_1,..,x_n \sim \mathcal{G}\left(\alpha + \frac{n}{2}, \frac{\sum(x_i-\overline{x})^2}{2} + \frac{nn_0}{n+n_0}(\mu_0\overline{x})^2\right)$$

Illustration avec un jeu de données

Ici $\tau = \dfrac{1}{\sigma^2}$ et on simule les $x_i \sim \mathcal{N}(9, 1.5^2)$ On choisit les a priori suivantes : $\mu | x, \sigma^2 \sim \mathcal{N}\left(8, \dfrac{\sigma^2}{0.5}\right)$ et $\sigma^2 \sim \mathcal{IG}(1, 1)$



Modèle de régression linéaire simple

On considére maintenant le modéle de régression suivant : $Y = X\beta + \epsilon$

avec $Y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$; $X = \begin{pmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{pmatrix}$; $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$; $\epsilon = \begin{pmatrix} \epsilon_1 \\ \dots \\ \epsilon_n \end{pmatrix}$ et $\epsilon_i \sim \mathcal{N}\left(0, \sigma^2\right)$

Vraisemblance :

$$\mathcal{L}\left(Y|X, \beta, \sigma^2\right) \propto \left(\sigma^2\right)^{\frac{-n}{2}} \exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)^T(Y - X\beta)\right)$$

Lois a priori :

$$\pi\left(\sigma^2\right) \sim IG\left(\frac{v_0}{2}, \frac{v_0 \ s_0}{2}\right)$$

$$\pi\left(\beta|\sigma^2\right) \sim \mathcal{N}\left(\mu_0, \sigma^2\Lambda_0^{-1}\right)$$

lois a posteriori

Connaissant la vraisemblance et nos lois a priori on peut calculer les lois a posteriori :

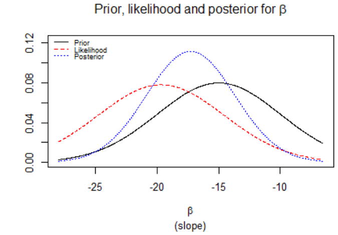
$$\beta|\ \sigma^2, Y, X \sim \mathcal{N}\left(\mu_n, \sigma^2\Lambda_n^{-1}\right)$$

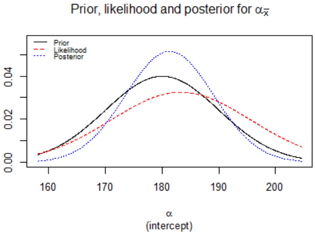
$$\Lambda_n = \left(X^T X + \Lambda_0\right)$$

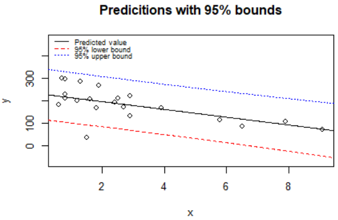
$\mu_n = (\Lambda_n)^{-1}\left(X^T X\hat{\beta} + \Lambda_0\mu_0\right)$ $\widehat{\beta}$ l'estimateur des moindres carrés .

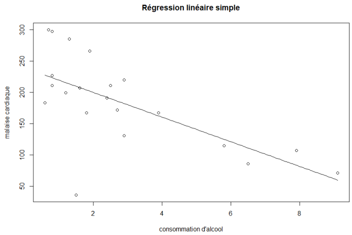
$$\sigma^2\ |Y, X \sim IG(a_n, b_n)\ a_n = \frac{v_0 + n}{2} b_n = \frac{v_0 s_0^2\left(Y^T Y + \mu_0^T\Lambda_0\mu_0 - \mu_n^T\Lambda_n\mu_n\right)}{2}$$

Illustration avec un jeu de données









Régréssion linéaire : $\hat{\alpha} = 239.147$ $\hat{\beta} = -19.683$ $E(\alpha|x, y) = 181.3$ et $E(\beta|x, y) = -17.29$.

Modèle de régression semi-paramétrique

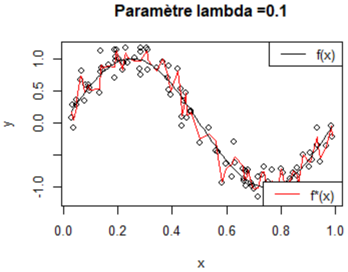
$$y_i = f(x_i) + \epsilon_i \quad \epsilon_i \sim \mathcal{N}\left(0, \sigma^2\right)$$

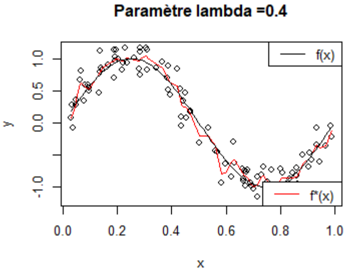
Avec une approche classique, on peut obtenir f comme le minimiseur de la somme des carrés pénalisée suivante :

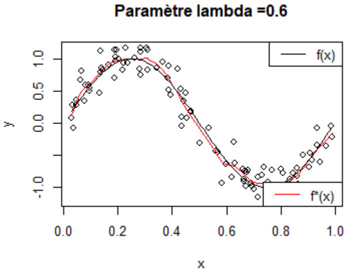
$$J(f) = \sum_{i=1}^{n} (y_i - f(x_i))^2 \; + \; \lambda \int \left(f''(x)\right)^2 dx$$

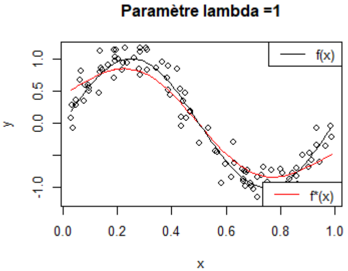
λ est un paramètre de lissage utilisé pour contrôler l'allure de la courbe et ainsi obtenir des courbes plus ou moins lisses.

Influence de λ









Approche bayésienne

$f = (f(x_1), ...f(x_n))$ $y = (y_1, ..., y_n)$

$y|f, \sigma^2 \sim \mathcal{N}_m \left(f, \sigma^2 I\right)$

Loi a priori : $f \sim \mathcal{N}_m \left(0; K^- \tau^2\right)$ K^- est l'inverse généralisé de K et $K^- = B\Omega^- B^T$

$\pi \left(\sigma^2\right) \sim I\mathcal{G}\ (a, b)$

Lois a posteriori : $f|y, \sigma^2 \sim \mathcal{N}_m \left(Sy, \sigma^2 S\right)$ Avec $S = B\left(B^T B + \lambda\Omega\right)^{-1} B^T$

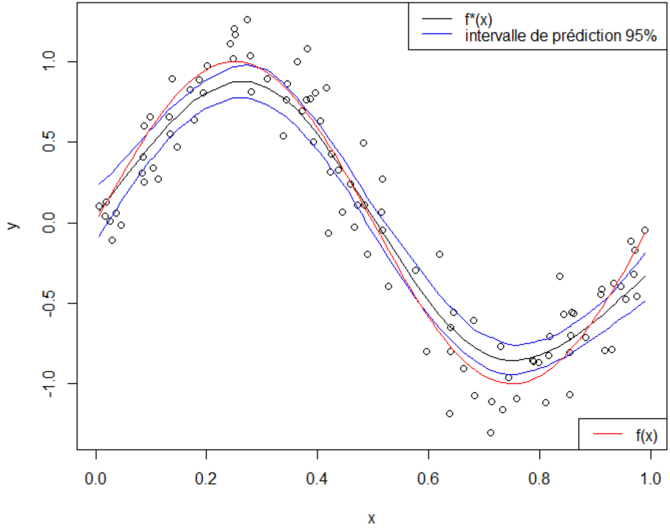
$\sigma^2|f, y \sim IG\left(\frac{n}{2} + a, \frac{1}{2}(y - f)^T (y - f) + b\right)$

Algorithme de Gibbs

Avant d'appliquer notre algorithme on génére des données $X \sim \mathcal{U}(0,1)$ Puis on applique aux points X la fonction $f(x) = \sin(2\pi x)$ et on ajoute une perturbation $\epsilon \sim \mathcal{N}(0, 0.04)$

- On initialise f_0 et σ_0^2
- Pour $i = 1, ...N$
  - On Calcul la matrice $S(\lambda)$
  - On simule $f_i | \sigma_{i-1}^2, y \sim \mathcal{N}_m \left(S(\lambda) y, \sigma^2 S(\lambda) \right)$
  - On simule $\sigma_i^2 | f_i, y \sim IG \left(\frac{n}{2} + a, \frac{1}{2}(y - f)^T (y - f) + b \right)$
  - On calcul $\lambda = \sigma_i^2 / \tau_i^2$

Sortie Graphique



Application aux données climatiques

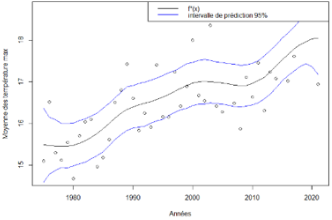


Figure 9: Régression bayésienne semi-paramétrique pour les températures maximales en fonction des années à Lyon.

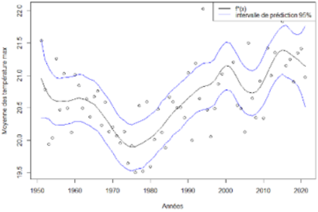


Figure 10: Régression bayésienne semi-paramétrique pour les températures maximales en fonction des années à Brindisi.

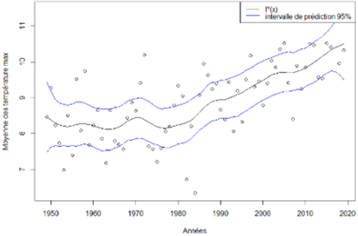


Figure 12: Régression bayésienne semi-paramétrique pour les températures maximales en fonction des années à Helsinki.

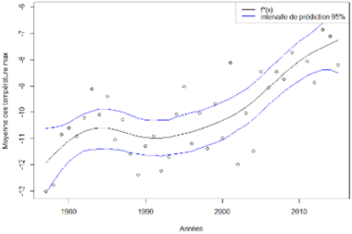


Figure 13: Régression bayésienne semi-paramétrique pour les températures maximales en fonction des années à Cape Sterlegov.