

# Chapitre 8

## Méthode des moindres carrés

Lorsque l'on lie plusieurs variables, il en existe deux types : les **variables explicatives**  $X$  et les **variables expliquées**  $Y$ .

Il existe plusieurs types d'ajustement (Tab. 8.1).

Nom du modèle	Équation associée
L'ajustement linéaire	$Y = a_1 X$
L'ajustement quadratique	$Y = a_2 X^2 + a_1 X + a_0$
L'ajustement cubique	$Y = a_3 X^3 + a_2 X^2 + a_1 X + a_0$
L'ajustement du quatrième degré	$Y = a_4 X^4 + a_3 X^3 + a_2 X^2 + a_1 X + a_0$
L'ajustement du $n$ ième degré	$Y = \sum_{i=0}^n a_i X^i$
L'hyperbole	$Y = \frac{1}{a_0 + a_1 X}$ ou $\frac{1}{Y} = a_0 + a_1 X$
La fonction exponentielle	$Y = ab^X$ ou $\ln Y = \ln a + (\ln b) X = a_1 X + a_0$
La fonction puissance	$Y = aX^b$ ou $\ln Y = \ln a + bX$
La fonction exponentielle modifiée	$Y = ab^X + g$
La fonction puissance modifiée	$Y = aX^b + g$
La fonction de Gompertz	$Y = pq^{b^X}$ ou $\ln Y = \ln p + b^X \ln q = ab^X + g$
La fonction de Gompertz modifiée	$Y = pq^{b^X} + h$
La fonction logistique	$Y = \frac{1}{ab^X + g}$ ou $\frac{1}{Y} = ab^X + g$ ou $Y = a_0 + a_1 \ln X + a_2 \ln^2 X$

TABLE 8.1 – Les différents types d'ajustement avec la méthode des moindres carrés

## 8.1 Méthode graphique d'ajustement d'une courbe

Un jugement personnel permet d'obtenir rapidement l'ajustement d'un ensemble de données. On dit que l'on a **esquissé une courbe ajustement**.

Si on connaît la forme de l'équation de la courbe, on peut obtenir les constantes qui figurent dans l'équation en choisissant autant de points sur la courbe qu'il y a de constantes.

- Si la courbe est une droite, deux points sont nécessaires.
- Si la courbe est une parabole, trois points sont nécessaires.

Toutefois, chaque observateur obtiendra une courbe différente. C'est pour contrer ce phénomène qu'une méthode analytique d'ajustement fut mise au point : la méthode des moindres carrés.

## 8.2 Méthode des moindres carrés

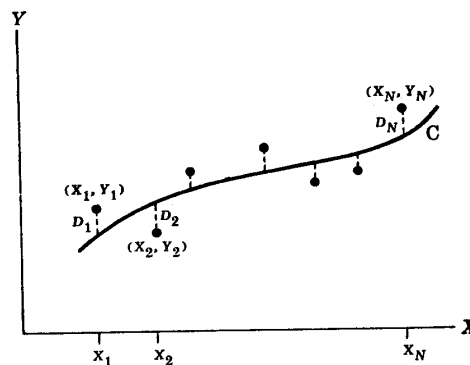


FIGURE 8.1 – Illustration de la méthode des moindres carrés

La figure n° 8.1 illustre l'ensemble de points  $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$ . Pour une valeur donnée de  $X$ , soit  $X_1$ , il y aura une différence entre la valeur  $Y_1$  et la valeur correspondante déterminée à partir de la courbe. On appelle cette différence  $D_1$ ;  $D_1$  est quelquefois appelée l'**écart**, l'**erreur** ou le **résidu**. Elle peut être positive, négative ou nulle. De la même façon, on obtient pour  $X_2, X_3, \dots, X_N$  des écarts  $D_2, D_3, \dots, D_N$ .

La somme  $D_1^2 + D_2^2 + \dots, D_N^2$  mesure « l'efficacité de l'ajustement » de la courbe  $C$  pour les données étudiées.

- Si elle est petite, l'ajustement est bon.
- Si elle est grande, l'ajustement est mauvais.

De fait, parmi toutes les courbes qui approchent un ensemble de données, celle qui donne le meilleur ajustement est celle qui vérifie la propriété suivante : la somme  $D_1^2 + D_2^2 + \dots + D_N^2$  doit être minimale. Pour ce, on pose :

$$J = D_1^2 + D_2^2 + \dots + D_N^2 = \sum_{i=1}^N D_i^2 \quad (8.1)$$

Dans le cas d'une droite, on aura  $J = \sum_{i=1}^N [Y_i - (a_0 - a_1 X_i)]^2$ . La méthode des moindres carrés généralisée fut inventée par A.-M. Legendre en 1806. Lorsque la dépendance entre  $Y$  et  $X$  est régie par une fonction  $f$ , avec  $f$  dépendant de certains paramètres, la méthode des moindres carrés consiste à trouver les paramètres pour :  $J = \sum_{i=1}^N [Y_i - f(X_i)]^2$  avec  $(X_i, Y_i)$  sont les points observés. Adrien-Marie Legendre (1752-1833)

Pour obtenir un minimum, il faut dériver partiellement  $J$  en fonction des coefficients recherchés en supposant que chaque dérivée partielle est nulle pour que l'erreur sur  $J$  soit minimale pour chaque coefficient  $a_i$ .

$$\left\{ \begin{array}{l} \frac{\partial J}{\partial a_0} = 0 \\ \frac{\partial J}{\partial a_1} = 0 \\ \dots \\ \frac{\partial J}{\partial a_i} = 0 \\ \dots \\ \frac{\partial J}{\partial a_N} = 0 \end{array} \right. \quad (8.2)$$

Dans le cas linéaire, les deux dérivées partielles valent :

$$\left\{ \begin{array}{l} \frac{\partial J}{\partial a_0} = -2 \sum_{i=1}^N [Y_i - (a_0 + a_1 X_i)] \\ \frac{\partial J}{\partial a_1} = -2 \sum_{i=1}^N [Y_i - (a_0 + a_1 X_i)] X_i \end{array} \right. \quad (8.3)$$

$$\Leftrightarrow \left\{ \begin{array}{l} \sum_{i=1}^N Y_i - \sum_{i=1}^N (a_0 + a_1 X_i) = 0 \\ \sum_{i=1}^N X_i Y_i - \sum_{i=1}^N (a_0 + a_1 X_i) X_i = 0 \end{array} \right. \quad (8.4)$$

$$\Leftrightarrow \left\{ \begin{array}{l} \sum_{i=1}^N Y_i = a_0 N + a_1 \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i Y_i = a_0 \sum_{i=1}^N X_i + a_1 \sum_{i=1}^N X_i^2 \end{array} \right. \quad (8.5)$$

Lorsqu'une courbe vérifie cette propriété, on dit qu'elle **ajuste les données au sens des moindres carrés**.

- Si l'ajustement est une droite, elle est appelée la **droite des moindres carrés**.
- Si l'ajustement est une parabole, elle est appelée la **parabole des moindres carrés**.

**Remarque.** On peut définir un autre type de courbe d'ajustement en considérant les projections orthogonales des points sur la courbe, au lieu des projections verticales ou horizontales, mais cela est assez rarement utilisé. En revanche, cet autre choix sert lors des analyses multivariées dites factorielles

## 8.3 Exemple de quelques ajustements

### 8.3.1 La droite des moindres carrés

Soit un nuage de points  $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$ , alors on peut ajuster l'équation  $Y = a_0 + a_1 X$  avec  $a_0$  et  $a_1$  déterminés en résolvant simultanément les équations suivantes :

$$\begin{cases} \sum_{i=1}^N Y_i = a_0 N + a_1 \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i Y_i = a_0 \sum_{i=1}^N X_i + a_1 \sum_{i=1}^N X_i^2 \end{cases} \quad (8.6)$$

Le système précédent est souvent appelé **équations normales de la droite des moindres carrés**.

On peut calculer  $a_0$  et  $a_1$  de la manière suivante :

$$\begin{cases} a_0 = \frac{(\sum_{i=1}^N Y_i)(\sum_{i=1}^N X_i^2) - (\sum_{i=1}^N X_i)(\sum_{i=1}^N X_i Y_i)}{N(\sum_{i=1}^N X_i^2) - (\sum_{i=1}^N X_i)^2} \\ a_1 = \frac{N(\sum_{i=1}^N X_i Y_i) - (\sum_{i=1}^N X_i)(\sum_{i=1}^N Y_i)}{N(\sum_{i=1}^N X_i^2) - (\sum_{i=1}^N X_i)^2} \end{cases} \quad (8.7)$$

Pour obtenir une droite des moindres carrés, on peut raccourcir les calculs en centrant les données sur les moyennes de chacune des coordonnées,  $\bar{X}$  et  $\bar{Y}$ , par le changement de variables suivant :

$$\begin{cases} X' = X - \bar{X} \\ Y' = Y - \bar{Y} \end{cases} \quad (8.8)$$

car il est clair que la droite des moindres carrés passe par le point  $(\bar{X}, \bar{Y})$  que l'on appelle barycentre ou centre de gravité du nuage de points.

### 8.3.2 La parabole des moindres carrés

Soit un nuage de points  $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$ , alors on peut ajuster l'équation  $Y = a_0 + a_1 X + a_2 X^2$  avec  $a_0$ ,  $a_1$  et  $a_2$  déterminées en résolvant simultanément les équations suivantes :

$$\begin{cases} \sum_{i=1}^N Y_i = a_0 N + a_1 \sum_{i=1}^N X_i + a_2 \sum_{i=1}^N X_i^2 \\ \sum_{i=1}^N X_i Y_i = a_0 \sum_{i=1}^N X_i + a_1 \sum_{i=1}^N X_i^2 + a_2 \sum_{i=1}^N X_i^3 \\ \sum_{i=1}^N X_i^2 Y_i = a_0 \sum_{i=1}^N X_i^2 + a_1 \sum_{i=1}^N X_i^3 + a_2 \sum_{i=1}^N X_i^4 \end{cases} \quad (8.9)$$

## 8.4 Problèmes mettant en jeu plus de deux variables

La méthode d'ajustement par les moindres carrés se généralise facilement avec plusieurs de deux variables. Par exemple, on peut considérer qu'il existe une relation fonctionnelle entre les trois variables  $X$ ,  $Y$  et  $Z$  et qu'elle est exprimée par l'équation linéaire :  $Z = a_0 + a_1 X + a_2 Y$ .

Par extension de la méthode des moindres carrés, on parle de l'approximation de données par un **plan des moindres carrés**. Si l'on veut estimer  $Z$  à partir de valeurs données pour  $X$  et  $Y$ , on appellera ce plan : **plan de régression de  $Z$  en  $X$  et  $Y$** . Les équations normales correspondant au plan des moindres carrés sont données par :

$$\begin{cases} \sum_{i=1}^N Z_i = a_0 N + a_1 \sum_{i=1}^N X_i + a_2 \sum_{i=1}^N Y_i \\ \sum_{i=1}^N X_i Z_i = a_0 \sum_{i=1}^N X_i + a_1 \sum_{i=1}^N X_i^2 + a_2 \sum_{i=1}^N X_i Y_i \\ \sum_{i=1}^N Y_i Z_i = a_0 \sum_{i=1}^N Y_i + a_1 \sum_{i=1}^N X_i Y_i + a_2 \sum_{i=1}^N Y_i^2 \end{cases} \quad (8.10)$$

On peut considérer des équations plus compliquées que l'équation linéaire. De telles équations correspondent à des surfaces dites **surfaces de régression**. Si le nombre de variables excède trois, on parle d'**hyperplans de régression**.



# **Bibliographie**