



« Un voyage de mille lieues commence par un premier pas. » — Lao Tseu

Avant de commencer : à propos de ce document

Ce devoir maison n'est pas une évaluation ordinaire. C'est une **invitation au voyage** dans un territoire mathématique que vous n'avez peut-être jamais exploré.

Quelques principes importants :

- **Ce DM est facultatif.** Tout travail rendu sera valorisé par la note maximale. Ce qui compte, c'est l'effort authentique et la curiosité, pas le fait d'avoir tout terminé. Un travail bâclé ne sera pas valorisé ; un travail sincère, même partiel, le sera toujours.
- **Ce document vous accompagnera peut-être plusieurs années.** Certaines questions qui vous semblent inaccessibles aujourd'hui deviendront limpides après quelques mois d'études supplémentaires. Gardez-le.
- **Un horizon raisonnable : 3-4 semaines.** Il n'y a pas de date limite stricte, mais visez cet horizon. Si vous avez besoin de plus de temps, parlez-en.
- **Si vous bloquez, écrivez.** Expliquez ce que vous avez essayé, où vous êtes bloqué, ce que vous pensez qu'il faudrait faire. C'est aussi du travail mathématique — et souvent le plus formateur.

Niveaux de difficulté

Les questions sont organisées en trois niveaux, comme les étapes d'une ascension :

△ Niveau I – Fondations

Calculs et vérifications. Accessible à tous avec les outils du cours.

▲ Niveau II – Construction

Compréhension des mécanismes. Demande un peu de recul et d'autonomie.

◆ Niveau III – Sommet

Démonstrations rigoureuses et questions ouvertes. Un défi stimulant pour les plus motivés, ou une invitation pour plus tard.

Conseil : Faites toutes les questions △ d'une partie avant de passer aux ▲. Les ◆ sont des défis : ne vous découragez pas si vous n'y arrivez pas cette année.

1805. Paris. Adrien-Marie Legendre publie une méthode révolutionnaire.

En science, on mesure des grandeurs : la température d'un liquide au fil du temps, la distance parcourue par un mobile, la taille d'une plante en fonction de l'engrais apporté. Ces mesures, portées dans un repère, forment un **nuage de points**.

Souvent, les points semblent *presque* alignés — mais jamais parfaitement, à cause des erreurs de mesure. Se pose alors une question fondamentale : **quelle est la « meilleure » droite qui résume ces données ?**

En 1805, Legendre propose un critère précis : la meilleure droite est celle qui **minimise la somme des carrés des erreurs**. C'est la *méthode des moindres carrés*, encore utilisée aujourd'hui dans toutes les sciences.

Exercice 1: Question préliminaire | ★

△ Mise en situation

On a mesuré la longueur y (en cm) d'un ressort en fonction de la masse x (en kg) suspendue à son extrémité. Voici les résultats :

Masse x_i (kg)	1	2	3	4	5
Longueur y_i (cm)	2	3	5	6	9

- Placer ces cinq points $M_1(1; 2)$, $M_2(2; 3)$, $M_3(3; 5)$, $M_4(4; 6)$, $M_5(5; 9)$ dans un repère orthogonal.
- Ces points sont-ils parfaitement alignés ? Justifier.
- (Réflexion) : À votre avis, quelle droite passant par l'origine ($y = ax$) « résumerait le mieux » ces données ? Proposer une valeur de a à l'intuition.

1. Partie I – La notion d'erreur

Pour comparer différentes droites candidates, il faut d'abord définir ce qu'est une « erreur ».

Définition : erreur verticale

Soit une droite $d : y = ax$ et un point $M_i(x_i; y_i)$. L'**erreur verticale** de M_i par rapport à d est la différence :

$$e_i = y_i - ax_i$$

C'est la distance (signée) entre le point M_i et la droite, mesurée verticalement.

Exercice 2: Calcul des erreurs | ★★

On teste la droite $d : y = 2x$.

△ Q1. Pour chaque point M_i , calculer $e_i = y_i - 2x_i$ et compléter le tableau :

Point	$M_1(1; 2)$	$M_2(2; 3)$	$M_3(3; 5)$	$M_4(4; 6)$	$M_5(5; 9)$
$ax_i = 2x_i$					
$e_i = y_i - 2x_i$					

△ Q2. Calculer la somme $e_1 + e_2 + e_3 + e_4 + e_5$. Que constate-t-on ?

△ Q3. Pourquoi la somme des erreurs n'est-elle pas un bon indicateur de la qualité d'une droite ?

Indication : des erreurs positives et négatives peuvent se compenser.

△ Q4. Pour éviter les compensations, on élève chaque erreur au carré. Compléter :

Point	M_1	M_2	M_3	M_4	M_5	Somme
$e_i^2 = (y_i - 2x_i)^2$						$S =$

2. Partie II – La fonction erreur

On cherche maintenant la valeur de a qui rend la somme des erreurs au carré la plus petite possible.

Exercice 3: Construire $S(a)$ | ★★★

On considère une droite quelconque $d_a : y = ax$ passant par l'origine. Pour chaque point $M_i(x_i; y_i)$, l'erreur est $e_i(a) = y_i - ax_i$, et on définit :

$$S(a) = \sum_{i=1}^5 (y_i - ax_i)^2 = (y_1 - ax_1)^2 + (y_2 - ax_2)^2 + \cdots + (y_5 - ax_5)^2$$

▲ Q5. Remplacer x_i et y_i par les valeurs numériques de nos 5 points et écrire $S(a)$:

$$S(a) = (2 - a)^2 + (3 - 2a)^2 + (5 - 3a)^2 + (6 - 4a)^2 + (9 - 5a)^2$$

▲ Q6. Développer chaque carré. On rappelle que $(u - v)^2 = u^2 - 2uv + v^2$.

1. $(2 - a)^2 =$

2. $(3 - 2a)^2 =$

3. $(5 - 3a)^2 =$

4. $(6 - 4a)^2 =$

5. $(9 - 5a)^2 =$

▲ Q7. En additionnant les cinq expressions, montrer que :

$$S(a) = 55a^2 - 184a + 155$$

Vérification intermédiaire :

- Coefficient de a^2 : $1 + 4 + 9 + 16 + 25 = \dots$
- Coefficient de a : $-2(1 \times 2 + 2 \times 3 + 3 \times 5 + 4 \times 6 + 5 \times 9) = \dots$
- Terme constant : $4 + 9 + 25 + 36 + 81 = \dots$

▲ Q8. $S(a)$ est une fonction du second degré en a , de la forme $\alpha a^2 + \beta a + \gamma$.

1. Identifier α , β et γ .
2. Le coefficient α est-il positif ou négatif? En déduire si la parabole admet un minimum ou un maximum.

▲ Q9. On rappelle que le minimum d'une fonction $f(a) = \alpha a^2 + \beta a + \gamma$ (avec $\alpha > 0$) est atteint pour :

$$a_0 = -\frac{\beta}{2\alpha}$$

Calculer a_0 pour notre fonction S . Donner le résultat sous forme de fraction irréductible.

▲ Q10. Tracer la droite $y = a_0 x$ sur le repère de la question préliminaire. Cette droite vous semble-t-elle un bon résumé du nuage de points ?

3. Partie III – Généralisation

On cherche maintenant à comprendre la méthode en toute généralité.

Exercice 4: La formule générale | ★★★★★

On dispose de n points $M_1(x_1; y_1)$, $M_2(x_2; y_2)$, $\dots$, $M_n(x_n; y_n)$.

◆ Q11. Montrer que pour n points quelconques :

$$S(a) = \left(\sum_{i=1}^n x_i^2 \right) a^2 - 2 \left(\sum_{i=1}^n x_i y_i \right) a + \sum_{i=1}^n y_i^2$$

Indication : développer $(y_i - ax_i)^2$ puis sommer.

◆ Q12. En déduire que le coefficient a_0 qui minimise $S(a)$ vaut :

$$a_0 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

Vérifier que l'on retrouve bien la valeur trouvée à la question 9 avec nos 5 points.

♦ **Q13. (Ouverture)**

Et si l'on cherchait la meilleure droite de la forme $y = ax + b$ (qui ne passe pas forcément par l'origine)? On aurait alors :

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

C'est une fonction de **deux variables** a et b . Sa minimisation sera étudiée en première (ou en prépa!) — mais le principe reste le même : annuler les dérivées.

Annexe A – Rappels sur les fonctions du second degré

Rappel : extremum d'un trinôme

Soit $f(x) = \alpha x^2 + \beta x + \gamma$ un trinôme du second degré.

- Si $\alpha > 0$, la parabole est tournée vers le haut : f admet un **minimum**.
- Si $\alpha < 0$, la parabole est tournée vers le bas : f admet un **maximum**.

Dans les deux cas, l'extremum est atteint pour $x_0 = -\frac{\beta}{2\alpha}$.

La valeur de l'extremum est $f(x_0) = \gamma - \frac{\beta^2}{4\alpha}$.

Exemple : $f(x) = 3x^2 - 12x + 7$. On a $\alpha = 3 > 0$ donc f admet un minimum, atteint pour $x_0 = -\frac{-12}{2 \times 3} = 2$.

La valeur minimale est $f(2) = 3(4) - 12(2) + 7 = 12 - 24 + 7 = -5$.

Annexe B – Code Python (optionnel)

Objectif : Implémenter la formule trouvée en Python et tracer le résultat.

Exercice de programmation. Compléter le code ci-dessous.

Listing 1 – Calcul du coefficient optimal

```
1 def coefficient_optimal(x, y):
2     """
3     Calcule le coefficient a optimal pour y = ax
4     par la methode des moindres carres.
5     Entree : deux listes x et y de meme longueur.
6     Sortie : le reel a qui minimise S(a).
7     """
8     numerateur = ... # A completer : somme des x_i * y_i
9     denominateur = ... # A completer : somme des x_i ** 2
10    return numerateur / denominateur
11
12 # Donnees experimentales
13 x = [1, 2, 3, 4, 5]
14 y = [2, 3, 5, 6, 9]
15
16 a = coefficient_optimal(x, y)
17 print(f"Coefficient optimal : a = {a}")
18 print(f"Fraction : a = 92/55")
```

Bonus : Tracer le nuage de points et la droite optimale.

Listing 2 – Tracé avec matplotlib

```

1 import matplotlib.pyplot as plt
2
3 x = [1, 2, 3, 4, 5]
4 y = [2, 3, 5, 6, 9]
5 a = 92 / 55 # coefficient optimal
6
7 # Nuage de points
8 plt.scatter(x, y, color='red', label='Points experimentaux')
9
10 # Droite optimale
11 t = [0, 6]
12 plt.plot(t, [a * ti for ti in t], color='blue',
13           label=f'y = {a:.2f}x')
14
15 plt.xlabel('Masse (kg)')
16 plt.ylabel('Longueur (cm)')
17 plt.legend()
18 plt.grid(True)
19 plt.title('Regression lineaire par les moindres carres')
20 plt.show()

```

Fin du devoir

« La théorie des moindres carrés est, de toutes les branches de l'analyse, celle qui a été le plus utile aux sciences d'observation. »

— Adrien-Marie Legendre (1752–1833)