Exercice supplémentaire – Modèles linéaires

Par Pierre-Alexandre Veilleux

Question

Considérez le modèle suivant :

$$Y_t = (\alpha + \beta) + \beta * {X_t}^2 + 2 * \beta * X_t + \varepsilon_t$$

Trouvez $\hat{\beta}$ et $\hat{\alpha}$ de façon à minimiser l'erreur quadratique moyenne.

Réponse

On a

$$Y_t = (\alpha + \beta) + \beta * {X_t}^2 + 2 * \beta * X_t + \varepsilon_t$$

$$Y_t = \alpha + \beta * (X_t + 1)^2 + \varepsilon_t$$

Et on veut minimiser

$$\sum_{t=1}^{n} {\varepsilon_t}^2 = \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2 = \sum_{t=1}^{n} (Y_t - \hat{\alpha} - \hat{\beta}(X_t + 1)^2)^2$$

On a donc

$$Eq. 1: \; -2 * \sum_{t=1}^{n} (Y_t - \hat{\alpha} - \hat{\beta}(X_t + 1)^2) = 0$$

et

$$Eq. 2: \; -2 * \sum_{t=1}^{n} (Y_t - \hat{\alpha} - \hat{\beta}(X_t + 1)^2) * (X_t + 1)^2 = 0$$

Ainsi, par l'*Eq. 1,*

$$\bar{Y} - \hat{\beta} * \frac{1}{n} \sum_{t=1}^{n} (X_t + 1)^2 = \hat{\alpha}$$

et par l'*Eq. 2,*

$$\sum_{t=1}^{n} Y_t * (X_t + 1)^2 - \sum_{t=1}^{n} \hat{\alpha} * (X_t + 1)^2 = \sum_{t=1}^{n} \hat{\beta}(X_t + 1)^4$$

$$\sum_{t=1}^{n} Y_t * (X_t + 1)^2 - \sum_{t=1}^{n} (\bar{Y} - \hat{\beta} * \frac{1}{n} \sum_{t=1}^{n} (X_t + 1)^2) * (X_t + 1)^2 = \sum_{t=1}^{n} \hat{\beta}(X_t + 1)^4$$

$$\frac{\sum_{t=1}^{n} Y_t * (X_t + 1)^2 - n * \bar{Y} * \frac{1}{n} \sum_{t=1}^{n} (X_t + 1)^2}{\sum_{t=1}^{n} (X_t + 1)^4 - n * (\frac{1}{n} \sum_{t=1}^{n} (X_t + 1)^2)^2} = \hat{\beta}$$

Une autre solution beaucoup plus simple serait de définir $Z_t = (X_t + 1)^2$. Ce faisant, nous retombons sur le modèle de régression linéaire simple que nous connaissons bien et nous pouvons utiliser les formules développées dans les notes de cours. Évidemment, le résultat selon les deux approches est identique.