Méthodes factorielles

David Hébert
hebert.iut@gmail.com
Aurélie Nassiet
aurelie.nassiet@gmail.com

2023

Table des matières

Table des matières **1**

1 Produit scalaire **2**

2 Orthogonalité **6**

3 Bases orthonormales **9**

4 Projection **13**

5 Angles **17**

6 Matrices **19**

7 ACP **24**

1. Produit scalaire

Définition

On appelle **produit scalaire réel** sur un espace vectoriel E toute application, $f : E \times E \longrightarrow \mathbb{R}$ vérifiant les conditions suivantes :

(P0) **Positivité :** $\forall \overrightarrow{x} \in E,\ f(\overrightarrow{x}, \overrightarrow{x}) \geqslant 0$

(P1) **Séparation :** $\forall \overrightarrow{x} \in E,\ f(\overrightarrow{x}, \overrightarrow{x}) = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{0}$

(P2) **Symétrie :** $\forall \overrightarrow{x}, \overrightarrow{y} \in E,\ f(\overrightarrow{x}, \overrightarrow{y}) = f(\overrightarrow{y}, \overrightarrow{x})$

(P3) **Bilinéarité :** $\forall a, b \in \mathbb{R}, \forall \overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z} \in E,\ f(a\overrightarrow{x} + b\overrightarrow{y}, \overrightarrow{z}) = af(\overrightarrow{x}, \overrightarrow{z}) + bf(\overrightarrow{y}, \overrightarrow{z})$

Lorsqu'il n'y a pas d'ambiguïté, on a l'habitude de noter un produit scalaire $f(\overrightarrow{x}, \overrightarrow{y}) = \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle$.

Exercice

Les applications suivantes définissent-elles des produits scalaire sur $\mathbb{R}^2$.

1. $\varphi_1\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \sqrt{x_1^2 + y_1^2 + x_2^2 + y_2^2}$
2. $\varphi_2\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 4x_1y_1 - x_2y_2$
3. $\varphi_3\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 4x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2.$

Exercice

Soit f un produit scalaire sur un espace vectoriel E. Montrer en utilisant les axiomes des produits scalaire et uniquement ces axiomes, que pour tout $\overrightarrow{x} \in E$, $f\left(\overrightarrow{x}, \overrightarrow{0}\right) = 0$.

Définition

Un $\mathbb{R}$ espace vectoriel E de dimension finie muni d'un produit scalaire est appelé un **espace euclidien**.

Lorsque l'espace E est de dimension infinie, on dira qu'il est *préhilbertien* (ça ne sera pas l'objet de ce cours).

L'adjectif *euclidien* fait référence à Euclide, célèbre savant de la Grèce antique, qui serait le *premier* penseur des éléments de géométrie plane usuelle.

Pour des vecteurs $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $\overrightarrow{x} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, le produit scalaire usuelle $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = x_1y_1 + x_2y_2$ fait de $\mathbb{R}^2$ un espace euclidien.

Définition

On appel **norme** sur un espace vectoriel E toute application $N : E \to \mathbb{R}_+$ vérifiant les conditions suivantes :

(N1) **Séparation.** $\forall \overrightarrow{x} \in E,\ N(\overrightarrow{x}) = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{0}$

(N2) **Absolue homogénéité.** $\forall \lambda \in \mathbb{R}, \overrightarrow{x} \in E,\ N(\lambda\overrightarrow{x}) = |\lambda| N(\overrightarrow{x})$

(N3) **Inégalité du triangle.** $\forall \overrightarrow{x}, \overrightarrow{y} \in E,\ N(\overrightarrow{x} + \overrightarrow{y}) \leqslant N(\overrightarrow{x}) + N(\overrightarrow{y})$

Lorsqu'il n'y a pas d'ambiguïté, on a l'habitude de noter une norme $N(\overrightarrow{x}) = \|\overrightarrow{x}\|$.

Proposition

Soit E un espace vectoriel de dimension finie.

1. Soit $\langle \bullet \mid \bullet \rangle$ un produit scalaire sur E. Alors

$$\begin{array}{ccc} \mathsf{E} & \longrightarrow & \mathbb{R}_+ \\ \overrightarrow{x} & \longmapsto & \sqrt{\langle \overrightarrow{x} \mid \overrightarrow{x} \rangle} \end{array}$$

   définie une norme sur E.

2. Soit $\mathsf{N}(\bullet)$ une norme sur E. Alors

$$\begin{array}{ccc} \mathsf{E} \times \mathsf{E} & \longrightarrow & \mathbb{R} \\ (\overrightarrow{x}, \overrightarrow{y}) & \longmapsto & \dfrac{\mathsf{N}(\overrightarrow{x} + \overrightarrow{y})^2 - \mathsf{N}(\overrightarrow{x})^2 - \mathsf{N}(\overrightarrow{y})^2}{2} \end{array}$$

   définie un produit scalaire sur E.

Démonstration. Exercice - à jumeler avec le lemme suivant. □

Lemme : Soient $\langle \bullet \mid \bullet \rangle$ un produit scalaire sur un espace vectoriel E et pour tout $\overrightarrow{x} \in \mathsf{E}$, $\mathsf{N}(\overrightarrow{x}) = \sqrt{\langle \overrightarrow{x} \mid \overrightarrow{x} \rangle}$ alors

$$\forall \overrightarrow{x}, \overrightarrow{y} \in \mathsf{E},\ \mathsf{N}(\overrightarrow{x} + \overrightarrow{y})^2 = \mathsf{N}(\overrightarrow{x})^2 + 2\left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \mathsf{N}(\overrightarrow{y})^2$$

Démonstration.

$$\begin{aligned} \mathsf{N}(\overrightarrow{x} + \overrightarrow{y})^2 &= \left\langle \overrightarrow{x} + \overrightarrow{y} \mid \overrightarrow{x} + \overrightarrow{y} \right\rangle \\ &= \left\langle \overrightarrow{x} \mid \overrightarrow{x} + \overrightarrow{y} \right\rangle + \left\langle \overrightarrow{y} \mid \overrightarrow{x} + \overrightarrow{y} \right\rangle \\ &= \left\langle \overrightarrow{x} + \overrightarrow{y} \mid \overrightarrow{x} \right\rangle + \left\langle \overrightarrow{x} + \overrightarrow{y} \mid \overrightarrow{y} \right\rangle \\ &= \left\langle \overrightarrow{x} \mid \overrightarrow{x} \right\rangle + \left\langle \overrightarrow{y} \mid \overrightarrow{x} \right\rangle + \left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \left\langle \overrightarrow{y} \mid \overrightarrow{y} \right\rangle \\ &= \left\langle \overrightarrow{x} \mid \overrightarrow{x} \right\rangle + \left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \left\langle \overrightarrow{y} \mid \overrightarrow{y} \right\rangle \\ &= \left\langle \overrightarrow{x} \mid \overrightarrow{x} \right\rangle + 2\left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \left\langle \overrightarrow{y} \mid \overrightarrow{y} \right\rangle \\ &= \mathsf{N}(\overrightarrow{x})^2 + 2\left\langle \overrightarrow{x} \mid \overrightarrow{y} \right\rangle + \mathsf{N}(\overrightarrow{y})^2 \end{aligned}$$

□

Par exemple la norme euclidienne sur $\mathbb{R}^2$ donne pour un vecteur $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\|\overrightarrow{x}\| = \sqrt{x_1^2 + x_2^2}$

§ Qu'est-ce qu'un cercle ? Lorsque l'on pose cette question, le commun des mortels imagine une roue, ronde...

Considérons $\mathbb{R}^2$ comme un espace euclidien avec sa norme usuelle.

Le cercle de centre $(0,0)$ et de rayon 1, noté $\mathcal{C}$ est l'ensemble des points, que nous associerons à leur vecteur, de norme 1. Les plus savant dirons que $\mathcal{C}$ est l'ensemble des vecteurs $\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tel que $x_1^2 + x_2^2 = 1$ soit encore $\sqrt{x_1^2 + x_2^2} = 1$ soit encore $\|\overrightarrow{x}\| = 1$.

En fait la *vraie* définition d'un cercle centré en l'origine de rayon 1 est effectivement $\left\{ \overrightarrow{x} \in \mathbb{R}^2 \middle| \|\overrightarrow{x}\| = 1 \right\}$ et cela dessine un joli rond... avec la norme classique.

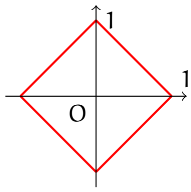
Donnons une description littérale de cette définition de cercle : c'est l'ensemble des points à une distance de 1 de l'origine.

Mais il existe d'autre distance, d'autre norme, comme par exemple la norme euclidienne associée au produit scalaire :

$$\forall \overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \overrightarrow{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},\ \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \frac{(|x_1+y_1|+|x_2+y_2|)^2 - (|x_1|+|x_2|))^2 - (|y_1|+|y_2|)^2}{2}$$

(exercice : vérifier qu'il s'agit bien d'un produit scalaire) dont on a $\|\overrightarrow{x}\| = |x_1| + |x_2|$.

Avec cette manière de *mesurer les distances* et en gardant notre définition qu'un cercle de rayon 1 est l'ensemble des points de norme 1, on peut affirmer sans aucune erreur que la figure ci-contre est un cercle.



Exercice

Considérons le produit scalaire de Chebyshev :

$$\forall \overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \overrightarrow{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},\ \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \frac{\max(|x_1+y_1|,|x_2+y_2|)^2 - \max(|x_1|,|x_2|))^2 - \max(|y_1|,|y_2|)^2}{2}$$

1. Vérifier qu'il s'agit bien d'un produit scalaire.
2. Donner l'expression de la norme associée.
3. Dessiner, pour cette norme, le cercle de centre $(0,0)$ et de rayon 1.

Exercice

Soit $\overrightarrow{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\overrightarrow{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ deux vecteurs de l'espace vectoriel $\mathbb{R}^n$ de dimension $n \in \mathbb{N}_{>0}$. Montrer que $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \sum_{i=1}^{n} x_i y_i$ est un produit scalaire sur $\mathbb{R}^n$.

Définition

Considérons $\mathbb{R}^n$ l'espace vectoriel sur $\mathbb{R}$ de dimension $n \in \mathbb{N}_{>0}$. On appelle **produit scalaire canonique** l'application

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (\overrightarrow{x}, \overrightarrow{y}) & \longmapsto & \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = {}^t\overrightarrow{x}\,\overrightarrow{y} \end{array}$$

Sa norme associée est appelé la *norme euclidienne* canonique.

La norme euclidienne canonique mesure les distances au sens usuelle. Par exemple, la taille du vecteur $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est $\sqrt{2}$.

Exercice

Considérons $\|\bullet\|$ la norme euclidienne canonique sur $\mathbb{R}^2$. Montrer que $\|\overrightarrow{x} + \overrightarrow{y}\| = \|\overrightarrow{x}\| + \|\overrightarrow{y}\|$ si et seulement si $\overrightarrow{x}$ et $\overrightarrow{y}$ sont colinéaire.

Théorème Inégalité de Cauchy-Schwarz

Soit $\langle \bullet \mid \bullet \rangle$ un produit scalaire sur un espace vectoriel de dimension finie et $\|\bullet\|$ sa norme associée.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in E, \ |\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle| \leqslant \|\overrightarrow{x}\| \, \|\overrightarrow{y}\|$$

Démonstration. POur tout $t \in \mathbb{R}$, l'expression $\|\overrightarrow{x} + t\overrightarrow{y}\|^2$ est un polynôme de degré 2 en t. Précisément, on a

$$\|\overrightarrow{x} + t\overrightarrow{y}\|^2 = \|\overrightarrow{x}\|^2 + 2t\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle + t^2 \|\overrightarrow{y}\|^2$$

Ce polynôme étant positif ou nul, son discriminant ne peut être strictement positif ainsi $(2\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle)^2 - 4\|\overrightarrow{x}\|^2 \|\overrightarrow{y}\|^2 \leqslant 0$. Le passage à la racine et la simplification par 4 permet de conclure. □

Théorème Inégalités de Minkowski

Soit $\|\bullet\|$ une norme sur un espace vectoriel E de dimension finie.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in E, \ \Big| \|x\| - \|y\| \Big| \leqslant \|\overrightarrow{x} + \overrightarrow{y}\| \leqslant \|\overrightarrow{x}\| + \|\overrightarrow{y}\|$$

Démonstration. L'inégalité de droite est l'inégalité du triangle (axiome (N3)). Pour l'inégalité de gauche on observe que $\|\overrightarrow{x}\| = \|\overrightarrow{x} + \overrightarrow{y} - \overrightarrow{y}\| \leqslant \|\overrightarrow{x} + \overrightarrow{y}\| + \|\overrightarrow{y}\|$ en utilisant l'inégalité du triangle. Soit en réécrivant $\|\overrightarrow{x}\| - \|\overrightarrow{y}\| \leqslant \|\overrightarrow{x} + \overrightarrow{y}\|$.

En changeant le rôle de $\overrightarrow{x}$ et $\overrightarrow{y}$ on conclut. □

Proposition Égalité du parallélogramme

Soit $\|\bullet\|$ une norme sur un espace vectoriel E de dimension finie.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in E, \ \|\overrightarrow{x} + \overrightarrow{y}\|^2 + \|\overrightarrow{x} - \overrightarrow{y}\|^2 = 2\left(\|\overrightarrow{x}\|^2 + \|\overrightarrow{y}\|^2\right)$$

Démonstration. D'après le précédent lemme on a $\|\overrightarrow{x} + \overrightarrow{y}\|^2 = \|\overrightarrow{x}\|^2 + 2\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle + \|\overrightarrow{y}\|^2$. En remplaçant $\overrightarrow{y}$ par $-\overrightarrow{y}$ on a également $\|\overrightarrow{x} - \overrightarrow{y}\|^2 = \|\overrightarrow{x}\|^2 - 2\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle + \|\overrightarrow{y}\|^2$. L'addition de ces deux égalités prouve l'égalité du parallélogramme. □

Corollaire Identité de polarisation

Soit $\|\bullet\|$ une norme sur un espace vectoriel E de dimension finie.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in E, \ \|\overrightarrow{x} + \overrightarrow{y}\|^2 - \|\overrightarrow{x} - \overrightarrow{y}\|^2 = 4\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle$$

Démonstration. Exercice □

2. Orthogonalité

Définition

Soient $\vec{x}$ et $\vec{y}$ deux vecteurs d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$.
On dira que $\vec{x}$ et $\vec{y}$ sont **orthogonaux** si $\left\langle \vec{x} \mid \vec{y} \right\rangle = 0$. On notera dans ce cas, $\vec{x} \perp \vec{y}$.

Par exemple dans $\mathbb{R}^2$ munit de sa structure euclidienne canonique les vecteurs $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sont orthogonaux.

Définition

Soit A un sous-ensemble non vide d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$. On note $A^\perp$ l'ensemble

$$A^\perp = \left\{ \vec{x} \in E \middle| \forall \vec{a} \in A, \vec{x} \perp \vec{a} \right\}$$

En d'autre terme $A^\perp$ est l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de A.

Proposition

Soient A et B des sous-ensembles non vide d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$.

(i). L'ensemble $A^\perp$ est un sous-espace vectoriel de E.

(ii). $A \subseteq B \Rightarrow B^\perp \subseteq A^\perp$

(iii). $A \subseteq \left(A^\perp\right)^\perp$

(iv). $A^\perp = \mathtt{Vect}(A)^\perp = \mathtt{Vect}\left(A^\perp\right)$

(v). On a $\{\vec{0}\}^\perp = E$ et $E^\perp = \{\vec{0}\}$

Démonstration.

(i). Effectuons la vérification rapide. Naturellement pour tout $\vec{a} \in A$, $\left\langle \vec{0} \middle| \vec{a} \right\rangle = 0$ donc $\vec{0} \in A^\perp$. Soient $\vec{x}, \vec{y} \in A^\perp$ et $\lambda \in \mathbb{R}$, vérifions que $\vec{x} + \lambda \vec{y} \in A^\perp$. Soit $\vec{a} \in A$ alors par la propriété de la bilinéarité des produits scalaires on a $\left\langle \vec{x} + \lambda \vec{y} \mid \vec{a} \right\rangle = \left\langle \vec{x} \mid \vec{a} \right\rangle + \lambda \left\langle \vec{y} \mid \vec{a} \right\rangle = 0 + \lambda 0 = 0$. CQFD.

(ii). Soit $\vec{x} \in B^\perp$, pour observer que $\vec{x} \in A^\perp$ nous devons vérifier que pour tout $\vec{a} \in A$, $\left\langle \vec{x} \mid \vec{a} \right\rangle = 0$ or $A \subseteq B$, c'est à dire que tout vecteur de A est un vecteur de B mais pour tout vecteur $\vec{a} \in B$ on a $\left\langle \vec{x} \mid \vec{a} \right\rangle = 0$. CQFD.

(iii). Par définition de $A^\perp$ on a

$$\forall \vec{x} \in A^\perp, \forall \vec{a} \in A, \left\langle \vec{x} \mid \vec{a} \right\rangle = 0$$

Par symétrie du produit scalaire, on peut récrire $\left\langle \vec{a} \mid \vec{x} \right\rangle = 0$ en particulier $\vec{a}$ est orthogonale à tous les vecteurs de $A^\perp$, c'est à dire que $\vec{a} \in \left(A^\perp\right)^\perp$. Soit $A \subseteq \left(A^\perp\right)^\perp$. CQFD.

(iv). On rappel que $\mathtt{Vect}(A)$ est le plus petit sous-espace vectoriel (au sens de l'inclusion) contenant les éléments de A. La définition de $\mathtt{Vect}(A)$ est l'ensemble des vecteurs $\vec{x}$ tel qu'il existe une famille finie $(\vec{a}_i)_i$ de vecteurs de A et des réels λ_i tel que $\vec{x} = \sum_i \lambda_i \vec{a}_i$.

Puisque $A^\perp$ est un espace vectoriel d'après la proposition (i), $A^\perp = \mathtt{Vect}\left(A^\perp\right)$.

Naturellement $A \subseteq \mathtt{Vect}(A)$ donc d'après la proposition (ii) on a $\mathtt{Vect}(A)^\perp \subseteq A^\perp$.
Il reste à montrer l'inclusion inverse, c'est à dire que si $\vec{x} \in A^\perp$ alors $\vec{x} \in \mathtt{Vect}(A)^\perp$, c'est à dire qu'il faut observer que pour toute combinaison linéaire $\vec{z} = \sum_i \lambda_i \vec{a}_i \in \mathtt{Vect}(A)$ alors $\left\langle \vec{x} \mid \vec{z} \right\rangle = 0$.

Mais la bilinéarité donne $\left\langle \vec{x} \mid \vec{z} \right\rangle = \sum_i \lambda_i \left\langle \vec{x} \mid \vec{a}_i \right\rangle$ et $\left\langle \vec{x} \mid \vec{a}_i \right\rangle = 0$ par définition de $\vec{x}$. CQFD.

(v). C'est une conséquence triviale de l'égalité $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{0} \right\rangle = 0$.

□

Exercice

On munit $\mathbb{R}^2$ de sa structure canonique d'espace euclidien.

Répondre par vrai ou par faux.

1. Si $\overrightarrow{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ et $\overrightarrow{y} = \begin{pmatrix} a \\ b \end{pmatrix}$ tel que $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = 0$ alors $a = -2b$.
2. Si $\overrightarrow{y}$ vérifie qu'il existe $\overrightarrow{x} \in E$ tel que $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = 0$ alors $\overrightarrow{y} = 0$.
3. Si $\overrightarrow{x}$ vérifie $\langle \overrightarrow{x} \mid \overrightarrow{x} \rangle = 0$ alors $\overrightarrow{x} = 0$.
4. Si $\overrightarrow{y}$ vérifie que pour tout $\overrightarrow{x} \in E$ on a $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = 0$ alors $\overrightarrow{y} = 0$.
5. Si $\overrightarrow{y}$ et $\overrightarrow{z}$ vérifient qu'il existe $\overrightarrow{x} \in E$ tel que $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \langle \overrightarrow{x} \mid \overrightarrow{z} \rangle$ alors $\overrightarrow{y} = \overrightarrow{z}$.
6. Si $\overrightarrow{y}$ et $\overrightarrow{z}$ vérifient que pour tout $\overrightarrow{x} \in E$ on a $\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \langle \overrightarrow{x} \mid \overrightarrow{z} \rangle$ alors $\overrightarrow{y} = \overrightarrow{z}$.
7. Si $\overrightarrow{x} \neq \overrightarrow{0}$ alors pour tout $\overrightarrow{y} \in E$, $\dfrac{\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle}{\langle \overrightarrow{x} \mid \overrightarrow{x} \rangle} \overrightarrow{x} \in \mathtt{Vect}(\overrightarrow{x})$
8. Si $\overrightarrow{x} \neq \overrightarrow{0}$ alors pour tout $\overrightarrow{y} \in E$, $\dfrac{\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle}{\langle \overrightarrow{x} \mid \overrightarrow{x} \rangle} \overrightarrow{y} \in \mathtt{Vect}(\overrightarrow{x})$

Théorème Pythagore

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in E, \qquad \overrightarrow{x} \perp \overrightarrow{y} \Leftrightarrow \left\| \overrightarrow{x} + \overrightarrow{y} \right\|^2 = \left\| \overrightarrow{x} \right\|^2 + \left\| \overrightarrow{y} \right\|^2$$

Démonstration. D'après le précédent lemme, on a $\left\| \overrightarrow{x} + \overrightarrow{y} \right\|^2 = \left\| \overrightarrow{x} \right\|^2 + 2\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle + \left\| \overrightarrow{y} \right\|^2$. CQFD. □

Le théorème de Pythagore fait parti des théorème les plus démontré au monde. Il existe en effet plusieurs centaines de démonstration différentes. En voici donc une de plus.

Définition

Soient A et B deux sous-espace vectoriels d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$. On dira que A et B sont **orthogonaux** si

$$\forall \overrightarrow{a} \in A, \ \forall \overrightarrow{b} \in B, \qquad \overrightarrow{a} \perp \overrightarrow{b}$$

Par exemple dans $\mathbb{R}^2$ munit de sa structure euclidienne canonique les droites $\mathtt{Vect}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ et $\mathtt{Vect}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ sont orthogonaux.

Proposition

Soient A et B deux sous-espace vectoriels orthogonaux d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$. Alors A et B sont en somme directe. On dit que cette somme directe est orthogonale et on note $A \overset{\perp}{\oplus} B$.

Démonstration. Il suffit de montrer que $A \cap B = \{\overrightarrow{0}\}$. Soit $\overrightarrow{x} \in A \cap B$ alors $\left\| \overrightarrow{x} \right\|^2 = \langle \overrightarrow{x} \mid \overrightarrow{x} \rangle = 0$ (le premier $\overrightarrow{x}$ vu dans A, le second dans B). D'après l'axiome de séparation des produits scalaire $(P1)$ on a $\overrightarrow{x} = \overrightarrow{0}$. □

Théorème Unicité de l'orthogonale

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien. Si $A \overset{\perp}{\oplus} B = E$ alors $A = B^\perp$.

Démonstration.

$A \subseteq B^\perp$**.** Soit $\overrightarrow{a} \in A$ et $\overrightarrow{b} \in B$ alors par définition de $A \overset{\perp}{\oplus} B$, $\left\langle \overrightarrow{a} \,\middle|\, \overrightarrow{b} \right\rangle = 0$ ce qui prouve $\overrightarrow{a} \in B^\perp$ et donc que $A \subseteq B^\perp$.

$B^\perp \subseteq A$**.** Soit $\overrightarrow{x} \in B^\perp$. Puisque $E = A \overset{\perp}{\oplus} B$ alors $\overrightarrow{x} = \overrightarrow{a} + \overrightarrow{b}$ pour $\overrightarrow{a} \in A$ et $\overrightarrow{b} \in B$ mais $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{b} \right\rangle = \left\langle \overrightarrow{a} \,\middle|\, \overrightarrow{b} \right\rangle + \left\langle \overrightarrow{b} \,\middle|\, \overrightarrow{b} \right\rangle$. Sauf que puisque $\overrightarrow{x} \in B^\perp$, $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{b} \right\rangle = 0$ et puisque A et B sont orthogonaux, $\left\langle \overrightarrow{a} \,\middle|\, \overrightarrow{b} \right\rangle = 0$. Ainsi $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{b} \right\rangle = \left\langle \overrightarrow{a} \,\middle|\, \overrightarrow{b} \right\rangle + \left\langle \overrightarrow{b} \,\middle|\, \overrightarrow{b} \right\rangle$ s'écrit en $0 = 0 + \left\| \overrightarrow{b} \right\|^2$ mais par l'hypothèse de séparation des normes implique que $\overrightarrow{b} = \overrightarrow{0}$. Finalement $\overrightarrow{x} = \overrightarrow{a} \in A$.

□

3. Bases orthonormales

Théorème

Dans un espace euclidien, toute famille de vecteurs non nuls deux à deux orthogonaux est libre.

Démonstration. Exercice □

Corollaire

Dans un espace euclidien de dimension $n \in \mathbb{N}_{>0}$, toute famille de n vecteurs non nuls deux à deux orthogonaux est une base.

Démonstration. D'après le théorème précédent les n vecteurs sont libre. Puisque la dimension de l'espace est n alors ces vecteurs sont nécessairement générateurs. Ils forment donc une base de l'espace euclidien. □

Définition

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de dimension n et $\mathcal{B} = \{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ une base.

(i). On dira que $\mathcal{B}$ est une **base orthogonale** si $\left\langle \overrightarrow{e_i} \mid \overrightarrow{e_j} \right\rangle = 0$ pour tout $i \neq j$.

(ii). On dira que $\mathcal{B}$ est une **base orthonormale** si c'est une base orthogonale et si de plus $\left\| \overrightarrow{e_i} \right\| = 1$ pour tout i.

Par exemple les vecteurs $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ forme une base orthogonale mais non orthonormale.

Proposition Normalisation

Si $\{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ est une base orthogonale d'un espace euclidien alors $\left\{ \frac{\overrightarrow{e_1}}{\left\|\overrightarrow{e_1}\right\|}, \dots, \frac{\overrightarrow{e_n}}{\left\|\overrightarrow{e_n}\right\|} \right\}$ est une base orthonormale.

Démonstration. Triviale. □

Théorème Procédé d'orthonormalisation de Gram-Schmidt

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}_{>0}$ et $\{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ une base de E.

Il existe une base orthonormale $\{\overrightarrow{f_1}, \dots, \overrightarrow{f_n}\}$ de E tel que pour tout i, $\overrightarrow{f_i} \in \mathtt{Vect}\left(\overrightarrow{e_1}, \dots, \overrightarrow{e_i}\right)$.

Démonstration. D'après la proposition précédente, il suffit de trouver une base orthogonale, il suffira alors de la normaliser pour conclure.

On construit itérativement les vecteurs $\overrightarrow{f_i}$.

- On prend $\overrightarrow{f_1} = \overrightarrow{e_1}$.
- On prend $\overrightarrow{f_2} = \overrightarrow{e_2} - \dfrac{\left\langle \overrightarrow{f_1} \mid \overrightarrow{e_2} \right\rangle}{\left\|\overrightarrow{f_1}\right\|^2} \overrightarrow{f_1}$. Alors par construction et par bilinéarité on a

$$\left\langle \overrightarrow{f_1} \mid \overrightarrow{f_2} \right\rangle = \left\langle \overrightarrow{e_1} \mid \overrightarrow{e_2} \right\rangle - \frac{\left\langle \overrightarrow{f_1} \mid \overrightarrow{e_2} \right\rangle}{\left\|\overrightarrow{f_1}\right\|^2} \left\langle \overrightarrow{f_1} \mid \overrightarrow{f_1} \right\rangle = 0$$

.

- On prend $\overrightarrow{f_3} = \overrightarrow{e_3} - \frac{\left\langle \overrightarrow{f_1} \mid \overrightarrow{e_3} \right\rangle}{\left\| \overrightarrow{f_1} \right\|^2} \overrightarrow{f_1} - \frac{\left\langle \overrightarrow{f_2} \mid \overrightarrow{e_3} \right\rangle}{\left\| \overrightarrow{f_2} \right\|^2} \overrightarrow{f_2}$. On procédant comme précédemment on montre sans peine que $\left\langle \overrightarrow{f_1} \mid \overrightarrow{f_3} \right\rangle = \left\langle \overrightarrow{f_2} \mid \overrightarrow{f_3} \right\rangle = 0$
- $\cdots$
- On prend
$$\overrightarrow{f_p} = \overrightarrow{e_p} - \sum_{k=1}^{p-1} \frac{\left\langle \overrightarrow{f_k} \mid \overrightarrow{e_p} \right\rangle}{\left\| \overrightarrow{f_k} \right\|^2} \overrightarrow{f_k}$$
et on vérifie sans peine l'orthogonalité.

□

Par exemple considérons la base de l'espace euclidien canonique $\mathbb{R}^3$ formé des vecteurs $\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\overrightarrow{e_2} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ et $\overrightarrow{e_3} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et appliquons le procédé de Gram-Schmidt.

- On pose $\overrightarrow{f_1} = \overrightarrow{e_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
- On a $\left\| \overrightarrow{f_1} \right\|^2 = 3$, $\left\langle \overrightarrow{f_1} \mid \overrightarrow{e_2} \right\rangle = 4$ et $\left\langle \overrightarrow{f_1} \mid \overrightarrow{e_3} \right\rangle = 2$
- On pose $\overrightarrow{f_2} = \overrightarrow{e_2} - \frac{4}{3} \overrightarrow{f_1} = \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}$
- On a $\left\| \overrightarrow{f_2} \right\|^2 = \frac{2}{3}$, $\left\langle \overrightarrow{f_2} \mid \overrightarrow{e_2} \right\rangle = \frac{2}{3}$ et $\left\langle \overrightarrow{f_2} \mid \overrightarrow{e_3} \right\rangle = \frac{1}{3}$
- On pose $\overrightarrow{f_3} = \overrightarrow{e_3} - \frac{2}{3} \overrightarrow{f_1} - \frac{\frac{1}{3}}{\frac{2}{3}} \overrightarrow{f_2} = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$

Par construction ces trois vecteurs formes une base orthogonale. Il suffit des les normaliser pour conclure : une base orthonormale de $\mathbb{R}^3$ est

$$\left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{\sqrt{3}}{\sqrt{2}} \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}, \sqrt{2} \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right\}$$

Exercice

Considérons l'espace vectoriel $\mathbb{R}^3$ munit de sa structure euclidienne canonique. On considère les vecteurs $\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\overrightarrow{e_2} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $\overrightarrow{e_3} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$. Vérifiez que la famille $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}\}$ est une base de $\mathbb{R}^3$ et appliquer le procédé d'orthonormalisation de Gram-Schmidt pour en extraire une base orthonormale.

Corollaire

Tout espace euclidien distinct de $\{\overrightarrow{0}\}$ admet une base orthonormale.

Démonstration. Tout espace vectoriel de dimension fini admet une base d'après le théorème de la base incomplète ; le procédé d'orthonormalisation de Gram-Schmidt permet de conclure. □

Proposition

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}_{>0}$ et $\mathcal{B} = \{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ une base orthonormale.

(i). $\forall \overrightarrow{x} \in E, \ \overrightarrow{x} = \sum_{i=1}^{n} \langle \overrightarrow{x} \mid \overrightarrow{e_i} \rangle \overrightarrow{e_i}$

En d'autre terme les coordonnés de $\overrightarrow{x}$ dans la base $\mathcal{B}$ sont $(\langle \overrightarrow{x} \mid \overrightarrow{e_i} \rangle)_i$

(ii). Soient $(x_i)_i$ et $(y_i)_i$ les coordonnées de vecteurs $\overrightarrow{x}$ et $\overrightarrow{y}$ de E exprimés dans la base $\mathcal{B}$, alors

$$\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \sum_{i=1}^{n} x_i y_i$$

(iii). Soit $(x_i)_i$ les coordonnées d'un vecteur $\overrightarrow{x}$ de E exprimés dans la base $\mathcal{B}$, alors

$$\|\overrightarrow{x}\|^2 = \sum_{i=1}^{n} x_i^2$$

Démonstration. Exercice □

Théorème

Soit F un sous espace vectoriel de dimesnion p d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$ de dimension $n \geqslant p$.

(i). $\dim(F^\perp) = n - p$

(ii). $F \overset{\perp}{\oplus} F^\perp = E$

(iii). $(F^\perp)^\perp = F$

Démonstration.

Soit $\{\overrightarrow{e_1}, \dots, \overrightarrow{e_p}\}$ une base de F. En appliquant le procédé de Gram-Schmidt, on déterminer une base orthonormale $\left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}\right\}$ de F. D'après le théorème de la base incomplète, on trouve une base $\left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}, \overrightarrow{e_{p+1}}, \dots, \overrightarrow{e_n}\right\}$ de E. Si on applique l'algorithme de Gram-Schmidt à cette base, on trouve une base $\left\{\overrightarrow{f_1}', \dots, \overrightarrow{f_p}', \overrightarrow{f_{p+1}}, \dots, \overrightarrow{f_n}\right\}$. Mais d'après le processus itératif, puisque $\left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}\right\}$ est déjà orthonormale on en déduit que $\overrightarrow{f_i}' = \overrightarrow{f_i}$.

En conclusion $\left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}, \overrightarrow{f_{p+1}}, \dots, \overrightarrow{f_n}\right\}$ est une base orthonormale de E et $\left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}\right\}$ est une base orthonormale de F. Montrons que $\left\{\overrightarrow{f_{p+1}}, \dots, \overrightarrow{e_n}\right\}$ est une base de $F^\perp$. Ceci prouvera que $\dim(F^\perp) = n - p$ et aussi que $F \overset{\perp}{\oplus} F^\perp = E$.

Naturellement c'est une famille libre, il suffit de montrer que c'est une famille génératrice. Soit $\overrightarrow{x} \in F^\perp$, d'après la proposition précédente, $\overrightarrow{x} = \sum_{i=1}^{n} \left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle \overrightarrow{f_i}$ mais puisque $\overrightarrow{x}$ est orthogonale à F, on en déduit que $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle = 0$ pour $i \in [\![1; p]\!]$.

Finalement il reste à observer que $(F^\perp)^\perp = F$. Nous avons déjà observer que $F \subseteq (F^\perp)^\perp$. Montrons l'inclusion inverse.

Soit $\overrightarrow{x} \in \left(F^{\perp}\right)^{\perp}$. Puisque $\overrightarrow{x}$ est aussi un vecteur de E, on peut, comme précédemment, écrire $\overrightarrow{x} = \sum_{i=1}^{n} \left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle \overrightarrow{f_i}$. Puisque $\overrightarrow{x}$ est un vecteur de l'orthogonale à l'orthogonale de F on en déduit que $\left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle = 0$ pour $i \in [\![p+1, n]\!]$. CQFD. $\square$

4. Projection

Définition

Soit E un espace vectoriel et $p : E \to E$ une application linéaire. On dira que p est un **projecteur** si $p \circ p = p$.

En d'autre terme, pour tout $\vec{x} \in E$, $p(p(x)) = p(x)$.

Exercice

Dans $\mathbb{R}^2$, vérifier que l'application défini par $p\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ 0 \end{pmatrix}$ est une application linéaire et un projecteur.

Lemme : Soit p un projecteur d'un espace vectoriel E et $\vec{y} \in \mathrm{Im}(p)$ alors $p(\vec{y}) = \vec{y}$.

Démonstration. Puisque $\vec{y} \in \mathrm{Im}(p)$, il existe $\vec{x} \in E$ tel que $\vec{y} = p(\vec{x})$. Composons par p pour obtenir : $p(\vec{y}) = p(p(\vec{x}))$ mais puisque p est un projecteur $p(p(\vec{x})) = p(\vec{x}) = \vec{y}$. □

Théorème

Soient $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien et $p : E \to E$ un projecteur. Les conditions suivantes sont équivalentes.

(i). $E = \mathrm{Ker}(p) \overset{\perp}{\oplus} \mathrm{Im}(p)$

(ii). $\forall \vec{x}_1, \vec{x}_2 \in E,\ \langle p(\vec{x}_1) \mid \vec{x}_2 \rangle = \langle \vec{x}_1 \mid p(\vec{x}_2) \rangle$

Sous ces conditions on dit que p est une **projection orthogonale**.

Démonstration.

(i) $\Rightarrow$ (ii). Puisque $E = \mathrm{Ker}(p) \overset{\perp}{\oplus} \mathrm{Im}(p)$ on a $\vec{x}_i = \vec{e_i} + \vec{f_i}$ pour $\vec{e_i} \in \mathrm{Ker}(p)$ et $\vec{f_i} \in \mathrm{Im}(p)$. Alors, en utilisant le lemme on a $p(\vec{x}_i) = p(\vec{e_i} + \vec{f_i}) = p(\vec{e_i}) + p(\vec{f_i})$ par linéarité de p et $p(\vec{e_i}) = \vec{0}$ donc $p(\vec{x}_i) = p(\vec{f_i}) = \vec{f_i}$ d'après le lemme précédent. Alors puisque nous supposons que le noyau et l'image sont orthogonaux, on a d'une part $\langle p(\vec{x}_1) \mid \vec{x}_2 \rangle = \left\langle \vec{f_1} \,\middle|\, \vec{e_2} + \vec{f_2} \right\rangle = \left\langle \vec{f_1} \,\middle|\, \vec{e_2} \right\rangle + \left\langle \vec{f_1} \,\middle|\, \vec{f_2} \right\rangle = \left\langle \vec{f_1} \,\middle|\, \vec{f_2} \right\rangle$ et d'autre part $\langle \vec{x}_1 \mid p(\vec{x}_2) \rangle = \left\langle \vec{e_1} + \vec{f_1} \,\middle|\, \vec{f_2} \right\rangle = \left\langle \vec{e_1} \,\middle|\, \vec{f_2} \right\rangle + \left\langle \vec{f_1} \,\middle|\, \vec{f_2} \right\rangle = \left\langle \vec{f_1} \,\middle|\, \vec{f_2} \right\rangle$.

(ii) $\Rightarrow$ (i). Soit $\vec{e} \in \mathrm{Ker}(p)$ et $\vec{y} \in \mathrm{Im}(p)$, Montrons que $\langle \vec{e} \mid \vec{y} \rangle = 0$. Par définition de l'image, il existe $\vec{x} \in E$ tel que $\vec{y} = p(\vec{x})$ alors

$$\langle \vec{e} \mid \vec{y} \rangle = \langle \vec{e} \mid p(\vec{x}) \rangle = \langle p(\vec{e}) \mid \vec{x} \rangle = \left\langle \vec{0} \,\middle|\, \vec{x} \right\rangle = 0$$

Il suffit maintenant de vérifier que tout vecteur de E s'écrit comme la somme d'un vecteur de $\mathrm{Ker}(p)$ et d'un vecteur de $\mathrm{Im}(p)$. On pose $\vec{x} = (\vec{x} - p(\vec{x})) + p(\vec{x})$. Naturellement $p(\vec{x}) \in \mathrm{Im}(p)$ et on observe que $p(\vec{x} - p(\vec{x})) = p(\vec{x}) - p(p(\vec{x})) = p(\vec{x}) - p(\vec{x}) = \vec{0}$. CQFD.

□

Exercice

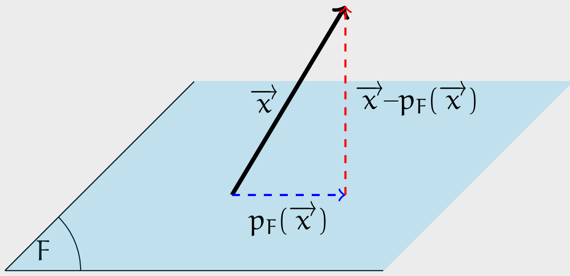
On munit $\mathbb{R}^2$ de sa structure euclidienne canonique. On considère l'application $p\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x+y}{2} \end{pmatrix}$.

S'agit-il d'une projection orthogonale ?

Corollaire

Soit $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien et F un sous espace vectoriel de E. Il existe une unique projection orthogonale p_F tel que $\operatorname{Im}(p_F) = F$ et $\operatorname{Ker}(p_F) = F^{\perp}$.

En particulier, pour tout $\overrightarrow{x} \in E$, $\overrightarrow{x} - p_F(\overrightarrow{x}) \in F^{\perp}$



Démonstration. Nous avons vu que $E = F \overset{\perp}{\oplus} F^{\perp}$, c'est à dire que pour tout vecteur $\overrightarrow{x} \in E$ il existe une unique décomposition $\overrightarrow{x} = \overrightarrow{f} + \overrightarrow{e}$ où $\overrightarrow{f} \in F$ et $\overrightarrow{e} \in F^{\perp}$. On pose $p_F(\overrightarrow{x}) = \overrightarrow{f}$. Par construction et unicité de la décomposition, c'est clairement une application linéaire. C'est aussi trivialement un projecteur puisque si $\overrightarrow{f} \in F$ alors $p_F(\overrightarrow{f}) = \overrightarrow{f}$. Par construction son image est F et d'après le théorème précédent, son noyau est $\operatorname{Ker}(p_F) = \operatorname{Im}(p_F)^{\perp} = F^{\perp}$. □

Proposition

Soient $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien et F un sous-espace vectoriel de dimension p. Soit $\mathcal{B} = \left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}\right\}$ une base orthonormale de F. Alors

$$\forall \overrightarrow{x} \in E,\ p_F(\overrightarrow{x}) = \sum_{i=1}^{p} \left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle \overrightarrow{f_i}$$

Démonstration. On peut compléter la base $\mathcal{B}$ en une base orthornormale $\mathcal{B}' = \left\{\overrightarrow{f_1}, \dots, \overrightarrow{f_p}, \overrightarrow{f_{p+1}}, \dots, \overrightarrow{f_n}\right\}$ de E de sorte que $\texttt{Vect}\left(\overrightarrow{f_{p+1}}, \dots, \overrightarrow{f_n}\right) = F^{\perp}$ (conséquence de l'égalité $E = F \overset{\perp}{\oplus} F^{\perp}$ et du procédé de Gram-Schmidt). Puisque $p_F(\overrightarrow{f_i}) = \overrightarrow{f_i}$ pour tout $i \in [\![1;p]\!]$ et $p_F(\overrightarrow{f_i}) = \overrightarrow{0}$ pour tout $i \in [\![p+1;n]\!]$ et puisque $\overrightarrow{x} = \sum_{i=1}^{n} \left\langle \overrightarrow{x} \,\middle|\, \overrightarrow{f_i} \right\rangle \overrightarrow{f_i}$, la linéarité de p permet de conclure. □

Cette proposition permet de déterminer l'expression de n'importe quelle projection orthogonale.

Par exemple, considérons $\mathbb{R}^3$ munit de sa structure euclidienne classique et considérons le plan $P = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x + y = 0 \right\}$. On laisse le soin au lecteur de vérifier qu'il s'agit bien d'un sous-espace vectoriel de $\mathbb{R}^3$ dont une base est $\texttt{Vect}\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$.

Appliquons à cette base le procédé de Gram-Schmidt. On pose $\overrightarrow{e_1} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $\overrightarrow{e_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \dfrac{\left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \,\middle|\, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\|^2} =$

$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Surprise, la base est déjà orthogonale. La base orthonormée est alors $\overrightarrow{f_1} = \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $\overrightarrow{f_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

D'après le théorème, on a

$$\begin{aligned} p_F\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) &= \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \frac{x-y}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{x-y}{2} \\ -\frac{x-y}{2} \\ z \end{pmatrix} \end{aligned}$$

Exercice

On se place dans $\mathbb{R}^2$ munit de sa structure euclidienne classique. Soit $\overrightarrow{u} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ et $\overrightarrow{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. On pose $U = \mathtt{Vect}(\overrightarrow{u})$ et $V = \mathtt{Vect}(\overrightarrow{v})$

1. Déterminer l'expression de p_U et le projeté de $\overrightarrow{v}$ sur U.
2. Déterminer l'expression de p_V et le projeté de $\overrightarrow{u}$ sur V.
3. Illustrer la situation dans le plan.

Exercice

Soit $\mathbb{R}^3$ munit de sa structure euclidienne canonique. On considère $P = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \middle| x + y = 0 \right\}$.

1. Montrer que P est un sous-espace vectoriel de $\mathbb{R}^3$.
2. Déterminer $P^\perp$ ainsi qu'une base.
3. Donner l'expression de $p_{P^\perp}$.

Exercice

Soit $\mathbb{R}^3$ munit de sa structure euclidienne canonique. Soient $\overrightarrow{u} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $\overrightarrow{v} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ et $\overrightarrow{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

1. Montrer que $\{\overrightarrow{u}, \overrightarrow{v}\}$ est une famille libre de $\mathbb{R}^3$.
2. Donner une base orthonormale de $F = \mathtt{Vect}(\overrightarrow{u}, \overrightarrow{v})$.
3. Donner l'expression de p_F et l'image $p_F(\overrightarrow{w})$

§ Il peut être plus commode de considérer le procédé d'orthonormalisation de Gram-Schmidt à l'aide de projections orthogonales.

Soit $\mathcal{B} = \left\{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\right\}$ une base de E.

L'algorithme s'exprime alors comme :

- On pose $\overrightarrow{f_1} = \dfrac{\overrightarrow{e_1}}{\left\|\overrightarrow{e_1}\right\|}$ et $F_1 = \mathtt{Vect}\left(\overrightarrow{f_1}\right)$.
- On pose $\overrightarrow{f_2}' = \overrightarrow{e_2} - p_{F_1}(\overrightarrow{e_2})$, $\overrightarrow{f_2} = \dfrac{\overrightarrow{f_2}'}{\left\|\overrightarrow{f_2}'\right\|}$ et $F_2 = \mathtt{Vect}\left(\overrightarrow{f_1}, \overrightarrow{f_2}\right)$
- On pose $\overrightarrow{f_3}' = \overrightarrow{e_3} - p_{F_2}(\overrightarrow{e_3})$, $\overrightarrow{f_3} = \dfrac{\overrightarrow{f_3}'}{\left\|\overrightarrow{f_3}'\right\|}$ et $F_3 = \mathtt{Vect}\left(\overrightarrow{f_1}, \overrightarrow{f_2}, \overrightarrow{f_3}\right)$.
- $\cdots$
- On pose $\overrightarrow{f_k}' = \overrightarrow{e_k} - p_{F_{k-1}}(\overrightarrow{e_k})$, $\overrightarrow{f_k} = \dfrac{\overrightarrow{f_k}'}{\left\|\overrightarrow{f_k}'\right\|}$ et $F_k = \mathtt{Vect}\left(\overrightarrow{f_1}, \dots, \overrightarrow{f_k}\right)$.
- $\cdots$

Définition

Soient $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de norme associée $\|\bullet\|$, $\overrightarrow{v} \in E$ et F un sous-espace vectoriel de E. On note

$$d(\overrightarrow{v}, F) = \underset{\overrightarrow{f} \in F}{\mathrm{Inf}} \left(\left\|\overrightarrow{v} - \overrightarrow{f}\right\|\right)$$

Ainsi $d(\overrightarrow{v}, F)$ représente la plus courte distance entre le vecteur $\overrightarrow{v}$ et l'espace F.

Théorème Meilleure approximation

Soient $(E, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de norme associée $\|\bullet\|$, F un sous-espace vectoriel de E, p_F la projection orthogonale sur F et $\overrightarrow{v} \in E$.

$$\forall \overrightarrow{f} \in F,\ \left\|\overrightarrow{v} - \overrightarrow{f}\right\| \geqslant \left\|\overrightarrow{v} - p_F(\overrightarrow{v})\right\|$$

En particulier $d(\overrightarrow{v}, F) = \left\|\overrightarrow{v} - p_F(\overrightarrow{v})\right\|$.

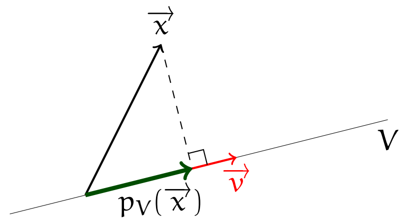
Démonstration. On sait que F et $F^\perp$ sont orthogonaux, de plus on sait que $\overrightarrow{v} - p_F(\overrightarrow{v}) \in F^\perp$, enfin on sait que $\overrightarrow{f} - p_F(\overrightarrow{v}) \in F$.

En conclusion $\overrightarrow{v} - p_F(\overrightarrow{v})$ et $\overrightarrow{f} - p_F(\overrightarrow{v})$ sont deux vecteur orthogonaux. D'après le théorème de Pythagore

$$\left\|\overrightarrow{v} - \overrightarrow{f}\right\|^2 = \left\|(\overrightarrow{v} - p_F(\overrightarrow{v})) + \left(p_F(\overrightarrow{v}) - \overrightarrow{f}\right)\right\|^2 = \left\|\overrightarrow{v} - p_F(\overrightarrow{v})\right\|^2 + \left\|p_F(\overrightarrow{v}) - \overrightarrow{f}\right\|^2 \geqslant \left\|\overrightarrow{v} - p_F(\overrightarrow{v})\right\|^2$$

$\square$

5. Angles



Considérons un vecteur $\vec{v}$ d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$. Notons $V = \mathtt{Vect}(\vec{v})$ la droite vectoriel dirigée par le vecteur $\vec{v}$. Comme il n'y a qu'un seul vecteur, il s'agit trivialement d'une base orthogonale. D'après le précédent chapitre, on peut exprimer la projection orthogonale sur V. Précisément

$$\forall \vec{x} \in E,\ p_V(\vec{x}) = \left\langle \vec{x} \,\middle|\, \frac{\vec{v}}{\|\vec{v}\|} \right\rangle \frac{\vec{v}}{\|\vec{v}\|} = \frac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v}$$

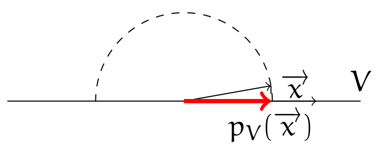
Calculons la taille du vecteur $p_V(\vec{x})$.

$$\begin{aligned}
\|p_V(\vec{x})\|^2 &= \langle p_V(\vec{x}) \mid p_V(\vec{x}) \rangle \\
&= \left\langle \frac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \,\middle|\, \frac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} \right\rangle \\
&= \left(\frac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{v}\|^2} \right)^2 \langle \vec{v} \mid \vec{v} \rangle \\
&= \left(\frac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{v}\|^2} \right)^2 \|\vec{v}\|^2 \\
&= \frac{\langle \vec{x} \mid \vec{v} \rangle^2}{\|\vec{v}\|^2}
\end{aligned}$$

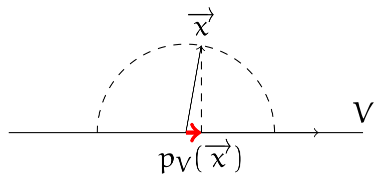
Ainsi $\|p_V(\vec{x})\| = \dfrac{|\langle \vec{x} \mid \vec{v} \rangle|}{\|\vec{v}\|}$, soit encore $\|p_V(\vec{x})\| = \|\vec{x}\| \times \left(\dfrac{|\langle \vec{x} \mid \vec{v} \rangle|}{\|\vec{x}\| \, \|\vec{v}\|} \right)$. Autrement dit, $\dfrac{|\langle \vec{x} \mid \vec{v} \rangle|}{\|\vec{x}\| \, \|\vec{v}\|}$ mesure la variation entre la taille de $p_V(\vec{x})$ et la taille de $\vec{x}$.

Observons cette variation.

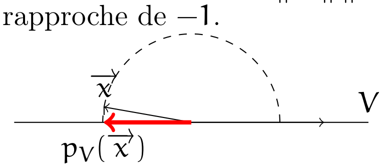
Lorsque $\vec{x}$ se rapproche de V, dans le sens du vecteur directeur $\vec{v}$, $\vec{x}$ et $p_V(\vec{x})$ sont presque confondu donc $\dfrac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{x}\| \, \|\vec{v}\|}$ se rapproche de 1.



Lorsque $\vec{x}$ se rapproche de l'orthogonale de V, $p_V(\vec{x})$ est proche de l'origine donc $\dfrac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{x}\| \, \|\vec{v}\|}$ se rapproche de 0.



Lorsque $\vec{x}$ se de V mais dans le sens opposé à $\vec{v}$, $\vec{x}$ et $p_V(\vec{x})$ sont presque confondu donc $\dfrac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{x}\| \, \|\vec{v}\|}$ se rapproche de -1.



Ces observations nous amène à penser que $\dfrac{\langle \vec{x} \mid \vec{v} \rangle}{\|\vec{x}\| \, \|\vec{v}\|}$ dépend uniquement de l'*écartement* entre $\vec{x}$ et $\vec{v}$. Ceci motive la définition suivante.

Définition

Soient $\vec{x}$ et $\vec{y}$ deux vecteurs non nuls d'un espace euclidien $(E, \langle \vec{x} \mid \vec{y} \rangle)$. Le **cosinus** entre $\vec{x}$ et $\vec{y}$, noté $\cos(\vec{x}, \vec{y})$ est défini par la formule

$$\cos(\vec{x}, \vec{y}) = \frac{\langle \vec{x} \mid \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|}$$

Ce cosinus correspond, dans le plan, au cosinus bien connu des lycéens, ce que permet d'observer l'exercice suivant.

Exercice

On munit $\mathbb{R}^2$ de sa structure euclidienne canonique. Soit $\overrightarrow{e} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $E = \mathtt{Vect}(\overrightarrow{e})$, $M = \begin{pmatrix} x \\ y \end{pmatrix}$ un point du plan distinct de l'origine et H le projeté orthogonal de M sur l'axe des abscisses.

1. Donner l'expression de p_E.
2. Donnez les coordonnées du points H.
3. En déduire les distances OM et OH.
4. Montrer que $\cos\left(\overrightarrow{e}, p_E\left(\overrightarrow{OM}\right)\right) = \dfrac{OH}{OM}$ (on retrouve bien la fameuse formule *cosinus = coté adjacent sur hypoténuse*).

Proposition

Soient $\overrightarrow{x}$ et $\overrightarrow{y}$ deux vecteurs non nuls d'un espace euclidien $(E, \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle)$.

(i). $\cos(\overrightarrow{x}, \overrightarrow{y}) = \cos(\overrightarrow{y}, \overrightarrow{x})$

(ii). $\forall \lambda \in \mathbb{R}^*, \ \cos(\lambda\overrightarrow{x}, \overrightarrow{y}) = \mathrm{sg}(\lambda)\cos(\overrightarrow{x}, \overrightarrow{y})$

(iii). $\cos(\overrightarrow{x}, \overrightarrow{y}) = 0 \iff \overrightarrow{x} \perp \overrightarrow{y}$

(iv). $\cos(\overrightarrow{x}, \overrightarrow{y}) \in [-1, 1]$.

(v). $|\cos(\overrightarrow{x}, \overrightarrow{y})| = 1 \iff \exists \lambda \in \mathbb{R}^*, \ \overrightarrow{x} = \lambda\overrightarrow{y}$

($\mathrm{sg}(\lambda)$ désigne le signe de λ)

Démonstration. Exercice □

Théorème

Soit $\overrightarrow{v}$ un vecteur de norme 1 d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$ et $V = \mathtt{Vect}(\overrightarrow{v})$. Alors

$$\forall \overrightarrow{x} \in E - \{\overrightarrow{0}\}, \ p_V(\overrightarrow{x}) = \cos(\overrightarrow{x}, \overrightarrow{v}) \left\|\overrightarrow{x}\right\| \overrightarrow{v}$$

Démonstration. adapté dans le cas où $\overrightarrow{v}$ n'est pas nécessairement normé. $p_V(\overrightarrow{x}) = \dfrac{\langle \overrightarrow{x} \mid \overrightarrow{v} \rangle}{\left\|\overrightarrow{v}\right\|^2}\overrightarrow{v} = \dfrac{\langle \overrightarrow{x} \mid \overrightarrow{v} \rangle}{\left\|\overrightarrow{x}\right\| \left\|\overrightarrow{v}\right\|^2} \left\|\overrightarrow{x}\right\| \overrightarrow{v} = \dfrac{\langle \overrightarrow{x} \mid \overrightarrow{v} \rangle}{\left\|\overrightarrow{x}\right\| \left\|\overrightarrow{v}\right\|} \left\|\overrightarrow{x}\right\| \dfrac{\overrightarrow{v}}{\left\|\overrightarrow{v}\right\|} = \cos(\overrightarrow{x}, \overrightarrow{v}) \left\|\overrightarrow{x}\right\| \dfrac{\overrightarrow{v}}{\left\|\overrightarrow{v}\right\|}$ □

6. Matrices

Souvent en algèbre (bi)linéaire (en dimension finie) on choisi de travailler avec les matrices. Cependant les matrices impliquent le choix d'une base.

Si $\mathcal{B}$ est une base d'un espace vectoriel E et $\overrightarrow{e} \in \mathsf{E}$, on notera $\overrightarrow{e_{\mathcal{B}}}$ l'expression de $\overrightarrow{e}$ dans la base $\mathcal{B}$.

On rappel qu'étant donné deux bases $\mathcal{B}$ et $\mathcal{B}'$, on note $\mathsf{Pass}(\mathcal{B}, \mathcal{B}')$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. En particulier on a $P\overrightarrow{e_{\mathcal{B}'}} = \overrightarrow{e_{\mathcal{B}}}$.

Définition

Soient $(\mathsf{E}, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}_{>0}$ et $\mathcal{B} = \{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ une base de E. On appel **matrice du produit scalaire** par rapport à $\mathcal{B}$, la matrice M où

$$\forall i, j \in [\![1; n]\!],\ M_{i,j} = \langle \overrightarrow{e_i} \mid \overrightarrow{e_j} \rangle$$

Considérons par exemple $\mathbb{R}^2$ munit de sa structure euclidienne canonique. On laisse le soin au lecteur de vérifier que $\mathcal{B} = \left\{ \overrightarrow{e_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \overrightarrow{e_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ forment une base de $\mathbb{R}^2$. Alors la matrice du produit scalaire dans cette base est $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Exercice

Considérons $\mathbb{R}^2$ munit de sa structure euclidienne canonique. Déterminer la matrice du produit scalaire dans les bases suivantes après avoir montré qu'il s'agit bien de basse de $\mathbb{R}^2$

1. $\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$
2. $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$
3. $\mathcal{C} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$
4. $\mathcal{D} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$
5. $\mathcal{E} = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$

Proposition

Soit $n \in \mathbb{N}_{>0}$.La matrice du produit scalaire canonique dans $\mathbb{R}^n$ munit de sa structure euclidienne canonique est Id_n.

Démonstration. Exercice □

Théorème

Soient $(\mathsf{E}, \langle \bullet \mid \bullet \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}_{>0}$, $\mathcal{B} = \{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ une base de E et M la matrice du produit scalaire par rapport à $\mathcal{B}$.

$$\forall \overrightarrow{x}, \overrightarrow{y} \in \mathsf{E},\ \langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = {}^t\overrightarrow{x}_{\mathcal{B}} M \overrightarrow{y}_{\mathcal{B}}$$

Démonstration. Exercice □

Reprenons l'exemple précédent. On vérifie que $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2\vec{e_1} + (x_1 - x_2)\vec{e_2} = \begin{pmatrix} x_2 \\ x_1 - x_2 \end{pmatrix}_{\mathcal{B}}$.

On a alors

$$\begin{aligned}
{}^t\vec{x}_{\mathcal{B}} M \vec{y}_{\mathcal{B}} &= \begin{pmatrix} x_2 & x_1 - x_2 \end{pmatrix}_{\mathcal{B}} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_2 \\ y_1 - y_2 \end{pmatrix}_{\mathcal{B}} \\
&= \begin{pmatrix} x_2 & x_1 - x_2 \end{pmatrix}_{\mathcal{B}} \begin{pmatrix} 2y_2 + (y1 - y2) \\ y_2 + (y_1 - y_2) \end{pmatrix} \\
&= \begin{pmatrix} x_2 & x_1 - x_2 \end{pmatrix}_{\mathcal{B}} \begin{pmatrix} y_2 + y1 \\ y_1 \end{pmatrix} \\
&= x_2(y_2 + y_1) + (x_1 - x_2)y_1 \\
&= x_2y_2 + x_2y_1 + x_1y_1 - x_2y_1 \\
&= x_2y_2 + x_1y_1
\end{aligned}$$

On retrouve bien l'expression du produit scalaire canonique.

Proposition

Soit M la matrice du produit scalaire dans une base $\mathcal{B}$ quelconque d'un espace vectoriel euclidien $(E, \langle \bullet \mid \bullet \rangle)$.

(i) **Symétrie.** ${}^tM = M$

(ii) **Positivité.** $\forall \vec{x} \in E,\ {}^t\vec{x}_{\mathcal{B}} A \vec{x}_{\mathcal{B}} \geqslant 0$

(iii) **Séparation.** $\forall \vec{x} \in E,\ {}^t\vec{x}_{\mathcal{B}} A \vec{x}_{\mathcal{B}} = 0 \iff \vec{x}_{\mathcal{B}} = \vec{0}_{\mathcal{B}}$

(iv) **Inversibilité.** $M \in GL_n(\mathbb{R})$.

Démonstration. Hormis la dernière proposition, toutes les formules sont l'expression matricielle des axiomes d'un produit scalaire.

Pour démontrer que $M \in GL_n(\mathbb{R})$, c'est à dire que M est inversible, il suffit de montrer qu'elle définie une application linéaire injective (c'est une conséquence du théorème du rang). Soit $\vec{x}_{\mathcal{B}} \in \mathrm{Ker}(M)$ alors $M\vec{x}_{\mathcal{B}} = \vec{0}$ donc $0 = {}^t\vec{x}_{\mathcal{B}} M \vec{x}_{\mathcal{B}} = \langle \vec{x} \mid \vec{x} \rangle = \|\vec{x}\|^2$ et donc $\vec{x} = \vec{0}$ par la propriété de la séparation. $\square$

Intéressons nous aux changement de base.

Lemme : Soit $A, B \in \mathrm{Mat}_n(\mathbb{R})$.

$$A = B \iff \left(\forall \vec{x}, \vec{y} \in E,\ {}^t\vec{x}_{\mathcal{B}} A \vec{y}_{\mathcal{B}} = {}^t\vec{x}_{\mathcal{B}} B \vec{y}_{\mathcal{B}}\right)$$

Démonstration. Le sens $\Rightarrow$ est trivial, il suffit de montrer $\Leftarrow$ mais cela découle de l'observation que ${}^t\vec{e_i} A \vec{e_j} = A_{i,j}$ pour $\vec{e_k}$ les vecteurs de la base la base canonique. $\square$

Théorème

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un espace euclidien $(E, \langle \bullet \mid \bullet \rangle)$, $P = \mathrm{Pass}(\mathcal{B}, \mathcal{B}')$, $M_{\mathcal{B}}$ la matrice du produit scalaire par rapport à la base $\mathcal{B}$ et $M_{\mathcal{B}'}$ la matrice du produit scalaire par rapport à la base $\mathcal{B}'$.

$$M_{\mathcal{B}'} = {}^tPM_{\mathcal{B}}P$$

Démonstration. Soient $\vec{x}, \vec{y} \in E$. On rappel que $P\vec{x}_{\mathcal{B}'} = \vec{x}_{\mathcal{B}}$

$$\langle \overrightarrow{x} \mid \overrightarrow{y} \rangle = \langle \overrightarrow{x}_{\mathcal{B}} \mid \overrightarrow{y}_{\mathcal{B}} \rangle = \langle \overrightarrow{x}_{\mathcal{B}'} \mid \overrightarrow{y}_{\mathcal{B}'} \rangle \implies {}^t\overrightarrow{x}_{\mathcal{B}} M_{\mathcal{B}} \overrightarrow{y}_{\mathcal{B}} = {}^t\overrightarrow{x}_{\mathcal{B}'} M_{\mathcal{B}'} \overrightarrow{y}_{\mathcal{B}'}$$
$$\implies {}^t(P\overrightarrow{x}_{\mathcal{B}'}) M_{\mathcal{B}} (P\overrightarrow{y}_{\mathcal{B}'}) = {}^t\overrightarrow{x}_{\mathcal{B}'} M_{\mathcal{B}'} \overrightarrow{y}_{\mathcal{B}'}$$
$$\implies {}^t\overrightarrow{x}_{\mathcal{B}'} {}^tP M_{\mathcal{B}} P \overrightarrow{y}_{\mathcal{B}'} = {}^t\overrightarrow{x}_{\mathcal{B}'} M_{\mathcal{B}'} \overrightarrow{y}_{\mathcal{B}'}$$
$$\implies {}^t\overrightarrow{x}_{\mathcal{B}'} \left({}^tP M_{\mathcal{B}} P\right) \overrightarrow{y}_{\mathcal{B}'} = {}^t\overrightarrow{x}_{\mathcal{B}'} M_{\mathcal{B}'} \overrightarrow{y}_{\mathcal{B}'}$$

On conclut avec le précédent lemme. □

Définition

On dira qu'une matrice $A \in \mathrm{Mat}_n(\mathbb{R})$ est **orthogonale** si ${}^tAA = \mathrm{Id}_n$.

Proposition

(i). Si A est une matrice orthogonale $\mathtt{det}(A) = \pm 1$

(ii). Si A est une matrice orthogonale alors A est inversible et $A^{-1} = {}^tA$.

(iii). Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthogonales d'un espace euclidien alors $\mathrm{Pass}(\mathcal{B}, \mathcal{B}')$ est orthogonale.

Démonstration.

(i). Par définition on a ${}^tAA = \mathrm{Id}_n$ alors $\mathtt{det}\left({}^tAA\right) = \mathtt{det}\left(\mathrm{Id}_n\right)$ soit encore $\mathtt{det}\left({}^tA\right)\mathtt{det}\left(A\right) = 1$ et donc $\mathtt{det}(A)^2 = 1$ puisque une matrice et sa transposée ont le même déterminant.

(ii) En particulier puisque ce determinant est non nul alors la matrice est inversible. L'unicité de l'inverse prouve que $A^{-1} = {}^tA$.

(iii). La matrice d'un produit scalaire par rapport à une base orthonormé est l'identité. Le théorème précédent donne alors $\mathrm{Id}_n = {}^tP\mathrm{Id}_nP = {}^tPP$.

□

Exercice

Déterminer des réels a, b et c pour que $\begin{pmatrix} \frac{1}{a} & b \\ \frac{1}{a} & c \end{pmatrix}$ soit orthogonale.

Lemme : Toute matrice symétrique réelle admet au moins une valeur propre réelle.

Démonstration. D'après le théorème de de Gauss d'Alembert, le polynôme caractéristique $\chi_A(X) = \mathtt{det}(A - X\mathrm{Id})$ d'un matrice symétrique A admet au moins une solution complexe. Soit λ une solution. Montrons que λ est réel. Soit $\overrightarrow{x}$ un vecteur propre (éventuellement complexe).

$$\begin{array}{rcll}
A\overrightarrow{x} & = & \lambda\overrightarrow{x} & \\
\overline{A\overrightarrow{x}} & = & \overline{\lambda\overrightarrow{x}} & \text{On conjugue} \\
{}^t\overline{\overrightarrow{x}}\,{}^t\overline{A} & = & \overline{\lambda}\,{}^t\overline{\overrightarrow{x}} & \text{On transpose} \\
{}^t\overline{\overrightarrow{x}}\,{}^t\overline{A}\overrightarrow{x} & = & \overline{\lambda}\,{}^t\overline{\overrightarrow{x}}\overrightarrow{x} & \text{On multiplie à droite par } \overrightarrow{x} \\
{}^t\overline{\overrightarrow{x}}A\overrightarrow{x} & = & \overline{\lambda}\,{}^t\overline{\overrightarrow{x}}\overrightarrow{x} & \text{La matrice } A \text{est symétrique et réelle} \\
{}^t\overline{\overrightarrow{x}}\lambda\overrightarrow{x} & = & \overline{\lambda}\,{}^t\overline{\overrightarrow{x}}\overrightarrow{x} & \text{Le vecteur } \overrightarrow{x} \text{ est un vecteur propre de la valeur propre } \lambda \\
\lambda\left\|\overrightarrow{x}\right\|^2 & = & \overline{\lambda}\left\|\overrightarrow{x}\right\|^2 & \text{On observe que } {}^t\overline{\overrightarrow{x}}\overrightarrow{x} = \left\|\overrightarrow{x}\right\|^2 \\
\lambda & = & \overline{\lambda} & \text{On simplifie}
\end{array}$$

□

Théorème — Théorème spectrale

Toute matrice symétrique réelle est diagonalisable dans une base orthonormé.

Précisément, pour toute matrice symétrique A (${}^tA = A$), il existe une matrice orthogonale P (${}^tPP = \mathtt{Id}$) tel que ${}^tPAP$ soit une matrice diagonale.

Démonstration. Munissons $\mathbb{R}^n$ de sa structure d'espace euclidien canonique. En particulier la base canonique $\mathcal{B}$ est une base orthonormale.

D'après le lemme précédent, A possède une valeur propre réelle. Notons λ cette valeur propre et $\overrightarrow{x}$ un vecteur propre que l'on peut choisir normalisé. Notons $X = \mathtt{Vect}(\overrightarrow{x})$ et considérons, par l'intermédiaire du procédé de Gram-Schmidt, une base orthonormé $\mathcal{B}' = \{\overrightarrow{x}, \overrightarrow{f_1}, \overrightarrow{f_{n-1}}\}$ de $\mathbb{R}^n = X \overset{\perp}{\oplus} X^\perp$. Puisque les bases $\mathcal{B}$ et $\mathcal{B}'$ sont orthonormales, $P = \mathtt{Pass}(\mathcal{B}, \mathcal{B}')$ est orthogonale.

Soit B l'expression de la matrice A dans la base $\mathcal{B}'$ alors $B = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$. En effet, la première colonne est la définition de vecteur propre, tandis que pour la première ligne, cela vient du fait que $\left\langle A\overrightarrow{f_k} \,\middle|\, \overrightarrow{x} \right\rangle = {}^t\left(A\overrightarrow{f_k}\right)\overrightarrow{x} = {}^t\overrightarrow{f_k}{}^tA\overrightarrow{x} = {}^t\overrightarrow{f_k}A\overrightarrow{x} = {}^t\overrightarrow{f_k}\lambda\overrightarrow{x} = \lambda\,{}^t\overrightarrow{f_k}\overrightarrow{x} = \lambda\left\langle \overrightarrow{f_k} \,\middle|\, \overrightarrow{x} \right\rangle = \lambda 0 = 0$.

Puisque A est symétrique, il de va même pour B et naturellement aussi pour C.

En raisonnant par récurrence, on peut trouver une base orthogonale de vecteur propre pour C. □

Corollaire

Soit A une matrice symétrique de $\mathtt{Mat}_n(\mathbb{R})$ et $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ ses valeurs propres.

$$\forall \overrightarrow{x} \in \mathbb{R}^n,\ \lambda_1 \left\|\overrightarrow{x}\right\|^2 \leqslant {}^t\overrightarrow{x}A\overrightarrow{x} \leqslant \lambda_1 \left\|\overrightarrow{x}\right\|^2$$

Démonstration. D'après le théorème précédent, quitte à changer de base, on peut supposer que A est diagonale $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Notons x_i les coordonnées de $\overrightarrow{x}$ dans cette base orthonormé de diagonalisation. Alors ${}^t\overrightarrow{x}A\overrightarrow{x} = \sum_{i=1}^{n} \lambda_i x_i^2$. L'organisation des valeurs propres permet de conclure. □

Corollaire

Soit A une matrice symétrique de $\mathtt{Mat}_n(\mathbb{R})$ et $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ ses valeurs propres.

$$\operatorname*{Sup}_{\left\|\overrightarrow{x}\right\|=1} \left({}^t\overrightarrow{x}A\overrightarrow{x}\right) = \lambda_n$$

De plus le $\mathtt{Sup}$ est atteint pour les vecteurs de l'espace propre $\operatorname{Ker}(A - \lambda_n \mathtt{Id}_n)$

Démonstration. Triviale □

Exercice

1. Justifier que la matrice $A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}$ admet une base orthonormée de $\mathbb{R}^3$ comme vecteurs propres.

2. Monter que les valeurs propres de cette matrice sont 3, 6 et 9.
3. Déterminer une matrice P telle que ${}^tPAP$ soit diagonale.

Exercice

1. Justifier que la matrice $A = \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$ admet une base orthonormée de $\mathbb{R}^3$ comme vecteurs propres.
2. Monter que les valeurs propres de cette matrice sont 3 et -3.
3. Déterminer une base de $\mathbb{R}^3$ composée de vecteurs propres de A.

Exercice

Soit $E = \mathbb{R}^4$ est muni de sa structure euclidienne canonique.

On considère le sous-espaces F de E défini par $F = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in E \;\middle|\; x + z + t = 0 \wedge x - y + z = 0 \right\}$

Déterminer une base orthonormale de F

7. ACP

Il est de plus en plus fréquent pour le statisticien de se retrouver confronté à des jeux de données de grande dimension, dont il faudra extraire des informations pertinentes.

L'objectif d'une analyse en composantes principales (ACP) est d'obtenir une représentation approchée d'un nuage de points dans un sous-espace de dimension plus faible tout en perdant le moins d'information possible.

Les méthodes d'analyse factorielle contribuent également à améliorer la lisibilité des données, en mettant en évidence des **relations entre variables ou entre groupes de variables**, ainsi qu'en permettant de visualiser **les relations entre données ou entre variables projetées** dans l'espace de faible dimension des facteurs les plus significatifs.

Le point de départ d'une analyse factorielle est un jeu de données concernant n individus pour lesquels on détient des observations pour p variables. Ce jeu de données peut alors être représenté sous la forme d'un nuage de points de $\mathbb{R}^p$.

Lorsque p est grand, il devient alors difficile de visualiser correctement ce nuage de points, et les méthodes d'analyses factorielles que nous allons détailler ici permettent alors de représenter la projection ce ce nuage de points dans un sous-espace de dimension plus petite, mais en gardant le maximum d'information.

Cette situation peut être résumée dans une matrice $(x_{i,j})$ communément appelée la **matrice des données**

$$X = \begin{pmatrix} x_{1,1} & \cdots & x_{1,p} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,p} \end{pmatrix}$$

Ainsi dans la notation x_{ij}, l'indice i fait référence à l'individu i et l'indice j au caractère j.

Notons $c_j = (x_{1,j}, \dots, x_{n,j})$ la variable statistique du caractère j et $e_i = (x_{i,1}, \dots, x_{i,p})$ la variable statistique de tous les caractères de l'individu i.

Ainsi la matrice $(x_{i,j})$ est la matrice dont les lignes sont les e_i et les colonnes c_j.

Pour illustrer les propos de ce chapitre nous allons utiliser l'exemple suivant.
Pour un groupe de 60 étudiants, on dispose de leur note en anglais et de mathématiques.

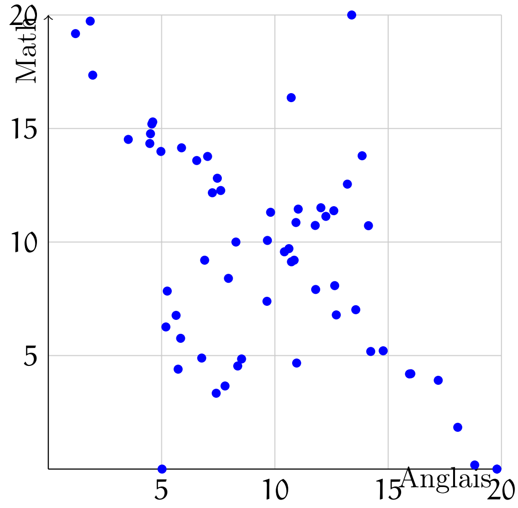
Étudiant 1	10.73	9.13
Étudiant 2	3.53	14.52
Étudiant 3	5.88	14.15
Étudiant 4	10.42	9.57
Étudiant 5	4.51	14.77
Étudiant 6	7.61	12.27
Étudiant 7	12.71	6.79
Étudiant 8	18.82	0.18
Étudiant 9	4.48	14.34
Étudiant 10	15.94	4.19
Étudiant 11	14.78	5.21
Étudiant 12	7.03	13.77
Étudiant 13	4.97	13.99
Étudiant 14	18.07	1.84
Étudiant 15	4.57	15.2
Étudiant 16	4.61	15.29
Étudiant 17	19.8	0
Étudiant 18	17.21	3.91
Étudiant 19	11.8	7.91
Étudiant 20	7.46	12.81

Étudiant 21	16.0	4.2
Étudiant 22	13.57	7.02
Étudiant 23	1.96	17.35
Étudiant 24	1.85	19.73
Étudiant 25	14.23	5.18
Étudiant 26	6.55	13.59
Étudiant 27	8.28	10.0
Étudiant 28	1.2	19.18
Étudiant 29	9.67	10.07
Étudiant 30	10.62	9.71
Étudiant 31	5.84	5.76
Étudiant 32	10.93	10.86
Étudiant 33	8.36	4.54
Étudiant 34	7.24	12.17
Étudiant 35	5.25	7.84
Étudiant 36	13.2	12.55
Étudiant 37	10.72	16.36
Étudiant 38	11.78	10.73
Étudiant 39	6.9	9.2
Étudiant 40	5.02	0

Étudiant 41	6.77	4.89
Étudiant 42	13.85	13.8
Étudiant 43	8.53	4.85
Étudiant 44	5.64	6.77
Étudiant 45	12.6	11.38
Étudiant 46	14.13	10.72
Étudiant 47	7.95	8.4
Étudiant 48	7.41	3.34
Étudiant 49	5.73	4.4
Étudiant 50	12.25	11.13
Étudiant 51	12.03	11.51
Étudiant 52	9.65	7.39
Étudiant 53	12.64	8.08
Étudiant 54	11.03	11.45
Étudiant 55	10.85	9.2
Étudiant 56	5.19	6.26
Étudiant 57	9.81	11.31
Étudiant 58	10.96	4.67
Étudiant 59	7.8	3.66
Étudiant 60	13.39	20

Avec nos notations, $X = \begin{pmatrix} 10.73 & 9.13 \\ 3.53 & 14.52 \\ \vdots & \vdots \\ 7.8 & 3.66 \\ 13.39 & 20 \end{pmatrix}$ est une matrice à 60 lignes et 2 colonnes et par exemple $e_{30} = (10.62 \quad 9.71)$, $e_{52} = (9.65 \quad 7.39)$.

Dans ce cas très particulier de la dimension $p = 2$, nous pouvons représenter ces données dans un repère.



> **Théorème**
>
> Soit $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données. Alors ${}^tXX$ est une matrice diagonalisable et toutes ses valeurs propres sont strictement positive.

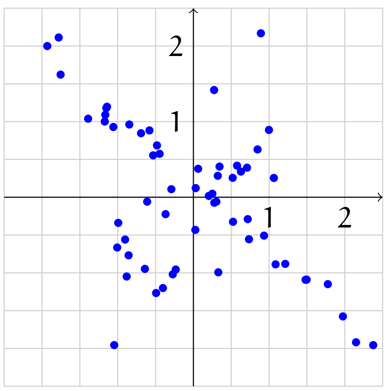
Démonstration. Admise $\square$

Avec le données de notre exemple, on a ${}^tXX = \begin{pmatrix} 6578.5797 & 4818.7964 \\ 4818.7964 & 6808.7059 \end{pmatrix}$.

Il est d'accoutumé en statistique de travailler avec des données normalisées. C'est à dire de considérer la matrice, que nous noterons dans la suite encore X pour ne pas alourdir les notations

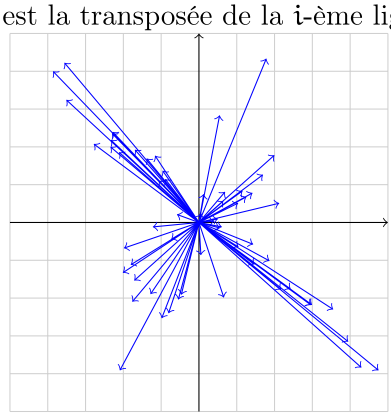
$$X = \left(\frac{x_{i,j} - \overline{c_j}}{\sigma_{c_j}}\right)_{i,j}$$

Avec notre exemple, nous avons :



En d'autre terme, sans perdre en généralité, on peut supposer que les colonnes des X sont des variables statistiques de moyenne nulles et d'écart-type 1.

En particulier, nous pouvons considérer les individus e_i comme des vecteurs de $\overrightarrow{e_i} \in \mathbb{R}^p$. Dans l'absolue $\overrightarrow{e_i}$ est la transposée de la i-ème ligne de X.



Dans la suite, on munit $\mathbb{R}^p$ de sa structure euclidienne canonique.

> **Exercice**
>
> Soit $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données et $\overline{X}$ la matrice des données centrées et réduites.
>
> 1. Montrer que $\left(\frac{1}{n}{}^tXX\right)_{i,j} = \mathrm{corr}(c_i, c_j)$.
> 2. En déduire que $\left(\frac{1}{n}{}^t\overline{XX}\right)_{i,j} = \mathrm{cov}(c_i, c_j)$.

3. En déduire que $\left(\frac{1}{n}{}^t\overline{X}\overline{X}\right)_{i,i} = 1$.
4. En déduire que $\mathrm{tr}\left({}^t\overline{X}\overline{X}\right) = np$.

Avec notre exemple, on a $\frac{1}{60}{}^t\overline{X}\overline{X} = \begin{pmatrix} 1 & -0.4849 \\ -0.4849 & 1 \end{pmatrix}$

Définition

Soit $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(e_i)_{i\in[\![1,n]\!]}$. On appel **inertie absolue** ou **inertie totale** de X, noté $I(X)$, le nombre

$$I(X) = \frac{1}{n}\sum_{i=1}^{n} \left\|\overrightarrow{e_i}\right\|^2$$

On pourra vérifier que l'inertie absolue de notre exemple est $I(X) = 2$. En fait c'est toujours le cas, on a toujours $I(X) = p$. Nous le démontrerons plus tard.

On note parfois l'inertie absolue $I_{\mathbb{R}^p}$. Cela donne du sens à la définition suivante.

Définition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(e_i)_{i\in[\![1,n]\!]}$ et V un sous espace vectoriel de $\mathbb{R}^p$. On appel **inertie relative** à V le nombre

$$I_V(X) = \frac{1}{n}\sum_{i=1}^{n} \left\|p_V(\overrightarrow{e_i})\right\|^2$$

En d'autre terme, au lieu de regarder les individus dans $\mathbb{R}^p$, on les regarde dans un espace plus petit au travers de leur projection. On ne calcule plus les normes des individus dans l'absolue de tout l'espace, comme dans l'inertie totale, mais leur projeté sur l'espace considéré.

Proposition Pythagore

Soit $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(e_i)_{i\in[\![1,n]\!]}$.

(i). $\forall V \subseteq \mathbb{R}^p$, $I_V(X) \leqslant I(X)$

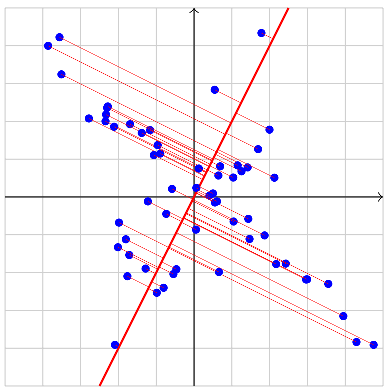
(ii). $\forall V \subseteq \mathbb{R}^p$, $I(X) = I_V(X) + I_{V^\perp}(X)$

(iii). $\forall V \subseteq \mathbb{R}^p$, $I_{V^\perp}(X) = \frac{1}{n}\sum_{i=1}^{n} \left\|\overrightarrow{e_i} - p_V(\overrightarrow{e_i})\right\|^2$

Démonstration. Il s'agit de la reformulation du théorème de Pythagore : $\overrightarrow{e_i} = \left(\overrightarrow{e_i} - p_V(\overrightarrow{e_i})\right) + p_V(\overrightarrow{e_i})$ $\square$

Avec notre exemple, considérons $\overrightarrow{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ et $V = \mathtt{Vect}\left(\overrightarrow{v}\right)$ alors l'inertie relative à V est $I_V(X) = 0.6120739427123439$.

L'idée de l'ACP est de trouver un sous-espace vectoriel V (dont la dimension est fixée *a priori* et est en général 1, 2 ou 3 en fonction des besoins du datascientiste) de sorte que l'inertie relative à V, soit la plus proche de l'inertie totale. D'après la proposition précédente, il s'agit donc de trouver $V \subseteq \mathbb{R}^p$ *simple* tel que $I_V(X)$ soit le plus grand possible ou de manière équivalente $I_{V^\perp}(X)$ soit le plus petit possible.



Proposition

Soit $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données.
Si $\overrightarrow{v} \in E$ un vecteur unitaire et $V = \mathtt{Vect}(\overrightarrow{v})$ alors

$$I_V(X) = \frac{1}{n}\langle \overrightarrow{v} \mid {}^t X X \overrightarrow{v} \rangle$$

Démonstration.

D'une part on a

$$\begin{aligned}
I_V(X) &= \frac{1}{n}\sum_{i=1}^{n} \left\| p_V(\overrightarrow{e_i}) \right\|^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \langle \overrightarrow{e_i} \mid \overrightarrow{v} \rangle^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \left(\sum_{j=1}^{p} x_{i,j} v_j \right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p}\sum_{l=1}^{p} x_{i,k} v_k x_{i,l} v_l \\
&= \frac{1}{n}\sum_{k=1}^{p}\sum_{l=1}^{p}\sum_{i=1}^{n} x_{i,k} v_k x_{i,l} v_l \\
&= \frac{1}{n}\sum_{k=1}^{p}\sum_{l=1}^{p} \left(\sum_{i=1}^{n} x_{i,k} x_{i,l} \right) v_k v_l \\
&= \frac{1}{n}\sum_{k=1}^{p}\sum_{l=1}^{p} \mathrm{cov}(c_k, c_l) v_k v_l
\end{aligned}$$

D'autre part, on rappel que $({}^t XX)_{i,j} = \mathrm{cov}(c_i, c_j)$. On a alors

$$\begin{aligned}
\langle \overrightarrow{v} \mid {}^t XX\overrightarrow{v} \rangle &= \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix} \middle| \begin{pmatrix} \mathrm{cov}(c_1,c_1) & \cdots & \mathrm{cov}(c_1,c_p) \\ \vdots & \ddots & \vdots \\ \mathrm{cov}(c_p,c_1) & \cdots & \mathrm{cov}(c_p,c_p) \end{pmatrix} \right. \\
&= \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix} \middle| \begin{pmatrix} \vdots \\ \sum_{l=1}^{p} \mathrm{cov}(c_i, c_l) v_l \\ \vdots \end{pmatrix} \right\rangle \\
&= \sum_{k=1}^{p}\sum_{l=1}^{p} \mathrm{cov}(c_k, c_l) v_l v_k
\end{aligned}$$

□

Avec notre exemple, prenons $\overrightarrow{v} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$. On a ${}^t XX = \begin{pmatrix} 60 & -29.0945 \\ -29.0945 & 60 \end{pmatrix}$ et $\frac{1}{n}\langle \overrightarrow{v} \mid {}^t XX\overrightarrow{v} \rangle =$
$\frac{1}{60}\left\langle \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \middle| \begin{pmatrix} 60 & -29.0945 \\ -29.0945 & 60 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \right\rangle = 0.6120739427123443 = I_V(X)$

Théorème

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données, $\overrightarrow{v}$ un vecteur propre unitaire de valeur propre λ de la matrice $\frac{1}{n}{}^t XX$.

$$I_V(X) = \lambda$$

Démonstration. $I_V(X) = \frac{1}{n}\langle \overrightarrow{v} \mid {}^t XX\overrightarrow{v} \rangle = \left\langle \overrightarrow{v} \middle| \frac{1}{n}{}^t XX\overrightarrow{v} \right\rangle = \langle \overrightarrow{v} \mid \lambda\overrightarrow{v} \rangle = \lambda \left\| \overrightarrow{v} \right\|^2 = \lambda$ □

Dans notre exemple, nous avons $\frac{1}{60}{}^t XX = \begin{pmatrix} 1 & -0.4849 \\ -0.4849 & 1 \end{pmatrix}$ dont on peut montrer que 0.51509243 et 1.48490757 sont des valeurs propres de vecteurs propres unitaire respectif $\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ et $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$. On laisse le soin au lecteur de vérifier les calculs.

Corollaire

Soient $X \in \mathtt{Mat}_{n,p}(\mathbb{R})$ une matrice des données et $\lambda_1, \ldots, \lambda_p$ des valeurs propres de $\frac{1}{n}{}^t XX$. L'inertie totale de X est la somme des valeurs propre

$$I(X) = \sum_{i=1}^{p} \lambda_i$$

Démonstration. C'est une conséquence de l'égalité $I(X) = I_V(X) + I_{V^\perp}(X)$ avec le précédent théorème. □

Avec notre exemple, on vérifie bien que $I(X) \simeq 2 \simeq 0.515 + 1.485$.

Corollaire

Soit $X \in \mathtt{Mat}_{n,p}(\mathbb{R})$ une matrice des données. Alors $I(X) = p$

Démonstration. L'inertie $I(X)$ est la somme des valeurs propres, c'est à dire la trace de la diagonalisation de la matrice de variance-covariance $\frac{1}{n}{}^t XX$. La trace est invariante par changement de base.

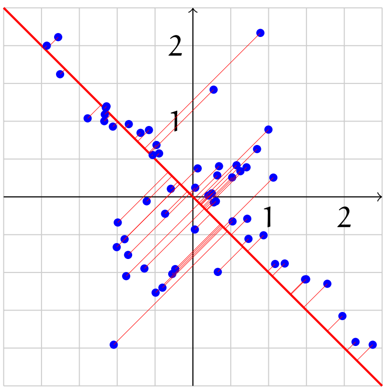
Puisque les données sont centrée et réduite les variances, les éléments diagonaux de la matrice de variance-covariance, valent 1. □

Définition

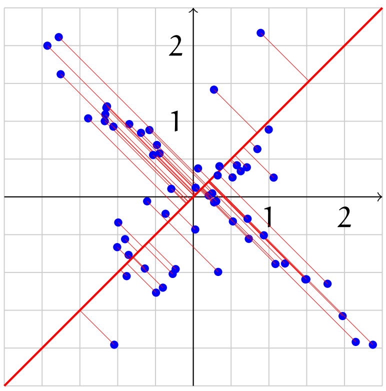
Soient $X \in \mathtt{Mat}_{n,p}(\mathbb{R})$ une matrice des données, $\lambda_1, \cdots, \lambda_p$ les valeurs propres de $A = \frac{1}{n}{}^t XX$ (comptées avec leur multiplicité) et $\overrightarrow{v}_i$ leur vecteurs propres associés.

On appel les droites $V_i = \mathtt{Vect}\left(\overrightarrow{v}_i\right)$ les **axes factoriels** ou **axes principales d'inertie**.

Dans notre exemple les axes factoriels sont $\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ et $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$



$I_V(X) \simeq 1.485$



$I_V(X) \simeq 0.515$

On rappel que l'idée est de trouver une sous-espace vectoriel V tel que $I_V(X)$ soit le plus grand possible (ou $I_{V^\perp}(X)$ est le plus petit possible).

D'après les résultats précédent, si on considère les vecteurs propres $\overrightarrow{v}_i$ de la matrices $A = \frac{1}{n}{}^t XX$ associés aux valeurs propres ordonnées $\lambda_1 \geqslant \cdots \geqslant \lambda_p$ et $V_i = \mathtt{Vect}\left(\overrightarrow{v}_i\right)$ alors $I_{V_1} = \lambda_1 = \max(\lambda_i) = \max(I_{V_i}(X))$. On appel usuellement cette droite **axe principale**.

§ En pratique il y a deux manières de construire les axes.

Ascendante. On choisit la valeur propre λ la plus grande. L'axe principal $\overrightarrow{v}$ de cette valeur propre porte donc principalement l'inertie totale.

Dans notre exemple, l'axe principale est porté par le vecteur $\overrightarrow{v} = \begin{pmatrix} 0.7071 \\ -0.7071 \end{pmatrix}$. En particulier le calcul de $d_i = 0.7071 \times \mathsf{Note}_{\mathsf{Anglais},i} - 0.7071 \times \mathsf{Note}_{\mathsf{Math},i}$ permet de garder l'essentiel de l'information d'un étudiant.

Descendante. On choisit d'enlever la valeur propre λ la plus petite. Si $\overrightarrow{v}$ est un vecteur propre de cette valeur propre et $V = \mathtt{Vect}(\overrightarrow{v})$, on choisit de projeter sur $V^{\perp}$.

Définition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données, $\lambda_1, \cdots, \lambda_p$ les valeurs propres de $A = \frac{1}{n}{}^tXX$ (comptées avec leur multiplicité), $\overrightarrow{v}_i$ leur vecteurs propres associés et $V_i = \mathtt{Vect}(\overrightarrow{v}_i)$.

On appel **contribution relative de l'axe** V_j le réel $cr_j = \dfrac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$.

Dans notre exemple la contribution relative de l'axe principale $\overrightarrow{v} = \begin{pmatrix} 0.7071 \\ -0.7071 \end{pmatrix}$ est $cr_1 = \dfrac{1.485}{2} \simeq 74,25\%$. Il faut donc comprendre que environ 75% de l'information d'un individu se retrouve sur l'axe principale $V = \mathtt{Vect}(\overrightarrow{v})$.

Définition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(\overrightarrow{e_i})_{i\in[\![1;n]\!]}$, $\mathcal{B} = \{\overrightarrow{v}_1, \ldots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^tXX$ (comptées avec leur multiplicité) et $V_i = \mathtt{Vect}(\overrightarrow{v}_i)$ les axes factoriels.

On appel **contribution relative de l'individu** $\overrightarrow{e_i}$ **sur l'axe** V_j le nombre $cr_j(\overrightarrow{e_i}) = \dfrac{\langle \overrightarrow{e_i}_{\mathcal{B}} \mid \overrightarrow{v}_{j\mathcal{B}} \rangle^2}{\|\overrightarrow{e_i}_{\mathcal{B}}\|^2}$

Reprenons notre exemple, la base $\mathcal{B}$ de diagonalisation est $\overrightarrow{f_1} = \begin{pmatrix} -0.7071 \\ -0.7071 \end{pmatrix}$ et $\overrightarrow{f_2} = \begin{pmatrix} 0.7071 \\ -0.7071 \end{pmatrix}$. Considérons par exemple l'individu $\overrightarrow{e_{30}} = \begin{pmatrix} 0.2504 \\ 0.0464 \end{pmatrix}$, dont on peut démontrer que $\overrightarrow{e_{30}} \simeq -0.2099\overrightarrow{f_1} + 0.1442\overrightarrow{f_2}$.

Ainsi on a $cr_1(\overrightarrow{e_{30}}) \simeq \dfrac{\left\langle \begin{pmatrix} -0.2099 \\ 0.1442 \end{pmatrix} \middle| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle^2}{\left\| \begin{pmatrix} 0.2099 \\ 0.1442 \end{pmatrix} \right\|^2} \simeq 0.6793$. De même $cr_2(\overrightarrow{e_{30}}) \simeq 0.3207$

Proposition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(\overrightarrow{e_i})_{i\in[\![1;n]\!]}$, $\mathcal{B} = \{\overrightarrow{v}_1, \ldots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^tXX$ (comptées avec leur multiplicité) et $V_i = \mathtt{Vect}(\overrightarrow{v}_i)$ les axes factoriels.

$$cr_j(\overrightarrow{e_i}) = \cos(\overrightarrow{e_i}, \overrightarrow{v}_j)^2 = \frac{\langle \overrightarrow{e_i} \mid \overrightarrow{v}_j \rangle^2}{\|\overrightarrow{e_i}\|^2}$$

Démonstration. Notons P la matrice de passage de la base canonique à la base orthonormale des vecteurs propre. Alors P est une matrice orthogonale, c'est à dire ${}^{t}PP = \mathrm{Id}_p$. De plus $\overrightarrow{x}_{\mathcal{B}} = P\overrightarrow{x}$.

$$
\begin{aligned}
cr_j(\overrightarrow{e_i}) &= \frac{\left\langle \overrightarrow{e_i}_{\mathcal{B}} \mid \overrightarrow{v}_{j\mathcal{B}} \right\rangle^2}{\left\| \overrightarrow{e_i}_{\mathcal{B}} \right\|^2} \\
&= \frac{{}^{t}\overrightarrow{e_i}_{\mathcal{B}} \overrightarrow{v}_{j\mathcal{B}}{}^2}{{}^{t}\overrightarrow{e_i}_{\mathcal{B}} \overrightarrow{e_i}_{\mathcal{B}}} \\
&= \frac{{}^{t}(P\overrightarrow{e_i})(P\overrightarrow{v}_j)^2}{{}^{t}(P\overrightarrow{e_i})(P\overrightarrow{e_i})} \\
&= \frac{{}^{t}\overrightarrow{e_i}\,{}^{t}PP\overrightarrow{v}_j{}^2}{{}^{t}\overrightarrow{e_i}\,{}^{t}PP\overrightarrow{e_i}} \\
&= \frac{{}^{t}\overrightarrow{e_i}\overrightarrow{v}_j{}^2}{{}^{t}\overrightarrow{e_i}\overrightarrow{e_i}} \\
&= \frac{\left\langle \overrightarrow{e_i} \mid \overrightarrow{v}_j \right\rangle^2}{\left\| \overrightarrow{e_i} \right\|^2} \\
&= \frac{\left\langle \overrightarrow{e_i} \mid \overrightarrow{v}_j \right\rangle^2}{\left\| \overrightarrow{e_i} \right\|^2 \left\| \overrightarrow{v}_j \right\|^2} \\
&= \cos^2(\overrightarrow{e_i}, \overrightarrow{v}_j)
\end{aligned}
$$

□

Plus la (racine carré) contribution relative d'une donnée est proche de 1 plus sa projection sur l'axe principale est proche de la donnée.

Définition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(\overrightarrow{e_i})_{i \in [\![1;n]\!]}$, $\mathcal{B} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \dfrac{1}{n}{}^{t}XX$ et P la matrice de passage de la base canonique à $\mathcal{B}$.

Posons $Y = XP$. Les colonnes de Y sont appelées les **composantes principales**.

Les composantes principales correspondent donc aux caractères étudiés dans la base des axes factoriels. Avec notre exemple, nous obtenons la matrice Y

−0.247	−0.143
1.718	0.249
1.279	−0.081
−0.132	−0.157
1.594	0.052
0.722	−0.09
−0.912	−0.126
−2.876	−0.163
1.536	0.12
−1.82	−0.276
−1.481	−0.235
1.035	−0.214
1.405	0.091
−2.511	−0.282
1.647	−0.02
1.653	−0.04
−3.063	−0.297
−2.069	−0.443
−0.6	−0.141
0.825	−0.145

−1.828	−0.287
−1.019	−0.301
2.387	0.094
2.752	−0.235
−1.396	−0.14
1.088	−0.109
0.281	0.131
2.779	−0.049
0.064	−0.107
−0.144	−0.21
0.062	1.149
−0.027	−0.428
−0.528	0.914
0.768	−0.015
0.462	0.942
−0.152	−1.046
0.809	−1.196
−0.185	−0.548
0.39	0.473
−0.643	2.123

−0.217	1.123
−0.077	−1.335
−0.511	0.841
0.242	1.034
−0.225	−0.777
−0.571	−0.932
0.102	0.418
−0.548	1.244
−0.118	1.365
−0.204	−0.684
−0.113	−0.703
−0.324	0.287
−0.713	−0.303
0.042	−0.531
−0.256	−0.173
0.242	1.182
0.222	−0.311
−0.935	0.469
−0.565	1.134
0.903	−2.164

Pur alléger les notations, nous noterons encore c_j la variable statistique des composantes principale (les colonnes de Y).

Proposition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données, $\mathcal{B} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^t XX$, P la matrice de passage de la base canonique à $\mathcal{B}$ et $Y = XP = (c_1, \dots, c_p)$.

(i). $\forall j \in [\![1;p]\!]$, $\overline{c_j} = 0$

(ii). $\forall j \in [\![1;p]\!]$, $\sigma^2_{c_j} = \lambda_j$

(iii). $\forall i, j \in [\![1;p]\!]$, $i \neq j$, $\mathrm{cov}(c_i, c_j) = 0$

Démonstration.

(i). Par construction la i-ème coordonnée du vecteur c_j, est $\sum_{k=1}^{p} x_{i,k} P_{k,j}$ alors

$$\begin{aligned}
\overline{c_j} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p} x_{i,k}P_{k,j} \\
&= \frac{1}{n}\sum_{k=1}^{p}\sum_{i=1}^{n} x_{i,k}P_{k,j} \\
&= \sum_{k=1}^{p}\left(\frac{1}{n}\sum_{i=1}^{n} x_{i,k}\right) P_{k,j} \\
&= \sum_{k=1}^{p}(0)\, P_{k,j} \\
&= 0
\end{aligned}$$

(ii) **et** (iii). Les vecteurs colonnes de la matrice P sont les vecteurs $\overrightarrow{v}_j$ de la base orthonormale de vecteurs propres. En particulier $c_j = X\overrightarrow{v}_j$. Ainsi

$$\begin{aligned}
\mathrm{cov}(c_i, c_j) &= \mathrm{cov}(X\overrightarrow{v}_i, X\overrightarrow{v}_j) \\
&= \frac{1}{n}\langle X\overrightarrow{v}_i \mid X\overrightarrow{v}_j\rangle \\
&= \frac{1}{n}{}^t(X\overrightarrow{v}_i)(X\overrightarrow{v}_j) \\
&= \frac{1}{n}{}^t\overrightarrow{v}_i{}^t XX\overrightarrow{v}_j \\
&= {}^t\overrightarrow{v}_i A\overrightarrow{v}_j \\
&= \langle \overrightarrow{v}_i \mid A\overrightarrow{v}_j\rangle \\
&= \langle \overrightarrow{v}_i \mid \lambda_j\overrightarrow{v}_j\rangle \\
&= \lambda_j\langle \overrightarrow{v}_i \mid \overrightarrow{v}_j\rangle
\end{aligned}$$

où $A = \frac{1}{n}{}^t XX$ est diagonale dans la base $\mathcal{B}$.

□

Proposition

Soient $X \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice des données, $\mathcal{B} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^t XX$.

Pour tout i et j distinct entre 1 et n, notons $U_{i,j} = \left(V_i \overset{\perp}{\oplus} Vj\right)^{\perp}$.

Pour tout $\overrightarrow{x} \in \mathbb{R}^p$, il existe des uniques $\alpha_{i,j}(\overrightarrow{x}) \in \mathbb{R}$, $\beta_{i,j}(\overrightarrow{x}) \in \mathbb{R}$ et $\overrightarrow{u}_{i,j}(\overrightarrow{x}) \in U$ tel que

$$\overrightarrow{x} = \alpha_{i,j}(\overrightarrow{x})\overrightarrow{v}_i + \beta_{i,j}(\overrightarrow{x})\overrightarrow{v}_j + \overrightarrow{u}_{i,j}(\overrightarrow{x})$$

De plus $\alpha_{i,j}(\overrightarrow{x}) = \langle \overrightarrow{x} \mid \overrightarrow{v}_i \rangle$ et $\beta_{i,j}(\overrightarrow{x}) = \langle \overrightarrow{x} \mid \overrightarrow{v}_j \rangle$

Démonstration. L'existence de la décomposition est une conséquence de l'égalité $\mathbb{R}^p = V_i \overset{\perp}{\oplus} V_j \overset{\perp}{\oplus} U_{i,j}$. Finalement la description de $\alpha_{i,j}$ est une conséquence de la définition de la projection □

Théorème

Soient $X \in \mathtt{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(\overrightarrow{e_k})_{k\in[\![1;n]\!]}$, $\mathcal{B} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^tXX$ et $V_i = \mathtt{Vect}(\overrightarrow{v}_i)$ les axes factoriels. Notons P la matrice passage de la base canonique à $\mathcal{B}$.

$$\alpha_{i,j}(\overrightarrow{e_k}) = \sqrt{\lambda_i}P_{k,i} \qquad \beta_{i,j}(\overrightarrow{e_k}) = \sqrt{\lambda_j}P_{k,j}$$

Démonstration. Distinguons les colonnes de X en les notant c_j et les colonnes des $Y = XP$ en les notant d_j. On rappel que ${}^tPAP = D$ ou encore $AP = PD$ où $A = \frac{1}{n}{}^tXX$ et D est la matrice diagonale des valeurs propres.

D'une part on a $\alpha_{i,j}(\overrightarrow{e_k}) = \langle \overrightarrow{e_k} \mid \overrightarrow{v}_i \rangle = \sum_{a=1}^{p} x_{k,a}P_{a,i} = (XP)_{k,i} = Y_{k,i} = \mathtt{corr}(c_k, d_i)$. Calculons cette corrélation.

$$\begin{aligned}
\mathtt{cov}(c_i, d_j) &= \mathtt{cov}(c_i, (XP)_j) \\
&= \frac{1}{n}\sum_{k=1}^{n} x_{k,i}(XP)_{k,j} \\
&= \frac{1}{n}\sum_{k=1}^{n} x_{k,i}\left(\sum_{a=1}^{p} x_{k,a}P_{a,j}\right) \\
&= \frac{1}{n}\sum_{a=1}^{p}\sum_{k=1}^{n} x_{k,i}x_{k,a}P_{a,j} \\
&= \sum_{a=1}^{p}\left(\frac{1}{n}\sum_{k=1}^{n} x_{k,i}x_{k,a}\right)P_{a,j} \\
&= \sum_{a=1}^{p} A_{i,a}P_{a,j} \\
&= (AP)_{i,j} \\
&= (PD)_{i,j} \\
&= \sum_{a=1}^{p} P_{i,a}D_{a,j} \\
&= P_{i,j}D_{j,j} \\
&= \lambda_j P_{i,j}
\end{aligned}$$

Ainsi $\mathtt{corr}(c_i, d_j) = \dfrac{\mathtt{cov}(c_i, d_j)}{\sigma_{c_i}\sigma_{d_j}} = \dfrac{\lambda_j P_{i,j}}{\sqrt{\lambda_j}} = \sqrt{\lambda_j}P_{i,j}$ □

Définition

Soient $X \in \mathtt{Mat}_{n,p}(\mathbb{R})$ une matrice des données de ligne $(\overrightarrow{e_k})_{k \in [\![1;n]\!]}$, $\mathcal{B} = \{\overrightarrow{v}_1, \dots, \overrightarrow{v}_p\}$ une base orthonormale de vecteur propre de $A = \frac{1}{n}{}^tXX$ et $V_i = \mathtt{Vect}\left(\overrightarrow{v}_i\right)$ les axes factoriels.

On appel **cercle des corrélations** suivant les axes V_i et V_j la représentation des vecteurs $\begin{pmatrix} \alpha_{i,j}(\overrightarrow{e_k}) \\ \beta_{i,j}(\overrightarrow{e_k}) \end{pmatrix}$ dans $\mathbb{R}^2$

L'exemple fil rouge de ce chapitre a atteint ses limites car les données sont déjà en dimension 2. Le principe du cercle des corrélation est de projeter les points sur un plan significatif. Dans notre cas, il ne s'agit (malheureusement) que d'une rotation des données [1].

Généralement, on représente les anciennes variables en dimension 2, à l'aide de ce que l'on appelle un cercle des corrélations. Plus précisément, on choisit deux axes E_j et E_k , et l'on trace les points dont les coordonnées sont les corrélations de chacune des variables avec les $j$ème et $k$ème composantes. principales.

Remarque :

- Plus une variable est proche du cercle de corrélation, mieux elle est représentée par le plan considérée.

- Si deux variables proches du cercle de corrélation (et donc bien représentées dans le plan considéré) sont proches l'une de l'autre, alors elles sont elle-mêmes fortement corrélées positivement.

- Inversement, deux variables proches du cercle de corrélation mais symétriquement opposées par rapport à l'origine seront fortement corrélées négativement.

1. On peut démontrer que les changements de base orthogonales dans $\mathbb{R}^2$ sont soient des rotations soient des symétries.