Université de Strasbourg
M1 - Magistère
Statistique - études de cas

Ségolen Geffray
geffray@math.unistra.fr
Année 2015/2016

Sujet 15: Estimation des coefficients d'une équation différentielle

Soit un échantillon i.i.d. $(Y_1, \ldots, Y_n)$ qui est une observation bruitée d'un système différentiel. Considérons que, pour $i = 1, \ldots, n$, le vecteur $Y_i = (Y_i^{(1)}, \ldots, Y_i^{(d)})^\top \in \mathbb{R}^d$ satisfait l'équation d-dimensionnelle suivante:

$$Y_i = m_\beta(x_i) + \varepsilon_i$$

dans laquelle le paramètre $\beta = (\beta_1, \ldots, \beta_q) \in \mathbb{R}^q$ est inconnu et à estimer, les vecteurs aléatoires $\varepsilon_i = (\varepsilon_i^{(1)}, \ldots, \varepsilon_i^{(d)})^\top \in \mathbb{R}^d$ sont i.i.d. centrés et admettent des moments d'ordre deux, et la fonction $m_\beta(.) = (m_\beta^{(1)}(.), \ldots, m_\beta^{(d)}(.))^\top$ de $\mathbb{R}$ dans $\mathbb{R}^d$ est solution du système différentiel suivant:

$$m'_\beta(x) = F(m_\beta(x), \beta).$$

Les x_i sont ici déterministes et équi-répartis dans l'intervalle d'observation noté $[0, T]$.
Pour estimer β, on propose ici une estimation en deux temps:

- 1ère étape: estimer une version "débruitée" des Y_i en lissant les observations Y_i au moyen d'un noyau noté $K(.)$. Formellement, un noyau est défini comme étant une fonction $K : \mathbb{R} \to \mathbb{R}$ intégrable et telle que $\int_{\mathbb{R}} K(u)du = 1$. Un noyau est dit positif lorsque $K(.) \geq 0$ et symétrique lorsque $K(u) = K(-u)$ pour tout $u \in \mathbb{R}$. L'estimateur (Gasser et Müller, 1979) des composantes de la fonction $m(.)$ est donné pour $j = 1, \ldots, d$ et $x \in \mathbb{R}$ par

$$\widehat{m}_{n,h}^{(j)}(x) = \sum_{i=1} w_{i,h}^{(j)}(x) Y_i^{(j)}$$

où les poids $w_{i,h}^{(j)}(x)$ sont donnés par

$$w_{i,h}^{(j)}(x) = \frac{1}{h} \int_{s_{i-1}}^{s_i} K\left(\frac{x-t}{h}\right) dt,$$

avec $s_0 = 0 < s_1 < \ldots < s_n = T$ où $\max |s_i - s_{i-1}| = O\left(\frac{1}{n}\right)$. L'estimateur des composantes de la fonction $m'(.)$ est donné pour $j = 1, \ldots, d$ et $x \in \mathbb{R}$ par

$$\widehat{m}_{n,h}^{\prime(j)}(x) = \sum_{i=1} w_{i,h}^{\prime(j)}(x) Y_i^{(j)}$$

où les poids $w_{i,h}^{\prime(j)}(x)$ sont donnés par

$$w_{i,h}^{\prime(j)}(x) = \frac{1}{h^2} \int_{s_{i-1}}^{s_i} K\left(\frac{x-t}{h}\right) dt.$$

La quantité $h > 0$ s'appelle la fenêtre ou le paramètre de lissage. Plus h est grand, plus l'estimateur est lisse et plus sa variance diminue. A l'inverse, lorsque h est petit, plus

l'estimateur est rugueux et plus son biais diminue. C'est un hyperparamètre au sens où on ne l'estime pas mais on le calibre. Il existe des méthodes de choix automatiques.
NB: des conditions techniques, portant sur le noyau et la fenêtre, sont en réalité requises pour obtenir des propriétés de convergence optimales des estimateurs.
La fonction `glkerns` du package `lokern` du logiciel `R` implémente l'estimation de la fonction $m(.)$ et de sa dérivée $m'(.)$ et propose un choix automatique de la fenêtre.

- 2ème étape: estimer β en minimisant une distance entre $\widehat{m'}_{n,h}(.)$ et $F(\widehat{m}_{n,h}(.), \beta)$. Par exemple, on peut proposer:

$$\widehat{\beta}_n = \arg\min_{\beta} \int_0^T \|\widehat{m'}_{n,h}(t) - F(\widehat{m}_{n,h}(t), \beta)\|^2 w(t)dt \tag{1}$$

où $w(.)$ est une fonction de poids choisie par l'utilisateur et où $\|.\|$ désigne la norme euclidienne de $\mathbb{R}^d$. Le choix le plus simple est $w(.) \equiv 1$. La fonction $w(.)$ peut être utilisée pour diminuer les effets de bord, par exemple en prenant

$$w(t) = \widetilde{w}_{\delta,\lambda}\left(1.05\frac{(t - T/2)}{T/2}\right)$$

avec

$$\widetilde{w}_{\delta,\lambda}(t) = \begin{cases} 0 & \text{si } |t| \geq 1 \\ 1 & \text{si } |t| \leq \delta \\ \exp\left(-\frac{\lambda}{(|t|-1)^2}\exp\left(-\frac{\lambda}{(|t|-\delta)^2}\right)\right) & \text{si } \delta < |t| < 1 \end{cases}$$

pour $0 < \delta, \lambda < 1$ fixés. Un choix peut être $\lambda = 0.5$ et $\delta = 0.7$.
L'estimation de β s'obtient donc en résolvant le problème d'optimisation non-linéaire en (1). Si cette étape est trop exigeante du point de vue computationnel, on peut approximer le critère à minimiser au moyen d'une somme de Riemann.

Exercice 1.
Dans chacun des cas suivants, générer un échantillon i.i.d. de taille n satisfaisant le modèle proposé, en utilisant les paramètres proposés.

1. Soit l'équation différentielle du 1er ordre:

$$m'(x) = \beta m(x).$$

Pour $i = 1, \ldots, n$, générer les Y_i définis par

$$Y_i = \exp(\beta x_i) + \varepsilon_i, \quad \varepsilon_i \sim_{i.i.d.} \mathcal{N}(0, \sigma^2).$$

Utiliser les valeurs $\beta = 1$ et tour à tour $\sigma^2 \in \{0.1, 0.2, 0.3\}$.

2. Soit l'équation différentielle du 2nd ordre:

$$m''(x) = -\beta^2 m(x)$$

qui est équivalente au système suivant de deux équations différentielles du premier ordre:

$$\begin{cases} m_1'(x) = \beta m_2(x), \\ m_2'(x) = -\beta m_1(x). \end{cases}$$

Pour $i = 1, \ldots, n$, générer les vecteurs $Y_i = (Y_i^{(1)}, Y_i^{(2)})^\top$ définis par

$$\begin{cases} Y_i^{(1)} = \cos(\beta x_i) + \varepsilon_i^{(1)}, \quad \varepsilon_i^{(1)} \sim_{i.i.d.} \mathcal{N}(0, \sigma_1^2), \\ Y_i^{(2)} = -\sin(\beta x_i) + \varepsilon_i^{(2)}, \quad \varepsilon_i^{(2)} \sim_{i.i.d.} \mathcal{N}(0, \sigma_2^2). \end{cases}$$

Utiliser les valeurs $\beta = 1$ et tour à tour $\sigma_1^2, \sigma_2^2 \in \{0.1, 0.2, 0.3\}$.

3. Soit l'équation différentielle du 1er ordre:

$$m'(x) = \beta_1 m(x) + \beta_2.$$

Pour $i = 1, \ldots, n$, générer les Y_i définis par

$$Y_i = \frac{1}{\beta_1} \exp(\beta_1 x_i) - \beta_2 + \varepsilon_i, \quad \varepsilon_i \sim_{i.i.d.} \mathcal{N}(0, \sigma^2).$$

Utiliser tour à tour les valeurs $\beta = (\beta_1, \beta_2)^\top \in \{(1,1)^\top, (1,2)^\top, (2,1)^\top\}$ et $\sigma^2 \in \{0.1, 0.2, 0.3\}$.

4. Soit le système de deux équations différentielles du premier ordre:

$$\begin{cases} m_1'(x) = \beta_1 \beta_2 m_2(x), \\ m_2'(x) = -\dfrac{\beta_1}{\beta_2} m_1(x). \end{cases}$$

Pour $i = 1, \ldots, n$, générer les vecteurs $Y_i = (Y_i^{(1)}, Y_i^{(2)})^\top$ définis par

$$\begin{cases} Y_i^{(1)} = \cos(\beta_1 x_i) + \varepsilon_i^{(1)}, \quad \varepsilon_i^{(1)} \sim_{i.i.d.} \mathcal{N}(0, \sigma_1^2), \\ Y_i^{(2)} = -\dfrac{1}{\beta_2} \sin(\beta_1 x_i) + \varepsilon_i^{(2)}, \quad \varepsilon_i^{(2)} \sim_{i.i.d.} \mathcal{N}(0, \sigma_2^2). \end{cases}$$

Utiliser les valeurs $\beta = (\beta_1, \beta_2)^\top \in \{(1,1)^\top, (1,2)^\top, (2,1)^\top\}$ et tour à tour $\sigma_1^2, \sigma_2^2 \in \{0.1, 0.2, 0.3\}$.

Estimer la valeur du paramètre $\beta = (\beta_1, \ldots, \beta_q)^\top \in R^q$ par $\widehat{\beta}_n$ selon la méthode proposée plus haut à partir des données générées. Utiliser tout à tour les tailles d'échantillon $n = 100, 250, 500, 1000, 1500$. Répéter ce travail $M = 1000$ fois. Présenter l'estimateur médian, le biais estimé des composantes de $\widehat{\beta}_n$, et l'écart-type estimé des composantes de $\widehat{\beta}_n$.