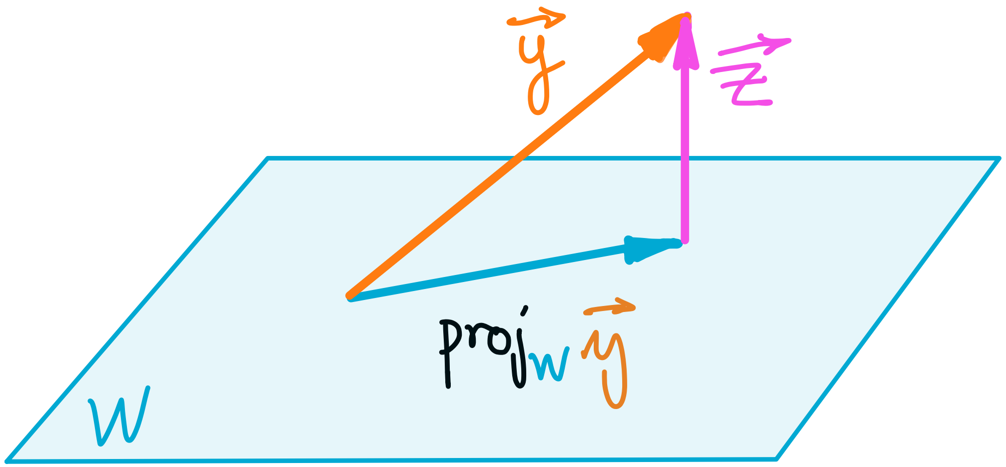
Théorème de la meilleure approximation.

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ et $\vec{y} \in \mathbb{R}^n$. On a :

$$\|\vec{y} - \mathrm{proj}_W \vec{y}\| < \|\vec{y} - \vec{w}\|$$

pour tout $\vec{w} \in W$, $\vec{w} \neq \mathrm{proj}_W \vec{y}$.



Le vecteur $\mathrm{proj}_W \vec{y}$ est appelé meilleure approximation (quadratique) de $\vec{y}$ par un élément de W.

Preuve.

Comme $\vec{y} - \text{proj}_W \vec{y} \in W^{\perp}$

et $\vec{w} - \text{proj}_W \vec{y} \in W$ pour tout $\vec{w} \in W$

nous pouvons écrire

$$\vec{y} - \vec{w} = (\vec{y} - \text{proj}_W \vec{y}) + (\text{proj}_W \vec{y} - \vec{w})$$

et le théorème de Pythagore nous donne

$$\|\vec{y} - \vec{w}\|^2 = \|\vec{y} - \text{proj}_W \vec{y}\|^2 + \underbrace{\|\text{proj}_W \vec{y} - \vec{w}\|^2}_{>0 \text{ si } \vec{w} \neq \text{proj}_W \vec{y}}$$

$$> \|\vec{y} - \text{proj}_W \vec{y}\|^2$$

d'où $\|\vec{y} - \vec{w}\| > \|\vec{y} - \text{proj}_W \vec{y}\|$ pour tout $\vec{w} \in W, \vec{w} \neq \text{proj}_W \vec{y}$. ■

Nous venons de voir que $\mathrm{proj}_W \vec{y}$ peut se calculer facilement si l'on dispose d'une base orthogonale de W.

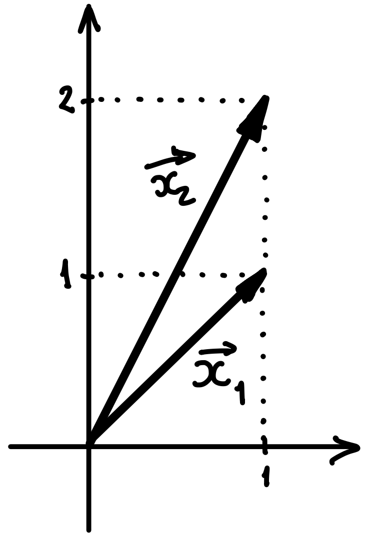
Question. Comment trouver une telle base ?

Réponse :

6.4. Le procédé d'orthogonalisation de Gram-Schmidt.

Considérons un premier exemple (très simple) :

Les vecteurs $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ et $\vec{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ne sont pas orthogonaux car $\vec{x}_1 \cdot \vec{x}_2 = 1+2 = 3 \neq 0$.

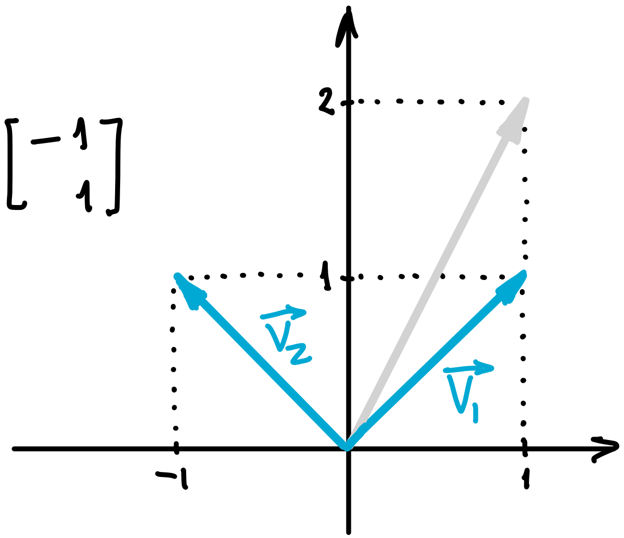


Soit $W = \mathrm{Vect}\{\vec{x}_1, \vec{x}_2\}$ (ici $W = \mathbb{R}^2$)

On pose $\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ et $W_1 = \mathrm{Vect}\{\vec{v}_1\}$

On a $\mathrm{proj}_{W_1} \vec{x}_2 = \dfrac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 = \dfrac{1+2}{1+1} \vec{v}_1 = \dfrac{3}{2} \vec{v}_1$

et $\vec{x}_2 - \mathrm{proj}_{W_1} \vec{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{3}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$



Le vecteur $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ est orthogonal à $\vec{v}_1$

car $\vec{v}_2 \cdot \vec{v}_1 = -1 + 1 = 0$.

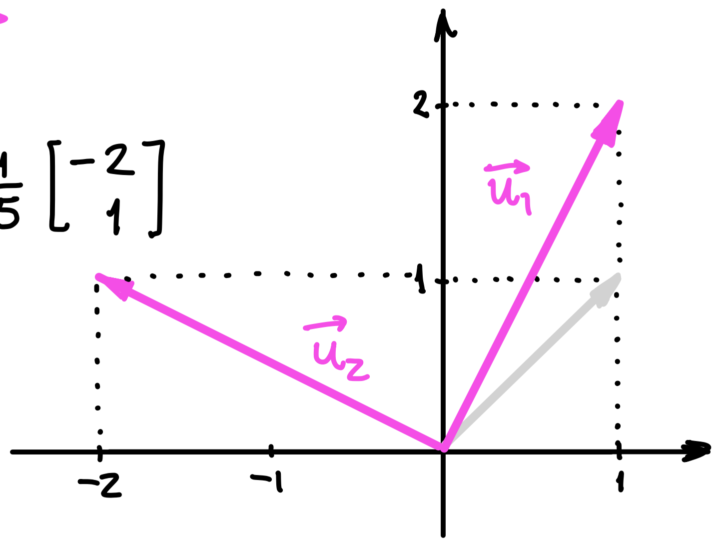
Par conséquent $\{\vec{v}_1, \vec{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ est une base orthogonale de W.

Remarque.

Si l'on pose $\vec{u}_1 = \vec{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ et $W_1 = \text{Vect}\{\vec{u}_1\}$

On a $\text{proj}_{W_1} \vec{x}_1 = \dfrac{\vec{x}_1 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \dfrac{1+2}{1+4} \vec{u}_1 = \dfrac{3}{5} \vec{u}_1$

et $\vec{x}_1 - \text{proj}_{W_1} \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \dfrac{3}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/5 \\ -1/5 \end{bmatrix} = -\dfrac{1}{5} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$



Le vecteur $\vec{u}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ est orthogonal à $\vec{u}_1$

car $\vec{u}_1 \cdot \vec{u}_2 = -2 + 2 = 0$.

Par conséquent $\{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ est une autre base orthogonale de W.

Attention :

L'ordre dans lequel on choisit les vecteurs est important.

Procédé d'orthogonalisation de Gram-Schmidt.

1. On se donne une base $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ de W.

2. On pose $\vec{v}_1 = \vec{x}_1$ et on note $W_1 = \mathrm{Vect}\{\vec{v}_1\}$ $(= \mathrm{Vect}\{\vec{x}_1\})$

3. On calcule $\mathrm{proj}_{W_1} \vec{x}_2$ et $\vec{x}_2 - \mathrm{proj}_{W_1} \vec{x}_2$

4. On pose $\vec{v}_2 = \vec{x}_2 - \mathrm{proj}_{W_1} \vec{x}_2$ (ou un multiple de ce vecteur de sorte que $\vec{v}_2$ ait des composantes entières).

   et on note $W_2 = \mathrm{Vect}\{\vec{v}_1, \vec{v}_2\}$ $(= \mathrm{Vect}\{\vec{x}_1, \vec{x}_2\})$

   Par construction, $\vec{v}_1 \cdot \vec{v}_2 = 0$ et $\{\vec{v}_1, \vec{v}_2\}$ est une base orthogonale de W_2.

5. On calcule $\text{proj}_{W_2}\vec{x}_3$ et $\vec{x}_3 - \text{proj}_{W_2}\vec{x}_3$

6. On pose $\vec{v}_3 = \vec{x}_3 - \text{proj}_{W_2}\vec{x}_3$ (ou un multiple de ce vecteur de sorte que $\vec{v}_3$ ait des composantes entières).

   et on note $W_3 = \text{Vect}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ $(= \text{Vect}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\})$

   Par construction, $\vec{v}_3 \cdot \vec{v}_1 = 0$ et $\vec{v}_3 \cdot \vec{v}_2 = 0$ et $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ est une base orthogonale de W_3.

7. On continue de la sorte jusqu'à la dernière étape :

   On pose $\vec{v}_k = \vec{x}_k - \text{proj}_{W_{k-1}}\vec{x}_k$ (ou un multiple de ce vecteur de sorte que $\vec{v}_k$ ait des composantes entières).

   L'ensemble $\{\vec{v}_1, \dots, \vec{v}_k\}$ est une base orthogonale de W.

Exemple.

Soient $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{x}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$

Trouver une base orthogonale de $W = \text{Vect}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$

On pose $\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ et $W_1 = \text{Vect}\{\vec{v}_1\}$ $(= \text{Vect}\{\vec{x}_1\})$

On calcule $\text{proj}_{W_1} \vec{x}_2 = \dfrac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 = \dfrac{-1+2+0+0+1}{1+1+0+1+1} \vec{v}_1 = \dfrac{2}{4} \vec{v}_1 = \dfrac{1}{2} \vec{v}_1$

et $\vec{x}_2 - \text{proj}_{W_1} \vec{x}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \dfrac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 3/2 \\ 1 \\ -1/2 \\ 1/2 \end{bmatrix} = \dfrac{1}{2} \begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}$

On pose $\vec{v}_2 = 2(\vec{x}_2 - \text{proj}_{W_1} \vec{x}_2) = \begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}$ et $W_2 = \text{Vect}\{\vec{v}_1, \vec{v}_2\}$

On vérifie: $\vec{v}_1 \cdot \vec{v}_2 = -3+3+0-1+1=0$

On calcule $\text{proj}_{W_2} \vec{x}_3 = \dfrac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \dfrac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2$

$$= \frac{0+1+0+1+2}{1+1+0+1+1} \vec{v}_1 + \frac{0+3+2-1+2}{9+9+4+1+1} \vec{v}_2 = \vec{v}_1 + \frac{1}{4} \vec{v}_2$$

et $\vec{x}_3 - \text{proj}_{W_2} \vec{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{4} \begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ -3/4 \\ 1/2 \\ 1/4 \\ 3/4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \\ 3 \end{bmatrix}$

On pose $\vec{v}_3 = 4 \left(\vec{x}_3 - \text{proj}_{W_2} \vec{x}_3 \right) = \begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \\ 3 \end{bmatrix}$

On vérifie: $\vec{v}_1 \cdot \vec{v}_3 = -1-3+0+1+3=0$ et $\vec{v}_2 \cdot \vec{v}_3 = 3-9+4-1+3=0$

Par conséquent, $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \\ 3 \end{bmatrix} \right\}$ est une base orthogonale de W.

Factorisation QR

Théorème. Soit A une matrice de taille $m \times n$ dont les colonnes sont linéairement indépendantes (c'est-à-dire, telle que $\operatorname{rang}(A) = n$)

Alors la matrice A peut s'écrire sous la forme

$$\boxed{A = QR}$$

où Q est une matrice de taille $m \times n$ dont les colonnes forment une base orthonormale de $\operatorname{Col}(A)$

et R est une matrice de taille $n \times n$ triangulaire supérieure inversible à coefficients diagonaux strictement positifs.

Comme par hypothèse $\text{rang}(A) = n$, les n colonnes de A forment une base $\{\vec{a}_1, \ldots, \vec{a}_n\}$ de $\text{Col}(A)$.

Le procédé d'orthogonalisation de Gram-Schmidt nous fournit une base orthogonale $\{\vec{v}_1, \ldots, \vec{v}_n\}$ de $\text{Col}(A)$

Par conséquent, les vecteurs $\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$, $\vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|}$, $\ldots$, $\vec{u}_n = \frac{\vec{v}_n}{\|\vec{v}_n\|}$ forment une base orthonormale de $\text{Col}(A)$.

On pose $\boxed{Q = [\vec{u}_1 \ \vec{u}_2 \ \cdots \ \vec{u}_n]}$

Comme par construction $Q^T Q = I_n$ nous avons:

$$Q^T A = Q^T(QR) = (Q^T Q)R = R$$

d'où $\boxed{R = Q^T A}$

Exemple.

Trouver une factorisation QR de $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

Nous avons trouvé que

$$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \\ 3 \end{bmatrix} \right\}$$

est une base orthogonale de Col(A).

$$\left.\begin{aligned} \|\vec{v}_1\|^2 &= 1+1+0+1+1 = 4 \\ \|\vec{v}_2\|^2 &= 9+9+4+1+1 = 24 \\ \|\vec{v}_3\|^2 &= 1+9+4+1+9 = 24 \end{aligned}\right\} \Rightarrow \{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \left\{ \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2\sqrt{6}}\begin{bmatrix} -3 \\ 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \frac{1}{2\sqrt{6}}\begin{bmatrix} -1 \\ -3 \\ 2 \\ 1 \\ 3 \end{bmatrix} \right\}$$

est une base orthonormale de Col(A).

On pose $$Q = \frac{1}{2\sqrt{6}} \begin{bmatrix} \sqrt{6} & -3 & -1 \\ \sqrt{6} & 3 & -3 \\ 0 & 2 & 2 \\ \sqrt{6} & -1 & 1 \\ \sqrt{6} & 1 & 3 \end{bmatrix}$$

et $$R = Q^T A = \frac{1}{2\sqrt{6}} \begin{bmatrix} \sqrt{6} & \sqrt{6} & 0 & \sqrt{6} & \sqrt{6} \\ -3 & 3 & 2 & -1 & 1 \\ -1 & -3 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$= \frac{1}{2\sqrt{6}} \begin{bmatrix} 4\sqrt{6} & 2\sqrt{6} & 4\sqrt{6} \\ 0 & 12 & 6 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\Rightarrow \quad R = \begin{bmatrix} 2 & 1 & 2 \\ 0 & \sqrt{6} & \frac{1}{2}\sqrt{6} \\ 0 & 0 & \frac{1}{2}\sqrt{6} \end{bmatrix}$$

Remarque.

Tout comme la factorisation LU, la factorisation QR est très utile dans la résolution numérique de systèmes $A\vec{x}=\vec{b}$, où A est une matrice de taille $m\times n$ avec m,n grands. La factorisation QR est aussi utilisée pour le calcul des valeurs propres de la matrice carrée A.

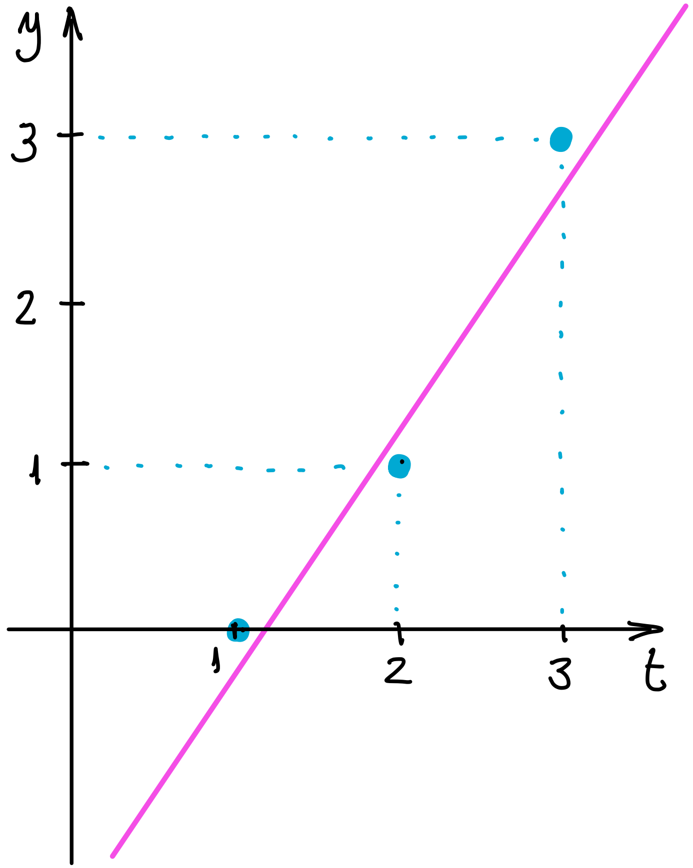
6.5. La méthode des moindres carrés.

Motivation.

Considérons trois points du plan :

$$(1,0),\ (2,1),\ (3,3)$$

On constate qu'ils ne sont pas alignés !



Question.

Quelle est l'équation de la droite du plan la plus "proche" de ces points ?

Soit $y = x_1 + x_2 t$ l'équation cherchée, où x_1, x_2 sont à déterminer.

On aimerait que les trois points donnés soient sur la droite :

$$
\begin{array}{ll}
(t_1, y_1) = (1, 0) : & \\
(t_2, y_2) = (2, 1) : & \\
(t_3, y_3) = (3, 3) : &
\end{array}
\left\{
\begin{array}{l}
0 = x_1 + 1\, x_2 \\
1 = x_1 + 2\, x_2 \\
3 = x_1 + 3\, x_2
\end{array}
\right.
$$

Equation matricielle associée : $A\vec{x} = \vec{b}$ où $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ et $\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$

matrice augmentée :

$$
\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \end{array}\right]
\underset{\substack{L_2 \to L_2 - L_1 \\ L_3 \to L_3 - L_1}}{\sim}
\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 3 \end{array}\right]
\underset{L_3 \to L_3 - 2L_2}{\sim}
\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]
$$

ce qui confirme que les points ne sont pas alignés.