Université de Strasbourg **Techniques d'Analyse Numérique (TAN1)**
L3 Mathématiques (S5) **Année 2019-2020**

TD-TP6. Problème des moindres carrés et régression polynomiale

Rappels : Problème des moindres carrés.

On s'intéresse ici au problème suivant :

$$\begin{cases} \text{Trouver } \bar{X} \in \mathbb{R}^n \text{ tel que} \\ A\bar{X} = b, \text{ avec } A \in \mathcal{M}_{p\times n}(\mathbb{R}) \text{ et } b \in \mathbb{R}^p. \end{cases} \tag{1}$$

• Si $n = p$ et si A est inversible, il y a une unique solution $\bar{X} = A^{-1}b$ que l'on peut trouver numériquement avec une méthode directe (LU), une méthode itérative (Jacobi, Gauss-Seidel...) ou en reformulant ce problème par un problème de minimisation (méthodes de gradient et de descente).

• Par contre, si A n'est pas inversible ou n'est pas carrée (par exemple, $p > n$ le cas le plus fréquent en pratique), un tel $\bar{X}$ n'existera pas. Le problème des moindres carrés consiste alors à résoudre le problème précédent dans un sens «aussi satisfaisant que possible». C'est-à-dire qu'on va chercher $\bar{X}_2 \in \mathbb{R}^n$ minimisant le résidu en norme 2 : $X \mapsto ||AX - b||_2$ (ou plus précisément, son carré). Le problème des moindres carrés revient donc à s'intéresser au problème suivant (à la place du problème (1) initial)

$$\begin{cases} \text{Trouver } \bar{X}_2 \in \mathbb{R}^n \text{ tel que} \\ ||A\bar{X}_2 - b||_2^2 = \min\{||AX - b||_2^2, \text{ avec } X \in \mathbb{R}^n\}. \end{cases} \tag{2}$$

Exercice 1 : Régression polynomiale

Le but de cet exercice est de donner un exemple de problème qui conduit à la résolution des moindres carrés.

Considérons un ensemble de points $(x_i, y_i)_{i\in[\![1,p]\!]}$. On cherche le polynôme de degré $n-1$ qui approche au mieux les points. On note $P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + ... + \alpha_{n-1}x^{n-1}$ ce polynôme.

Méthode 1

1. On suppose que $p > n$, et on va chercher à faire passer P exactement par les points (x_i, y_i). Écrire ce problème sous la forme d'un système linéaire $AX = b$ avec $X = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{pmatrix}$.
2. On suppose que les $(x_i, y_i)_{i\in[\![1,p]\!]}$ valent $\{(0, 1), (1, 9), (3, 9), (4, 21)\}$ et qu'on cherche un polynôme de degré 1. Que valent p et n dans ce cas ? Donner explicitement la matrice A et le vecteur b.
3. Montrer qu'un tel polynôme n'existe pas.

Méthode 2

On va chercher alors un polynôme de degré $n-1$ qui passe au plus près des points $(x_i, y_i)_{i\in[\![1,p]\!]}$ sans pour autant passer exactement par ces points. On cherche plus précisément à minimiser le carré de l'erreur des ordonnées en norme 2 :

$$E(\alpha_0, \alpha_1, ..., \alpha_{n-1}) = \sum_{i=1}^{p} |y_i - P(x_i)|^2 = \sum_{i=1}^{p} |y_i - (\alpha_0 + \alpha_1 x_i + ... + \alpha_{n-1}x_i^{n-1})|^2.$$

On s'intéresse donc au problème suivant

$$\begin{cases} \text{Trouver } \bar{\alpha}_0, \bar{\alpha}_1, ..., \bar{\alpha}_{n-1} \in \mathbb{R} \text{ tels que} \\ E(\bar{\alpha}_0, \bar{\alpha}_1, ..., \bar{\alpha}_{n-1}) = \min\{E(\alpha_0, \alpha_1, ..., \alpha_{n-1}), \text{ avec } \alpha_0, \alpha_1, ..., \alpha_{n-1} \in \mathbb{R}\}. \end{cases} \tag{3}$$

4. Réécrire le problème (3) sous la forme d'un problème de moindres carrés (2). Que valent la matrice de Vandermonde A et le vecteur b ? Quelles sont leurs dimensions ?
5. On note $J(X) = ||AX - b||_2^2$ la fonction à minimiser du problème des moindres carrés déterminés à la question précédente. On définit χ le sous espace vectoriel tel que $\mathbb{R}^n = \chi \oplus \mathrm{Ker}(A)$. Montrer que J est coercice sur χ (c'est-à-dire que $J(X) \underset{||X|| \to \infty}{\longrightarrow} +\infty$ avec $X \in \chi$) et qu'elle est convexe sur $\mathbb{R}^n$ (et même strictement convexe sur χ).
6. En déduire que J admet un unique minimiseur sur χ, noté x_2 et que l'ensemble des solutions au problème des moindres carrés déterminé à la question 4 est l'ensemble $\{x_2 + y,\ y \in \mathrm{Ker}(A)\}$.
7. Montrer que chercher un minimimiseur de J sur $\mathbb{R}^n$ équivaut à chercher une solution de l'équation suivante qu'on appelle équation normale :
$$A^T AX = A^T b.$$
8. Montrer que cette solution est unique si $rang(A) = n$.
9. On reprend le jeu de données de la question 2 et on cherche un polynôme de degré 1. Que valent $A^T A$ et $A^T b$? Résoudre les équations normales dans ce cas.
10. En déduire le polynôme de degré 1 recherché.

Exercice 2 : Régression linéaire par python : cas d'une droite

On se donne les points suivants $(0, -1), (1, 0.2), (2, 0.9), (3, 2.1)$. On cherche la droite de régression linéaire : la droite passant «au mieux» par ces points au sens des moindres carrés, d'équation $y = \alpha_0 + \alpha_1 x$.

1. Positionner ces points sur un graphe à l'aide d'un marqueur.
2. Écrire la matrice A et le vecteur b de la régression linéaire par moindres carrés dans ce cas et les implémenter en python.
3. Le problème des moindres carrés (2) est résolu en python par la fonction `np.linalg.lstsq(A, b)`. Cette fonction ressort le vecteur $\bar{X}_2$ qui minimise la norme $||A\bar{X}_2 - b||_2$, ainsi que la valeur du résidu **au carré** $||A\bar{X}_2 - b||_2^2$. Pour avoir accès au vecteur $\bar{X}_2$, il faudra donc écrire `np.linalg.lstsq(A, b)[0]` et pour avoir accès au carré du résidu il faudra utiliser `np.linalg.lstsq(A, b)[1]`. Calculer les coefficients α_0 et α_1 de la droite de régression linéaire.
4. Tracer cette droite sur la figure. Calculer le résidu et le comparer à la valeur `np.linalg.lstsq(A, b)[1]`.

Exercice 3 : Régression linéaire par python : cas d'une parabole

On cherche maintenant la parabole $P(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2$ qui approche au mieux les points
$$(1, 0.3), (2.5, 1.1), (3.5, 1.5), (4, 2.0), (5, 3.2), (7, 6.6), (8.5, 8.6).$$

1. Écrire ce problème sous la forme d'un problème aux moindres carrés (2) en explicitant la matrice A et le vecteur b.
2. Résoudre ce problème en python et tracer sur la même figure les points par des marqueurs et la parabole de régression.
3. Le problème est-il le même si on cherche la parabole qui approche au mieux les points sous la forme $\tilde{P}(x) = \tilde{\alpha}_0 + \tilde{\alpha}_2 x^2$? Que valent la nouvelle matrice $\tilde{A}$ et le nouveau vecteur $\tilde{b}$? Les coefficients $\tilde{\alpha}_0$ et $\tilde{\alpha}_2$ sont-ils identiques aux α_0 et α_2 ?
4. Tracer la parabole de $\tilde{P}$ sur la figure.
5. Quelle parabole a le meilleur résidu ?