

Chapitre 6 : Orthogonalité et moindres carrés

But

Rappelons le résultat suivant : Si $A \in M_{n \times n}(\mathbb{R})$, on a les équivalences suivantes :

- 1) $A\vec{x} = \vec{b}$ est compatible $\forall \vec{b} \in \mathbb{R}^n$
- 2) $\text{Im}(A) = \mathbb{R}^n$
- 3) A est inversible

Mais : si la matrice A est singulière, on peut trouver des vecteurs $\vec{b} \in \mathbb{R}^n$ tels que

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

soit incompatible.

Plus généralement, si $A \in M_{m \times n}(\mathbb{R})$, il peut exister des vecteurs $\vec{b} \in \mathbb{R}^m$ pour lesquels

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

soit incompatible.

Dans ces cas, on va essayer d'approximer $\vec{b}$ en cherchant $\vec{x} \in \mathbb{R}^n$ avec

$$A\vec{x} \simeq \vec{b}$$

Pour cela, on doit définir une notion de distance entre vecteurs.

6.1 Produit scalaire

Définition 58 (produit scalaire usuel dans $\mathbb{R}^n$).

Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$. On définit le *produit scalaire usuel* $\vec{u} \cdot \vec{v}$ (parfois aussi noté $\langle \vec{u}, \vec{v} \rangle$) par

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{i=1}^n u_i v_i \quad \text{où}$$

$\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ sont les composantes dans la base canonique de $\mathbb{R}^n$.

On $\vec{u} \cdot \vec{v} \in \mathbb{R}$, d'où le nom de produit "scalaire".

Remarque

$$\vec{u} \cdot \vec{v} = (\underbrace{u_1 \dots u_n}_{1 \times n}) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{n \times 1} = \sum_{i=1}^n u_i v_i$$

donc $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$ (mult. matricielle)

Théorème 56. Soient $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ et soit $\alpha \in \mathbb{R}$. Alors

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (commutativité)
2. $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$ (distributivité)
3. $(\alpha \vec{u}) \cdot \vec{v} = \alpha (\vec{u} \cdot \vec{v})$ (associativité mixte)
4. $\vec{u} \cdot \vec{u} \geq 0$ et $\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$.

Preuve

1 à 3 en exercice

4. Soit $\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$. Alors $\vec{u} \cdot \vec{u} = \sum_{i=1}^n u_i^2 \geq 0$.

$$\begin{aligned} \text{Cette somme est nulle} &\Leftrightarrow u_i^2 = 0 \quad \forall 1 \leq i \leq n \\ &\Leftrightarrow u_i = 0 \quad \forall 1 \leq i \leq n \\ &\Leftrightarrow \vec{u} = \vec{0}. \end{aligned}$$

Généralisation

Dans un EV quelconque, il n'est pas forcément possible de définir une notion de produit scalaire généralisé.

Définition 59 (produit scalaire).

Soient V un espace vectoriel et $u, v \in V$. On dira qu'une application

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R} \\ (u, v) \mapsto \langle u, v \rangle$$

est un *produit scalaire* si les conditions suivantes sont satisfaites :

1. $\langle u, v \rangle = \langle v, u \rangle$
2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
3. $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ pour tout $\alpha \in \mathbb{R}$
4. $\langle u, u \rangle \geq 0$ avec $\langle u, u \rangle = 0$ si et seulement si $u = 0_V$.

Si V admet un tel produit scalaire, on parle d'espace préhilbertien, ou d'espace euclidien dans le cas où V est de dimension finie. Beaucoup de notions et résultats liés au produit scalaire usuel dans $\mathbb{R}^n$ peuvent alors être généralisés.

Norme d'un vecteur et distance dans $\mathbb{R}^n$

Définition 60 (norme).

Soit $\vec{v} \in \mathbb{R}^n$. La *norme* de $\vec{v}$ est le scalaire donné par

$$\|\vec{v}\| = \sqrt{v_1^2 + \dots + v_n^2} \in \mathbb{R}_+$$

$$\|\vec{v}\| \geq 0$$

Remarque La norme coïncide avec la notion de longueur usuelle.

$$\begin{aligned} (-1)^2 &= 1 & \sqrt{1} &= 1 \\ (+1)^2 &= 1 \end{aligned}$$

Propriétés

$$1) \quad \|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$$

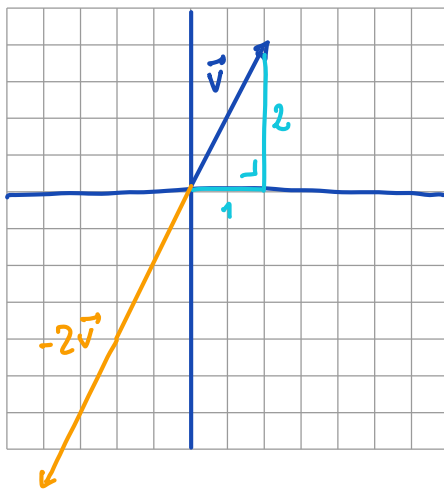
$$2) \quad \|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$$

$$\begin{aligned} 3) \quad \text{Soit } \alpha \in \mathbb{R}. \text{ Alors } \|\alpha \vec{v}\| &= \sqrt{\alpha^2 v_1^2 + \dots + \alpha^2 v_n^2} \\ &= \sqrt{\alpha^2} \cdot \sqrt{v_1^2 + \dots + v_n^2} = |\alpha| \cdot \|\vec{v}\| \in \mathbb{R}_+ \end{aligned}$$

Exemples

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \|\vec{v}\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

$$\begin{aligned} -2\vec{v} &= \begin{pmatrix} -2 \\ -4 \end{pmatrix} \quad \|-2\vec{v}\| = \sqrt{(-2)^2 + (-4)^2} = \sqrt{20} = \sqrt{4 \cdot 5} \\ &= 2 \cdot \sqrt{5} = |-2| \cdot \|\vec{v}\| \end{aligned}$$



Pythagore $\Rightarrow$

$$\|\vec{v}\| = \sqrt{1^2 + 2^2}$$

$$\|\vec{v}\|^2 = 1^2 + 2^2$$

Définition 61 (unitaire).

Un vecteur $\vec{v} \in \mathbb{R}^n$ est dit *unitaire* si $\vec{v}$ est de norme 1.

Remarque

Soit $\vec{v} \in \mathbb{R}^n$. Si $\vec{v} \neq \vec{0}$, alors $\|\vec{v}\| > 0$ et on peut considérer le vecteur

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\|\vec{v}\|} \cdot \vec{v}. \quad \text{Alors } \vec{u} \text{ est unitaire et colinéaire à } \vec{v}:$$

$$\|\vec{u}\| = \left\| \frac{1}{\|\vec{v}\|} \cdot \vec{v} \right\| = \frac{1}{\|\vec{v}\|} \cdot \|\vec{v}\| = 1 \in \mathbb{R}.$$

Exemple $\vec{v} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$ $\|\vec{v}\| = \sqrt{1^2 + (-3)^2 + 2^2} = \sqrt{14}$

$\vec{v}$ n'est pas unitaire, mais

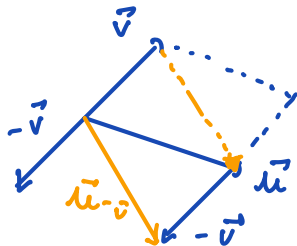
$$\vec{u} = \frac{1}{\sqrt{14}} \vec{v} \quad \text{e'est.}$$

Définition 62 (distance).

Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$. La *distance* entre $\vec{u}$ et $\vec{v}$ est définie par la norme du vecteur $\vec{u} - \vec{v}$. On la note

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

Remarque $\text{dist}(\vec{u}, \vec{v}) = 0 \Leftrightarrow \|\vec{u} - \vec{v}\| = 0$
 $\Leftrightarrow \vec{u} - \vec{v} = \vec{0}$
 $\Leftrightarrow \vec{u} = \vec{v}$



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

Exemples

1) $\vec{u} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \left\| \begin{pmatrix} 4 \\ -1 \end{pmatrix} \right\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

2) Si $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$, alors

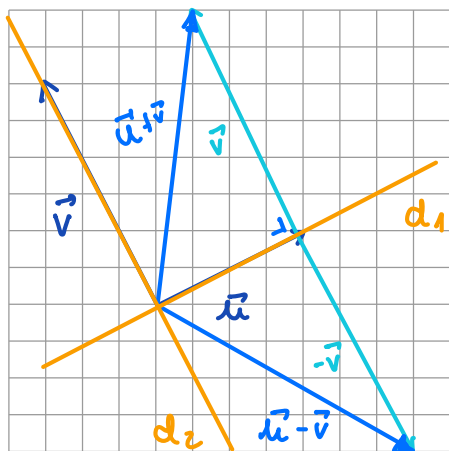
$$\text{dist}(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$$

Par ex $\vec{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $\vec{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\text{dist}(\vec{u}, \vec{v}) = \sqrt{1^2 + 1^2} = \sqrt{2}$.

6.2 Orthogonalité

Dans $\mathbb{R}^2$, considérons les vecteurs $\vec{u}$ et $\vec{v}$, ainsi que $d_1 = \text{span}\{\vec{u}\}$ et $d_2 = \text{span}\{\vec{v}\}$ les deux droites vectorielles associées.

But : Trouver une condition sur $\vec{u}$ et $\vec{v}$ pour que les deux droites vectorielles soient orthogonales.



On peut construire un Δ isocèle si et seulement si :
 $d_1 \perp d_2$.

$$d_1 \perp d_2 \Leftrightarrow \|\vec{u} + \vec{v}\| = \|\vec{u} - \vec{v}\|$$

$$\Leftrightarrow \|\vec{u} + \vec{v}\|^2 = \|\vec{u} - \vec{v}\|^2$$

$$\|\vec{u} + \vec{v}\|^2 = (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v})$$

$$\|\vec{u} - \vec{v}\|^2 = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$$

$$d_1 \perp d_2 \Leftrightarrow (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$$

$$\Leftrightarrow \cancel{\vec{u} \cdot \vec{u}} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \cancel{\vec{v} \cdot \vec{v}} = \cancel{\vec{u} \cdot \vec{u}} - \vec{u} \cdot \vec{v} - \vec{v} \cdot \vec{u} + \cancel{\vec{v} \cdot \vec{v}}$$

$$\Leftrightarrow 2 \vec{u} \cdot \vec{v} = -2 \vec{u} \cdot \vec{v}$$

$$\Leftrightarrow 4 \vec{u} \cdot \vec{v} = 0$$

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

En résumé,

$$\underline{d_1 \perp d_2 \Leftrightarrow \vec{u} \cdot \vec{v} = 0}$$

Définition 63 (vecteurs orthogonaux).

Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$. Les vecteurs $\vec{u}$ et $\vec{v}$ sont *orthogonaux* si et seulement si

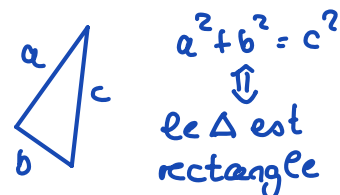
$$\vec{u} \cdot \vec{v} = 0$$

Notation: $\vec{u} \perp \vec{v}$

Remarque: $\vec{0}$ est orthogonal à tous les vecteurs de $\mathbb{R}^n$.

Théorème 57 (de Pythagore). Soient $\vec{u}, \vec{v} \in \mathbb{R}^n$ alors

$$\vec{u} \perp \vec{v} \Leftrightarrow \|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$



Preuve

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

Orthogonal à un sous-espace vectoriel de $\mathbb{R}^n$

Soit $W = \text{span} \left\{ \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\} \in \mathbb{R}^2$. On cherche l'ens. de tous les vecteurs orthogonaux à W : Soit $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2$.

$$\begin{aligned} \vec{u} \perp W &\Leftrightarrow \vec{u} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 0 \Leftrightarrow 3u_1 + 2u_2 = 0 \\ &\Leftrightarrow u_1 = -\frac{2}{3}u_2 \Leftrightarrow \vec{u} = t \begin{pmatrix} -2/3 \\ 1 \end{pmatrix} \\ &\Leftrightarrow \vec{u} = s \begin{pmatrix} -2 \\ 3 \end{pmatrix} \end{aligned}$$

Ainsi: $\text{span} \left\{ \begin{pmatrix} -2 \\ 3 \end{pmatrix} \right\} \perp W$

Définition 64 (ensemble orthogonal).

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

1. Un vecteur $\vec{v} \in \mathbb{R}^n$ est dit *orthogonal* à W s'il est orthogonal à tous les vecteurs du sous-espace W . $\vec{v} \cdot \vec{w} = 0$ par tout $\vec{w} \in W$
2. On définit l'orthogonal de W par

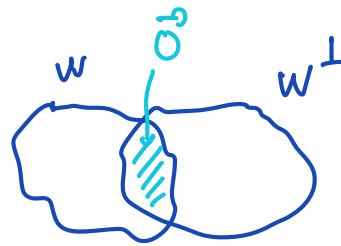
$$W^\perp = \left\{ \vec{v} \in \mathbb{R}^n \mid \vec{v} \cdot \vec{w} = 0 \quad \forall \vec{w} \in W \right\}$$

$W^\perp$ est l'ens. de tous les vecteurs orthogonaux à W .

Théorème 58. Soit W un sous-espace vectoriel de $\mathbb{R}^n$. Alors

1. $W^\perp$ est un SEV de $\mathbb{R}^n$.

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



Rappel: Si E et F sont deux ensembles,

$$E \cap F = \{p \mid p \in E \text{ et } p \in F\} \quad (\text{intersection})$$

1. en exercice

Preuve

2. $\vec{0} \in W$ et $\vec{0} \in W^\perp$ (car ce sont des SEV)

Soit $\vec{v} \in W \cap W^\perp$. Alors $\vec{v} \in W$ et $\vec{v} \in W^\perp$.

$$\Rightarrow \vec{v} \cdot \vec{w} = 0 \quad \forall \vec{w} \in W$$

$$\text{En particulier } \vec{v} \cdot \vec{v} = 0, \text{ donc } \vec{v} = \vec{0}$$

$$\text{Donc } \{\vec{0}\} \subset W \cap W^\perp$$

$$\text{et } W \cap W^\perp \subset \{\vec{0}\}$$

$$\text{d'où } W \cap W^\perp = \{\vec{0}\}.$$

Remarque On a le résultat suivant :

Si W est un SEV de $\mathbb{R}^n$, alors

$$\dim W + \dim W^\perp = n$$

(voir série à venir)

Théorème 59. Soit A une matrice $m \times n$ donnée par $A = (\vec{a}_1 \cdots \vec{a}_n)$
où $\vec{a}_i \in \mathbb{R}^m$. Soient

$$\text{Ker}(A) = \{ \vec{u} \in \mathbb{R}^n \mid A\vec{u} = \vec{0} \} \subseteq \mathbb{R}^n$$

$$\text{Im}(A) = \text{span} \{ \vec{a}_1, \dots, \vec{a}_n \} \subseteq \mathbb{R}^m$$

$$\text{Lgn}(A) = \text{span} \{ \text{Lgn}_1(A)^T, \dots, \text{Lgn}_m(A)^T \} \subseteq \mathbb{R}^n$$

Alors

$$1. \quad \text{Lgn}(A)^\perp = \text{Ker}(A)$$

$$2. \quad \text{Im}(A)^\perp = \text{Ker}(A^T)$$

Preuve

$$1) \quad \text{Lgn}_i(A)^T = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \in \mathbb{R}^n \text{ par déf.}$$

$$\vec{z} \in \text{Lgn}(A)^\perp \Leftrightarrow \vec{z} \in \text{Ker}(A). \text{ En effet}$$

$$\vec{z} \in \text{Ker}(A) \Leftrightarrow A\vec{z} = \vec{0} = \begin{pmatrix} \text{Lgn}_1(A) \\ \vdots \\ \text{Lgn}_m(A) \end{pmatrix} \vec{z} = \vec{0}$$

$$\Leftrightarrow \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \vec{0}$$

$$\Leftrightarrow \begin{pmatrix} a_{11}z_1 + \dots + a_{1n}z_n \\ \vdots \\ a_{m1}z_1 + \dots + a_{mn}z_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Leftrightarrow a_{i1}z_1 + \dots + a_{in}z_n = 0 \quad \forall 1 \leq i \leq m$$

$$\Rightarrow \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = 0 \quad \forall 1 \leq i \leq m$$

$$\text{Donc } \vec{z} \in \text{Ker}(A) \Leftrightarrow \text{Lgn}_i(A)^T \cdot \vec{z} = 0 \quad \forall 1 \leq i \leq m$$

$$\Leftrightarrow \vec{z} \perp \text{span} \{ \text{Lgn}_1(A)^T, \dots, \text{Lgn}_m(A)^T \}$$

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

2) On applique 1) à A^T au lieu de A .

$$\text{Ker}(A^T) = \text{Lgn}(A^T)^\perp. \text{ Or on sait que } \text{Lgn}(A^T) = \text{Im}(A).$$

Exemple