Licence de Biologie, 3e semestre
S. Vinatier Compléments de Mathématiques

5. Méthode des moindres carrés

L'objet de cette méthode est de fournir un outil d'interprétation de données. Plus précisément, lorqsu'on dispose de données dépendant de deux paramètres x et y, on peut les représenter dans le plan muni d'un repère, en marquant x en abscisse et y en ordonnée ; si le "nuage de points" qu'on obtient a l'allure d'une droite, on veut savoir quelle est l'équation de cette droite, c'est-à-dire quelle loi relie les deux paramètres de la mesure. C'est ce que la *méthode des moindres carrés* permet d'obtenir.

Exercice 1 (Par deux points distincts du plan, passe une droite et une seule)
Trouver l'équation cartésienne (du type : $y = ax + b$) de la droite passant par les points de coordonnées $(-1\,;4)$ et $(2\,;-2)$. Il s'agit donc de déterminer la *pente* a (qu'on appelle aussi le *cœfficient directeur*) et l'*ordonnée à l'origine* b. La tracer dans le plan muni d'un repère.
Toutes les droites du plan ont-elles une équation de ce type ?

Dès qu'on se donne trois points, il n'y a plus forcément une droite qui les relie. On cherche alors la droite passant "au plus près" des points donnés, en un sens que l'on va préciser. On commence par mesurer :

Exercice 2 (la "distance verticale" d'un point à une droite)
On considère la droite D d'équation $y = ax + b$ et le point M de coordonnées (x_0, y_0).

(a) Quelle est l'ordonnée du point de D d'abscisse x_0 ? En déduire que la *distance verticale* de M à D (c'est-à-dire la longueur du segment vertical qui relie M à D) est :

$$d(M, D) = |y_0 - ax_0 - b| \ .$$

(b) En déduire la "distance verticale" des points de coordonnées $(-1\,;1)$, $(0\,;4)$ et $(1\,;-1)$ à la droite de l'exercice 1. Calculer la somme des carrés de ces distances.

(c) Faire de même avec la droite d'équation $y = -x + 4/3$. Laquelle des deux droites passe-t-elle le plus près des trois points ?

Dans la méthode des moindres carrés, on cherche la droite (représentée par son équation $y = ax + b$) qui minimise la somme des carrés des distances verticales des points à la droite. Donnons-nous n points de coordonnées respectives (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n). La distance du i-ème point à la droite d'équation $y = ax + b$ est $|y_i - ax_i - b|$, donc la somme des carrés des distances est :

$$S = (y_1 - ax_1 - b)^2 + (y_2 - ax_2 - b)^2 + \cdots + (y_n - ax_n - b)^2 \ .$$

Ici, les x_i et y_i $(1 \leq i \leq n)$ sont des données, et on cherche les valeurs de a et b qui minimisent S. On voit donc S comme une fonction de deux variables : $S = S(a, b)$, et on cherche le couple (a_0, b_0) en lequel S prend sa valeur minimale.

La théorie des fonctions de deux variables donne une *condition nécessaire* pour que S atteigne un *extremum* (minimum ou maximum) en un couple (a_0, b_0). Si $f(a, b)$ est une fonction des deux variables a et b, on définit ses *dérivées partielles* par rapport à chacune des variables :

- $\dfrac{\partial f}{\partial a}(a,b)$ est la dérivée de f par rapport à a, en considérant b comme une constante ;
- $\dfrac{\partial f}{\partial b}(a,b)$ est la dérivée de f par rapport à b, en considérant a comme une constante.

Par exemple, si $g(a,b) = 3a^2 - 2ab$, on a : $\dfrac{\partial g}{\partial a}(a,b) = 6a - 2b$, $\dfrac{\partial g}{\partial b}(a,b) = -2a$.

Les dérivées partielles sont à priori des fonctions des deux variables a et b, mais elles peuvent être constantes par rapport à l'une ou l'autre des variables (tout comme la dérivée de la fonction $x \mapsto 2x$ est constante).

Pour qu'une fonction à deux variables $f(a,b)$ atteigne un extremum en (a_0, b_0), il faut que ses deux dérivées partielles s'annulent en (a_0, b_0) :

$$\frac{\partial f}{\partial a}(a_0, b_0) = 0 = \frac{\partial f}{\partial b}(a_0, b_0) \ .$$

Cette condition est **nécessaire** mais pas toujours **suffisante** : dans l'exemple ci-dessus, elle devient $6a_0 - 2b_0 = 0 = -2a_0$, donc $a_0 = 0 = b_0$. Le seul extremum **possible** de g est donc $g(0,0) = 0$. Cependant, 0 n'est pas un maximum car g prend des valeurs positives autour de $(0,0)$: $g(a,0) = 3a^2 \geq 0$; et ce n'est pas un minimum car g prend des valeurs négatives autour de $(0,0)$: $g(a,2a) = -a^2 \leq 0$. En définitive, la fonction g n'a pas d'extremum.

Exercice 3 (Dérivées partielles et extremum)
Pour chacune des fonctions suivantes :

$$h(a,b) = (a+1)^2 + (3-b)^2 \,, \quad i(a,b) = (2a+b)^2 \,, \quad j(a,b) = \cos a + \sin b \,, \quad k(a,b) = \frac{a+b}{a-b} \ .$$

(a) calculer les dérivées partielles par rapport aux deux variables a et b ;

(b) trouver les couples (a,b) pour lesquels la fonction est susceptible d'atteindre un extremum ; s'agit-il effectivement d'un extremum (minimum ou maximum) ?

Exercice 4 (Retour à $S(a,b)$)

(a) On considère d'abord la fonction $f_i(a,b) = (y_i - ax_i - b)^2$. Calculer ses deux dérivées partielles.

(b) En déduire :

$$\frac{\partial S}{\partial a} = 2x_1(ax_1 + b - y_1) + \cdots + 2x_n(ax_n + b - y_n) \,, \ \frac{\partial S}{\partial b} = 2(ax_1 + b - y_1) + \cdots + 2(ax_n + b - y_n) \ .$$

(c) Montrer que la condition nécessaire sur (a,b) pour que S y prenne un extremum est :

$$\begin{cases} (x_1 + \cdots + x_n)a & + & nb & = & y_1 + \cdots + y_n \\ (x_1^2 + \cdots + x_n^2)a & + & (x_1 + \cdots + x_n)b & = & x_1y_1 + \cdots + x_ny_n \end{cases}$$

On peut montrer que pour la fonction S, la condition nécessaire est suffisante ; de plus, le seul extremum de S est un minimum. Les solutions du système de deux équations à deux inconnues ci-dessus fournissent donc les cœfficients a et b de la droite qui minimise la somme des carrés des distances verticales.

Exercice 5 (Applications)
Trouver l'équation de la droite passant au plus près des points suivants :

(a) $(2\,;5)$, $(3\,;9)$, $(4\,;15)$, $(5\,;21)$.

(b) $(4\,;3)$, $(15\,;16)$, $(30\,;13)$, $(100\,;70)$, $(200\,;90)$.

On calculera à chaque fois $x_1 + \cdots + x_n$, $y_1 + \cdots + y_n$, $x_1^2 + \cdots + x_n^2$ et $x_1y_1 + \cdots + x_ny_n$, où n désigne le nombre de points dont on dispose, puis on écrira le système dont la solution (a,b) fournit l'équation de la droite.