MAT380

Algèbre Linéaire

Jean-Michel Lemay
jean-michel.lemay@etsmtl.ca

Chapitre 8 : Orthogonalité et moindres carrés

- 8.1 Rappel sur le produit scalaire
- 8.2 Complément orthogonal
- 8.3 Projection orthogonale
- 8.4 Bases orthogonales
- 8.5 Méthode des moindres carrés

Objectifs d’apprentissage

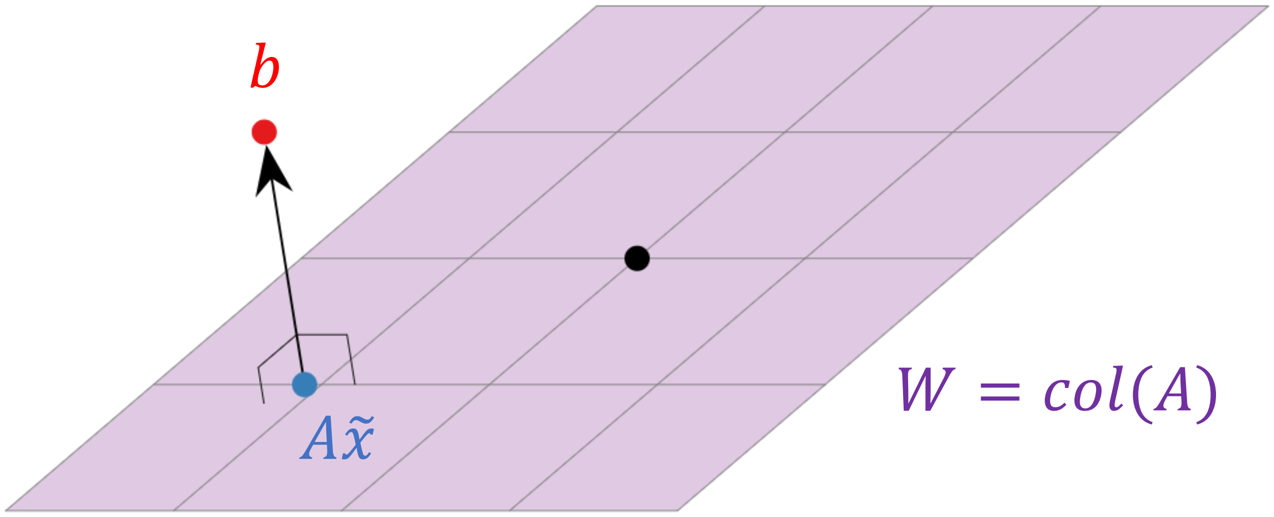
- Calculer des compléments orthogonaux
- Calculer des décompositions orthogonales
- Appliquer le procédé de Gram-Schmidt
- Résoudre des problèmes à l’aide de la méthode des moindres carrés

Mots clés : Complément orthogonal, projection orthogonale, décomposition orthogonale, base orthonormale, méthode des moindres carrés

Orthogonalité et moindres carrés

Si un système d'équations linéaires $Ax = b$ n'a pas de solution, peut-on trouver une approximation $\tilde{x}$ pour que $A\tilde{x}$ soit le plus proche possible de b ? $(A\tilde{x} \approx b)$

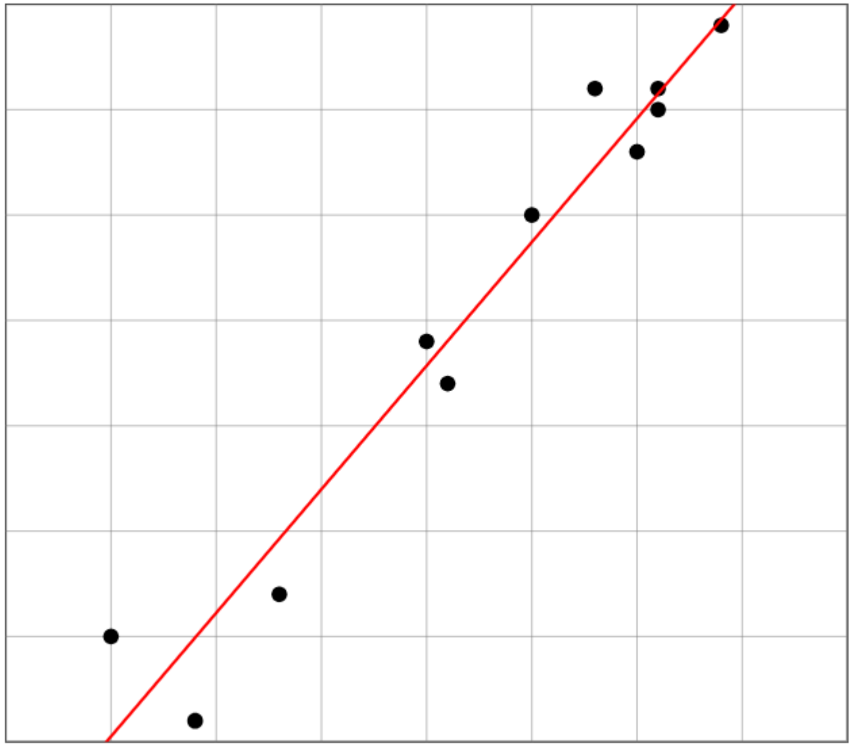
Ceci demande de trouver le point d'un sous-espace le plus près possible de b.



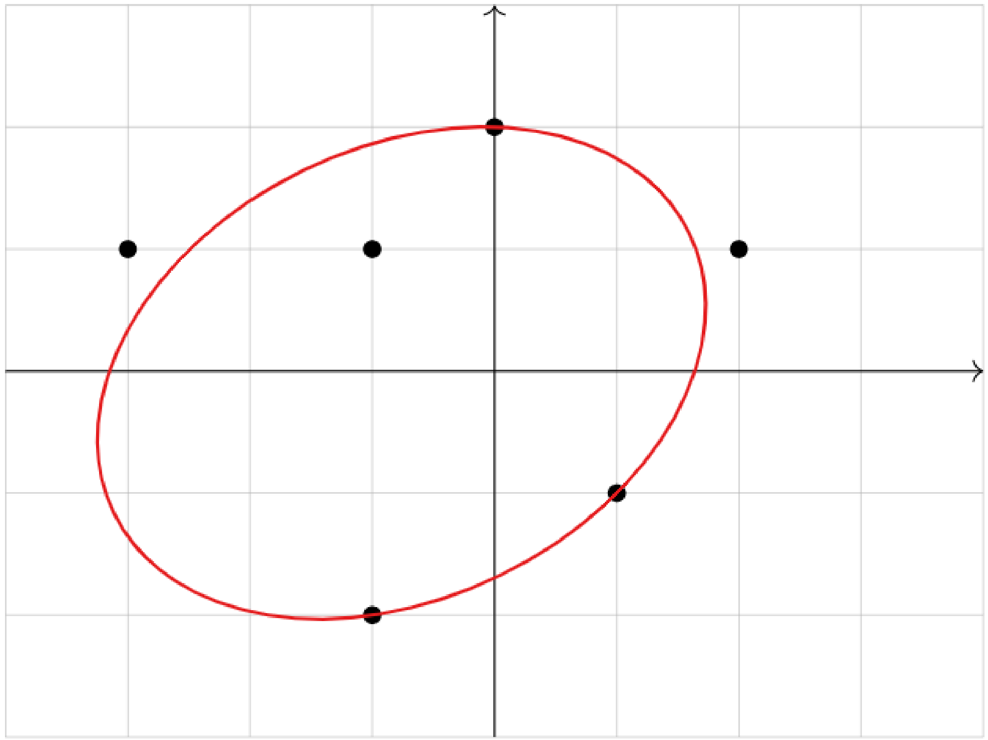
Modélisation de données

Plusieurs problèmes de modélisation de données peuvent être du type $Ax \approx b$

Régression linéaire



Orbites elliptiques



8.1 Rappel sur le produit scalaire

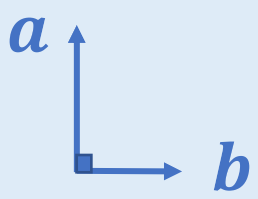
Rappels : Soient $\boldsymbol{a} = (a_1, \ldots, a_n)^T$ et $\boldsymbol{b} = (b_1, \ldots, b_n)^T$ deux vecteurs de $\mathbb{R}^n$.

- $\boldsymbol{a} \cdot \boldsymbol{b} = \|\boldsymbol{a}\| \|\boldsymbol{b}\| \cos(\theta)$ où θ est le plus petit angle entre les deux vecteurs.
- $\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$
- dotp(a,b) sur la TI
- $\|\boldsymbol{a}\| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} = \sqrt{a_1^2 + \cdots + a_n^2}$
- $dist(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{b} - \boldsymbol{a}\|$
- Vecteur $\boldsymbol{v}$ est *unitaire* si : $\|\boldsymbol{v}\| = \boldsymbol{1}$
- $\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|}$ est toujours unitaire pour $\boldsymbol{a} \neq \boldsymbol{0}$

Deux propriétés importantes

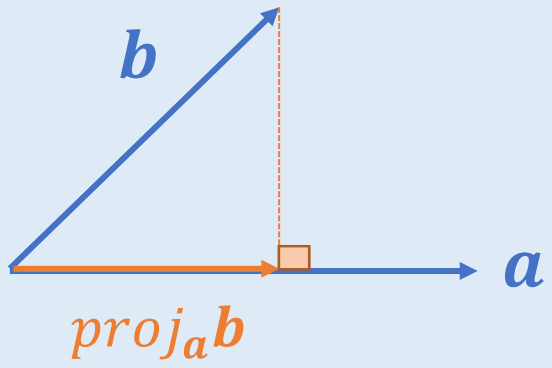
- Deux vecteurs sont orthogonaux (perpendiculaires) si et seulement si leur produit scalaire est nul.

$$\boldsymbol{a} \perp \boldsymbol{b} \Leftrightarrow \boldsymbol{a} \cdot \boldsymbol{b} = 0$$



- La projection de $\boldsymbol{b}$ sur $\boldsymbol{a}$ est donnée par

$$proj_{\boldsymbol{a}}\boldsymbol{b} = \left(\frac{\boldsymbol{b} \cdot \boldsymbol{a}}{\boldsymbol{a} \cdot \boldsymbol{a}}\right)\boldsymbol{a}$$



Exemple 1 :

$$\boldsymbol{u} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad \boldsymbol{v} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad \boldsymbol{w} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

a) Donnez un vecteur unitaire dans la direction de $\boldsymbol{u}$.

b) Déterminez si les vecteurs sont orthogonaux.

c) Calculez la projection de **w** sur **v**.

8.2 Complément orthogonal

On s'intéresse dans cette section à une notion importante pour la suite :

Comment trouver tous les vecteurs qui sont orthogonaux à un ensemble de vecteurs donnés ?

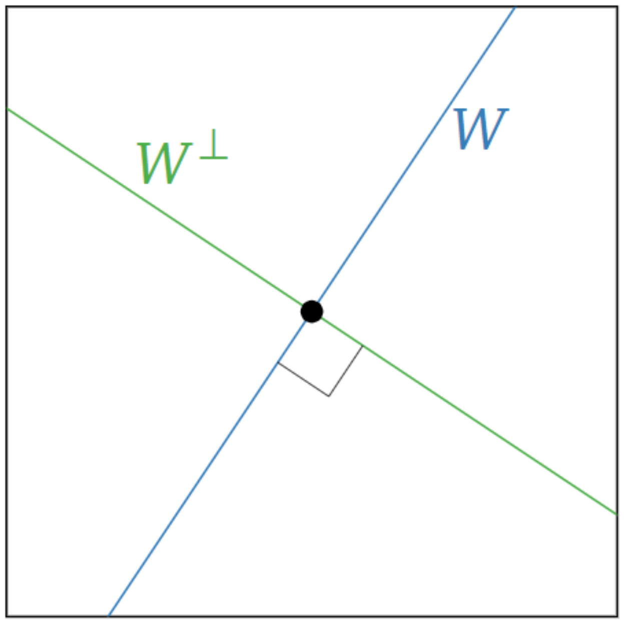
Définition : Soit W un sous-espace vectoriel de $\mathbb{R}^n$. Le **complément orthogonal** de W est le sous-espace vectoriel* contenant tous les vecteurs orthogonaux à ceux de W :

$$W^\perp = \{\boldsymbol{v} \in \mathbb{R}^n |\, \boldsymbol{v} \cdot \boldsymbol{w} = 0 \;\; \forall \boldsymbol{w} \in W\}$$

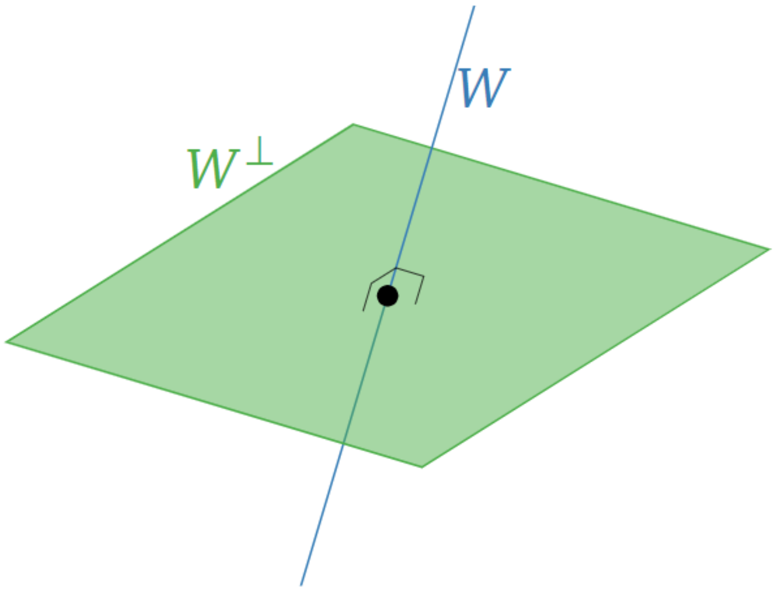
***Remarque :** Il n'est pas évident à priori que $W^\perp$ est un sous-espace vectoriel. On pourrait le montrer.

Exemples de compléments orthogonaux :

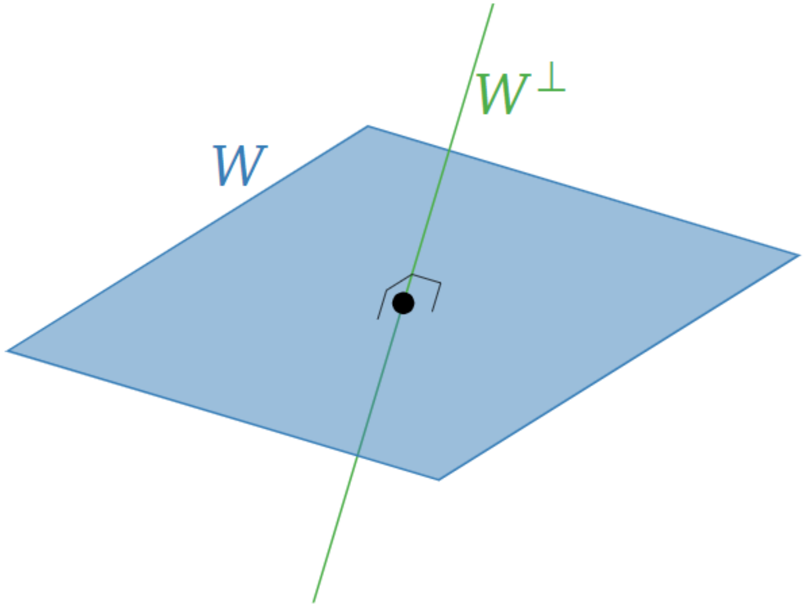
W : droite dans $\mathbb{R}^2$
$W^\perp$: droite dans $\mathbb{R}^2$



W : droite dans $\mathbb{R}^3$
$W^\perp$: plan dans $\mathbb{R}^3$



W : plan dans $\mathbb{R}^3$
$W^\perp$: droite dans $\mathbb{R}^3$



Remarques : $(W^\perp)^\perp = W$ et $\dim W + \dim W^\perp = \dim \mathbb{R}^n$.

Comment calculer le complément orthogonal ?

Théorème :

$$\text{Si } W = col(A) \text{ alors } W^{\perp} = Nul(A^T).$$

Remarque : *Calculer $Nul(A^T)$ revient à résoudre un SEL.*

De plus, si W est engendré par les vecteurs $v_1, v_2, \dots, v_n$, on peut former une matrice $A = \begin{pmatrix} | & & | \\ \boldsymbol{v_1} & \dots & \boldsymbol{v_n} \\ | & & | \end{pmatrix}$ de telle sorte que $W = col(A)$.

Preuve :

Preuve : Soit $W = col(A)$ et $A = \begin{pmatrix} | & & | \\ \boldsymbol{v_1} & \dots & \boldsymbol{v_n} \\ | & & | \end{pmatrix}$.

Alors,

$$x \in Nul(A^T) \iff A^T x = 0$$

$$\iff \begin{pmatrix} - & \boldsymbol{v_1^T} & - \\ & \vdots & \\ - & \boldsymbol{v_n^T} & - \end{pmatrix} \begin{pmatrix} | \\ \boldsymbol{x} \\ | \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\iff \begin{pmatrix} \boldsymbol{v_1} \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{v_n} \cdot \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\iff \boldsymbol{x} \perp \boldsymbol{v_i} \ \forall \boldsymbol{i}$$

$$\iff \boldsymbol{x} \in W^{\perp}$$

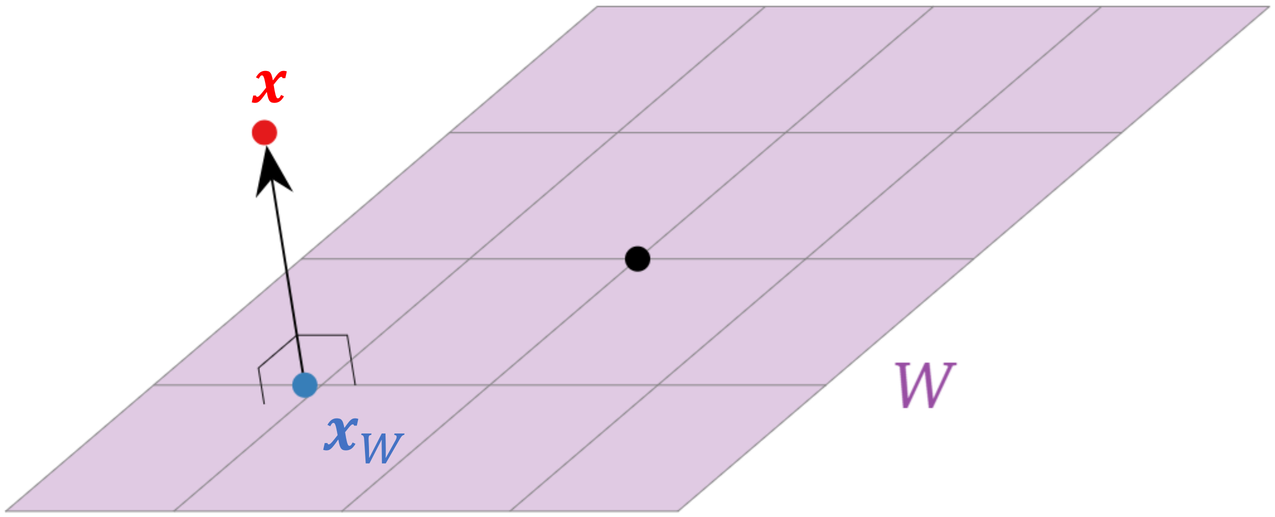
Exemple 2 :

a) $\textit{Trouvez } W^{\perp} \textit{ si } W = span\left\{ \begin{pmatrix} 1 \\ 7 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \right\}.$

b) $\textit{Trouvez tous les vecteurs orthogonaux à } \boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$

8.3 Projection orthogonale

Comment calculer le $\boldsymbol{x}_W \in W$ le plus proche de $\boldsymbol{x} \in \mathbb{R}^n$?



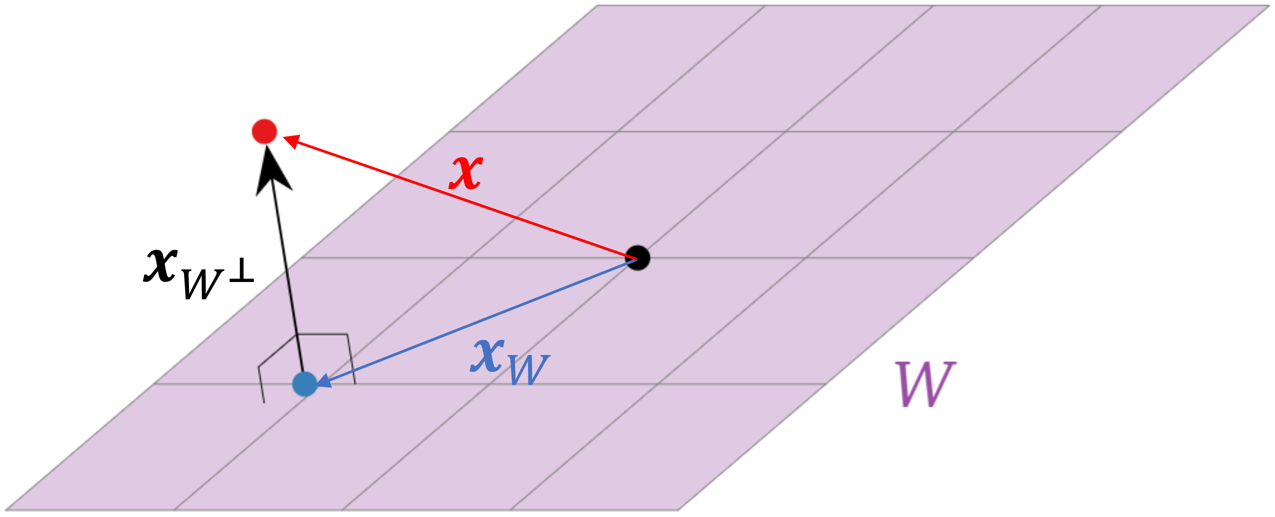
Définition : Soit W un sous-espace vectoriel de $\mathbb{R}^n$ et $\boldsymbol{x} \in \mathbb{R}^n$. La projection orthogonale de $\boldsymbol{x}$, notée $\boldsymbol{x}_W$, est le vecteur de W le plus près de $\boldsymbol{x}$.

Décomposition orthogonale

Théorème : *Soit W un sous-espace de $\mathbb{R}^n$* et $\boldsymbol{x} \in \mathbb{R}^n$. Il existe une unique décomposition orthogonale de $\boldsymbol{x}$:

$$\boldsymbol{x} = \boldsymbol{x}_W + \boldsymbol{x}_{W^\perp}$$

Où $\boldsymbol{x}_{W^\perp} \in W^\perp$ et $dist(\boldsymbol{x}, \boldsymbol{x}_W) = \|\boldsymbol{x}_{W^\perp}\|$.



Exemple 3 : Soit W qui correspond au plan xy dans $\mathbb{R}^3$.

a) À quoi correspond $W^\perp$?

b) Donnez la décomposition orthogonale de $\boldsymbol{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ par rapport à W.

c) Donnez la décomposition orthogonale de $\boldsymbol{b} = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}$ par rapport à W.

d) Donnez la décomposition orthogonale de $\boldsymbol{c} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ par rapport à W.

Calcul de x_W

Théorème : *Si* $W = col(A)$*, alors*

$$A^T A\boldsymbol{c} = A^T \boldsymbol{x}$$

Peut se résoudre pour $\boldsymbol{c}$ et donne la projection orthogonale via $\boldsymbol{x}_W = A\boldsymbol{c}$.

Algorithme : *Si* $W = span\{\boldsymbol{v_1}, \dots, \boldsymbol{v_k}\}$ *et* $\boldsymbol{x} \in \mathbb{R}^n$*. Alors* $x_W = ?$ *et* $x_{W^\perp} = ?$

1. *Construire la matrice* $A = \begin{pmatrix} | & & | \\ \boldsymbol{v_1} & \dots & \boldsymbol{v_k} \\ | & & | \end{pmatrix}$ *pour que* $W = col(A)$*.*
2. *Calculer* $A^T A$ *et* $A^T \boldsymbol{x}$*.*
3. *Résoudre* $A^T A\boldsymbol{c} = A^T \boldsymbol{x}$ *pour* $\boldsymbol{c}$*. Choisir un* $\boldsymbol{c}$ *s'il y a plusieurs solutions.*
4. *Alors* $\boldsymbol{x}_W = A\boldsymbol{c}$ *et* $\boldsymbol{x}_{W^\perp} = \boldsymbol{x} - \boldsymbol{x}_W$*.*

Preuve : Soit $W = col(A)$ alors $W^{\perp} = Nul(A^T)$.

Alors $x = x_W + x_{W^{\perp}}$ donne que

$$x_{W^{\perp}} = x - x_W \in Nul(A^T)$$

$$A^T(x - x_W) = 0$$

$$A^T x = A^T x_W$$

Mais $x_W \in col(A)$ alors il existe $c \in \mathbb{R}^n$ tel que $x_W = Ac$. Donc,

$$A^T x = A^T Ac.$$

Exemple 4 : Soit W la droite engendrée par le vecteur $\boldsymbol{u} = \binom{3}{2}$.

a) Calculez la décomposition orthogonale de $\boldsymbol{a} = \binom{-6}{4}$ par rapport à W.

b) Calculez la projection $proj_{\boldsymbol{u}}(\boldsymbol{a})$.

c) Que remarque-t-on en comparant les deux réponses ?

Exemple 5 : Soit W le plan engendré par $\boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ et $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

a) Calculez la décomposition orthogonale de $\boldsymbol{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ par rapport à W.

b) Calculez la distance entre $\boldsymbol{x}$ et W.

8.4 Bases orthogonales

Les projections s'expriment beaucoup plus facilement lorsqu'on travaille avec des vecteurs orthogonaux.

Définition : Un ensemble de vecteurs non-nuls $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ est dit

- **Orthogonal** si $\boldsymbol{v_i} \cdot \boldsymbol{v_j} = 0 \quad \forall i, j,$

- **Orthonormal** si en plus chaque vecteur est unitaire : $\|\boldsymbol{v_i}\| = 1 \quad \forall i.$

Remarques :
- *Un ensemble orthogonal est automatiquement linéairement indépendant.*
- *On peut facilement normaliser chaque vecteur en divisant par sa norme :* $\frac{\boldsymbol{v_i}}{\|\boldsymbol{v_i}\|}$

Exemple 6 : Considérons la base canonique de $\mathbb{R}^3$:

$$\boldsymbol{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

a) Vérifiez que c'est une base orthogonale.

b) Vérifiez que c'est une base orthonormale.

Procédé de Gram-Schmidt

Comment obtenir une base orthogonale ?

Soit $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ *une base de* W. *Alors on peut construire*

$$\mathbf{u}_1 = \mathbf{w}_1$$

$$\mathbf{u}_2 = \mathbf{w}_2 - \frac{\mathbf{w}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1$$

$$\mathbf{u}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{w}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$$

$$\vdots$$

$$\mathbf{u}_k = \mathbf{w}_k - \sum_{i=1}^{k-1} proj_{\mathbf{u}_i}(\mathbf{w}_k)$$

de telle sorte que $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ *est une base orthogonale.*

(Animation)

Exemple 7 : Voici une base d'un sous-espace W.

$$\boldsymbol{v_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{v_2} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

a) Construisez une base orthogonale de W.

b) Construisez une base orthonormale de W.

c) Ajoutez $\boldsymbol{v_3} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$ et utilisez le procédé Gram-Schmidt pour construire une base orthonormale de $\mathbb{R}^3$.

Projection orthogonale (bis)

Théorème **:** *Soit* $\{u_1, \dots, u_k\}$ *une base orthogonale de* W. *Alors la projection orthogonale* $\boldsymbol{x}_W$ *peut sa calculer par :*

$$\boldsymbol{x}_W = \frac{\boldsymbol{x} \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{x} \cdot \boldsymbol{u}_2}{\boldsymbol{u}_2 \cdot \boldsymbol{u}_2}\boldsymbol{u}_2 + \dots + \frac{\boldsymbol{x} \cdot \boldsymbol{u}_k}{\boldsymbol{u}_k \cdot \boldsymbol{u}_k}\boldsymbol{u}_k$$

Remarques :

- *Chaque terme est une projection de x sur un vecteur de la base.*
- *Si la base est orthonormale, alors les dénominateurs sont tous 1 :*

$$\boldsymbol{x}_W = (\boldsymbol{x} \cdot \boldsymbol{u}_1)\boldsymbol{u}_1 + (\boldsymbol{x} \cdot \boldsymbol{u}_2)\boldsymbol{u}_2 + \dots + (\boldsymbol{x} \cdot \boldsymbol{u}_k)\boldsymbol{u}_k$$

Exemple 8 : Soit W le plan engendré par

$$\boldsymbol{w_1} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \boldsymbol{w_2} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

a) Vérifiez que ces vecteurs forment une base orthogonale.

b) Calculez la projection orthogonale de $\boldsymbol{x} = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}$ sur W.

Exemple 9 : Voici une base orthogonale de $\mathbb{R}^3$:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

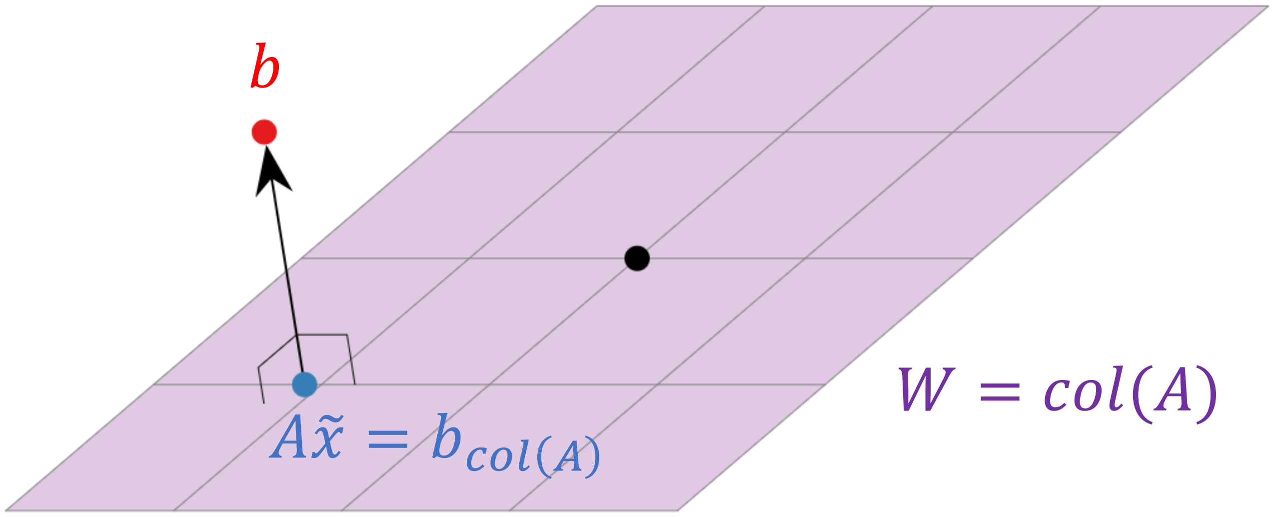
Calculez $[\boldsymbol{x}]_{\mathcal{B}}$ si $\boldsymbol{x} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$.

Suggestion : Pensez à $\mathcal{B}$ comme une base d'un sous-espace vectoriel et calculez la projection orthogonale de $\boldsymbol{x}$ sur ce sous-espace.

8.5 Méthode des moindres carrés

Si un système d'équations linéaires $Ax = b$ n'a pas de solution, peut-on trouver une approximation $\tilde{x}$ pour que $A\tilde{x}$ soit le plus proche possible de b ? $(A\tilde{x} \approx b)$

Ceci demande de trouver le point d'un sous-espace le plus près possible de b.



Exemple 10 : Soient trois points (1,1), (2,1) et (3,3). On cherche la droite qui modélise le mieux le nuage de points.

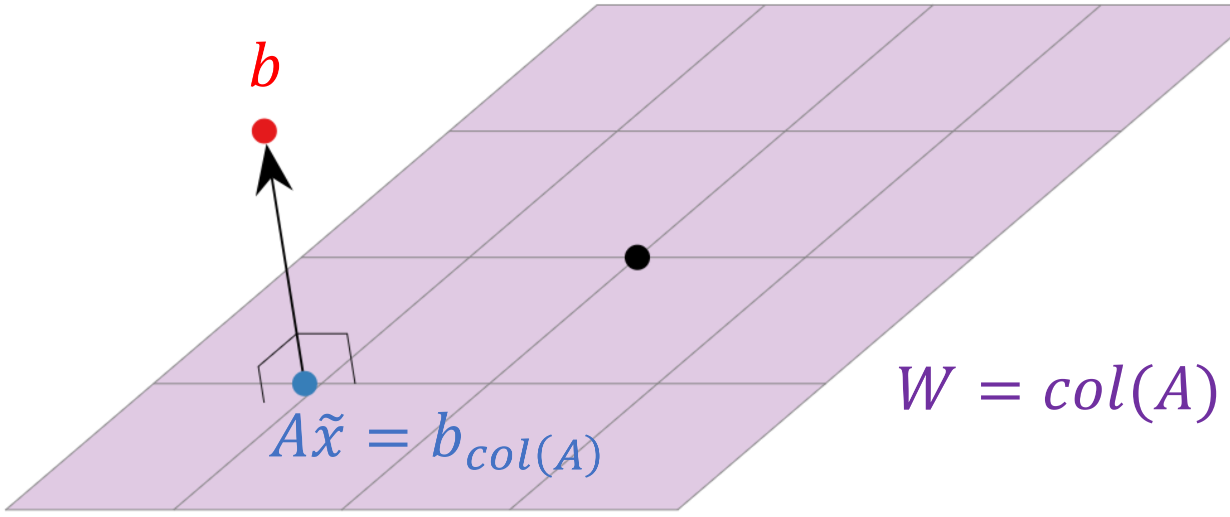
a) Représentez graphiquement les trois points. Existe-t-il une droite qui passe par les trois points ?
b) Posons $y = mx + c$ l'équation de la droite cherchée. Créez un SEL de trois équations pour m et c à l'aide des trois points.
c) Écrivez le SEL sous la forme $A\boldsymbol{x} = \boldsymbol{b}$. Remarquez qu'on cherche $\boldsymbol{x}$.
d) Échelonnez le système. Existe-t-il une solution ?
e) On en déduit que $\boldsymbol{b}$ n'appartient pas à $col(A)$. Trouvez une base de $col(A)$ et calculez la projection orthogonale $\boldsymbol{b}_{col(A)}$.
f) Résolvez le système $A\boldsymbol{x} = \boldsymbol{b}_{col(A)}$. On note la solution $\tilde{x}$.
g) Représentez graphiquement la droite $y = mx + c$ qui correspond à $\tilde{x}$ dans le même graphique qu'en a).

Remarques : On cherche la droite $y = mx + c$ modélisant les points (x_i, y_i).

Si $A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} c \\ m \end{pmatrix}$ et $\boldsymbol{b} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ donne un système incompatible,

alors on cherche une approximation $A\boldsymbol{x} = \begin{pmatrix} c + mx_1 \\ c + mx_2 \\ c + mx_3 \end{pmatrix} \approx \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \boldsymbol{b}$.

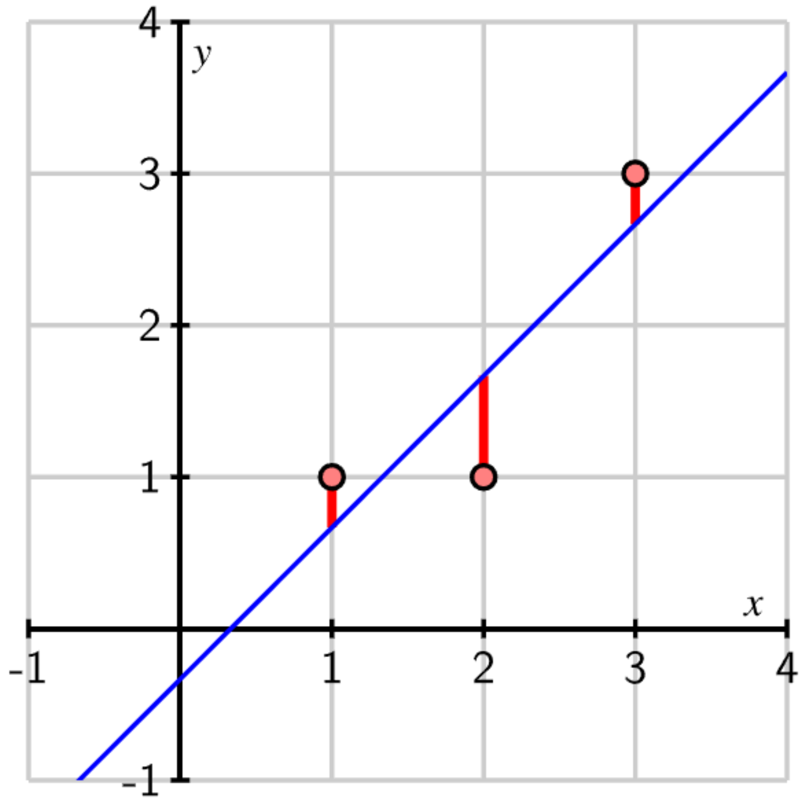
La projection orthogonale $\boldsymbol{b}_{col(A)}$ nous a permis de trouver le $\tilde{\boldsymbol{x}}$ qui minimise $\mathrm{dist}(A\tilde{\boldsymbol{x}}, \boldsymbol{b})$.



Qu'est-ce que ça représente « minimiser dist$(A\widetilde{\boldsymbol{x}}, \boldsymbol{b})$ » ?

$$\text{dist}(\boldsymbol{b}, A\widetilde{\boldsymbol{x}}) = \|\boldsymbol{b} - A\widetilde{\boldsymbol{x}}\| = \sqrt{(\boldsymbol{y_1} - (\boldsymbol{m}\boldsymbol{x_1} + \boldsymbol{c}))^2 + (\boldsymbol{y_2} - (\boldsymbol{m}\boldsymbol{x_2} + \boldsymbol{c}))^2 + (\boldsymbol{y_3} - (\boldsymbol{m}\boldsymbol{x_3} + \boldsymbol{c}))^2}$$

On **minimise** la somme du **carré des distances** entre les points (x_i, y_i) données et les points de la droite choisie : $(x_i, mx_i + c)$. D'où le nom : méthode des moindres carrés.



Méthode des moindres carrés

Comment trouver la meilleure approximation $Ax \approx b$ ***?***

La solution moindres carrés $\tilde{\boldsymbol{x}}$ *qui fournit la meilleure approximation s'obtient à l'aide d'une projection orthogonale* $\boldsymbol{b}_{col(A)}$ *et en résolvant* $A\tilde{\boldsymbol{x}} = \boldsymbol{b}_{col(A)}$.

On a vu deux façons de calculer des projections orthogonales.

Méthode 1 *: Résoudre* $A^T A\tilde{\boldsymbol{x}} = A^T\boldsymbol{b}$.

Méthode 2 *: Si les colonnes de A forment une base orthogonale* $\{\boldsymbol{u_1}, \dots, \boldsymbol{u_k}\}$, *alors*

$$\tilde{\boldsymbol{x}} = \left(\frac{\boldsymbol{b}\cdot\boldsymbol{u_1}}{\boldsymbol{u_1}\cdot\boldsymbol{u_1}}, \frac{\boldsymbol{b}\cdot\boldsymbol{u_2}}{\boldsymbol{u_2}\cdot\boldsymbol{u_2}}, \dots, \frac{\boldsymbol{b}\cdot\boldsymbol{u_k}}{\boldsymbol{u_k}\cdot\boldsymbol{u_k}}\right)^T.$$

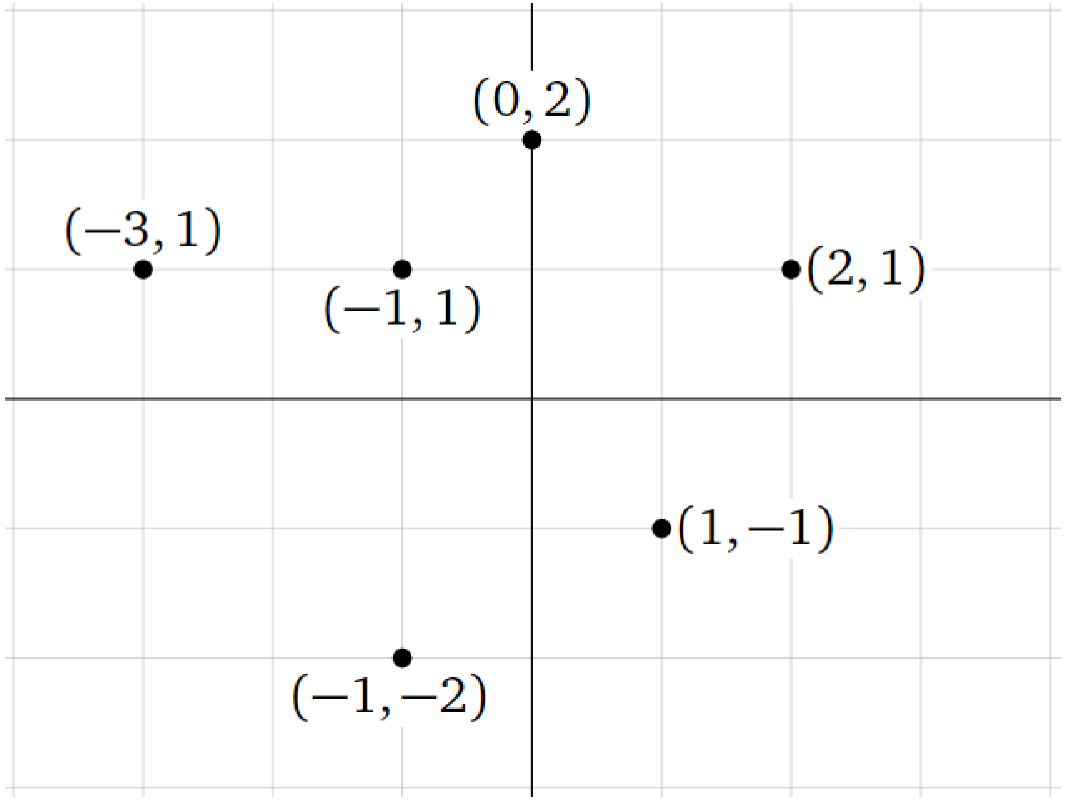
Remarque: *La première méthode fonctionne toujours, tandis que la seconde à une condition.*

Exemple 11 : Trouver la solution moindre carré au système $A\boldsymbol{x} = \boldsymbol{b}$ où

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 4 \end{pmatrix}.$$

Gauss a développé la méthode des moindres carrés pour prédire la trajectoire elliptique de l'astéroïde Cérès alors qu'elle passait derrière le soleil en 1801. Regardons comment la méthode fonctionne dans un exemple similaire.

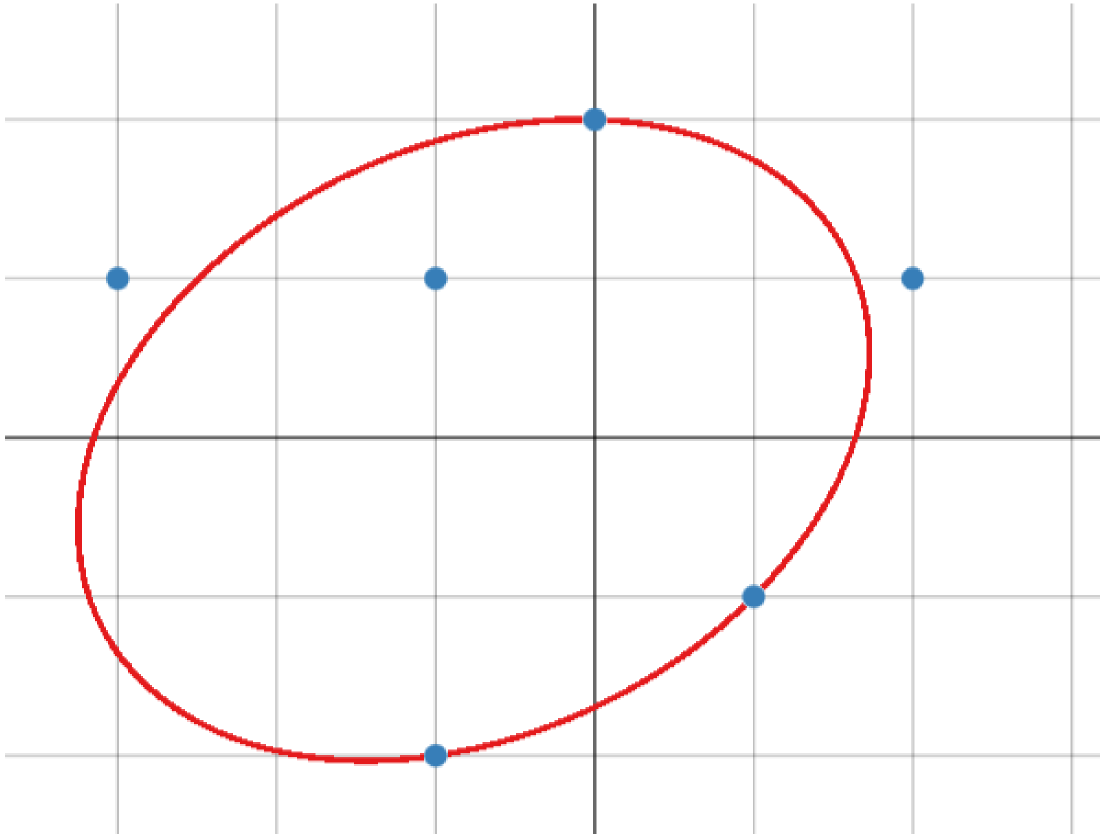
Exemple 12 : Trouver la trajectoire elliptique d'un objet dont on a mesuré les positions suivantes : $(0,2)$; $(2,1)$; $(1,-1)$; $(-1,-2)$; $(-3,-1)$ et $(-1,-1)$.

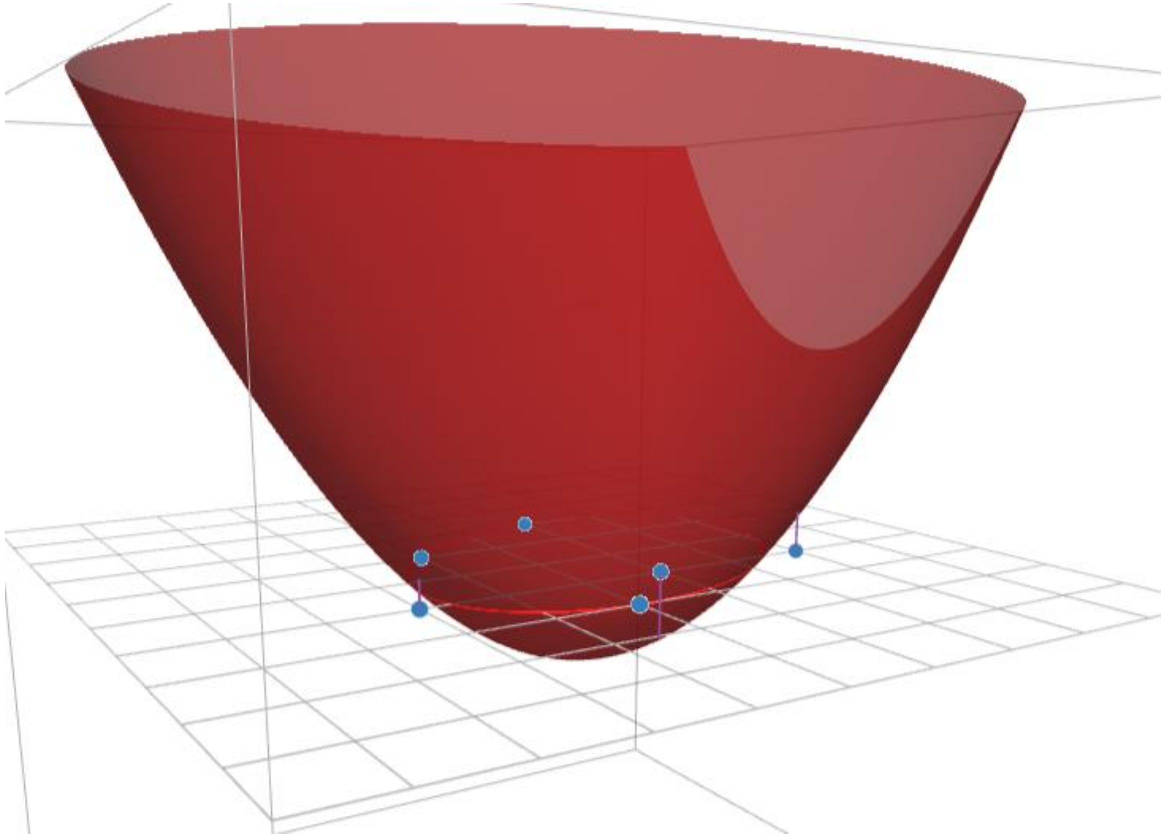


Note : L'équation d'une ellipse peut s'écrire sous la forme :

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

Remarque : L’ellipse trouvée minimise la distance verticale entre le paraboloïde et les points données.





Graphiques disponibles à la fin de cette page