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 $\vec{u}^T = [u_1 \ u_2 \cdots u_n]$ 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

$$\boxed{\vec{u}\bullet\vec{v} = \vec{u}^T\vec{v}}$$

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

On le note aussi $<\vec{u},\vec{v}>$.

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 :

$$\boxed{\vec{u}\bullet\vec{v} = \vec{u}^T\vec{v} = [u_1 \cdots u_n]\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \sum_{j=1}^{n} u_jv_j}$$

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 : $\vec{u} \cdot \vec{u} = 0 \iff \vec{u} = \vec{0}$

Preuve.

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

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

2. $\vec{u} \cdot (\vec{v} + \vec{w}) \overset{\text{déf}}{=} \vec{u}^T (\vec{v} + \vec{w}) = \vec{u}^T \vec{v} + \vec{u}^T \vec{w} = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$

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. $\vec{u} \cdot \vec{u} = \underbrace{u_1^2}_{\geqslant 0} + \underbrace{u_2^2}_{\geqslant 0} + \ldots + \underbrace{u_n^2}_{\geqslant 0} \geqslant 0$

   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. $\blacksquare$

Conséquence.

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

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

$$\boxed{\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}} \geqslant 0$$

Remarque. Pour des calculs théoriques, on utilise plutôt la formule $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$.

Propriété. 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

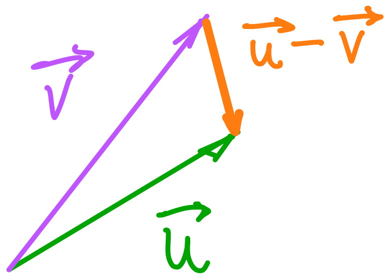
$$\frac{1}{\|\vec{u}\|}\vec{u} = \frac{\vec{u}}{\|\vec{u}\|}$$

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

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

Geométriquement:



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

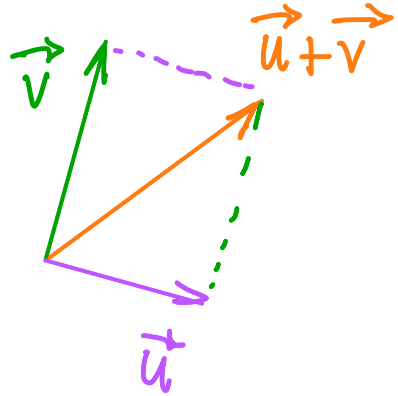
$$\boxed{\vec{u} \cdot \vec{v} = 0}$$

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 :

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

Preuve.



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

d'où : $\vec{u} \cdot \vec{v} = 0 \iff \|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$ ∎

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**.

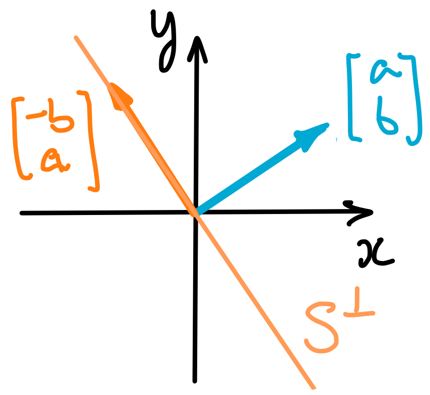
L'ensemble de tous les vecteurs orthogonaux à S est appelé **complément orthogonal de S**, noté $S^\perp$ (se lit « S orthogonal ») :

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

Exemples.

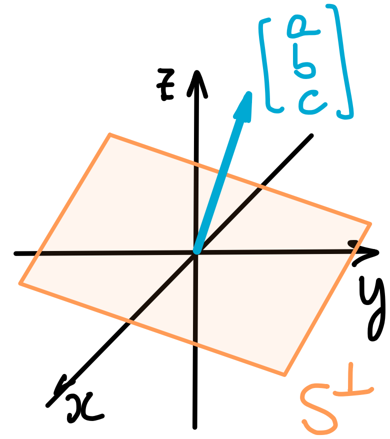
$n=2$: Si $S=\left\{\begin{bmatrix} a \\ b \end{bmatrix}\right\}$, alors

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



$n=3$: 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 $\vec{v} \in S$

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

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

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

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

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

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

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

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

$$\boxed{W^\perp = S^\perp}$$

Autrement dit,

$\vec{x} \in W^\perp \Longleftrightarrow \vec{x}$ 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 $j \in \{1, \ldots, k\}$

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 $W = \mathrm{Vect}\{\vec{v_1}, \dots, \vec{v_k}\}$, 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 $\vec{x} \in W^\perp$ $\blacksquare$

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

$$\boxed{W \subset (W^\perp)^\perp}$$

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 :

$$\boxed{W = (W^\perp)^\perp}$$

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

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

Preuve.

$$\vec{x} \in \mathrm{Nul}(A) \iff A\vec{x} = \vec{0}$$

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

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

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

$$\iff \vec{x} \in (\mathrm{Lgn}(A))^{\perp}$$

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

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

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

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

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

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

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, $\boxed{\mathrm{Lgn}(A) = W}$

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

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

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

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

Exemple

Soit $W = \mathrm{Vect}\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ avec $\vec{v_1} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} -1 \\ 5 \\ 4 \end{bmatrix}$ et $\vec{v_3} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

Déterminer $W^\perp$.

Nous avons vu que pour trouver une base de W on peut construire une matrice A dont les lignes sont $\vec{v_1}, \vec{v_2}, \vec{v_3}$. Autrement dit, $\boxed{W = \mathrm{Lgn}(A)}$

$$A = \begin{bmatrix} 2 & 3 & 5 \\ -1 & 5 & 4 \\ 3 & -2 & 1 \end{bmatrix} \underset{L_1 \to L_1 + L_2}{\sim} \begin{bmatrix} 1 & 8 & 9 \\ -1 & 5 & 4 \\ 3 & -2 & 1 \end{bmatrix} \underset{\substack{L_2 \to L_2 + L_1 \\ L_3 \to L_3 - 3L_1}}{\sim} \begin{bmatrix} 1 & 8 & 9 \\ 0 & 13 & 13 \\ 0 & -26 & -26 \end{bmatrix} \sim \begin{bmatrix} 1 & 8 & 9 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = R$$

Ainsi, $\dim W = 2$ et $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ est une base de W.

système homogène associé : $\begin{cases} x + z = 0 \\ y + z = 0 \end{cases} \iff \begin{cases} x = -z \\ y = -z \end{cases}$

solution générale : $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, avec $t \in \mathbb{R}$ $\Rightarrow$ $\boxed{W^\perp = \mathrm{Vect}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}}$