Produit scalaire.

1. Produit scalaire réel.

2. Orthogonalité.

3. Projections orthogonales.

4. Espaces euclidiens.

5. Automorphismes orthogonaux et matrices orthogonales.

6. Espaces euclidiens de dimension 2 ou 3.

Produit scalaire.

1. Produit scalaire réel.

Définition 1.1 : produit scalaire sur un $\mathbb{R}$-espace vectoriel, espace préhilbertien réel

Soit E un $\mathbb{R}$-espace vectoriel.

On dit que φ est un produit scalaire sur E si et seulement si φ est une forme bilinéaire symétrique positive, non dégénérée, soit encore :

- φ est une application de $E\times E$ dans $\mathbb{R}$,
- φ est bilinéaire :
  $\forall (x,y) \in E\times E$, $\forall (x',y') \in E\times E$, $\forall (\lambda,\mu) \in \mathbb{R}^2$,
  $\varphi(\lambda.x + \mu.y,x') = \lambda.\varphi(x,x') + \mu.\varphi(y,x')$, et :
  $\varphi(x,\lambda.x' + \mu.y') = \lambda.\varphi(x,x') + \mu.\varphi(x,y')$,
- φ est symétrique : $\forall (x,y) \in E\times E$, $\varphi(x,y) = \varphi(y,x)$,
- φ est positive : $\forall x \in E$, $\varphi(x,x) \geq 0$,
- φ est définie : $\forall x \in E$, $(\varphi(x,x) = 0) \Rightarrow (x = 0)$.

On dit alors que (E,φ) est un espace préhilbertien réel.

Théorème 1.1 : exemples classiques dans le cas réel

Les applications suivantes définissent des produits scalaires sur les espaces vectoriels indiqués :

- $\forall (x,y) \in (\mathbb{R}^n)^2$, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, $(x,y) \mapsto \sum_{i=1}^{n} x_i.y_i$, dans $\mathbb{R}^n$.
- $\forall (f,g) \in (C^0([a,b],\mathbb{R})^2$, $(f,g) \mapsto \int_a^b f(t).g(t).dt$, dans $C^0([a,b], \mathbb{R})$, où $[a,b]$ est un segment inclus dans $\mathbb{R}$.
- $\forall (A,B) \in \mathcal{M}_n(\mathbb{R})^2$, $(A,B) \mapsto \mathrm{tr}({}^tA.B)$.

Démonstration :

- L'application proposée (notons-la $(.|.)$) est correctement définie de $(\mathbb{R}^n)^2$ dans $\mathbb{R}$.

Il est clair qu'elle est symétrique puisque : $\forall (x,y) \in (\mathbb{R}^n)^2$, $(x|y) = \sum_{i=1}^{n} x_i.y_i = \sum_{i=1}^{n} y_i.x_i = (x|y)$.

Elle est de plus linéaire par rapport à sa première variable puisque :
$\forall x \in \mathbb{R}^n$, $\forall (y,y') \in \mathbb{R}^n$, $\forall (\lambda,\lambda') \in \mathbb{R}^2$, $(x|\lambda.y +\lambda'.y') = \lambda.(x|y) + \lambda'.(x|y')$, sans difficulté.
Enfin, elle est non dégénérée, puisque si pour : $x \in \mathbb{R}^n$, on a : $\forall y \in \mathbb{R}^n$, $(x|y) = 0$, en particulier pour :

$y = x$, et cela donne : $\sum_{i=1}^{n} x_i^2 = 0$, et tous les x_i étant positifs, on conclut à : $\forall 1 \leq i \leq n$, $x_i = 0$, soit : $x = 0$.

- De même, l'application (notée encore $(.|.)$) est correctement définie de E^2 dans $\mathbb{R}$, (puisque, en notant E l'espace vectoriel $C^0([a,b],\mathbb{R})$), pour tout : $(f,g) \in E^2$, $f.g$ est continue est à valeurs réelles sur $[0,1]$.
De plus, elle est symétrique de façon immédiate.
Puis elle est linéaire par rapport à sa deuxième variable (du fait de la linéarité de l'intégrale sur $[a,b]$).

Enfin, si pour : $f \in E$, on a : $(f|g) = 0$, en particulier pour : $g = f$, ce qui donne : $\int_a^b f(t)^2.dt = 0$.

Mais comme la fonction f^2 est continue et positive sur $[a,b]$, on en déduit bien que : $f^2 = 0$, puis : $f = 0$.
- Enfin, l'application

Théorème 1.2 : inégalité de Cauchy-Schwarz

Soit E un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire $(.|.)$.

Alors : $\forall (x,y) \in E^2$, $|(x|y)| \leq \sqrt{(x|x)}.\sqrt{(y|y)}$.

Démonstration :

Soit ψ la fonction définie de $\mathbb{R}$ dans $\mathbb{R}$ par : $\forall t \in \mathbb{R}$, $\psi(t) = (x+t.y|x+t.y)$.

Puisque le produit scalaire est une forme positive, ψ est également à valeurs dans $\mathbb{R}^+$.

Par ailleurs : $\forall t \in \mathbb{R}$, $\psi(t) = (x|x) + 2.t(x|y) + t^2.(y|y)$, en utilisant la bilinéarité et la symétrie de $(.|.)$.

Distinguons alors deux cas :

- $(y|y) = 0$.

Dans ce cas, ψ est une fonction affine de t qui reste positive : elle est donc constante et : $(x|y) = 0$.

On a bien alors : $0 = |(x|y)| \leq \sqrt{(x|x)}.\sqrt{(y|y)} = 0$.

- $(y|y) \neq 0$, et donc : $(y|y) > 0$.

Dans ce cas ψ est une fonction polynomiale du second degré.
Comme elle reste positive, elle admet au plus une racine réelle (double) et son discriminant est négatif ou nul, soit : $\Delta = (x|y)^2 - (x|x).(y|y) \leq 0$, ce qui donne à nouveau : $|(x|y)| \leq \sqrt{(x|x)}.\sqrt{(y|y)}$.

Définition 1.2 : forme bilinéaire symétrique non dégénérée

Soit φ une forme bilinéaire symétrique sur un $\mathbb{R}$-espace vectoriel E.
On dit que la forme φ est non dégénérée si et seulement si :
$\forall$ x $\in$ E, ($\forall$ y $\in$ E, φ(x,y) = 0) $\Rightarrow$ (x = 0).

Théorème 1.3 : équivalence non dégénérée $\Leftrightarrow$ définie

Soit E un $\mathbb{R}$-espace vectoriel, et soit φ une forme bilinéaire symétrique sur E.
Alors φ est non dégénérée si et seulement si φ est définie.
On peut donc remplacer « non dégénérée » par « définie » dans la définition d'un produit scalaire.

Démonstration :

On peut commencer par remarquer que dans la preuve de l'inégalité de Cauchy-Schwarz, on n'utilise à aucun moment le fait que le produit scalaire est une forme définie.
Puis on travaille par double implication.

- [$\Rightarrow$]

Si φ est non dégénérée, soit : x $\in$ E, tel que : φ(x,x) = 0.
On constate alors, d'après l'inégalité de Cauchy-Schwarz, que : $\forall$ y $\in$ E, φ(x,y) = 0.
Or φ étant non dégénérée, on en déduit que : x = 0, et φ est définie.

- [$\Leftarrow$]

Si φ est maintenant supposée définie et si x est tel que : $\forall$ y $\in$ E, φ(x,y) = 0, alors en particulier pour : y = x, et on en déduit que : φ(x,x) = 0.
Or φ étant supposée définie, on conclut bien que : x = 0.

Théorème 1.4 : cas d'égalité dans l'inégalité de Cauchy-Schwarz pour un produit scalaire

Soit (E, (.|.)) un espace préhilbertien réel.
Alors : $\forall$ (x,y) $\in$ E^2, $|(x|y)| \leq \sqrt{(x|x)}.\sqrt{(y|y)}$, si et seulement si (x,y) est liée.

Démonstration :

Si l'on reprend la démonstration de l'inégalité de Cauchy-Schwarz, on constate que, si cette inégalité devient une égalité, alors :

- dans le cas où : $(y|y) = 0$, alors y est nul puisque φ comme produit scalaire est une forme définie,

donc : 0.x + 1.y = 0.

- dans le cas où : $(y|y) \neq 0$, cela signifie que le discriminant de la fonction qui avait servi

d'intermédiaire est nul, et donc que le trinôme noté ψ admet une racine double.
Autrement dit, dans les deux cas : $\exists$ λ $\in$ $\mathbb{R}$, $(x + t.y|x + t.y) = 0$, et donc le vecteur (x + λ.y) est nul, ce qui s'écrit encore : 1.x + λ.y = 0.
Dans tous les cas, on constate bien que (x,y) est liée.

Définition 1.3 et théorème 1.5 : norme et distance associée à un produit scalaire, inégalité de Minkowski

Soit (E, (.|.)) un espace préhilbertien réel.
Alors l'application $\|\,.\,\|$ définie par : $\forall$ x $\in$ E, x $\mapsto$ $\|x\| = \sqrt{(x,x)}$, est une norme sur E, appelée norme associée au produit scalaire (.|.).
De même, l'application d définie sur E^2 par : $\forall$ (x,y) $\in$ E^2, $d(x, y) = \|x - y\|$, est appelée distance

associée à la norme $\|\,.\,\|$ (ou au produit scalaire (.|.)).

Enfin, l'inégalité triangulaire vérifiée par $\|\,.\,\|$ est appelée inégalité de Minkowski.

Démonstration :

Il y a donc quatre points à vérifier.

• pour tout vecteur x de E, la quantité N(x) existe puisque $\varphi(x,x)$ est un réel positif, et : $N(x) \in \mathbb{R}^+$.

• si pour : $x \in E$, on a : N(x) = 0, alors : $\varphi(x,x) = 0$, et φ comme produit scalaire est une forme définie, donc : x = 0.

• pour : $x \in E$, et : $\lambda \in \mathbb{R}$ (ou $\mathbb{C}$), $N(\lambda.x) = \sqrt{\varphi(\lambda.x,\lambda.x)} = \sqrt{|\lambda|^2.\varphi(x,x)} = |\lambda|.\sqrt{\varphi(x,x)} = |\lambda|.N(x)$.

• enfin, pour : $(x,y) \in E^2$, on a dans le cas réel :

$N(x+y)^2 = \varphi(x+y,x+y) = \varphi(x,x) + 2.\varphi(x,y) + \varphi(y,y) \le N(x)^2 + 2.N(x).N(y) + N(y)^2 = (N(x) + N(y))^2$, d'où l'inégalité triangulaire pour N, et dans le cas complexe :

$N(x+y)^2 = \varphi(x,x) + 2.\mathrm{Re}(\varphi(x,y)) + \varphi(y,y) \le N(x)^2 + 2.|\varphi(x,y)| + N(y)^2 \le N(x)^2 + 2.N(x).N(y) + N(y)^2$, et on termine comme dans le cas réel avec l'inégalité triangulaire.

Théorème 1.6 : égalités dites « de polarisation »

Soit (E, (.|.)) un espace préhilbertien réel et $\|\,.\,\|$ la norme associée à (.|.).

On a les égalités suivantes : $\forall$ $(x,y) \in E^2$,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2.(x|y),$$

$$(x|y) = \frac{1}{2}.(\|x+y\|^2 - \|x\|^2 - \|y\|^2) = \frac{1}{4}.(\|x+y\|^2 - \|x-y\|^2).$$

Démonstration :

Il suffit d'écrire : $\forall$ $(x,y) \in E^2$,

$\|x+y\|^2 = (x+y|x+y) = \|x\|^2 + \|y\|^2 + 2.(x|y)$, et : $\|x-y\|^2 = (x-y|x-y) = \|x\|^2 + \|y\|^2 - 2.(x|y)$,

d'où les égalités proposées.

2. Orthogonalité.

Définition 2.1 : vecteurs orthogonaux, vecteurs unitaires (ou normés), famille orthogonale, orthonormale

Soit (E, (.|.)) un espace préhilbertien réel et $\|\,.\,\|$ la norme associée à (.|.).

Deux vecteurs x et y de E sont dits orthogonaux pour (.|.) si et seulement si : (x|y) = 0.

De même, un vecteur x de E est dit unitaires (ou normé) si et seulement si : $\|x\| = 1$.

Soit $(x_i)_{i\in I}$, une famille de vecteurs de E.

On dit que la famille est orthogonale pour (.|.) si et seulement si : $\forall$ $(i,j) \in I^2$, $(i \neq j) \Rightarrow ((x_i|x_j) = 0)$.

On dit que la famille est orthonormale pour (.|.) si et seulement si elle est orthogonale et si de plus, tous les vecteurs de la famille sont normés.

Théorème 2.1 : liberté d'une famille orthogonale ou orthonormale

Soit (E, (.|.)) un espace préhilbertien réel et $\|\,.\,\|$ la norme associée à (.|.).

Soit $(x_i)_{i\in I}$, une famille de vecteurs de E.

Si la famille est orthogonale pour (.|.) et ne comporte pas le vecteur nul, elle est libre.

Si la famille est orthonormale pour (.|.), elle est libre.

Démonstration :

Comme les combinaisons linéaires qu'on peut envisager ne comporte toujours qu'au plus un nombre fini de coefficients non nuls, on peut noter la famille $(x_i)_{1\le i\le n}$, pour un entier n donné.

Alors, si pour : $(\lambda_i)_{1\le i\le n} \in \mathbf{K}^n$, on a : $\lambda_1.x_1 + \ldots + \lambda_n.x_n = 0$, alors :

$\forall$ $1 \le i \le n$, $(x_i|\lambda_1.x_1 + \ldots + \lambda_n.x_n) = 0$, et par linéarité par rapport à la deuxième variable, on obtient :

$\forall$ $1 \le i \le n$, $\lambda_i.\|x_i\|^2 = 0$, (puisqu la famille est orthogonale).

Donc si tous les x_i sont non nuls, on conclut bien au fait que : $\forall$ $1 \le i \le n$, $\lambda_i = 0$, et la famille est libre.

Si maintenant on considère une famille orthonormale, elle est orthogonale et ne contient pas le vecteur nul, donc elle est libre.

Théorème 2.2 : de Pythagore

Soit (E,(.|.)) un espace préhilbertien réel et $\|\,.\,\|$ la norme associée à (.|.).

Soit $(x_i)_{1\le i\le n}$, une famille finie orthogonale de vecteurs de E.

Alors : $\|x_1+\ldots+x_n\|^2 = \|x_1\|^2+\ldots+\|x_n\|^2$.

Démonstration :

On peut par exemple procéder par récurrence sur n.
Pour : n = 2, si (x_1, x_2) est une famille orthogonale de E, le théorème 2.7 donne bien le résultat voulu.
Si maintenant on suppose le résultat vrai pour tout famille orthogonale de n vecteurs de E (avec : $n \ge 2$), alors étant donné une famille orthogonale $(x_1, \ldots, x_{n+1})$ de vecteurs de E, la famille $(x_1 + \ldots + x_n, x_{n+1})$ est une famille de deux vecteurs de E, orthogonale car: $(x_1 + \ldots + x_n|x_{n+1}) = (x_1|x_{n+1}) + \ldots + (x_n|x_{n+1}) = 0$.

Donc : $\|x_1+\ldots+x_{n+1}\|^2 = \|x_1+\ldots+x_n\|^2+\|x_{n+1}\|^2 = \|x_1\|^2+\ldots+\|x_n\|^2+\|x_{n+1}\|^2$.

Et la récurrence est terminée.

Théorème 2.3 : procédé d'orthogonalisation et d'orthonormalisation de Gram-Schmidt

Soit (E, (.|.)) un espace préhilbertien réel et $\|\,.\,\|$ la norme associée à (.|.).

Soit $(x_i)_{1\le i\le n}$, une famille libre de vecteurs de E.
On peut construire une famille $(e_1, \ldots, e_n)$ de vecteurs de E telle que :

- $\forall\ 1 \le p \le n$, $e_p \ne 0$,
- $\forall\ 1 \le p \le n$, $e_p \in \mathrm{Vect}(x_1, \ldots, x_p)$,
- $\forall\ 1 \le i \ne j \le n$, $(e_i|e_j) = 0$.

On a dans ce cas : $\forall\ 1 \le p \le n$, $\mathrm{Vect}(x_1, .., x_p) = \mathrm{Vect}(e_1, \ldots, e_p)$.
Par ailleurs, il existe une unique famille $(\varepsilon_1, \ldots, \varepsilon_n)$ de vecteurs de E, telle que :

- $\forall\ 1 \le p \le n$, $\varepsilon_p \in \mathrm{Vect}(x_1, \ldots, x_p)$,
- $\forall\ 1 \le i \ne j \le n$, $(\varepsilon_i|\varepsilon_j) = 0$,
- $\forall\ 1 \le p \le n$, $\|\varepsilon_p\| = 1$,
- $\forall\ 1 \le p \le n$, $(\varepsilon_p|x_p) > 0$.

On a encore dans ce cas : $\forall\ 1 \le p \le n$, $\mathrm{Vect}(x_1, .., x_p) = \mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_p)$.

Démonstration :

• orthogonalisation de la famille : on procède à la construction par récurrence en montrant en même temps que la famille obtenue vérifie bien : $\forall\ 1 \le p \le n$, $\mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Vect}(x_1, \ldots, x_p)$.
On doit alors prendre e_1 colinéaire à x_1, et il est possible de trouver e_1 tel que : $e_1 \ne 0$, et : $e_1 \in \mathrm{Vect}(x_1)$.
Il suffit pour cela de choisir par exemple : $e_1 = x_1$.
On a alors en plus (et quelque soit en fait le choix de e_1) : $\mathrm{Vect}(e_1) = \mathrm{Vect}(x_1)$.
Supposons maintenant que pour : $1 \le p \le n - 1$, on ait construit une famille $(e_1, \ldots, e_p)$ telle qu'indiqué et supposons alors de plus que : $\mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Vect}(x_1, \ldots, x_p)$.
On cherche alors e_{p+1} tel que : $e_{p+1} \in \mathrm{Vect}(x_1, \ldots, x_{p+1})$, donc sous la forme : $e_{p+1} = \lambda_1.x_1 + \ldots + \lambda_{p+1}.x_{p+1}$.
Mais puisqu'on a de plus supposé que : $\mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Vect}(x_1, \ldots, x_p)$, on peut alors écrire e_{p+1} sous la forme : $e_{p+1} = \lambda'_1.e_1 + \ldots + \lambda'_p.e_p + \lambda_{p+1}.x_{p+1}$.
On veut de plus que la famille $(e_1, \ldots, e_{p+1})$ soit orthogonale, ce qui s'écrit :
$\forall\ 1 \le i \le p$, $(e_i|e_{p+1}) = 0 = \lambda'_i.(e_i|e_p) + \lambda_{p+1}.(e_i|x_{p+1})$,

et comme les e_i sont non nuls, pour : $1 \le i \le p$, il est équivalent d'écrire : $\forall\ 1 \le i \le p$, $\lambda'_i = -\dfrac{(e_i|x_{p+1})}{\|e_i\|^2}.\lambda_{p+1}$.

Autrement dit, le vecteur e_{p+1} est nécessairement de la forme : $e_{p+1} = \lambda_{p+1}.\left[x_{p+1} - \sum_{i=1}^{p}\dfrac{(e_i|x_{p+1})}{\|e_i\|^2}.e_i\right]$.

Or le vecteur dans le crochet est non nul (puisque la famille $(x_{p+1}, e_1, \ldots, e_p)$ est libre) donc il est possible de trouver un vecteur e_{p+1} répondant aux conditions imposées (en prenant : $\lambda_{p+1} \ne 0$).
Enfin, on a : $\mathrm{Vect}(e_1, \ldots, e_{p+1}) \subset \mathrm{Vect}(e_1, \ldots, e_p, x_{p+1}) = \mathrm{Vect}(x_1, \ldots, x_{p+1})$, vu l'écriture de e_{p+1}.

Et comme de même, on a : $x_{p+1} = \sum_{i=1}^{p} \frac{(e_i|x_{p+1})}{\|e_i\|^2}.e_i + \frac{1}{\lambda_{p+1}}.e_{p+1}$,
on a aussi : $\text{Vect}(x_1, \ldots, x_{p+1}) \subset \text{Vect}(e_1, \ldots, e_{p+1})$, d'où l'égalité, et ceci quelque soit le choix de e_{p+1}.
La récurrence est donc terminée et l'existence d'une famille orthogonalisée établie.
• orthonormalisation de la famille.
Si on veut orthonormaliser la famille, alors la famille $(\varepsilon_1, \ldots, \varepsilon_n)$ est une de celles construites précédemment.
Montrons qu'à chaque étape de la construction, il n'y a qu'un vecteur qui répond aux conditions.
Pour ε_1, on doit normer e_1 et vérifier ensuite que : $(\varepsilon_1|x_1) > 0$.
Cela donne donc : $\varepsilon_1 = \pm\frac{x_1}{\|x_1\|}$, puis : $\varepsilon_1 = \frac{x_1}{\|x_1\|}$, pour que le produit scalaire soit positif.
Si maintenant, pour : $1 \leq p \leq n - 1$, on suppose tous les vecteurs ε_i, pour : $1 \leq i \leq p$, construits de façon unique, alors les vecteurs ε_i sont colinéaires aux vecteurs e_i précédents et ε_{p+1} s'écrit :

$$\varepsilon_{p+1} = \lambda_{p+1}.\left[x_{p+1} - \sum_{i=1}^{p} \frac{(e_i|x_{p+1})}{\|e_i\|^2}.e_i\right] = \lambda_{p+1}.y_{p+1}.$$

La fait que ε_{p+1} doit être normé laisse deux possibilités pour λ_{p+1} (puisque le vecteur y_{p+1} est non nul, comme on l'a vu précédemment), à savoir : $\lambda_{p+1} = \pm\frac{1}{\|y_{p+1}\|}$.
Si on fixe alors λ_{p+1} à l'une de ces valeurs, alors : $1 = (\varepsilon_{p+1}|\varepsilon_{p+1}) = \lambda_{p+1}.(\varepsilon_{p+1}|y_{p+1}) = \lambda_{p+1}.(\varepsilon_{p+1}|x_{p+1})$, puisque ε_{p+1} est orthogonal par construction aux vecteurs $(\varepsilon_1, \ldots, \varepsilon_p)$.
Donc $(\varepsilon_{p+1}|x_{p+1})$ est non nul, et il suffit enfin de choisir : $\lambda_{p+1} = +\frac{1}{\|y_{p+1}\|}$, pour garantir que le produit scalaire $(\varepsilon_{p+1}|x_{p+1})$ est strictement positif.
Finalement, le vecteur ε_{p+1} ainsi obtenu est le seul possible et il convient (au rang p+1).
Par récurrence, on vient bien de démontrer l'existence et l'unicité de la famille orthonormale annoncée.

***Définition 2.2* et théorème 2.4 : orthogonal d'une partie d'un espace vectoriel**
Soit $(E, (.|.))$ un espace préhilbertien réel.
Soit A une partie de E.
On appelle orthogonal de A, noté $A^\perp$, l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de A, soit : $A^\perp = \{x \in E, \forall y \in A, (x|y) = 0\}$.
Pour toute partie A de E, $A^\perp$ est un sous-espace vectoriel de E.
En particulier : $E^\perp = \{0\}$, et : $\{0\}^\perp = E$.

Démonstration :
• Il est clair que $A^\perp$ est inclus dans E.
De plus : $\forall y \in A, (0|y) = 0$, et : $0 \in A^\perp$.
Enfin : $\forall (x,x') \in (A^\perp)^2, \forall (\lambda,\lambda') \in \mathbb{R}^2, \forall y \in A, (\lambda.x + \lambda'.x'|y) = \lambda.(x|y) + \lambda'(x'|y) = 0$, (en rajoutant une barre de conjugaison sur les scalaires dans le cas complexe, ce qui ne change pas le résultat), et : $(\lambda.x+\lambda'.x') \in A^\perp$.
Finalement, $A^\perp$ est bien un sous-espace vectoriel de E.
• De plus si : $x \in E^\perp$, alors en particulier x est orthogonal lui-même et : $(x|x) = \|x\|^2 = 0$, d'où : $x = 0$.
• De même, tout vecteur est orthogonal à 0.

Définition 2.3 : sous-espaces vectoriels orthogonaux, supplémentaires orthogonaux
Soit $(E, (.|.))$ un espace préhilbertien réel.
Soient F et G des sous-espaces vectoriels de E.
On dit que F et G sont orthogonaux si et seulement si : $\forall (x,y) \in F\times G, (x|y) = 0$.
On note alors parfois : $F \perp G$.
On dit que F et G sont supplémentaires orthogonaux dans E si et seulement si F et G sont supplémentaires dans E et orthogonaux, et on note alors : $E = F \overset{\perp}{\oplus} G$.

Théorème 2.5 : cas de sous-espaces vectoriels de dimension finie

Soit (E, (.|.)) un espace préhilbertien réel.

Soient F et G des sous-espaces vectoriels de E de dimension finie, de bases respectives $\mathcal{B}_F$ et $\mathcal{B}_G$.

F et G sont orthogonaux si et seulement si tout vecteur de $\mathcal{B}_F$ est orthogonal à tout vecteur de $\mathcal{B}_G$.

Démonstration :

On peut procéder par double implication.

• [⇒] si tout vecteur de F est orthogonal à tout vecteur de G, c'est en particulier vrai pour les vecteurs de $\mathcal{B}_F$ et de $\mathcal{B}_G$.

• [⇐] puisque : F = Vect($\mathcal{B}_F$), et : G = Vect($\mathcal{B}_G$), la bilinéarité du produit scalaire montre qu'alors, tout vecteur de F est orthogonal à tout vecteur de G si c'est vrai pour les vecteurs de $\mathcal{B}_F$ et de $\mathcal{B}_G$.

Théorème 2.6 et définition 2.4 : somme directe orthogonale de sous-espaces vectoriels

Soit (E, (.|.)) un espace préhilbertien réel.

Soient F_1, …, F_n des sous-espaces vectoriels de E.

Si les sous-espaces vectoriels sont orthogonaux deux à deux, alors la somme $[F_1 + \ldots + F_n]$ est directe.

On dit alors que les sous-espaces vectoriels sont en somme directe orthogonale, et on la note encore :

$F_1 \overset{\perp}{\oplus} F_2 \overset{\perp}{\oplus} \ldots \overset{\perp}{\oplus} F_n$, ou : $\overset{n}{\underset{i=1}{\overset{\perp}{\oplus}}} F_i$.

Démonstration :

Soit un vecteur par exemple de l'intersection : $(F_1 + \ldots + F_{n-1}) \cap F_n$.

Alors on peut l'écrire : $x = x_1 + \ldots + x_{n-1} = x_n$, chaque x_i appartenant à l'espace F_i correspondant.

On constate que : $(x|x) = (x_n|x_1 + \ldots + x_{n-1}) = (x_n|x_1) + \ldots + (x_{n-1}|x_n) = 0$,

et ce résultat est identique en intervertissant les rôles de espaces.

Finalement, la somme est bien directe.

3. Projections orthogonales.

Définition 3.1 : projecteur orthogonal

Soit (E, (.|.)) un espace préhilbertien réel.

Soient F et G des sous-espaces vectoriels supplémentaires de E et p le projecteur de E sur F dans la direction G.

On dit que p est un projecteur orthogonal de E si : $G = F^\perp$.

Théorème 3.1 : supplémentaire orthogonal d'un sous-espace vectoriel de dimension finie

Soit (E, (.|.)) un espace préhilbertien réel de dimension finie ou non.

Soit F un sous-espace vectoriel de E de dimension finie m.

Alors $F^\perp$ est un supplémentaire de F dans E, appelé supplémentaire orthogonal de F dans E.

Démonstration :

On va montrer que tout vecteur de E peut se décomposer de façon unique comme somme d'un vecteur de F et de $F^\perp$.

Pour cela, soit $(e_1, \ldots, e_m)$ une base orthonormale $\mathcal{B}_F$ de F et soit x un vecteur de E.

Si x peut s'écrire : $x = x_F + x_\perp$, avec : $x_F \in F$, et : $x_\perp \in F^\perp$, alors :

$\exists (x_1, \ldots, x_p) \in \mathbf{K}^p$, $x_F = x_1.e_1 + \ldots + x_m.e_m$, et $x_\perp$ est orthogonal à tout vecteur de F.

En particulier : $\forall\ 1 \le i \le p$, $(e_i|x) = (e_i|x_F) + (e_i|x_\perp) = x_i$, puisque la base $\mathcal{B}_F$ est orthonormale.

Donc x_F ne peut valoir que : $x_F = \sum_{i=1}^{m} (e_i|x).e_i$, et : $x_\perp = x - x_F$.

Réciproquement, cette unique décomposition possible convient car on a bien :

• $x_F \in F$, puisque $\mathcal{B}_F$ engendre F,

• $x = x_F + x_\perp$, par construction,

• $\forall\ 1 \le k \le m$, $(e_k|x_\perp) = (e_k|x) - (e_k|x_F) = (e_k|x) - (e_k|x) = 0$, à nouveau parce que la base $\mathcal{B}_F$ de F est orthonormale et en utilisant la linéarité du produit scalaire par rapport à sa première variable.

Le troisième point permet de conclure plus généralement que : $\forall\ y \in F$, $(y|x_\perp) = 0$.

Théorème 3.2 : expression de la projection orthogonale sur un sous-espace vectoriel de dimension finie

Soit (E, (.|.)) un espace préhilbertien réel ou complexe de dimension finie ou non.
Soit F un sous-espace vectoriel de E de dimension finie m.
Soit $(e_1, \ldots, e_m)$ une base orthonormale de F et p_F la projection orthogonale de E sur F.

Alors pour tout vecteur x de E, on a : $p_F(x) = \sum_{i=1}^{m} (e_i|x).e_i$.

Démonstration :

On vient de montrer qu'un vecteur x de E pouvait se décomposer suivant la somme directe : $E = F \oplus F^{\perp}$,

en : $x = \sum_{i=1}^{m} (e_i|x).e_i + x_{\perp}$, puisque la base choisie dans F est orthonormale.

Dans ce cas, la projection orthogonale $p_F(x)$ de x sur F s'écrit : $p_F(x) = \sum_{i=1}^{m} (e_i|x).e_i$.

Définition 3.2 et théorème 3.3 : distance d'un vecteur à un sous-espace vectoriel de dimension finie

Soit (E, (.|.)) un espace préhilbertien réel ou complexe de dimension finie ou non et $\|\,.\,\|$ la norme associée.
Soit F un sous-espace vectoriel de E de dimension finie p.
Pour un vecteur x de E, on appelle distance de x à F, notée d(x,F), la quantité : $d(x,F) = \inf_{z \in F} \|x - z\|$.
Cette valeur est atteinte en un unique vecteur de F qui est $p_F(x)$, la projection orthogonale de x sur F.
On a donc : $d(x,F) = \|x - p_F(x)\|$.

Démonstration :

Tout d'abord, pour un vecteur x de E, l'ensemble $\{\|x - z\|, z \in F\}$ est non vide (il suffit de prendre : z = 0) et il est constitué de réels positifs, donc il est minoré et il admet une borne inférieure.
De plus, en notant x_F la projection orthogonale $p_F(x)$ de x sur F, on a :
$\forall z \in F$, $\|x - z\|^2 = \|x - x_F + x_F - z\|^2$.
Or : $(x - x_F) \in F^{\perp}$, et : $(x_F - z) \in F$, donc ces deux vecteurs sont orthogonaux et le théorème de Pythagore peut s'appliquer pour donner : $\forall z \in F$, $\|x - z\|^2 = \|x - x_F\|^2 + \|x_F - z\|^2$.
Il est alors clair que : $\forall z \in F$, $\|x - z\| \geq \|x - x_F\|$, et : $\forall z \in F$, $(z \neq x_F) \Rightarrow (\|x - z\| > \|x - x_F\|)$.
Donc la valeur $\|x - x_F\|$ est le plus petit élément de l'ensemble $\{\|x - z\|, z \in F\}$, donc aussi sa borne inférieure, et cette dernière n'est atteinte qu'en : $z = x_F = p_F(x)$.

Théorème 3.4 : inégalité de Bessel

Soit (E, (.|.)) un espace préhilbertien réel de dimension finie ou non et $\|\,.\,\|$ la norme associée.
Soit F un sous-espace vectoriel de E de dimension finie m.
Soit $(e_1, \ldots, e_m)$ une base orthonormale de F.

Pour tout vecteur x de E, on a : $\sum_{i=1}^{m} |(e_i|x)|^2 \leq \|x\|^2$.

Démonstration :

Soit x un vecteur de E, et : $x = \sum_{i=1}^{m} (e_i|x).e_i + x_{\perp}$, sa décomposition suivant la somme : $E = F \oplus F^{\perp}$, puisque la base proposée dans E est orthonormale.
Alors les deux vecteurs de cette décomposition étant orthogonaux, le théorème de Pythagore s'applique à nouveau et la base de F étant orthonormale, l'expression de la norme au carré est alors canonique et :

$$\|x\|^2 = \left\|\sum_{i=1}^{m} (e_i|x).e_i\right\|^2 + \|x_{\perp}\|^2 \geq \left\|\sum_{i=1}^{m} (e_i|x).e_i\right\|^2 = \sum_{i=1}^{m} |(e_i|x)|^2 .$$

4. Espaces euclidiens.

Définition 4.1 : espace vectoriel euclidien

On appelle espace euclidien un espace vectoriel réel de dimension finie, muni d'un produit scalaire.

Théorème 4.1 : existence de bases orthonormales dans les espaces euclidiens

Soit (E, (.|.)) un espace vectoriel euclidien.

Il existe dans E des bases orthonormales pour (.|.).

Démonstration :

Il suffit de considérer une base de E (donc une famille libre) et de lui appliquer le procédé d'orthonormalisation de Schmidt pour obtenir une nouvelle famille libre de n vecteurs de E (puisque orthogonale sans le vecteur nul) donc une base orthonormale de E.

Théorème 4.2 : de la base incomplète orthogonale ou orthonormale

Soit (E, (.|.)) un espace vectoriel euclidien.

Soit $(x_1, \dots, x_p)$ une famille orthogonale de vecteurs de E, ne comportant pas le vecteur nul.

Alors il est possible de compléter la famille $(x_1, \dots, x_p)$ en une base orthogonale de E.

De même, si $(e_1, \dots, e_p)$ est une famille orthonormale de vecteurs de E, il est possible de compléter cette famille en une base orthonormale de E.

Démonstration :

On utilise cette fois le théorème de la base incomplète appliqué à la famille $(x_1, \dots, x_p)$ et une base de E.

La famille complétée peut alors être orthogonalisée suivant le procédé de Schmidt qui donne pour les p premiers vecteurs, les vecteurs $x_1, \dots, x_p$ eux-mêmes.

Dans le cas d'une famille $(e_1, \dots, e_p)$ de départ, orthonormale, on applique l'orthonormalisation pour obtenir une base de E orthonormale, dont les premiers vecteurs sont encore $(e_1, \dots, e_p)$.

Théorème 4.3 : caractérisation matricielle des bases orthogonales ou orthonormales

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n.

Soit : $\mathscr{B} = (e_i)$, une base de E.

On note M la matrice du produit scalaire dans la base (e_i) définie par : $\forall\ 1 \leq i,j \leq n,\ m_{i,j} = (e_i|e_j)$.

La matrice M est alors symétrique.

On a de plus les équivalences suivantes, que E soit un espace vectoriel euclidien ou hermitien :

- ($\mathscr{B}$ est orthogonale) $\Leftrightarrow$ (M est diagonale),
- ($\mathscr{B}$ est orthonormale) $\Leftrightarrow$ ($M = I_n$).

Plus généralement, si M est la matrice d'une forme bilinéaire symétrique d'un espace euclidien E, alors M est symétrique.

Démonstration :

Puisque : $\forall\ (i,j) \in \mathbb{N}_n^2$, $m_{i,j} = (e_i|e_j) = (e_j|e_j) = m_{j,i}$, la matrice M est bien symétrique.

De plus, les deux résultats suivants sont presque immédiats puisque :

- ($\mathscr{B}$ est orthogonale) $\Leftrightarrow$ ($\forall\ 1 \leq i \neq j \leq n$, $(e_i|e_j) = 0 = m_{i,j}$) $\Leftrightarrow$ (M est diagonale).
- ($\mathscr{B}$ est orthonormale) $\Leftrightarrow$ ($\forall\ 1 \leq i,j \leq n$, $(e_i|e_j) = \delta_{i,j}$) $\Leftrightarrow$ ($M = I_n$).

Théorème 4.4 : expression matricielle du produit scalaire

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n, et soit : $\mathscr{B} = (e_i)$, une base de E.

Soient x et y des vecteurs de E, et X et Y les matrices colonnes de leurs coordonnées dans $\mathscr{B}$.

Alors : $(x|y) = {}^tX.M.Y$.

Si $\mathscr{B}$ est orthonormale, on a de plus :

- $(x|y) = {}^tX.Y = \sum_{i=1}^{n} x_i.y_i$ (forme canonique d'un produit scalaire réel dans une base orthonormale),
- $x = \sum_{i=1}^{n} (e_i|x).e_i$,
- $\|x\|^2 = \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} (e_i|x)^2$.

Démonstration :

• Pour le premier point, il suffit de développer :

On note pour commencer : Y' = M.Y, et on a : $\forall$ 1 ≤ i ≤ n, $y'_i = \sum_{k=1}^{n} m_{i,k}.y_k$.

Puis ^{t}X.M.Y est une matrice à une ligne et une colonne et son seul coefficient vaut alors :

$\sum_{i=1}^{n} x_i.y'_i = \sum_{i=1}^{n}\sum_{k=1}^{n} m_{i,k}.x_i y_k$.

D'autre part, en utilisant la bilinéarité du produit scalaire, on obtient également :

$(x|y) = \sum_{i=1}^{n} x_i.(e_i|y) = \sum_{i=1}^{n} x_i.\sum_{k=1}^{n} y_k.(e_i|e_k) = \sum_{i=1}^{n}\sum_{k=1}^{n} x_i.y_k.m_{i,k}$,

et les deux expressions sont égales si on confond la matrice 1×1 avec son unique coefficient.

• Si la base est orthonormale alors : $\forall$ 1 ≤ i,j ≤ n, $m_{i,j} = \delta_{i,j}$, la matrice M vaut I_n, et on a bien :

$(x|y) = {}^{t}X.Y = \sum_{i=1}^{n} x_i.y_i$.

• Soit : x = $x_1.e_1$ + … + $x_n.e_n$, la décomposition de x suivant la base orthonormale $\mathscr{B}$.

On peut alors calculer : $(e_i|x) = \sum_{j=1}^{n} x_j.(e_i|e_j) = x_i$, d'où l'expression de x.

• Enfin : $\|x\|^2 = (x|x) = \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} (e_i|x)^2$, vu les expressions trouvées précédemment.

Théorème 4.5 : dimension de l'orthogonal d'un sous-espace vectoriel dans un espace euclidien

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n.
Soit F un sous-espace vectoriel de E.
Alors : dim($F^{\perp}$) = n – dim(F).
De plus : $(F^{\perp})^{\perp}$ = F.

Démonstration :

• Puisque F est de dimension finie, $F^{\perp}$ est un supplémentaire de F dans E et : dim($F^{\perp}$) + dim(F) = n.

• De plus : $\forall$ x ∈ F, $\forall$ y ∈ $F^{\perp}$, $(x|y) = 0$, donc : x ∈ $(F^{\perp})^{\perp}$, soit : F ⊂ $(F^{\perp})^{\perp}$.

Et comme : dim($(F^{\perp})^{\perp}$) = n – dim($F^{\perp}$) = n – (n – dim(F)) = dim(F), on en déduit que : F = $(F^{\perp})^{\perp}$.

Théorème 4.6 : caractérisation en dimension finie des supplémentaires orthogonaux

Soit (E, (.|.)) un espace préhilbertien réel.
Soient F et G des sous-espaces vectoriels de E.
F et G sont supplémentaires orthogonaux dans E si et seulement si on obtient une base orthogonale de E en réunissant une base orthogonale de F et une base orthogonale de G.

Démonstration :

Travaillons par double implication :

• [⇒]

Si F et G sont supplémentaires orthogonaux dans E, alors en réunissant des bases orthogonales de F et de G, on obtient tout d'abord une base de E, et tous les vecteurs de la base de F étant orthogonaux à tous ceux de la base de G, la base de E obtenue est bien orthogonale.

• [⇐]

Si la réunion de deux bases orthogonales de F et de G donne une base orthogonale de E, alors les vecteurs de la base de F et ceux de G sont orthogonaux entre eux et F et G sont orthogonaux d'après le théorème 2.4.
De plus, il est clair que F et G sont supplémentaires dans E.

Théorème 4.7 : représentation d'une forme linéaire à l'aide du produit scalaire

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n.
Soit φ* une forme linéaire sur E.
Alors il existe un unique vecteur : a ∈ E, tel que : $\forall$ x ∈ E, $\varphi^*(x) = (a|x)$.

Démonstration :

Soit : $\mathscr{B}$ = (e_1, …, e_n), une base orthonormale de E.

Alors : ∃ (a_1, …, a_n) ∈ $\mathbb{R}^n$, ∀ x ∈ E, $x = \sum_{i=1}^{n} x_i.e_i$, $\varphi^*(x) = \sum_{i=1}^{n} a_i.x_i$.

Le vecteur : $a = \sum_{i=1}^{n} a_i.e_i$, est alors bien tel que : ∀ x ∈ E, $\varphi^*(x) = (a|x)$, puisque $\mathscr{B}$ est orthonormale.

Puis si a' répond également au problème, alors : ∀ x ∈ E, $\varphi^*(x) = (a|x) = (a'|x)$, d'où : $(a - a'|x) = 0$.

On a donc : (a – a') ∈ $E^\perp$, et : a – a' = 0, d'où : a = a'.

Théorème 4.8 : vecteur normal à un hyperplan d'un espace euclidien

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n.
Soit H un hyperplan de E, et : $\mathscr{B}$ = (e_1, …, e_n), une base orthonormale de E.
Une équation de H dans $\mathscr{B}$ fournit alors un vecteur orthogonal à H, dit aussi vecteur normal à H.

Démonstration :

Si : $\mathscr{B}$ = (e_1, …, e_n), est une base orthonormale de E, et si : $a_1.x_1 + \ldots + a_n.x_n = 0$, est une équation de H dans $\mathscr{B}$, alors : H = ker(φ*), où φ* est la forme linéaire définie par :

∀ x ∈ E, $x = \sum_{i=1}^{n} x_i.e_i$, $\varphi^*(x) = \sum_{i=1}^{n} a_i.x_i$.

Le vecteur : $a = \sum_{i=1}^{n} a_i.e_i$, est alors orthogonal à H puisque : ∀ x ∈ H, $\varphi^*(x) = 0 = (a|x)$.

5. Automorphismes orthogonaux et matrices orthogonales.

Définition 5.1 et théorème 5.1 : endomorphisme orthogonal dans un espace vectoriel euclidien

Soit (E, (.|.)) un espace vectoriel euclidien et $\|\,.\,\|$ la norme associée.

Soit : u ∈ $\mathscr{L}$(E).
On dit que u est un endomorphisme orthogonal de E si et seulement si u conserve la norme ou le produit scalaire de E, c'est-à-dire : ∀ (x,y) ∈ E^2, $(u(x)|u(y)) = (x|y)$, ou : ∀ x ∈ E, $\|u(x)\| = \|x\|$.

Ces deux définitions sont bien équivalentes.

Démonstration :

Si u conserve le produit scalaire, alors :

∀ x ∈ E, $\|u(x)\|^2 = (u(x)|u(x)) = (x|x) = \|x\|^2$, et u conserve bien la norme.

Si par ailleurs, u conserve la norme, alors l'expression du produit scalaire à l'aide de normes (par le biais des identités dites « de polarisation » (théorème 1.6)) et la linéarité de u montrent que u conserve le produit scalaire.

Théorème 5.2 : bijectivité des endomorphismes orthogonaux en dimension finie

Soit (E, (.|.)) un espace vectoriel euclidien et $\|\,.\,\|$ la norme associée.

Soit : u ∈ $\mathscr{L}$(E), un endomorphisme orthogonal de E.
Alors u est bijectif, et donc tout endomorphisme orthogonal d'un espace vectoriel euclidien est un automorphisme : on parle donc d'automorphisme orthogonal.
De plus, u ne peut admettre comme valeur propre que ±1.

Démonstration :

Puisque l'on travaille dans un espace euclidien, il est de dimension finie et il suffit de démontrer que l'endomorphisme u est injectif.

Or : ∀ x ∈ E, (x ∈ ker(u)) ⇔ (u(x) = 0) ⇔ ($\|x\| = 0$) ⇔ (x = 0).

Supposons maintenant que λ est valeur propre (donc réelle) de u, et soit x un vecteur propre associé.
Alors : u(x) = λ.x, et : $\|u(x)\| = \|\lambda.x\| = |\lambda|.\|x\|$, mais aussi : $\|u(x)\| = \|x\|$, et comme x est non nul : |λ| = 1.

Définition 5.2 : matrice orthogonale

Soit : $n \geq 1$.

On dit que : $A \in \mathcal{M}_n(\mathbb{R})$, est une matrice orthogonale si et seulement si : ${}^tA.A = I_n$, ou : $A.{}^tA = I_n$.

Théorème 5.3 : caractérisation des matrices orthogonales par leurs vecteurs lignes ou colonnes

Soit : $n \geq 1$.

Une matrice : $A \in \mathcal{M}_n(\mathbb{R})$, est orthogonale si et seulement si ses colonnes, considérées comme des vecteurs de $\mathbb{R}^n$ (ou ses lignes) forment une base orthonormale de $\mathbb{R}^n$, pour le produit scalaire canonique de $\mathbb{R}^n$.

De plus, si : $A \in \mathcal{M}_n(\mathbb{R})$, est orthogonale, alors : $\det(A) = \pm 1$.

Démonstration :

Soit : $A \in \mathcal{M}_n(\mathbb{R})$.

On peut tout d'abord remarquer que :

$({}^tA.A = I_n) \Leftrightarrow (A^{-1} = {}^tA) \Leftrightarrow (A.{}^tA = I_n) \Leftrightarrow$ (A orthogonale).

• Puis, si on note B la matrice : $B = {}^tA.A$, ses coefficients sont :

$\forall\ 1 \leq i,j \leq n$, $b_{i,j} = (C_i|C_j)$, où le produit scalaire est ici le produit scalaire canonique de $\mathbb{R}^n$, et où on a noté $C_1, \ldots, C_n$ les colonnes de A, considérées comme des vecteurs de $\mathbb{R}^n$.

Donc A est orthogonale si et seulement si : $B = I_n$, donc si et seulement si la famille (formée de n vecteurs) $(c_1, \ldots, c_n)$ est orthonormale dans $\mathbb{R}^n$, donc en forme une base orthonormale.

• De même, en notant $L_1, \ldots, L_n$ les lignes de A, et : $B' = A.{}^tA$, on démontre l'autre équivalence.

• Enfin, si A est orthogonale, alors : $\det(I_n) = 1 = \det({}^tA.A) = \det({}^tA).\det(A) = (\det(A))^2$.

Théorème 5.4 : caractérisations des automorphismes orthogonaux

Soit $(E, (.|.))$ un espace vectoriel euclidien et $\|\,.\,\|$ la norme associée.

Soit : $u \in \mathcal{L}(E)$.

• u est un automorphisme orthogonal si et seulement si sa matrice dans une base orthonormale de E est une matrice orthogonale.

• u est un automorphisme orthogonal si et seulement si l'image d'une base orthonormale de E par u est une autre base orthonormale.

En particulier, si u est un automorphisme orthogonal, alors : $\det(u) = \pm 1$.

Démonstration :

Soit A la matrice de u dans une base $\mathcal{B}$ orthonormale de E.

• Si A est orthogonale, alors : $\forall$ x $\in$ E, $(u(x)|u(x)) = {}^t(A.X).(A.X)$, en notant X la matrice des coordonnées de X dans la base $\mathcal{B}$ et puisque cette même base est orthonormale.

Donc : $(u(x)|u(x)) = {}^tX.{}^tA.A.X = {}^tX.I_n.X = {}^tX.X = (x|x)$, et u est bien un automorphisme orthogonal.

• Si maintenant on suppose u orthogonal, alors : $\forall$ (x,y) $\in$ E^2, on a : $(u(x)|u(y)) = (x|y)$.

En notant X et Y les matrices colonnes des coordonnées de X et de Y dans $\mathcal{B}$, on obtient :

$\forall$ (X,Y) $\in \mathcal{M}_n(\mathbb{R})^2$, ${}^tX.{}^tA.A.Y = {}^tX.Y$.

Notons alors : $B = {}^tA.A$, et soit X_k la matrice colonne formée de 0 sauf celui de la ligne k qui vaut 1.

Alors : $\forall\ 1 \leq k,l \leq n$, ${}^tX_k.{}^tA.A.X_l = {}^tX_k.B.X_l = b_{k,l}$, et : ${}^tX_k.X_l = \delta_{k,l}$.

On en déduit que : $B = I_n$, et donc : ${}^tA.A = I_n$, autrement dit A est orthgonale.

• Soit maintenant : $\mathcal{B} = (e_1, \ldots e_n)$, une base orthonormale de E.

Si u est orthogonal alors par conservation du produit scalaire, la famille $(u(e_1), \ldots, u(e_n))$ est clairement une famille orthonormale (donc une base) de E.

Et si réciproquement, on suppose que $(u(e_1), \ldots, u(e_n))$ est orthonormale, alors :

$\forall$ x $\in$ E, $x = \sum_{i=1}^{n} x_i.e_i$, $\|u(x)\|^2 = \left(\sum_{i=1}^{n} x_i.u(e_i) \middle| \sum_{j=1}^{n} x_j.u(e_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i.x_j.(u(e_i)|u(e_j)) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i.x_j.\delta_{i,j}$,

et donc : $\|u(x)\|^2 = \sum_{i=1}^{n} x_i^2 = \|x\|^2$.

Puisque u conserve la norme, c'est donc bien un automorphisme orthogonal.

• En particulier, si u est orthogonal et si A est sa matrice représentative dans une base orthonormale de l'espace, alors : det(u) = det(A) = ±1.

Théorème 5.5 : automorphisme orthogonal et sous-espaces vectoriels stables

Soit (E, (.|.)) un espace vectoriel euclidien.
Soient u un automorphisme orthogonal de E et F un sous-espace vectoriel de E, stable par u.
Alors $F^\perp$ est également stable par u.

Démonstration :

Puisque F est stable par u, on a : u(F) ⊂ F, et u étant bijectif, donc conservant la dimension : u(F) = F.
Pour : x ∈ $F^\perp$, on peut écrire : ∀ y ∈ F, ∃ y' ∈ F, y = u(y'), et : $(u(x)|y) = (u(x)|u(y')) = (x|y) = 0$.
Donc : u(x) ∈ $F^\perp$, et $F^\perp$ est donc également stable par u.

Théorème 5.6 et définition 5.3 : les groupes (O(E),o), (O(n),×), (SO(E),o) et (SO(n),×)

Soit (E, (.|.)) un espace vectoriel euclidien de dimension n et $\|\,.\,\|$ la norme associée.

- Alors l'ensemble des endomorphismes orthogonaux de E, noté O(E), forme un groupe pour la loi o, appelé groupe orthogonal de E et sous-groupe de (Gl(E),o).
- De même, l'ensemble O(n) des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$, forme un groupe pour la loi ×, appelé groupe orthogonal d'ordre n, et sous-groupe de (Gl(n),×).
- Par ailleurs, les éléments de O(E) dont le déterminant vaut 1 forment un sous-groupe de O(E) appelé Groupe spécial orthogonal de E, de même les matrices de O(n) dont le déterminant vaut 1 forment un sous-groupe de O(n) appelé groupe spécial orthogonal d'ordre n.
- Enfin, si on fixe une base orthonormale $\mathcal{B}$ de E, l'application de $\mathcal{L}$(E) dans $\mathcal{M}_n(\mathbb{R})$, qui à un endomorphisme u associe sa matrice dans la base $\mathcal{B}$, induit un isomorphisme de groupes de (O(E),o) dans (O(n),×).

Démonstration :

- O(E) :

On a pour commencer : O(E) ⊂ Gl(E).
De plus il est clair que id_E conserve la norme donc c'est un automorphisme orthogonal, et : O(E) ≠ ∅.
Puis, si u et v conservent la norme dans E, alors : ∀ x ∈ E, $\|uov(x)\| = \|u(v(x))\| = \|v(x)\| = \|x\|$.
Donc : ∀ (u,v) ∈ $O(E)^2$, uov ∈ O(E).
Enfin, pour tout u dans O(E), u est bijectif et : ∀ x ∈ E, $\|u^{-1}(x)\| = \|u(u^{-1}(x))\| = \|x\|$, soit : u^{-1} ∈ O(E).

- O(n) :

De la même façon, toute matrice orthogonale étant inversible, on a : O(n) ⊂ Gl(n).
De plus : ${}^tI_n.I_n = I_n$, donc : I_n ∈ O(n), et : O(n) ≠ ∅.
Puis : ∀ (A,B) ∈ $O(n)^2$, ${}^t(A.B).(A.B) = {}^tB.{}^tA.A.B = {}^tB.I_n.B = I_n$, et : (A.B) ∈ O(n).
Enfin : ∀ A ∈ O(n), $A^{-1} = {}^tA$, et : ${}^t(A^{-1}).A^{-1} = {}^t({}^t(A)).{}^tA = A.{}^tA = I_n$, et : A^{-1} ∈ O(n).

- SO(E) :

On a : id_E ∈ O(E), puisque : $\det(id_E) = 1$, et : SO(E) ≠ ∅.
Puis : ∀ (u,v) ∈ $SO(E)^2$, alors : $\det(u \circ v) = \det(u).\det(v) = 1$, et : uov ∈ SO(E).
Enfin, si : u ∈ SO(E), $\det(u^{-1}) = (\det(u))^{-1} = 1$, et : u^{-1} ∈ SO(E).
SO(E) est bien un sous-groupe de O(E).

- On montre de la même façon que SO(n) forme un sous-groupe de O(n).
- L'application proposée est bien une application de O(E) dans O(n), d'après le théorème 5.4.

C'est un morphisme de groupes multiplicatifs, car : ∀ (u,v) ∈ $O(E)^2$, mat(uov,$\mathcal{B}$) = mat(u,$\mathcal{B}$).mat(v,$\mathcal{B}$).
Enfin, c'est bien une application injective puisque l'application de $\mathcal{L}$(E) dans $\mathcal{M}_n(\mathbb{R})$ l'est, et elle est surjective puisque pour une matrice orthogonale donnée A, l'endomorphisme u de E tel que :
A = mat(u,$\mathcal{B}$), est orthogonal, toujours d'après le théorème 5.4.

Théorème 5.7 : matrice de passage entre bases orthonormales

Soit (E, (.|.)) un espace vectoriel euclidien et $\|\,.\,\|$ la norme associée.
Soit $\mathcal{B}$ une base orthonormale de E et $\mathcal{B}$' une autre base de E.
Si on note P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}$', alors $\mathcal{B}$' est orthonormale si et seulement si P est orthogonale.

Démonstration :

Soit donc $\mathscr{B}$ une bas orthonormale de E et $\mathscr{B}'$ une autre base de E en notant P la matrice de passage de l'une à l'autre.
Les coefficients de P sont (en colonne) les coordonnées des vecteurs de $\mathscr{B}'$ exprimés dans la base $\mathscr{B}$.
Or dans ce cas : $\forall (j,j') \in \{1, \dots, n\}^2$,
${}^tC_j.C_{j'} = (C_j|C_{j'})$, pour le produit scalaire canonique (des colonnes C_j et $C_{j'}$ de P) dans $\mathbb{R}^n$,
${}^tC_j.C_{j'} = \sum_{k=1}^{n} p_{k,j}.p_{k,j'} = (e'_j|e'_{j'})$, où les e'_k sont les vecteurs de $\mathscr{B}'$, et où le produit scalaire est cette fois celui dans E : en effet, la base $\mathscr{B}$ étant orthonormale, la forme que prend le produit scalaire est encore canonique.
Il est alors clair que : $({}^tP.P = I_n) \Leftrightarrow (\forall (j,j') \in \{1, \dots, n\}^2, {}^tC_j.C_{j'} = \delta_{i,j} = (e'_j,e'_{j'})) \Leftrightarrow (\mathscr{B}'$ orthonormale).

6. Espaces euclidiens de dimension 2 ou 3.

Dans ce paragraphe, E désigne un espace euclidien de dimension 2 ou 3.

Définition 6.1 : orientation d'un espace vectoriel, orientation induite par un sous-espace vectoriel

• Une orientation de E correspond au choix d'une base de E décrétée comme directe.
Toute base de E est alors soit directe, soit indirecte, suivant que le déterminant de la matrice de passage de cette nouvelle base à la base de référence est positif ou négatif.
• Si F est un sous-espace vectoriel de E (un plan ou une droite), orienter F permet alors de définir une orientation induite dans tout sous-espace supplémentaire de F dans E de la façon suivante :
si $\mathscr{B}_F$ est une base directe de F, une base d'un supplémentaire G de F sera directe si la concaténation de $\mathscr{B}_F$ et de $\mathscr{B}_G$ est une base directe de E.

Exemple :

Soit $\mathbb{R}^3$ muni de son produit scalaire canonique et de son orientation canonique (la base canonique est directe).
Soit : $\Delta = \mathrm{Vect}(e_1)$, avec : $e_1 = (1,1,1)$), et muni de l'orientation donnée par ce vecteur.
Alors : $\Pi = \mathrm{Vect}(e_2,e_3)$, où : $(e_2,e_3) = ((1,1,0), (1,0,1))$, est un supplémentaire de Δ dans $\mathbb{R}^3$ (puisque la réunion des deux bases de Π et de Δ forme une base de $\mathbb{R}^3$), et (e_2,e_3) est une base indirecte de Π dans l'orientation induite par Δ car la matrice de passage P de la base canonique à (e_1,e_2,e_3) vérifie : $\det(P) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{vmatrix} = -1$.

Remarques :

• Si on prend au départ une base orthonormale comme base directe et qu'on examine les autres bases orthonormales de E, alors le déterminant de la matrice de passage vaut toujours ± 1.
• Orienter une droite (dans le plan ou l'espace) correspond donc à la donnée d'un vecteur de norme 1.
• Orienter un plan dans $\mathbb{R}^3$ revient alors à choisir une base orthonormale du plan comme base directe, mais de façon équivalente à choisir un vecteur de norme 1 et normal au plan.
En effet l'orientation de $\mathbb{R}^3$ induit, par le choix de ce vecteur normal au plan, une orientation dans le plan.

Théorème 6.1 et définition 6.2 : produit mixte

Soit E est un espace euclidien de dimension 2 ou 3 orienté.
Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases orthonormales directes de E.
Alors :
• si E est de dimension 2 : $\forall (u,v) \in E^2$, $\det_{\mathscr{B}}(u,v) = \det_{\mathscr{B}'}(u,v)$,
• si E est de dimension 3 : $\forall (u,v,w) \in E^3$, $\det_{\mathscr{B}}(u,v,w) = \det_{\mathscr{B}'}(u,v,w)$.
Cette valeur invariante est notée : $[u,v] = \det_{\mathscr{B}}(u,v)$, ou : $[u,v,w] = \det_{\mathscr{B}}(u,v,w)$, suivant le cas, et est appelé produit mixte de u et de v ou de u, v et w, suivant le cas.

Démonstration :

Pour un vecteur x de E de matrice de coordonnées X et X' dans $\mathscr{B}$ et $\mathscr{B}'$, on a : $X = P.X'$, où P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.
Si on note A et A' les matrices des coordonnées de (u,v) (ou de (u,v,w)) dans les bases $\mathscr{B}$ et $\mathscr{B}'$, on a alors, en utilisant un produit par blocs : $A = P.A'$.

Donc : $\det_{\mathscr{B}}(u,v) = \det(A) = \det(P).\det(A') = +1.\det(A') = \det_{\mathscr{B}'}(u,v)$, de même si E est de dimension 3.

Théorème 6.2 et définition 6.3 : produit vectoriel de deux vecteurs en dimension 3

Soit E est un espace euclidien de dimension 3 orienté.
Soient u et v deux vecteurs de E.
Il existe un unique vecteur w de E, noté $u \wedge v$ et appelé produit vectoriel de u et de v tel que :
$\forall x \in E$, $[u,v,x] = (w|x)$.

Démonstration :

L'application de E dans $\mathbb{R}$ définie par : $x \mapsto [u,v,x]$, est une forme linéaire sur E, donc le théorème 4.6 garantit qu'il existe un unique vecteur w dans E tel que : $\forall x \in E$, $[u,v,x] = (w|x)$.

Théorème 6.3 : propriétés du produit vectoriel

Soit E est un espace euclidien de dimension 3 orienté.
Le produit vectoriel a les propriétés suivantes :

- il est bilinéaire : $\forall (u,u',v,v') \in E^4$, $\forall (\lambda,\lambda') \in \mathbb{R}^2$,
  $u \wedge (\lambda.v + \lambda'.v') = \lambda.u \wedge v + \lambda'.u \wedge v'$, et : $(\lambda.u + \lambda'.u') \wedge v = \lambda.u \wedge v + \lambda'.u' \wedge v$,
- il est alterné : $\forall (u,v) \in E^2$, $v \wedge u = -u \wedge v$,
- $\forall (u,v) \in E^2$, $u \wedge v$ est orthogonal à u et à v.
- $\forall (u,v) \in E^2$, $(u \wedge v = 0) \Leftrightarrow$ ((u,v) est liée) $\Leftrightarrow$ (u et v colinéaires).
- $\forall (u,v) \in E^2$, ((u,v) libre) $\Leftrightarrow$ (u et v non colinéaires) $\Leftrightarrow$ ((u,v, $u \wedge v$) est une base directe de E.

Démonstration :

- Soient u, v, et v' dans E, λ et λ' dans $\mathbb{R}$.

Alors : $\forall x \in E$, $(u \wedge (\lambda.v + \lambda'.v')|x) = [u, \lambda.v + \lambda'.v', x] = \lambda.[u,v,x] + \lambda'.[u,v',x] = \lambda.(u \wedge v|x) + \lambda'.(u \wedge v'|x)$,
d'où : $(u \wedge (\lambda.v + \lambda'.v')|x) = (\lambda.u \wedge v + \lambda'.u \wedge v'|x)$, et par unicité : $u \wedge (\lambda.v + \lambda'.v') = \lambda.u \wedge v + \lambda'.u \wedge v'$.

- On montre de même la relation similaire par rapport à la première variable.
- Soient u et v dans E.

Alors : $\forall x \in E$, $(v \wedge u|x) = [v,u,x] = -[u,v,x] = -(u \wedge v|x) = (-u \wedge v|x)$, et par unicité : $v \wedge u = -u \wedge v$.

Puis : $(u \wedge v|u) = [u,v,u] = 0$, de même pour v, ce qui montre que $u \wedge v$ est orthogonal à u et à v.

De plus :

- si (u,v) est liée, alors : $\forall x \in E$, $(u \wedge v|x) = [u,v,x] = 0$, et comme 0 vérifie aussi : $(0|x) = 0$, par unicité on a : $u \wedge v = 0$,
- si (u,v) est libre alors : $\exists w \in E$, tel que (u,v,w) est une base de E, et : $(u \wedge v|w) = [u,v,w] \neq 0$, et donc le vecteur $u \wedge v$ est non nul.

Dans ce dernier cas, on a alors : $[u,v,u \wedge v] = (u \wedge v|u \wedge v) = \|u \wedge v\|^2 > 0$, et la base (u,v, $u \wedge v$) est une base directe de E.

Théorème 6.4 : expression du produit vectoriel dans une base orthonormale directe

Soit E est un espace euclidien de dimension 3 orienté muni d'une base orthonormale directe $\mathscr{B}$.
Soient u et v deux vecteurs de E de coordonnées respectives (x,y,z) et (x',y',z') dans $\mathscr{B}$.
Alors $u \wedge v$ a pour coordonnées dans $\mathscr{B}$ $(y.z' - y'.z, z.x' - z'.x, x.y' - y.x')$.

Démonstration :

Si on note (a,b,c) les coordonnées de $u \wedge v$ dans : $\mathscr{B}$ = (i,j,k), alors :

$$a = (u \wedge v|i) = [u,v,i] = \begin{vmatrix} x & x' & 1 \\ y & y' & 0 \\ z & z' & 0 \end{vmatrix} = y.z' - z.y'.$$

On obtient les deux autres coordonnées de la même façon avec j et k.

Théorème 6.5 : expression géométrique du produit vectoriel

Soit E est un espace euclidien de dimension 3.

Soient u et v deux vecteurs non colinéaires de E et Π le plan engendré par u et v.
Si on oriente Π à l'aide d'un vecteur unitaire n normal à Π (donc dirigeant et orientant : Δ = $\Pi^\perp$), alors :
$u \wedge v = \|u\|.\|v\|.\sin(\theta).n$, où θ est l'angle orienté (u,v).

Démonstration :

Notons : $i = \dfrac{u}{\|u\|}$, et j tel que (i,j) soit une base orthonormale du plan telle que (i,j,n) soit directe.

Alors : $u = \|u\|.i$, et : ∃ θ ∈]0,π[, $v = \|v\|.(\cos(\theta).i + \sin(\theta).j)$.

En effet, avec : $v = \|v\|.(x.i + y.j)$, alors : $x^2 + y^2 = \dfrac{\|v\|^2}{\|v\|^2} = 1$, et : ∃ ! θ ∈ [0,2.π], $x = \cos(\theta)$, $y = \sin(\theta)$.

Dans ce cas, θ est bien l'angle orienté (u,v) et : $u \wedge v = \|u\|.i \wedge (\|v\|.(\cos(\theta).i + \sin(\theta).j)) = \|u\|.\|v\|.\sin(\theta).n$.

Remarque :

Pour trois vecteurs u, v et w de E, espace euclidien de dimension 3 orienté on a :

- $\|u \wedge v\|$ représente la surface du parallélogramme défini par les vecteurs u et v,
- $|[u,v,w]|$ représente le volume du parallélépipède construit sur les vecteurs u, v et w.

Théorème 6.6 : éléments de O(2) : matrices orthogonales 2×2

Soit : A ∈ O(2), une matrice orthogonale de taille 2×2.

- si : det(A) = +1, il existe un réel θ tel que : $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$.
- si : det(A) = – 1, il existe un réel θ tel que : $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.

Démonstration :

Notons : $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ ∈ O(2).

Alors dans tous les cas : $a^2 + b^2 = c^2 + d^2 = 1$, et : a.c + b.d = 0, puisque les vecteurs colonnes de A constituent une base orthonormale de $\mathbb{R}^2$ muni de son produit scalaire canonique.
Donc il existe θ et θ' réels tels que : a = cos(θ), b = sin(θ), c = sin(θ'), d = cos(θ').
De plus, on a aussi : cos(θ).sin(θ') + sin(θ).cos(θ') = 0 = sin(θ + θ'), soit : θ + θ' = k.π, avec : k ∈ ℤ.

On peut donc écrire : $A = \begin{pmatrix} \cos(\theta) & -(-1)^k.\sin(\theta) \\ \sin(\theta) & (-1)^k.\cos(\theta) \end{pmatrix}$.

Enfin :

- si : det(A) = +1, alors : $(-1)^k = +1$, et : $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$,
- si : det(A) = – 1, alors : $(-1)^k = -1$, et : $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.

Remarque :

Si on identifie E au plan complexe ℂ, en représentant des vecteurs par leur affixe, alors la rotation vectorielle d'angle θ est représentée par l'application : $z \mapsto z.e^{i.\theta}$.

Théorème 6.7 : automorphismes orthogonaux d'un espace vectoriel euclidien de dimension 2

Soit (E, (.|.)) un espace vectoriel euclidien de dimension 2, orienté.
Soit u un automorphisme orthogonal de E et A la matrice de u dans une base : $\mathscr{B}$ = (i,j), orthonormale directe de E.

- si : det(u) = +1, alors il existe un réel θ tel que : $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$.

Si u n'est pas l'identité de E, u est la rotation de E d'angle θ, θ étant donné par : 2.cos(θ) = tr(u).

• si : det(u) = – 1, alors il existe un réel θ tel que : $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.

u est alors la symétrie orthogonale par rapport à la droite D d'équation : $x.\sin\left(\frac{\theta}{2}\right) - y.\cos\left(\frac{\theta}{2}\right) = 0$, et l'angle de D avec la droite dirigée par i est égal à $\frac{\theta}{2}$.

Démonstration :

Si u est un automorphisme orthogonal de E alors sa matrice représentative A dans une base orthonormale de E est alors orthogonale et on se trouve dans l'un des deux cas précédents.

De plus, si :

• det(u) = +1, alors : det(A) = +1, et : ∃ θ ∈ ℝ, $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$.

Distinguons alors deux cas.

Si : θ = 0 (2.π), alors u est l'identité de E sinon u est une rotation de E d'angle θ, et on obtient par : tr(A) = 2.cos(θ) = tr(u).

Si par ailleurs :

• det(u) = – 1, alors : det(A) = – 1, et : ∃ θ ∈ ℝ, $A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.

Puisqu'alors A est symétrique réelle, elle est diagonalisable et ses valeurs propres (réelles) λ_1 et λ_2 vérifient, en utilisant trace et déterminant : $\lambda_1 + \lambda_2 = 0$, et : $\lambda_1.\lambda_2 = -1$.

Donc λ_1 et λ_2 valent 1 et – 1.

La matrice de u dans une base de vecteurs propres est alors $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, et : $u^2 = id_E$.

u est donc bien une symétrie vectorielle.

Comme enfin, les espaces propres de u sont orthogonaux, c'est une symétrie orthogonale.

La droite par rapport à laquelle s'opère cette symétrie est l'ensemble des vecteurs invariants de u donnés par : A.X = 0, soit : $\begin{cases} (\cos(\theta) - 1).x + \sin(\theta).y = 0 \\ \sin(\theta).x - (\cos(\theta) + 1).y = 0 \end{cases}$, ou : $\begin{cases} -2.\sin\left(\frac{\theta}{2}\right).\left[\sin\left(\frac{\theta}{2}\right).x - \cos\left(\frac{\theta}{2}\right).y\right] = 0 \\ +2.\cos\left(\frac{\theta}{2}\right).\left[\sin\left(\frac{\theta}{2}\right).x - \cos\left(\frac{\theta}{2}\right).y\right] = 0 \end{cases}$.

Comme enfin, soit le sinus, soit le cosinus mis en facteur est non nul, on en déduit bien que la droite invariante est la droite D dont l'équation dans la base $\mathscr{B}$ est bien celle annoncée.

Le vecteur $\left(\cos\left(\frac{\theta}{2}\right), \sin\left(\frac{\theta}{2}\right)\right)$ est alors directeur de D et l'angle de ce vecteur avec i est bien $\frac{\theta}{2}$.

Théorème 6.8 : produit de rotations, commutativité de SO(2) et SO(E), pour E de dimension 2

Le groupe SO(2) est commutatif.

Si E est un espace vectoriel euclidien de dimension 2, le groupe SO(E) est commutatif.

En particulier, deux matrices orthogonales de déterminant +1 commutent, tout comme deux rotations dans un plan euclidien.

Démonstration :

Si A et A' sont deux matrices de rotation d'angle θ et θ', alors :

$$A.A' = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} . \begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix} = \begin{pmatrix} \cos(\theta + \theta') & -\sin(\theta + \theta') \\ \sin(\theta + \theta') & \cos(\theta + \theta') \end{pmatrix} = A'.A.$$

Si r et r' sont deux rotations de E, de matrices respectives A et A' dans une base orthonormale directe de E, alors ror' est la rotation d'angle (θ+θ').

Théorème 6.9 : endomorphismes orthogonaux d'un espace vectoriel euclidien de dimension 3

Soit (E, (.|.)) un espace vectoriel euclidien de dimension 3, orienté.

Soit u un automorphisme orthogonal de E.
• si : det(u) = +1, et si u n'est pas l'identité de E, alors u admet une droite Δ de vecteurs invariants.
Soit alors Π le plan orthogonal à Δ.
- Si : tr(u) = – 1, alors u est la symétrie orthogonale par rapport à Δ (ou rotation d'angle π autour de Δ).
- Si : tr(u) $\neq$ – 1, alors pour x un vecteur non nul de Π (ou orthogonal à Δ), le vecteur : $k = x \wedge u(x)$, est un vecteur non nul de Δ.
Si dans ce cas, on oriente Δ avec k et Π avec l'orientation induite par celle de Δ, u est alors la rotation d'axe Δ et d'angle $+\theta$, où θ est donné par la relation : $2.\cos(\theta) + 1 = tr(u)$.
• si : det(u) = – 1, et si u n'est pas $-id_E$, alors l'ensemble des vecteurs changés en leur opposé par u est une droite Δ de E.
Notons encore Π le plan orthogonal à Δ.
- Si : tr(u) = +1, u est la symétrie orthogonale par rapport à Π.
- Si : tr(u) $\neq$ +1, alors pour x un vecteur non nul de Π (ou orthogonal à Δ), le vecteur : $k = x \wedge u(x)$, est encore un vecteur non nul de Δ.
Si dans ce cas, on oriente Δ avec k,et Π avec l'orientation induite par celle de Δ, alors u est la composée de la rotation d'axe Δ et d'angle $+\theta$, où θ est donné par la relation : $2.\cos(\theta) - 1 = tr(u)$, et de la symétrie orthogonale de E par rapport à Π.

Démonstration :

Tout d'abord, le polynôme caractéristique P_u de u est de degré 3 à coefficients réels, donc il admet une racine réelle au moins.
Distinguons alors deux cas :
• P_u admet trois racines réelles et ça ne peut être que 1 ou – 1,
• P_u admet une racine réelle et deux racines complexes conjuguées : $\lambda_1 = \pm 1$, λ, et $\bar{\lambda}$.
Le produit des racines de P_u vaut par ailleurs det(u).
Deux cas alors se présentent :
• det(u) = +1.
Les racines de P_u sont donc (+1, +1, +1) ou (+1, – 1, – 1), si les trois racines sont réelles, et $(1, \lambda, \bar{\lambda})$, avec : $|\lambda| = 1$, s'il y a des racines non réelles.
On constate donc que 1 est toujours valeur propre de u.
Si u n'est pas l'identité, 1 est alors racine simple de P_u et l'espace propre associé est de dimension 1 donc c'est une droite Δ de E.
Notons alors Π le plan orthogonal à Δ.
Alors P est stable par u, car : $\forall (x,y) \in \Pi\times\Delta$, $(u(x)|y) = (u(x)|u(y)) = (x|y) = 0$, donc : $u(x) \in \Delta^{\perp} = \Pi$.
Or l'endomorphisme u' induit par u dans Π est encore une isométrie de Π (puisqu'il conserve lui aussi la norme, des vecteurs de Π cette fois).
Dans une base orthonormale : $\mathscr{B} = (e_1, e_2, e_3)$, de E adaptée alors à la décomposition : $E = \Pi\oplus\Delta$, on a :

$mat(u,\mathscr{B}) = \begin{pmatrix} mat(u', B') & 0 \\ 0 & 1 \end{pmatrix}$, où $\mathscr{B}'$ est une base orthonormale de Π.

On constate alors que : det(u) = det(u'), et donc : det(u') = +1.
Le théorème 7.2 montre alors que :

$\exists \theta \in \mathbb{R}$, $mat(u',B') = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$, et : $mat(u,\mathscr{B}) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Si maintenant : tr(u) = – 1, cela signifie que : $2.\cos(\theta) + 1 = -1$, et : $\pi = 0\ (2.\pi)$.

La matrice de u dans la base précédente est alors : $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, et u est bien la symétrie orthogonale par rapport à la droite Δ.

Si en revanche : tr(u) $\neq$ – 1, alors u' est la rotation de Π d'angle $+\theta$ si Π est orienté avec la base précédente, l'angle θ étant donné par : $tr(u) = tr(mat(u,\mathscr{B})) = 2.\cos(\theta) + 1$.
Plus généralement, si x est un vecteur non nul de Π, alors u(x) est non colinéaire à x puisque sinon u' admettrait des valeurs propres réelles ce qui n'est pas la cas.
Donc la famille $(x, u(x), x\wedge u(x))$ est une famille libre de E, donc une base de E, directe.

Si on oriente alors Δ avec : $k = x \wedge u(x)$, et Π avec l'orientation induite par ce choix sur Δ, la famille $(x, u(x))$ est une base directe de Π et la rotation dans le plan Π s'effectue dans le sens direct.
Reste à examiner le cas :

• $\det(u) = -1$.

Les racines de P_u sont alors : $(-1, -1, -1)$, $(-1, +1, +1)$, ou : $(-1, \lambda, \bar{\lambda})$.
-1 est alors toujours valeur propre de u.
Si u n'est pas $-id_E$, cette valeur propre est simple et l'espace propre associé est de dimension 1, soit une droite vectorielle Δ de E.
Notons encore Π le plan orthogonal à Δ, et s la symétrie orthogonale de E par rapport à Π.

Pour une base $\mathscr{B}$ adaptée à la décomposition : $E = \Pi \oplus \Delta$, on a alors : $\mathrm{mat}(s,\mathscr{B}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, et la composée sou est une isométrie de E, positive puisque : $\det(s) = \det(u) = -1$, Δ étant de plus invariante par cette isométrie.
D'après ce qu'on vient juste d'établir, on peut en déduire que :

$\exists\ \theta \in \mathbb{R}$, $\mathrm{mat}(sou,\mathscr{B}) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$, et : $\mathrm{mat}(u,\mathscr{B}) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Dans le cas où : $\cos(u) = -1$, u est alors $-id_E$, mais ce cas est écarté ici, sinon deux cas se présentent :
 $tr(u) = +1$, (soit : $\cos(\theta) = 1$) et u est alors la symétrie orthogonale par rapport à Δ,
 $tr(u) \neq +1$, (soit : $\cos(\theta) \neq 1$) et u est la composée de s et d'une rotation d'axe Δ.
Les éléments géométriques de cette rotation s'obtiennent alors comme précédemment, notamment pour tout ce qui concernent les différentes orientations d'espaces, avec comme différence le fait que son angle est cette fois donné par : $tr(u) = tr(\mathrm{mat}(u,\mathscr{B})) = 2.\cos(\theta) - 1$.

Théorème 6.10 : éléments de O(3) : matrices orthogonales 3×3

Soit : $A \in O(3)$, une matrice orthogonale de taille 3×3.

• si : $\det(A) = +1$, alors : $\exists\ \theta \in \mathbb{R}$, $P \in O(3)$, $A = P.\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.{}^tP$.

• si : $\det(A) = -1$, alors : $\exists\ \theta \in \mathbb{R}$, $P \in O(3)$, $A = P.\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{pmatrix}.{}^tP$.

Démonstration :

Si on appelle u l'endomorphisme canoniquement associé (dans $\mathbb{R}^3$ muni de son produit scalaire canonique), u est un endomorphisme orthogonal de $\mathbb{R}^3$, donc il existe comme on l'a vu dans la démonstration du théorème 7.3 une base orthonormale de $\mathbb{R}^3$ dans laquelle la matrice de u est d'un des deux types proposés (la différence se faisant par l'examen du déterminant de u).
Il suffit alors d'appeler P la matrice de passage de la base canonique de $\mathbb{R}^3$ à la base $\mathscr{B}$ évoquée au-dessus : cette matrice est orthogonale, d'où le résultat annoncé.

7. Réduction des endomorphismes symétriques, des matrices symétriques réelles.

Définition 7.1 : endomorphisme symétrique

Soit (E, (.|.)) un espace vectoriel euclidien.
On dit qu'un endomorphisme u de E est symétrique si et seulement si : $\forall\ (x,y) \in E^2$, $(u(x)|y) = (x|u(y))$.

Théorème 7.1 : caractérisation matricielle des endomorphismes symétriques

Soit (E, (.|.)) un espace vectoriel euclidien.
Soit : $u \in \mathscr{L}(E)$.

L'endomorphisme u est symétrique si et seulement si sa matrice représentative dans toute base **orthonormale** de E est symétrique.

Démonstration :

La démonstration est similaire à celle du théorème 5.4.

Soient : $\mathscr{B} = (e_1, \ldots, e_n)$, une base orthonormale de E, et : A = mat(u,$\mathscr{B}$).

• Si A est symétrique, alors pour x et y dans E, de matrices colonnes de coordonnées dans $\mathscr{B}$ notées X et Y, on a : $(u(x)|y) = {}^t(A.X).Y = {}^tX.{}^tA.Y = {}^tX.A.Y = {}^tX.(A.Y) = (x|u(y))$, et u est symétrique.

• Si u est symétrique alors :

$$\forall\ 1 \le i, j \le n,\ (e_i|u(e_j)) = \left(e_i \Big| \sum_{k=1}^{n} a_{k,j}.e_k\right) = \sum_{k=1}^{n} a_{k,j}.(e_i|e_k) = a_{i,j}, \text{ et : } (u(e_i)|e_j) = (e_j|u(e_i)) = a_{j,i}.$$

Et puisque : $(e_i|u(e_j)) = (u(e_i)|e_j)$, on en déduit que : $a_{i,j} = a_{j,i}$, et A est symétrique.

Théorème 7.2 : valeurs propres d'une matrice symétrique réelle, d'un endomorphisme symétrique

Soit : $n \ge 1$.

Soit : $A \in \mathscr{M}_n(\mathbb{R})$, une matrice symétrique réelle.

• Alors A, considérée comme élément de $\mathscr{M}_n(\mathbb{C})$, est telle que toutes les racines de son polynôme caractéristique sont réelles et donc a toutes ses valeurs propres réelles.

• De même, soit (E, (.|.)) un espace vectoriel euclidien, et soit : $u \in \mathscr{L}(E)$, symétrique.

Toutes les racines du polynôme caractéristique de u sont réelles.

Démonstration :

• Soit λ une racine de P_A éventuellement complexe, et X un vecteur propre (dans $\mathbb{C}^n$) associé à λ.

Alors : $A.X = \lambda.X$.

On peut écrire :

${}^t\overline{X}.A.X = \lambda.{}^t\overline{X}.X = \lambda.\sum_{k=1}^{n}|x_k|^2$, mais aussi, puisque A est symétrique et réelle et vérifie : ${}^t\overline{A} = A$,

$${}^t\overline{X}.A.X = {}^t\overline{X}.{}^t\overline{A}.X = {}^t\overline{(A.X)}.X = {}^t\overline{(\lambda.X)}.X = \overline{\lambda}.{}^t\overline{X}.X = \overline{\lambda}.\sum_{k=1}^{n}|x_k|^2.$$

Et comme X est un vecteur propre, il est non nul, donc : $\sum_{k=1}^{n}|x_k|^2 \neq 0$, et : $\lambda = \overline{\lambda}$.

Donc toute racine de P_A est réelle.

• De même, si u est un endomorphisme symétrique de E, sa matrice dans n'importe quelle base orthonormale de E est symétrique réelle, donc les racines de P_u qui sont celles de P_A sont toutes réelles.

Théorème 7.3 : orthogonalité des espaces propres d'un endomorphisme symétrique

Soit (E, (.|.)) un espace vectoriel euclidien, et soit : $u \in \mathscr{L}(E)$, symétrique.

Si λ et μ sont des valeurs propres distinctes de u, les espaces propres associés $E_\lambda(u)$ et $E_\mu(u)$ sont orthogonaux.

Démonstration :

• Toutes les valeurs propres de u sont réelles.

• Soient λ et μ deux valeurs propres distinctes de u, et x et y des éléments de $E_\lambda(u)$ et $E_\mu(u)$.

Alors : $(x|u(y)) = (x|\mu.y) = \mu.(x|y)$, mais aussi : $(x|u(y)) = (u(x)|y) = (\lambda.x|y) = \lambda.(x|y)$.

Or : $\lambda \neq \mu$, donc : $(x|y) = 0$, et les sous-espaces propres $E_\lambda(u)$ et $E_\mu(u)$ sont bien orthogonaux.

Théorème 7.4 : diagonalisabilité des endomorphismes symétriques

Soit (E, (.|.)) un espace vectoriel euclidien, et soit : $u \in \mathscr{L}(E)$, symétrique.

Il existe une base orthonormale de E dans laquelle l'endomorphisme u est diagonalisable.

E est de plus la somme directe orthogonale des sous-espaces propres de u.

Démonstration :

On effectue cette démonstration par récurrence sur la dimension n de l'espace.

Si u est un endomorphisme symétrique d'un espace euclidien de dimension 1, il est évidemment diagonalisable puisque sa matrice dans toute base de E est de taille 1×1, donc diagonale.

Supposons le résultat acquis pour tout endomorphisme symétrique d'un espace vectoriel de dimension : $n \geq 1$, donnée.
Soit alors u un endomorphisme symétrique d'un espace euclidien E de dimension (n+1).
Puisque P_u est scindé dans $\mathbb{R}$, u admet au moins une valeur propre λ et un vecteur propre associé e_1.
Notons alors E' le supplémentaire orthogonal de $Vect(e_1)$ dans E.

- E' est stable par u :

En effet : $\forall\ x \in E'$, $(e_1|u(x)) = (u(e_1)|x) = \lambda.(e_1|x) = 0$, puisque x est orthogonal à e_1.
Donc : $u(x) \in Vect(e_1)^{\perp}$, et : $u(x) \in E'$.

- Notons alors u' l'endomorphisme induit par u dans E' : l'espace E' est de dimension n puisque supplémentaire dans E de $Vect(e_1)$.
- L'endomorphisme u' de E' est symétrique car : $\forall\ (x,y) \in E'^2$, $(x|u'(y)) = (x|u(y)) = (u(x)|y) = (u'(x)|y)$.

Donc il existe une base orthonormale $(e_2, \ldots, e_{n+1})$ de E' formée de vecteurs propres de u'.
Mais ces vecteurs sont aussi vecteurs propres de u et en réunissant cette famille à e_1, on obtient une base orthonormale $(e_1, \ldots, e_{n+1})$ de E, formée de vecteurs propres de E, ce qui termine la récurrence.

Théorème 7.5 : diagonalisabilité des matrices symétriques réelles

Soit : $n \geq 1$.
Soit : $A \in \mathcal{M}_n(\mathbb{R})$, une matrice symétrique réelle.
Alors A est diagonalisable et il est possible de la diagonaliser par l'intermédiaire d'une matrice orthogonale.

Démonstration :

Si u désigne l'endomorphisme de $\mathbb{R}^n$ canoniquement associé à A, il est symétrique pour le produit scalaire canonique de $\mathbb{R}^n$, puisque la base canonique est orthonormale pour ce produit scalaire et A est symétrique réelle.
Donc u est diagonalisable et ses espaces propres sont orthogonaux.
Si on considère alors une base de vecteurs propres de u formée de la réunion de bases orthonormales des différents espaces propres de u, on obtient une base de vecteurs propres de u qui est une base orthonormale.
La matrice de passage P de la base canonique à cette base de vecteurs propres est alors orthogonale et si on note enfin D la matrice $P^{-1}.A.P$, elle est diagonale, autrement dit on vient de diagonaliser A par l'intermédiaire d'une matrice orthogonale.