

ANR Report - Target registration error distribution

Guillaume POTIER

04 septembre 2019

Introduction

Dans son papier Fitzpatrick montre :

$$E(TRE^2(r)) = E(FLE^2) \times \left(\frac{1}{N} + \frac{1}{3} \times \sum_{i=1}^3 \frac{d_i^2}{f_i^2} \right)$$

Nous voulons vérifier si ce résultat est compatible du point de vue de la régression linéaire.

Developpement dimension 1

Soit le modèle de régression linéaire simple :

$$y = X \times \beta + \epsilon$$

avec $X = \begin{bmatrix} 1 & x \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ et $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Nous pouvons estimer β et ϵ par $\hat{\beta}$ et $\hat{\epsilon}$ obtenus par l'estimateur des moindres carrés. Soit un point y_0 suivant le modèle de régression linéaire :

$$y_0 = X_0 \times \beta + \epsilon_0$$

avec $X_0 = \begin{bmatrix} 1 & x_0 \end{bmatrix}$ et $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$. Connaissant x_0 nous pouvons calculer \hat{y}_0 une estimation de y_0 :

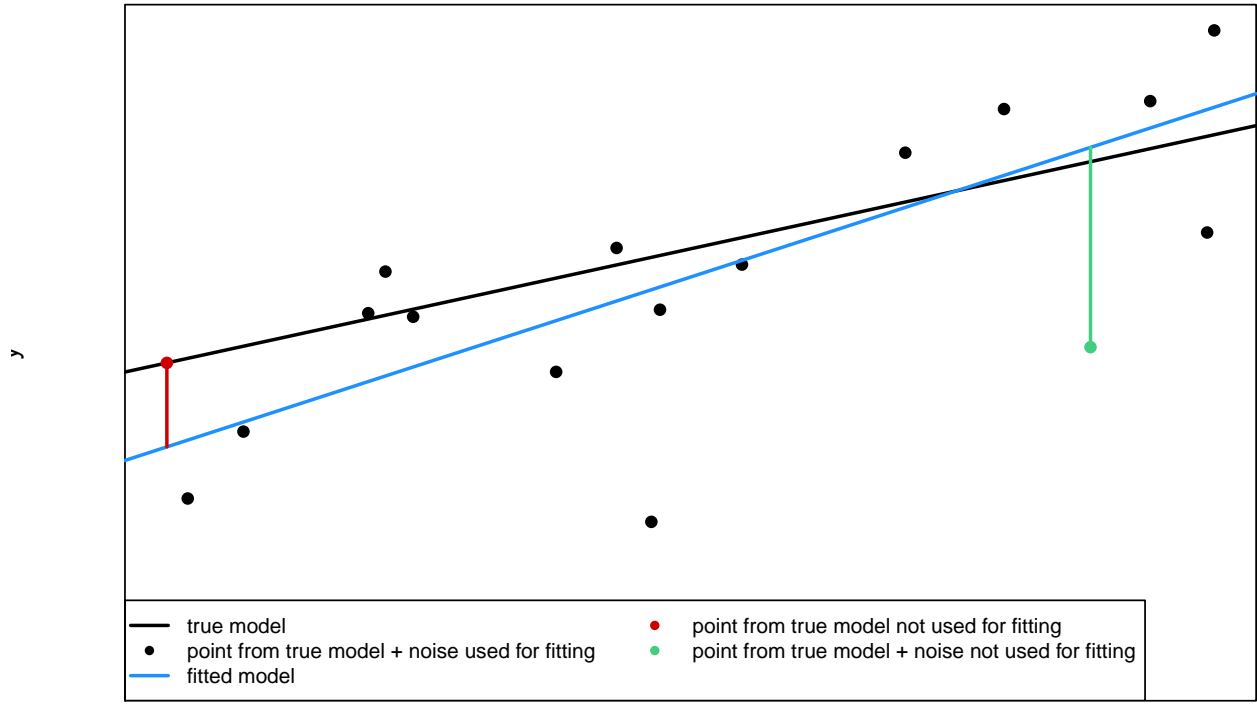
$$\hat{y}_0 = X_0 \times \hat{\beta}$$

Par ailleurs puisque y_0 suit le modèle de régression nous avons :

$$\epsilon_0 = y_0 - X_0 \times \beta$$

Que nous pouvons approcher par :

$$\hat{\epsilon}_0 = y_0 - X_0 \times \hat{\beta}$$



x

Sur la figure ci-dessus, la droite noire représente le modèle. Les points noirs représentent les points issus du modèle avec un bruit gaussien. La droite bleue est le modèle estimé par les moindres carrés à partir des points noirs. Le point rouge est un point issu du modèle sans bruit. Le point vert est un point issu du modèle avec un bruit gaussien. Le segment vert correspond à ϵ_0 . Le TRE de Fitzpatrick est défini comme la distance entre un point sur l'image cible et sa prédiction à partir du point associé sur l'image source. Le TRE est donc donné par $|\epsilon_0|$. Puisque $\epsilon_0 \sim \mathcal{N}(0, \sigma^2)$, alors $|\epsilon_0|$ suit une loi demi-normale de moyenne $\frac{\sigma\sqrt{2}}{\sqrt{\pi}}$ et de variance $\sigma^2 \left(1 - \frac{2}{\pi}\right)$.

Developpement dimension p

Soit le modèle de régression linéaire multivarié :

$$Y_{n \times p} = X_{n \times (p+1)} \times \beta_{(p+1) \times p} + \epsilon_{n \times p}$$

Avec $\epsilon \sim \mathcal{N}(0, \Sigma)$

Nous pouvons estimer β et ϵ par $\hat{\beta} = (X'X)^{-1}X'Y$ et $\hat{\epsilon} = Y - X\hat{\beta}$ obtenus par l'estimateur des moindres carrés. Soit un point y_0 suivant le modèle de régression linéaire :

$$y_0 = X_0 \times \beta + \epsilon_0$$

Nous avons :

$$\epsilon_0 = y_0 - X_0 \times \hat{\beta}$$

Et :

$$\hat{\epsilon}_0 \sim \mathcal{N}_p \left(0, \left(1 + X_0'(Z'Z)^{-1}X_0 \right) \Sigma \right)$$

En reprenant le raisonnement développé en dimension 1, nous voulons connaître la distribution de $\|\hat{\epsilon}_0\|^2$.

$$\|\hat{\epsilon}_0\|^2 = \sum_{i=1}^p \hat{\epsilon}_{0i}^2$$

Puisque $\hat{\epsilon}_0 \sim \mathcal{N}_p \left(0, \left(1 + X_0'(Z'Z)^{-1}X_0 \right) \Sigma \right)$, alors $\forall i \in [1, p]^{\mathbb{N}}, \hat{\epsilon}_{0i} \sim \mathcal{N}(0, \sigma_i^2)$. Si $\Sigma = \sigma^2 I_p$, alors les $\hat{\epsilon}_{0i}$ sont indépendants et identiquement distribués et :

$$\|\hat{\epsilon}_0\|^2 \sim \sigma^2 \left(1 + X_0'(Z'Z)^{-1}X_0 \right) \chi_p^2$$

Alors :

$$E(\|\hat{\epsilon}_0\|^2) = p\sigma^2 \left(1 + X_0'(Z'Z)^{-1}X_0 \right)$$

Et :

$$Var(\|\hat{\epsilon}_0\|^2) = 2p\sigma^4 \left(1 + X_0'(Z'Z)^{-1}X_0 \right)^2$$

De la même façon :

$$\|\hat{\epsilon}_0\| \sim \sqrt{\sigma^2 \left(1 + X_0'(Z'Z)^{-1}X_0 \right)} \chi_p$$

Alors :

$$E(\|\hat{\epsilon}_0\|) = \sqrt{\sigma^2 \left(1 + X_0'(Z'Z)^{-1}X_0 \right)} \times \sqrt{2} \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})}$$

Et :

$$Var(\|\hat{\epsilon}_0\|) = \sigma^2 \left(1 + X_0'(Z'Z)^{-1}X_0 \right) \times \left(p - \left(\sqrt{2} \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})} \right)^2 \right)$$

En dimension 1 nous avons trouvé que $|\hat{\epsilon}_0|$ suit une loi demi-normale. La loi demi-normale est équivalente à une loi du χ à 1 degré de liberté. Le développement en dimension p reste donc compatible avec le développement en dimension 1.