

1 Poids faible aux points aberrants

1.1 Introduction

LOESS est une méthode de régression non paramétrique. Elle utilise la régression linéaire des moindres carrés pondérés et crée une fonction qui décrit la variation des données, point par point, en considérant de manière plus importante les données les plus proches de ce point.

L'objectif est d'ajuster $\theta = [\theta_0, \theta_1]$ pour minimiser les moindres carrés pondérés soit :

$$\sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i))^2 \quad (1)$$

Avec :

- θ_0 et θ_1 les paramètres inconnus dont la valeur permet d'effectuer la procédure d'ajustement
- $(\theta_0 + \theta_1 x_i)$ la coordonnée en y prédite par la méthode en fonction de ces paramètres.
- $1 > w_i = \exp\left(-\frac{(x_i - x)^2}{2\rho^2}\right) > 0$ le poids Gaussien où ρ est considéré dans notre cas comme le paramètre de lissage.

Les poids est donc donné en utilisant le calcul des moindres carrés, on a par conséquent :

- Plus de poids à des points près du point cible x
- Moins de poids à des points plus loin de x

Autrement, si la différence $|x_i - x|$ est petite, alors le poids w_i est proche de 1, et, si dans le cas contraire, elle est grande, alors w_i est proche de 0. Notre modèle après la méthode est ajusté, ne retenant que le point du modèle qui est proche du point cible. La procédure se répète pour chaque point par ordre croissant des abscisses.

1.2 Exemple

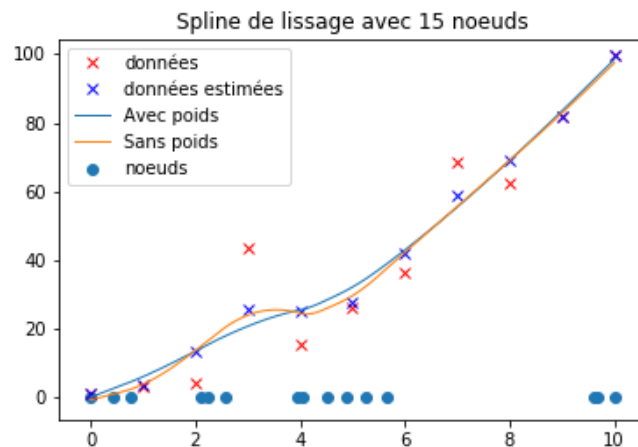


FIGURE 1 – Exemple du résultat

Dans cet exemple, nous avons en rouge une spline de lissage et en bleu une autre spline de lissage utilisant la méthode LOESS. Nous avons également les valeurs initiales de l'échantillon en rouge et celle ajustées en bleu. On remarque que la méthode LOESS considère moins les valeurs pouvant être considérées comme aberrantes ou bruitées.

1.3 Procédure

Partons de l'expression (1) avec x et y des vecteurs de taille m . Appelons cette expression S en fonction de θ , nous avons :

$$S(\theta) = \sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i))^2$$

$$\frac{\partial S}{\partial \theta_0} = -2 \sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i))$$

$$\frac{\partial S}{\partial \theta_1} = -2 \sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i)) x_i$$

Puis :

$$\frac{\partial S}{\partial \theta_0} = \sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i)) = 0$$

$$\iff \sum_{i=1}^m w_i \theta_0 + \sum_{i=1}^m w_i \theta_1 x_i = \sum_{i=1}^m w_i y_i \quad \text{Eq. (1)}$$

$$\frac{\partial S}{\partial \theta_1} = \sum_{i=1}^m w_i (y_i - (\theta_0 + \theta_1 x_i)) x_i = 0$$

$$\iff \sum_{i=1}^m w_i \theta_0 + \sum_{i=1}^m w_i \theta_1 x_i x_i = \sum_{i=1}^m w_i y_i x_i \quad \text{Eq. (2)}$$

En écrivant les équations Eq. (1) et Eq. (2) sous forme de matrice $\mathbf{A}\theta = \mathbf{b}$ nous obtenons :

$$\begin{bmatrix} \sum w_i & \sum w_i x_i \\ \sum w_i x_i & \sum w_i x_i x_i \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} \sum w_i y_i \\ \sum w_i y_i x_i \end{bmatrix}$$

$$\iff \mathbf{A}\theta = \mathbf{b}$$

$$\iff \theta = \mathbf{A}^{-1}\mathbf{b}$$

Il nous suffit donc de résoudre cette matrice.

1.4 Bilan

On a donc une méthode fonctionnelle s'adaptant à tout type de fonction mais peu devenir coûteuse en terme de temps lorsque l'échantillon devient grand.

2 Bibliographie

Xavier Bourret Sicotte, Locally Weighted Linear Regression (Loess)
Wikipedia, Local Regression
Wikipedia, Méthode des moindres carrés