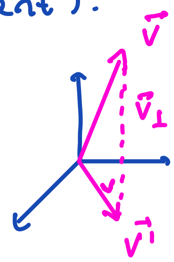
6.3 Familles orthogonales et projections orthogonales

But : utiliser les proj. orthogonales pour trouver $\vec{x}$ t.q. $A\vec{x} \cong \vec{b}$ (dans le cas d'un système inconsistant).

Projections:



$\vec{v}'$ est la proj. orthogonale de $\vec{v}$ sur Oxy

$\rightsquigarrow$ décomposition de $\vec{v}$ et $\underbrace{\vec{v}_{\parallel}}_{\vec{v}'}$ et $\vec{v}_\perp$

Définition 65 (famille orthogonale)**.**
Une famille de vecteurs $\{\vec{u}_1, \ldots, \vec{u}_p\}$ de $\mathbb{R}^n$ est dite *orthogonale* si

$$\vec{u}_i \cdot \vec{u}_j = 0 \qquad \forall \quad 1 \leq i,j \leq p \text{ avec } i \neq j$$

Exemples

1) $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ $\vec{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ $\vec{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

$\vec{v}_1 \cdot \vec{v}_2 = 0$
$\vec{v}_2 \cdot \vec{v}_3 = 0$
$\vec{v}_1 \cdot \vec{v}_3 = 1$

$\Rightarrow \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ n'est pas orthogonale

2) $\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ $\vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ $\vec{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

Alors $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ est orthogonale.

Remarque : c'est aussi une famille libre de $\mathbb{R}^3$.

Théorème 60. *Soit* $\{\vec{u}_1, \ldots, \vec{u}_p\}$ *une famille orthogonale de* $\mathbb{R}^n$ *formée de vecteurs* $\vec{u}_i$ *non nuls pour tout* $1 \leq i \leq p$. *Alors la famille*

$\{\vec{u}_1, \ldots, \vec{u}_p\}$ est lin. indépendante.

Remarque: On a forcément $p \leq n$!

Preuve Montrons que $\alpha_1 \vec{u}_1 + \ldots + \alpha_p \vec{u}_p = \vec{0}$ n'admet que la solution triviale. Soit $1 \leq i \leq p$.

$$0 = (\underbrace{\alpha_1 \vec{u}_1 + \ldots + \alpha_p \vec{u}_p}_{\vec{0}}) \cdot \vec{u}_i$$

$$= \alpha_1 \vec{u}_1 \cdot \vec{u}_i + \ldots + \alpha_i \vec{u}_i \cdot \vec{u}_i + \ldots + \alpha_p \vec{u}_p \cdot \vec{u}_i$$

$$= \alpha_i \|\vec{u}_i\|^2 \qquad \text{d'où} \qquad \alpha_i \|\vec{u}_i\|^2 = 0$$

donc soit $\alpha_i = 0$ soit $\|\vec{u}_i\|^2 = 0$

Comme $\vec{u}_i \neq \vec{0}$, $\|\vec{u}_i\|^2 > 0$ et $\alpha_i = 0$.

Finalement $\alpha_1 = \ldots = \alpha_p = 0 \Rightarrow \{\vec{u}_1, \ldots, \vec{u}_p\}$ est lin. indép.

Remarque $(\vec{u}_1, \dots, \vec{u}_p)$ est une base de $\mathrm{span}\{\vec{u}_1, \dots, \vec{u}_p\}$, mais ce n'est pas forcément une base de $\mathbb{R}^n$ ($p \le n$!)

Définition 66 (base orthogonale)**.**
Soit W un sous-espace vectoriel de $\mathbb{R}^n$. On appelle *base orthogonale* de W toute base qui est composée d'une famille orthogonale.

Exemples

1) p. 158 $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ est une base non orthogonale de $\mathbb{R}^3$

2) p. 158 $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est " " orthogonale de $\mathbb{R}^3$

3) La base canonique de $\mathbb{R}^n$ est une base orthogonale :

$$\vec{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^e \text{ ligne} \qquad \vec{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j^e \text{ ligne.}$$

Pour $i \ne j$, $\vec{e}_i \cdot \vec{e}_j = 0$. De plus $\vec{e}_i \cdot \vec{e}_i = 1$

$\|\vec{e}_i\|^2 = 1$, $\|\vec{e}_i\| = 1$

Application

Pour trouver les composantes d'un vecteur dans une base orthogonale, on peut utiliser soit les techniques de changement de base ou de résolution de système (échelonnement d'une matrice augmentée) classiques, soit utiliser les propriétés du produit scalaire :

Soit $B = \left(\underset{\vec{u_1}}{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}, \underset{\vec{u_2}}{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}, \underset{\vec{u_3}}{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}} \right)$ la base orthogonale

vue ci-dessus et $\vec{v} = \begin{pmatrix} 6 \\ -2 \\ 11 \end{pmatrix} = [\vec{v}]_{\mathcal{E}}$

On cherche $[\vec{v}]_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$, donc

$\vec{v} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \alpha_3 \vec{u_3}$.

On peut soit résoudre $\left(\begin{array}{ccc|c} 1 & 0 & 1 & 6 \\ 1 & 0 & -1 & -2 \\ 0 & 1 & 0 & 11 \end{array} \right)$, soit trouver

$P_{\mathcal{E}B} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$, puis $P_{\mathcal{E}B}^{-1} = P_{B\mathcal{E}}$ et calculer

$[\vec{v}]_B = P_{B\mathcal{E}} [\vec{v}]_{\mathcal{E}}$. On trouve $\alpha_1 = 2$, $\alpha_2 = 11$, $\alpha_3 = 4$

Méthode avec l'orthogonalité :

$$\vec{v} \cdot \vec{u_i} = (\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \alpha_3 \vec{u_3}) \cdot \vec{u_i} \qquad 1 \leq i \leq 3$$

$$= \alpha_i \|\vec{u_i}\|^2$$

$$\Rightarrow \quad \alpha_i = \frac{\vec{v} \cdot \vec{u_i}}{\|\vec{u_i}\|^2} = \frac{\vec{v} \cdot \vec{u_i}}{\vec{u_i} \cdot \vec{u_i}}$$

$$\alpha_1 = \frac{\vec{v} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}} = \frac{\begin{pmatrix} 6 \\ -2 \\ 11 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}} = \frac{4}{2} = 2$$

$$\alpha_2 = \frac{11}{1} = 11 \qquad \alpha_3 = \frac{8}{2} = 4$$

Théorème 61. *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $(\vec{u}_1, \ldots, \vec{u}_p)$ une base orthogonale de W. Alors* $\forall \vec{w} \in W$, on a

$\vec{w} = \alpha_1 \vec{u}_1 + \ldots + \alpha_p \vec{u}_p$ avec

$$\alpha_i = \frac{\vec{w} \cdot \vec{u}_i}{\vec{u}_i \cdot \vec{u}_i} \quad \forall\ 1 \leq i \leq p$$

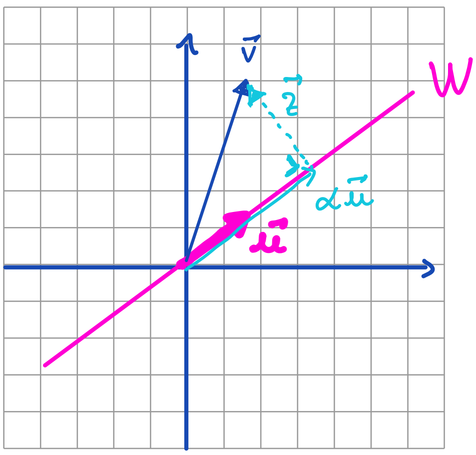
Preuve Soit $1 \leq i \leq p$

$$\vec{w} \cdot \vec{u}_i = \left(\sum_{k=1}^{p} \alpha_k \vec{u}_k \right) \cdot \vec{u}_i = \sum_{k=1}^{p} \alpha_k \vec{u}_k \cdot \vec{u}_i$$

$$= \alpha_i \vec{u}_i \cdot \vec{u}_i \Rightarrow \alpha_i = \frac{\vec{w} \cdot \vec{u}_i}{\vec{u}_i \cdot \vec{u}_i}$$

Conclusion: Avec une base orthogonale, on a une méthode alternative à l'échelonnement de matrices.

Projections orthogonales



$\vec{u} \in \mathbb{R}^2$, $\vec{u} \neq \vec{0}$

$W = \text{Span}\{\vec{u}\}$ et $\vec{v} \in \mathbb{R}^2$

on cherche une décomposition

$$\vec{v} = \alpha \vec{u} + \vec{z}$$

avec $\vec{z} \perp \vec{u}$.

Si $\vec{v} \in W$, alors $\vec{z} = \vec{0}$.

On cherche à déterminer $\alpha \in \mathbb{R}$ et $\vec{z} \in W^{\perp}$.

$\vec{z} = \vec{v} - \alpha \vec{u}$. Comme $\vec{u} \perp \vec{z}$, on a

$$\vec{z} \perp \vec{u} \Leftrightarrow \vec{z} \cdot \vec{u} = 0 \Leftrightarrow (\vec{v} - \alpha \vec{u}) \cdot \vec{u} = 0$$
$$\Leftrightarrow \vec{v} \cdot \vec{u} - \alpha \vec{u} \cdot \vec{u} = 0$$

Donc $\alpha = \dfrac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \in \mathbb{R}$.

Pour finir : $$\alpha \vec{u} = \underbrace{\frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}}_{\in \mathbb{R}} \vec{u} \in W$$

On appelle $\alpha\vec{u}$ la proj. orth. de $\vec{v}$ sur W.

Notation: $$\text{proj}_W(\vec{v}) = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$$

$\vec{z} = \vec{v} - \text{proj}_W(\vec{v})$ est la composante de $\vec{v}$ orthogonale à $\vec{u}$.

$\vec{z} \in W^{\perp}$.

Exemple $\vec{v} = \begin{pmatrix}-1\\2\end{pmatrix}$ $W = \text{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$

$\vec{u} = \begin{pmatrix}1\\1\end{pmatrix}$

$$\text{proj}_W(\vec{v}) = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u} = \frac{\begin{pmatrix}-1\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}}{\begin{pmatrix}1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1/2\\1/2\end{pmatrix}$$

$$\vec{z} = \vec{v} - \text{proj}_W(\vec{v}) = \begin{pmatrix}-1\\2\end{pmatrix} - \begin{pmatrix}1/2\\1/2\end{pmatrix} = \begin{pmatrix}-3/2\\3/2\end{pmatrix} \in W^{\perp}.$$

Définition 67 (famille orthonormale)**.**
On dira qu'une famille $\{\vec{u}_1, \ldots, \vec{u}_p\}$ de vecteurs de $\mathbb{R}^n$ est *orthonormale* si

1. la famille est orthogonale,
2. tous les vecteurs de la famille sont des vecteurs unitaires. $\|\vec{u}_i\| = 1 \ \forall 1 \leq i \leq p$

On dira alors que $(\vec{u}_1, \ldots, \vec{u}_p)$ est une *base orthonormale* de

Exemples

1) $(\vec{e}_1, \ldots, \vec{e}_n)$ est une base orthonormale de $\mathbb{R}^n$.

2) $\vec{v}_1 = \begin{pmatrix}3\\2\\2\end{pmatrix}$ $\vec{v}_2 = \begin{pmatrix}-2\\3\\0\end{pmatrix}$ $\vec{v}_3 = \begin{pmatrix}-6\\-4\\13\end{pmatrix}$

$\left.\begin{aligned}\vec{v}_1\cdot\vec{v}_2 &= -6+6 = 0\\ \vec{v}_1\cdot\vec{v}_3 &= -18-8+26 = 0\\ \vec{v}_2\cdot\vec{v}_3 &= 12-12+0 = 0\end{aligned}\right\}$ fam. orthogonale

On "normalise"

$\|\vec{v}_1\| = \sqrt{3^2+2^2+2^2} = \sqrt{17}$ $\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{17}}\vec{v}_1$ etc.

$\|\vec{v}_2\| = \sqrt{13}$

$\|\vec{v}_3\| = \sqrt{221}$ d'où une base $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ orthonormale.

Théorème 62. *Soit $U = (\vec{u}_1 \cdots \vec{u}_n) \in M_{m\times n}(\mathbb{R})$ une matrice. Alors les colonnes de U sont orthonormales si et seulement si*

$$\underset{(n\times m)}{U^T}\ \underset{(m\times n)}{U} = I_n$$

Remarque: $U^T U \neq U U^T$ et en générale $UU^T = I_m$

Preuve

$$(U^T U)_{ii} = \sum_{j=1}^{m} (U^T)_{ij} (U)_{ji} = \sum_{j=1}^{m} U_{ji} U_{ji}$$

$$= \sum_{j=1}^{m} u_{ji}^2 = \|\vec{u}_i\|^2$$

$$\underset{(i\neq k)}{(U^T U)_{ik}} = \sum_{j=1}^{m} (U^T)_{ij} (U)_{jk} = \sum_{j=1}^{m} u_{ji} u_{jk}$$

$$= \vec{u}_i \cdot \vec{u}_k$$

Exemple $U = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}_{3\times 2}$

$$U^T U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

Donc $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$ est une base orthonormale du Span $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$.

$UU^T \neq I_3$.

Théorème 63. *Soit $U \in M_{m\times n}(\mathbb{R})$ une matrice dont les colonnes sont orthonormales. Soient $\vec{v}, \vec{w} \in \mathbb{R}^n$. Alors*

1. $\|U\vec{v}\| = \|\vec{v}\|$
2. $(U\vec{v})\cdot(U\vec{w}) = \vec{v}\cdot\vec{w}$
3. $U\vec{v} \cdot U\vec{w} = 0 \Leftrightarrow \vec{v}\cdot\vec{w} = 0$

Preuve: Il suffit de prouver 2.

2. $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ et $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, $U = (\vec{u}_1 \dots \vec{u}_n)$

$$U\vec{v}\cdot U\vec{w} = (\vec{u}_1 \dots \vec{u}_n)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot (\vec{u}_1 \dots \vec{u}_n)\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$= \left(\sum_{i=1}^{n} v_i \vec{u}_i\right) \cdot \left(\sum_{j=1}^{n} w_j \vec{u}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \underbrace{(v_i w_j)}_{\in \mathbb{R}} \underbrace{\vec{u}_i \cdot \vec{u}_j}$$

$$= \sum_{i=1}^{n} (v_i w_i) \underbrace{\vec{u}_i \cdot \vec{u}_i}_{=1} = \sum_{i=1}^{n} v_i w_i = \vec{v} \cdot \vec{w}$$

Remarques

1. Les propriétés ci-dessus montrent que l'application linéaire $\vec{v} \mapsto U\vec{v}$ conserve les longueurs et l'orthogonalité.
2. On verra le lien entre le produit scalaire de deux vecteurs et l'angle formé par ces derniers dans une série d'exercices.

Généralisation de la projection à $\mathbb{R}^n$

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ ainsi qu'un vecteur $\vec{v} \notin W$. Alors la projection de $\vec{v}$ sur W, notée $\text{proj}_W(\vec{v})$ est

1. l'unique vecteur de W tel que
2. l'unique vecteur de W qui

Théorème 64. *Soit W un sous-espace vectoriel de $\mathbb{R}^n$. Alors tout vecteur $\vec{v} \in \mathbb{R}^n$ peut s'écrire de façon unique comme la somme de deux vecteurs*

Plus précisément, si $W = \text{span}\{\vec{u}_1, \ldots, \vec{u}_p\}$ où $(\vec{u}_1, \ldots, \vec{u}_p)$ est une base orthogonale de W, on a

Interprétation géométrique

Remarques

1. Le vecteur $\text{proj}_W(\vec{v})$ ne dépend pas du choix de la base orthogonale de W.
2. Si la base $(\vec{u_1}, \ldots, \vec{u_p})$ est orthonormale, alors

Exemple

Théorème 65 (de la meilleure approximation)**.** *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $\vec{v} \in \mathbb{R}^n$. Soit* $\text{proj}_W(\vec{v})$ *la projection orthogonale de $\vec{v}$ sur W. Alors* $\text{proj}_W(\vec{v})$ *est la meilleure approximation de $\vec{v}$ dans W, autrement dit, on a*

Illustration

Théorème 66. *Soit W un sous-espace vectoriel de $\mathbb{R}^n$ et supposons que $(\vec{u}_1, \ldots, \vec{u}_p)$ est une base orthonormale de W. Alors pour tout $\vec{v} \in \mathbb{R}^n$, on a*

Preuve

Remarques

Définition 68 (matrice orthogonale)**.**
Une matrice $U \in M_{n\times n}(\mathbb{R})$ inversible telle que

Remarques

6.4 Procédé de Gram-Schmidt

Le procédé de Gram-Schmidt permet d'obtenir une base orthogonale (ou même orthonormée) pour tout sous-espace vectoriel W de $\mathbb{R}^n$. Une telle base est nécessaire pour pouvoir calculer la projection orthogonale $\mathrm{proj}_W(\vec{v})$ d'un vecteur $\vec{v} \in \mathbb{R}^n$ sur W.

Exemple d'orthogonalisation d'une base

Théorème 67 (de Gram-Schmidt). *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $(\vec{v}_1, \ldots, \vec{v}_p)$ une base de W. Alors on peut construire une base orthogonale de W selon le procédé suivant :*

Exemple

6.5 Factorisation QR

Si les colonnes d'une matrice $A \in M_{m\times n}(\mathbb{R})$ sont linéairement indépendantes, on peut leur appliquer le procédé de Gram-Schmidt. Cela revient à trouver une factorisation de la matrice A :

Théorème 68 (Factorisation QR)**.** *Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice dont les colonnes sont linéairement indépendantes. Alors il existe une factorisation $A = QR$ où*

1.

2.

Remarques

Exemple

6.6 Méthode des moindres carrés

Exemple

Définition 69 (solution au sens des moindres carrés).
Soient $A \in M_{m\times n}(\mathbb{R})$ une matrice et $\vec{b} \in \mathbb{R}^m$. On appelle *solution au sens des moindres carrés* de $A\vec{x} = \vec{b}$ le vecteur

Démarche

Suite de l'exemple

L’équation normale

Théorème 69. *L'ensemble des solutions au sens des moindres carrés de $A\vec{x} = \vec{b}$ est égal à l'ensemble des solutions de*

On appelle cette équation *l'équation normale*.

Remarque

Exemples

Théorème 70. *Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice. Alors les propriétés suivantes sont équivalentes*

1.

2.

3.

Lien avec la factorisation QR

Théorème 71. *Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice dont les colonnes sont linéairement indépendantes. Considérons la décomposition $A = QR$. Alors pour tout $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ admet une unique solution au sens des moindres carrés donnée par*

6.7 Droite de régression

Supposons qu'on dispose d'un certain nombre de données expérimentales du type $(x_1, y_1), \ldots, (x_n, y_n)$ et qu'on cherche à établir une formule permettant de prédire les valeurs d'une variable en fonction de l'autre variable. Ce type de situation est étudié en analyse statistique. On résout souvent le problème à l'aide d'une méthode qui fait appel aux moindres carrés.

On s'intéresse ici à la relation la plus élémentaire entre deux variables x et y, relation du premier degré du type $y = ax + b$. Lorsqu'on dispose d'un nuage de points $(x_1, y_1), \ldots, (x_n, y_n)$, on cherche à déterminer les paramètres a et b qui rendent la droite $y = ax + b$ aussi proche que possible des points expérimentaux.

Exemple

Remarques