

Procédé de Gram-Schmidt, Méthode des moindres carrés

Cours #11



MTH1008 – Hiver 2026

Nathan Allaire, Théo Denorme, Sacha Benarroch-Lelong

- 1 Processus de Gram-Schmidt
- 2 Méthode des moindres carrés
- 3 Application - Moindres carrés

1 Processus de Gram-Schmidt

2 Méthode des moindres carrés

3 Application - Moindres carrés

Processus de Gram-Schmidt

Le **processus d'orthogonalisation de Gram-Schmidt (GS)** est une méthode itérative permettant d'obtenir une base **orthogonale** à partir d'une base quelconque.

L'idée est la suivante :

- 1 On considère E un s.e.v de $\mathbb{R}^n$, et $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$ une base de E .
- 2 On va construire petit à petit (vecteur par vecteur) une famille **orthogonale** à partir de $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p\}$.
- 3 Une fois cette famille complétée (c'est à dire après p itérations), on aura une base **orthogonale** de E .

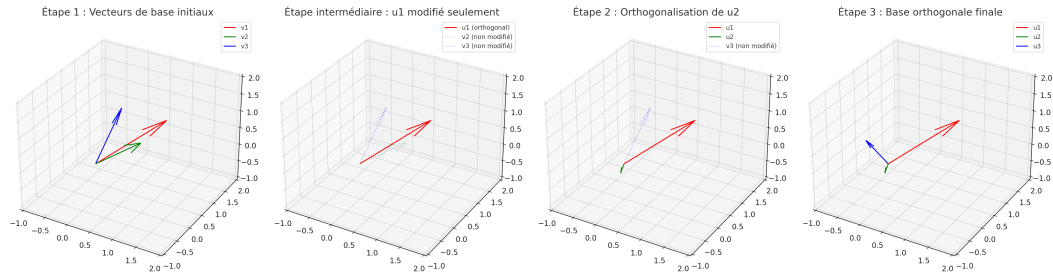


Figure – Exemple de la méthode de GS

Rappel (cours 10) :

Soit U un sev de $\mathbb{R}^n$, $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ une base orthogonale de U et $\mathbf{y}$ un vecteur de $\mathbb{R}^n$.

La projection orthogonale de $\mathbf{y}$ sur U est le vecteur :

$$\hat{\mathbf{y}} = \text{Proj}_U(\mathbf{y}) = \underbrace{\frac{\langle \mathbf{y}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1}_{\text{Proj}_{\text{Vect}\{\mathbf{u}_1\}}(\mathbf{y})} + \underbrace{\frac{\langle \mathbf{y}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2}_{\text{Proj}_{\text{Vect}\{\mathbf{u}_2\}}(\mathbf{y})} + \dots + \underbrace{\frac{\langle \mathbf{y}, \mathbf{u}_p \rangle}{\langle \mathbf{u}_p, \mathbf{u}_p \rangle} \mathbf{u}_p}_{\text{Proj}_{\text{Vect}\{\mathbf{u}_p\}}(\mathbf{y})}$$

De plus, le vecteur $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$ est orthogonal à tous les vecteurs de U .

La famille $U' = \{\mathbf{u}_1, \dots, \mathbf{u}_p, \mathbf{z}\}$ est donc une famille orthogonale de dimension $p + 1$.

On peut continuer ce processus jusqu'à ce que $p = n$ en prenant à chaque fois un vecteur $\mathbf{y} \notin \text{Vect}\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$

Algorithme : méthode de GS

Soit $\{\mathbf{y}_1, \dots, \mathbf{y}_p\}$ une base d'un sev V de $\mathbb{R}^n$. L'algorithme suivant produit une base orthogonale de V :

1 Choisir $\mathbf{u}_1 = \mathbf{y}_1$

2 pour $k \in [2, p]$, choisir

$$\mathbf{u}_k = \mathbf{y}_k - \underbrace{\sum_{i=1}^{k-1} \frac{\mathbf{u}_i \cdot \mathbf{y}_k}{\mathbf{u}_i \cdot \mathbf{u}_i} \times \mathbf{u}_i}_{\text{Proj}_{\text{Vect}\{\mathbf{u}_1, \dots, \mathbf{u}_{k-1}\}}(\mathbf{y}_k)}$$

3 La nouvelle famille $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ est une base orthogonale de V :
 $\text{Vect}\{\mathbf{u}_1, \dots, \mathbf{u}_p\} = \text{Vect}\{\mathbf{y}_1, \dots, \mathbf{y}_p\}$.

Q : comment construire une base **orthonormée** à partir de $\{\mathbf{y}_1, \dots, \mathbf{y}_p\}$?

Exemple :

Donner une base orthonormée de Vect $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$

Théorème de factorisation QR

Soit $A \in \mathbb{R}^{m \times n}$. Si les colonnes de A sont linéairement indépendantes, alors il existe

- $Q \in \mathbb{R}^{m \times n}$: dont les colonnes sont orthonormées
- $R \in \mathbb{R}^{n \times n}$: triangulaire supérieure inversible dont les coefficients diagonaux sont strictement positifs

telles que $A = QR$.

→ les éléments diagonaux de R correspondent aux normes des colonnes de Q .

Trouver une factorisation QR de $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Indice : $A = QR \implies Q^T A = Q^T QR$

1 Processus de Gram-Schmidt

2 Méthode des moindres carrés

3 Application - Moindres carrés

Parfois, il n'est pas possible de résoudre le système $A\mathbf{x} = \mathbf{b}$: on dit alors que le système est incompatible.

Cela dit, on peut essayer de trouver un vecteur $\mathbf{x}'$ tel que $A\mathbf{x}'$ soit "presque égal" à $\mathbf{b}$.

On peut même essayer de trouver le **meilleur** $\hat{\mathbf{x}}$ tel que $A\hat{\mathbf{x}}$ soit le plus proche possible de $\mathbf{b}$, c'est à dire que pour tout $\mathbf{x}$ de $\mathbb{R}^n$,

$$0 < \|A\hat{\mathbf{x}} - \mathbf{b}\| \leq \|A\mathbf{x} - \mathbf{b}\| \quad (1)$$

Autrement dit, on cherche à résoudre :

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\| \quad (2)$$

Définition

Soit $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$.

On appelle **solution au sens des moindres carrés** tout vecteur $\hat{x}$ tel que

$$\|A\hat{x} - b\| = \min_{x \in \mathbb{R}^n} \|Ax - b\|. \quad (3)$$

→ on recherche $\hat{b} \in \text{Im } A$ (avec $\hat{b} = A\hat{x}$) qui minimise la distance entre $A\hat{x}$ et b .

D'après le cours n°10,

$$\hat{b} = \text{Proj}_{\text{Im } A}(b) \quad (4)$$

Comme $\hat{b}$ appartient à $\text{Im } A$, le système $Ax = \hat{b}$ est **compatible**.

→ L'ensemble des solutions au sens des moindres carrés est donc l'ensemble des solutions de $Ax = \hat{b}$.

Supposons que $\mathbf{x}$ vérifie $A\mathbf{x} = \hat{\mathbf{b}}$, avec $\hat{\mathbf{b}} = \text{Proj}_{\text{Im } A}(\mathbf{b})$.

D'après le th. de la projection orthogonale (cours 10), le vecteur $\mathbf{b} - \hat{\mathbf{b}}$ est orthogonal à $\text{Im } A$.

Donc $\mathbf{b} - \hat{\mathbf{b}}$ est orthogonal à chacune des colonnes de A :

$$A^T(\mathbf{b} - \hat{\mathbf{b}}) = \mathbf{0}$$

$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$$

$$A^T\mathbf{b} - A^T A\mathbf{x} = \mathbf{0}$$

En somme, toute solution au sens des moindres carrés $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ vérifie l'équation :

$$A^T A\mathbf{x} = A^T\mathbf{b}. \quad (5)$$

Maintenant qu'on a prouvé que toute solution au sens des moindres carrés est une solution de l'équation normale il faut prouver l'autre sens, c'est à dire montrer que toute solution de l'équation normale est une solution au sens des moindres carrés :

Soit $\mathbf{x}$ qui vérifie $A^T A \mathbf{x} = A^T \mathbf{b}$ alors :

$$A^T A \mathbf{x} = A^T \mathbf{b} \Leftrightarrow A^T A \mathbf{x} - A^T \mathbf{b} = \mathbf{0}$$

$$A^T (A \mathbf{x} - \mathbf{b}) = \mathbf{0}$$

$$A^T (A \mathbf{x} - \hat{\mathbf{b}}) = \mathbf{0}.$$

On a que $A \mathbf{x} - \hat{\mathbf{b}} \in \text{Ker } A^T$, de plus comme $A \mathbf{x} \in \text{Im } A$ et $\hat{\mathbf{b}} \in \text{Im } A$ on a que $A \mathbf{x} - \hat{\mathbf{b}} \in \text{Im } A$ donc :

$$A \mathbf{x} - \hat{\mathbf{b}} \in (\text{Ker } A^T) \cap (\text{Im } A) = \{\mathbf{0}\} \text{ car } \text{Ker } A^T \perp \text{Im } A$$

Donc $A \mathbf{x} - \hat{\mathbf{b}} = \mathbf{0}$ et enfin on a bien que :

$$A \mathbf{x} = \hat{\mathbf{b}}$$

Définition

Soit $A \in \mathbb{R}^{m \times n}$ et $\mathbf{b} \in \mathbb{R}^m$ On appelle **équations normales** associées au système $A\mathbf{x} = \mathbf{b}$ les n équations :

$$A^T A \mathbf{x} = A^T \mathbf{b}. \quad (6)$$

Théorème

L'ensemble des solutions au sens des moindres carrés de l'équation $A\mathbf{x} = \mathbf{b}$ est égal à l'ensemble non-vidé des solutions des équations normales de $A\mathbf{x} = \mathbf{b}$.

Déterminer les solutions au sens des moindres carrés dans les cas suivants :

$$\mathbf{1} \quad A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, b = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$$

$$\mathbf{2} \quad A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}$$

Quel est l'ensemble des solutions au sens des moindres carrés de l'équation $Ax = b$ lorsque le système est **compatible** ?

Théorème

Soit $A \in \mathbb{R}^{m \times n}$. Les propositions suivantes sont équivalentes :

- 1 Pour tout vecteur $\mathbf{b} \in \mathbb{R}^m$, l'équation $A\mathbf{x} = \mathbf{b}$ admet une unique solution au sens des moindres carrés,
- 2 A est de plein rang colonne,
- 3 $A^T A$ est inversible.

Dans le cas où l'une de ces propositions est vérifiée,

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}.$$

Q : que vaut $(A^T A)^{-1} A^T \mathbf{b}$ lorsque A est inversible ?

Note : on appelle la matrice $(A^T A)^{-1} A^T$ la pseudo-inverse de A .

Définition

Soit $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$. On appelle **erreur quadratique** associée à la solution au sens des moindres carrés $\hat{x}$ la quantité :

$$e = \|A\hat{x} - b\|.$$

Elle désigne la distance entre la b et la meilleure approximation de b par un vecteur de $\text{Im } A$.

Questions :

- 1 que vaut e si $b \in \text{Im } A$?
- 2 existe-t-il un $\bar{x}$ tel que $\|A\bar{x} - b\| < e$?

Résoudre $A\hat{\mathbf{x}} = \mathbf{b}$

On se souvient que l'on recherche la meilleure solution de $A\mathbf{x} = \mathbf{b}$, et l'on sait que cette dernière est en fait la solution du système **compatible** : $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, où $\hat{\mathbf{b}} = \text{Proj}_{\text{Im } A}(\mathbf{b})$.

On peut donc résoudre ce système pour obtenir l'ensemble des solutions au sens des moindres carrés.

→ inconvénient : on a besoin d'une base orthogonale de $\text{Im } A$ pour projeter.

Résoudre avec la factorisation QR

Soit A une matrice de plein rang colonne (i.e ses colonnes sont linéairement indépendantes), et $A = QR$ une factorisation QR de A .

Pour tout $\mathbf{b} \in \mathbb{R}^m$, l'équation $A\mathbf{x} = \mathbf{b}$ admet une **unique** solution au sens des moindres carrés, donnée par la formule :

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b} \quad (7)$$

→ inconvénient : on a besoin de Q et R .

- 1 Processus de Gram-Schmidt
- 2 Méthode des moindres carrés
- 3 Application - Moindres carrés**

Après des recherches, un géologue déduit que l'érosion des sols dans une région donnée (en mm/an) est uniquement dépendante des précipitations annuelles et mesure les données suivantes :

Précipitations annuelles (dm)	Erosion (mm/an)
2	2.4
4	4.9
6	6.5
8	9.3

Table – Tableau représentant l'érosion des sols en fonction des précipitations annuelles

- 1 Tracer les 4 points sur un même graphique
- 2 En supposant que la relation est linéaire, écrire le problème sous la forme d'un SEL
- 3 Ce système est-il compatible ?
- 4 En utilisant la méthode des moindres carrés, établir un modèle qui donne l'érosion prévue en fonction des précipitations annuelles
- 5 D'après votre modèle, quelle sera l'érosion prévue s'il pleut 5.5dm en une année ?

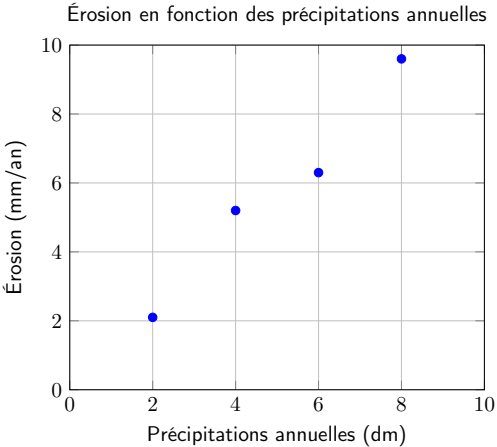


Figure – Graphique des points observés de précipitations et d'érosion

- 1 Voir graphique
- 2 On écrit les 4 couples (P, E) en suposant que la relation qui les lie est linéaire :
 $e(x) = ax + b$

$$a \times 2 + b = 2.4$$

$$a \times 4 + b = 4.9$$

$$a \times 6 + b = 6.5$$

$$a \times 8 + b = 9.3$$

$$\text{Donc on a } A = \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{bmatrix} \text{ et } B = \begin{bmatrix} 2.4 \\ 4.9 \\ 6.5 \\ 9.3 \end{bmatrix} \text{ avec } x = \begin{bmatrix} a \\ b \end{bmatrix}$$

3 Non, car A est de rang 2 au maximum, et on a au moins 3 équations linéairement indépendantes.

4 On recherche une solution au sens des moindres carrés en passant par les équations normales : $A^T A x = A^T b$

$$\text{Ici, } A^T A = \begin{bmatrix} 120 & 20 \\ 20 & 4 \end{bmatrix} \text{ et } A^T b = \begin{bmatrix} 137.8 \\ 23.1 \end{bmatrix}$$

Les colonnes de $A^T A$ sont linéairement indépendantes, donc la solution est

$$\hat{x} = (A^T A)^{-1} A^T b = \begin{bmatrix} 1.115 \\ 0.2 \end{bmatrix}$$

5 $1.115 \times 5.5 + 0.2 = 6.815$

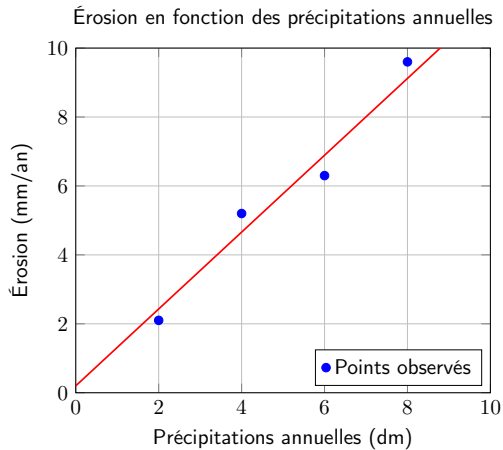


Figure – Graphique des points observés de précipitations et d'érosion avec la droite de régression

Exercices récapitulatifs

Répondez par vrai ou faux en justifiant :

1. On ne peut pas effectuer le procédé de Gram-Schmidt (GS) sur la famille

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

2. Effectuer GS sur une base ne contenant qu'un vecteur non nul la laisse inchangée.

3. Soit $B = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ une base de $\mathbb{R}^3$. Effectuer GS sur B transforme $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ en

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

4. Dans la décomposition QR d'une matrice A , il est possible que $Q = R$.

5. Une décomposition QR (effectuée avec GS) de $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ est $A = QR$ avec $Q = I_3$

et $R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

6. Dans la décomposition QR d'une matrice, la matrice R est toujours carrée.
7. Si A est une matrice inversible alors l'ensemble des solutions au sens des moindres carrés de $Ax = b$ contient une infinité de points.
8. Si $Ax = b$ possède une solution, alors l'ensemble des solutions de l'équation normale est égale à l'ensemble des solutions de $Ax = b$.
9. Si $A = [a] \in \mathbb{R}^{1 \times 1}$ ($a \neq 0$) alors l'ensemble des solutions au sens des moindres carrés de $Ax = [b]$ avec $b \in \mathbb{R}$ est exactement l'ensemble des points qui minimisent le polynôme $p(x) = a^2x^2 - 2bax + b^2$.
10. (**difficile**) Si A est une matrice carrée non-inversible alors l'ensemble des solutions au sens des moindres carrés contient une infinité de points.
11. Soit $A \in \mathbb{R}^{n \times n}$ tel que 1 est une valeur propre de A . Soit $b \in \text{Ker}(A - I)$, alors $\min_{x \in \mathbb{R}^n} \|Ax - b\| = 0$.

Soit $A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 3 & 7 \end{bmatrix}$

- 1) A est-elle de plein rang colonne ?
- 2) Effectuez la décomposition QR de A avec GS.

Une usine de cookies souhaite effectuer des estimations du prix du chocolat au kilo, pour savoir si l'entreprise restera rentable au cours des années. On note le couple

$(x_i, y_i) = (\text{année}, \text{prix au kilo du chocolat})$, l'usine a dans sa base de données les couples $(1, 10)$, $(2, 20)$ et $(4, 15)$ qui correspondent au prix au kilo du chocolat des années 1, 2 et 4. Elle souhaite savoir quel sera environ le prix au kilo du chocolat à l'année 5.

Un des employés ayant suivi le cours d'algèbre de Nathan, Théo et Sacha a eu une idée : "il faudrait trouver une droite qui passe au plus proche possible des points $(1, 10)$, $(2, 20)$ et $(4, 15)$!" .

Si on formalise ça, cela revient à trouver les paramètres a et b de la droite $y = ax + b$ tel que :

$$\|a \times 1 + b - 10\| + \|a \times 2 + b - 20\| + \|a \times 4 + b - 15\|$$

Est minimale.

L'employé très malin et passionné par les mathématiques (car le cours était trop intéressant) s'est rendu compte que cela revenait exactement au même que de trouver a et b qui permettent d'atteindre :

$$\min_{a,b \in \mathbb{R}} \left\| \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} 10 \\ 20 \\ 15 \end{bmatrix} \right\|$$

Malheureusement sur l'horloge 17h a sonné, la journée de l'employé s'est terminée et il est en congés pour 2 semaines. L'usine a besoin de vous pour trouver a et b tels que la droite d'équation $y = ax + b$ est une approximation de l'évolution du prix du chocolat en fonction des années.

Trouvez a et b puis estimez le prix du chocolat au kilo à l'année 5.

Soit $A \in \mathbb{R}^{m \times n}$ et Q, R tel que $A = QR$ est la décomposition QR de A .

- 1) Soit $x \in \text{Ker } R$ montrez que $x \in \text{Ker } A$ et que $Rx \in \text{Ker } Q$.
- 2) Soit $x \in \text{Ker } A$ montrez que $x \in \text{Ker } R$.
- 3) Pourquoi $\text{Ker } Q = \{0\}$?

Soit $A = [v_1 \ v_2]$ avec $v_2 = \alpha v_1$ ($\alpha \in \mathbb{R}$ et $v_1, v_2 \in \mathbb{R}^n$).

- 1) Montrez que $\text{Im } A = \text{Vect } \{v_1\}$.
- 2) Soit $b \in \mathbb{R}^2$, exprimez $\hat{b} = \text{Proj}_{\text{Im } A}(b)$ en fonction de b et v_1 .
- 3) Justifiez pourquoi il existe $\hat{x} \in \mathbb{R}^2$ tel que $\hat{b} = A\hat{x}$.
- 4) Montrez que $(b - \hat{b}) \cdot v_1 = 0$ et déduisez que $(b - \hat{b}) \cdot v_2 = 0$.
- 5) Ainsi $A^\top(b - \hat{b}) = 0$, montrez que $A^\top b = A^\top A\hat{x}$.
- 6) Pourquoi $A^\top b \in \text{Im } A^\top A$?
- 7) L'équation normale de $\begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ est elle compatible ? Si oui donnez l'ensemble de ses solutions.