Devoir maison n°10

À rendre le 16/01/2023

Exercice 1

Dans cet exercice, se place dans $\mathbb{R}^n$ munit du produit scalaire canonique.

On considère n points de $\mathbb{R}^2$, $M_i : \begin{pmatrix} x_i \\ y_i \end{pmatrix}$ pour $i \in [\![1, n]\!]$, et on cherche une droite $\mathcal{D} : y = ax + b$ qui serait une "meilleure approximation affine" du nuage de point $\{M_i,\ i \in [\![1, n]\!]\}$. Pour cela on cherche à minimiser la quantité (dépendant de $a, b \in \mathbb{R}$).

$$S(a,b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

qui représente la somme des carrés des distances obtenues par projection verticale. C'est la méthode des moindres carrés.

On considère $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ et $U = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$. Ces trois vecteurs sont fixés pour l'exercice.

1. Pour $a, b \in \mathbb{R}$ fixés, exprimer $S(a, b)$ sous forme d'une norme faisant intervenir X, Y, U.
2. On note $F = \mathrm{Vect}(U, X)$. À quelle condition sur les points M_i, l'espace F est-il un plan de $\mathbb{R}^n$?
3. Soient $a, b \in \mathbb{R}$. On note $u_{a,b} = aX + bU$ et on suppose que $Y - u_{a,b} \perp F$. Interpréter $u_{a,b}$ en fonction de F et Y.
4. On considère le même vecteur $u_{a,b}$ qu'à la question précédente. Montrer que si $v \in F$, alors $\|Y - v\| \geqslant \|Y - u_{a,b}\|$ en utilisant le théorème de Pythagore.
5. On cherche maintenant $a, b \in \mathbb{R}$ tels que $Y - u_{a,b} \perp F$. Montrer que $Y - u_{a,b} \perp F$ si et seulement si
$$\begin{cases} a \sum\limits_{i=1}^{n} x_i + nb = \sum\limits_{i=1}^{n} y_i \\ a \sum\limits_{i=1}^{n} x_i^2 + b \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i \end{cases}$$
6. ($\star$) Montrer que si les M_i n'ont pas tous la même abscisse, alors le système de la question précédente (d'inconnues a et b) est un système de Cramer. [1]
7. En notant A la matrice du système précédent, exprimer l'unique solution $\begin{pmatrix} a \\ b \end{pmatrix}$ du système en fonction de A et de produits scalaires.
8. ($\star$) Écrire en python une fonction `regression(X, Y)` qui prend comme arguments une liste d'abscisses `X` (pas toutes égales) et une liste d'ordonnées `Y` de même longueur, et retournant les flottants a, b correspondants aux points dont les coordonnées sont données.
   On pourra utiliser `numpy.linalg.solve` pour résoudre le système considéré.
9. ($\star$) Tracer sur un même graphique le nuage de point et la droite calculée, pour des points que vous aurez généré.

1. On pourra aller réviser l'inégalité de Cauchy-Schwartz, ainsi que son cas d'égalité.

Indications

1. Cette somme est la norme au carré d'un certain vecteur. Lequel ?
2. Un plan est de dimension...
3. À quel espace appartient $u_{a,b}$?
4. Classiquement, on raisonne sur les normes au carré.
5. Traduire $u_{a,b} \perp F = \text{Vect}(X, U)$.
6. On peut parler de la matrice de la question suivante.
7. Écrire le système sous forme d'égalité de matrices faisant intervenir A.
8. Construire d'abord les matrices.
9. Trouver des points dans un TP de SI ou bien les générer aléatoirement (mais pas trop, pour qu'ils soient presque alignés).