Démonstration. Supposons que $c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2} + \ldots + c_k\overrightarrow{u_k} = \overrightarrow{0}$.

En prenant le produit scalaire avec $\overrightarrow{u_i}$:

$$0 = \overrightarrow{u_i} \cdot (c_1\overrightarrow{u_1} + \ldots + c_k\overrightarrow{u_k}) = c_i(\overrightarrow{u_i} \cdot \overrightarrow{u_i}) = c_i\|\overrightarrow{u_i}\|^2$$

Comme $\overrightarrow{u_i} \neq \overrightarrow{0}$, on a $\|\overrightarrow{u_i}\|^2 > 0$, donc $c_i = 0$ pour tout i. □

Corollaire 6.22

Si $\{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_k}\}$ est un ensemble orthogonal **de vecteurs non nuls** dans $\mathbb{R}^n$, alors c'est une base de Vect $(\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_k})$.

Définition 6.23 *Base orthogonale et orthonormée*

Si W est un sous-espace vectoriel de $\mathbb{R}^n$, de dimension k et $\mathcal{B} = (\overrightarrow{u_1}, \ldots, \overrightarrow{u_k})$ une base de W.
On dit que cette base est :
- *orthogonale* si $\overrightarrow{u_i} \cdot \overrightarrow{u_j} = 0$ pour $i \neq j$
- *orthonormée* si elle est orthogonale et $\|\overrightarrow{u_i}\| = 1$ pour tout i

Exemples. 1. La base canonique est une base orthonormée de $\mathbb{R}^n$.

2. $\mathcal{B} = \left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)$ est une base orthogonale de $\mathbb{R}^3$. Cependant, cette base n'est pas orthonormée car $\left\|\begin{bmatrix}1\\1\\0\end{bmatrix}\right\| = \sqrt{2} \neq 1$.

3. $\mathcal{B}' = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)$ est une base orthonormée de $\mathbb{R}^3$.

Théorème 6.24 *Coordonnées dans une base orthogonale*

Soit $\mathcal{B} = (\overrightarrow{u_1}, \ldots, \overrightarrow{u_n})$ une base orthogonale de $\mathbb{R}^n$.
Pour tout vecteur $\overrightarrow{w} \in \mathbb{R}^n$, les coordonnées de $\overrightarrow{w}$ dans la base $\mathcal{B}$ sont :

$$\overrightarrow{w} = \frac{\overrightarrow{w} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}}\overrightarrow{u_1} + \frac{\overrightarrow{w} \cdot \overrightarrow{u_2}}{\overrightarrow{u_2} \cdot \overrightarrow{u_2}}\overrightarrow{u_2} + \cdots + \frac{\overrightarrow{w} \cdot \overrightarrow{u_n}}{\overrightarrow{u_n} \cdot \overrightarrow{u_n}}\overrightarrow{u_n}$$

Si la base est orthonormée, la formule se simplifie :

$$\overrightarrow{w} = (\overrightarrow{w} \cdot \overrightarrow{u_1})\overrightarrow{u_1} + (\overrightarrow{w} \cdot \overrightarrow{u_2})\overrightarrow{u_2} + \cdots + (\overrightarrow{w} \cdot \overrightarrow{u_n})\overrightarrow{u_n}$$

Démonstration. Puisque $\mathcal{B}$ est une base de $\mathbb{R}^n$, tout vecteur $\overrightarrow{w} \in \mathbb{R}^n$ s'écrit de manière unique comme combinaison linéaire des vecteurs de la base :

$$\overrightarrow{w} = c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2} + \cdots + c_n\overrightarrow{u_n}$$

Pour trouver le coefficient c_i, calculons le produit scalaire de $\overrightarrow{w}$ avec $\overrightarrow{u_i}$:

$$\begin{aligned}\overrightarrow{w} \cdot \overrightarrow{u_i} &= (c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2} + \cdots + c_n\overrightarrow{u_n}) \cdot \overrightarrow{u_i} \\ &= c_1(\overrightarrow{u_1} \cdot \overrightarrow{u_i}) + c_2(\overrightarrow{u_2} \cdot \overrightarrow{u_i}) + \cdots + c_n(\overrightarrow{u_n} \cdot \overrightarrow{u_i})\end{aligned}$$

Puisque la base est orthogonale, on a $\overrightarrow{u_j} \cdot \overrightarrow{u_i} = 0$ pour tout $j \neq i$. Tous les termes s'annulent sauf celui correspondant à $j = i$:

$$\overrightarrow{w} \cdot \overrightarrow{u_i} = c_i(\overrightarrow{u_i} \cdot \overrightarrow{u_i})$$

D'où

$$c_i = \frac{\overrightarrow{w} \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}$$

Si la base est orthonormée, alors $\overrightarrow{u_i} \cdot \overrightarrow{u_i} = \|\overrightarrow{u_i}\|^2 = 1$ pour tout i, donc la formule se simplifie en :

$$c_i = \overrightarrow{w} \cdot \overrightarrow{u_i}$$

□

Exemple. Soit la base orthogonale $\mathcal{B} = (\overrightarrow{u_1}, \overrightarrow{u_2}) = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$ de $\mathbb{R}^2$.

Pour $\overrightarrow{w} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$, calculons ses coordonnées :

$$\begin{aligned}\overrightarrow{w} \cdot \overrightarrow{u_1} &= 5 \cdot 1 + 3 \cdot 1 = 8 \\ \overrightarrow{u_1} \cdot \overrightarrow{u_1} &= 1^2 + 1^2 = 2 \\ \overrightarrow{w} \cdot \overrightarrow{u_2} &= 5 \cdot 1 + 3 \cdot -1 = 2 \\ \overrightarrow{u_2} \cdot \overrightarrow{u_2} &= 1^2 + (-1)^2 = 2\end{aligned}$$

Donc $\overrightarrow{w} = \frac{8}{2}\overrightarrow{u_1} + \frac{2}{2}\overrightarrow{u_2} = 4\overrightarrow{u_1} + \overrightarrow{u_2}$. Donc $[\overrightarrow{w}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

6.5 Projections orthogonales

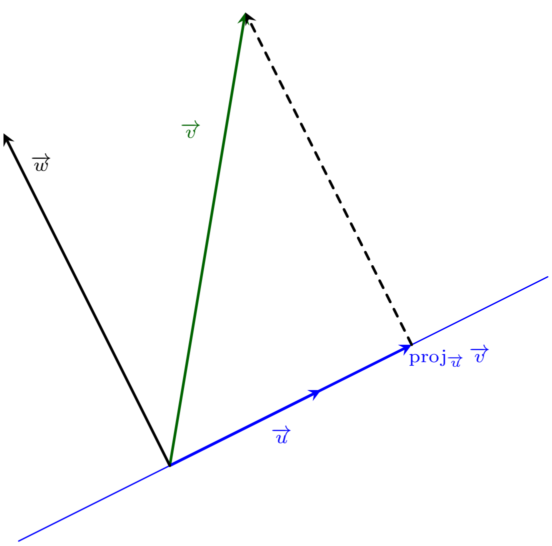
Définition 6.25 *Projection orthogonale sur un vecteur*

Soit $\overrightarrow{u}$ un vecteur non nul de $\mathbb{R}^n$. La *projection orthogonale* d'un vecteur $\overrightarrow{v}$ sur $W = \text{Vect}(\overrightarrow{u})$ est :

$$\text{proj}_W \overrightarrow{v} = \text{proj}_{\overrightarrow{u}} \overrightarrow{v} = \frac{\overrightarrow{v} \cdot \overrightarrow{u}}{\overrightarrow{u} \cdot \overrightarrow{u}}\overrightarrow{u}$$

Remarques 6.6.0.26.

1. *La projection* $\mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}$ *est le vecteur de la droite engendrée par* $\overrightarrow{u}$*, qui est le plus proche de* $\overrightarrow{v}$*.*
2. *Le vecteur* $\overrightarrow{w}=\overrightarrow{v}-\mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}$ *est orthogonal à* $\overrightarrow{u}$*. Donc* $\overrightarrow{v}=\mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}+\overrightarrow{w}$ *est la somme d'un vecteur colinéaire à* $\overrightarrow{u}$ *et d'un vecteur orthogonal à* $\overrightarrow{u}$*.*



Exemples. Projetons $\overrightarrow{v}=\begin{bmatrix}3\\4\end{bmatrix}$ sur $\overrightarrow{u}=\begin{bmatrix}1\\2\end{bmatrix}$:

$$\begin{aligned}\overrightarrow{v}\cdot\overrightarrow{u}&=3\cdot1+4\cdot2=11\\ \overrightarrow{u}\cdot\overrightarrow{u}&=1^2+2^2=5\\ \mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}&=\frac{11}{5}\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}\frac{11}{5}\\ \frac{22}{5}\end{bmatrix}\end{aligned}$$

Vérifions que $\overrightarrow{w}=\overrightarrow{v}-\mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}=\begin{bmatrix}3\\4\end{bmatrix}-\begin{bmatrix}\frac{11}{5}\\ \frac{22}{5}\end{bmatrix}=\begin{bmatrix}\frac{4}{5}\\ \frac{-2}{5}\end{bmatrix}$ est orthogonal à $\overrightarrow{u}$:

$$\overrightarrow{w}\cdot\overrightarrow{u}=\frac{4}{5}\cdot1+\frac{-2}{5}\cdot2=\frac{4-4}{5}=0$$

Théorème 6.27

Soit $\overrightarrow{u}\in\mathbb{R}^n$ non nul, l'application $\begin{aligned}\mathrm{proj}_{\overrightarrow{u}}:\ \mathbb{R}^n&\longrightarrow\mathbb{R}^n\\ \overrightarrow{v}&\longmapsto\mathrm{proj}_{\overrightarrow{u}}\overrightarrow{v}\end{aligned}$ est linéaire.

Démonstration. La linéarité se montre en utilisant la linéarité du produit scalaire. □

Remarque 6.6.0.28. $\text{proj}_{\vec{u}}\,\vec{v} = \vec{0} \Leftrightarrow \vec{u}\cdot\vec{v} = 0 \Leftrightarrow \vec{v} \in \vec{u}^{\perp}$ *donc* $\text{Ker}(\text{proj}_{\vec{u}}) = \vec{u}^{\perp}$

Par ailleurs, $\text{proj}_{\vec{u}}\,\vec{v}$ *est colinéaire à* $\vec{u}$ *pour tout* $\vec{v}$ *donc* $\text{Im}(\text{proj}_{\vec{u}}) = \text{Vect}\,(\vec{u})$*.*

Enfin, on a

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

car $\vec{u}^T\vec{v}$ *est un nombre réel.*

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

Donc la matrice canoniquement associée à cette application linéaire est

$$A_{\text{proj}_{\vec{u}}} = \frac{1}{\vec{u}\cdot\vec{u}}\vec{u}\,\vec{u}^T = \frac{1}{\vec{u}^T\vec{u}}\vec{u}\,\vec{u}^T.$$

Exemples. Soit $\vec{u} = \begin{bmatrix}1\\2\end{bmatrix}$ dans $\mathbb{R}^2$.

On a $\vec{u}\cdot\vec{u} = 1^2 + 2^2 = 5$, donc

$$\begin{aligned} A_{\text{proj}_{\vec{u}}} &= \frac{1}{\vec{u}\cdot\vec{u}}\vec{u}\,\vec{u}^T \\ &= \frac{1}{5}\begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}1 & 2\end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix}1 & 2\\2 & 4\end{bmatrix} \\ &= \begin{bmatrix}\frac{1}{5} & \frac{2}{5}\\ \frac{2}{5} & \frac{4}{5}\end{bmatrix} \end{aligned}$$

Image : Les colonnes de la matrice sont colinéaires, donc

$$\text{Im}(\text{proj}_{\vec{u}}) = \text{Vect}\left(\begin{bmatrix}\frac{1}{5}\\ \frac{2}{5}\end{bmatrix}, \begin{bmatrix}\frac{2}{5}\\ \frac{4}{5}\end{bmatrix}\right) = \text{Vect}\,(\vec{u}).$$

Par exemple, si $\vec{v}_1 = \begin{bmatrix}3\\1\end{bmatrix}$:

$$\text{proj}_{\vec{u}}\,\vec{v}_1 = \frac{3+2}{5}\begin{bmatrix}1\\2\end{bmatrix} = \frac{5}{5}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix} = \vec{u} \in \text{Vect}\,(\vec{u})$$

Même calcul avec la matrice :

$$A_{\text{proj}_{\vec{u}}}\vec{v}_1 = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3+2}{5} \\ \frac{6+4}{5} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

De plus, si $\vec{v} = \lambda\vec{u} = \begin{bmatrix} \lambda \\ 2\lambda \end{bmatrix}$ appartient déjà à Vect $(\vec{u})$, alors

$$\text{proj}_{\vec{u}}\,\vec{v} = \frac{\lambda + 4\lambda}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \lambda\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \vec{v}$$

Donc la projection laisse donc invariants tous les vecteurs de Vect $(\vec{u})$.

Noyau : $\text{Ker}(\text{proj}_{\vec{u}}) = \vec{u}^{\perp}$**.**

Un vecteur $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ appartient à $\vec{u}^{\perp}$ si et seulement si :

$$\vec{v}\cdot\vec{u} = x + 2y = 0 \quad \Leftrightarrow \quad x = -2y$$

Donc $\vec{u}^{\perp} = \text{Vect}\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right)$.

Vérifions avec $\vec{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \vec{u}^{\perp}$:

$$\text{proj}_{\vec{u}}\,\vec{v}_2 = \frac{-2+2}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{0}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Donc $\vec{v}_2 \in \text{Ker}(\text{proj}_{\vec{u}})$.

Avec la matrice :

$$A_{\text{proj}_{\vec{u}}}\vec{v}_2 = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-2+2}{5} \\ \frac{-4+4}{5} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Plus généralement, pour tout $\vec{w} = t\begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \vec{u}^{\perp}$ avec $t \in \mathbb{R}$:

$$\text{proj}_{\vec{u}}\,\vec{w} = \frac{-2t+2t}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Décomposition orthogonale de $\mathbb{R}^2$.

Tout vecteur $\vec{v} \in \mathbb{R}^2$ se décompose de manière unique en :

$$\vec{v} = \text{proj}_{\vec{u}}\,\vec{v} + (\vec{v} - \text{proj}_{\vec{u}}\,\vec{v})$$

où $\text{proj}_{\vec{u}}\,\vec{v} \in \text{Vect}\,(\vec{u})$ et $\vec{v} - \text{proj}_{\vec{u}}\,\vec{v} \in \vec{u}^{\perp}$.

Par exemple, avec $\vec{v}_3 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$:

$$\text{proj}_{\vec{u}}\,\vec{v}_3 = \frac{5+6}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{11}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{11}{5} \\ \frac{22}{5} \end{bmatrix}$$

$$\vec{v}_3 - \text{proj}_{\vec{u}}\,\vec{v}_3 = \begin{bmatrix} 5 \\ 3 \end{bmatrix} - \begin{bmatrix} \frac{11}{5} \\ \frac{22}{5} \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ \frac{-7}{5} \end{bmatrix} = \frac{7}{5}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

On vérifie que $\begin{bmatrix} 2 \\ -1 \end{bmatrix} \in \vec{u}^{\perp}$ car $2 + 2(-1) = 0$.

Ainsi :

$$\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \underbrace{\frac{11}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\in \text{Vect}(\vec{u})} + \underbrace{\frac{7}{5}\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\in \vec{u}^{\perp}}$$

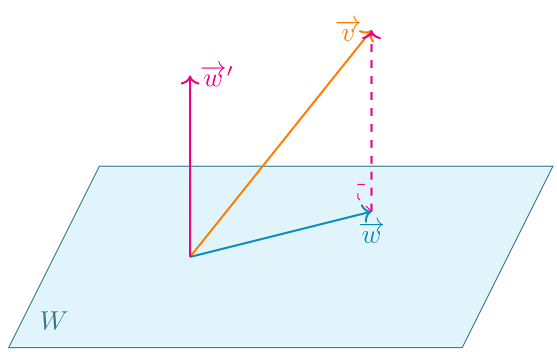
Remarque importante 6.29

Attention à ne pas confondre $\vec{u}^T\vec{u}$ qui est un réel, et $\vec{u}\,\vec{u}^T$ qui est une matrice de taille $n \times n$!

Soit maintenant $W \subset \mathbb{R}^n$ un sous-espace vectoriel de dimension $k \geq 2$. On aimerait écrire un vecteur arbitraire $\vec{v} \in \mathbb{R}^n$ comme la somme d'un vecteur $\vec{w} \in W$ et d'un vecteur $\vec{w}' \in W^{\perp}$:

$$\vec{v} = \vec{w} + \vec{w}'$$

Géométriquement :



Définition 6.30 *Projection orthogonale sur un sous-espace*

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$ et soit $(\overrightarrow{u_1}, \ldots, \overrightarrow{u_k})$ une base orthogonale de W.
La *projection orthogonale* de $\overrightarrow{v}$ sur W est :

$$\operatorname{proj}_W \overrightarrow{v} = \frac{\overrightarrow{v} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}} \overrightarrow{u_1} + \ldots + \frac{\overrightarrow{v} \cdot \overrightarrow{u_k}}{\overrightarrow{u_k} \cdot \overrightarrow{u_k}} \overrightarrow{u_k}.$$

Exemples. Considérons $\mathbb{R}^3$ et le sous-espace

$$W = \operatorname{Vect}(\overrightarrow{u}_1, \overrightarrow{u}_2) \quad \text{où} \quad \overrightarrow{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

On vérifie que

$$\overrightarrow{u}_1 \cdot \overrightarrow{u}_2 = 1 \cdot (-1) + 1 \cdot 1 + 0 \cdot 0 = -1 + 1 + 0 = 0,$$

donc $(\overrightarrow{u}_1, \overrightarrow{u}_2)$ est une base orthogonale de W.
Prenons

$$\overrightarrow{v} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

Nous allons calculer $\operatorname{proj}_W \overrightarrow{v}$ avec la définition 6.30.
On commence par les produits scalaires :

$$\overrightarrow{u}_1 \cdot \overrightarrow{u}_1 = 1^2 + 1^2 + 0^2 = 2, \quad \overrightarrow{u}_2 \cdot \overrightarrow{u}_2 = (-1)^2 + 1^2 + 0^2 = 2,$$

$$\overrightarrow{v} \cdot \overrightarrow{u}_1 = 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 0 = 5, \quad \overrightarrow{v} \cdot \overrightarrow{u}_2 = 2 \cdot (-1) + 3 \cdot 1 + 4 \cdot 0 = 1.$$

Donc

$$\begin{aligned}
\operatorname{proj}_W \overrightarrow{v} &= \frac{\overrightarrow{v} \cdot \overrightarrow{u}_1}{\overrightarrow{u}_1 \cdot \overrightarrow{u}_1} \overrightarrow{u}_1 + \frac{\overrightarrow{v} \cdot \overrightarrow{u}_2}{\overrightarrow{u}_2 \cdot \overrightarrow{u}_2} \overrightarrow{u}_2 \\
&= \frac{5}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{5}{2} \\ \frac{5}{2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{-1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}.
\end{aligned}$$

Remarquons que le vecteur

$$\vec{v} - \text{proj}_W \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

est orthogonal à W.

Remarque 6.6.0.31. *Si $(\vec{u_1}, \ldots, \vec{u_k})$ est orthonormée, alors*

$$\text{proj}_W \vec{v} = (\vec{v} \cdot \vec{u_1})\vec{u_1} + \ldots + (\vec{v} \cdot \vec{u_k})\vec{u_k}.$$

Théorème 6.32 *Théorème de la projection orthogonale*

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$. Tout vecteur $\vec{v} \in \mathbb{R}^n$ s'écrit de manière unique sous la forme :

$$\vec{v} = \vec{w} + \vec{w}' \quad \text{où } \vec{w} \in W \text{ et } \vec{w}' \in W^\perp$$

De plus, $\vec{w} = \text{proj}_W \vec{v}$.

Démonstration. Soit $(\vec{u_1}, \ldots, \vec{u_k})$ est une base orthogonale de W.

Par définition de la projection, le vecteur $\vec{w} = \text{proj}_W \vec{v}$ appartient à W car il s'écrit comme combinaison linéaire des vecteurs $\vec{u_1}, \ldots, \vec{u_k}$.

À voir : $\vec{w}' = \vec{v} - \vec{w} \in W^\perp$.

Montrons que $\vec{w}' \cdot \vec{u_j} = 0$ pour $j \in \{1, \ldots, k\}$.

On a :

$$\begin{aligned}
\vec{w}' \cdot \vec{u_j} &= (\vec{v} - \vec{w}) \cdot \vec{u_j} \\
&= \vec{v} \cdot \vec{u_j} - \vec{w} \cdot \vec{u_j} \\
&= \vec{v} \cdot \vec{u_j} - \left(\frac{\vec{v} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}} \vec{u_1} + \frac{\vec{v} \cdot \vec{u_2}}{\vec{u_2} \cdot \vec{u_2}} \vec{u_2} + \ldots + \frac{\vec{v} \cdot \vec{u_k}}{\vec{u_k} \cdot \vec{u_k}} \vec{u_k} \right) \cdot \vec{u_j} \\
&= \vec{v} \cdot \vec{u_j} - \frac{\vec{v} \cdot \vec{u_j}}{\vec{u_j} \cdot \vec{u_j}} (\vec{u_j} \cdot \vec{u_j}) - \sum_{i \neq j} \frac{\vec{v} \cdot \vec{u_i}}{\vec{u_i} \cdot \vec{u_i}} \underbrace{(\vec{u_i} \cdot \vec{u_j})}_{=0} \\
&= \vec{v} \cdot \vec{u_j} - \vec{v} \cdot \vec{u_j} \\
&= 0
\end{aligned}$$

Il nous reste à montrer que l'écriture est unique. Supposons qu'on a deux décompositions :

$$\vec{v} = \vec{w}_1 + \vec{w}'_1 = \vec{w}_2 + \vec{w}'_2 \quad \text{avec } \vec{w}_1, \vec{w}_2 \in W \text{ et } \vec{w}'_1, \vec{w}'_2 \in W^\perp$$

Cela implique que $\vec{w}_2 - \vec{w}_1 = \vec{w}'_1 - \vec{w}'_2$

Ainsi, le vecteur $\vec{u} = \vec{w}_2 - \vec{w}_1 = \vec{w}'_1 - \vec{w}'_2 \in W \cap W^\perp$

Donc $\overrightarrow{u} = \overrightarrow{0}$.

Or, on a les équivalences

$$\overrightarrow{u} = \overrightarrow{0} \Leftrightarrow \begin{cases} \overrightarrow{w}_2 - \overrightarrow{w}_1 = \overrightarrow{0} \\ \overrightarrow{w}'_1 - \overrightarrow{w}'_2 = \overrightarrow{0} \end{cases} \Leftrightarrow \begin{cases} \overrightarrow{w}_2 = \overrightarrow{w}_1 \\ \overrightarrow{w}'_2 = \overrightarrow{w}'_1 \end{cases}$$

□

Corollaire 6.33 $\mathrm{Im}(\mathrm{proj}_W) = W$

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

$$\mathrm{Im}(\mathrm{proj}_W) = W.$$

Démonstration.
— $W \subset \mathrm{Im}(\mathrm{proj}_W)$
Pour tout $\overrightarrow{w} \in W$, on a $\mathrm{proj}_W \overrightarrow{w} = \overrightarrow{w}$ (car $\overrightarrow{w} = \overrightarrow{w} + \overrightarrow{0}$ avec $\overrightarrow{0} \in W^\perp$), donc $\overrightarrow{w} \in \mathrm{Im}(\mathrm{proj}_W)$.
— $\mathrm{Im}(\mathrm{proj}_W) \subset W$
Réciproquement, $\mathrm{proj}_W \overrightarrow{v} \in W$ pour tout $\overrightarrow{v} \in \mathbb{R}^n$, donc $\mathrm{Im}(\mathrm{proj}_W) \subset W$.

□

Corollaire 6.34 $(W^\perp)^\perp = W$

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

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

Démonstration. Nous avons déjà montré $W \subset (W^\perp)^\perp$.
Il reste à montrer $(W^\perp)^\perp \subset W$.
Soit donc $\overrightarrow{v} \in (W^\perp)^\perp \subset \mathbb{R}^n$.
Le théorème de la projection orthogonale nous donne :

$$\overrightarrow{v} = \overrightarrow{w} + \overrightarrow{w}', \quad \text{avec } \overrightarrow{w} \in W \text{ et } \overrightarrow{w}' \in W^\perp.$$

Comme $\overrightarrow{w}' \in W^\perp$ et $\overrightarrow{v} \in (W^\perp)^\perp$, nous avons

$$\overrightarrow{v} \cdot \overrightarrow{w}' = 0.$$

Mais

$$0 = \overrightarrow{v} \cdot \overrightarrow{w}' = (\overrightarrow{w} + \overrightarrow{w}') \cdot \overrightarrow{w}' = \overrightarrow{w} \cdot \overrightarrow{w}' + \overrightarrow{w}' \cdot \overrightarrow{w}'.$$

Or $\overrightarrow{w} \cdot \overrightarrow{w}' = 0$ car $\overrightarrow{w} \in W$ et $\overrightarrow{w}' \in W^{\perp}$, donc

$$\overrightarrow{w}' \cdot \overrightarrow{w}' = 0.$$

Donc $\overrightarrow{w}' = \overrightarrow{0}$, et donc

$$\overrightarrow{v} = \overrightarrow{w} \in W.$$

On a donc $(W^{\perp})^{\perp} \subset W$, et finalement

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

□

Remarques 6.6.0.35. *1. L'unicité de la décomposition* $\overrightarrow{v} = \overrightarrow{w} + \overrightarrow{w}'$ *montre que la projection* $\mathrm{proj}_W \overrightarrow{v}$ *ne dépend que du sous-espace vectoriel* W *et non pas de la base (orthogonale) de* W *choisie.*

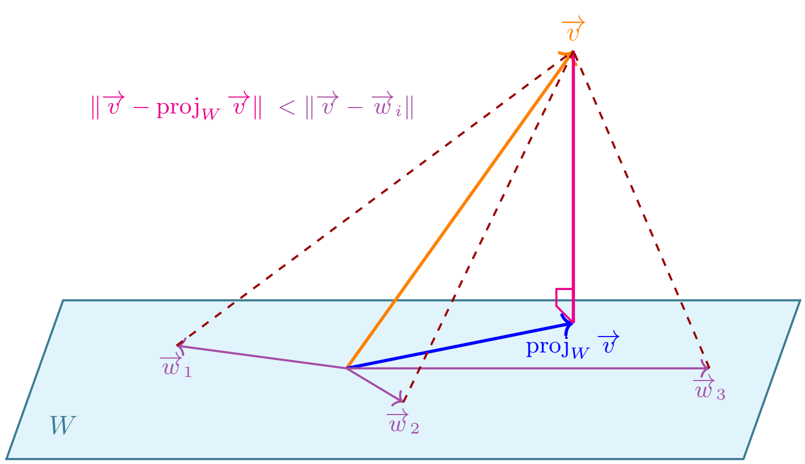
2. Si $\overrightarrow{v} \in W$*, alors* $\mathrm{proj}_W \overrightarrow{v} = \overrightarrow{v}$*.*

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

Théorème 6.36 *Théorème de la meilleure approximation*

Soit W un sous-espace de $\mathbb{R}^n$ et $\overrightarrow{v} \in \mathbb{R}^n$.
Alors $\|\overrightarrow{v} - \mathrm{proj}_W \overrightarrow{v}\| < \|\overrightarrow{v} - \overrightarrow{w}\|$ pour tout $\overrightarrow{w} \in W$ avec $\overrightarrow{w} \neq \mathrm{proj}_W \overrightarrow{v}$.

Autrement dit, $\mathrm{proj}_W \overrightarrow{v}$ est la meilleure approximation de $\overrightarrow{v}$ par un élément de W.



Démonstration. On fixe $\overrightarrow{v} \in \mathbb{R}^n$ et on note $\mathrm{proj}_W \overrightarrow{v}$ sa projection sur W.

D'après le théorème de la projection orthogonale, on a

$$\overrightarrow{v} = \mathrm{proj}_W \overrightarrow{v} + \overrightarrow{w}', \quad \text{avec } \mathrm{proj}_W \overrightarrow{v} \in W \text{ et } \overrightarrow{w}' \in W^{\perp}.$$

En particulier,
$$\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v} = \overrightarrow{w}' \in W^\perp.$$

Pour tout $\overrightarrow{w} \in W$, on a aussi
$$\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w} \in W,$$

puisque différence de deux vecteurs de W.

Ainsi, pour tout $\overrightarrow{w} \in W$,
$$\overrightarrow{v} - \overrightarrow{w} = (\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}) + (\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w}),$$

où $\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v} \in W^\perp$ et $\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w} \in W$ sont orthogonaux.

Par le théorème de Pythagore, on obtient donc
$$\|\overrightarrow{v} - \overrightarrow{w}\|^2 = \|\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}\|^2 + \|\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w}\|^2.$$

Si $\overrightarrow{w} \neq \operatorname{proj}_W \overrightarrow{v}$, alors $\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w} \neq \overrightarrow{0}$, donc
$$\|\operatorname{proj}_W \overrightarrow{v} - \overrightarrow{w}\|^2 > 0,$$

et par conséquent
$$\|\overrightarrow{v} - \overrightarrow{w}\|^2 > \|\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}\|^2.$$

En prenant la racine des deux côtés, on obtient pour tout $\overrightarrow{w} \in W$ avec $\overrightarrow{w} \neq \operatorname{proj}_W \overrightarrow{v}$:
$$\|\overrightarrow{v} - \overrightarrow{w}\| > \|\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}\|.$$

Cela montre que $\operatorname{proj}_W \overrightarrow{v}$ est le vecteur de W le plus proche de $\overrightarrow{v}$. □

Cela motive la définition suivante.

Définition 6.37 *Distance d'un vecteur à un sous-espace*

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$. Soit $\overrightarrow{v} \in \mathbb{R}^n$ un vecteur quelconque. On définit la *distance* entre $\overrightarrow{v}$ et W, notée $\operatorname{dist}(\overrightarrow{v}, W)$, par
$$\operatorname{dist}(\overrightarrow{v}, W) = \left\|\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}\right\|$$

Exemple. Considérons $\mathbb{R}^3$ et le sous-espace
$$W = \operatorname{Vect}(\overrightarrow{e}_1, \overrightarrow{e}_2) \quad \text{où} \quad \overrightarrow{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \ \overrightarrow{e}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}.$$

C'est le plan (x, y), et $(\overrightarrow{e}_1, \overrightarrow{e}_2)$ est une base orthonormée de W.

Prenons

$$\overrightarrow{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

1. Projection de $\overrightarrow{v}$ sur W.

Comme W est engendré par une base orthonormée, on a

$$\mathrm{proj}_W \overrightarrow{v} = (\overrightarrow{v} \cdot \overrightarrow{e}_1)\overrightarrow{e}_1 + (\overrightarrow{v} \cdot \overrightarrow{e}_2)\overrightarrow{e}_2.$$

Or

$$\overrightarrow{v} \cdot \overrightarrow{e}_1 = 1, \qquad \overrightarrow{v} \cdot \overrightarrow{e}_2 = 2,$$

d'où

$$\mathrm{proj}_W \overrightarrow{v} = 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

On a donc la décomposition orthogonale

$$\overrightarrow{v} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}_{\in W} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}}_{\in W^\perp}.$$

2. Distance de $\overrightarrow{v}$ à W.

Par définition,

$$\mathrm{dist}(\overrightarrow{v}, W) = \|\overrightarrow{v} - \mathrm{proj}_W \overrightarrow{v}\| = \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\| = 3.$$

3. Comparaison avec d'autres vecteurs de W.

Prenons par exemple deux vecteurs de W différents de $\mathrm{proj}_W \overrightarrow{v}$:

$$\overrightarrow{w}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \overrightarrow{w}_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}.$$

Alors

$$\|\overrightarrow{v} - \overrightarrow{w}_1\| = \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14} > 3,$$

et

$$\|\overrightarrow{v} - \overrightarrow{w}_2\| = \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \right\| = \sqrt{(-1)^2 + 2^2 + 3^2} = \sqrt{14} > 3.$$

Donc parmi ces vecteurs de W, c'est bien

$$\operatorname{proj}_W \overrightarrow{v} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

qui est le plus proche de $\overrightarrow{v}$.

Exemple. Considérons $\mathbb{R}^3$ et le sous-espace

$$W = \operatorname{Vect}(\overrightarrow{u}_1, \overrightarrow{u}_2) \quad \text{où} \quad \overrightarrow{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \overrightarrow{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

On vérifie que

$$\overrightarrow{u}_1 \cdot \overrightarrow{u}_2 = 1 \cdot 1 + 1 \cdot (-1) + 1 \cdot 0 = 0,$$

donc $(\overrightarrow{u}_1, \overrightarrow{u}_2)$ est une base orthogonale de W (non orthonormée).

Prenons maintenant

$$\overrightarrow{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Projection de $\overrightarrow{v}$ sur W.

Comme la base de W est orthogonale, la définition 6.30 donne

$$\operatorname{proj}_W \overrightarrow{v} = \frac{\overrightarrow{v} \cdot \overrightarrow{u}_1}{\overrightarrow{u}_1 \cdot \overrightarrow{u}_1} \overrightarrow{u}_1 + \frac{\overrightarrow{v} \cdot \overrightarrow{u}_2}{\overrightarrow{u}_2 \cdot \overrightarrow{u}_2} \overrightarrow{u}_2.$$

On calcule les produits scalaires :

$$\overrightarrow{u}_1 \cdot \overrightarrow{u}_1 = 1^2 + 1^2 + 1^2 = 3, \qquad \overrightarrow{u}_2 \cdot \overrightarrow{u}_2 = 1^2 + (-1)^2 + 0^2 = 2,$$

$$\overrightarrow{v} \cdot \overrightarrow{u}_1 = 0 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 = 1, \qquad \overrightarrow{v} \cdot \overrightarrow{u}_2 = 0 \cdot 1 + 0 \cdot (-1) + 1 \cdot 0 = 0.$$

D'où

$$\operatorname{proj}_W \overrightarrow{v} = \frac{1}{3} \overrightarrow{u}_1 + 0 \cdot \overrightarrow{u}_2 = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}.$$

Distance de $\overrightarrow{v}$ à W.

La définition de la distance donne

$$\operatorname{dist}(\overrightarrow{v}, W) = \|\overrightarrow{v} - \operatorname{proj}_W \overrightarrow{v}\| = \left\| \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \right\| = \left\| \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \right\|.$$

On a
$$\|\overrightarrow{v} - \mathrm{proj}_W \overrightarrow{v}\|^2 = \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{1}{9} + \frac{1}{9} + \frac{4}{9} = \frac{6}{9} = \frac{2}{3},$$
donc
$$\mathrm{dist}(\overrightarrow{v}, W) = \sqrt{\frac{2}{3}}.$$

Théorème 6.38

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$. L'application $\begin{array}{rccl} \mathrm{proj}_W : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & \overrightarrow{v} & \longmapsto & \mathrm{proj}_W \overrightarrow{v} \end{array}$ est linéaire.

Démonstration. De même que pour la projection sur une droite, la linéarité se montre en utilisant la linéarité du produit scalaire. □

Théorème 6.39 $\mathrm{Ker}(\mathrm{proj}_W) = W^\perp$

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$. On a $\mathrm{Ker}(\mathrm{proj}_W) = W^\perp$

Démonstration. On a les équivalences suivantes,
$$\begin{aligned} \mathrm{proj}_W \overrightarrow{v} = \overrightarrow{0} &\iff \frac{\overrightarrow{v} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}} \overrightarrow{u_1} + \cdots + \frac{\overrightarrow{v} \cdot \overrightarrow{u_k}}{\overrightarrow{u_k} \cdot \overrightarrow{u_k}} \overrightarrow{u_k} = \overrightarrow{0} \\ &\iff \overrightarrow{v} \cdot \overrightarrow{u_j} = 0 \text{ pour tout } j \in \{1, \ldots, k\} \text{ car les } u_j \text{ constituent une base de } W \\ &\iff \overrightarrow{v} \in W^\perp \end{aligned}$$

Donc $\mathrm{Ker}(\mathrm{proj}_W) = W^\perp$. □

Comme d'autre part, $\mathrm{Im}(\mathrm{proj}_W) = W$ d'après le corollaire 6.33, le théorème du rang nous donne donc que pour tout sous-espace vectoriel W de $\mathbb{R}^n$,
$$\dim(W^\perp) = n - \dim(W).$$
C'est aussi vrai pour $W = \{\overrightarrow{0}\}$.

Par construction,

— si $\overrightarrow{v} \in W$, alors $\mathrm{proj}_W \overrightarrow{v} = \overrightarrow{v}$. Ainsi $\overrightarrow{v}$ est un vecteur propre de proj_W de valeur propre associée $\lambda = 1$;

— si $\overrightarrow{v} \in W^\perp$, alors $\mathrm{proj}_W \overrightarrow{v} = \overrightarrow{0}$. Ainsi $\overrightarrow{v}$ est un vecteur propre de proj_W de valeur propre associée $\lambda = 0$.

Soit $\left(\vec{b}_1, \ldots, \vec{b}_k\right)$ une base de W (pas forcément orthogonale) et soit $\{\vec{v}_1, \ldots, \vec{v}_{n-k}\}$ une base de $W^\perp$ (pas forcément orthogonale). L'ensemble

$$\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_k, \vec{v}_1, \ldots, \vec{v}_{n-k}\right)$$

est une base de $\mathbb{R}^n$ et la matrice de proj_W par rapport à la base $\mathcal{B}$ est diagonale :

$$A^{\mathcal{B}}_{\mathrm{proj}_W} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} k \text{ fois} \\ \left.\vphantom{\begin{matrix}0\\ \ddots\\ 0\end{matrix}}\right\} n-k \text{ fois} \end{array}$$

Théorème 6.40 *Matrice canoniquement associée à* proj_W

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $k > 0$ et $(\vec{u}_1, \ldots, \vec{u}_k)$ une base orthonormée de W. On note

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_k \end{bmatrix}.$$

Alors, la matrice canoniquement associée à proj_W est UU^T.

Avant de démontrer ce théorème, nous allons énoncer et démontrer un théorème utile :

Théorème 6.41

Soit $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix} \in \mathcal{M}_{m,n}(\mathbb{R})$ Les n colonnes de U sont orthonormées si et seulement si

$$U^T U = I_n.$$

Démonstration. ($\Rightarrow$)

Par définition du produit matriciel, l'élément de la i-ème ligne et j-ème colonne de $U^T U$ est

$$(U^T U)_{ij} = \vec{u}_i^T \vec{u}_j = \vec{u}_i \cdot \vec{u}_j.$$

Comme les colonnes de U sont orthonormées. Alors

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

Donc $U^T U = I_n$.

($\Leftarrow$)

Réciproquement, supposons que $U^T U = I_n$. Alors, pour tous i, j,

$$\vec{u}_i \cdot \vec{u}_j = (U^T U)_{ij} = (I_n)_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

Donc $\overrightarrow{u}_i \cdot \overrightarrow{u}_j = 0$ si $i \neq j$ et $\overrightarrow{u}_i \cdot \overrightarrow{u}_i = 1$ pour tout i, ce qui signifie que les colonnes de U sont orthonormées. □

Démonstration. du théorème 6.40
Comme $(\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k)$ est une base de W et

$$U = \begin{bmatrix} \overrightarrow{u}_1 & \overrightarrow{u}_2 & \cdots & \overrightarrow{u}_k \end{bmatrix},$$

nous avons $W = \mathrm{Im}(U)$, car $\mathrm{Im}(U)$ désigne le sous-espace vectoriel engendré par les colonnes de U.

Par conséquent (propriété 6.16),

$$W^\perp = (\mathrm{Im}(U))^\perp = \mathrm{Ker}(U^T).$$

Le théorème 6.32 de la projection orthogonale nous dit que tout vecteur $\overrightarrow{v} \in \mathbb{R}^n$ s'écrit de manière unique

$$\overrightarrow{v} = \mathrm{proj}_W \overrightarrow{v} + \overrightarrow{w}', \quad \text{où } \mathrm{proj}_W \overrightarrow{v} \in W \text{ et } \overrightarrow{w}' \in W^\perp.$$

Comme $\mathrm{proj}_W \overrightarrow{v} \in W = \mathrm{Im}(U)$, on a

$$\mathrm{proj}_W \overrightarrow{v} = U\overrightarrow{x} \quad \text{pour un certain } \overrightarrow{x} \in \mathbb{R}^k.$$

Comme $\overrightarrow{w}' \in W^\perp = \mathrm{Ker}(U^T)$, on a

$$U^T\overrightarrow{w}' = \overrightarrow{0}.$$

Ainsi,

$$U^T\overrightarrow{v} = U^T(\mathrm{proj}_W \overrightarrow{v} + \overrightarrow{w}') = U^T \mathrm{proj}_W \overrightarrow{v} + U^T\overrightarrow{w}' = U^TU\overrightarrow{x} + \overrightarrow{0},$$

d'où

$$U^T\overrightarrow{v} = U^TU\overrightarrow{x}.$$

Comme les k colonnes de U forment un ensemble orthonormé, on a

$$U^TU = I_k,$$

ce qui donne

$$\overrightarrow{x} = U^T\overrightarrow{v} = \begin{bmatrix} \overrightarrow{u}_1^T\overrightarrow{v} \\ \vdots \\ \overrightarrow{u}_k^T\overrightarrow{v} \end{bmatrix} = \begin{bmatrix} \overrightarrow{v} \cdot \overrightarrow{u}_1 \\ \vdots \\ \overrightarrow{v} \cdot \overrightarrow{u}_k \end{bmatrix}.$$

On en déduit

$$\begin{aligned} \mathrm{proj}_W \overrightarrow{v} &= U\overrightarrow{x} \\ &= (\overrightarrow{v} \cdot \overrightarrow{u}_1)\overrightarrow{u}_1 + \cdots + (\overrightarrow{v} \cdot \overrightarrow{u}_k)\overrightarrow{u}_k \\ &= U\big(U^T\overrightarrow{v}\big) \\ &= (UU^T)\,\overrightarrow{v}. \end{aligned}$$

□

Remarque importante 6.42

Attention à ne pas confondre UU^T et U^TU !

6.6 Matrices orthogonales

Théorème 6.43

Soit $U \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice de taille $m \times n$ dont les colonnes sont orthonormées. Si $\vec{v}, \vec{w} \in \mathbb{R}^n$, alors :

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

Dans ce cas, l'application linéaire $T : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ \vec{v} & \longmapsto & U\vec{v} \end{array}$ conserve les longueurs et l'orthogonalité et elle est injective.

Démonstration. Comme les colonnes de U sont orthonormées, on sait que

$$U^TU = I_n.$$

1. Pour la norme, on calcule

$$\begin{aligned}\|U\vec{v}\|^2 &= (U\vec{v}) \cdot (U\vec{v}) = (U\vec{v})^T(U\vec{v}) \\ &= \vec{v}^TU^TU\vec{v} = \vec{v}^TI_n\vec{v} = \vec{v}^T\vec{v} = \|\vec{v}\|^2.\end{aligned}$$

En prenant la racine, on obtient $\|U\vec{v}\| = \|\vec{v}\|$.

On a donc que T est injective, puisque $U\vec{v} = \vec{0} \Leftrightarrow \|U\vec{v}\| = 0 \Leftrightarrow \|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$.

2. Pour le produit scalaire,

$$\begin{aligned}(U\vec{v}) \cdot (U\vec{w}) &= (U\vec{v})^T(U\vec{w}) \\ &= \vec{v}^TU^TU\vec{w} \\ &= \vec{v}^TI_n\vec{w} \\ &= \vec{v}^T\vec{w} = \vec{v} \cdot \vec{w}.\end{aligned}$$

3. Le point 2 donne directement

$$(U\vec{v}) \cdot (U\vec{w}) = \vec{v} \cdot \vec{w},$$

d'où

$$(U\vec{v}) \cdot (U\vec{w}) = 0 \iff \vec{v} \cdot \vec{w} = 0.$$

$\square$

Exemple. Soit

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix}$$

Les colonnes

$$\overrightarrow{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{et} \quad \overrightarrow{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

sont orthonormées.

Prenons $\overrightarrow{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ et $\overrightarrow{w} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ dans $\mathbb{R}^2$.

On calcule d'abord :

$$U\overrightarrow{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} \\ 1 \end{bmatrix}$$

$$U\overrightarrow{w} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -3 \end{bmatrix}$$

Conservation de la norme.

$$\|\overrightarrow{v}\| = \sqrt{3^2 + 1^2} = \sqrt{10}$$

$$\|U\overrightarrow{v}\| = \sqrt{\left(\frac{3}{\sqrt{2}}\right)^2 + \left(\frac{3}{\sqrt{2}}\right)^2 + 1^2} = \sqrt{\frac{9}{2} + \frac{9}{2} + 1} = \sqrt{10}$$

Donc $\|U\overrightarrow{v}\| = \|\overrightarrow{v}\|$.

De même :

$$\|\overrightarrow{w}\| = \sqrt{1^2 + (-3)^2} = \sqrt{10}$$

$$\|U\overrightarrow{w}\| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + (-3)^2} = \sqrt{\frac{1}{2} + \frac{1}{2} + 9} = \sqrt{10}$$

Donc $\|U\overrightarrow{w}\| = \|\overrightarrow{w}\|$.

Conservation du produit scalaire.

$$\overrightarrow{v} \cdot \overrightarrow{w} = 3 \cdot 1 + 1 \cdot (-3) = 3 - 3 = 0$$

$$(U\overrightarrow{v}) \cdot (U\overrightarrow{w}) = \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + 1 \cdot (-3) = \frac{3}{2} + \frac{3}{2} - 3 = 0$$

Donc $(U\overrightarrow{v}) \cdot (U\overrightarrow{w}) = \overrightarrow{v} \cdot \overrightarrow{w}$.

L'application linéaire $T : \mathbb{R}^2 \to \mathbb{R}^3$ définie par $T(\overrightarrow{x}) = U\overrightarrow{x}$ conserve donc les longueurs et l'orthogonalité. Les colonnes de U sont indépendantes, donc T est injective.

Définition 6.44 *Matrice orthogonale*

On dit qu'une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si A est inversible et $A^{-1} = A^T$.

Remarque 6.6.0.45. *Comme par définition, nous avons $AA^T = \mathbb{I}_n$, les colonnes de A sont orthonormées, d'après le théorème 6.41, et forment donc une base orthonormale de $\mathbb{R}^n$.*

C'est aussi le cas des lignes de A, puisque $A^TA = \mathbb{I}_n$

Exemples.

1. **Matrice identité.** Pour tout $n \geq 1$, la matrice identité $\mathbb{I}_n$ est orthogonale car $\mathbb{I}_n^T\mathbb{I}_n = \mathbb{I}_n$.
2. **Matrice de rotation dans $\mathbb{R}^2$.** Pour tout angle $\theta \in \mathbb{R}$, la matrice

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

est orthogonale. En effet :

$$R_\theta^T R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Par exemple, pour $\theta = \frac{\pi}{4}$:

$$R_{\pi/4} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

3. **Matrice de réflexion.** La matrice

$$S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

représente une réflexion par rapport à l'axe des x et est orthogonale car $S^TS = S^2 = \mathbb{I}_2$. Plus généralement, pour $a, b \in \mathbb{R}$ avec $a^2 + b^2 = 1$, la matrice de réflexion par rapport à la droite engendrée par $\begin{bmatrix} a \\ b \end{bmatrix}$ est :

$$H = \begin{bmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{bmatrix}$$

qui est orthogonale.

4. **Matrice de permutation.** La matrice

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

qui permute circulairement les coordonnées est orthogonale car $P^TP = \mathbb{I}_3$.

5. **Matrices de Householder.** Pour tout vecteur unitaire $\overrightarrow{u} \in \mathbb{R}^n$ (avec $\|\overrightarrow{u}\| = 1$), la matrice

$$H = \mathbb{I}_n - 2\overrightarrow{u}\,\overrightarrow{u}^T$$

est orthogonale. Par exemple, pour $\overrightarrow{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ dans $\mathbb{R}^3$:

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

6. **Matrice diagonale à coefficients ± 1.** La matrice

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

est orthogonale car $D^T D = D^2 = \mathbb{I}_3$.

7. **Produit de matrices orthogonales.** Si A et B sont des matrices orthogonales de même taille, alors AB est aussi orthogonale car :

$$(AB)^T(AB) = B^T A^T AB = B^T I_n B = B^T B = \mathbb{I}_n$$

6.7 Le procédé d'orthogonalisation de Gram-Schmidt

Le procédé de Gram-Schmidt permet de transformer une base quelconque en une base orthogonale (ou orthonormée).

Méthode 6.46 *Procédé de Gram-Schmidt*

Soit $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_k}\}$ un ensemble linéairement indépendant dans $\mathbb{R}^n$.

On construit une base orthogonale $\{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_k}\}$ de $\operatorname{Vect}(\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_k})$ comme suit :

$$\begin{aligned}
\overrightarrow{u_1} &= \overrightarrow{v_1} \\
\overrightarrow{u_2} &= \overrightarrow{v_2} - \operatorname{proj}_{\overrightarrow{u_1}} \overrightarrow{v_2} \\
\overrightarrow{u_3} &= \overrightarrow{v_3} - \operatorname{proj}_{\overrightarrow{u_1}} \overrightarrow{v_3} - \operatorname{proj}_{\overrightarrow{u_2}} \overrightarrow{v_3} \\
&\vdots \\
\overrightarrow{u_k} &= \overrightarrow{v_k} - \sum_{i=1}^{k-1} \operatorname{proj}_{\overrightarrow{u_i}} \overrightarrow{v_k}
\end{aligned}$$

Pour obtenir une base orthonormée, on normalise : $\overrightarrow{q_i} = \dfrac{\overrightarrow{u_i}}{\|\overrightarrow{u_i}\|}$.

Remarques 6.6.0.47. Pourquoi le procédé fonctionne-t-il ? *À chaque étape j, le vecteur $\overrightarrow{u}_j$ est construit en retirant de $\overrightarrow{v}_j$ toutes ses composantes dans les directions $\overrightarrow{u}_1, \ldots, \overrightarrow{u}_{j-1}$.*

Ce qui reste est donc orthogonal à tous les vecteurs précédents. De plus, comme $\overrightarrow{v}_j \notin \operatorname{Vect}(\overrightarrow{v}_1, \ldots, \overrightarrow{v}_{j-1})$ (par indépendance linéaire), le vecteur $\overrightarrow{u}_j$ est nécessairement non nul.

Espace engendré *Pour tout $j \in \{1, \ldots, k\}$, on a :*

$$\operatorname{Vect}(\overrightarrow{u}_1, \ldots, \overrightarrow{u}_j) = \operatorname{Vect}(\overrightarrow{v}_1, \ldots, \overrightarrow{v}_j)$$

En particulier, $\operatorname{Vect}(\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k) = \operatorname{Vect}(\overrightarrow{v}_1, \ldots, \overrightarrow{v}_k)$.

Exemple. Orthogonalisons la base $\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right)$ de $\mathbb{R}^3$.

Étape 1 : $\overrightarrow{u_1} = \overrightarrow{v_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

Étape 2 :

$$\operatorname{proj}_{\overrightarrow{u_1}} \overrightarrow{v_2} = \frac{\overrightarrow{v_2} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}} \overrightarrow{u_1} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\overrightarrow{u_2} = \overrightarrow{v_2} - \operatorname{proj}_{\overrightarrow{u_1}} \overrightarrow{v_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

Étape 3 :

$$\operatorname{proj}_{\overrightarrow{u_1}} \overrightarrow{v_3} = \frac{\overrightarrow{v_3} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}} \overrightarrow{u_1} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\operatorname{proj}_{\overrightarrow{u_2}} \overrightarrow{v_3} = \frac{\overrightarrow{v_3} \cdot \overrightarrow{u_2}}{\overrightarrow{u_2} \cdot \overrightarrow{u_2}} \overrightarrow{u_2} = \frac{1/2}{3/2} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\overrightarrow{u_3} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

On obtient la base orthogonale $\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix} \right)$.

Normalisation : Pour obtenir une base orthonormée, calculons les normes et normalisons :

$$\|\overrightarrow{u}_1\| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$$

$$\|\overrightarrow{u}_2\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + 1^2} = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2}$$

$$\|\overrightarrow{u}_3\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = \frac{2\sqrt{3}}{3}$$

La base orthonormée est donc :

$$\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{\sqrt{6}}{6} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \frac{\sqrt{3}}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right)$$

Exemple. Orthogonalisons l'ensemble $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ de $\mathbb{R}^4$ par le procédé de Gram–Schmidt.

On note

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Étape 1 :

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Étape 2 :

$$\operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

On calcule

$$\vec{v}_2 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 = 1, \qquad \vec{u}_1 \cdot \vec{u}_1 = 1^2 + 1^2 = 2,$$

d'où

$$\operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}.$$

On pose alors

$$\vec{u}_2 = \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

On vérifie que $\vec{u}_2 \cdot \vec{u}_1 = 0$.

Étape 3 :

$$\text{proj}_{\vec{u}_1} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

Comme

$$\vec{v}_3 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 0 \cdot 0 + 1 \cdot 0 = 1,$$

on obtient

$$\text{proj}_{\vec{u}_1} \vec{v}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}.$$

Pour la projection sur $\vec{u}_2$:

$$\text{proj}_{\vec{u}_2} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2.$$

On calcule

$$\vec{v}_3 \cdot \vec{u}_2 = 1 \cdot \frac{1}{2} + 0 \cdot \left(-\frac{1}{2}\right) + 0 \cdot 1 + 1 \cdot 0 = \frac{1}{2},$$

$$\vec{u}_2 \cdot \vec{u}_2 = \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2 + 0^2 = \frac{1}{4} + \frac{1}{4} + 1 = \frac{3}{2},$$

donc

$$\frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = \frac{1/2}{3/2} = \frac{1}{3},$$

et

$$\text{proj}_{\vec{u}_2} \vec{v}_3 = \frac{1}{3} \vec{u}_2 = \frac{1}{3} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/6 \\ -1/6 \\ 1/3 \\ 0 \end{bmatrix}.$$

On pose enfin

$$\begin{aligned} \vec{u}_3 &= \vec{v}_3 - \text{proj}_{\vec{u}_1} \vec{v}_3 - \text{proj}_{\vec{u}_2} \vec{v}_3 \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/6 \\ -1/6 \\ 1/3 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{6} \\ 0 - \frac{1}{2} + \frac{1}{6} \\ 0 - 0 - \frac{1}{3} \\ 1 - 0 - 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix}. \end{aligned}$$

On vérifie que

$$\vec{u}_1 \cdot \vec{u}_3 = 0 \quad \text{et} \quad \vec{u}_2 \cdot \vec{u}_3 = 0,$$

donc les vecteurs $\vec{u}_1, \vec{u}_2, \vec{u}_3$ sont deux à deux orthogonaux.

On a ainsi obtenu une base orthogonale de Vect $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$:

$$\mathcal{B}_{\text{ortho}} = \left(\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/3 \\ -1/3 \\ -1/3 \\ 1 \end{bmatrix} \right).$$

Si l'on souhaite une base orthonormée, il suffit de normaliser chacun des vecteurs $\vec{u}_1, \vec{u}_2, \vec{u}_3$.

6.8 Factorisation QR

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 et n très grands. Elle peut aussi fournir un moyen rapide de calculer des valeurs propres d'une matrice.

Définition 6.48 *Factorisation QR*

Soit A une matrice $m \times n$ dont les colonnes sont linéairement indépendantes.
La *factorisation QR* de A est : $A = QR$
où :
- Q est une matrice $m \times n$ dont les colonnes forment une base orthonormée de $\text{Im}(A)$
- R est une matrice $n \times n$ triangulaire supérieure inversible avec des éléments diagonaux strictement positifs

Théorème 6.49 *Théorème de factorisation QR*

Toute matrice A de taille $m \times n$ avec $\text{rang}(A) = n$ admet une factorisation QR unique.

Démonstration. On suppose que $A \in \mathcal{M}_{m,n}(\mathbb{R})$ vérifie $\text{rang}(A) = n$. Les n colonnes de A sont alors linéairement indépendantes et forment une base $(\vec{a}_1, \ldots, \vec{a}_n)$ de $\text{Im}(A)$, le sous-espace vectoriel engendré par les colonnes de A.

En appliquant le procédé d'orthogonalisation de Gram–Schmidt à cette base, on obtient une base orthogonale $(\vec{u}_1, \ldots, \vec{u}_n)$ de $\text{Im}(A)$.

En normalisant chaque vecteur,

$$\vec{q}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|}, \quad \vec{q}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|}, \quad \ldots, \quad \vec{q}_n = \frac{\vec{u}_n}{\|\vec{u}_n\|},$$

on obtient une base orthonormale $(\vec{q}_1, \ldots, \vec{q}_n)$ de $\text{Im}(A)$.

On pose alors

$$Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \end{bmatrix}.$$

Par construction, les colonnes de Q sont orthonormées, donc

$$Q^T Q = I_n.$$

D'autre part, le procédé de Gram–Schmidt montre que chaque colonne $\vec{a}_j$ de A s'écrit comme combinaison linéaire de $\vec{q}_1, \ldots, \vec{q}_j$ seulement. On peut donc écrire

$$\vec{a}_j = r_{1j}\vec{q}_1 + r_{2j}\vec{q}_2 + \cdots + r_{jj}\vec{q}_j$$

pour certains coefficients r_{ij}. En rassemblant ces coefficients dans une matrice $R \in \mathcal{M}_n(\mathbb{R})$, on obtient une matrice triangulaire supérieure (car $r_{ij} = 0$ pour $i > j$), et

$$A = QR.$$

En multipliant cette égalité à gauche par Q^T, on a

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

On en déduit bien

$$R = Q^T A.$$

De plus, pour chaque j, on a

$$r_{jj} = \vec{q}_j \cdot \vec{a}_j = \|\vec{u}_j\| > 0,$$

Unicité de la factorisation.

Supposons $A = Q_1R_1 = Q_2R_2$ deux factorisations QR. Alors $Q_2^\top Q_1 R_1 = R_2$. En posant $S = Q_2^T Q_1$ matrice orthogonale $n \times n$, on a $R_2 = SR_1$, donc $S = R_2 R_1^{-1}$. Or R_1 et R_2 sont triangulaires supérieures, donc S aussi (produit de triangulaires supérieures). Une matrice orthogonale triangulaire supérieure est forcément diagonale avec des entrées diagonales égales à ± 1. La positivité des diagonales de R_1, R_2 impose alors $S = \mathbb{I}_n$, donc $Q_1 = Q_2$ et $R_1 = R_2$. $\square$

Remarques 6.6.0.50. *La condition* $\operatorname{rang}(A) = n$ *est équivalente à dire que les colonnes de* A *sont linéairement indépendantes, ce qui garantit que tous les vecteurs* $\vec{u}_j$ *obtenus par Gram-Schmidt sont non nuls, et donc que* R *est inversible avec des coefficients diagonaux strictement positifs.*

Méthode 6.51 *Construction de la factorisation QR*

1. Appliquer le procédé de Gram-Schmidt aux colonnes de A pour obtenir une base orthonormée $(\vec{q_1}, \ldots, \vec{q_n})$
2. Former $Q = \begin{bmatrix} \vec{q_1} & \cdots & \vec{q_n} \end{bmatrix}$
3. Calculer $R = Q^T A$ (matrice triangulaire supérieure)

Exemple. Soit $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$. Trouvons sa factorisation QR.

Les colonnes de A sont $\vec{a_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ et $\vec{a_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. Elles ne sont pas colinéaires, donc $\operatorname{rang}(A) = 2$.

Par Gram-Schmidt :

$$\overrightarrow{u_1} = \overrightarrow{a_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\overrightarrow{q_1} = \frac{\overrightarrow{u_1}}{\|\overrightarrow{u_1}\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\overrightarrow{u_2} = \overrightarrow{a_2} - (\overrightarrow{a_2} \cdot \overrightarrow{q_1})\overrightarrow{q_1} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\overrightarrow{q_2} = \frac{\overrightarrow{u_2}}{\|\overrightarrow{u_2}\|} = \frac{1}{\sqrt{3/2}} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

Donc : $Q = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & 2/\sqrt{6} \end{bmatrix}$

Et : $R = Q^T A = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{bmatrix}$

Exemple. Soit

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}.$$

On cherche une factorisation $A = QR$.

Les colonnes de A sont

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}.$$

Étape 1 : procédé de Gram–Schmidt.

Premier vecteur.

$$\vec{u}_1 = \vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \|\vec{u}_1\| = \sqrt{1^2 + 1^2} = \sqrt{2}, \qquad \vec{q}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Deuxième vecteur.

$$\operatorname{proj}_{\vec{u}_1} \vec{a}_2 = \frac{\vec{a}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

On calcule

$$\vec{a}_2 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 + 1 \cdot 0 = 1, \quad \vec{u}_1 \cdot \vec{u}_1 = 2,$$

d'où

$$\operatorname{proj}_{\vec{u}_1} \vec{a}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

On définit

$$\vec{u}_2 = \vec{a}_2 - \operatorname{proj}_{\vec{u}_1} \vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Sa norme vérifie

$$\|\vec{u}_2\|^2 = \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2 + 0^2 + 1^2 = \frac{5}{2}, \quad \|\vec{u}_2\| = \sqrt{\frac{5}{2}} = \frac{\sqrt{10}}{2},$$

et donc

$$\vec{q}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{2}{\sqrt{10}} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 2 \end{bmatrix}.$$

Troisième vecteur.

$$\vec{u}_3 = \vec{a}_3 - \operatorname{proj}_{\vec{u}_1} \vec{a}_3 - \operatorname{proj}_{\vec{u}_2} \vec{a}_3.$$

On calcule d'abord

$$\operatorname{proj}_{\vec{u}_1} \vec{a}_3 = \frac{\vec{a}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

Comme

$$\vec{a}_3 \cdot \vec{u}_1 = 0 \cdot 1 + 1 \cdot 1 + (-1) \cdot 0 + 2 \cdot 0 + 0 \cdot 0 = 1,$$

on obtient

$$\operatorname{proj}_{\vec{u}_1} \vec{a}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Puis

$$\operatorname{proj}_{\vec{u}_2} \vec{a}_3 = \frac{\vec{a}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\, \vec{u}_2.$$

On a

$$\vec{a}_3 \cdot \vec{u}_2 = 0 \cdot \frac{1}{2} + 1 \cdot \left(-\frac{1}{2}\right) + (-1) \cdot 1 + 2 \cdot 0 + 0 \cdot 1 = -\frac{3}{2},$$

$$\vec{u}_2 \cdot \vec{u}_2 = \frac{5}{2} \quad \Rightarrow \quad \frac{\vec{a}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = -\frac{3}{5},$$

et donc

$$\operatorname{proj}_{\vec{u}_2} \vec{a}_3 = -\frac{3}{5} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/10 \\ 3/10 \\ -3/5 \\ 0 \\ -3/5 \end{bmatrix}.$$

Ainsi

$$\begin{aligned} \vec{u}_3 &= \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -3/10 \\ 3/10 \\ -3/5 \\ 0 \\ -3/5 \end{bmatrix} \\ &= \begin{bmatrix} -1/5 \\ 1/5 \\ -2/5 \\ 2 \\ 3/5 \end{bmatrix}. \end{aligned}$$

Sa norme vérifie

$$\|\vec{u}_3\|^2 = \left(-\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(-\frac{2}{5}\right)^2 + 2^2 + \left(\frac{3}{5}\right)^2 = \frac{23}{5}, \quad \|\vec{u}_3\| = \sqrt{\frac{23}{5}} = \frac{\sqrt{115}}{5}.$$

Donc

$$\vec{q}_3 = \frac{\vec{u}_3}{\|\vec{u}_3\|} = \frac{5}{\sqrt{115}} \begin{bmatrix} -1/5 \\ 1/5 \\ -2/5 \\ 2 \\ 3/5 \end{bmatrix} = \frac{1}{\sqrt{115}} \begin{bmatrix} -1 \\ 1 \\ -2 \\ 10 \\ 3 \end{bmatrix}.$$

On a ainsi construit une base orthonormée

$$(\vec{q}_1, \vec{q}_2, \vec{q}_3)$$

de $\operatorname{Im}(A)$.

Étape 2 : matrice Q.

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{115}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} & \frac{1}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & -\frac{2}{\sqrt{115}} \\ 0 & 0 & \frac{10}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & \frac{3}{\sqrt{115}} \end{bmatrix}.$$

Étape 3 : matrice $R = Q^T A$.
Les coefficients de R sont $r_{ij} = \overrightarrow{q}_i \cdot \overrightarrow{a}_j$. On obtient

$$R = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{10}}{2} & -\frac{3}{\sqrt{10}} \\ 0 & 0 & \frac{\sqrt{115}}{5} \end{bmatrix}.$$

On vérifie que $A = QR$ et que Q a des colonnes orthonormées et R est triangulaire supérieure : c'est une factorisation QR de A.

6.9 La méthode des moindres carrés

La méthode des moindres carrés permet de trouver la « meilleure » solution approximative d'un système d'équations linéaires qui n'a pas de solution exacte.

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 ?