Simultaneous confidence bands for contrasts among several non-linear regression curves

Xiaolei Lu, Graduate University for Advanced Studies
Satoshi Kuriki, The Institute of Statistical Mathematics

Spurrier (1999, JASA) constructed exact simultaneous confidence bands for linear regression curves. Jamshidian, et al. (2010, CSDA) proposed a simulation-based method when the explanatory variable is restricted to an interval and the design matrices for each group may be different. Naiman's (1986, AS) gives a method for constructing conservative Scheffé-type simutaneous confidence bands single curvilinear regression model over a finite intervals. We consider constructing simultaneous confidence bands for all contrasts of k non-linear regression models by means of the volume-of-tube method (e.g., Takemura and Kuriki (2002, AAP)).

Suppose that observations (x_j, y_{ij}) are available from k groups, and for each group we assume non-linear regression models

$$y_{ij} = \beta_i^T f(x_j) + \epsilon_{ij}, \quad i = 1, \ldots, k, \;\; j = 1, \ldots, n.$$

Here, random errors ϵ_{ij} are assumed to be independently distributed as the normal distribution $N(0, \sigma(x_j)^2)$, where $\sigma(x)$ is a known function, and $f(x)$ is a $p \times 1$ known vector function. The least square estimator $\widehat{\beta}_i$ distributed normally with mean of β_i and covariance matrix Σ, here

$$\Sigma = \left(\sum_{j=1}^{n} \sigma(x_j)^{-2} f(x_j) f(x_j)^T \right)^{-1}$$

is the inverse of the $p \times p$ information matrix.

Let $\mathcal{C}$ be the set of vectors $c = (c_1, \ldots, c_k)^T$ such that $\sum_{i=1}^{k} c_i = 0$. The focus of this paper is to construct $1 - \alpha$ level simultaneous confidence bands for all the contrasts among the k regression models over a given finite interval $\mathcal{X} = (l, u)$ or infinite interval of the covariate for all $x \in \mathcal{X}$ and $c \in \mathcal{C}$. Specifically, we consider simultaneous confidence bands $\sum_{i=1}^{k} c_i y_i(x) \in \sum_{i=1}^{k} c_i \widehat{y}_i(x)) \pm b_\alpha \sqrt{\mathrm{Var}(\sum_{i=1}^{k} c_i(\widehat{y}_i(x) - E(\widehat{y}_i(x)))}$, for all $x \in \mathcal{X}$ and $c \in \mathcal{C}$. Here, $\widehat{y}_i(x) = \widehat{\beta}_i^T f(x)$, $y_i(x) = E(\widehat{y}_i(x)) = \beta_i^T f(x)$.

For this purpose, we need to find $b = b_\alpha$ satisfying the equation

$$P\left(\max_{x \in \mathcal{X}, c \in \mathcal{C}} \frac{|\sum_{i=1}^{k} c_i(\widehat{y}_i(x) - E(\widehat{y}_i(x)))|}{\sqrt{\mathrm{Var}(\sum_{i=1}^{k} c_i(\widehat{y}_i(x) - E(\widehat{y}_i(x)))}} \geq b \right) = P\left(\max_{q \in \Gamma} \langle \xi, q \rangle \geq b \right) = \alpha,$$

where $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}$, $\xi_i = \Sigma^{-\frac{1}{2}}(\widehat{\beta}_i - \beta_i)$, $\Gamma = \{c \otimes \psi(x) |\; x \in \mathcal{X},\; c \in \mathcal{C}\}$, $\psi(x) = \frac{\Sigma^{\frac{1}{2}} f(x)}{\|\Sigma^{\frac{1}{2}} f(x)\|} \in \mathbb{S}^{p-1}$. Here $\otimes$ denotes kronecker product. We caculate this probability by means of the volume of tube method.

Tube formula is given by

$$\begin{aligned} P\left(\max_{q \in \Gamma} \langle \xi, q \rangle \geq b \right) = & \frac{\Gamma(\frac{k}{2})}{\sqrt{\pi}\Gamma(\frac{k-1}{2})} \mathrm{Vol}(M)\{P(\chi_k^2 > b^2) - P(\chi_{k-2}^2 > b^2)\} \\ & + P(\chi_{k-1}^2 > b^2) + O\left(b^{n-2} e^{-\frac{1}{2}(1+\tan^2\theta_c)} \right) \end{aligned}$$

as $b \to \infty$, where $M = \{\pm\psi(x) |\; x \in \mathcal{X}\} \subset \mathbb{S}^{p-1}$, $\mathrm{Vol}(M)$ denotes the length of M, and θ_c is the criticle radius of Γ.

The critical radius θ_c of Γ is given by

$$\tan^2\theta_c = \inf_{(x,\theta) \neq (\tilde{x}, \tilde{\theta})} \frac{(1 - \alpha s)^2}{1 - s^2 - \max\{0, \alpha k\}^2},$$

where $s = \psi^{'}(x)\psi(\tilde{x})$, $\alpha = h^{'}(\theta)h(\tilde{\theta})$, $k = \sqrt{m}t$, $t = \psi_x^{'}(x)\psi(\tilde{x})$ and $m = (\psi_x(x)^{'}\psi_x(x))$. We find that large critical radius makes good accuracy.

Illustrative numerical data analyses to compare several growth curves are demonstrated.