Droites approchant un nuage de points

Dans un plan rapporté à un repère orthonormal, on considère les trois points M_1, M_2, M_3 de coordonnées respectives $(x_1 = 0, y_1 = 0)$, $(x_2 = 1, y_2 = 1)$, $(x_3 = -2, y_3 = 0)$.

On va étudier par différentes méthodes l'ajustement du “nuage” (M_1, M_2, M_3) par une droite.

On désigne par :

— δ une direction donnée du plan.

— D une droite non parallèle à δ, d'équation $y = ax + b$.

On projette les points M_1, M_2, M_3 sur D dans la direction δ.

On note respectivement m_1, m_2, m_3 les points obtenus.

1. Dans cette question, la direction δ est celle de l'axe des ordonnées Oy.

   On cherche la droite D rendant minimale l'expression $f(a, b) = M_1m_1 + M_2m_2 + M_3m_3$.

   (a) Calculer les distances M_1m_1, M_2m_2, et M_3m_3.

   (b) Dans cette question, le nombre réel b est fixé.

   En discutant suivant b, étudier la fonction φ définie par $\varphi(x) = |2x - b| + |x + b - 1|$.

   Montrer que l'application φ passe par un minimum pour $x = \frac{b}{2}$.

   (c) En déduire l'existence et l'unicité d'un couple (a_1, b_1) minimisant $f(a, b)$.

   Identifier la droite D d'équation $y = a_1x + b_1$.

2. Dans cette question, la direction δ est encore celle de l'axe des ordonnées.

   On cherche la droite D minimisant l'expression $g(a, b) = \max(M_1m_1, M_2m_2, M_3m_3)$.

   (a) On définit les trois ensembles suivants :

   — l'ensemble E_1 des points $M(x, y)$ du plan tels que $|y| \leqslant \frac{1}{3}$.

   — l'ensemble E_2 des points $M(x, y)$ du plan tels que $|y - 2x| \leqslant \frac{1}{3}$.

   — l'ensemble E_3 des points $M(x, y)$ du plan tels que $|x + y - 1| \leqslant \frac{1}{3}$.

   Montrer que leur intersection $E_1 \cap E_2 \cap E_3$ se réduit au seul couple $\left(\frac{1}{3}, \frac{1}{3}\right)$.

   (b) Prouver l'existence et l'unicité d'une droite D minimisant $g(a, b)$.

3. Dans cette question, la direction δ est toujours celle de l'axe des ordonnées.

   On cherche la droite D rendant minimale $h(a, b) = (M_1m_1)^2 + (M_2m_2)^2 + (M_3m_3)^2$.

   Le nombre réel a étant fixé, montrer que la fonction $b \mapsto h(a, b)$ admet un minimum en un point unique que l'on précisera.

   En déduire l'existence et l'unicité d'une droite D minimisant $h(a, b)$.

4. Dans cette question, λ est un nombre réel donné, dictinct de a.

   La direction δ est celle de la droite d'équation $y = \lambda x$.

   On cherche la droite D minimisant l'expression $h_\lambda(a, b) = (M_1m_1)^2 + (M_2m_2)^2 + (M_3m_3)^2$.

   (a) Exprimer $h_\lambda(a, b)$ en fonction de a, b, λ.

(b) Dans cette question et la suivante, le nombre λ est fixé et différent de $\frac{2}{7}$.

a étant donné, pour quelle valeur de b la fonction $b \mapsto h_\lambda(a,b)$ est-elle minimale ?

En déduire que $h_\lambda(a,b)$ est minimal lorsque $\theta(a) = \dfrac{7a^2-4a+1}{(a-\lambda)^2}$ est minimal.

(c) Étudier les variations de la fonction θ.

(d) En déduire l'existence et l'unicité d'un couple (a_4, b_4) conduisant à la plus petite valeur possible pour $h_\lambda(a,b)$. On explicitera a_4 et b_4 en fonction de λ.

(e) Montrer qu'une telle droite D (d'équation $y = ax+b$) n'existe pas si $\lambda = \frac{2}{7}$.

(f) Toujours si $\lambda = \frac{2}{7}$, vérifier l'existence d'une droite D solution mais d'équation $x = \alpha$ (autrement dit, la quantité $h_\lambda(a,b)$ est minimum si on projette parallèlement à la direction $y = \lambda x$ et sur une droite D "verticale".)

5. Dans cette question, n est un entier strictement positif.

On se donne un "nuage" de n points $M_k(x_k, y_k)$, avec $1 \leqslant k \leqslant n$.

On suppose que les points M_k ne sont pas alignés sur une même droite. On pose

$$\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k,\quad \overline{y} = \frac{1}{n}\sum_{k=1}^{n} y_k,\quad \overline{x^2} = \frac{1}{n}\sum_{k=1}^{n} x_k^2,\quad \overline{y^2} = \frac{1}{n}\sum_{k=1}^{n} y_k^2,\quad \text{et}\quad \overline{xy} = \frac{1}{n}\sum_{k=1}^{n} x_k y_k.$$

On définit également $\mathrm{v}(x) = \overline{x^2} - \overline{x}^2$, $\mathrm{v}(y) = \overline{y^2} - \overline{y}^2$, et $\mathrm{cov}(x,y) = \overline{xy} - \overline{x}\,\overline{y}$.

Soit m_k la projection de M_k sur la droite D d'équation $y = ax+b$ parallèlement à Oy.

On se propose de trouver le couple (a,b) minimisant l'expression $H(a,b) = \sum_{k=1}^{n}(M_k m_k)^2$.

(a) En considérant $\sum_{k=1}^{n}(x_k - \overline{x})^2$, montrer que $\mathrm{v}(x) > 0$.

(b) Quand on fixe a, montrer que $H(a,b)$ est minimum quand $b = \overline{y} - a\overline{x}$.

Comment peut-on interpréter ce résultat ?

(c) Montrez que le calcul précédent conduit à $|\,\mathrm{cov}(x,y)| < \sqrt{\mathrm{v}(x)\ \mathrm{v}(y)}$.

(d) En déduire l'unique couple (a,b) minimisant la fonction H.

Quelle est la valeur de ce minimum ?

6. Dans cette question, on projette les n points M_k sur la droite D d'équation $y = ax+b$, parallèlement à la direction $y = \lambda x$ (où λ est un réel fixé distinct de a.)

Pour chaque λ, on cherche s'il existe un couple (a,b) minimisant $H_\lambda(a,b) = \sum_{k=1}^{n}(M_k m_k)^2$.

En s'inspirant de ce qui précède, étudier l'existence d'un tel couple (a,b) et le cas échéant donner la valeur obtenue pour le minimum de $H_\lambda(a,b)$.

Pour simplifier les notations, on pourra poser : $v = \mathrm{v}(x)$, $w = \mathrm{v}(y)$, $c = \mathrm{cov}(x,y)$.