
6. Orthogonalité et méthode des moindres carrés

6.1. Produit scalaire, longueur et orthogonalité.

Rappel. Si A est une matrice de taille $m \times n$
et B est une matrice de taille $n \times p$
alors AB est une matrice de taille $m \times p$

Soient maintenant $\vec{u}, \vec{v} \in \mathbb{R}^n$ deux vecteurs quelconques.

Comme $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ est une matrice de taille $n \times 1$

alors que $\vec{u}^T = [u_1 \ u_2 \ \dots \ u_n]$ est une matrice de taille $1 \times n$,

le produit matriciel $\vec{v} \vec{u}^T$ est une matrice de taille $n \times n$

et le produit matriciel $\vec{u}^T \vec{v}$ est une matrice de taille 1×1 ,
autrement dit, un nombre réel.

Définition. Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de $\mathbb{R}^n$. Le nombre réel

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$$

est appelé **produit scalaire (euclidien)** de $\vec{u}$ et $\vec{v}$.

On le note aussi $\langle \vec{u}, \vec{v} \rangle$.

Si $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ et $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, alors on a :

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = [u_1 \dots u_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{j=1}^n u_j v_j$$

Remarque. Il est possible de définir d'autres produits scalaires sur $\mathbb{R}^n$ mais le produit scalaire euclidien est le seul qui sera traité dans ce cours.

Théorème. Soient $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. On a :

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (commutativité)
2. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$ (distributivité)
3. $(\lambda \vec{u}) \cdot \vec{v} = \vec{u} \cdot (\lambda \vec{v}) = \lambda (\vec{u} \cdot \vec{v})$
4. $\vec{u} \cdot \vec{u} \geq 0$ pour tout $\vec{u} \in \mathbb{R}^n$

De plus : $\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$

Preuve.

1. On a $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$

et $\vec{v} \cdot \vec{u} = \vec{v}^T \vec{u} = v_1 u_1 + v_2 u_2 + \dots + v_n u_n$

2. $\vec{u} \cdot (\vec{v} + \vec{w}) \stackrel{\text{d\u00e9f}}{=} \vec{u}^T (\vec{v} + \vec{w}) = \vec{u}^T \vec{v} + \vec{u}^T \vec{w} = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$

3. $(\lambda \vec{u}) \cdot \vec{v} = (\lambda \vec{u})^T \vec{v} = \lambda \vec{u}^T \vec{v} = \lambda (\vec{u} \cdot \vec{v})$

$\vec{u} \cdot (\lambda \vec{v}) = \vec{u}^T (\lambda \vec{v}) = \lambda \vec{u}^T \vec{v} = \lambda (\vec{u} \cdot \vec{v})$

4. $\vec{u} \cdot \vec{u} = \underbrace{u_1^2}_{\geq 0} + \underbrace{u_2^2}_{\geq 0} + \dots + \underbrace{u_n^2}_{\geq 0} \geq 0$

Pour avoir une somme nulle, il faut que chaque terme soit nul:

$u_j^2 = 0$ pour tout j . Autrement dit, $u_j = 0$ pour tout j . ■

Cons\u00e9quence.

$$\vec{u} \cdot (\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_p \vec{v}_p) = \alpha_1 \vec{u} \cdot \vec{v}_1 + \alpha_2 \vec{u} \cdot \vec{v}_2 + \dots + \alpha_p \vec{u} \cdot \vec{v}_p$$

D\u00e9finition. Soit $\vec{u} \in \mathbb{R}^n$. La **norme (euclidienne)** ou **longueur** du vecteur $\vec{u}$ est le nombre

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} \geq 0$$

Remarque. Pour des calculs th\u00e9oriques, on utilise plut\u00f4t la formule $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$.

Propri\u00e9t\u00e9. Si $\vec{u} \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, alors

$$\|\lambda \vec{u}\| = |\lambda| \|\vec{u}\|$$

Définition. On dit qu'un vecteur $\vec{u} \in \mathbb{R}^n$ est **unitaire** (ou **unité**) si $\|\vec{u}\| = 1$.

Remarque. Si $\vec{u}$ est un vecteur non nul, alors le vecteur

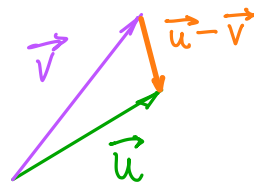
$$\frac{1}{\|\vec{u}\|} \vec{u} = \frac{\vec{u}}{\|\vec{u}\|}$$

est un vecteur unitaire.

Définition. Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$. On définit la **distance** entre $\vec{u}$ et $\vec{v}$ par

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

Geométriquement:



Définition. Deux vecteurs $\vec{u}$ et $\vec{v}$ sont **orthogonaux** (entre eux) si

$$\vec{u} \cdot \vec{v} = 0$$

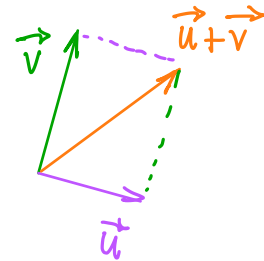
Remarque. Comme $\vec{0} \cdot \vec{v} = 0$ pour tout $\vec{v} \in \mathbb{R}^n$, le vecteur $\vec{0}$ est orthogonal à tout vecteur de $\mathbb{R}^n$.

Théorème de Pythagore. Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$. On a l'équivalence:

$$\vec{u} \text{ et } \vec{v} \text{ sont orthogonaux} \Leftrightarrow \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$

Preuve.

$$\begin{aligned} \text{On a } \|\vec{u} + \vec{v}\|^2 &= (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) \\ &= \vec{u} \cdot (\vec{u} + \vec{v}) + \vec{v} \cdot (\vec{u} + \vec{v}) \\ &= \vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v} \\ &= \|\vec{u}\|^2 + 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2 \end{aligned}$$



$$\text{d'où: } \vec{u} \cdot \vec{v} = 0 \Leftrightarrow \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 \quad \blacksquare$$

Définition. Soit S un sous-ensemble de $\mathbb{R}^n$.

Si $\vec{x} \in \mathbb{R}^n$ est orthogonal à tous les vecteurs $\vec{v} \in S$, alors on dit que $\vec{x}$ est orthogonal à S .

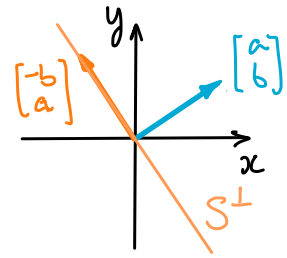
L'ensemble de tous les vecteurs orthogonaux à S est appelé complément orthogonal de S , noté $S^\perp$ (se lit "S orthogonal") :

$$S^\perp = \{ \vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{v} = 0 \text{ pour tout } \vec{v} \in S \}$$

Exemples.

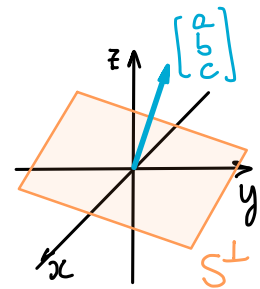
$n=2$: Si $S = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \right\}$, alors

$$S^\perp = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : ax + by = 0 \right\}$$



$n=3$: Si $S = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\}$, alors

$$S^\perp = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : ax + by + cz = 0 \right\}$$



Proposition. Si S est un sous-ensemble de $\mathbb{R}^n$, alors $S^\perp$ est un sous-espace vectoriel de $\mathbb{R}^n$.

Preuve.

On a $\vec{0} \in S^\perp$ car $\vec{0} \cdot \vec{v} = 0$ pour tout $\vec{v} \in S$

Soient $\vec{x}, \vec{y} \in S^\perp$ et $\lambda \in \mathbb{R}$

A voir : $\vec{x} + \vec{y} \in S^\perp$ et $\lambda \vec{x} \in S^\perp$

Par hypothèse : $\vec{x} \cdot \vec{v} = 0$ pour tout $\vec{v} \in S$

$$\vec{y} \cdot \vec{v} = 0 \text{ pour tout } \vec{v} \in S$$

$$\text{d'où : } (\vec{x} + \vec{y}) \cdot \vec{v} = \vec{x} \cdot \vec{v} + \vec{y} \cdot \vec{v} = 0 + 0 = 0 \Rightarrow \vec{x} + \vec{y} \in S^\perp$$

$$(\lambda \vec{x}) \cdot \vec{v} = \lambda (\vec{x} \cdot \vec{v}) = \lambda 0 = 0 \Rightarrow \lambda \vec{x} \in S^\perp \quad \blacksquare$$

Remarque.

La proposition précédente nous dit que même si S n'est pas un sous-espace vectoriel, son complément orthogonal $S^\perp$ en est un !

Proposition. Soit W un sous-espace vectoriel de $\mathbb{R}^n$. On a

$$W \cap W^\perp = \{\vec{0}\}$$

Preuve.

Comme W et $W^\perp$ sont des sous-espaces vectoriels de $\mathbb{R}^n$, ils contiennent le vecteur zéro $\vec{0}$. Par conséquent, $\vec{0} \in W \cap W^\perp$.

Soit $\vec{x} \in W \cap W^\perp$. Comme $\vec{x} \in W$ et $\vec{x} \in W^\perp$, on a $\vec{x} \cdot \vec{x} = 0$, d'où $\vec{x} = \vec{0}$ est le seul élément de $W \cap W^\perp = \{\vec{0}\}$ ■

Proposition. Soient $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ k vecteurs quelconques de $\mathbb{R}^n$.

Si $S = \{\vec{v}_1, \dots, \vec{v}_k\}$ et $W = \text{Vect } S$, alors on a :

$$W^\perp = S^\perp$$

Autrement dit,

$\vec{x} \in W^\perp \Leftrightarrow \vec{x}$ est orthogonal à un système de générateurs de W

Preuve.

On va montrer $W^\perp \subset S^\perp$ et $S^\perp \subset W^\perp$, d'où l'égalité.

$W^\perp \subset S^\perp$: Si $\vec{x} \in W^\perp$, alors $\vec{x} \cdot \vec{w} = 0$ pour tout $\vec{w} \in W$

d'où $\vec{x} \cdot \vec{v}_j = 0$ pour tout $j \in \{1, \dots, k\}$

autrement dit, $\vec{x} \in S^\perp$.

$S^\perp \subset W^\perp$: Si $\vec{x} \in S^\perp$, alors $\vec{x} \cdot \vec{v}_j = 0$ pour tout $j \in \{1, \dots, k\}$
Comme $W = \text{Vect}\{\vec{v}_1, \dots, \vec{v}_k\}$, tout vecteur $\vec{w} \in W$ s'écrit:

$$\vec{w} = \alpha_1 \vec{v}_1 + \dots + \alpha_k \vec{v}_k$$

On a donc

$$\begin{aligned}\vec{x} \cdot \vec{w} &= \vec{x} \cdot (\alpha_1 \vec{v}_1 + \dots + \alpha_k \vec{v}_k) \\ &= \alpha_1 (\vec{x} \cdot \vec{v}_1) + \dots + \alpha_k (\vec{x} \cdot \vec{v}_k) \\ &= \alpha_1 0 + \dots + \alpha_k 0 = 0\end{aligned}$$

d'où $\vec{x} \in W^\perp$

■

Proposition. Soit W un sous-espace vectoriel de $\mathbb{R}^n$. On a

$$W \subset (W^\perp)^\perp$$

Preuve.

Soit $\vec{w} \in W$ un vecteur quelconque.

Comme $\vec{w}$ est orthogonal à tout vecteur de $W^\perp$, nous avons $\vec{w} \in (W^\perp)^\perp$.

■

Remarque. Nous allons montrer plus tard qu'on a en fait égalité:

$$W = (W^\perp)^\perp$$

Théorème. Soit A une matrice de taille $m \times n$. Alors on a :

$$(\text{Lgn}(A))^{\perp} = \text{Nul}(A) \subset \mathbb{R}^n$$

Preuve.

$$\vec{x} \in \text{Nul}(A) \Leftrightarrow A\vec{x} = \vec{0}$$

$$\Leftrightarrow \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^m$$

$$\Leftrightarrow \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Leftrightarrow \vec{x}$ est orthogonal à chaque ligne de A .

$$\Leftrightarrow \vec{x} \in (\text{Lgn}(A))^{\perp}$$

d'où l'égalité $(\text{Lgn}(A))^{\perp} = \text{Nul}(A)$ ■

Corollaire. Soit A une matrice de taille $m \times n$. Alors on a :

$$(\text{Col}(A))^{\perp} = \text{Nul}(A^T) \subset \mathbb{R}^m$$

Preuve. Le théorème précédent appliqué à la matrice A^T nous donne :

$$(\text{Lgn}(A^T))^{\perp} = \text{Nul}(A^T)$$

Comme $\text{Lgn}(A^T) = \text{Col}(A)$ nous avons le résultat. ■

Méthode pour déterminer $W^\perp$:

Soit $W \subset \mathbb{R}^n$ un sous-espace vectoriel de dimension $0 < k \leq n$.

1. Trouver une base $\{\vec{w}_1, \dots, \vec{w}_k\}$ de W
2. Construire la matrice A de taille $k \times n$ dont les lignes sont les vecteurs $\vec{w}_1, \dots, \vec{w}_k$.

Par construction, $\text{Lgn}(A) = W$

Comme $\{\vec{w}_1, \dots, \vec{w}_k\}$ est libre, on a $\text{rang}(A) = k$.

3. Résoudre le système homogène $A\vec{x} = \vec{0}$ pour trouver

$$W^\perp = \text{Nul}(A)$$

Le théorème du rang nous donne $\dim(W^\perp) = n - k$.

6.2. Ensembles orthogonaux.

Définition. Un ensemble $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ de vecteurs de $\mathbb{R}^n$ est appelé **ensemble orthogonal** si les vecteurs $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ sont orthogonaux deux à deux:

$$\vec{u}_j \cdot \vec{u}_l = 0 \quad \text{pour tout } j, l \in \{1, \dots, k\} \text{ tels que } j \neq l$$

Exemples.

1. L'ensemble $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$ est orthogonal car $2 \cdot 3 + 3 \cdot (-2) = 0$.
2. L'ensemble $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ est orthogonal car
 $1 \cdot 1 + 2 \cdot (-2) + 3 \cdot 1 = 0$, $1 \cdot 0 + 2 \cdot 0 + 3 \cdot 0 = 0$ et $1 \cdot 0 - 2 \cdot 0 + 1 \cdot 0 = 0$
3. Si $E = \{\vec{e}_1, \dots, \vec{e}_n\}$ est la base canonique de $\mathbb{R}^n$ alors tout sous-ensemble de E est orthogonal.

Théorème. Si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble orthogonal de vecteurs **non nuls**, alors $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble linéairement indépendant.

Conséquence. Si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble orthogonal de vecteurs **non nuls**, alors $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base de $\text{Vect}\{\vec{u}_1, \dots, \vec{u}_k\}$.

Preuve.

Soit $\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_k \vec{u}_k = \vec{0}$.

A voir: $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_k = 0$.

On a $\vec{u}_1 \cdot (\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_k \vec{u}_k) = \vec{u}_1 \cdot \vec{0} = 0$

d'où $\alpha_1 (\vec{u}_1 \cdot \vec{u}_1) + \alpha_2 (\vec{u}_1 \cdot \vec{u}_2) + \dots + \alpha_k (\vec{u}_1 \cdot \vec{u}_k) = 0$

Comme $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble orthogonal, on a $\vec{u}_i \cdot \vec{u}_j = 0$

pour tout $j \in \{2, \dots, k\}$, ce qui implique:

$$\alpha_1 (\vec{u}_1 \cdot \vec{u}_1) = 0$$

Comme $\vec{u}_1$ est non nul, on a $\vec{u}_1 \cdot \vec{u}_1 > 0$ et $\alpha_1 = 0$.

Un calcul analogue nous donne $\alpha_2 = 0, \dots, \alpha_k = 0$



Définition. Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $0 < k \leq n$.

On dit que l'ensemble $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base orthogonale de W

si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base de W qui est un ensemble orthogonal.

Exemple.

La base canonique de $\mathbb{R}^n$ est une base orthogonale de $\mathbb{R}^n$.

Le théorème suivant nous dit que lorsqu'on dispose d'une base orthogonale de W , le calcul des coordonnées d'un vecteur $\vec{w} \in W$ dans cette base est très simple :

Théorème. Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $0 < k \leq n$.

Soit $\{\vec{u}_1, \dots, \vec{u}_k\}$ une base orthogonale de W . Alors tout élément $\vec{w} \in W$ s'écrit de manière unique :

$$\vec{w} = \frac{\vec{w} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{w} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 + \dots + \frac{\vec{w} \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} \vec{u}_k$$

Preuve.

Soit $\vec{w} \in W$. Comme $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base de W on a

$$\vec{w} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_k \vec{u}_k \quad (\text{écriture unique})$$

$$\begin{aligned} \text{d'où } \vec{w} \cdot \vec{u}_1 &= (\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_k \vec{u}_k) \cdot \vec{u}_1 \\ &= \alpha_1 \underbrace{\vec{u}_1 \cdot \vec{u}_1} + \alpha_2 \underbrace{\vec{u}_2 \cdot \vec{u}_1}_{=0} + \dots + \alpha_k \underbrace{\vec{u}_k \cdot \vec{u}_1}_{=0} \end{aligned}$$

$$\text{Ainsi, } \vec{w} \cdot \vec{u}_1 = \alpha_1 \vec{u}_1 \cdot \vec{u}_1$$

$$\text{ce qui nous donne: } \alpha_1 = \frac{\vec{w} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}$$

On trouve $\alpha_2, \dots, \alpha_k$ de manière analogue en calculant

$$\vec{w} \cdot \vec{u}_j, \text{ avec } j \in \{2, \dots, k\}$$

■

Exemple.

Les vecteurs $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ et $\vec{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ sont orthogonaux et non nuls.

Ainsi, $B = \{\vec{u}_1, \vec{u}_2\}$ est une base orthogonale de $\mathbb{R}^2$.

Soit $\vec{w} = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$. Nous avons:

$$\begin{aligned} \vec{w} &= \frac{\vec{w} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{w} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 \\ &= \frac{5 \cdot 1 + 9 \cdot 1}{1^2 + 1^2} \vec{u}_1 + \frac{5 \cdot 1 + 9 \cdot (-1)}{1^2 + (-1)^2} \vec{u}_2 \\ &= \frac{14}{2} \vec{u}_1 + \frac{-4}{2} \vec{u}_2 = 7\vec{u}_1 - 2\vec{u}_2 \end{aligned}$$

$$\Rightarrow [\vec{w}]_B = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$$

Calcul alternatif:

$$\vec{w} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 \Leftrightarrow \left[\begin{array}{cc|c} 1 & 1 & 5 \\ 1 & -1 & 9 \end{array} \right] \sim \left[\begin{array}{cc|c} 1 & 1 & 5 \\ 0 & -2 & 4 \end{array} \right] \sim \dots \sim \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & -2 \end{array} \right]$$

Définition. Un ensemble $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ de vecteurs de $\mathbb{R}^n$ est appelé **ensemble orthonormal** si :

- $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ est un ensemble orthogonal
- les vecteurs $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k$ sont unitaires :
$$\|\vec{u}_j\| = 1, \text{ pour } j \in \{1, \dots, k\}$$

Exemples.

1. L'ensemble $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$ **n'est pas** orthonormal.
2. L'ensemble $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ **n'est pas** orthonormal.
3. Si $E = \{\vec{e}_1, \dots, \vec{e}_n\}$ est la base canonique de $\mathbb{R}^n$ alors tout sous-ensemble de E est orthonormal.

Remarque.

Par construction, si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble orthonormal, alors les vecteurs $\vec{u}_1, \dots, \vec{u}_k$ sont non nuls et de ce fait, $\{\vec{u}_1, \dots, \vec{u}_k\}$ est **toujours** linéairement indépendant.

Par conséquent, si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est un ensemble orthonormal, alors $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base de $\text{Vect}\{\vec{u}_1, \dots, \vec{u}_k\}$.

Définition. Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $0 < k \leq n$.

On dit que l'ensemble $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une **base orthonormale** (ou **base orthonormée**) de W si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base de W qui est un ensemble orthonormal.

Exemple

La base canonique de $\mathbb{R}^n$ est une base orthonormale de $\mathbb{R}^n$.

Remarque.

Si $\{\vec{w}_1, \dots, \vec{w}_k\}$ est une base orthogonale de W , alors l'ensemble $\{\vec{u}_1, \dots, \vec{u}_k\}$, où $\vec{u}_j = \frac{\vec{w}_j}{\|\vec{w}_j\|}$ pour $j=1, \dots, k$, est une base orthonormale de W .

Théorème. Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $0 < k \leq n$.

Soit $\{\vec{u}_1, \dots, \vec{u}_k\}$ une base orthonormale de W . Alors tout élément $\vec{w} \in W$ s'écrit de manière unique:

$$\vec{w} = (\vec{w} \cdot \vec{u}_1) \vec{u}_1 + (\vec{w} \cdot \vec{u}_2) \vec{u}_2 + \dots + (\vec{w} \cdot \vec{u}_k) \vec{u}_k$$

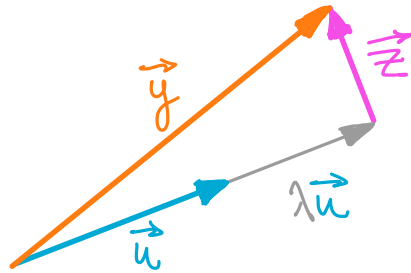
6.3. Projection orthogonale

Soit $\vec{u} \in \mathbb{R}^n$ un vecteur non nul fixé.

On aimerait écrire un vecteur arbitraire $\vec{y} \in \mathbb{R}^n$ comme la somme d'un multiple de $\vec{u}$ et d'un vecteur orthogonal à $\vec{u}$:

$$\vec{y} = \lambda \vec{u} + \vec{z} \text{ avec } \vec{z} \cdot \vec{u} = 0 \text{ et } \lambda \in \mathbb{R}$$

Géométriquement :



Comme $\vec{z} = \vec{y} - \lambda \vec{u}$

nous avons $\vec{u} \cdot \vec{z} = \vec{u} \cdot (\vec{y} - \lambda \vec{u})$

$$0 = \vec{u} \cdot \vec{y} - \lambda (\vec{u} \cdot \vec{u})$$

d'où $\lambda = \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}}$

Ainsi, le vecteur $\vec{y} \in \mathbb{R}^n$ peut s'écrire :

$$\vec{y} = \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \vec{u} + \vec{z}$$

Définition. Soit $\vec{u} \in \mathbb{R}^n$ un vecteur non nul.

La projection orthogonale de $\vec{y} \in \mathbb{R}^n$ sur $W = \text{Vect}\{\vec{u}\}$, notée $\text{proj}_W \vec{y}$ ou $\text{proj}_{\vec{u}} \vec{y}$, est le vecteur

$$\text{proj}_W \vec{y} = \text{proj}_{\vec{u}} \vec{y} = \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \vec{u}$$

Comme par construction le vecteur $\vec{z} = \vec{y} - \text{proj}_{\vec{u}} \vec{y}$ est orthogonal à W , nous avons $\vec{z} \in W^\perp$ et nous pouvons exprimer le vecteur $\vec{y}$ comme une somme d'un vecteur dans W et d'un vecteur dans $W^\perp$:

$$\vec{y} = \underbrace{\text{proj}_{\vec{u}} \vec{y}}_{\in W} + \underbrace{\vec{z}}_{\in W^\perp}$$

Remarque.

Soit $\vec{u} \in \mathbb{R}^n$ un vecteur non nul et $W = \text{Vect}\{\vec{u}\}$.

On peut montrer que l'application $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie par

$$T(\vec{v}) = \text{proj}_{\vec{u}} \vec{v} = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \vec{u}$$

est linéaire (voir exercice 15, série 12). Elle est appelée la projection orthogonale sur $W = \text{Vect}\{\vec{u}\}$.

Comme $T(\vec{v}) = \vec{0} \Leftrightarrow \vec{u} \cdot \vec{v} = 0 \Leftrightarrow \vec{v} \in W^\perp$

nous avons $\text{Ker}(T) = W^\perp$

Comme $T(\vec{v})$ est un multiple de $\vec{u}$ pour tout $\vec{v} \in \mathbb{R}^n$

nous avons $\text{Im}(T) = W$

L'associativité du produit matriciel nous permet d'écrire :

$$\underbrace{(\vec{u}\vec{u}^T)}_{\text{matrice de taille } n \times n} \underbrace{\vec{v}}_{\text{nombre réel}} = \vec{u}(\underbrace{\vec{u}^T \vec{v}}_{\text{nombre réel}}) = \vec{u}(\vec{u} \cdot \vec{v}) = (\vec{u} \cdot \vec{v}) \vec{u}, \text{ car } \vec{u} \cdot \vec{v} \in \mathbb{R}$$

Ainsi,

$$T(\vec{v}) = \text{proj}_{\vec{u}} \vec{v} = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{1}{\vec{u} \cdot \vec{u}} (\vec{u}\vec{u}^T) \vec{v}$$

Par conséquent, la matrice canoniquement associée à l'application linéaire $T = \text{proj}_{\vec{u}}$ est :

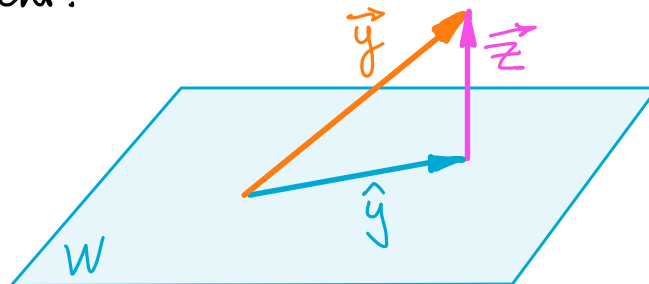
$$A_T = \frac{1}{\vec{u} \cdot \vec{u}} \vec{u}\vec{u}^T = \frac{\vec{u}\vec{u}^T}{\vec{u}^T \vec{u}}$$

Soit maintenant $W \subset \mathbb{R}^n$ un sous espace vectoriel de dimension $k \geq 2$,

On aimerait écrire un vecteur arbitraire $\vec{y} \in \mathbb{R}^n$ comme la somme d'un vecteur $\hat{\vec{y}} \in W$ et d'un vecteur $\vec{z} \in W^\perp$:

$$\vec{y} = \hat{\vec{y}} + \vec{z}$$

Géométriquement :



Théorème de la projection orthogonale.

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$.

Tout vecteur $\vec{y} \in \mathbb{R}^n$ s'écrit de manière unique sous la forme :

$$\vec{y} = \hat{y} + \vec{z} \quad \text{où } \hat{y} \in W \text{ et } \vec{z} \in W^\perp.$$

De plus, si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base orthogonale de W , alors nous avons une formule explicite pour calculer $\hat{y}$:

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 + \dots + \frac{\vec{y} \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} \vec{u}_k \quad (*)$$

Le vecteur $\hat{y}$ est appelé la projection orthogonale de $\vec{y} \in \mathbb{R}^n$ sur W et on le note $\text{proj}_W \vec{y}$.

Preuve.

Par construction, le vecteur $\hat{y}$ appartient à W car il s'écrit comme combinaison linéaire des vecteurs $\{\vec{u}_1, \dots, \vec{u}_k\}$.

A voir: $\vec{z} = \vec{y} - \hat{y} \in W^\perp$

Il faut donc montrer $\vec{z} \cdot \vec{u}_j = 0$, pour $j \in \{1, \dots, k\}$

On a

$$\begin{aligned} \vec{z} \cdot \vec{u}_1 &= (\vec{y} - \hat{y}) \cdot \vec{u}_1 = \vec{y} \cdot \vec{u}_1 - \hat{y} \cdot \vec{u}_1 \\ &= \vec{y} \cdot \vec{u}_1 - \left(\frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 + \dots + \frac{\vec{y} \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} \vec{u}_k \right) \cdot \vec{u}_1 \\ &= \vec{y} \cdot \vec{u}_1 - \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} (\vec{u}_1 \cdot \vec{u}_1) - \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} (\underbrace{\vec{u}_2 \cdot \vec{u}_1}_{=0}) - \dots - \frac{\vec{y} \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} (\underbrace{\vec{u}_k \cdot \vec{u}_1}_{=0}) \\ &= \vec{y} \cdot \vec{u}_1 - \vec{y} \cdot \vec{u}_1 = 0 \end{aligned}$$

Un calcul analogue nous donne

$$\vec{z} \cdot \vec{u}_j = 0, \text{ pour } j \in \{2, \dots, k\}$$

Il nous reste à montrer que l'écriture est unique.

Supposons qu'en a une autre décomposition:

$$\vec{y} = \hat{y}_1 + \vec{z}_1 \text{ avec } \hat{y}_1 \in W \text{ et } \vec{z}_1 \in W^\perp$$

$$\text{On a donc } \hat{y} + \vec{z} = \hat{y}_1 + \vec{z}_1$$

$$\text{d'où } \underbrace{\hat{y} - \hat{y}_1}_{\in W} = \underbrace{\vec{z}_1 - \vec{z}}_{\in W^\perp}$$

$$\text{Ainsi, le vecteur } \vec{u} = \hat{y} - \hat{y}_1 = \vec{z}_1 - \vec{z} \in W \cap W^\perp$$

$$\text{d'où l'unicité: } \vec{u} = \vec{0} \Leftrightarrow \begin{cases} \hat{y} - \hat{y}_1 = \vec{0} \\ \vec{z}_1 - \vec{z} = \vec{0} \end{cases} \Leftrightarrow \begin{cases} \hat{y} = \hat{y}_1 \\ \vec{z} = \vec{z}_1 \end{cases} \quad \blacksquare$$

Corollaire. Si W est un sous-espace vectoriel de $\mathbb{R}^n$, alors

$$(W^\perp)^\perp = W$$

Preuve Nous avons déjà montré $W \subset (W^\perp)^\perp$.

Il reste à montrer $(W^\perp)^\perp \subset W$.

Soit donc $\vec{y} \in (W^\perp)^\perp \subset \mathbb{R}^n$.

Le théorème de la projection orthogonale nous donne:

$$\vec{y} = \hat{y} + \vec{z}, \text{ avec } \hat{y} \in W \text{ et } \vec{z} \in W^\perp$$

Comme $\vec{z} \in W^\perp$ et $\vec{y} \in (W^\perp)^\perp$, nous avons $\vec{y} \cdot \vec{z} = 0$

$$\text{d'où } 0 = \vec{y} \cdot \vec{z} = (\hat{y} + \vec{z}) \cdot \vec{z} = \hat{y} \cdot \vec{z} + \vec{z} \cdot \vec{z} = 0 + \vec{z} \cdot \vec{z}$$

Ainsi $\vec{z} \cdot \vec{z} = 0$ et $\vec{z} = \vec{0}$, ce qui implique $\vec{y} = \hat{y} \in W$ ■

Remarques.

- L'unicité de la décomposition $\vec{y} = \vec{\hat{y}} + \vec{z}$ montre que la projection $\text{proj}_W \vec{y}$ ne dépend que du sous-espace vectoriel W et non pas de la base (orthogonale) de W choisie.
- Si $\vec{y} \in W$, alors $\text{proj}_W \vec{y} = \vec{y}$.

Ceci nous donne un critère pratique pour caractériser l'appartenance d'un vecteur à un sous-espace vectoriel.

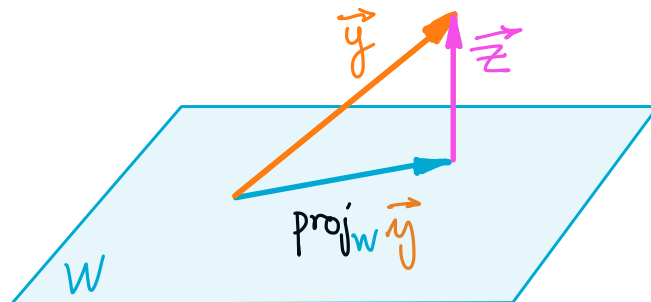
Définition.

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$.

Soit $\vec{y} \in \mathbb{R}^n$ un vecteur quelconque.

On définit la distance entre $\vec{y}$ et W , notée $\text{dist}(\vec{y}, W)$, par :

$$\text{dist}(\vec{y}, W) = \|\vec{y} - \text{proj}_W \vec{y}\| = \|\vec{z}\|$$



Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$,

On peut montrer que l'application $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie par

$$T(\vec{v}) = \text{proj}_W \vec{v}$$

est linéaire. Elle est appelée la **projection orthogonale** sur W .

$$\text{Comme } T(\vec{v}) = \vec{0} \Leftrightarrow \frac{\vec{v} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \dots + \frac{\vec{v} \cdot \vec{u}_k}{\vec{u}_k \cdot \vec{u}_k} \vec{u}_k = \vec{0}$$

$$\Leftrightarrow \vec{v} \cdot \vec{u}_j = 0 \text{ pour tout } j \in \{1, \dots, k\}$$

$$\Leftrightarrow \vec{v} \in W^\perp$$

nous avons $\text{Ker}(T) = W^\perp$

D'autre part, $\text{Im}(T) = W$.

Comme $\dim(W) = k$, le théorème du rang nous donne $\dim(W^\perp) = n - k$

Par construction,

- Si $\vec{v} \in W$, alors $\text{proj}_W \vec{v} = \vec{v}$. Ainsi $\vec{v}$ est un vecteur propre de proj_W de valeur propre associée $\lambda = 1$.

- Si $\vec{v} \in W^\perp$, alors $\text{proj}_W \vec{v} = \vec{0}$. Ainsi $\vec{v}$ est un vecteur propre de proj_W de valeur propre associée $\lambda = 0$.

Soit $\{\vec{b}_1, \dots, \vec{b}_k\}$ une base de W (pas forcément orthogonale)

et soit $\{\vec{v}_1, \dots, \vec{v}_{n-k}\}$ une base de $W^\perp$ (pas forcément orthogonale)

L'ensemble $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_k, \vec{v}_1, \dots, \vec{v}_{n-k}\}$ est une base de $\mathbb{R}^n$

et la matrice de $T = \text{proj}_W$ par rapport à la base $\mathcal{B}$ est diagonale:

$$A_{\mathcal{B}}^T = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ 0 & & & \ddots & \\ & & & & 0 \end{bmatrix} \begin{matrix} \left. \begin{matrix} 1 & \dots & 1 \end{matrix} \right\} k \text{ fois} \\ \left. \begin{matrix} 0 & \dots & 0 \end{matrix} \right\} n-k \text{ fois} \end{matrix}$$

Théorème. Si $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base orthonormale de W , alors

$$\text{proj}_W \vec{y} = (\vec{y} \cdot \vec{u}_1) \vec{u}_1 + (\vec{y} \cdot \vec{u}_2) \vec{u}_2 + \dots + (\vec{y} \cdot \vec{u}_k) \vec{u}_k$$

De plus, si U est la matrice de taille $n \times k$ suivante:

$$U = [\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_k]$$

alors nous avons

$$\text{proj}_W \vec{y} = (UU^T) \vec{y} \quad \text{pour tout } \vec{y} \in \mathbb{R}^n$$

Autrement dit, la matrice $A = UU^T$ est la matrice canoniquement associée à la projection orthogonale sur W

$$T = \text{proj}_W: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

Preuve. Le théorème de la projection orthogonale nous dit que tout vecteur $\vec{y} \in \mathbb{R}^n$ s'écrit de manière unique

$$\vec{y} = \hat{y} + \vec{z}, \quad \text{où } \hat{y} = \text{proj}_W \vec{y} \in W \quad \text{et} \quad \vec{z} \in W^\perp.$$

De plus, comme $\{\vec{u}_1, \dots, \vec{u}_k\}$ est une base orthonormale de W , nous trouvons:

$$\hat{y} = \text{proj}_W \vec{y} = (\vec{y} \cdot \vec{u}_1) \vec{u}_1 + (\vec{y} \cdot \vec{u}_2) \vec{u}_2 + \dots + (\vec{y} \cdot \vec{u}_k) \vec{u}_k$$

Comme $U = [\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_k]$ et $W = \text{Vect} \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ nous avons:

$$W = \text{Col}(U).$$

Par conséquent,

$$W^\perp = (\text{Col}(U))^\perp = \text{Nul}(U^T)$$

Comme $\hat{y} \in W = \text{Col}(U)$, on a

$$\hat{y} = U\vec{x} \text{ pour un certain } \vec{x} \in \mathbb{R}^k$$

Comme $\vec{z} \in W^\perp = \text{Nul}(U^T)$, on a $U^T\vec{z} = \vec{0}$.

$$\begin{aligned} \text{Ainsi } U^T\vec{y} &= U^T(\hat{y} + \vec{z}) = U^T\hat{y} + U^T\vec{z} = U^T(U\vec{x}) + \vec{0} \\ &= (U^TU)\vec{x} \end{aligned}$$

Comme les k colonnes de U forment un ensemble orthonormal, on a $U^TU = I_k$ ce qui nous donne

$$\vec{x} = U^T\vec{y}$$

d'où

$$\text{proj}_W \vec{y} = \hat{y} = U\vec{x} = U(U^T\vec{y}) = (UU^T)\vec{y} \quad \blacksquare$$

Matrices orthogonales.

Soit U une matrice de taille $m \times n$:

$$U = [\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_n], \text{ avec } \vec{u}_j \in \mathbb{R}^m \text{ pour tout } j \in \{1, \dots, n\}$$

Théorème.

Les n colonnes d'une matrice U de taille $m \times n$ sont orthonormées si et seulement si $U^TU = I_n$.

Attention.

La matrice UU^T est une matrice de taille $m \times m$ et est associée à la projection orthogonale sur le sous-espace vectoriel $W = \text{Vect}\{\vec{u}_1, \dots, \vec{u}_n\} \subset \mathbb{R}^m$.

Théorème.

Soit U une matrice de taille $m \times n$ dont les colonnes sont orthonormées.

Si $\vec{x}, \vec{y} \in \mathbb{R}^n$, alors :

1. $\|U\vec{x}\| = \|\vec{x}\|$
2. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$
3. $(U\vec{x}) \cdot (U\vec{y}) = 0$ si et seulement si $\vec{x} \cdot \vec{y} = 0$.

Preuve. Voir Exercice 11, série 13. ■

Conséquence.

L'application linéaire $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ définie par $T(\vec{x}) = U\vec{x}$ conserve les longueurs et l'orthogonalité.

Définition.

Soit A une matrice carrée de taille $n \times n$.

On dit que A est une **matrice orthogonale** si A est inversible et

$$A^{-1} = A^T$$

Remarque.

Comme par définition nous avons $A^T A = I_n$, les colonnes de A sont orthonormées et forment une **base orthonormale** de $\mathbb{R}^n$.

De plus, comme $A A^T = I_n$, les colonnes de A^T (c-à-d.

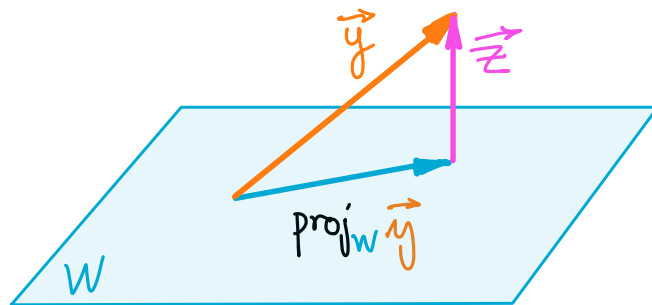
les lignes de A) sont aussi orthonormées et forment une autre base orthonormale de $\mathbb{R}^n$.

Théorème de la meilleure approximation.

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

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

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



Le vecteur $\text{proj}_W \vec{y}$ 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

$$\begin{aligned} \|\vec{y} - \vec{w}\|^2 &= \|\vec{y} - \text{proj}_W \vec{y}\|^2 + \|\text{proj}_W \vec{y} - \vec{w}\|^2 \\ &> \|\vec{y} - \text{proj}_W \vec{y}\|^2 \quad \underbrace{> 0 \text{ si } \vec{w} \neq \text{proj}_W \vec{y}} \end{aligned}$$

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 $\text{proj}_W \vec{y}$ 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 $W = \text{Vect} \{ \vec{x}_1, \vec{x}_2 \}$ (ici $W = \mathbb{R}^2$)

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

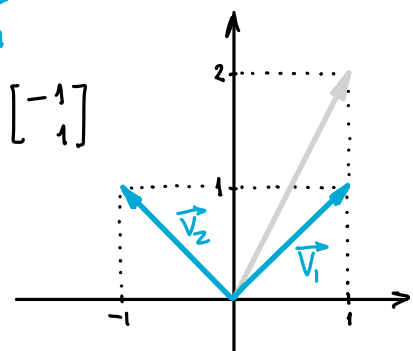
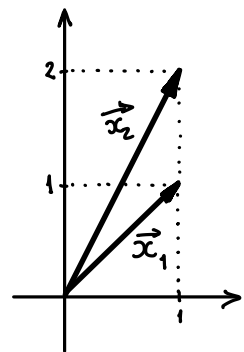
On a $\text{proj}_{W_1} \vec{x}_2 = \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 = \frac{1+2}{1+1} \vec{v}_1 = \frac{3}{2} \vec{v}_1$

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

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.

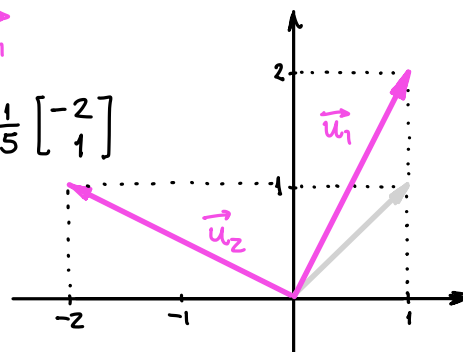
Si l'on pose $\vec{u}_1 = \vec{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ et $W_1 = \text{Vect}\{\vec{u}_1\}$

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

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

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 $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ de W .
 2. On pose $\vec{v}_1 = \vec{x}_1$ et on note $W_1 = \text{Vect}\{\vec{v}_1\}$ ($= \text{Vect}\{\vec{x}_1\}$)
 3. On calcule $\text{proj}_{W_1} \vec{x}_2$ et $\vec{x}_2 - \text{proj}_{W_1} \vec{x}_2$
 4. On pose $\vec{v}_2 = \vec{x}_2 - \text{proj}_{W_1} \vec{x}_2$ (ou un multiple de ce vecteur de sorte que $\vec{v}_2$ ait des composantes entières).
et on note $W_2 = \text{Vect}\{\vec{v}_1, \vec{v}_2\}$ ($= \text{Vect}\{\vec{x}_1, \vec{x}_2\}$)
- 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.

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

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

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

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

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

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

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

$$\begin{aligned} \text{On calcule } \text{proj}_{W_2} \vec{x}_3 &= \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 \\ &= \frac{0 + 1 + 0 + 1 + 2}{1 + 1 + 1 + 0 + 1} \vec{v}_1 + \frac{0 + 3 + 2 - 1 + 2}{9 + 9 + 4 + 1 + 1} \vec{v}_2 = \vec{v}_1 + \frac{1}{4} \vec{v}_2 \end{aligned}$$

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

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

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 $\text{rang}(A) = n$)

Alors la matrice A peut s'écrire sous la forme

$$A = QR$$

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

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 $\{\vec{a}_1, \dots, \vec{a}_n\}$ de $\text{Col}(A)$.

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

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

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

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

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

d'où $R = Q^T A$

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 $\text{Col}(A)$.

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

est une base orthonormale de $\text{Col}(A)$.

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

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

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

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.

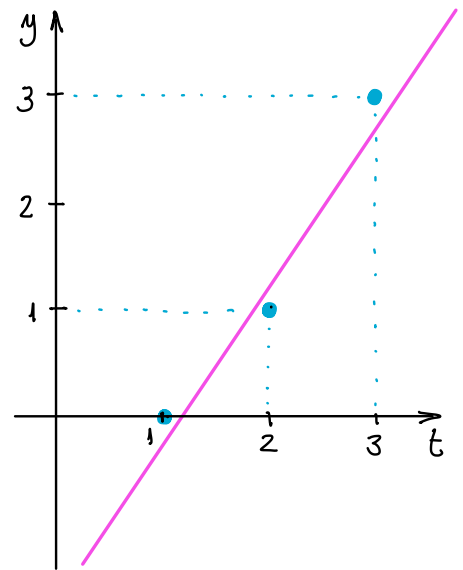
Considérons trois points du plan :

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

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 :

$$\begin{aligned} (t_1, y_1) = (1, 0) : & \begin{cases} 0 = x_1 + 1 x_2 \\ 1 = x_1 + 2 x_2 \\ 3 = x_1 + 3 x_2 \end{cases} \\ (t_2, y_2) = (2, 1) : & \\ (t_3, y_3) = (3, 3) : & \end{aligned}$$

Équation 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 :

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

ce qui confirme que les points ne sont pas alignés.

Considérons maintenant le cas général:

Soit A une matrice de taille $m \times n$ et $\vec{b} \in \mathbb{R}^m$.

Nous avons vu que le système $A\vec{x} = \vec{b}$ possède des solutions si et seulement si $\vec{b} \in \text{Col}(A)$.

Si $m > n$, alors il est fort probable que cela n'arrive pas.

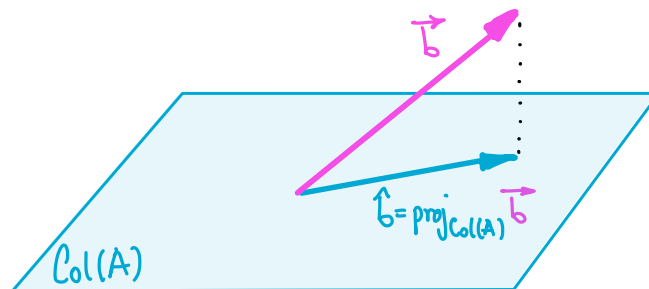
En pratique, on s'intéresse à des situations où $m \gg n$ (m beaucoup plus grand que n).

En particulier, si $\vec{b} \notin \text{Col}(A)$ alors $A\vec{x} = \vec{b}$ ne possède pas de solution!

On prend

$$\hat{\vec{b}} = \text{proj}_{\text{Col}(A)} \vec{b}$$

Géométriquement:



Comme par construction $\hat{\vec{b}} \in \text{Col}(A)$, le système $A\vec{x} = \hat{\vec{b}}$ est consistant et il existe au moins un $\hat{\vec{x}} \in \mathbb{R}^n$ tel que

$$A\hat{\vec{x}} = \hat{\vec{b}}.$$

Définition.

Soit A une matrice de taille $m \times n$ et $\vec{b} \in \mathbb{R}^m$.

On dit que $\hat{x} \in \mathbb{R}^n$ est une solution au sens des moindres carrés de l'équation $A\vec{x} = \vec{b}$ si

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{y}\| \quad \text{pour tout } \vec{y} \in \mathbb{R}^n.$$

La distance $\|\vec{b} - A\hat{x}\|$ est appelée erreur de l'approximation.

Remarque.

Si $\vec{b} \in \text{Col}(A)$ alors $\hat{x} \in \mathbb{R}^n$ n'est rien d'autre qu'une solution de l'équation $A\vec{x} = \vec{b}$

Si $\vec{b} \notin \text{Col}(A)$ alors $\hat{x} \in \mathbb{R}^n$ n'est pas une solution de $A\vec{x} = \vec{b}$.

Question. Comment déterminer $\hat{x}$?

Par construction, $\vec{b} - \hat{b} \in (\text{Col}(A))^\perp = \text{Nul}(A^T)$.

Par conséquent, $A^T(\vec{b} - \hat{b}) = \vec{0} \iff A^T\vec{b} = A^T\hat{b}$

Comme $\hat{x} \in \mathbb{R}^n$ est tel que $A\hat{x} = \hat{b}$ nous avons donc :

$$(A^T A) \hat{x} = A^T \vec{b}$$

Définition.

Soit A une matrice de taille $m \times n$ et $\vec{b} \in \mathbb{R}^m$.

Le système $(A^T A)\vec{x} = A^T \vec{b}$ est appelé système normal associé au système $A\vec{x} = \vec{b}$.

Remarque.

Si $\hat{x} \in \mathbb{R}^n$ est une solution au sens des moindres carrés de $A\vec{x} = \vec{b}$, alors $\hat{x}$ est une solution du système normal $(A^T A)\vec{x} = A^T \vec{b}$ associé.

Théorème.

L'ensemble des solutions au sens des moindres carrés de l'équation $A\vec{x} = \vec{b}$ est égal à l'ensemble (non-vide) des solutions du système normal $(A^T A)\vec{x} = A^T \vec{b}$ associé.

Preuve.

Il reste à montrer que si $\hat{x} \in \mathbb{R}^n$ est une solution de $(A^T A)\vec{x} = A^T \vec{b}$, alors $\hat{x}$ est une solution au sens des moindres carrés de $A\vec{x} = \vec{b}$.

Si on a $A^T A \hat{x} = A^T \vec{b}$, alors $A^T (A \hat{x} - \vec{b}) = \vec{0}$ et

Par conséquent, $\vec{b} - A \hat{x} \in \text{Nul}(A^T) = (\text{Col}(A))^\perp$.

Comme $\vec{b} = \underbrace{(\vec{b} - A \hat{x})}_{\in (\text{Col}(A))^\perp} + \underbrace{A \hat{x}}_{\in \text{Col}(A)}$

l'unicité de la décomposition orthogonale nous dit que

$$A \hat{x} = \text{proj}_{\text{Col}(A)} \vec{b} = \vec{b}$$

autrement dit, $\hat{x}$ est une solution au sens des moindres carrés de $A\vec{x} = \vec{b}$. ■

Exemple.

Soient $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ et $\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$

Trouver une solution au sens des moindres carrés du système $A\vec{x} = \vec{b}$.

On a $A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$

et $A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$

matrice augmentée associée au système normal:

$$\left[\begin{array}{cc|c} 3 & 6 & 4 \\ 6 & 14 & 11 \end{array} \right] \xrightarrow{L_2 \rightarrow L_2 - 2L_1} \left[\begin{array}{cc|c} 3 & 6 & 4 \\ 0 & 2 & 3 \end{array} \right] \xrightarrow{L_1 \rightarrow L_1 - 3L_2} \left[\begin{array}{cc|c} 3 & 0 & -5 \\ 0 & 2 & 3 \end{array} \right] \sim \left[\begin{array}{cc|c} 1 & 0 & -5/3 \\ 0 & 1 & 3/2 \end{array} \right]$$

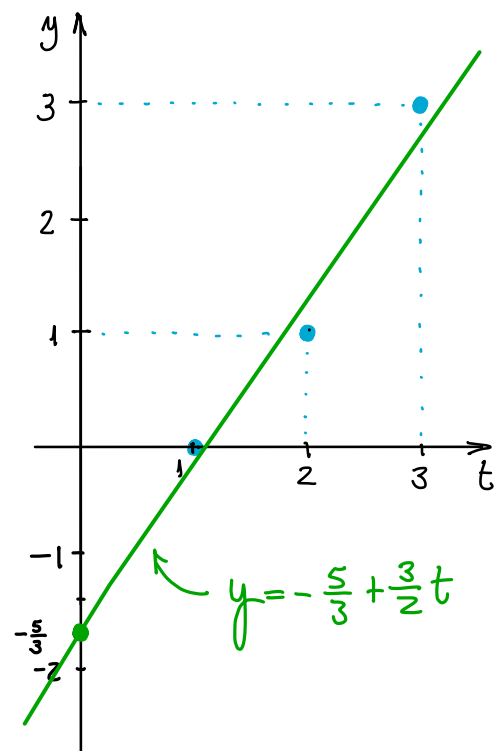
solution: $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} -5/3 \\ 3/2 \end{bmatrix}$

Remarque.

La matrice A et le vecteur $\vec{b}$ de l'exemple sont associés au problème de recherche de la droite la plus "proche" des points $(1,0), (2,1), (3,3)$. Ainsi,

$$y = \hat{x}_1 + \hat{x}_2 t = -\frac{5}{3} + \frac{3}{2} t$$

est la droite la plus "proche" des points $(1,0), (2,1), (3,3)$, appelée droite de régression linéaire ou droite des moindres carrés.



De manière générale, la droite de régression linéaire pour les points

$$(t_1, y_1), (t_2, y_2), \dots, (t_N, y_N) \in \mathbb{R}^2$$

peut être trouvée à l'aide de la solution au sens des moindres carrés du système $A\vec{x} = \vec{b}$ où

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_N \end{bmatrix} \in M_{N,2}(\mathbb{R}) \quad \text{et} \quad \vec{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \in \mathbb{R}^N$$

Si $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$ est la solution au sens des moindres carrés, alors l'équation de la droite de régression linéaire est:

$$y = \hat{x}_1 + \hat{x}_2 t$$

Exemple.

$$\text{Soit } A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{et} \quad \vec{b} = \begin{bmatrix} 4 \\ 4 \\ 5 \\ 5 \\ 7 \end{bmatrix}$$

On constate que le système $A\vec{x} = \vec{b}$ n'a pas de solution.

Trouver une solution au sens des moindres carrés du système $A\vec{x} = \vec{b}$.

$$\text{On a} \quad A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

$$\text{et} \quad A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 5 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 25 \\ 7 \end{bmatrix}$$

$$\text{d'où} \quad \left[\begin{array}{cc|c} 5 & 0 & 25 \\ 0 & 10 & 7 \end{array} \right] \sim \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 7/10 \end{array} \right]$$

$$\Rightarrow \text{solution: } \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 7/10 \end{bmatrix}$$

Conséquence.

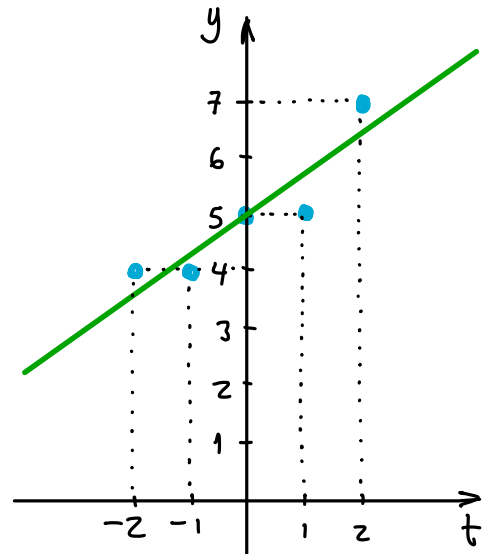
La droite

$$y = \hat{x}_1 + \hat{x}_2 t = 5 + \frac{7}{10} t$$

est la droite de régression linéaire

pour les points

$$(-2, 4), (-1, 4), (0, 5), (1, 5), (2, 7)$$



Théorème. Soit A une matrice de taille $m \times n$.

Les affirmations suivantes sont équivalentes:

1. $A\vec{x} = \vec{b}$ admet une unique solution au sens des moindres carrés pour tout $\vec{b} \in \mathbb{R}^m$.

2. Les colonnes de A sont linéairement indépendantes

3. $\text{rang}(A) = n$

4. $\text{rang}(A^T A) = n$

5. La matrice $A^T A$ est inversible

Dans ce cas, la solution au sens des moindres carrés s'écrit

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

Remarque.

Pour montrer l'équivalence entre 3 et 4 on peut utiliser :

Si A est une matrice de taille $m \times n$, alors $\text{rang}(A^T A) = \text{rang}(A)$

(voir exercice 16, série 13).

Théorème.

Soit A une matrice de taille $m \times n$ dont les colonnes sont linéairement indépendantes.

Soit $A = QR$ une factorisation QR de A . La solution au sens des moindres carrés de $A\vec{x} = \vec{b}$ est unique et s'écrit :

$$\hat{\vec{x}} = R^{-1} Q^T \vec{b}$$

Preuve.

Comme les colonnes de A sont linéairement indépendantes, la solution au sens des moindres carrés est unique :

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

Par hypothèse, A peut s'écrire sous la forme $A = QR$

où Q est une matrice de taille $m \times n$ telle que $Q^T Q = I_n$

et R est une matrice de taille $n \times n$ triangulaire supérieure inversible.

On a $A^T = R^T Q^T$ et $A^T A = (R^T Q^T)(QR) = R^T (Q^T Q) R = R^T R$,

d'où $(A^T A)^{-1} A^T = (R^T R)^{-1} R^T Q^T = R^{-1} (R^T)^{-1} R^T Q^T = R^{-1} Q^T$

et $\hat{\vec{x}} = R^{-1} Q^T \vec{b}$



Calcul alternatif de la projection orthogonale.

Nous avons défini $\hat{\vec{b}} = \text{proj}_{\text{Col}(A)} \vec{b}$

Comme $\hat{\vec{x}} \in \mathbb{R}^n$ est tel que $A\hat{\vec{x}} = \hat{\vec{b}}$, pour trouver la projection orthogonale de $\vec{b}$ sur $\text{Col}(A)$ il suffit de trouver une solution $\hat{\vec{x}}$ du système normal $(A^T A)\hat{\vec{x}} = A^T \vec{b}$ et la multiplier par A :

$$\text{proj}_{\text{Col}(A)} \vec{b} = A\hat{\vec{x}}$$

plutôt que de construire une base orthogonale à l'aide du procédé d'orthogonalisation de Gram-Schmidt.

(voir exercice 8, série 14).

Complément: (hors programme)

Théorème.

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$.

Soit $\{\vec{w}_1, \dots, \vec{w}_k\}$ une base quelconque de W (pas forcément orthogonale). Soit A la matrice de taille $n \times k$ suivante.

$$A = [\vec{w}_1 \ \vec{w}_2 \ \dots \ \vec{w}_k]$$

alors la matrice canoniquement associée à l'application

linéaire $T = \text{proj}_W: \mathbb{R}^n \rightarrow \mathbb{R}^n$ est

$$A_T = A(A^T A)^{-1} A^T$$

Preuve.

Comme $\{\vec{w}_1, \dots, \vec{w}_k\}$ est une base de W et $A = [\vec{w}_1 \ \vec{w}_2 \ \dots \ \vec{w}_k]$, nous avons $W = \text{Col}(A)$.

Par conséquent, $W^\perp = (\text{Col}(A))^\perp = \text{Nul}(A^T)$

Le théorème de la décomposition spectrale nous dit que tout vecteur $\vec{y} \in \mathbb{R}^n$ s'écrit de manière unique

$$\vec{y} = \hat{\vec{y}} + \vec{z}, \text{ où } \hat{\vec{y}} \in W \text{ et } \vec{z} \in W^\perp.$$

Comme $\hat{\vec{y}} \in W = \text{Col}(A)$, on a $\hat{\vec{y}} = A \vec{x}$, avec $\vec{x} \in \mathbb{R}^k$

Comme $\vec{z} \in W^\perp = \text{Nul}(A^T)$, on a $A^T \vec{z} = \vec{0}$.

$$\text{Ainsi } A^T \vec{y} = A^T (\hat{\vec{y}} + \vec{z}) = A^T \hat{\vec{y}} + A^T \vec{z} = A^T A \vec{x} + \vec{0}$$

$$\text{d'où } A^T \vec{y} = A^T A \vec{x}$$

Comme $A^T A$ est une matrice carrée de taille $k \times k$

et $\text{rang}(A^T A) = \text{rang}(A)$ (voir exercice 16, série 13)

la matrice $A^T A$ est inversible car $\text{rang}(A^T A) = k$.

Nous avons donc $\vec{x} = (A^T A)^{-1} A^T \vec{y}$, d'où

$$\begin{aligned} \text{proj}_W \vec{y} &= \hat{\vec{y}} = A \vec{x} \\ &= A (A^T A)^{-1} A^T \vec{y} \end{aligned}$$

■