Université de M'sila
Faculté des Mathématiques de l'Informatique
Département des mathématiques
Licence 3ième année
Module: Optimisation sans contraintes (2020/2021)

Méthode des moindres carrés

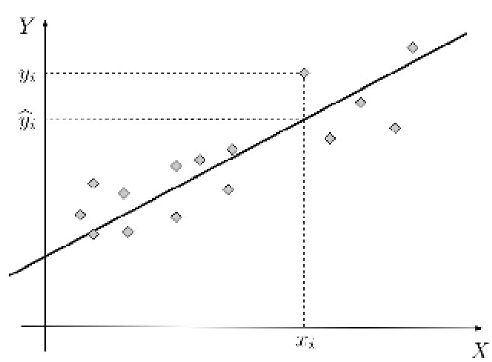
Soient $\{(x_1, y_1), ..., (x_m, y_m)\}$ une famille de m couples qui peuvent être répresenter dans le plan $\mathbb{R}^2$. On supposera que $m \geq 2$ et les x_i ne sont pas tous égaux. Les données $\{(x_i, y_i) : 1 \leq i \leq m\}$ s'appelle un *nuage* de m points. La méthode des moindres carrées consiste à rechercher une relation pouvant éventuellement exister entre les x_i et les y_i sous la forme

$$y = f(x),$$

cette opération s'appelle un *ajustement.*

On peut penser à beaucoup d'ajustements:
⧫ **Ajustement linéaire** $f(x) = ax + b$.
⧫ **Ajustement non linéaire** :
ajustement exponentiel : $f(x) = be^{\alpha x}$
ajustement puissance : $f(x) = bx^a$
ajustement logarithmique : $f(x) = a \ln(x) + b$
ajustement polynomial, par ex : $f(x) = ax^2 + bx + c$.

Ajustement linéaire. Si la fonction f est affine on parle d'une *régression linéaire.* On essaie d'approcher le nuage par une droite $f(x) = ax + b$. L'objectif donc de la méthode des moindres carrés est de rechercher parmi toutes les droites, celle qui rend minimale la somme des carrés des écarts des valeurs y_i et $f(x_i) = ax_i + b$.



Soit e_i l'écart de y_i et $f(x_i)$, on va s'intéresser à résoudre le problème de minimisation suivante

$$\begin{cases} \min g(a, b) \\ (a, b) \in \mathbb{R}^2 \end{cases}$$

où

$$g(a,b)=\sum_{i=1}^{m}e_i^2=\sum_{i=1}^{m}\left(y_i-(ax_i+b)\right)^2.$$

Remarque.

Quand un point du nuage est au dessus de la droite, l'écart $e_i=y_i-(ax_i+b)$ est positif et quand il en dessous de la droite, l'écart $e_i=y_i-(ax_i+b)$ sera négatif. Pour cette raison, on fait la somme des carrés des distances et on détrmine les coefficients a et b pour que cette somme soit minimale. Une autre raison d'utiliser le carré est de garantir la différentiabilité de fonction objectif.

Théorème.

Soient $\{(x_i,y_i):1\leq i\leq m\}$ un nuage de m points où $m\geq 2$ et les x_i ne sont pas tous égaux. Le problème de minimisation

$$\begin{cases}\min g(a,b)\\(a,b)\in\mathbb{R}^2\end{cases}$$

admet une solution globale unique.

Preuve.

Les points critiques de $g(a,b)$

$$\nabla g(a,b)=\begin{pmatrix}\frac{\partial g}{\partial a}(a,b)\\\frac{\partial g}{\partial b}(a,b)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

c'est à dire

$$\begin{cases}-2\sum_{i=1}^{m}x_i\left(y_i-ax_i-b\right)=0\\-2\sum_{i=1}^{m}\left(y_i-ax_i-b\right))=0\end{cases}$$

d'où

$$\begin{cases}\sum_{i=1}^{m}x_iy_i-a\sum_{i=1}^{m}x_i^2-b\sum_{i=1}^{m}x_i=0\\\sum_{i=1}^{m}y_i-a\sum_{i=1}^{m}x_i-bn=0\end{cases}$$

On pose $\sum_{i=1}^{m}x_iy_i=S_{xy}$, $\sum_{i=1}^{m}x_i^2=S_{x^2}$, $\sum_{i=1}^{m}x_i=S_{x^2}$ et $\sum_{i=1}^{m}y_i=S_y$, le systéme se réécrit

$$\begin{cases}S_{xy}-aS_{x^2}-bS_x=0\\S_y-aS_x-bn=0\end{cases}\Rightarrow\begin{cases}aS_{x^2}+bS_x=S_{xy}\\aS_x+bn=S_y\end{cases}$$

le déterminant du ce système

$$\begin{aligned}\begin{vmatrix}S_{x^2}&S_x\\S_x&n\end{vmatrix}&=S_x^2-nS_{x^2}\\&=\left(\sum_{i=1}^{m}x_i\right)^2-n\left(\sum_{i=1}^{m}x_i^2\right)\end{aligned}$$

comme $h(x) = x^2$ est strictement convexe

$$\begin{aligned} h\left(\sum_{i=1}^{m} \frac{1}{n} x_i\right) &< \sum_{i=1}^{m} \frac{1}{n} h(x_i) \Rightarrow \frac{1}{n^2}\left(\sum_{i=1}^{m} x_i\right)^2 < \sum_{i=1}^{m} \frac{1}{n} x_i^2 \\ &\Rightarrow \left(\sum_{i=1}^{m} x_i\right)^2 < n \sum_{i=1}^{m} x_i^2 \end{aligned}$$

c'est à dire

$$\begin{vmatrix} S_{x^2} & S_x \\ S_x & n \end{vmatrix} < 0$$

donc il existe une solution unique

$$\begin{cases} a^* = \frac{nS_{xy} - S_y S_x}{nS_{x^2} - S_x^2} \\ b^* = \frac{S_{x^2} S_y - S_x S_{xy}}{nS_{x^2} - S_x^2} \end{cases}$$

On a donc un point critique unique, il reste à voir si c'est bien un minimum. La matrice hessienne est

$$\begin{aligned} \nabla^2 g(a,b) &= \begin{pmatrix} \frac{\partial^2 g}{\partial a^2}(a,b) & \frac{\partial^2 g}{\partial b \partial a}(a,b) \\ \frac{\partial^2 g}{\partial a \partial b}(a,b) & \frac{\partial^2 g}{\partial b^2}(a,b) \end{pmatrix} \\ &= \begin{pmatrix} 2\sum_{i=1}^{m} x_i^2 & 2\sum_{i=1}^{m} x_i \\ 2\sum_{i=1}^{m} x_i & 2n \end{pmatrix} \end{aligned}$$

On a

$$\det \nabla^2 g(a^*, b^*) = 4\left(n \sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2\right) > 0$$

et

$$Tr\left(\nabla^2 g(a^*, b^*)\right) = 2n + 2\sum_{i=1}^{m} x_i^2 > 0$$

Donc, le point trouvé (a^*, b^*) est bien un minimum glaobal. En fin, la droite d'équation

$$f(x) = a^* x + b^*$$

minimise bien $g(a,b)$. ■

Ajustement exponentiel : $f(x) = be^{\alpha x}$ avec $b > 0$. Dans ce cas, on pose

$$y = be^{ax}$$

ce qui imlique

$$\begin{aligned} \ln(y) &= \ln(be^{ax}) \\ &\Rightarrow \ln(y) = \ln(b) + \ln(e^{ax}) \\ &\Rightarrow \ln(y) = \ln(b) + ax \end{aligned}$$

Posons $Y = \ln(y)$ et $B = \ln(b)$. On ontient donc

$$Y = ax + B$$

c'est à dire une régression linéaire.

Ajustement puissance : $f(x) = bx^a$ avec $b > 0$ et $x > 0$. Donc

$$\begin{array}{rcl} \ln(y) & = & \ln(bx^a) \\ & \Rightarrow & \ln(y) = \ln(b) + \ln(x^a) \\ & \Rightarrow & \ln(y) = \ln(b) + a\ln(x) \end{array}$$

On pose $Y = \ln(y)$, $X = \ln(x)$ et $B = \ln(b)$. On obtient donc

$$Y = aX + B,$$

c'est à dire une régression linéaire.

Ajustement logarithmique : $f(x) = a\ln(x) + b$ avec $x > 0$. On pose $X = \ln(x)$. Donc

$$y = aX + b,$$

qui est une relation de type affine.